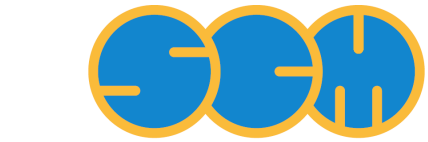


Scientific Computing & Modelling

ADF Manual

ADF Program System
Release 2014

Scientific Computing & Modelling NV
Vrije Universiteit, Theoretical Chemistry
De Boelelaan 1083; 1081 HV Amsterdam; The Netherlands
WWW: www.scm.com
E-mail: support@scm.com

Table of Contents

Preface

ADF (Amsterdam Density Functional) is a Fortran program for calculations on atoms and molecules (in gas phase or solution). It can be used for the study of such diverse fields as molecular spectroscopy, organic and inorganic chemistry, crystallography and pharmacochemistry. A separate program BAND is available for the study of periodic systems: crystals, surfaces, and polymers. The COSMO-RS program is used for calculating thermodynamic properties of (mixed) fluids.

The underlying theory is the Kohn-Sham approach to Density-Functional Theory (DFT). This implies a one-electron picture of the many-electron systems but yields in principle the exact electron density (and related properties) and the total energy.

If ADF is a new program for you we recommend that you carefully read Chapter 1, section 1.3 'Technical remarks, Terminology', which presents a discussion of a few ADF-typical aspects and terminology. This will help you to understand and appreciate the output of an ADF calculation.

ADF has been developed since the early 1970s (at that time called HFS, later AMOL, see also Refs. [308-310]), mainly by the two theoretical chemistry groups of, respectively, the Vrije Universiteit in Amsterdam (http://www.chem.vu.nl/en/research/division-theoretical-chemistry/index.asp) and the University of Calgary, Canada (http://www.cobalt.chem.ucalgary.ca/group/master.html). Other researchers have also contributed. As a major research tool of these academic development groups, ADF is in continuous development and retains a firm basis in the academic world.

Maintenance and distribution of the commercial (export) version of the program is done by Scientific Computing & Modelling NV (SCM) (http://www.scm.com), a company based in Amsterdam, formally split off from the theoretical chemistry group in Amsterdam but practically still very much a part of it. Documentation such as User manuals, Installation instructions, Examples, Theoretical documents can be found at the SCM web site.

Publications based on research with ADF should include appropriate references to the program. We recommend that references are made both to the program itself and to publications related to its development and structure. See the References document, available at the SCM web site.

Release 2014

In comparison to ADF 2013, the ADF 2014 release offers the following new functionality:

Functionality

- DFT-MBD dispersion corrected XC functional
- intensity selected excitation energies
- NMR spin-spin couplings with subsystem DFT
- distance difference restraints in optimizations

Analysis

- Charge model 5

Accuracy

- improved density fitting with radial spline functions and Zlm' can be used for most properties
- stricter settings for distance cut-offs in calculating Hartree-Fock exchange integrals
- NMR chemical shifts: spin-orbit gauge correction term

Speed

- scalable SCF
- COSMO runs in parallel

Default settings changed

- default density fitting scheme changed to ZlmFit instead of STO fit
- NMR chemical shifts uses unscaled ZORA instead of scaled ZORA

Apart from this new functionality and performance improvements, certain bugs have been fixed.

A more extended list of 'what is new or different' can be found in the Release Notes document.

1 GENERAL

1.1 Introduction

The installation of the Amsterdam Density Functional program package (ADF) on your computer is explained in the Installation manual. This User's Guide describes how to use the program, how input is structured, what files are produced, and so on. Some special applications of ADF are described in the Examples document.

Where references are made to the operating system (OS) and to the file system on your computer the terminology of UNIX type OSs is used and a hierarchical structure of *directories* is assumed.

The ADF package is in continuous development to extend its functionality and applicability, to increase its efficiency and user-friendliness, and of course to correct errors. We appreciate comments and suggestions for improvement of the software and the documentation.

Characterization of ADF

Functionality

- Single Point calculation
- Geometry Optimization
- Transition States
- Frequencies and thermodynamic properties
- Tracing a Reaction Path
- Computation of any electronic configuration
- Excitation energies, oscillator strengths, transition dipole moments, (hyper)polarizabilities, Van der Waals dispersion coefficients, CD spectra, ORD, using Time-Dependent Density Functional Theory (TDDFT)
- ESR (EPR) g-tensors, A-tensors, NQCCs
- NMR chemical shifts and spin-spin coupling constants
- Various other molecular properties
- Treatment of large systems and environment by the QM/MM (Quantum Mechanics / Molecular Mechanics) hybrid approach.

Applicability

All elements of the periodic table can be used (Z = 1-118). For each of the elements, the database contains basis sets of different sizes, ranging from minimal to high quality. Special basis sets are provided for relativistic calculations within the ZORA approach and for response calculations that require additional diffuse basis functions.

Model Hamiltonian

- A choice of Density Functionals, both for the Local Density Approximation (LDA), for the Generalized Gradient Approximation (GGA), for hybrid functionals (not for all properties available), and for meta-GGA functionals (not for all properties available) are available.
- Spin: restricted or unrestricted
- Relativistic effects: scalar approximation and spin-orbit (double-group symmetry), using the (now recommended) ZORA or the (previously used) Pauli formalism
- Environment: Solvent Effects, Homogeneous Electric Field, Point Charges (Madelung Fields), QM/MM method

Analysis

- Decomposition of the bond energy in 'chemical' components (steric interaction, Pauli repulsion, orbital interactions...)
- Representation of data (Molecular Orbital coefficients, Mulliken Populations) in terms of the constituent chemical fragments in the molecule, along with the conventional representation in elementary basis functions
- Atomic charge determination by Hirshfeld analysis and by Voronoi analysis, multipole derived charges, along with the classical Mulliken populations, and Mayer bond orders

Technical

- The implementation is based upon a highly optimized numerical integration scheme for the evaluation of matrix elements of the Fock operator, property integrals involving the charge density, etc. The code has been vectorized and parallelized.
- Basis functions are Slater-Type Orbitals (STOs). A database is available with several basis sets for each atom in the periodic table of elements.
- The Coulomb potential is evaluated via an accurate fitting of the charge density.
- A frozen core facility is provided for an efficient treatment of the inner atomic shells.
- Extensive use is made of point group symmetry. Most of the commonly encountered symmetry groups are available.
- Linear scaling techniques are used to speed up calculations on large molecules

Fragments

ADF has a fragment oriented approach: the poly-atomic system to be computed is conceptually built up from fragments, the molecular one-electron orbitals are calculated as linear combinations of fragment orbitals, and final analyzes of e.g. the bonding energy are in terms of fragment properties. The fragments may be single atoms or larger moieties.

When you compute a system in terms of its constituent fragments, these fragments must have been computed before and their properties must be passed on to the current calculation. This is done by attaching *fragment files*, which contain the necessary information. A fragment file is simply the standard result file of an ADF calculation on that fragment.

When using Basic Atoms as fragments, you do not need to create the fragment files yourself. Instead, you may use the Basis key, and ADF will create the required fragment files automatically. We therefore recommend this feature for starting ADF users.

Basic atoms

Obviously there must be a set of fundamental fragments that are not defined in terms of smaller fragments. Therefore ADF has two modes of execution: the normal mode, using fragments, and the create mode, in which a fundamental fragment is generated. Such a fundamental fragment *must* be a single atom, spherically symmetric and spin-restricted (i.e. spin-α and spin-β orbitals are spatially identical, they are equally occupied, and fractional occupations are applied, if necessary, to distribute the electrons equally over symmetry-degenerate states). Such a fundamental fragment is denoted a *basic atom*. The basic atoms are the smallest building blocks from which any 'real' calculations are started.

One should realize that the basic atoms are artificial objects that are convenient in the computational approach but that do not necessarily represent real atoms very well (in fact, usually not at all). The bonding energy of a molecule with respect to basic atoms, for instance, should be corrected for this discrepancy in order to get a decent comparison against experimental data. See ref. [1] for a discussion and for examples of applicable values.

A basic atom is computed in the conventional way. The one-electron orbitals are determined as linear combinations of basis functions; the frozen core approximation may be applied for the inner atomic states; a particular type of density functional can be chosen, et cetera. You may have, for instance, different basic Copper atoms by using different basis sets, by choosing different levels of frozen core approximations, or by applying different density functionals.

Slater-type basis sets

ADF uses Slater-Type Orbitals (STO's) as basis functions. Slaters can display the correct nuclear cusp and asymptotic decay.

$$f(\mathbf{r}) = Y_{lm} r^n e^{-\zeta r}$$

The center of the function is at a nucleus, the Y_{lm} are spherical harmonics, and the exponential factor ζ (zeta) determines the long-range decay of the function.

ADF has a database, which is included in the distribution, with thoroughly tested basis set files, ranging in quality from single-zeta to quadruple-zeta basis sets with various diffuse and polarization functions. All-electron and frozen-core basis sets are available for all elements, including lanthanides and actinides. The frozen-core approximation can be used to considerably reduce the computation time for systems with heavy nuclei, in a controlled manner.

Automatic mode

If you are using 'Basic Atom' fragments only, you do not need to prepare the corresponding fragment files yourself. Instead, add the BASIS block key to the ADF input, and ADF will generate all the required fragment files for you. This makes your job scripts and ADF inputs simpler, it ensures that consistent options for the create runs and molecular runs are used, and you will be sure that the fragment files used have been created by the same release of ADF.

1.2 Feature List

Model Hamiltonians

- XC energy functionals and potentials
  - LDA, GGA, meta-GGA, model potentials
  - (meta-)hybrid, range-separated
  - dispersion corrected
- Relativistic effects (ZORA and spin-orbit coupling)
- Solvents and other environments
  - COSMO, QM/MM, pdb2adf, DIM/QM, SCRF, FDE, 3D-RISM, QUILD
- Homogeneous electric field and point charges

Structure and Reactivity

- Geometry Optimizations
- Linear Transit, Transition States, CI-NEB, TSRC
- Intrinsic Reaction Coordinate
- Excited state optimizations with TDDFT gradients

Optimizations can be done in Cartesian, internal, and delocalized coordinates.
Various restraints and constraints (1,2) can be imposed.

Hessians are available analytically for most GGAs, and numerically otherwise. Preoptimization is possible with DFTB.

Spectroscopic properties

- Vibrational Spectroscopy
  - IR frequencies and intensities
  - Mobile Block Hessian (MBH), Vibrational Circular Dichroism (VCD)
  - Raman intensities
  - Resonance Raman from frequency-dependent polarizabilities or excited state gradients
  - vibrational Raman optical activity (VROA)
  - Franck-Condon Factors
- Excitation energies: UV/Vis spectra, X-ray absorption, CD, MCD
  - UV/Vis spectra, oscillator strengths, open shell excitations, core excitations, spin-orbit coupled excitations
  - vibrationally resolved electronic spectra
  - excited state optimizations
  - CD spectra, MCD
- (Hyper-)Polarizabilities, dispersion coefficients, ORD, magnetizabilities, Verdet constants
  - frequency-dependent (hyper)polarizabilities, lifetime effects
  - van der Waals dispersion coefficients
  - optical rotatory dispersion (ORD)
  - magnetizability
  - Verdet constants, Faraday terms
- NMR
  - chemical shifts
  - spin-spint couplings
- ESR (EPR)
  - g-tensors (g-factor)
  - A-tensor (hyperfine interaction)
  - zero-field splitting (ZFS, D-tensor)
- Nuclear quadrupole interaction (EFG), ESR Q-tensor
- Mössbauer, NRVS

Charge transport properties

- charge transfer integrals
- Non-self-consistent Green's function calculation

Analysis

- Fragments
- Bond energy decomposition, ETS-NOCV
- Advanced charge density and MO analysis
  - Mulliken, Multipole-derived charges
  - Hirshfeld charges, Voronoi deformation density, CM5 charges
  - bond orders: Mayer, Nalewajski-Mrozek
  - Bader (QT-AIM)
  - NB0 6.0
  - (partial) DOS
- Molecular symmetry,Schönfliess symbols and symmetry labels

Accuracy and Efficiency

- Slater-type basis sets
    - Z = 1 to 118, all electron, frozen-core, nonrelativistic and relativistic
    - SZ, DZ, DZP, TZP, TZ2P, QZ4P, even-tempered, diffuse
- Integration scheme
- Parallelization
- Linear scaling / distance cut-offs
- Density fit and frozen core approximation
- SCF convergence: simple damping, DIIS, EDIIS, ADIIS, LISTi, ARH

1.3 Technical remarks, Terminology

A few words about ADF as regards its technical setup and the names and abbreviations used in this manual. References to these will be made in the discussion of output and print switches.

Density functional theory

The underlying theory of the ADF package is the Kohn-Sham approach to the Density-Functional Theory (DFT). Kohn-Sham DFT is an important first-principles computational method to predict chemical properties accurately and to analyze and interpret these in convenient and simple chemical terms.

The reasons for its popularity and success are easy to understand. In the first place, the DFT approach is in principle exact. In particular, the Kohn-Sham method implies a one-electron picture of the many-electron systems but yields in principle the exact electron density (and related properties) and the total energy. The exact exchange-correlation (XC) functional is unknown, but the currently available XC functionals provide in most cases already a 'chemical' accuracy of a few kcal/mol for binding energies. Moreover, the quest for more accurate ones based on a more detailed understanding of their essential properties is continuing.

In the past two decades, computational chemistry has evolved from a curiosity of theoreticians into a mainstream tool used by all types of chemists, physicists and engineers who have an interest in research and development. In that time Density Functional Theory has become the dominant method for modeling chemistry at the molecular level.

In the second place, it preserves at all levels of approximation the appealing one-electron molecular orbital (MO) view on chemical reactions and properties. The computed orbitals are suitable for the typical MO-theoretical analyses and interpretations. The KS method effectively incorporates all correlation effects.

In the third place, it is a relatively efficient computational method, and its fundamental scaling properties do not deteriorate when methodological precision is increased, in particular, when a more accurate XC functional is applied. Recent research paves the way to implementations that scale only linearly with the system size. This brings within reach the treatment by fundamental quantum chemical methods of systems with hundreds, maybe even thousands of atoms.

DFT gives superior accuracy to Hartree-Fock theory and semi-empirical approaches, and it is well suited for molecules containing metal atoms. In contrast to conventional ab initio methods (MP2, CI, CC), it enables accurate treatment of systems with several hundreds of atoms (or several thousands with QM/MM).

Text is mostly taken from: *Chemistry with ADF*, G. te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler J. Comp. Chem. **22** (2001) 931.

The Kohn-Sham MO model

The basic postulate in Kohn-Sham DFT is that we can apply a one-electron formulation to the system of N interacting electrons by introducing a suitable local potential $V_{XC}(\mathbf{r})$, in addition to any external potentials $V_{ext}(\mathbf{r})$ and the Coulomb potential of the electron cloud $V_C(\mathbf{r})$, and solving:

$$[T + V_{ext}(\mathbf{r}) + V_C(\mathbf{r}) + V_{XC}(\mathbf{r})] \varphi_i (\mathbf{r}) = \varepsilon_i \varphi_i (\mathbf{r})$$

Here T is the kinetic energy operator. The potential V_{XC} is the functional derivative with respect to the density ρ of $E_{XC}[\rho]$, the exchange-correlation energy functional. The one-electron molecular orbitals (MOs) φ_i with corresponding orbital energies ε_i define the exact electronic charge density and give, in principle, access to all properties because these are expressible as functional of the density, in particular the energy. Moreover, they provide an intuitively appealing view of the system as being built from independent-electron orbitals with all ensuing interpretations. The exact form of the exact energy density $E_{XC}(\mathbf{r})$, representing and incorporating all exchange and correlation (XC) effects is unknown. From general principles one can formulate conditions on what $E_{XC}(\mathbf{r})$ should look like, and several, more and more advanced expressions have been advocated for it in the literature. Their application to real systems has been impressively successful, and it seems likely that a further increase of accuracy is a matter of time.

Basis functions and orbitals

Let us make a clear distinction between (basis) *functions* and *orbitals*, even where these phrases are sometimes mixed up in the traditional terminology. Orbitals are always specific combinations of the basis functions. Orbitals are related to the computed eigenfunctions of some Fock operator or Hamiltonian occurring in the run or in a related preceding calculation. Functions are merely the elementary mathematical entities in which the orbitals are expressed. A Slater Type Orbital (STO), for instance is a function, not an orbital.

The physical meaning of one-electron orbitals in DFT has often been questioned. We believe that they are useful quantities for interpretation, just like the HF orbitals. For a recent discussion see [2].

See also
ADF-GUI tutorial: basis set effects
ADF-GUI reference: basis sets

Cartesian function sets, spurious components

ADF employs Slater-type exponential basis functions centered on the atoms. Such a function consists of an exponential part $\exp(-ar)$ and a polynomial pre-factor $r^{kr}x^{kx}y^{ky}z^{kz}$. A function *set* is characterized by its radial behavior (the exponential part and the power of r, kr) and by its angular momentum quantum number *l*. The functions in such a set consist of all possible combinations $x^{kx}y^{ky}z^{kz}$, such that kx+ky+kz=*l*. These are denoted the *Cartesian* spherical harmonics.

The Cartesian function sets are very suitable for computational manipulations, but they have a drawback. By inspection it is easily verified that a *d*-set consists of 6 Cartesian functions, while there can of course be only 5 true *d*-type functions among them: one (linear combination) of them is in fact an *s*-type function ($x^2+y^2+z^2$). Similarly, there are 10 *f*-type Cartesian functions, 3 of which are in fact *p*-functions. And so on. In ADF all such lower-*l* (combinations of) functions are projected out of the basis and not employed. As a consequence the basis set *size* in the sense of the number of degrees of freedom and hence the number of possible eigenfunctions of the Fock operator is smaller than the number of expansion coefficients that refer to the primitive (Cartesian) basis functions.

The abbreviation BAS is used for references to the elementary Cartesian basis functions.

Frozen core: Core Orbitals and Core Functions

To speed up the computation the innermost atomic shells are kept frozen. The frozen Core Orbitals (CO), which are solutions of a large-basis all-electron calculation on the isolated atom, are expressed in an auxiliary set of (Slater-type) basis functions cor-bas, distinct from the valence set. The core basis set and the expansion coefficients that give the COs expressed in them are stored in the database data files.

Orthogonality of the valence Molecular Orbitals (MO) to the COs is achieved with the help of so-called Core Functions (CF). These functions are included in the valence set but they are not additional degrees of freedom. Each of the normal valence functions is combined with a linear combination of all CFs in the molecule in such a way that the transformed function (cbas) is orthogonal to all frozen COs in the molecule. There are exactly as many CFs as COs so the orthogonality condition for all valence basis functions amounts to the solution of a linear system where the number of conditions equals the number of parameters.

This aspect once more increases the discrepancy between the number of expansion coefficients of an MO and the number of MOs: the expansion coefficients in the most elementary bas representation run over all bas functions, including the CFs among them. At some places there may, alternatively, be expansions in the core-orthogonalized BAS functions, CBAS, where the CFs do not count anymore: they are included implicitly in the cbas functions.

Symmetry

The Overlap and Fock matrices become block-diagonal by using symmetry-adapted combination of the (C)BAS functions, such that each such function transforms under the symmetry operators as one of the subspecies of the irreducible representations (irrep) of the symmetry group. Symmetry adapted functions are denoted (C)SBAS.

For a given irrep and subspecies not all elementary basis functions can participate in the symmetry adapted combinations. For instance, for an atom in a reflection plane a basis function that is antisymmetric with respect to the reflection cannot be part of any symmetric combination of functions. In particular for higher symmetries the number of BAS functions that are relevant for a subspecies may be considerably smaller than the total number of BAS functions. This is used to cut down expansion lengths, both as used internally in the computation and construction of the Fock matrix, and in printed output. The printed expansion coefficients (in the bas representation) refer only to the participating BAS functions. A defining list of them is printed at an early stage of the run for each of the subspecies.

Orthonormal basis

It is often computationally convenient to use an orthonormal basis. This is constructed from the CSBAS basis by a Lowdin orthogonalization procedure. The resulting symmetry-adapted orthonormal basis is denoted low.

The MOs are computed by diagonalization of the Fock matrix in the LOW representation. The resulting eigenvectors are easily transformed back to any other representation whenever suitable, such as for instance to the primitive cartesian bas representation (including the CFs).

Fragments

Except in Create mode, where a *basic atom* is constructed, the system is built up from fragments and the corresponding fragment files are attached to the run. The program reads from the files the fragment MOs and these are used as (compound) *basis functions* for the molecular calculation. The fragment MOs are called Fragment Orbitals (FO).

FOs belong of course to one of the symmetry representations of the fragment, but not necessarily to a symmetry representation of the new molecule. The FOs are therefore combined into symmetry-adapted combinations, SFOs, to serve as a symmetry-adapted basis in the molecule. These combinations may involve one or more FOs from the same fragment and/or from different fragments. In the latter case the fragments must be symmetry related by one of the operators of the molecule. Symmetry related fragments must of course be identical, apart from their spatial location: they must be of the same fragment *type*.

FOs are naturally orthogonal to the Core Orbitals of their own fragment, but not necessarily to COs of other fragments. By a suitable combination of the SFOs with all CFs in the molecule we obtain the core-orthogonalized symmetry-adapted CSFOs.

The CSFOs can be transformed to an orthonormal basis by a Lowdin transformation. The resulting basis is called low, as above.

Summary of functions and orbitals

In Create mode the (conceptual) approach is:

BAS → (core-orthogonalization) → CBAS → (symmetry) → CSBAS → (orthonormality) → LOW → (Fock diagonalization) → MO

In Fragment mode:

FO (=MO from fragment file) → (symmetry) → SFO → (core-orth.) → CSFO → (orthonormality) → LOW → (Fock diagonalization) → MO

Acronyms:

`BAS`

Elementary cartesian basis functions, consisting of a radial part (exponential factor and power of r) and an angular part (cartesian spherical harmonic). The complete BAS set contains spurious lower-*l* combinations; these combinations are projected out and not used in the calculation. The BAS set contains also Core Functions.

`SBAS`

Symmetry-adapted combination of BAS functions.

`CF`

Core Function, part of the bas set. The CFs do not represent degrees of freedom in the basis set but serve only to ensure orthogonalization of the valence space to all frozen Core Orbitals.

`CBAS`

Core-orthogonalized elementary basis functions: the true valence (not-CF) BAS functions transformed by adding a suitable combination of the CFs. The total number of CBAS + the total number of of CFs equals the total number of BAS.

`CSBAS`

Symmetry-adapted combination of cbas functions.

`CO`

Frozen Core Orbitals, expressed as linear combinations of an auxiliary corbas basis set. The corbas set plays no role in the further discussion. The corbas functions are *not* the CFs.

The number of COs equals the number of CFs.

`LOW`

Lowdin orthonormalized symmetry-adapted core-orthogonalized basis. In Create mode they are derived directly from the BAS functions, in Fragment mode from the Fragment Orbitals, which are themselves of course expressible in the BAS set.

`FO`

Fragment Orbital: the MO of a fragment calculation, now used as a *basis function* in the molecule of which the fragment is part.

`SFO`

Symmetry adapted combination of FOs.

`CSFO`

Core-orthogonalized SFO.

Fit functions

Using Slater-type basis functions yields awkward multi-center integrals in the evaluation of the Coulomb potential. We therefore first need to find an approximate density-representation for which the Coulomb integral can be evaluated efficiently. This procedure is commonly referred to as density fitting. The default density fitting procedure in ADF is described in Ref. [379].

An alternative density fitting approach (STOFIT) employs an auxiliary set of *fit* functions, see also Ref. [308]. Like the basis functions, the fit functions are Slater-type exponential functions centered on the atoms. The true density, a sum of *products* of basis functions, is then replaced (approximated) by a linear combination (not products!) of the fit functions. The combination coefficients are called the fit *coefficients*.

$$\rho(r) = \sum_i c_i f_i(r) \quad (1.2.1)$$

The Poisson equation for the fit functions is easily solved, yielding the (approximate) Coulomb potential as an expansion in fit *potential* functions $f_i^c(r)$

$$f_i^c(r) = \int f_i(r') / |r-r'| \, dr \quad (1.2.2)$$

$$V_{\text{Coulomb}}(\rho(r)) \approx \sum_i c_i f_i^c(r) \quad (1.2.3)$$

In the SCF procedure the fit coefficients are computed by a least-squares minimization of

$$\int (\rho_{\text{exact}}(r) - \rho_{\text{fit}}(r))^2 \, dr = \min \quad (1.2.4)$$

with the constraint that ρ_{fit} contain the correct number of electrons. ρ_{exact} is defined as the sum of occupied orbitals (squared and multiplied by the appropriate occupation number). The accuracy of the fit approximation is important and the fit set plays a role similar to the basis set: too few functions (or badly chosen function characteristics) yield inferior results and there is also such a thing as the fit set limit. The fit functions on an atom are consequently an integral part of the definition of the basic atom and they are included in the Create data files. Fortunately, the size of the fit set does not determine the computational effort in such a drastic way as the size of the basis set does. We have chosen therefore to use always fair (though not extreme) fit sets, with the purpose that the effect of fit-incompleteness should in all cases be

small enough to be ignored compared with basis set effects, numerical integration errors and Density Functional deficiencies. This does of course depend somewhat on the computed molecule and the studied properties, so a general guarantee cannot be given and, as with basis set effects, one should always have an open eye for possible problems and check the pertaining information in the output file.

One of the most important properties of a molecule is its energy, or its bonding energy with respect to the constituent fragments. The fit incompleteness introduces two types of errors. The first is that, since the Coulomb potential is only approximated, the SCF solution itself, i.e. the set of self-consistent Molecular Orbitals and their energy eigenvalues may be slightly wrong, yielding an error in the charge density and hence in the energy. Since the energy is to first order stable with respect to changes in the mo coefficients this error in the energy can be assumed very small. The second type of error derives from the computation of the energy from the (self-consistent) charge density, via the Coulomb potential. Let

$$\rho \equiv \rho_{exact}(r) = \rho_{fit}(r) + \delta(r) \quad (1.2.5)$$

and

$$V_{fit}(r) = \int \rho_{fit}(r') /|r-r'|\, dr \quad (1.2.6)$$

For the Coulomb energy of the charge density we have

$$2E_{Coul}(r) = \iint \rho(r)\, \rho(r') /|r-r'|\, drdr' = \int \rho(r)\, V_{fit}(r)\, dr + \iint \rho(r)\, \delta(r') /|r-r'|\, drdr' =$$

$$\int V_{fit}(r)\, [\rho(r)+\delta(r)]\, dr + \iint \delta(r)\, \delta(r') /|r-r'|\, drdr' \quad (1.2.7)$$

from which we see that the fit error is corrected to first order (by adding the fit deficiency δ(r) to the exact charge density when integrating against the fit potential) and that only a second order term remains that cannot be evaluated, the last term in the right-hand-side of (1.2.7).

A fair impression of the fit quality and the importance of the second order error term is obtained by checking

a) the size of the first order correction term $\int V_{fit}(r)\, \delta(r)\, dr$ and *b*) the norm of the deficiency function, $\int \delta^2(r)\, dr$. Both are printed in standard output, at the end of the output of the SCF procedure computational report. They are usually very small, which gives some confidence that the second order fit error can be ignored.

Three-step build-up of the bonding

The approach of ADF is based on fragments. This applies not only in the analysis at the end of the computation but also in the set-up of the program. The computation of the molecule from its constituent fragments takes place in three steps, and these are reflected in the analysis of bond energy components.

First, the (free, unrelaxed) fragments are placed at their positions in the molecule. This implies an *electrostatic* interaction: for each fragment the Coulomb interaction of its undisturbed charge density with the fields of the other fragments.

Next, the Pauli exclusion principle is applied. Even without considering self-consistency the one-electron orbitals of the combined fragments cannot represent a correct one-determinant wave function because the orbitals of different fragments are not orthogonal to one another. The program performs an orthonormalization of the occupied Fragment Orbitals to obtain an antisymmetrized product. This implies a change in the total molecular charge density from the sum-of-fragments to what is called the sum-of-*orthogonalized*-fragments. The corresponding (repulsive) energy term is evaluated separately and is called *Exchange* repulsion, or alternatively *Pauli* repulsion. The phrase *orthogonal(ized) fragments*, if you find it elsewhere in this manual or in the source code of ADF, refers to this aspect. The sum of Pauli repulsion and electrostatic interaction is called the *steric interaction*.

The third phase is the relaxation to self-consistency, with of course the ensuing contributions to the bond energy.

Transition State procedure

This phrase stands for an analysis method described in ref. [3] and has no relation to transition states in chemical reactions. An extensive discussion of bond energy analysis by ADF is given in [4, 5]

The energy associated with a change in charge density, say the relaxation to self-consistency from the sum-of-orthogonal-fragments, can be computed by subtracting final and initial energies, each obtained from the corresponding charge density. For purposes of analysis the change in energy de can be reformulated as

$$dE = \int d\mathbf{r} \{ [\rho_{final}(\mathbf{r}) - \rho_{initial}(\mathbf{r})] \times \int_{\rho_{initial}}^{\rho_{final}} d\rho\, F[\rho(\mathbf{r})] \} \quad (1.2.8)$$

F(ρ) is the Fock operator belonging to the charge density ρ

By writing the density difference $\rho_{final} - \rho_{initial}$ a summation over contributions from the different irreducible representations Γ of the molecular symmetry group, an expression is obtained that lends itself for a decomposition of the bond energy into terms from the different symmetry representations:

$$dE = \sum_{\Gamma} \int d\mathbf{r} \{ \Delta\rho^{\Gamma}(\mathbf{r}) \times \int_{\rho_{initial}}^{\rho_{final}} d\rho\, F[\rho(\mathbf{r})] \} \quad (1.2.9)$$

The integral of the Fock operator over the charge density is now approximated by a weighted summation (in fact, a Simpson integration):

$$\int_{\rho_{initial}}^{\rho_{final}} d\rho\, F[\rho] \approx 1/6\, F(\rho_{initial}) + 2/3\, F(\rho_{average}) + 1/6\, F(\rho_{final}) \quad (1.2.10)$$

$$\rho_{average} = 1/2\, \rho_{initial} + 1/2\, \rho_{final} \quad (1.2.11)$$

The term with the Fock operator due to the average charge density has given rise to the phrase *transition state*. To avoid confusion we will often refer to it as to the transition *field.*

The approximate integral (1.2.10) involves two errors. The first, rather obvious, is the approximation of the exact integral in (1.2.9) by the weighted sum in (1.2.10). Except in pathological cases this approximation is highly accurate.

The second error comes from the fact that the Coulomb and XC potentials in the Fock operator are computed from the *fit* density. This is only an approximation to the true density, while in the original bond-energy expression (energy due to the final density minus energy due to the initial density) no potentials occur and the *exact* charge density can be used. As mentioned before, these fit-related errors are usually small. For the XC potential the true density can be used if one includes the keyword `EXACTDENSITY`.

All such errors in the total bonding energy are easily corrected by comparing the summation over the Γs with the correct value for the total bonding interaction term. The difference is simply added to the total bond

energy, so no true error remains. We only have a (correction) term that can't be split in contributions from the distinct symmetry representations. In the printed bond energy analysis such small corrections are 'distributed' over the other terms by scaling the other terms such that their sum is the correct total value.

1.4 Running the program

Execution of ADF

When ADF has been installed you can run it by supplying appropriate input and starting the 'adf' script, located in $ADFBIN. This script sets up some environment variables, and parses the input to see if anything special needs to be done. For example, if the BASIS key is used the adf script will also execute commands to make the appropriate fragment files. You can use this run script both for the serial and parallel versions of the program. For other programs in the package, there are similar run scripts ('band', 'dirac', and so on).

Running the program using the run script involves the following steps:

- Construct an ASCII input file, say in.
- Run the program by typing (under UNIX):
  `$ADFBIN/adf {-n nproc} <in >out`
  The part between curly brackets is optional, so the shortest application has the format
  `$ADFBIN/adf <in >out`
  Note that the run files in the $ADFHOME/examples directory are UNIX scripts which are executes with:
  `run >out`
- Move / copy relevant result files (in particular TAPE21) to the directory where you want to save them, and give them appropriate names.
- Inspect the standard output file out to verify that all has gone well.

During the run you may inspect the logfile, to see how far the program has proceeded, or whether you should interrupt the calculation.

In the above scheme adf is the name of the run script that invokes the adf.exe program executable. During the installation the script has been put in the same directory where the program executables are generated: $ADFBIN. You may have moved it to another place, or renamed it. We recommend that you adjust your $PATH variable so that you can omit $ADFBIN from the execution command.

To run another program from the ADF suite, just use the appropriate program run script.

The input for the program is read from standard input, and output is written to standard output, redirected in the example above to in and out, respectively.

The part between square brackets is optional and is only meaningful for a parallel program version. The -n flag specifies the number of parallel processes (nproc) to use. If omitted the default applies, which is the value of the environment variable $NSCM, if such variable exist, otherwise it is defined by installation parameters (in the $ADFHOME/settings file, see the Installation Manuals).

The program run scripts have, in fact, more flags and arguments, for special usage. You can get a survey by typing

```
$ADFBIN/adf -h
```

Parallel execution

If a parallel version of ADF has been installed you should be aware of a few special aspects of running ADF in parallel. Partially this depends on the platform and on the installation settings.

First of all, you may specify (by command-line options in the run-script and/or by defining suitable environment variables) explicitly how many parallel processes are to be used. Secondly, you should realize that most of the files that you would have in a single-node run are in a parallel run distributed over the parallel processes. Some parts of the file may be identical across the processes while other parts are not and would only after a recombination yield the data of the corresponding single-node file. The normal result files, (standard output, the logfile and the binary result file TAPE21) are complete at the master process.

How to set up a parallel calculation can be found in the Installation Manual.

Files

The ADF program may generate several output / result files, along with the standard output file. The most important one is TAPE21 (.t21 file), the general result file. TAPE21 contains relevant information about the outcome of the calculation. You may want to store this file somewhere under an appropriate name, for future usage. The meaning of any other files that are produced are explained later in this User's Guide.

Any files produced by the program are generated in the local (working) directory where the calculation runs. If you want to keep them, make sure to move them after the calculation has finished to wherever you want to store them.

Files attached to the job, such as fragment files, are by default also assumed to exist in the local directory. You must take care to move or copy required files to that directory before starting the calculation, or to provide via input adequate information to the program where to find the files. In many cases you can specify a complete path to the file.

Most files that are generated by the program, in particular the standard result file that can be used as a fragment file in other calculations, are *binary* files. A binary file should usually not be moved from one machine to another, i.e. it may not be readable by another machine than the one that generated the file, unless the two machines are of the same type. The ADF package provides utilities to convert the ADF binary result files from binary to ASCII, and vice versa, so that you don't have to regenerate your fragment libraries when going to another machine. See Appendix 5.5

TAPE21 and logfile

Two of the files that are produced by ADF deserve special attention. The first is the general result file TAPE21 (.t21 files). It is a binary file that contains a lot of information about the calculation, such as the one-electron orbitals expressed in the basis functions. It can be used as a fragment file for subsequent calculations (although only TAPE21 files from spin-*restricted* calculations can be used as fragment files). Like all files produced by the program, it is generated in the directory where the job runs. Having done a calculation, you will usually store TAPE21 somewhere under a suitable name so that you can later reuse it, as a fragment file, for a restart, to feed it to an analysis program, and so on.

The second is an ASCII log file, called logfile. It accumulates messages from ADF into a (brief) summary of the run. You can inspect it during the calculation to check how far the calculation has proceeded, whether there are important warnings and so on. At the end of the run this file is copied to the tail of the normal standard output file.

Standard output

ADF is a program that lends itself particularly well for chemical analysis. This is a direct result of the fragment-based approach, where properties of the molecule are related to the properties of the constituent fragments, which is precisely how the chemist thinks. Molecular Orbitals are (optionally) analyzed extensively as how they are composed from occupied and virtual fragment orbitals. This inherently implies a large amount of output. Even computations on small molecules may produce startlingly many pages of output. This is not necessarily so because you can regulate the production of output in detail. Obviously,

some kind of *default* production of output had to be implemented. The field of ADF users is so wide and diverse that it is hard to satisfy everybody as regards this default level of output. Depending on your purposes the automatic settings, which determine how much output is generated without instructions to the contrary, may yield boringly many numbers that you just skip through in search for the one value you're interested in, or it may be widely insufficient. Therefore, take notice of the possibilities to regulate output.

Above all, however get familiar with the analysis tools that ADF provides to see in what ways these may help to interpret your results. In a later chapter a global description of output is given as it is normally produced. The chapter below gives an introduction in some of the essential features of ADF, which may be sufficiently different from what you are used to in other Quantum Chemistry codes to deserve your attention.

File names during parallel runs

The adf proces of a kid normally runs in a separate directory.

Standard output is a special: the parent writes its normal ('print') output to standard output while the *kids* each write to a file KidOutput.

1.5 ADF-GUI

The graphical user interface ADF-GUI enables all users to set up complicated calculations with a few mouse clicks, and provides graphical representations of calculated data fields, see the ADF-GUI overview tutorials, and advanced ADF-GUI tutorials.

2 INPUT

2.1 Minimal input

We will start now with a discussion of the input file for ADF. First a minimal input is discussed, next an extensive list of all input options is described.

Most keys in the input file for ADF are optional. Default values apply for omitted keys. Assuming that the defaults are sensible, short input files can often be used. We will examine first the minimal input that is required to run ADF. Having read that part, you can start to do calculations.

The following input will run a geometry optimization on water, using a (almost) minimal input:

```
ATOMS
 O  0  0  0
 H  1  1  0
 H -1  1  0
End

Basis
End

Geometry
End
```

This is the input for the ADF program. You need to store it in a file, and pass it as standard input to ADF.

For example, assume you have stored the above input in a file in. Also assume that the $ADFBIN directory is in your $PATH. Then you run ADF using the following command:

```
adf <in >out
```

ADF will run, and the resulting output will be stored in the file out. If you examine the contents of this file, you will find that ADF has actually run three times: two create runs, and one geometry optimization. The fragment files produced by the create runs are saved in t21.H and t21.O, for hydrogen and oxygen respectively.

2.2 Structure of the input

Much of the general remarks about input for ADF apply also to related property and analysis programs, which are also described in this document.

Delimiters

An input record may contain several items. The general rule is that each sequence of characters that does not contain a delimiter is an entity. Delimiters in this context are: 1) the blank or space character ' ', 2) the comma ',' and 3) the equal sign '='.

It is assumed throughout that only characters of the Fortran character set are used.

DO NOT USE TABS IN THE INPUT FILE! The program may *not* see them as delimiters and the effects are hard to predict!

Uppercase and lowercase

Virtually all input items are case-insensitive, but take notice of the obvious exceptions: names of files and directories are case-sensitive.

Keywords

Input for ADF is structured by keywords, in short: keys. A key is a string of characters that does not contain a delimiter (blank, comma or equal sign). Keys are not case sensitive. Input is read until either the end-of-file condition (*eof*) becomes true, or until a record `end input` is encountered, whichever comes first. (`end input` is not a key.)

Key-controlled input occurs with two formats. In the first you have only one record, which contains both the key and - depending on the case - associated data: the key *argument*:

```
KEY argument
```

The whole part of the line that follows after the key is the argument. It may consist of more than one item.

The alternative format is a sequence of records, collectively denoted as a key *block*. The first record of the block gives the key (which may have an argument). The block is closed by a record containing (only) the word end. The other records in the block constitute the *data block*, and provide information related to the key.

```
KEY {argument}
 data record
 data record
 ...(etc.)...
 ...
end
```

In this manual, when items are optional, such as the argument in the scheme above, they are typed enclosed in curly brackets {}. The { and } characters themselves are not part of the item. Different allowed / eligible values are separated by a bar (|). The keywords are usually typed in small capitals, subkeys in *italic small capitals*.

Structures like 'key=value' should be read as: type 'key=' as such, followed by a suitable value.

Block type keys may have subkeys in their data block. The subkeys may themselves also be block type keys. The data blocks of block type subkeys, however, do not end with end, but with subend:

```
KEY {argument}
 data
 data
 subkey {argument}
    subkey data
    subkey data
    ...
 subend
 data
 data
 ...
end
```

Layout features such as an open line, indentation, or the number of spaces between items are not significant.

The format to be used for a key is not optional: each admissible key corresponds to one specific format. As a general rule, the block keys control lists of data, such as atomic position coordinates.

A few special keys can have either format. For such keys the format actually in effect depends on the presence of the argument: the block type applies in absence of the argument. The block type applies also when an argument is present that ends with a continuation symbol. The continuation symbol is the ampersand (&) or, alternatively, two contiguous plus-characters preceded by at least one blank (++):

```
KEY {argument} &
 data
 data
end
```

The various types of keys are referred to respectively as *simple* keys, *block* keys, and *general* keys.

A considerable number of keys can be used to specify the geometry, the model Hamiltonian, cf. the Density Functional, the precision of the calculation, and so on. The order in which keys occur in the input file is immaterial, except that a few special keys determine how input data is interpreted, such as the unit-of-length for atomic coordinates. These *interpretation keys* must be used before the pertaining data in input occur. This will be mentioned explicitly again where they are discussed.

The items that can be addressed with keys and the keys themselves are listed in the Index.

Irrelevant keys, misspelling of keys

Specification of a key that is not relevant in the calculation will go unnoticed. Similarly, if you misspell a key such that it is not recognized, the incorrectly labeled input data will be ignored and the program will proceed as if the intended key had not occurred. This results in the application of pre-defined default values or in an error abort, depending on the case. Therefore, whenever the output suggest that part of your input has been ignored, check the spelling.

In this context we stress again: be alert on TAB characters: don't use them at all.

ADF may recognize a key if it is spelled incompletely, that is, if only some initial substring is given, and also if redundant characters are typed after the end of the key. The reason is that often only a small initial part of the true keyname is checked against the input items. Don't rely on this, however: it is not formally supported and it may get disabled in a next release without further notice.

We advise therefore to stick to the correct key names. In particular, you must avoid to use different abbreviated or elongated forms when a key occurs more than once in input: ADF will likely assume that you want to indicate distinct keys and it will associate only one of them with the key you had in mind.

Interpretation of Input

ADF has two special keys that regulate the specification and interpretation of numerical data in input. These keys, and related aspects, are convenient for the formatting of input.

The position of the interpretation keys in the input file is significant! Therefore, to avoid problems and misunderstandings: before supplying any numerical data, specify first (if at all) the keys units and define (see below).

Units of length and angle

Geometric lengths and angles are in units defined by:

```
UNITS
 length Angstrom / Bohr
 angle Degree / Radian
end
```

Angstrom and Bohr, respectively Degree and Radian, are recognized strings. Each of the subkeys is optional, as is the key UNITS itself. Defaults: Angstrom for lengths, and Degree for angles.

The position of the key UNITS in input is significant as regards the evaluation of expressions (see the paragraph on constants and functions below). In other respects its position plays no role. To avoid mistakes one should place units as early as possible in input (if at all).

Expressions

ADF supports the use of arithmetic *expressions*, *functions*, and *constants* to represent numerical data. This can be convenient for the input of, for instance, atomic positions when these would most easily be represented in terms of 1/3, sin(360/5), et cetera. Using expressions and functions is easier, avoids the tedious typing of long decimal expansions and solves the question of precision (how many digits should be supplied?).

The standard arithmetic operands in Fortran (+ - * / **) can be applied in expressions, together with parentheses where suitable.

Blanks are allowed and ignored, but they are interpreted as separators, i.e. as denoting the end of an expression, whenever the part until the blank can be evaluated as a correct expression. For instance 3* 4 will be interpreted as 12, but 3 *4 will be interpreted as 3, followed by a character *, followed in turn by the number 4.

All numbers and results are interpreted and handled as being of type real, but whenever the result is a whole number (allowing for very small round-off) it will be recognized and accepted as an integer when such data is required.

Constants and functions

The user may define *constants* and *functions* in the input file, and apply them subsequently in expressions. The input file is read sequentially and *constants and functions must be defined before they can be used*.

The argument list of a function must be enclosed in parentheses and the arguments, if more than one, separated by commas.

The following functions are predefined in ADF and can be used directly in input:

sin, cos, tan, asin, acos, atan, exp, log, sqrt, nint. Each of them has one argument. log is the natural logarithm (base *e*).

No constants are predefined.

The angular argument to the trigonometric functions cos, sin, tan is in the unit for angles as defined by units, *provided the unit has been set before it is applied*. For the result of the inverse trigonometric functions the same holds.

Constants and functions can be defined with the block key DEFINE:

```
DEFINE
 angle=54
```

```
 ab = sin(angle/3)
 s13 = 14*sqrt(2)
 func(x,y,z) = x*ab+y**2-y*z
end
```

The constants angle, ab, and s13 are defined together with a function func, using the predefined functions sin and sqrt. These can then be applied to assign values elsewhere in input.

In the example above, the constant angle is used in the definition of ab, and ab is used in turn to define func; these constructions are allowed because angle is defined before ab, and ab is defined before func.

The replacement of constants, functions, and other expressions by their numerical values may considerably increase the *length* of the input record, in particular when real values are being generated (by the parser) in the standard format E22.14. Take care that the resulting record does not exceed 80 characters. The program will abort or may run into an error if this is violated.

The input-reading routine applies the constants and functions wherever it is allowed to do so. To prevent any unwanted replacements in the input file you should avoid very short identifiers for constants and functions.

Warning example:

```
DEFINE
 A=3.18
 C=4.12
end
...
atoms
 C 0.00 1.05 -3.22
...
```

The program will apply the definition of the variable C and read:

```
DEFINE
 A=3.18
 C=4.12
end
...
atoms
 4.12 0.00 1.05 -3.22
...
```

Avoid single-character identifiers!

Strings

Quotes can be used to designate strings, i.e. (parts of) records which are not to be parsed for expressions, but which should be taken as they are. The quotes themselves are ignored, i.e. removed by the parser. Two consecutive quotes inside a string are interpreted to denote the (single) quote character as a part of the string.

Where does parsing apply?

Replacing pre-defined variables and expressions by their value is applied only to keys that carry numerical data. For example: atoms, define, units. However, it is *not* applied to keys that carry electronic occupation numbers.

Note that when parsing applies to a given key the whole record of the key (key + argument) and its data block are parsed. The parsing then applies to all items, even those that in themselves have no numerical meaning (for instance, the atom type names in the atoms data block are scanned and must of course then not be 'defined' as identifiers with a numerical value.

Constants vs. geometric parameters

Note carefully the difference between constants defined with define and identifiers that are used for atomic coordinates in the data blocks of atoms and geovar. Constants defined under define are merely symbols for, and exactly equivalent to, certain numerical values, whereas the coordinate identifiers carry implications such as the distinction between frozen and optimization coordinates. *Constants* affect only the input *after* their definition and the location of their definition in the input file is significant. Geometric *identifiers* only relate to the data blocks of atoms and geovar respectively and the relative order in which the keys atoms and geovar occur is irrelevant.

Link-in Input files

Part of the input file can be put into a separate ASCII file, which can be addressed from the (standard) input stream:

```
INLINE inlinefile
```

inlinefile must be the name of the auxiliary ASCII file (including its path, absolute or relative to the run-directory). When inline is encountered in the input file, ADF opens the specified file and continues reading from that file as if it were in-line expanded into the input file. When the end-of-file is encountered reading resumes from the original file.

The contents of the inlinefile must *not* end with end input, unless you wish to terminate all input reading at that point.

InLine may occur any number of times in the input file. Use of inline may also be nested (up to 10 levels): the key INLINE may be used in the inlinefile in the same fashion as in the standard input file.

The inline feature makes it easy to pack your preferred settings that are not matched by the program's defaults in one file and use them in every run with a minimum of input-typing effort. Obvious applications are output control (print) settings and precision parameters.

Note: you can *not* use inline to store parallel settings, not even by using inline on the first line of your input and placing the parallel keyword on the first line of the inlinefile: before opening the inlinefile and expanding it into the inputfile, the program has already detected that the first line of input does not specify the parallel settings.

Title, comment, layout of input

```
TITLE Title
```

Title may be any string. The program combines it (that is, the first approximately 50 characters) with date and time of the job to construct the *job identification.* The job identification is used to stamp an identification on result files, which will be read and printed if such a file is used again, for instance as a fragment file.

The job identification will also be echoed in the output header to identify the current run. By default the date and time are combined with a dummy string. In Create mode the title is first read from the data file that supplies the basis functions etc and can then be overwritten via input.

Note that, contrary to some other programs, ADF does *not* take the first input record as a title. Typing your title as the first record, *without starting the record with the keyword* title, may produce very strange results: ADF will try to interpret the first word on that line as a keyword, possibly abbreviated!

You can put more remarks in the input file to be echoed in the standard output file; these will not become part of the job identification:

```
COMMENT
  text
  ...
end
```

The text records are copied to the output header, directly after the job identification. Expressions are not parsed and constants or functions are not replaced: it is a straightforward copy.

The key COMMENT may occur any number of times; all text blocks are printed in the output header with a blank line between any two text blocks.

Layout of input

Empty records and leading blanks in records are allowed and ignored, and can be used to enhance clarity and readability of the input file for human readers.

An exclamation mark (!) is interpreted by the input reading routine as denoting the end-of-line. Instead of the exclamation mark you may also use a double colon (::). The part of the line after the exclamation mark (double colon) - including the ! or :: itself - is ignored. In this way one can include comments and clarifying remarks, which will not be echoed in the output header (compare the key COMMENT).

2.3 Coordinates, basis sets, fragments

See also
ADF-GUI tutorial: building molecules, basis set effects, fragments
GUI manual: basis sets, ghost atoms, nuclear model

Atomic coordinates

The input of (initial) atomic positions as Cartesian coordinates has been mentioned already in the minimal-input example in Chapter 2.1. Alternatively they may be given in z-matrix form.

```
ATOMS {Cartesian / Zmatrix / MOPAC}
 {N} Atom Coords {F=Fragment}
 ...
End
```

`Cartesian or Zmatrix or MOPAC`

Specifies the type of coordinates. Default (no specification) is Cartesian. Instead of Zmatrix you may also type internal.
MOPAC is a special variety: the subsequent records in the data block are MOPAC style Z-matrix input for the atomic system, see example below.

`N`

This is an optional integer by which you may number the atoms. The numbers should be 1,2,3, et cetera if any reference is made to them in other parts of input. The reason for this restriction is that ADF

numbers the atoms internally according to their occurrence in the input file and it applies this internal numbering when any subsequent references are interpreted.

`Atom`

The name of an *atom type*. It must begin with the standard one- or two-character symbol for the chemical element: H, He, Li, and so on. Optionally it may be appended by .text, where text is any string (not containing delimiters). Examples: H, Mn.3, Cu.dz-new.

Dummy atoms may be useful in the construction of a Z-matrix, for instance to obtain a set of internal coordinates that reflect the symmetry of the molecule better. They may also be useful in a Z-matrix to avoid an ill-defined dihedral angle, which occurs when three (almost) co-linear atoms span either of the two planes that define the angle. In geometry optimizations this must absolutely be avoided if such internal coordinates are used as optimization parameters.
Dummy atoms are input with the chemical symbol xx. XX-type atoms can be inserted in the list of atoms like any other atom types. The name (xx) can have a suffix of the form *.text.* No fragment files must be supplied for dummies. There are no symmetry constraints on the positions of the dummies. The dummies serve only to set up the Z-matrix in a proper way.

`Coords`

This specifies he coordinates of the atom. If Cartesian coordinates are used the x, y, z values must be given. For Z-matrix coordinates you put first the three connection numbers, then the values of the bond length, bond angle and dihedral angle. Example:

Ge 2 1 5 2.1 95.3 24.8

defines that a Germanium atom is located with a distance 2.1 Angstrom from the second atom in the input list, that the angle (Ge-atom2-atom1) is 95.3 degrees and that the dihedral angle between the planes (Ge-atom2-atom1) and (atom2-atom1-atom5) is 24.8 degrees.

To avoid any confusion as regards the direction (sign) of the dihedral angle, here is the definition used in ADF: Let the connection numbers for an atom P refer to the atoms Q, R and S, in that order. Choose a local coordinate frame such that Q is at the origin, R on the positive z-axis and S in the xz-plane with a positive x-value. The three Z-matrix coordinates bond length, bond angle and dihedral angle of P are then precisely its spherical coordinates r, q, and -f: the distance to the origin, the angle that PQ makes with the positive z-axis (0.π) and the *negative* of the angle that the projection of PQ on the xy-plane makes with the positive x-axis (0..2π, or -π..+π).

The connection numbers and internal coordinate values of the first atom in a Z-matrix have no meaning. Similarly, the second atom requires only a bond-length specification and the third atom only a bond length and a bond angle. However, for each atom three connection numbers are read from input and interpreted, and you must therefore supply *zeros* for them if they don't refer to any atoms. The corresponding meaningless Z-matrix *coordinate* values can be omitted. More in general: missing coordinate values are set to zero (also for Cartesian coordinates input). Z-matrix values that are meaningless because they correspond to zero connection numbers are ignored, whatever their value is in the input file.

In a Z-matrix definition the three reference atoms, with respectively 3, 2, and 1 connection numbers equal to zero, do not have to be the first three in the input list. The program will scan the list for any atom that has 3 connection numbers zero, then for one that has only a bond length specification, etc. If the Z-matrix is not properly defined, for instance if more than one atom occurs with all three connection numbers equal to zero, or when not every atom is somehow connected to all others, the program will abort.

`F=Fragment`

Specifies that the atom belongs to a particular fragment. The fragment name must be of the form fragtype/n, where fragtype is the name of one of the types of fragments in the molecule. The integer n, after the slash, counts the individual fragments of that type. The numbering suffix /n is not required if there is only one fragment of that type.

When f=fragment is omitted altogether, the fragment type is taken to be the *atom type* that was specified earlier on the same line. (The numbering /n is then added automatically by the program, by counting the number of times that this single-atom fragment type occurs in the list of atoms).

`Mopac`

The MOPAC style input requires that the records in the data block have the following format:

atomtype distance idist angle iangle dihedral idehedral

The three internal coordinate values (distance, angle, dihedral) are each followed directly by the connection number.

Atom type is not identical to *chemical element*: an atom type is defined by all characteristics of the basic atom to which it in fact refers: the nuclear charge, the basis functions, the frozen core, the density functional and any other features that were applied in generating that basic atom.

As mentioned before, the point group symmetry specified in input with a Schönfliess type symbol puts restrictions on the orientation of the atomic system. Unless the input-specified symmetry equals the true symmetry of the nuclear frame (in which case ADF will adjust the orientation of the molecule, if necessary), the user must take care of this by supplying the *Cartesian* coordinates (in the appropriate orientation). If a subgroup of the true nuclear symmetry is used and Z-matrix format is used for the coordinates, the program will place the atoms in the standard Z-matrix frame: first atom at the origin, second on the positive x-axis, third in the xy-plane with positive y-value.

Dummy atoms may be placed asymmetrically. If the atomic coordinates are input as Cartesians, any dummy atoms are irrelevant. Their coordinates will be printed but otherwise they are ignored.

Input items are generally case insensitive. Exceptions are the names of files and directories. Since (to be discussed below) the name of the fragment type as it is defined under atoms (explicitly with the f=option, or implicitly as the name of the atom type) might also directly indicate the fragment file, the specification of fragment types is in principle case-sensitive. Errors may occur if you are sloppy in this respect.

However, you must not give different fragment types names that differ only by case: at various places in the program fragment type names are compared in a case-insensitive way.

Mixed Cartesian and Z-matrix coordinates

The key ATOMS can also be used to supply coordinates in a format that gives the values for the cartesian coordinates *and* the connection matrix, which defines a Z-matrix.

```
ATOMS ZCart
 {N} Atom Coords {F=Fragment}
 ...
End
```

`ZCart`

Signals this particular format for the coordinates

`Coords`

As for Z-matrix input: three integers and three real values. The integers are the connection numbers that define the Z-matrix structure, but the reals are the *Cartesian* coordinates.

With ZCart input, the z-matrix is internally generated from the Cartesian coordinates and the connection numbers.

This feature is convenient when for instance Cartesian coordinates are easily available but you want to run a Geometry Optimization in *internal* coordinates, for which a Z-matrix structure is required.

The zcart option comes in handy also to satisfy symmetry-related orientation requirements when you basically wish to use Z-matrix coordinates.

With zcart input the program defines the *type* of coordinates in the input file as *Cartesian*. This is significant in Geometry Optimizations, where the optimization variables are by default taken as the input coordinate type.

Orientation of Local Atomic Coordinates

As discussed before the atomic positions are input with the key ATOMS. One option has thus far not been mentioned: the possiblity to redefine the local coordinate frame of an atom.

```
ATOMS {type of coordinates}
  {n} atomname coordinates {F=fragment} {Z=xx yy zz} {X=xx yy zz}
  ...
end
```

Except for the z= option all aspects have been examined already before.

`z=xx yy zz`

defines a reorientation of the local atomic z-axis; it is interpreted as a direction vector with components (xx,yy,zz) pointing away from the atom. In the local, reoriented frame the local atomic x-axis will be rotated to the plane defined by the directions of the molecular z-axis and the local atomic z-axis.
This feature can be used only for single-atom fragments (otherwise it is ignored). Its purpose is to give more flexibility in the analysis of the final molecular orbitals in terms of the atomic orbitals. In such a case it may be very helpful to redefine the orientation of say the p-orbitals of an atom. For instance, you may orient all p-orbitals towards the origin by specifying for each atom z= -x -y -z (with x,y,z the coordinates of that atom).
By default the local and molecular z-axes are identical.

`x=xx yy zz`

defines a reorientation of the local atomic x-axis; it is interpreted as a direction vector with components (xx,yy,zz) pointing away from the atom. Together with the z vector this defines the xz plane. The y axis is then given by the vector product z * x.
This is used for analysis (see orientation of the z-axis).

ASCII Output Files with Atomic Coordinates

You may want to have a special result file that contains the atomic coordinates corresponding to all the geometries processed in the calculation, for instance to feed it to a 'movie' generator to display the development of an optimization run. This is regulated with the key FILE:

```
FILE filetype filename { filetype2 filename2 }
```

`filetype`

Specifies the format of the output. Currently supported are three varieties: MOPAC, mol and xyz

`filename`

The file to which the output is written; the file should not yet exist. The name may include a full or relative path with respect to the directory where the calculation runs.

The same input record may contain any number of pairs-of-arguments, for instance to specify that both a mol-type *and* a xyz-type result file are to be generated. The key may also occur more than once in the input stream, in which case the argument lists are effectively all concatenated (by the program).

Basis sets and atomic fragments

Database of STO basis sets

The ADF package is equipped with a database to help you generate basic atoms. Each data file in the database contains a standard basis set (and related information) for the creation of one basic atom. The data files are relatively small ASCII files. You can easily inspect them. In Appendix 5.1 a definition is given of such a file. This enables you to create variations and construct your own adapted basis sets.

The basis functions used in ADF are commonly known as Slater Type Orbitals (STOs). A basis set can roughly be characterized by its size (single-, double-, triple-zeta; with or without polarization) and by the level of frozen core approximation. Initially, the only basis sets provided with ADF were those in the directories I, II, III, IV, V, which now have the more intuitive names SZ, DZ, DZP, TZP, and TZ2P, respectively. The increasing numbers point to an increase in size and quality. It is not possible to give a formally correct short general classification for each basis set directory. However, generally speaking we can say that SZ is a single-zeta basis set, DZ is a double zeta basis set, DZP is a double zeta polarized basis, TZP is a core double zeta, valence triple zeta, polarized basis set, and finally TZ2P is a core double zeta, valence triple zeta, doubly polarized basis. This explains the more intuitive names that are given for the basis sets. The names have also been changed since some of the basis sets have been modified substantially.

For small negatively charged atoms or molecules, like F^- or OH^-, use basis sets with extra diffuse functions, like they are available in the AUG or ET/QZ3P-nDIFFUSE directories, see description below. For example, the standard basis sets, or even the large ZORA/QZ4P basis set will often not be large enough for the accurate calculation of such anions.

In addition to the standard basis sets, the database contains directories with special basis sets:

TZ2P+For transition metals Sc-Zn and lanthanides (ZORA) only: as TZ2P, but with extra d-STO (3d metals), and extra f-STO (lanthanides, ZORA)

ZORA contains basis sets that should be used (exclusively) for relativistic calculations with the ZORA approach. Using 'normal' basis sets in a ZORA calculation may give highly inaccurate results, in particular for heavy elements. The same is classification is used for the directories ZORA/SZ-TZ2P as in the non-relativistic directories. The ZORA basis sets were added later because of the special requirements on basis sets for ZORA relativistic calculations, especially in the core region. The ZORA/QZ4P basis set can be loosely described as core triple zeta, valence quadruple zeta, with four sets of polarization functions.

ET contains several even tempered basis sets which enables one to go to the basis set limit, such as ET/ET-pVQZ, ET/ET-QZ3P, ET/ET-QZ3P-1DIFFUSE, ET/ET-QZ3P-2DIFFUSE, ET/ET-QZ3P-3DIFFUSE. The accuracy of the smallest basis set in this directory can loosely be described as quadruple zeta in the valence with three polarization functions added. This directory also contains basis sets with extra diffuse functions. In Response calculations one should use such large basis sets. Very

diffuse functions are absolutely necessary to get good results for excitation energies corresponding to high lying orbitals.

AUG contains several augmented standard basis sets which enables one to get reasonable results for excitation energies with relatively small basis sets, such as AUG/ASZ, AUG/ADZ, AUG/ADZP, AUG/ATZP, AUG/ATZ2P.

OLD contains basis sets that were contained in the 2.3 release, but that we feel should not be used anymore unless with great care, primarily because the involved frozen cores are too large to justify the frozen core approximation. In some cases we found uncomfortably large errors in equilibrium geometries resulting from such too-large cores.

Furthermore, you will find in the database:

Special/AE contains non-relativistic basis sets for all-electron calculations. However, these files cannot be used as such, because they don't contain any fit sets. Using basis sets without fit sets is pointless and is in fact not possible at all. (The usage and relevance of fit functions is explained later). Therefore, they serve as starting point for the development of (new) basis sets. For some of the all-electron sets appropriate fit sets have already been generated. The corresponding data base files can be found in the appropriate subdirectories SZ, DZ, DZP, et cetera.

Special/Vdiff contains non-relativistic basis sets that include very diffuse functions. These were recommended to be used for Response calculations. Very diffuse functions are absolutely necessary to get good results for excitation energies corresponding to high lying orbitals. Recommendation: use the even tempered basis sets in the ET directory, since these basis sets are better.

Special/MDC contains non-relativistic basis sets with optimized fit functions especially useful for accurate Multipole Derived Charges. These are available only for a limited number of basis sets.

Cerius contains data files that are used in the Cerius2-ADF graphical user interface.

Dirac contains the input files for the DIRAC auxiliary program (see the RELATIVISTIC keyword).

Band contains input files for the BAND program (see the BAND User's Guide)

ForceFields contains force field files to be used in the QM/MM functionality. Their structure and contents are described in the QM/MM manual. See also the pdb2adf utility (ADF QM/MM documentation), which transforms a PDB file into an ADF input file, for use with QM/MM.

The files in SZ/) (and ZORA/SZ/) have minimal basis sets: single-zeta without polarization. The exponents of the functions correspond to the standard STO-3g basis sets used in programs that employ Gaussian type basis functions. The frozen core approximation is applied, however, for the inner atomic shells. Type-SZ database files are provided only for the lighter elements, up to Kr.

The files in DZ can be characterized as double-zeta basis sets without polarization functions. A triple-zeta set is used for the 3d shells of the first row transition metals, the 4f shells of the Lanthanides, and the 5f shells of the Actinides. In all these cases a double-zeta set provides a rather poor expansion basis for the true (numerically computed) atomic orbital.

The basis sets in DZP are derived from DZ, extended with a polarization function. This type of basis sets is thus far provided only for the elements up to Ar, and for the 4p series Ga through Kr.

TZP contains triple-zeta basis sets. A polarization function is added for H through Ar and for Ga through Kr (from DZP).

TZ2P finally gives extended basis sets: triple-zeta with two polarization functions, for H through Ar and Ga through Kr (from DZP). Note that the TZ2P database files are provided only for the lighter elements, up to Kr. The ZORA/TZ2P database files are provided for all elements. Typically for all elements one polarization

function is added compared to the corresponding TZP basis set. Note, however, that TZ2P will not always give you extra basis functions for most lanthanide and actinide frozen core basis sets.

Multiple occurrences of one chemical element in the same basis set subdirectory correspond to different levels of the frozen core approximation. Manganese for instance may have a basis set for an atom with a frozen 2p shell and another one with a frozen 3p shell. The file names are self-explanatory: Mn.2p stands for a data file for Manganese with frozen core shells up to the 2p level. An all-electron basis set would correspond to a file that has no frozen-core suffix in its name.

Another type of multiple occurrence of one element in one database directory may be found when basis sets have been developed for different electronic configurations: the Slater-type basis sets are fitted then to numerical orbitals from runs with different occupation numbers. Currently this applies only for Ni (in database directories DZ, TZP and TZ2P), where basis sets are supplied for the d8s2 and the d9s1 configurations respectively. Since in earlier releases only the d8s2 variety was available, the names of the database files are Ni.2p (for d8s2) and Ni_d9.2p, and likewise Ni.3p and Ni_d9.3p.

As mentioned above, some all-electron basis sets are present in the basis set directories SZ through TZ2P, but not for all elements of the periodic table. The heavier elements, from Rb on, the non-relativistic all electron basis sets are missing. In the ZORA basis sets directory you will find all-electron basis sets for all elements (Z = 1-118), which also could be used in non-relativistic calculations. Note, however, that these basis sets were optimized for ZORA calculations, which means that non-relativistic calculations will not always give you the expected accuracy. Warning: the frozen core basis sets in the ZORA directory should never be used in non-relativistic calculations. Non-relativistically optimized basis sets for the heavier elements are provided in a separate directory AE, which contains basis sets of single-, double- and triple-zeta quality indicated respectively by suffixes 'sz', 'dz', and 'tz'. The files in Special/AE/ are not complete database files, because they don't contain fit sets (the usage and relevance of fit functions is explained later).

The development of fit sets and their testing is not a triviality. It is absolutely a bad idea to take a fit set from another database file, corresponding to some frozen core level, and use that in an all-electron basis set: this will give significant errors and make results worthless. In the ZORA directory one can find all-electron basis sets with good fits sets for the heavier elements.

References on basis sets

Older references for STO basis sets are Refs. [336-338]. More recent:
[320] E. van Lenthe and E.J. Baerends, *Optimized Slater-type basis sets for the elements 1-118.* Journal of Computational Chemistry **24**, 1142 (2003)
[321] D.P. Chong, E. van Lenthe, S.J.A. van Gisbergen and E.J. Baerends, *Even-tempered Slater-Type orbitals revisited: From Hydrogen to Krypton.* Journal of Computational Chemistry **25**, 1030 (2004)
[322] D.P. Chong, *Augmenting basis set for time-dependent density functional theory calculation of excitation energies: Slater-type orbitals for hydrogen to krypton.* Molecular Physics **103**, 749 (2005)

See also the paper by Raffennetti on design and optimization of even-tempered STO basis sets [317]. The paper by Del Chong describes completeness profiles as a visual tool in estimating the completeness of a basis set [318]. Finally, Zeiss and coworkers [319] describe field-induced polarization functions for STOs. These are useful for defining basis sets with diffuse functions for (hyper)polarizability and other property calculations.

The procedure for the usage and optimization of fit functions is described by Baerends et al. [308].

How TO make EVEN-tempered basis/fit sets?

The programs questbas, rafbas, etprog, etwrite, described below, have not yet been made generally available.

The standard basis sets that are provided are sufficiently flexible to accommodate the needs of almost every type of calculation of almost every user. Therefore, the first question should be: do I really need to make my own basis set? Almost always, the answer should be 'no', because of the availability of the directories (ZORA/) SZ-QZ4P (old names I-V) and basis set directories containing diffuse ET basis and fit sets. If however, you decide that these basis sets might be insufficiently reliable for your purpose, you can use the utilities described below.

At the moment, there are several restrictions to these utilities. (We currently do not yet make these utilities generally available as they have not been very extensively tested yet. People who think they may need these utilities should contact SCM). First, only starting points (basis sets for the occupied shells) are available for elements up to Kr. Also, we currently have some reservations about the ET basis sets from K-Kr, because of the relatively large basis set superposition errors that occur. Nevertheless, for accurate all-electron nonrelativistic basis sets up to Kr, the utilities may still prove a very useful tool. The tools have not yet been tested to work properly for the generation of basis sets suitable for ZORA calculations.

ET stands for even-tempered. This may apply to the basis and to the fit. In our ET basis sets, only 1s, 2p, 3d, and 4f functions occur. In the fit sets, 5g functions occur in addition to this. [Currently all ADF basis sets are restricted to f functions and fit sets are restricted to g functions as the highest l-value]. The exponents in ET basis sets are given by the simple formula:

$$\zeta = \alpha * \beta^{l} \quad \text{where } l=1,\ldots N$$

Here ζ is the exponent of the STO, b (which should be larger than 1) defines how far apart two consecutive exponents are, and a determines what the most diffuse exponent is. In principle, each l-value has its own set of α, β, N.

What is the basic idea behind the basis set utilities?

The basis set utilities generate ADF atomicdata files for ET basis and fit sets, using some simple input from the user. With the first utility, the user selects the ET basis set for the occupied shells that should form the starting point. These were developed and tested on atomic total and orbital energies by Prof. Del Chong during a sabbatical spent in Amsterdam. These basis sets are intended to be at least of the quality of basis V, but usually better. However, also some smaller ET basis sets were made for more economical calculations. After selecting the basis set for the occupied shell, the user has to specify whether additional polarization, diffuse, or contracted basis functions should be added (also tight functions for ZORA calculations can be added, but this has not been properly tested). To answer these questions requires of course some expertise in basis sets. It is therefore recommended for users with some experience in this area. However, suitable, safe defaults are defined and suggested in the scripts. The user also defines the quality of the fit set which is desired. Only the highest quality fit set has been thoroughly tested. The other options are not supported at the moment. We will now describe the input and output of the various utilities in some detail.

The utility questbas

This utility is available in $ADFBIN/questbas.exe. Running this executable results in the following questions:

Please provide the atomic number of the element

For Carbon, we type 6

As a next question a long list of default basis sets is presented. These correspond to predefined choices for the number of diffuse, tight, and polarization functions, and are used to automatically generate all kinds of ET basis sets. If any of these standard basis sets are chosen, the answers to all questions are filled in automatically.

Choose 0 for the general basis set procedure

SUGGESTED CHOICE: 0

Here, we follow the suggested choice which allows us to explain all options, and type 0.

Now the script tells us we have the choice between 4 different basis sets for the occupied shells. The basis sets which have been most thoroughly investigated at present are the ET basis sets in which the factor beta (β) [defining the spacing between two subsequent STOs] is fixed at the value 1.7 for all l-values in the basis. These basis sets form the starting point for the directories VI and VII. For carbon, the output tells us that we would start from a 5s4p basis with this choice. In another ET basis set, with a different design philosophy, β is variable (typically 1.55 for s, 1.7 for p, 2.0 for d), but α is fixed. These basis sets occur in a large (valence quadruple zeta, core double zeta) and medium (augmented double zeta) variety. In the case of carbon, they are of the size 6s4p and 4s3p. These variable β basis sets have been less intensively tested in molecular calculations than the fixed β variety. For that reason a final verdict has to be postponed as to which type is preferable. The fixed β seems the safest choice at present, because it has been used more in molecular calculations. The large variable β basis leads to names like VIB and VIIB. We type here 1 for the fixed β variety. As output we now get information on our intermediate basis set

As a starting point you have selected an even-tempered basis with following values for n, alpha, beta:

S: 6 0.6592211993160408 1.700000047683716

P: 4 0.5316027776326796 1.700000047683716

The script follows with the following question:

You can specify here how many additional fus you want for each l-value. Suggested: 0 0 0 0

This option allows the user to add functions to the final basis by keeping the endpoints, most diffuse and tightest functions, fixed and reducing the β value. This option is not often useful and we ignore it by typing 0 0 0 0

The next question concerns the ZORA option. As this has not yet been thoroughly tested we choose 0 for a nonrelativistic basis set.

The next question is if we want to add functions. This is definitely needed, because the basis without polarization functions is quite poor, so we type 1.

Then we are asked how many polarization functions should be added. As a default is suggested 0 2 1, meaning 2d and 1f polarization functions. This means that we get more polarization functions than in basis V if we use the default. The geometric mean of the exponent of the polarization functions is stored inside the program in a data statement. This gives the 'best' exponent for a single polarization function. If more than polarization function is used, this will be the middle of these polarization functions. The beta value for the polarization function is taken identical to the beta value of the highest occupied l-value.

After typing the suggested default of 0 2 1, we are asked if we want to add diffuse functions. Whether such functions are needed in your application is discussed elsewhere in this document. Let us add one diffuse p function by typing 0 1 0 0

We are then asked about additional tight functions to improving the description of the core region. In this example, we do not add such functions by typing 0 0 0 0

Then we are asked what size of fit set we require. We type 3 to get the largest fit set, the only thoroughly tested option.

Finally we are asked if we want to let the program add diffuse function to make the basis set suitable for (hyper)polarizability calculations. We follow the suggested choice by typing 0 0. Otherwise, the utility would add diffuse functions until the most diffuse function is more diffuse than a stored default diffuse value (provided by Prof. Del Chong, based on field-induced polarization functions).

The output of the questbas utility is given in the (local) file raf_in. This file gives the input for the rafbas utility described next.

The utility rafbas

The utility $ADFBIN/rafbas.exe reads its input from the local file 'raf_in' and writes its output to the local file 'et_in'. No further user input is required. This utility basically transforms the answers of the user to the questions posed in 'questbas' into a set of N, α, β values for both the basis and the fit.

The rafbas and the other utilities were originally also intended for generating so-called cusp-satisfying basis sets, which consist of (for the s functions) 1 1s functions to which ET 3s functions are added. However, some tests indicated that the core description of these basis sets, which was supposed to be very good, did not improve upon those of the normal ET basis sets and therefore the cusp-satisfying basis sets were abandoned.

The rafbas utility contains information about the beta values for the fit. For the default 'VERYLARGEFIT' option, the beta values were chosen such that the overlap between two successive fit functions does not exceed 0.95 for s and 0.90 for higher l it functions. Earlier experience with generating fit sets led to the conclusion that even larger fit sets have a large risk of becoming linearly dependent which causes numerical problems. The rafbas utility also adds the polarization function exponents. It contains a list of 'best' exponents for many elements. It further more handles the options to extend the basis by decreasing beta and to add first- and second-order FIPs (field-induced polarization functions) for the generation of basis sets for (hyper)polarizability calculations.

The utility etprog

Similarly to the 'rafbas' utility, 'etprog' reads from the local file 'et_in' and writes to the local file 'et_tmp'. It also prints some output which gives information on what the meaning is of the numbers in the file 'et_in'. The result in 'et_tmp' is almost of the form (N-values basis, α-values basis, β-values basis, N-values fit, α-values fit, β-values fit).

The main task of etprog is to generate a fit set corresponding to the basis set information generated by rafbas. To this purpose, it first generates the most diffuse and most contracted products of basis functions. This defines the range which should at least be well described by the fit set. However, the range is extended somewhat more to further increase the quality of the fit. Experience shows that, at least for light elements, this indeed leads to very small fit errors in molecular calculations, at the expense of a large fit. The beta values for the fit are read from the input file and originate from a data statement in the rafbas program.

The utility etwrite

No calculations are needed in the utility 'etwrite', it merely writes out the final information in a usable form to several files. This utility reads from the local file 'et_tmp' and writes to the file 'C' because this was the element we selected. This file 'C' is an atomicdata file as required for an ADF create run and can be used without further change. This file also contains information about the final basis and fit set, which can be used to uniquely specify this basis and fit in a publication. The output further consists of local files like 'SSTO.BAS.INP' which is intended as input for Chong's completeness profile utility $ADFBIN/ssto.exe. The result of that program is a datafile which can them be visualized with the gnuplot viewing program using the other output file 'gnu.in.basis' as an example. For the interpretation of the completeness profiles, we refer to Prof. Chong's publications on this subject.

An example of how all these scripts and programs can be combined is to be found in $ADFHOME/ examples/ basis_fit_utils/e_make_bas_Kr/run, which should be more or less self-explanatory using the information given here. The directory $ADFHOME/examples/basis_fit_utils also contains other scripts which can be of use in the generation and testing of basis and fit sets.

Available standard basis sets

Click on the element to see the available standard basis sets for a given element that are distributed with the ADF package. Note that a click will be towards the Products page on the www.scm.com website.

	1	2		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1 H																		2 He
2	3 Li	4 Be												5 B	6 C	7 N	8 O	9 F	10 Ne
3	11 Na	12 Mg												13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
4	19 K	20 Ca		21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
5	37 Rb	38 Sr		39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
6	55 Cs	56 Ba	La-Yb	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
7	87 Fr	88 Ra	Ac-No	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Uut	114 Fl	115 Uup	116 Lv	117 Uus	118 Uuo

Lanthanide elements	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb			
Actinide elements	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No			

Basis sets directories

Basis sets can be found in the directory $ADFHOME/atomicdata, for non-relativistic calculations in the subdirectories SZ, DZ, DZP, TZP, TZ2P, TZ2P+, for ZORA calculations in ZORA/SZ, ZORA/DZ, ZORA/DZP, ZORA/TZP, ZORA/TZ2P, ZORA/TZ2P+, ZORA/QZ4P, the augmented basis sets can be found in AUG/ASZ, AUG/ADZ, AUG/ADZP, AUG/ATZP, AUG/ATZ2P, and the even tempered basis sets in ET/ET-pVQZ, ET/ET-QZ3P, ET/ET-QZ3P-1DIFFUSE, ET/ET-QZ3P-2DIFFUSE, ET/ET-QZ3P-3DIFFUSE. All electron basis sets can be used in non-relativistic and ZORA calculations.

Basis sets acronyms

- SZ: single zeta
- DZ: double zeta
- DZP: double zeta + 1 polarization function
- TZP: valence triple zeta + 1 polarization function
- TZ2P: valence triple zeta + 2 polarization function
- TZ2P+: = TZ2P + extra d (3d metals) or extra f (lanthanides)
- pVQZ, QZ3P: valence quadruple zeta + 3 polarization function, even tempered
- QZ3P-nD: = QZ3P + n diffuse sets of s, p, d, and f functions, even tempered
- QZ4P: valence quadruple zeta + 4 polarization function, relativistically optimized
- ASZ, ADZ, ADZP, ATZP, ATZ2P: augmented for use in TDDFT

For small negatively charged atoms or molecules, like F^- or OH^-, use basis sets with extra diffuse functions, like the augmented or QZ3P-nD

All electron or frozen core

- element name (without suffix): all electron
- .1s frozen: 1s
- .2p frozen: 1s 2s 2p
- .3p frozen: 1s 2s 2p 3s 3p
- .3d frozen: 1s 2s 2p 3s 3p 3d
- .4p frozen: 1s 2s 2p 3s 3p 3d 4s 4p

- .4d frozen: 1s 2s 2p 3s 3p 3d 4s 4p 4d
- .4f frozen: 1s 2s 2p 3s 3p 3d 4s 4p 4d 4f
- .5p frozen: 1s 2s 2p 3s 3p 3d 4s 4p 4d 5s 5p (La-Lu)
- .5p frozen: 1s 2s 2p 3s 3p 3d 4s 4p 4d 4f 5s 5p (other)
- .5d frozen: 1s 2s 2p 3s 3p 3d 4s 4p 4d 4f 5s 5p 5d
- .6p frozen: 1s 2s 2p 3s 3p 3d 4s 4p 4d 4f 5s 5p 5d 6s 6p (Ac-Lr)
- .5f frozen: 1s 2s 2p 3s 3p 3d 4s 4p 4d 4f 5s 5p 5d 5f 6s 6p

Automatic mode

The following input will run a geometry optimization on water, using a (almost) minimal input:

```
ATOMS
 O  0  0  0
 H  1  1  0
 H -1  1  0
End

Basis
End

Geometry
End
```

The ATOMS block key specifies the starting geometry.

The GEOMETRY key instructs ADF to perform a geometry optimization.

The BASIS block key instructs ADF to run the appropriate create runs automatically, using default values for the basis sets to use. For the XC potentials invoked with the MODEL subkey (cf. XC input block) the XC potential in the Create run will be a GGA potential rather than such a model potential, as these potentials cannot currently be applied in Create runs.

The Automatic mode will be used when the Basis key is present in the input:

```
BASIS
 Type bastyp
 Core coretyp
 Path apath
 Atom atompath
 ...
 FitType fittyp
 CreateOutput FileName
End
```

All subkeys are optional. For most calculations you need only to set the Type and Core subkeys.

`Type bastyp`

> `bastyp` is the type of basis set to use, and must correspond with the name of the directory as used within $ADFRESOURCES (=$ADFHOME/atomicdata), or within $ADFRESOURCES/ZORA for ZORA calculations.
> Valid standard basis set types are: SZ, DZ, DZP, TZP, TZ2P, and QZ4P. More valid basis set types can be found in the $ADFRESOURCES directory, such as the TZ2P+ basis set, the augmented basis sets AUG/ASZ, AUG/ADZ, AUG/ADZP, AUG/ATZP, AUG/ATZ2P, and the even tempered basis sets ET/ET-pVQZ, ET/ET-QZ3P, ET/ET-QZ3P-1DIFFUSE, ET/ET-QZ3P-2DIFFUSE, ET/ET-QZ3P-3DIFFUSE.
> ZORA will be included only for the standard basis set types: SZ, DZ, DZP, TZP, TZ2P, and QZ4P, and if

the calculation is a ZORA calculation. In case one of the standard basis set types is used, but no basis set of the specified type is available, ADF will try to use a larger basis set.
Default: DZ.

`Core coretyp`

`coretyp` is the type of frozen core to use. Allowed values: None, Small, Medium, Large.
If no basis set with core is available, an all electron basis set will be used.
If there is only one basis set with core, Small, Medium and Large are identical.
If there are two basis sets with core, Medium and Large are identical.
Default: Large.

`Path apath`

`apath` is an alternative directory with basis sets to use. ADF looks for appropriate basis sets only within this directory.
Default: $ADFRESOURCES

`Atom atompath`

In this subkey 'Atom' should be replaced by the name of the atomic fragment for which you want to specify the basis, for example 'O'.

Use this key to specifically select a basis set for this atom:

- an absolute path to a basis file (for example $ADFRESOURCES/DZ/O.1s)
- a relative path to a basis file (for example DZ/O.1s)
- a filename within the Type directory (for example O.1s)

An absolute path will always be used as specified.
A relative path is relative to the value of the PATH (or PATH/ZORA) subkey.
A filename is always relative to PATH/Type or PATH/ZORA/TYPE directory.

The relative path or filename will automatically switch to a ZORA basis set only in case of the standard basis set types: SZ, DZ, DZP, TZP, TZ2P, and QZ4P, and if the calculation is a ZORA calculation.
You can have one Atom subkey for each basic atom type in your input.

Since you pick explicitly the file to use, you are responsible for choosing a reasonable basis set.

`FitType fittyp`

`fittyp` is the type of auxiliary fit set to use, and must correspond with the name of the directory as used within $ADFRESOURCES, or within $ADFRESOURCES/ZORA for ZORA calculations.
Note that this is an expert option. For all atoms the fit set will be changed if this key is used. The fit set for a given atom is then taken from the all-electron basis set file for the same element as is the atom.
Typical usage: one might want a larger fit set than is present on the basis file, and then fittyp could be: ZORA/QZ4P. More valid fit set types can be found in the $ADFRESOURCES directory. If the requested directory or fit set is not available, then this option might fail without warning.
Default: the auxiliary fit set that is available in the requested basis file.

`CreateOutput FileName`

Use the CreateOutput option to change where the output from ADF create runs and the Dirac program goes. If it is not present, it will go to standard output. The special value 'NONE' for FileName makes it disappear, and any other value will be used as a file name in which to save the output.

Do not include the Fragments or Corepotentials keys when using the Basis key!

When the `Basis` key is present, ADF will first create fragment files for all the basic atom fragments found in the ATOMS key block. Normally this means that for each atom type in your molecule a fragment file will be created.

You may have different fragments with the same atom: add a dot and a name (without spaces) after the name of the element, as described in the ATOMS key. For example: H.1 and H.2. In this example two fragment files will be created: one for the H.1 fragment and one for the H.2 fragment. Using the ATOM subkey you may assign different basis sets to these fragments. Another consequence is that the H.1 and H.2 atoms will never be symmetry equivalent to each other.

Starting from ADF2006.01 the BASIS key recognizes elements denoted with Gh.atom in the ATOMS key as being ghost atoms. If one does not specifically select a basis set for this ghost atom, the all electron basis set for the atom is selected in the creation of the ghost atom using the type of basis set chosen with the BASIS key. The atom name must begin with the standard one- or two-character symbol for the chemical element: Gh.H, Gh.He, Gh.Li, and so on. Optionally it may be appended by .text, where text is any string (not containing delimiters). Examples: Gh.H, Gh.Mn.3, Gh.Cu.dz-new.

The basis set to use follows from the subkeys (or the default values), the XC potential follows from the XC block if it is present in the input.

In case of a relativistic calculation, the DIRAC program will also be run automatically, and the create runs will include the correct relativistic key and corresponding basis sets. For ZORA calculations, ADF first tries to locate a special ZORA basis set. If this does not succeed it will use a normal basis set if the required basis set does not use a frozen core.

The resulting fragment files will be named t21.atom, with 'atom' replaced by the names of the basic atoms present. In case of a relativistic calculation, the corepotentials will be stored on t12.rel.

Create mode

In Create mode the input file can be extremely simple. First, the geometry is trivial: one atom at the origin. Indeed, no coordinates etc. are read from input; any such items are ignored.

Second, the problem is computationally so simple that default settings for precision aspects, such as convergence criteria and levels of numerical integration accuracy, are internally defined to be much more stringent than in normal calculations. These aspects don't have to be looked after.

In Create mode you need only a one-line input file of the following form:

```
CREATE Atomtype Datafile
```

`Create`

is the keyword. The remainder of the record (atomtype datafile) is the argument.

`Atomtype`

is a name for the basic atom that you want to create. The program reads and interprets this name. Therefore, the name must begin with the standard chemical symbol (H, He, Li, ...) of the element to be created. Optionally the name may have an suffix of the form .text. The suffix begins with a period (.); the part after the period (text) is at your discretion as long as it does not contain a delimiter. A few examples:

appropriate names	*inappropriate names for an atom type*	
K	Si-with-core	: no period after the chemical symbol
Li.newbasis	$HOME/atomicdata/C.dzp	: not beginning with the chemical symbol
P.1992/Feb./30	Ga.nocore,smallbasis	: contains a comma (a delimiter)

Sodium.2s	: Sodium is not the *symbol* for this element (Na)

Examples of appropriate (left) and inappropriate (right) atom type names used with the keyword `create`.

`Datafile`

specifies the data file that contains the basis set and related items. It may contain a full path if the file does not reside in the working directory of the job.
The `datafile` part is optional. If you omit it, ADF assumes that the file name is identical to the atom type name, i.e.
`Create Atomtype`
is equivalent to and interpreted as
`Create Atomtype Atomtype`
In view of the restrictions that apply to the atom type name, the option to use the short form can only be used if the file name has the appropriate format.
To make the input file easier to understand for a human reader you may, for Datafile, also type file=Datafile, where file= must be typed as such, and datafile is the name of the file.

So you could have a very simple calculation as follows (the 'creation' of a Carbon atom);

```
$ADFBIN/adf << eor
 Create C.dzp
eor
```

The presence of the keyword `create` sets the computational mode of ADF to: *create a basic atom*. The argument (C.dzp) is then analyzed and found to have as initial part C, telling ADF that we'll be creating a Carbon atom. Since the file-specification part is missing, the data file with the basis set etc. must be the (local) file with the name C.dzp.

More often you will directly address a file (with the basis set) that is not local, but located in the database of your ADF package. The script could then be:

```
$ADFBIN/adf << eor
 Create C $ADFHOME/atomicdata/DZ/C.1s
eor
```

Here you address the file 'C.1s' in the database subdirectory DZ/ (this contains basis sets of double-zeta quality).

A considerable number of data files are included in the ADF database. To apply such a file for the creation of a basic atom:

Make a copy of the data file in the directory where you want to run the program. Since the standard data file names satisfy the requirements for atom type names you can now use the simplest option to use the `create` key:
Construct a one line input file *in* (`create` name-of-data-file-copy)
Run ADF by typing
`adf <in >out`
When the calculation has finished, give the result file TAPE21 a suitable name and move it to a directory where you build your database of fragment libraries.
Examine logfile and out to check that everything has gone well.

You may want to define alternative basic atoms, different from those in the standard ADF database, for instance to try out a different basis set developed by yourself. By inspection of one of the standard data files you can see what the contents of such a file should be. A complete description is given in Appendix 5.1.

You can also create basic atoms corresponding to so-called *Alternative Elements*, with for instance a non-integer nuclear charge or a different mass. See the next section.

Ghost Atoms & Non-standard Chemical Elements

The atom type names used under atoms (and in the create record) must begin with the standard chemical element symbol (H, He, Li...). The program uses this to deduce the nuclear charge and other elemental properties.

For the standard elements one can redefine the atomic mass (for instance to define a suitable isotope).

A more extensive feature is available to define an artificial chemical element with user-specified properties. Such new elements are denoted *Alternative Elements*; and may for instance have a non-integer nuclear charge. The chemical symbol of an Alternative Elements is Gh (for ghost) or J: either one is ok.

You can create J-type or Gh-type basic atoms and use them subsequently as fragments in a molecule.

Automatic mode

Starting from ADF2009.01 it is easy to make different isotopes or elements if one uses the ATOMPROPS key in combination with the BASIS key. Typical use would be for the nuclear mass.

```
ATOMPROPS
 Atom.name {m=mass} {q=Q}
 ...
End
```

`mass`

The atomic mass, in atomic mass units, which will then override the default value for the indicated chemical element. Can be used, for example, for the calculation of isotopes. If not supplied for a J-element it will be set to the atomic mass of the standard chemical element with nuclear charge A, where A equals Q rounded to the nearest integer, but not smaller than 1 and not larger than 118.

`Q`

The nuclear charge. The q= option *must* be used for a J-element.
It must *not* be used for *standard* chemical elements.

`name`

In Atom.name the first letter after the dot should be a capital.

Example with three different isotopes of hydrogen:

```
Atoms
 N         0.000000    0.000000    0.010272
 H        -0.471582   -0.816803    0.407861
 H.D       0.943163    0.000000    0.407861
 H.T      -0.471582    0.816803    0.407861
End
AtomProps
 H.D m=2.014101778
 H.T m=3.01604927
End
Basis
 Type TZP
End
```

Starting from ADF2006.01 the BASIS key recognizes elements denoted with Gh.atom in the ATOMS key as being ghost atoms. Thus for ghost atoms one does not need the ATOMPROPS keyword. When you use

ghost atoms within your ADF calculation ADF will read the basis file that you provide as usual. However, starting from ADF2009.01 the core section will be skipped automatically. That is, even if the basis file specifies a frozen core ADF will treat it as if no frozen core is present. Thus, one no longer needs to edit the basis files, or to switch to all-electron basis files to use ghost atoms. The atom name must begin with the standard one- or two-character symbol for the chemical element: Gh.H, Gh.He, Gh.Li, and so on. Optionally it may be appended by .text, where text is any string (not containing delimiters). Examples: Gh.H, Gh.Mn.3p, Gh.Cu.dz-new.

Create mode

For the standard elements one can redefine the atomic mass (for instance to define a suitable isotope).

```
CREATE H {m=value} datafile
```

`value`

The atomic mass, which will then override the default value for the indicated chemical element.

The nuclear charge of an Alternative Element is not pre-defined, and *must* therefore be specified in the Create run. The atomic mass is optionally supplied.

```
CREATE J.NewElement q=Q {m=mass} datafile
```

`J.NewElement`

The atom type name, beginning with the *alternative* chemical element symbol J. It has an (optional) suffix of the form .text, completely similar to the construction of atom type names from standard chemical element symbols.

`Q`

The nuclear charge. The q= option *must* be used for a J-element.
It must *not* be used for *standard* chemical elements.

`mass`

The atomic mass, in atomic mass units. If not supplied it will be set to the atomic mass of the standard chemical element with nuclear charge A, where A equals Q rounded to the nearest integer, but not smaller than 1 and not larger than 118.

`datafile`

The Create data file.

If you want to use the Alternative Element feature you'll have to construct your own Create data file, suited to the Alternative Element you have in mind. Appendix 5.1 describes the format of such a file.

Use as fragment

J-type basic atoms can be used like any other basic atoms to build up larger fragments and molecules. In fact, J (or Gh) can be considered just one more chemical symbol along with the 118 traditional ones. The element J has no pre-defined properties. Therefore you have to specify them where appropriate (c.f. the nuclear charge and atomic mass).

You may have different J-elements in a molecule, with different nuclear charges for instance. Yet, they must be denoted with the same chemical symbol J; the difference can only be made clear by the .text suffix in the atom type name.

It does no harm of course to make this suffix a concise but clear description of the main characteristics.

Nuclear Model

By default in ADF a point charge model is used for the nuclear charge distribution. In ADF2009.01 a spherical Gaussian nuclear charge distribution model has been implemented in ADF, see Ref. [270]. Nuclear finite size effects can have large effects on hyperfine interactions (ESR A-tensor, NMR spin-spin coupling) if heavy atoms like, for example, Mercury (Hg), are involved. In Ref. [270] it was written that the isotropic J-couplings (parameters in NMR spin-spin coupling) are typically reduced in magnitude by about 10 to 15 % for couplings between one of the heaviest NMR nuclei and a light atomic ligand, and even more so for couplings between two heavy atoms. This Ref. [270] gives more details on the parameters used in the Gaussian nuclear charge distribution model. Note that one needs basis sets with very tight functions to see any effect of using a finite size of the nucleus instead of a point nucleus. Such basis sets can be found for some elements in $ADFRESOURCES/ZORA/jcpl, which are basis sets especially designed for NMR spin-spin coupling calculations.

A Gaussian nuclear charge distribution will be used if one uses the NUCLEARMODEL key with:

```
NUCLEARMODEL gaussian
```

```
NUCLEARMODEL nuclearmodel
```

The argument nuclearmodel of the key NUCLEARMODEL can be 'pointcharge' (default) or 'gaussian'. It should be included in the Create run of an atomic calculation and in the molecular calculation. If the BASIS key is used it will be automatically added in the Create run of the atoms. If this key is absent a point charge nuclear model is used.

In the ADF output parameters will be shown for the Gaussian nuclear charge distribution if one includes in the input for ADF:

```
PRINT Nuclei
```

starting from ADF2013 ADF also uses a finite distribution of the nuclear magnetic dipole moment for the calculation of the A-tensor.

What basis set should I use in ADF?

This question is hard to answer in general, but a few general suggestions can be made. This will be split here into several subsections.

ZORA or nonrelativistic calculation?

The first question to ask is: am I going to do a ZORA calculation to include scalar relativistic effects? This is recommendable for systems containing heavy nuclei, but there is no objection to doing a ZORA calculation for a system with light atoms only, as the required CPU time does not increase very much if ZORA is used.

If you are doing a ZORA calculation, you will need the ZORA basis sets which can be found in $ADFHOME/atomicdata/ZORA. You may also use the all electron basis sets from the ET or AUG directory, but be aware that these were optimized to non-relativistic calculations. Currently the ZORA basis sets cover the entire periodic table and besides all electron basis sets offer a choice of frozen cores. At present the ZORA directory does not contain basis sets with very diffuse functions, which may be required in calculations for hyperpolarizabilities or high-lying excitation energies, but for the lighter elements (H-Kr) you can certainly use the all-electron basis sets from the ET or AUG directory. Warning: in a ZORA calculation use only the frozen core basis sets coming from the $ADFHOME/atomicdata/ZORA directory, or use all electron basis sets.

If you do not use ZORA, your basis sets should come from the directories SZ, DZ, DZP, TZP, TZ2P, or one of the ET or AUG basis sets. For many of the heavy elements only ZORA basis sets are available, but for such elements it would be inadvisable to do nonrelativistic calculations anyway. For light elements the ZORA and normal basis sets should be identical except for the description of the frozen core. Usually the ZORA basis sets contain much steeper basis and fit functions to accurately describe the core region.

Large or small molecule?

For standard calculations (energies, geometries, etc.) we recommend the following hierarchy of basis sets:
SZ < DZ < DZP < TZP < TZ2P < TZ2P+ < ET/ET-pVQZ < ZORA/QZ4P
where the largest and most accurate basis is on the right. Not all basis sets are available for all elements.

For small negatively charged atoms or molecules, like F^- or OH^-, basis sets with extra diffuse functions are needed, like they are available in the AUG or ET/QZ3P-nDIFFUSE directories. For example, the standard basis sets, or even the large ZORA/QZ4P basis set will often not be large enough for the accurate calculation of such anions.

In general it is advisable to use the best basis set that you can afford to use in terms of CPU time and memory. If you want to optimize the geometry or calculate the atomization energy of a diatomic molecule there is little reason not to use the very large ZORA/QZ4P basis, or (for light elements) a similarly large ET basis (we recommend the ET-pVQZ basis). If you are studying a molecule with 100 atoms or more, the use of such large basis sets does not only become prohibitive because of the required CPU time and memory, but it also is much less needed than for smaller systems. In medium-sized or large molecules even the moderately large basis sets will prove to be quite adequate because of the effect of basis set sharing. Each atom profits from the basis functions on its many neighbors. Additionally, if a large basis contains diffuse functions, linear dependency problems may occur. See also the input key DEPENDENCY. In many cases basis DZ or DZP will give acceptable accuracy for calculations on large systems. If you are inexperienced it may be prudent to test a few different basis sets to get a feel for the size of basis set effects. To get a rough idea for the size of various basis sets, we mention here the number of functions for all-electron basis sets from the directories ZORA/SZ up to ZORA/QZ4P. For carbon, the number of functions is 5 (basis ZORA/SZ), 10 (DZ), 15 (DZP), 19 (TZP), 26 (TZ2P), 43 (QZ4P). The same numbers for hydrogen are: 1 (SZ), 2 (DZ), 5 (DZP), 6 (TZP), 11(TZ2P, 21(QZ4P). These numbers arise because ADF uses 'pure' d and f functions. In other words, 5 instead of 6 d functions are used and 7 instead of 10 f functions. Note that especially the jump from TZ2P to QZ4P is quite steep.

In an overgeneralizing fashion we can state that the single zeta basis SZ is hardly ever sufficient to get more than a qualitative picture and should be used only when larger basis sets are not affordable. The double zeta basis DZ performs already much better and may give quite reasonable results, for example in geometry optimizations on large molecules. However, in more subtle situations, for example if hydrogen bonds are important, it is advisable to use at least one set of polarization functions. This is the double zeta polarized DZP basis set. Basis set TZP extends the valence space (but not the core space which remains double zeta) to triple zeta. In basis TZ2P an additional polarization function is added. For example, hydrogen gets a d polarization function in addition to its p polarization function and carbon gets an f polarization function on top of a d polarization function. Several tests have shown that often the second polarization function is of more use when it has an l value one higher than the first polarization function. This is reflected in the choice just described. If another polarization function is needed it is usually best to add another one of the lowest l-value (2p+1d for hydrogen, 2d+1f for carbon). This choice has been made in the ET basis ET-QZ3P. There, sometimes even three d polarization functions were added, for example 3 p functions for Be, and 3 d functions for S. The reason for this is that S can occur in hypervalent species such as SF_6, which put special demands on the basis set. In the case of Be, the unoccupied p level is so close in energy to the occupied ones that it is sometimes called a valence level. Symantics aside, it is clear that a proper description of the p level of Be is very important and it is therefore not strange to add a third p function. In general, the unoccupied levels for the atoms on the left side of the periodic table are close to the occupied ones. This makes it necessary to add a few extra functions for the lowest unoccupied levels in order to get a description which corresponds to the general level of accuracy one expects from the hierarchical basis set names SZ-TZ2P. The basis set quality for a particular subdirectory is now rather uniform throughout the periodic

system. At the same time we have attempted to increase the number of functions in a systematic fashion so that each element is described by at least as many functions of a particular l value as its predecessor.

The ZORA/QZ4P basis sets might be roughly described as core triple zeta, valence quadruple zeta, with 4 polarization functions (2 d and 2 f functions for C, 2 p and 2 d for H). The fit sets corresponding to these basis sets are also much larger than the fit sets found in basis sets SZ-TZ2P. If one has doubts about the adequacy of a fit set for a certain element, this can be tested by replacing its fit set by the large one from the QZ4P directory, see the subkey FitType of the key BASIS. In the ZORA/QZ4P basis sets, the choice for the exponents of the polarization functions was done in a systematic, but somewhat hand-waving manner. For this reason the exponents were always rounded to half integers. Also the geometric mean of the exponents usually does not coincide with the choices made in directories SZ-TZ2P and the ET basis sets. However, the fact that two polarization functions (with reasonable exponents) are present instead of a single one is far more important. A reasonable intermediate basis set, in size between TZ2P and QZ4P might be envisaged in which a single polarization function is added, as described above. This is roughly the choice for the polarization functions made in the ET directory ET-QZ3P.

Frozen core or all-electron?

In general we recommend the use of frozen core basis sets if available. Especially for the heavier atoms the number of functions is much smaller than for their all-electron counterparts. Our tests indicate that the error made by invoking the frozen core approximation is usually clearly smaller than the difference with respect to slightly higher quality basis sets. For the ZORA/QZ4P basis sets, only all-electron basis sets are available as these are intended for near basis set limit calculations only in which the CPU time is not a major concern.

Geometry optimizations involving atoms with a too large frozen core may give rise to numerical problems. In such cases it is recommendable to use a smaller frozen core. In previous occurrences we have removed such atomicdata files from the database.

For accurate results on properties like nuclear magnetic dipole hyperfine interactions (ESR), nuclear quadrupole coupling constants, and chemical shifts (NMR), all electron basis sets are needed on the interesting atoms. For such properties tight functions might be necessary for high accuracy, especially in a ZORA calculation.

Diffuse functions needed?

For small negatively charged atoms or molecules, like F^- or OH^-, basis sets with extra diffuse functions are needed, like they are available in the AUG or ET/QZ3P-nDIFFUSE directories. For example, the standard basis sets, or even the large ZORA/QZ4P basis set will often not be large enough for the accurate calculation of such anions.

For accurate results on properties like polarizabilities, hyperpolarizabilities, and high-lying excitation energies, also diffuse functions are needed. This is especially true for calculations on small molecules. In larger molecules the nature of the relevant virtuals is much more 'molecular', much less Rydberg-like, so that the normal basis sets may be sufficient. Basically all properties calculated through the RESPONSE keyword may require diffuse functions. If you use the EXCITATIONS keyword, the necessity of diffuse functions depends on the type of excitation you are interested in. The lowest excitations do not require diffuse functions, but Rydberg excitations do.

In case of diffuse basis functions the risk of linear dependency in the basis increases. This can be checked, and corrected for with the DEPENDENCY keyword. It is recommended to use this keyword for all calculations involving diffuse functions. A good default setting is

DEPENDENCY bas=1d-4

However, it may be advisable to experiment a bit with the parameter, especially if many linear dependent combinations of AOs are removed. Using too many diffuse functions on a large molecule will lead to insurmountable numerical problems. In such a case it is not only useless, but even harmful, to add many diffuse functions.

In the previous release only some basis sets were provided which contained diffuse functions. These were gathered in the directory Vdiff. Now several ET basis sets have been developed for the elements up to Ar containing some or many diffuse functions. We recommend to use these instead of the Vdiff directory. Most of these basis sets are quite large and not very suitable for large molecules.

In ADF2005.01 augmented basis sets were added in the AUG directory, especially devised for use in in TDDFT calculations, such that one can do a reasonable accurate calculation of excitation energies, with a relatively small basis set, see D.P. Chong [322].

Normal or even-tempered basis?

For normal calculations (these form the vast majority) we recommend the use of the optimized basis sets in the directories SZ-TZ2P and, for ZORA calculations, ZORA/SZ-QZ4P. These should be sufficient in accuracy for even very demanding users and are available for the entire periodic system (in the case of the ZORA basis sets). They are also available with a frozen core variety, which saves much CPU time.

The ET basis sets on the other hand are available only in all-electron form at the moment. Furthermore, most are pretty large (larger, but also better than TZ2P). Additionally, relatively large basis set superposition errors were detected for molecules containing atoms in the row K-Kr. For this reason we only recommend ET basis sets for the elements H-Ar at the moment. There they have yielded quite nice, near basis set limit, results for the G2 test set. For these light elements the ET basis sets can be comparable in quality to the ZORA/QZ4P basis, even though it is smaller. The ET basis sets are considered to be especially useful when diffuse functions are required. In that case it is very easy to adapt the original ET basis and fit set. The utilities provided for this in ADF will be described below in an Appendix. The ET basis set utilities will also prove useful for users who want to experiment with making their own basis sets, or who have very special demands on the basis and fit. The provided utilities automate much of the work needed to make new atomicdata files.

What accuracy do the basis sets give?

Tests on many diatomics were performed to test the various basis sets. We now document the results of some of these tests, in order to give a feeling for the quality that can be obtained from the various basis sets. See also van Lenthe and Baerends [320].

Summary of test results

Tests for nonrelativistic calculations on 36 diatomics containing oxygen, namely the oxides of the first 36 elements (H-Kr). All-electron basis sets were used. The ZORA/QZ4P basis set was used to define the basis set limit result. The numbers in the table refer to bonding energies in eV. Differences were taken between the QZ4P results and the results in smaller basis sets. By construction, the errors in the QZ4P column are zero.

	QZ4P	DZ	DZP	TZP	TZ2P
Average error	0.0	1.33	0.39	0.18	0.06
Average absolute error	0.0	1.33	0.39	0.18	0.06
Maximum error	0.0	2.84	1.07	0.41	0.13
Worst case	all	SO	BeO	FO	O_2

A few comments are in order to explain this table.

The oxides were used as a small test set because their equilibrium bond lengths are known in many cases. Also, they have a large influence on the electronic structure of the molecule, so that they also test the adequacy of the polarization functions.

The errors in the small basis sets are systematic, because the isolated atoms are described reasonably well, but the molecular energy is not deep enough. For this reason the average errors and average absolute errors are (nearly) always equal.

Test calculations on 100 diatomics containing oxygen, using all-electron ZORA basis sets. Many basis sets for (very) heavy elements are included here, which could not be included in the table above. The numbers have the same interpretation as above and are again in eV.

	QZ4P	DZ	DZ	TZP	TZP	TZ2P	TZ2P
	ae	fc	ae	fc	ae	fc	ae
Average error	0.00	0.95	1.07	0.20	0.20	0.05	0.05
Average absolute error	0.00	0.98	1.07	0.20	0.21	0.05	0.05
Maximum error	0.00	2.86	2.83	0.74	0.74	0.19	0.17
Worst case	all	SO	SO	UuoO	UuoO	ThO	UuoO

Again we place a few comments on these frozen core and all-electron results.

The trends are very similar to those in the previous table for the lighter elements.

The frozen core results are very satisfactory, as they are very close to the results with the corresponding all-electron basis sets. The error introduced by the frozen core approximation is typically much smaller than the one introduced by basis set incompleteness.

The average errors are quite comparable to those from the previous table. The heavier elements do not seem to be much more difficult than the lighter ones.

For heavy elements no reliable ET basis set is yet available for comparison.

More results, all-electron, nonrelativistic on roughly 140 different diatomics at experimental or 'reasonable' equilibrium geometries.

	QZ4P	DZ	TZP
Average error	0.00	0.89	0.11
Average absolute error	0.00	0.89	0.11
Maximum error	0.00	2.84	0.32
Worst case	all	SO	O_2

Only the nonrelativistic basis sets DZ and TZP are fairly complete for heavier elements.

Also for these general diatomics (not just oxides) the average and maximum errors have decreased substantially, especially for basis TZP.

Same table, but now for frozen core basis sets. In all these tests the smallest frozen core files were employed (i.e. the largest basis).

	QZ4P	DZ	TZP
Average error	0.00	0.73	0.13

Average absolute error	0.00	0.75	0.16
Maximum error	0.00	2.87	1.80
Worst case	all	SO	ThO

The frozen core approximation has little influence on the accuracy for the new basis DZ, but a somewhat larger effect on the new basis TZP. This is especially due to certain worst cases, such as ThO.

ZORA, all electron, over 240 diatomics

	QZ4P	DZ	TZP	TZ2P
Average error	0.00	0.70	0.11	0.02
Average absolute error	0.00	0.70	0.11	0.03
Maximum error	0.00	2.83	0.44	-0.16
Worst case	all	SO	I_2	Cr_2

The average error goes down very nicely from 0.70 to 0.11 to 0.03 eV when going from DZ to TZP to TZ2P. The average error in basis TZ2P is clearly below 1kcal/mol (the famous chemical accuracy). Errors due to deficiencies in current xc functionals are still much larger than this. As a consequence, the ZORA/TZ2P basis will be more than adequate for all standard calculations.

It is to be expected that these conclusions will not dramatically change if larger test molecules are used. Also for geometry optimizations the improved basis sets SZ-TZ2P and ZORA/SZ-TZ2P should be more than sufficient for all standard cases. The ZORA/QZ4P can be considered a very safe (though expensive) option for basis set limit calculations.

Molecular fragments

Fragment mode

In Fragment mode more input is required than in Create mode: you have to specify at least: (1) the atomic positions and (2) how the total system is built up from fragments. We recommended to specify also (3) the point group symmetry.

Example of an input file for the C2H4 molecule:

```
ATOMS
 C  0    0  .6685
 C  0    0 -.6685
 H  .927 0 -1.203
 H -.927 0 -1.203
 H  .927 0  1.203
 H -.927 0  1.203
end

fragments
 C TAPE21c.dzp
 H TAPE21h.dzp
end
```

```
symmetry D(2h)
end input
```

Three keys are used: atoms, fragments and symmetry. The first two are block keys.

`atoms`

defines the atomic positions: each record in the data block contains the chemical symbol of an atom followed by its Cartesian coordinates in Angstroms.
Z-matrix type input of atomic positions is also possible. This will be explained in a later section.

`fragments`

lists the fragment files each record contains a fragment *type* followed by the corresponding fragment *file*. In the example the files are *local* files. Files in other directories are addressed by giving the complete file path.
Note: if a *parallel* calculation is performed, be sure that each 'kid' finds the specified fragment files. This will usually require that the files are *not* local to the job, but first be moved to some shared volume, and that the references to the fragment files in the input contain the full path. An alternative is to ensure that the (local) files in the parent directory are copied first to the 'kid' directories before the parallel calculation starts.

`symmetry`

specifies the point group symmetry by a Schönfliess type symbol. Appendix 5.3 contains a complete list of all Schönfliess symbols that are recognized by ADF. If no symmetry is specified ADF will take the true symmetry of the nuclear frame as the *user-specified* symmetry. If (electric) fields are used, see later, the computed symmetry will take this into account. Note that the computed symmetry may not occur in the list of allowed symmetries (see Appendix 5.3), in which case you have to explicitly specify the (lower) point group symmetry you wish to apply.

The atomic coordinates must conform to the point group symmetry; the program will check this and abort if the atomic system does not have the specified symmetry. It is allowed, however, to specify a *lower* symmetry than what is actually present in the set of atomic positions. The *specified* symmetry determines how results are analyzed and how irreducible representations and subspecies are labeled. It also determines various algorithmic aspects: the program runs more efficiently with the highest possible symmetry.

The spatial orientation of the molecular coordinate system is not arbitrary. ADF requires for each pointgroup symmetry a specific standard orientation. In axial groups for instance, the main rotation axis must be the z-axis. This implies a restriction on how you can define the atomic coordinates under atoms. The orientation requirements for all point groups are listed in Appendix 5.3. If the specified symmetry equals the true symmetry of the nuclear frame ADF will adjust the input orientation of the molecule to the requirements (if necessary). If you have specified a subgroup of the true nuclear symmetry, no such orientation adjustment is carried out and the user has to make sure that his input data yield the correct orientation, lest an error will occur.

Restrictions apply to the symmetry (as specified) of the molecule, related to the symmetries of the fragments as they were stipulated in the preceding fragment calculations. All symmetry operators of the molecule that internally rotate or reflect a fragment but leave it at the same position in the molecule, must also be operators of the symmetry group in which the fragment has been computed. Furthermore, two fragments must not be symmetry-equivalent in the molecule only by an improper rotation. The implied internal reflection of the fragment must be one of the symmetry operators in the point group symmetry that is used in the fragment calculation *and* the molecular symmetry group must also contain a proper rotation that maps the two fragments onto each other.

The example of the C2H4 molecule implicitly assumes that all fragments are *single atom* fragments. When the fragments are larger the data records in the atoms key have to be extended: you must specify which atoms belong together in one fragment.

```
SYMMETRY T(D)
Atoms
 Ni 0     0     0
 C -1.06 -1.06  1.06 f=CO/1
 C -1.06  1.06 -1.06 f=CO/2
 C -1.06  1.06 -1.06 f=CO/3
 C  1.06 -1.06 -1.06 f=CO/4
 O  1.71  1.71  1.71 f=CO/1
 O -1.71 -1.71  1.71 f=CO/2
 O -1.71  1.71 -1.71 f=CO/3
 O  1.71 -1.71 -1.71 f=CO/4
End
Fragments
 CO TAPE21co.yesterday
 Ni t21ni.dzp
End
End Input
```

Another sample input file; using a single atom Ni fragment and four molecular CO fragments. The keys symmetry and fragments operate as before. Again we have two types of fragments (here: Ni and CO); for each of them, the fragment file is specified.

Under the key ATOMS the chemical symbols and the nuclear coordinates are listed. Added is the f=...-part; f stands here for fragment and tells the program that the carbon and oxygen atoms belong to CO fragments. The last part /n enumerates the individual CO fragments: here you define which C and O belong together in one CO fragment.

The record for Ni contains no f= part, implying the *default* for this atom: it is a fragment on its own. In the C2H4 example before the default applied to all atoms.

Note that one should use the f= part for symmetry equivalent fragments. In the next example, ADF assumes the fragments CO1, CO2, CO3, and CO4, to be of different fragment types, even though they are coming from the same TAPE21. Therefore ADF will assume symmetry NOSYM in the next calculation, and will not run in T(D) symmetry.

```
Atoms
 Ni 0     0     0
 C -1.06 -1.06  1.06 f=CO1
 C -1.06  1.06 -1.06 f=CO2
 C -1.06  1.06 -1.06 f=CO3
 C  1.06 -1.06 -1.06 f=CO4
 O  1.71  1.71  1.71 f=CO1
 O -1.71 -1.71  1.71 f=CO2
 O -1.71  1.71 -1.71 f=CO3
 O  1.71 -1.71 -1.71 f=CO4
End
Fragments
 CO1 TAPE21co.yesterday
 CO2 TAPE21co.yesterday
 CO3 TAPE21co.yesterday
 CO4 TAPE21co.yesterday
 Ni t21ni.dzp
End
End Input
```

There are more possibilities with the keys atoms and fragments. This is worked out later. The purpose of this section was to provide a quick and easy start.

Fragment files

The TAPE21 result files from the ADF computations on the fragments that constitute a molecule completely characterize these fragments. The fragment TAPE21 files must be attached as *fragment files*. This is achieved with the key FRAGMENTS. See also the next section for the relation between Atom type, Fragment type and Fragment file names.

```
FRAGMENTS {Directory}
 FragType FragFile
 FragType FragFile
 ...
end
```

`FragType`

One of the fragment *types* defined under atoms, either explicitly (f=fragtype/n) or implicitly (fragment type=atom type, if the f= option is not used).

`FragFile`

The fragment file: the standard TAPE21 result file from the computation of that fragment. The file name must contain the complete path relative to Directory (the argument of the key). By default, when no Directory is specified, this is the local directory where the job runs. You may therefore omit the directory and give simple (local) file names if all the files are present in the working directory of the job.

Obviously, FragFile is case sensitive. However, FragType is also treated as case sensitive; see also the ATOMS key discussion (f= option). The reason is that there are shortcuts possible to the effect that the FragType name (in the atoms block) is immediately interpreted as the name of the fragment file.

The key FRAGMENTS may be used any number of times in the input file. This is convenient if you employ a sizeable number of fragment files, with subsets located in different directories. You can then use the key separately for each directory, to avoid typing long path names for all the files. Fragtypes that occur in the fragments block(s), but that are not referred by atoms are ignored. No fragment files must be specified for dummy atoms (xx).

It is allowed to use one and the same fragment file for different fragment types. Example:

```
ATOMS
 C.1 x1 y1 z1
 C.2 x2 y2 z2
 ...
end
fragments
 C.1 TAPE21.c
 C.2 TAPE21.c
 ...
end
```

Two different atom types (and fragment types) C.1 and C.2 are defined. The properties of the two fragment types are now identical since they are characterized by the same fragment file, but from the program's point of view they are different and can therefore not be symmetry equivalent.

The reason you may want to specify different atom types will usually be related to analysis, in particular symmetry aspects. If you know in advance that the two atom types are not symmetry equivalent, or more

generally, that they play a rather different role in the molecule, it can enhance clarity of printed output to assign different atom type names to them. However, see the notes below.

A fragment file must not be the result file of a spin-unrestricted calculation. When you try to use such a fragment file, the program will detect it and abort with an error message. If you want to analyze a molecule in terms of unrestricted fragments, you should use restricted fragment files and apply the key FRAGOCCUPATIONS.

Suppose that you have done a calculation on a molecule *mol*, in which you have defined two different atom types for atoms of the same chemical element. Suppose furthermore, that you want to use that molecule now as a fragment in a new calculation.

You list under atoms all atoms of the molecule and you specify which atoms belong to the various fragments, among which the molecular fragment *mol*. The program will then have a problem deciding which atoms in your system are associated with the different atom types in the fragment. Normally, ADF analyzes this by comparing the chemical elements. That is not sufficient here because one chemical element corresponds with more than one type of atom in the *mol fragment* type. In such a case it is imperative to use *the same atom type names* in your new calculation as you used in the generation of the fragment. These names are stored in the fragment file, and they are printed in the output file of the calculation of *mol*.

The names of three items may be related to each other, depending on how you specify input: the *atom type*, the *fragment type*, and the *fragment file*.

The atom type is defined in the data block to atoms.

The fragment type is defined also in the data block to atoms: with the f= option. For records in the data block that don't have the f= option, the fragment type name is by definition identical to the atom type name.

The fragment file is defined in the data block to fragments, each record consisting of a fragment *type* name, followed by the fragment *file*. If a fragment type is not listed in the data block to fragments, so that no fragment file name is specified, the fragment *file* is by definition identical to the fragment *type* name.

2.4 Model Hamiltonians

See also
ADF-GUI tutorial: multi-level calculations, spin-orbit coupling
GUI manual: model Hamiltonians
QM/MM manual
Examples: special XC functionals, relativistic effects, solvents, other environments

Electronic Configuration

The next few keys can be used to specify the electronic configuration. If you don't specify any such keys, certain defaults will apply. In principle, the program will (by default) attempt to find the lowest-energy spin-restricted (one-determinant) state. If SCF convergence is problematic the program may wind up at an excited state, by which (in this context) we mean a one-determinant state with a higher energy than some other one-determinant state with the same net spin polarization. In worse cases the program may fail to converge to any state at all. It is good practice to *always* verify which configuration you actually have computed.

When you specify a particular configuration and/or net charge and/or net spin-polarization of the system, the program will try to compute accordingly, even if the data have no physical or chemical meaning. The program has no knowledge about the existence of materials and will simply try to carry out what you tell it to do.

Charge and Spin

Spin: restricted vs. unrestricted

```
UNRESTRICTED
```

Specifies that spin-α and spin-β MOs may be spatially different and may have different occupation numbers. The default (absence of the key) is spin-restricted. The key has no argument. In the case of Spin-Orbit coupling it means that Kramer's symmetry does not have to be satisfied, in which case the key UNRESTRICTED should be used in combination with the key NONCOLLINEAR or COLLINEAR.

The unrestricted mode roughly doubles the computational effort. The actual numbers of spin-α and spin-β electrons respectively are controlled by the keys charge and occupations.Not e carefully, that using *only* the keyword unrestricted, without either Charge or Occupations (or both) would not result in any spin polarization. This implies that you would effectively perform a spin-restricted calculation, but with increased computational effort. Therefore, the program will check that in an unrestricted calculation at least one of the keys Charge and Occupations is applied.

The unrestricted feature is equivalent with, in *ab-initio* terminology, (Spin-)Unrestricted-Hartree-Fock (UHF); the N-particle wave function is a single determinant and not necessarily an eigenfunction of the spin operator S^2.

A *restricted* calculation implies that the (spatial) orbitals *and* the occupation numbers are identical for spin-α and spin-β.

The Fock operator, both in an unrestricted and in a restricted run, commutes with the spin operator Sz, but not (unless accidentally) with S^2. The obtained one-determinant wave function may for instance be a mixture of a singlet and a triplet state.

In an unrestricted calculation the expectation value of S^2 is now computed in ADF (note 29 in [98]). The implementation of an evaluation of S^2 is not quite trivial. DFT is essentially a one-particle formalism, so the S-operator for the n-particle system has to be written out in single-particle operators [99]. The equations used in ADF to calculate the expectation value of S^2 can be found in Szabo and Ostlund [100]. Note that the so called exact value $(S_{exact})^2$, which is printed in the ADF output, is defined as $(S_{exact})^2 = (|N_a - N_b|/2)(|N_a - N_b|/2+1)$, where N_a and N_b are the number of spin-α and spin-β electrons, respectively. The expectation value of S^2 is not calculated in a Spin-Orbit coupled calculation.

Molecules that have been calculated using the unrestricted formalism cannot be employed as fragments. ADF will abort when you attach the TAPE21 result file from an unrestricted calculation as a fragment file.

A fair approximation to a computation with unrestricted fragments can be achieved with the key FRAGOCCUPATIONS. See also the Examples.

Unrestricted and Spin-Orbit Coupling

In the case of Spin-Orbit coupling there are two ways to do spin-polarized calculations, either using the collinear approximation or the noncollinear approximation [101, 102]. Using the unrestricted feature in order to assign different numbers of electrons to a and b spin, respectively, cannot be applied as such, if one includes Spin-Orbit coupling, since the electrons are not directly associated with spin-α and spin-β. For the collinear and noncollinear approximation one should use symmetry NOSYM, and each level can allocate 1 electron. Note that with the key CHARGE one should only specify one value, namely the total charge. One should not specify the spin-polarization.

```
COLLINEAR
```

This key is only relevant in the case of Spin-Orbit coupling. The key has no argument. See also the key NONCOLLINEAR.

In the collinear approximation in each point in space the spin-polarization has the same direction (default is in the direction of the z-axis). Kramer's symmetry does not have to be satisfied. Symmetry used in the calculation should be NOSYM. The default direction of the spin-polarization can be overruled using the key SOUX (this key has no argument) for spin-polarization only in the direction of the x-axis, and the key SOUY (this key has no argument) for spin-polarization only in the direction of the y-axis. Both keys SOUX and SOUY are only relevant in the case of Spin-Orbit coupling in combination with the key COLLINEAR.

```
NONCOLLINEAR
```

This key is only relevant in the case of Spin-Orbit coupling. The key has no argument. See also the key COLLINEAR.

In the noncollinear approximation in each point in space the spin-polarization can have a different direction. Kramer's symmetry does not have to be satisfied. Symmetry used in the calculation should be NOSYM.

Net Charge and Spin polarization

The net charge of the molecule and the net spin polarization can be controlled with the key CHARGE.

```
CHARGE {NetQ {ab}}
```

`NetQ`

The net total charge of the molecule

`ab`

The net total spin polarization: the number of spin-α electrons in excess of spin-β electrons. Specification is only meaningful in a spin-unrestricted calculation. However, specification is not meaningful in an unrestricted Spin-Orbit coupled calculation using the (non-)collinear approximation.

If the key is used, the first value in the argument is assigned to netQ, the net total charge, and the second to ab. If the key is not used at all, default values apply. The default for the net total charge is the *sum of fragment charges: not necessarily neutral!!* The fragment charges are the net total charges that were used in the fragment runs; this information is stored in the fragment files.

The default spin polarization is zero.

An unrestricted calculation with ab=0 (for instance by not specifying the spin polarization at all) is, in the case one does spin break the spin symmetry, in fact a restricted run: it should give exactly the same as the restricted calculation, but it will use more CPU time. If one does break the spin symmetry, for example with the key MODIFYSTARTPOTENTIAL or the SPINFLIP option in the key RESTART, the solution may also be a broken spin symmetry solution. For example one may want to start a calculation in broken symmetry with spin-α density on one fragment and spin-β density on another, e.g. in a spin-unrestricted calculation of H_2 at large separation.

Orbital occupations: electronic configuration, excited states

With the key OCCUPATIONS you can specify in detail the assignment of electrons to MOs

```
OCCUPATIONS Options
  {irrep orbitalnumbers
  irrep orbitalnumbers
```

```
   ...
End }
```

`Occupations`

is a *general* key: it has an argument or a data block. If you want to use both, the continuation code (&) must be appended at the end of the argument.

Aufbau, smearing, freezing

```
OCCUPATIONS Options
```

`Options`

May contain Keeporbitals, Smearq, Freeze, or Steep:

`Keeporbitals=NKeep`

Until SCF cycle Nkeep electrons are assigned to MOs according to the Aufbau principle, using at each cycle the then current orbital energies of the MOs. Thereafter the KeepOrbitals feature is applied. As soon as this is activated the program will on successive SCF cycles assign electrons to the MOs that maximally resemble - in spatial form - those that were occupied in a 'reference cycle number'. The default for Nkeep is 20, except:
a) When orbital occupations for MOs are specified explicitly in the data block of the occupations key, these apply throughout.
b) In a Create run fixed occupations are derived from a database in the program.
c) When electron smearing is explicitly turned on by the user (see the Smearq option below) Nkeep is by default 1,000,000 so the program will 'never' compare the spatial forms of MOs to determine the occupation numbers.
The 'reference cycle number' is by default the previous cycle, which will suppress jumps in the spatial occupations during the SCF development while at the other hand allowing the system to let the more-or-less-frozen configuration relax to self-consistency.

`Freeze`

Occurrence of this word in the option list specifies that the 'reference cycle number' will be the cycle number on which the KeepOrbitals feature is activated: during all subsequent SCF cycles the program will assign electrons to MOs that resemble the MOs of that specific SCF cycle. This may be used when the MOs of that cycle are already reasonably close to the final ones, and it will suppress unwanted step-by-step charge-transfers from occupied to empty orbitals that are very close in energy. By default this option is not active.

`Smearq=Smear1[,Smear2,Smear3,...,Smear10]`

Smear*N* is half the energy width (in hartrees) over which electrons are smeared out over orbitals that lie around the fermi level and that are close in energy. Smearing is a trick that may help when the SCF has problems converging. One should be well aware that the physical meaning of a result obtained with smeared occupations is unclear (to express it mildly). It may be useful to get over a hurdle in a geometry optimization.
By default the *initial* smear parameter is zero (i.e.: smearing is not applied). It is turned on automatically by the program when SCF convergence is found to be problematic, but only in an optimization-type application (simple optimization, linear transit, transition state) when the geometry is not yet converged.
You can rigorously prohibit any smearing by specifying it explicitly with value zero. More generally: specifying the smear parameter makes the program to apply it always, but always with the input-specified value.

When a comma-delimited list of values is specified, after SCF has converged, the next value from the list is picked and the SCF is continued. This way one can specify a list of gradually decreasing values to get sort of annealing effect. NOTE: No spaces are allowed when specifying a list of values for Smearq.

`Steep=Lambda[,Nmax]`

The occupation number for each orbitals are updated according to steepest-descent method (Ref: F. W. Averill and G. S. Painter, Phys. Rev. B **46**, 2498 (1992)). During an SCF cycle, the occupation number for each new orbital is initially determined by decomposing the old charge density with new orbitals. Then, the occupation numbers are modified so that the total energy of the system will decrease.
The Lambda parameter gives the coefficient for the charge transfer in 1/au unit. The second parameter, Nmax, is an additional limit for the amount of the charge transfer. Nmax would be useful for early steps of cycle when the Lambda parameter gives too large charge transfer. Too small Nmax results in irregular behavior in SCF convergence. In the case of difficult SCF convergence, you should make mixing and Lambda smaller. From our experience, Nmax=0.1 or 0.2 is usually OK.
This method should be used with turning off DIIS method (DIIS N=0), and the choice of the mixing parameter in SCF cycle is also important. This option is especially useful for systems with many quasi-degenerate orbitals around Fermi level. For instance, cluster models of surface systems usually suffer from dangling bonds and should be converged with this method. Note though that slow convergence is an intrinsic feature of this method so one should specify a large limit for the number of SCF cycles, say 500 or even 1000, depending on the cluster size.

Notes about the occupations options:

- When occupation numbers are explicitly defined via the block form of the OCCUPATIONS keyword (see next section), the Smearq option cannot be used.
- The aufbau principle does not determine or adjust the distribution of electrons over spin-α versus spin-β in an unrestricted calculation. This aspect is controlled by the key CHARGE and by any explicit occupations in the data block of occupations.
- When occupation numbers are not specified and no Smearing is specified either, the program will turn on smearing automatically when the SCF has serious convergence problems, in an attempt to overcome those problems, but only in a geometry optimization (including transition state, linear transit, etc.). If such happens the program restores the original situation (no smearing) at the start of each new SCF. In automatic smearing the smear parameter is initiated at 0.01 hartree and may be varied (by the program) between 0.001 and 0.1 hartree. The automatic use of smearing by the program can be prohibited by explicitly setting the smear option with value zero (Smearq=0).
- Smearing cannot be used in combination with the keeporbitals option. This option therefore also turns of *automatic* smearing in troublesome SCF 's during an optimization.

Explicit occupation numbers

```
OCCUPATIONS
  irrep orbitalnumbers
  irrep orbitalnumbers
  ...
End
```

`irrep`

The name of one of the irreducible representations (not a subspecies) of the point group of the system. See the Appendix 5.3 for the irrep names as they are used in ADF.

`orbitalnumbers`

A series of one or more numbers: the occupation numbers for one-electron *valence* orbitals in that irrep. The orbitals are ordered according to their energy eigenvalue; higher states than those listed get an occupation number zero.

For degenerate representations such as the 2-dimensional E-representations or the 3-dimensional T-representations, you must give the *total* occupation, i.e. the sum over the partner representations; ADF assigns each partner an occupation equal to the appropriate fraction of what appears here.

In an unrestricted calculation, two sequences of numbers must be specified for each irrep; the sequences are separated by a double slash (//). The first set of numbers is assigned to the spin-α orbitals, the second set to the spin-β orbitals. Example unrestricted calculation in symmetry NOSYM with two unpaired electrons:

```
OCCUPATIONS
  A 28 // 26
End
CHARGE 0 2
SYMMETRY NOSYM
```

Note that this is not meaningful in an unrestricted Spin-Orbit coupled calculation using the (non-)collinear approximation, where one should use one sequence of occupation numbers for each irrep.

Notes about the occupations data block:

- When specifying electron configurations, all valence electrons in the calculation must be explicitly assigned to MOs and the block form of the OCCUPATIONS keyword must be used. In this context the concept *valence electrons* and hence *valence orbitals* is not necessarily identical to what you may normally assume to be the valence space of an atom or molecule. The meaning of *valence* is here strictly defined as whatever electrons are outside the frozen core. It depends therefore on the level of frozen core approximation applied in the calculation. This traces back to the Create runs in which the basic atoms were generated that are now used to build the molecule.
- When for some irrep there is a rather long list of occupation numbers, corresponding to *consecutive fully occupied* states, you can combine these numbers and enter their sum instead: ADF knows the maximum occupation for an irrep, and when you put a larger number the program will split it up. For instance, if you give for the *p*-representation (in a single atom calculation):
  P 17 3
  ADF will interpret this as
  P 6 6 5 3
  i.e. the occupation number 17 is interpreted as denoting two fully occupied p-shells and the remaining five electrons in the next higher shell. This example also illustrates how to specify an excited state: here we have defined a hole in the third p-shell.
- Fractional occupation numbers in input are allowed. For a discussion of the interpretation of fractional occupation numbers see [103]. The program even allows you (technically) to use a non-integer total number of electrons, whatever the physical meaning of such a calculation is.
- The data block of occupations is not parsed (see the section Interpretation of Input). The program does not replace expressions by their value and it does not recognize constants or functions defined with the define key.
- In a numerical frequencies run (without the Symm argument) the symmetry used internally in the program is NOSYM, irrespective of any Schönfliess symbol in the input file. As a consequence the program will recognize only the A representation (the only irrep in nosym), but not the representations belonging to the input point group symmetry. (The symmetry in the equilibrium geometry, defined by the input Schönfliess symbol, is used to enhance efficiency and stability in the construction of the matrix of Force constants).

CHARGE vs. OCCUPATIONS

The contents of the data block of occupations, if used, defines the total number of valence electrons and hence the net total charge. In an unrestricted run it also defines the net spin polarization. If the key CHARGE is also used, the program will check that both specifications are consistent.

We strongly recommend to employ this and always specify the net total charge and spin polarization with charge whenever explicit occupation numbers are supplied with occupations, to that the program will check that your occupation numbers result in the total charge and spin polarization that you have in mind.

Create mode

In Create mode occupation numbers are predefined (see Appendix 5.2 Elements of the Periodic Table), and these are applied unless you specify occupations in input yourself. Conceivably this may result in a non-aufbau configuration. In Create mode the program always operates as if occupations were set in input.

Multiplet States

Calculations with ADF yield results for one-determinant electronic states, which are not always the 'true' states of the molecule. The evaluation of the correct multiplet energies is not trivial in this approach, see further below the section on multiplet energies. The point is to evaluate a specific multiplet state as a linear combination of selected one-determinant functions, each computed in the field of the so-called Average-of-Configuration (AOC). Typically, in an open shell system, the AOC is the spin-restricted system in which all orbitals in the open shell are degenerate and equally occupied. The AOC serves then as a fragment for the subsequent calculations, in which the different open shell orbitals are occupied differently by specifying the appropriate occupation numbers as explained below.

Important: in these follow-up calculations it is imperative that the results are obtained in the AOC field: no SCF convergence must be carried out, because we only want to assign the electrons differently, while keeping exactly the AOC orbitals. To achieve this, the follow-up calculations must use the keyword SCF, and the subkey iterations must be set to 0.

Since ADF requires that the point-group symmetry matches not only to the nuclear frame but also to the electronic charge density and MO occupations, these calculations must run in a lower pointgroup symmetry. Often you will also want to run the modified calculations spin-*un*restricted. For an example, see the set of sample runs that come with the package and the discussion in the Examples document.

The calculation of the one-determinant states based on the AOC reference state is controlled with the key SLATERDETERMINANTS: . It is a *general* key; it can be used as a simple key and requires an argument then. It can also be used as a block key. For this particular key it is not correct to specify an argument *and* a data block.

```
SLATERDETERMINANTS file
```

When used as a simple key, the argument must be a file (including the path). The file must be an ASCII file containing data in the same format as you would supply in the data block when using the key as block type key, see below. All information on the file until the *eof* must be suitable for the data block, but no record 'end' on the file must be specified: only the *contents* of the data block.

The block format:

```
SLATERDETERMINANTS file
 title1
 irrep occups
```

```
 irrep occups
 ....
 subend
 title2
 irrep occups
 .....
 subend
 title3
 ....
 subend
 ....
end
```

Each 'title' functions as a subkey, but is otherwise an arbitrary string to label the resulting one-determinant calculation. Each such subkey block contains the occupation numbers for a single one-determinant calculation. It is necessary that the calculation uses the reference AOC run as its only fragment file. The occupations in the subkey blocks must be re-arrangements of the AOC open-shell electrons. In the Slaterdeterminants calculation you must explicitly specify the pointgroup symmetry in which you want to run; this must be a lower symmetry than the AOC one, otherwise you couldn't rearrange the open shell electrons. See the sections below on multiplet energies. An sample run is included in Examples document.

Each 'irrep occups' record specifies the occupations for the indicated irrep in the usual way (see for instance the occupations key). The irrep labels must correspond to the (lower) point group symmetry used in the slaterdeterminants calculation. Note that in an unrestricted calculations, occupations numbers must be given for both spins, using the double slash (//) to separate the occupations for spin-α and spin-β.

In this setup, the program will for each of the subkey blocks under the slaterdeterminants key execute an SCF calculation with only one cycle, i.e. no convergence, where the start-up field is the fragment field, i.e. the AOC field. So all one-determinant states in this calculation are evaluated in the AOC field. The resulting energies for the distinctly computed one-determinant states can then be combined to the desired multiplet values, corresponding to how the multiplet states are combinations of the one-determinant states.

Multiplet energies

The energies of atomic and molecular multiplet states that correspond to a given electron configuration can be calculated approximately with the method suggested in ref. [323]. There it is indicated that it would not be justified to take an arbitrary configuration-state function (CSF), defined in general as a linear combination of determinants that has specific spin and space symmetry properties, and use the corresponding alpha and beta spin densities in a DFT energy expression. The same holds true for the densities corresponding to the average-of-configuration (see section 'DFT energy of a one-determinantal wavefunction', the 'procedure' notes).

Therefore, we restrict ourselves to just computing the DFT energies of single-determinant wavefunctions. Usually (but not always) this is sufficient information to obtain the multiplet energies. The procedure, which is explained in [323], requires knowledge of the CSFs belonging to a given configuration. This means that a multiplet state with specific L, M_L and S, M_S values has to be written as a linear combination of the determinant wavefunctions that belong to the given configuration. With the auxiliary program ASF (*A*df *S*ingle-determinants *F*ribourg, developed by Claude Daul in Fribourg, Switzerland) all the CSFs can be obtained, printed as linear combinations of the determinants [324]. The inverse transformation yields the determinants written as linear combinations of the CSFs.

It is often advantageous to search for CSFs that consist of one determinant only, since the energy of this determinant should correspond directly to the multiplet energy. Sometimes there is redundancy in the information and there may even be some inconsistency: two determinants may exist that both are CSFs belonging to the same multiplet state but yield somewhat different energies. We will illustrate this for the Carbon atom example treated below.

The discrepancies are a measure of 'error bars' associated with the theoretical multiplet energies. As a matter of fact, there are certain symmetry relations between the energies of the determinants of a configuration, calculated as the expectation value of the determinant for the full many-electron Hamiltonian. An example is the equal energy for the determinants of a p^1 configuration, whether the electron is placed in the p_0 ($=p_z$) orbital or in the p_{+1} ($=(p_x+ip_y)/\sqrt{2}$) orbital. This equality is not obtained with present-day density functionals, leaving an ambiguity ('error bar') in the determination of the energy. A more complete treatment of the symmetry relations between determinant energies is given in [324].

The auxiliary program ASF, that for finite point groups finds the CSFs as linear combinations of determinants, performs also a symmetry analysis of all the two-electron integrals for a configuration, reducing them to a minimum number of non-redundant ones. ASF expresses the energies of the multiplets in the non-redundant two-electron integrals. However (WARNING!), there have occasionally been found inconsistent results. A comparison to the results obtained by the procedure outlined in [323] may show significant differences and the latter seem more accurate and consistent.

DFT energy of a one-determinantal wavefunction

The determinant corresponds to a well defined ρ^α and ρ^β. Suppose we are dealing with a d^2 configuration and we wish to know the energy of

$$D_1 = |d_2^\alpha(1)\, d_1^\alpha(2)|$$

where d_m has the Y_{2m} angular part. This determinant is a CSF of the ^3F multiplet:

$$D_1 = |^3F;M_L=3;M_S=1>$$

We can easily transform to the real spherical harmonics that are used in ADF:

$$Z_{lm}^c = 1/\sqrt{2}\,(Y_l^{-m} + Y_l^{-m^*}) = 1/\sqrt{2}\,[Y_l^{-m} + (-1)^m Y_l^m]$$

$$Z_{lm}^s = i/\sqrt{2}\,(Y_l^{-m} + Y_l^{-m^*}) = i/\sqrt{2}\,[Y_l^{-m} + (-1)^m Y_l^m]$$

with back transformations:

$$Y_l^m = 1/\sqrt{2}\,(-1)^m\,[Z_{lm}^c + i\, Z_{lm}^s]$$

$$Y_l^{-m} = 1/\sqrt{2}\,[Z_{lm}^c - i\, Z_{lm}^s]$$

Here the superscripts c and s stand for the cosine, respectively sine type of combinations of exp(-imφ) and exp(imφ). This yields explicitly:

$$d_{z^2} = d_0$$

$$d_{xz} = 1/\sqrt{2}\,(d_{-1} - d_1)$$

$$d_{yz} = i/\sqrt{2}\,(d_{-1} + d_1)$$

$$d_{x^2-y^2} = 1/\sqrt{2}\,(d_{-2} + d_2)$$

$$d_{xy} = i/\sqrt{2}\,(d_{-2} - d_2)$$

$d_0 = d_{z^2}$

$d_1 = -1/\sqrt{2}\ (d_{xz} + i\ d_{yz})$

$d_{-1} = 1/\sqrt{2}\ (d_{xz} - i\ d_{yz})$

$d_2 = 1/\sqrt{2}\ (d_{x^2-y^2} + i\ d_{xy})$

$d_{-2} = 1/\sqrt{2}\ (d_{x^2-y^2} - i\ d_{xy})$

For D_1 we obtain:

$$\rho^{\alpha} = |d_2|^2 + |d_1|^2 = 1/2\ |d_{x^2-y^2}|^2 + 1/2\ |d_{xy}|^2 + 1/2\ |d_{xz}|^2 + 1/2\ |d_{yz}|^2$$

$$\rho^{\beta} = 0$$

The fractional occupations have to be used in order to generate the densities ρ^{α} and ρ^{β} and the corresponding density matrices P^{α} and P^{β}. The density matrices can be used to calculate the energy of D_1 (and ^3F) with respect to the energy of the 'master fragment', which should be the restricted atom with d^2 occupation. Other determinants of this configuration can be treated similarly to obtain more multiplet energies of the d^2 configuration.

Below is an example of how you can obtain determinant energies 'by hand', i.e. by carrying out a specific sequence of ADF calculations. ADF supports an automatic procedure to do this, using the input key `SLATERDETERMINANTS`, see the ADF User's Guide, the Examples document, and below.

Procedure

1 Determine a set of orbitals belonging to the given configuration. These orbitals are generated in what we call the average-of-configuration (AOC) calculation. This is a spin-restricted SCF calculation where the electrons of the configuration are distributed equally over the subspecies of the open shell irreps in order to retain the A1 symmetry of the total density in the symmetry group of the molecule. For instance, in case of an atomic d^2 configuration, the AOC calculation can be done in symmetry atom with occupation 2 in the d irrep. In case of an $t_{2g}^5\ e_g^1$ configuration of an octahedral complex, the AOC calculation requires an occupation of 5 electrons in the t_{2g}, and 1 electron in the e_g.
The result file TAPE21 of the AOC calculation has to be saved, to be used as a fragment file in the subsequent calculations.

2 The AOC is used as a fragment in all subsequent calculations that are performed to obtain single determinant energies. This means that those single determinant energies are always with respect to the AOC energy. This is a case where there is only one "fragment", which is actually the complete system, but in a different electronic configuration and in possibly a different symmetry group.

Suppose that a single determinant corresponds to spin-up and spin-down densities ρ^{α} and ρ^{β}, i.e. to specific spin-unrestricted occupations of the AOC orbitals. These densities ρ^{α} and ρ^{β} correspond to a symmetry group that will in general be a subgroup of the symmetry group of the molecule. For instance, the occupation $(p_{+1}\alpha)^1$ in the case of an atomic p^1 configuration corresponds to

$$\rho^{\alpha} = 1/2\ p_x^2 + 1/2\ p_y^2$$

with $D_{\infty h}$ symmetry.

To obtain the energy of the determinant wavefunction we must now perform *one* cycle (iterations= 0 in the block key SCF) of a spin-unrestricted calculation, with AOC as (the only) fragment with alpha and beta occupation numbers (using the input key occupations) such that ρ^{α} and ρ^{β} result. Note that the appropriate (lower) symmetry pointgroup must be specified in the input file.

Occasionally, the single determinant corresponds to a closed shell configuration in the appropriate lower symmetry, for instance determinant $D_{10} = |0^{+}\ 0^{-}|$ of the p^2 configuration of Carbon, with density $r=p_z^2$. In that case the one-cycle calculation can of course be spin-restricted.

N.B.1. One cycle will regenerate the SCF orbitals of AOC, if the same field is used as the converged AOC field. This will actually be the case because the starting potential is taken from the fragment TAPE21 file. The key modifystartpotential must not be used (the density should be distributed equally over the spins).

N.B.2. After diagonalization in the one-cycle run, the AOC orbitals have been obtained again and are occupied as specified. The ('bonding') energy is calculated from the resulting charge density.

Remarks:

- If one does not perform just one cycle, but instead converges the unrestricted calculation, the energy will be lowered by 'polarization' of the orbitals.
  It is theoretically not so clear what the status of such converged energies is.
  Usually the energy lowering is in the order of 0.1 eV, but it may be quite a bit larger.
- It is not necessary to use AOC as fragment in the single-determinant runs.
  It is also perfectly allowed to run all calculations (ground state, AOC, determinants) from one set of fragments, for instance the standard atomic fragments.
  Since we must arrange that the one-cycle determinant calculations use the AOC field, so as to reproduce the AOC orbitals, we must then supply the result file TAPE21 of the AOC run as a restart file, using the key restart; see the adf User's Guide.
  Of course, in such an approach the computed energies are with respect to another reference, for instance the restricted atoms.

Results for first period atoms

In one of the next sections tables are given for the energy lowering in going from the converged spherically symmetric spin-restricted atom (the 'master' fragment) to specific one-determinant wavefunctions with the orbital occupations as specified. Note that the p_x and p_y populations are always equal; only their sum is given. In many cases the determinant corresponds to a specific state, which is then given in the last column. For each atom, the first calculation is for the spherically symmetric spin-*un*restricted atom. These tables are now obsolete, all information needed to obtain the atomic reference energies, i.e. the groundstate multiplet energy with respect to the AOC, can be found in ref. [325].

Examples worked out for all first period atoms:

H: Configuration $(1s)^1$.
Only one determinant: |1sα(1)|

He: Configuration $(1s)^2$.
Closed shell.

Li: Configuration $(2s)^1$.
Only one determinant: |2sα(1)|

Be: Configuration $(2s)^2$.
Closed shell.

B: Configuration $(2p)^1$.

Ground multiplet ^2P.

$D_1 = |p_1\alpha(1)| = |^2P;M_L=1;M_S=1/2>$

$\rho^\alpha = |p_1|^2 = (-p_x/\sqrt{2} - ip_y/\sqrt{2})^* (-p_x/\sqrt{2} - ip_y/\sqrt{2})$

$= 1/2(p_x - ip_y)(p_x + ip_y) = 1/2(p_x^2 + p_y^2)$

The occupation numbers for D_1 are

$p_x^\alpha = p_y^\alpha = 1/2$; $p_z^\alpha = 0$; $p_x^\beta = p_y^\beta = p_z^\beta = 0$

Another determinant belonging to ^2P is
$D_2 = |..p_0\alpha(1)|$

with occupations $p_z^\alpha = 1$ and all other p-occupations zero. This determinant is 0.04 eV lower in energy than D_1 for LDA, but 0.15 eV for BP.

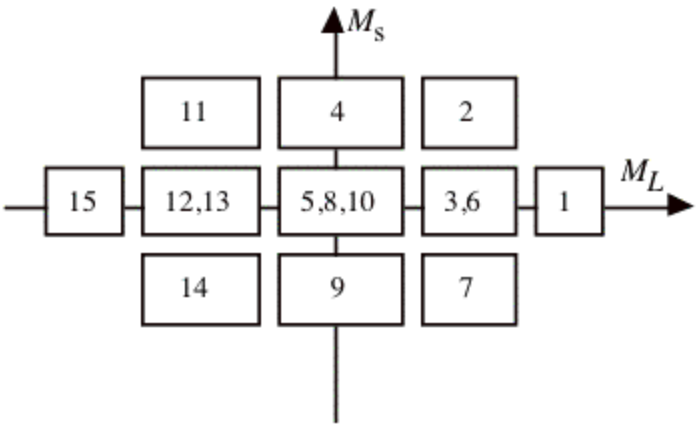
C: Configuration p^2.

Multiplet states are ^3P, ^1S and ^1D.

We use this atom as an example of the general procedure. First write down all determinants belonging to p^2 and group them according to M_S and M_L

($1^+ \equiv p_1\alpha$, ...)

Determinant	M_S	M_L
$D_1 = \|1^+ 1^-\|$	0	2
$D_2 = \|1^+ 0^+\|$	1	1
$D_3 = \|1^+ 0^-\|$	0	1
$D_4 = \|1^+ -1^+\|$	1	0
$D_5 = \|1^+ -1^-\|$	0	0
$D_6 = \|1^- 0^+\|$	0	1
$D_7 = \|1^- 0^-\|$	-1	1
$D_8 = \|1^- -1^+\|$	0	0
$D_9 = \|1^- -1^-\|$	-1	0
$D_{10} = \|0^+ 0^-\|$	0	0
$D_{11} = \|0^+ -1^+\|$	1	-1
$D_{12} = \|0^+ -1^-\|$	0	-1
$D_{13} = \|0^- -1^+\|$	0	-1
$D_{14} = \|0^- -1^-\|$	-1	-1
$D_{15} = \|-1^+ -1^-\|$	0	-2



The presence of a determinant with $M_L = 2$, $M_S = 0$ and no $M_L = 2$, $M_S > 0$ determinant indicates the presence of a ^1D multiplet, and $E(^1D) = E(D_1)$.

There is also a ^3P, the determinant with $M_S = 1$, $M_L = 1$. We should have $E(^3P) = E(D_2) = E(D_4)$. The two determinants D_3 and D_6 in the $M_S = 0$, $M_L = 1$ box will mix, and the solutions of the 2 by 2 secular problem will be $E(^1D)$ and $E(^3P)$. Since the sum of the eigenvalues is equal to the sum of the initial diagonal elements, we have $E(^1D) + E(^3P) = E(D_3) + E(D_6)$.

We should also have $E(D_3) + E(D_6) = E(D_1) + E(D_2)$. Such a relation provides a test on the consistency of the results.

Finally we have the $M_S = 0$, $M_L = 0$ block. The sum of the energies of D_5, D_8 and D_{10} should be the sum of the energies of ^1S, ^3P and ^1D. Since $E(^3P)$ and $E(^1D)$ are already known, $E(^1S)$ can be calculated.

In the appendix we first locate for C the spherical unrestricted atom. Next we have $E(D_4)$, yielding $E(^3P)$ = -1.345 eV (LDA + Becke). Next $E(D_2) = E(^3P)$ = -1.189 (always LDA + Becke). The difference is substantial: ~ 0.15

Next we have $E(D_3)$ = - 0.812. Since $E(D_6) = E(D_3)$, because $\rho^{\alpha}(D_6) = \rho^{\beta}(D_3)$ and $\rho^{\beta}(D_6) = \rho^{\alpha}(D_3)$, we should have $2E(D_3) = -1.624 = E(^1D) + E(^3P)$. Therefore $E(^1D)$ = -1.624 - (-1.345) = - 0.279 or $E(^1D)$ = -1.624 - (-1.189) = - 0.435.

These numbers can be checked against $E(D_1)$ which also should be $E(^1D)$:
$E(D_1)$ =+ 0.044. The discrepancy is large!

Finally, ^1S can be obtained:
$E(D_{10})$ = + 0.319
$(D_8) = E(D_1)$ = + 0.044
$E(D_5) = E(D_1)$ = + 0.044

So $0.407 = E(^1S) + E(^3P) + E(^1D)$.

Different results for $E(^1S)$ are obtained depending on the $E(^3P)$ and $E(^1D)$ we choose:

e.g. $E(^1S)$ = 0.407 -(-1.345) - (- 0.279) = 2.031

or $E(^1S)$ = 0.407 - (-1.189) - (0.044) = 1.552.

Comparing to experiment we might calculate the excitation energies w.r.t. $E(^3P)$:

	calculated	**experimental**	**HF**
$^3P \rightarrow {}^1D$:	1.066 to 1.389	1.26	1.55
$^3P \rightarrow {}^1S$:	2.741 to 3.376	2.684	3.78

N: Configuration p^3.

Ground multiplet ^4S, corresponds to the spherical unrestricted atom, energy -2.943 eV. Other multiplets: ^2P, ^2D. According to the printed output for configuration p^3 we have

$|^2D;M_L=2;M_S=1/2> = |p_1^{\alpha}\ p_1^{\beta}\ p_0^{\alpha}| = D_2$

$\rho^{\alpha} = 1/2\ p_x^2 + 1/2\ p_y^2 + p_z^2$

$\rho^{\beta} = 1/2\ p_x^2 + 1/2\ p_y^2$

$E(D_2)$ = - 0.745 according to the table in the Appendix (LDA + Becke)

The energy of D_1, with $\rho^{\alpha} = p_x^2 + p_y^2$, $\rho^{\beta} = p_z^2$, is E(|1A 2B 3A|) = -1.9702.

The energy of D_3, with $\rho^{\alpha} = 1/2\ p_x^2 + 1/2\ p_y^2 + p_z^2$, $\rho^{\beta} = p_z^2$ corresponding to |1A 2A 2B| or |2A 2B 3A|, is $E(D_3)$ = - 0.158.

Finally, D_4 has $\rho^{\alpha} = p_x^2 + p_y^2$, $\rho^{\beta} = 1/2\ p_x^2 + 1/2\ p_y^2$, corresponding to |1A 1B 3A| and |1A 3A 3B|, and $E(D_4)$ = - 0.109.

The M_L=1, M_S=1/2 determinants are |1A 1B 3A| and |1A 2A 2B|. Therefore $E(^2D) + E(^2P) = E(D_4) + E(D_3)$, so $E(^2P)$ = - 0.109 - 0.158 - (- 0.745) = + 0.478.

We can use D_1 in the M_L=0, M_S=1/2 block, from which we find $E(^4S) + E(^2D) + E(^2P) = 2E(D_2) =+ E(D_1)$.

Hence $E(^2P)$ = -1.490 - 1.9702 - (- 0.745) - (- 2.943) = + 0.2278.

O: Configuration p^4.

Multiplet states ^3P, ^1S, ^1D.

D_1, with $\rho^{\alpha} = p_x^2 + p_y^2 + p_z^2$, $\rho^{\beta} = p_z^2$ corresponds to |1A 2A 2B 3A| , the M_L=1, M_S=1 determinant of ^3P: $E(^3P)$ = -1.836

D_2 with $\rho^{\alpha} = p_x^2 + p_y^2 + p_z^2$, $\rho^{\beta} = 1/2\ p_x^2 + 1/2\ p_y^2$, corresponds to |1A 1B 2A 3A|, the determinant of ^3P: $E(^3P)$ = -1.568

D_3, with $\rho^{\alpha} = 1/2\ p_x^2 + 1/2\ p_y^2 + p_z^2$, $\rho^{\beta} = p_x^2 + p_y^2$, corresponds to |1A 1B 2A 3B|, and M_L=1, M_S=0 belonging to ^3P as well as ^1D.

F: Configuration p^5.

Ground multiplet ^2P.

As in B, we have two determinants with different energies belonging to this state:

$D_1 = |...(p_0^{\alpha})^1\ (p_0^{\beta})^0| \rightarrow E(D_1)$ = - 0.715.

$D_2 = |(p_{-1})^2\ (p_0)^2\ (p_1^{\alpha})^1\ (p_1^{\beta})^0| \rightarrow E(D_2)$ = - 0.467.

Ne: Configuration p^6.

Closed shell.

Ground and Excited State Multiplets

The computation of multiplets can be carried out with adf, using the input key `SLATERDETERMINANTS`.

The method described in [323] for the calculation of the energies of proper spin and spatial symmetry adapted Configuration State Functions is based on the calculation of the energies of single determinantal wavefunctions. The densities corresponding to those Slater determinants are inserted in the approximation used for the exchange-correlation energy.

The procedure as detailed above is somewhat involved. Moreover, one would like to have an easy procedure to calculate many determinants. This is particularly desirable if one wishes to calculate the energies of all Slater determinants of a given configuration, for instance if one wishes to calculate certain averages in view of the inconsistencies of the method.

We have implemented a semi-automatic procedure, using the key SLATERDETERMINANTS.

The general idea of this method is to first perform a restricted calculation in the symmetry that is appropriate for the molecule. This is called the average-of-configuration (AOC) calculation. This AOC calculation generates the orbitals which will be used in all the Slater determinants.

The AOC is the fragment that *must* be used in all subsequent calculations. The subsequent calculations are characterized by having the AOC as the (only) fragment, and by specifying the keyword SLATERDETERMINANTS. This key is a general key: it may be used as a simple key (it must then have an argument) or as a block type key (no argument, but a data block). If an argument is given this must be a file name. The named file should contain the occupations for the determinants (see below). If no file name is specified the key should be a block key and the occupations should be specified in the data block.

The required information, on file or in the data block, is the specification of the determinant or determinants that are to be calculated in the form of orbital occupation numbers for the AOC orbitals, using the irrep labels of the point group of the AOC calculation, see below for format. All specified determinants will be calculated, and the obtained energy will always be the energy difference with respect to the AOC. Default occupations for all subspecies of the AOC fragment are the occupations of the AOC itself. Therefore, only the open (modified) subspecies have to be specified.

One has to be careful with respect to the point group symmetry to use in the SLATERDETERMINANTS calculation. The density belonging to a specific determinant is usually lower than the AOC symmetry (which is the full symmetry group of the system). In that case this lower point group symmetry has to be specified in the SLATERDETERMINANTS calculation. Everything will always work fine if one just does not use any symmetry at all (nosym). However, if for reasons of computational efficiency one does want to use a subgroup of the system that corresponds to the determinant density or densities, this is perfectly possible. However, all the densities of the determinants specified *must* then have this (or a higher) symmetry. The program does not check on this, it is the user's responsibility to make sure that this condition is satisfied for all the determinants. The only check that is performed is that occupations of equivalent representations (subspecies of one irrep) in the lower point group of the SLATERDETERMINANTS run, that result from the specified occupations of the subspecies of the AOC symmetry, are equal.

Format of the input.

Important: in the SlaterDeterminants calculations it is imperative that the results are obtained in the AOC field: no SCF convergence must be carried out, because we only want to assign the electrons differently, while keeping the AOC orbitals exactly as they are. To achieve this, the calculations should use the keyword SCF, and the subkey iterations has to be set to 0 in the SCF data block.

Since adf requires that the point-group symmetry conforms not only to the nuclear frame but also to the electronic charge density and mo occupations, these calculations must run in a lower pointgroup symmetry. Often you will also want to run the modified calculations spin-*un*restricted.

For an example, see the set of sample runs that come with the package and the comments in the Examples document.

The calculation of the one-determinant states based on the AOC reference state is controlled with the key SLATERDETERMINANTS. It is a *general* key; it can be used as a simple key and requires an argument then. It can also be used as a block key. For this particular key it is not correct to specify an argument *and* a data block.

```
SLATERDETERMINANTS file
```

When used as a simple key, the argument must be a file (including the path). The file must be an ascii file containing data in the same format as you would supply in the data block when using the key as block type key, see below. All information on the file until the *eof* must be suitable for the data block, but no record 'end' on the file must be specified: only the *contents* of the data block.

The block format:

```
SLATERDETERMINANTS
   TITLE1
        irrep   occups
        irrep   occups
        ....
   SUBEND
   TITLE2
        irrep   occups
        ....
   SUBEND
   TITLE3
        irrep   occups
        ....
   SUBEND
   ....
END
```

Each 'title' functions as a subkey, but is otherwise an arbitrary string to label the resulting one-determinant calculation. Each such subkey block contains the occupation numbers for a single one-determinant calculation. It is necessary that the calculation uses the reference AOC run as its only fragment file. The occupations in the subkey blocks must be re-arrangements of the AOC open-shell electrons. In the `SLATERDETERMINANTS` calculation you must explicitly specify the pointgroup symmetry in which you want to run. The ρ^{α} and ρ^{β} densities of all determinants to be calculated must have this point group symmetry, or a higher symmetry.

Each 'irrep occups' record specifies the occupations for the indicated irrep in the usual way (see for instance the occupations key). The irrep labels must correspond to the AOC point group symmetry used in the AOC calculation, *not the point group symmetry used in the `SLATERDETERMINANTS` calculation!*. Note that in an unrestricted calculations, occupations numbers must be given for both spins, using the double slash (//) to separate the occupations for spin-alpha and spin-beta.

In this setup, the program will for each of the subkey blocks under the `SLATERDETERMINANTS` key execute an SCF calculation with only one cycle, i.e. no convergence, where the start-up field is the fragment field, i.e. the AOC field. So all one-determinant states in this calculation are evaluated in the AOC field. The resulting energies for the distinctly computed one-determinant states can then be combined to the desired multiplet values.

Example: Carbon p^2

```
SlaterDeterminants
C(p2) ALFA: s=1, px=py=2/3, pz=2/3; BETA: s=1, p=0 ! title
   S 1 // 1 ! irrep name and occupations
   P:x 0.666666666666666666 // 0 ! another irrep, et cetera
   P:y 0.666666666666666666 // 0
   P:z 0.666666666666666666 // 0
   D:z2 0 // 0
   D:x2-y2 0 // 0
   D:xy 0 // 0
   D:xz 0 // 0
   D:yz 0 // 0
SUBEND
C(p2) ALFA: S=1, px=py=1, pz=0; BETA: s=1 !next (Sl.Determinant) title
   S 1 // 1
   P:x 1 // 0
   P:y 1 // 0
```

```
   P:z 0 // 0
   D:z2 0 // 0
   D:x2-y2 0 // 0
   D:xy 0 // 0
   D:xz 0 // 0
   D:yz 0 // 0
SUBEND
C(p2) ALFA: s=1, px=py=0.5, pz=1; BETA: s=1
   S 1 // 1
   P:x 0.5 // 0
   P:y 0.5 // 0
   P:z 1 // 0
   D:z2 0 // 0
   D:x2-y2 0 // 0
   D:xy 0 // 0
   D:xz 0 // 0
   D:yz 0 // 0
SUBEND
C(p2) ALFA: s=1, px=py=0.5, pz=0; BETA: s=1, px=py=0, pz=1
   S 1 // 1
   P:x 0.5 // 0
   P:y 0.5 // 0
   P:z 0 // 1
   D:z2 0 // 0
   D:x2-y2 0 // 0
   D:xy 0 // 0
   D:xz 0 // 0
   D:yz 0 // 0
SUBEND
C(p2) ALFA: s=1, px=py=0.5, pz=0; BETA: s=1, px=py=0.5, pz=0
   S 1 // 1
   P:x 0.5 // 0.5
   P:y 0.5 // 0.5
   P:z 0 // 0
   D:z2 0 // 0
   D:x2-y2 0 // 0
   D:xy 0 // 0
   D:xz 0 // 0
   D:yz 0 // 0
SUBEND
C(p2) ALFA: s= 1, px=py=0, pz=1; BETA: s=1, px=py=0, pz=1
   S 1 // 1
   P:x 0 // 0
   P:y 0 // 0
   P:z 1 // 1
   D:z2 0 // 0
   D:x2-y2 0 // 0
   D:xy 0 // 0
   D:xz 0 // 0
   D:yz 0 // 0
SUBEND
```

In the example the AOC calculation is the Carbon atom in spherical symmetry (symmetry name atom).

Several spin states can be generated from this AOC set of orbitals, but they all have a lower symmetry than the AOC. In the example the point group $D_{\infty h}$ (DLIN) could be used in the `SLATERDETERMINANTS`

calculation. In $D_{\infty h}$ the *p* orbitals split into two sets, p_x and p_y occur in π_x and π_y respectively, so their occupations must be identical, and p_z is a Σ_u orbital.

In the data block of the `SLATERDETERMINANTS` key (or in the file) we now specify the occupations for the subspecies of the atom irreps of a specific Slater determinant and the program will sort out the corresponding occupations in the d(lin) symmetry.

In all cases the orbitals used for the energy calculation(s) will be the self-consistent AOC orbitals.

In the given example, the first set of occupations does not correspond to a Slater determinant, but is the spin-polarized spherical case with the *p* electrons evenly distributed over all components.

LDA results, with and without GGA (Becke-Perdew)

Energy changes (eV) for atoms going from restricted to (one-cycle) unrestricted. Results between parentheses are for *converged* unrestricted calculations)

All calculations have been performed in $D_{\infty h}$ symmetry, since p_x and p_y always had equal occupations and therefore could occur as π_u-x and π_u-y partners of the Π_u irrep.

El.	Occupations						LDA	LDA+Becke	BP
	alpha-spin			beta-spin					
	s	p_x+p_y	p_z	s	p_x+p_y	p_z			
H	1	0	0	0	0	0	-0.868 (-0.898)	-0.758 (-0.837)	-0.889 (-0.948)
Li	1	0	0	0	0	0	-0.231 (-0.235)	-0.195 (-0.207)	-0.249 (-0.256)
Be	1	0	0	1	0	0	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
B	1	2/3	1/3	1	0	0	-0.247 (-0.255)	-0.231 (-0.242)	-0.276 (-0.281)
	1	0	1	1	0	0	-0.295 (-0.321)	-0.436 (-0.474)	-0.448 (-0.485)
	1	1	0	1	0	0	-0.266 (-0.279)	-0.296 (-0.316)	-0.333 (-0.348)
C	1	4/3	2/3	1	0	0	-1.163 (-1.203)	-1.109 (-1.158)	-1.252 (-1.285)
	1	2	0	1	0	0	-1.152 (-1.211)	-1.271 (-1.345)	-1.372 (-1.436)
	1	1	1	1	0	0	-1.152 (-1.197)	-1.134 (-1.189)	-1.267 (-1.307)
	1	1	0	1	0	1	-0.462 (-0.506)	-0.726 (-0.812)	-0.778 (-0.868)
	1	1	0	1	1	0	0.159 (0.150)	0.039 (0.044)	0.087 (0.086)
	1	0	1	1	0	1	0.730 (0.668)	0.322 (0.319)	0.480 (0.450)
N	1	2	1	1	0	0	-2.936 (-3.032)	-2.827 (-2.943)	-3.101 (-3.190)
	1	2	0	1	0	1	-1.362 (-1.454)	-1.811 (-1.972)	-1.943 (-2.108)
	1	1	1	1	1	0	-0.581 (-0.618)	-0.688 (-0.745)	-0.746 (-0.801)
	1	1	1	1	0	1	0.178 (0.088)	-0.104 (-0.158)	-0.069 (-0.140)
	1	2	0	1	1	0	0.197 (0.135)	-0.077 (-0.109)	-0.011 (-0.053)
O	1	2	1	1	2/3	1/3	-1.400 (-1.477)	-1.361 (-1.447)	-1.480 (-1.552)
	1	2	1	1	0	1	-1.442 (-1.583)	-1.698 (-1.836)	-1.816 (-1.957)
	1	2	1	1	1	0	-1.422 (-1.515)	-1.470 (-1.568)	-1.590 (-1.678)
	1	1	1	1	2	0	-0.564 (-0.623)	-0.866 (-0.960)	-0.913 (-1.013)
	1	1	1	1	1	1	0.358 (0.321)	0.255 (0.237)	0.292 (0.266)
	1	2	0	1	2	0	1.323 (1.220)	0.825 (0.789)	0.992 (0.932)
F	1	2	1	1	4/3	2/3	-0.374 (-0.398)	-0.366 (-0.391)	-0.394 (-0.416)
	1	2	1	1	2	0	-0.323 (-0.409)	-0.605 (-0.686)	-0.627 (-0.715)
	1	2	1	1	1	1	-0.349 (-0.389)	-0.401 (-0.441)	-0.427 (-0.467)

Frozen core approximation

Frozen core vs. pseudopotentials

Pseudopotentials are not supported. The frozen core approximation is automatic in a normal (Fragment mode) calculation and is defined by the basic atomic fragments. The data file used in the Create run specifies the frozen core for the atom, which is then used in all molecules that incorporate that atomic fragment.

Core Potentials

In the standard approach the Coulomb potential and the charge density due to the atomic frozen core are computed from the frozen one-electron orbitals. ADF stores the computed core density and core potential for each atom type in the molecule on a file TAPE12. Alternatively, you may attach a file with (core) potentials and densities. The file must have the same structure as the standard TAPE12. It should contain one or more sections, each with the core information for one type of atom. With the key COREPOTENTIALS you specify the core file and (optionally) which sections pertain to the distinct atom types in the molecule. It is a general key that can be used as a simple key or as a block key.

Simple key:

```
COREPOTENTIALS corefile
```

Block key:

```
COREPOTENTIALS corefile &
  atomtype index
  atomtype index
  ...
end
```

`corefile`

The file with core potentials and charge densities. The name may contain a path.

`atomtype`

One of the atom type names as defined by atoms.

`index`

Points to the core section on the attached file that applies to the atom type. Different atom types may use the same section. A non-positive index tells the program that the atoms of that type don't have a frozen core. If the information on the corresponding fragment file (or data file in Create mode) indicates the contrary the program will abort with an error message.

If the key is used as a simple key (specifying only the core file) the sections on the file are associated with the atom types in order: the first section is used for the first atom type, et cetera. This is overruled by applying the block form. However, since the key *must* have the core file as argument, the block form requires that you apply the continuation symbol: an ampersand (&), separated from the core file name by at least one blank.

If you omit an atom type from the data block it gets a zero index (no core).

The attached file may contain more sections than used in the calculation, and the indices specified in the data block don't have to be in ascending order, consecutive, or cover a specific interval.

When a file with non-standard (e.g. relativistic) cores is attached and used in the calculation of an atom or molecule, and the result is used as fragment in a subsequent calculation, you should attach and use the same core potentials again. Otherwise, the program will internally compute the standard core potentials and hence implicitly employ another fragment than you may think, i.e. a fragment with other properties. ADF will not check anything in this respect and corepotentials should therefore be handled with great care.

The primary application of the corepotentials option is to include (scalar) relativistic corrections in the (frozen core part of the) Fock operator. The relativistic core potentials can be computed with the auxiliary program dirac (see the RELATIVISTIC keyword).

Spin-polarized start-up potential

The Coulomb and XC (exchange + correlation) potentials are computed from the fit approximation of the charge density (see Chapter 1.2).

The fit coefficients of this approximation for the first SCF cycle, needed to compute the first Fock matrix, are read from the fragment files: the start-up density is chosen as a sum-of-fragment-densities (fit approximations) and this combined density defines the initial potential.

In the SCF restart run the fit coefficients may be read from the attached TAPE21 file, see the key RESTART.

Spin-flip method for broken symmetries

Starting from ADF2009.01 it is possible to exchange alpha and beta electrons for selected atoms when performing a restart from a previous spin-unrestricted calculation.

In many cases, one wishes to perform a calculation of a low-spin complex where spin-density is positive on some atoms and negative on the others. It is usually very difficult to achieve SCF convergence if one starts from scratch. Sometimes, the MODIFYSTARTPOTENTIAL feature (see next section) helps with this problem but sometimes it does not. A more robust way is to first perform a high-spin calculation and then modify the resulting t21 file by "flipping" the spin on some atoms. This file then can be used to restart a subsequent low-spin calculation.

Such a "flipping" can now be performed during restart by specifying a SPINFLIP keyword in the RESTART input block as shown below:

```
RESTART high-spin.t21 &
! SpinFlip keyword is followed by atom numbers for
! which the flipping will be performed
  SPINFLIP 1
END
```

An example demonstrating the feature may be found in the Examples document.

Modify the starting potential

In some applications you may want to modify the initial fit coefficients (from the restart file or the fragment files), see also the previous section. This is achieved with the key MODIFYSTARTPOTENTIAL. It allows you to scale them in some way so as to represent user-chosen amounts of spin-α and spin-β fit density on some or all of the fragments. This will adjust the spin-α and spin-β initial potentials.

This option applies only to *unrestricted* calculations of course. It may be used to help the program find a particular state. This might, for instance, be hard to find otherwise due to the a-b symmetry in the start-up

situation. It may also be useful to speed up the SCF convergence in case you know what the final distribution of spin-α and spin-β density over the molecule will approximately be.

```
MODIFYSTARTPOTENTIAL {specification}
{ frag alfa // beta
 frag alfa // beta
 ...
end }
```

`ModifyStartPotential`

A *general* key: it has an argument *or* a data block.

`specification`

Must be *two numbers*, ASPIN and BSPIN, if provided at all. They specify the (relative) amounts of spin-α and spin-β fit density to define the spin-dependent potential at the first SCF cycle. The coefficients retrieved from the fragment files (or from the restart file in case of a SCF restart) are scaled accordingly. This will not affect the *total* amount of fit density: the absolute values of ASPIN and BSPIN play no role, only their ratio.
In case of a restart run the restart file must have been generated in a *restricted* calculation, while the continuation run must be an *unrestricted* one.

If no argument is given a data block must be supplied with records frag alfa // beta.
This is very much similar to the main option with ASPIN and BSPIN: you specify ASPIN and BSPIN now separately for each fragment. This involves somewhat more input but increases the possibilities to tune the initial potential. Again this can be applied only in an unrestricted calculation. It cannot be used in a restart: the affected fit coefficients are those from the fragment files, while in an SCF restart run these are ignored and replaced by the coefficients on the TAPE21 restart file.
Each line specifies a frag with its corresponding ASPIN and BSPIN fit partitioning. If frag is the name of a fragment *type*, the specified ASPIN-BSPIN is applied to all individual fragments of that type. Alternatively an *individual* fragment can be specified, using the format fragtype/n, where *n* is an index between one and the total number of fragments of that type. In such a case the ASPIN-BSPIN data applies only to that particular fragment while different values may be supplied for the other fragments of the same type.
It is allowed to specify for certain fragment types individual fragments and for other fragment types only the type. Duplicate specifications are not allowed; an individual fragment must not be specified if its fragment type is also specified as a whole.
If the data block form is used, only the fit coefficients of the referenced fragments are affected. For the not-referenced fragments the fit densities are used as they are defined on the corresponding fragment files.

The SCF convergence of a spin-unrestricted calculation usually improves when you start with potentials that correspond to the correct ratio of spin-α and spin-β electrons. By default ASPIN=BSPIN=0.5, as implied by the spin-restricted start density of the fragments or restricted molecule.

The total amount of fit density used on the first iteration is defined by the sum-of-fragment densities (or the density on the restart file). This may be different from the total nr. of electrons in the actual calculation. On the second SCF cycle the fit density will internally be normalized so as to represent the correct number of electrons.

The block-form of the key makes the start up of broken symmetry calculations easy. For example one may want to start a calculation in broken symmetry with spin-α density on one fragment and spin-β density on another, e.g. in a spin-unrestricted calculation of H_2 at large separation. It is particularly useful for larger systems, e.g. for magnetic coupling between spin-polarized magnetic centers, as in Fe-S complexes [111]: start with oppositely polarized Fe centers, but with, for instance, the remaining bridge and terminal ligands unpolarized. See also the N2+ sample run in the examples.

Unrestricted fragments

The fragments from which the molecule is built must be spin-restricted, that is: the fragment files must be result files of spin-restricted calculations. For purposes of analysis, however, it may be desirable in some applications to build your molecule from *un*restricted fragments. This can be simulated as follows.

You tell ADF that you want to *treat* the fragments as if they were unrestricted; this causes the program to duplicate the one-electron orbitals of the fragment: one set for spin-α and one set for spin-β. You can then specify occupation numbers for these spin-unrestricted fragments, and occupy spin-α orbitals differently from spin-β orbitals.

Of course, the unrestricted fragments that you use in this way, are not self-consistent: different numbers of spin-α and spin-β electrons usually result in different spatial orbitals and different energy eigenvalues for spin-α and spin-β when you go to self-consistency, while here you have spatially identical fragment orbitals. Nevertheless it is often a fair approximation which gives you a considerable extension of analysis possibilities.

In ADF2012 for hybrids, metaGGA's, and metahybrids the calculation of the Pauli repulsion term is implemented if one is simulating an unrestricted fragment with the key FRAGOCCUPATIONS.

```
FRAGOCCUPATIONS
 fragtype
   irrep spin-a // spin-b
   irrep spin-a // spin-b
   ...
 subend
 fragtype
   irrep spin-a // spin-b
   ...
 subend
 ...
end
```

`fragtype`

One of the fragment types and functions as a (block type) subkey. The data block for the subkey ends with the standard end code for block type subkeys (subend).

`irrep`

One of the irreducible representations (irreps) for the point group symmetry that was used in the computation of that fragment.

`spin-a // spin-b`

Two sequences of occupation numbers, which will be applied to the spin-α and spin-β versions of the Fragment Orbitals. The sequences must be separated by a double slash (//). See for comparison the specification of occupation numbers for the overall system (key CHARGE).

The sum of spin-α and spin-β occupations must, for each fragment orbital in each irrep separately, be equal to the total (restricted) occupation of that orbital as it is stored on the fragment file. In other words: you can only change the distribution over spin-α and spin-β electrons within one orbital.

(Without this restriction the spatial distribution of the total (sum over spins) fragment charge density would be changed, leading to an incorrect bonding energy analysis after the calculation).

The data block of FRAGOCCUPATIONS is not parsed for expressions and constants or functions defined under define. Any such items will not be recognized and not be replaced by their values.

Be aware that in more-dimensional irreps (e, t) the number of electrons in a fully occupied orbital is input as the dimension of the irrep times the one-electron orbital occupation. Compare the key OCCUPATIONS.

For irreps that are not mentioned in this input block, and hence for all irreps of fragment(type)s that are not mentioned at all, the spin-α and spin-β occupations will be set equal, which is of course what they in fact are on the (restricted) fragment file.

For an example of applying this option see [112].

Remove Fragment Orbitals

By default all fragment orbitals (the MOs of the fragment computation), which are stored on the fragment file, are used as basis functions for the overall molecule, see Chapter 1.2. You can remove one or more of these fragment orbitals from the basis set of the molecule. This may be useful for special analyzes, for instance to study the effect of deleting all virtual MOs of a particular fragment (CSOV analysis, Constrained Space Orbital Variation). It may also enhance the efficiency since you effectively reduce the size of the basis set, but you should be aware of the potential effects on the results.

```
REMOVEFRAGORBITALS
 fragtype
   subspecies nremove
   subspecies nremove
   ...
 subend
 fragtype
   subspecies nremove
   ...
 subend
 ....
 (etc.)
 ....
end
```

`fragtype`

One of the fragment types in the system. Any subset of the available fragment types can be used here as subkey. The subkeys are block type keys; their data blocks end subend.

`subspecies`

One of the subspecies of the irreducible representations of the point group symmetry that was used in the calculation of the fragment itself. This requires of course that one knows the symmetry that has been used for the fragment calculation.

`nremove`

The number of fragment orbitals of the pertaining representation that will not be used as basis functions for the overall system. The *highest* (in energy eigenvalue) nremove orbitals are discarded. You must not remove *occupied* fragment orbitals.

By default (omission of the key) all fragment orbitals are used in the basis set for the system.

Important Note

It is imperative that any removal of fragment orbitals will not break the symmetry of the molecule. This consideration is relevant when for instance two different subspecies of a fragment irrep contribute to different partner subspecies in one of the irreps of the molecule. In such a case, when one removes an

orbital in such a fragment subspecies, its partner orbital should also be removed. If this is violated an error may occur or the results will simply be wrong. Quite likely, the program will detect the error, but this may occur only in the final (analysis) stage of the calculation so that a lot of CPU time may have been wasted.

Example: consider a single-atom fragment, computed in atom symmetry, used as fragment in a c(lin) molecule and assume that the p:x and p:y fragment orbitals contribute to respectively the pi:x and pi:y subspecies of the molecule. Then, when you remove one or more p:x fragment orbitals, you should also remove the same number of p:y fragment orbitals. Practical cases may be more complicated and whenever you use this key, make sure that you've fully analyzed and understood how the fragment irreps combine into the molecular symmetry representations. Hint: run the molecule, without removing any fragment orbitals, and stop at an early stage after the program has computed and printed the build-up of the molecular SFOs from the fragment orbitals. To control early aborts via input, use the key STOPAFTER.

Density Functional

Exchange Correlation Potentials

The Density Functional, also called the exchange-and-correlation (XC) functional, consists of an LDA, a GGA part, a Hartree-Fock exchange part (hybrids), and a meta-GGA part (meta-GGA or meta-hybrid). LDA stands for the Local Density Approximation, which implies that the XC functional in each point in space depends only on the (spin) density in that same point. GGA stands for Generalized Gradient Approximation and is an addition to the LDA part, by including terms that depend on derivatives of the density. A hybrid GGA (for example B3LYP) stands for some combination of a standard GGA with a part of Hartree Fock exchange. A meta-GGA (for example TPSS) has a GGA part, but also depends on the kinetic energy density. A meta-hybrid (for example TPSSh) has GGA part, a part of Hartree-Fock exchange and a part that depends on the kinetic energy density. For these terms ADF supports a large number of the formulas advocated in the literature. For post-SCF energies only, ADF supports also various other meta-GGA functionals and more hybrid functionals.

The Perdew-Zunger self-interaction correction (SIC) was implemented [47-49] self-consistently using the Krieger-Li-Iafrate approximation to the optimized effective potential, and the Vosko-Wilk-Nusair (VWN) functional or gradient corrected density functionals. This approach is found to improve several properties, which are sometimes difficult to describe with standard DFT techniques, like for example some 'problematic' NMR chemical shifts, or some 'difficult' reaction barriers. Note that some of these problems can also be circumvented with hybrids.

Several asymptotically correct XC potentials have been implemented in ADF, such as the (now somewhat outdated) LB94 potential [15], the gradient-regulated asymptotic correction (GRAC) [16], and the statistical average of orbital potentials (SAOP) [244,17]. These can currently be used only for response property calculations, not for geometry optimizations. For spectroscopic properties, they usually give results superior to those obtained with LDA or GGA potentials, (see Ref.[18] for applications to (hyper)polarizabilities Cauchy coefficients, etc. of small molecules). This is particularly true if the molecule is small and the (high-lying) virtual orbitals are important for the property under study.

It was also shown that, simply using the orbital energies of the occupied Kohn-Sham orbitals of a SAOP calculation, quite good agreement with experiment vertical ionization potentials is obtained. This is true not only for the HOMO orbital energy, which should be identical to (minus) the experimental ionization potential with the exact XC potential, but also for lower-lying occupied orbital energies. The agreement becomes worse for deep-lying core orbital energies. A theoretical explanation and practical results are given in Ref. [19].

If a functional contains a part of Hartree-Fock exchange then the LDA, GGA, metaGGA, or MODEL key should not be used in combination with this key, and one should only specify one of HartreeFock, HYBRID

or MetaHYBRID. Dispersion can be added. Note that it is not recommended to use (part of the) Hartree-Fock exchange in combination with frozen cores, since at the moment the frozen core orbitals are not included in the Hartree Fock exchange operator. The implementation in ADF of the calculation of exact exchange (Hartree Fock exchange), which is also needed for the hybrid functionals, is based on work by Watson et al., Ref. [138]. In ADF one can do unrestricted Hartree-Fock (or hybrid or meta-hybrid) calculations, as long as one has integer occupation numbers (ROHF is not implemented in ADF, only UHF). Note that the DEPENDENCY key is switched on for Hartree-Fock exchange in order to circumvent numerical problems, see also the ADDDIFFUSEFIT key. You also need to the same XC-potential in the create run of the atoms, which is done automatically if you use the BASIS key.

The key that controls the Density Functional is XC, with sub keys *LDA* and *GGA* (or equivalently: *gradients*) to define the LDA and GGA parts of the functional, and *MODEL* in case one of the special 'model' XC potentials is required in stead of LDA or GGA. All subkeys are optional (need not be used). Some may occur twice in the data block.

```
XC
 {LDA LDA {Stoll}}
 {GGA GGA}
 {MetaGGA metagga}
 {Model MODELPOT [IP]}
 {HartreeFock}
 {OEP fitmethod {approximation}}
 {HYBRID hybrid {HF=HFpart}}
 {MetaHYBRID metahybrid}
 {XCFUN}
 {RANGESEP {GAMMA=X} {ALPHA=a} {BETA=b}}
 {DISPERSION [s6scaling] [RSCALE=r0scaling] [Grimme3] [BJDAMP] [PAR1=par1]  [PAR2=par2]
[PAR3=par3]  [PAR4=par4] }
 {DISPERSION dDsC}
 {DISPERSION UFF}
end
```

`LDA`

Defines the LDA part of the XC functional and can be any of the following:
`Xonly`: The pure-exchange electron gas formula. Technically this is identical to the Xalpha form (see next) with a value 2/3 for the X-alpha parameter.
`Xalpha`: the scaled (parameterized) exchange-only formula. When this option is used you may (optionally) specify the X-alpha *parameter* by typing a numerical value after the string Xalpha (separated by a blank). If omitted this parameter takes the default value 0.7
`VWN`: the parameterization of electron gas data given by Vosko, Wilk and Nusair (ref [20], formula version V). Among the available LDA options this is the more advanced one, including correlation effects to a fair extent.

`Stoll`

For the VWN or GL variety of the LDA form you may include Stoll's correction [21] by typing Stoll on the same line, after the main LDA specification. You must not use Stoll's correction in combination with the Xonly or the Xalpha form for the Local Density functional.

`PW92`: the parameterization of electron gas data given by Perdew and Wang (ref [288]).

`GGA`

Specifies the GGA part of the XC Functional, in earlier times often called the 'non-local' correction to the LDA part of the density functional. It uses derivatives (gradients) of the charge density. Separate choices can be made for the GGA exchange correction and the GGA correlation correction respectively. Both specifications must be typed (if at all) on the same line, after the GGA subkey.

For the exchange part the options are:
Becke: the gradient correction proposed in 1988 by Becke [22].
PW86x: the correction advocated in 1986 by Perdew-Wang [23].
PW91x: the exchange correction proposed in 1991 by Perdew-Wang [24]
mPWx: the modified PW91 exchange correction proposed in 1998 by Adamo-Barone [25]
PBEx: the exchange correction proposed in 1996 by Perdew-Burke-Ernzerhof [26]
RPBEx: the revised PBE exchange correction proposed in 1999 by Hammer-Hansen-Norskov [27]
revPBEx: the revised PBE exchange correction proposed in 1998 by Zhang-Wang [28]
mPBEx: the modified PBE exchange correction proposed in 2002 by Adamo-Barone [174]
PBEsolx: the PBEsol exchange correction proposed in 2008 by Perdew-Ruzsinszky-Csonka-Vydrov-Scuseria [285]
HTBSx: the HTBS exchange correction [437]
OPTX: the OPTX exchange correction proposed in 2001 by Handy-Cohen [29]
BEEx: the BEEx exchange correction proposed in 2005 by Mortensen-Kaasbjerg-Frederiksen-Nørskov-Sethna-Jacobsen [284]

For the correlation part the options are:
Perdew: the correlation term presented in 1986 by Perdew [30].
PBEc: the correlation term presented in 1996 by Perdew-Burke-Ernzerhof [26] .
PBEsolc: the PBEsol correlation correction proposed in 2008 by Perdew-Ruzsinszky-Csonka-Vydrov-Scuseria [285]
PW91c: the correlation correction of Perdew-Wang (1991), see [24].
LYP: the Lee-Yang-Parr 1988 correlation correction [31-33].

Some GGA options define the exchange and correlation parts in one stroke. These are:
BP86: this is equivalent to Becke + Perdew together.
PW91: this is equivalent to pw91x + pw91c together.
mPW: this is equivalent to mPWx + pw91c together.
PBE: this is equivalent to PBEx + PBEc together
RPBE: this is equivalent to RPBEx + PBEc together
revPBE: this is equivalent to revPBEx + PBEc together
mPBE: this is equivalent to mPBEx + PBEc together
PBEsol: this is equivalent to PBEsolx + PBEsolc together
HTBS: this is equivalent to HTBSx + PBEc together
BLYP: this is equivalent to Becke (exchange) + LYP (correlation).
OLYP: this is equivalent to OPTX (exchange) + LYP (correlation).
OPBE: this is equivalent to OPTX (exchange) + PBEc (correlation) [175].
XLYP: this is equivalent to XLYPx [172] (exchange, not available separately from LYP) + LYP (correlation).
BEE: this is equivalent to BEEx (exchange) + PBEc (correlation).
SSB-D: dispersion corrected functional by Swart-Solà-Bickelhaupt [286,287]. Single point only. Use METAGGA SSB-D in other cases.
S12g: dispersion corrected (Grimme-D3) functional by Swart, successor of SSB-D [367].
LB94: this refers to the XC functional of Van Leeuwen and Baerends [15].
KT1: this refers to the KT1 functional of Keal and Tozer [171].
KT2: this refers to the KT2 functional of Keal and Tozer [171].

The string GGA must contain not more than one of the exchange options and not more than one of the correlation options. If options are applied for both they must be separated by a blank or a comma.

MetaGGA

Specifies that a meta-GGA should be used during the SCF. All electron basis sets should be used. The meta-GGA can be one of the following:
M06-L: functional by Yan-Truhlar [223,224]

TPSS: functional by Tao-Perdew-Staroverov-Scuseria [246,247]
revTPSS: revised TPSS functional [438]
SSB-D: dispersion corrected GGA functional by Swart-Solà-Bickelhaupt [286,287]. Use GGA SSB-D for NMR calculations.

MODEL

Specifies that one of the less common XC potentials should be used during the SCF. These potentials specify both the exchange and the correlation part. No LDA, GGA, MetaGGA, HartreeFock, HYBRID or MetaHYBRID key should be used in combination with these keys. It is also not advised to use any energy analysis in combination with these potentials. For energy analysis we recommend to use one of the GGA potentials. It is currently not possible to do a Create run with these potentials. It is possible to do a one atom regular ADF calculation with these potentials though, using a regular TAPE21 file from an LDA or GGA potential as input.

LB94: this refers to the XC functional of Van Leeuwen and Baerends [15]. There are no separate entries for the Exchange and Correlation parts respectively of LB94. Usually the GRAC or SAOP potentials give results superior to LB94.

GRAC: the gradient-regulated asymptotic correction, which in the outer region closely resembles the LB94 potential [16]. It requires a further argument: the ionization potential [IP] of the molecule, in hartree units. This should be estimated or obtained externally, or calculated in advance from two GGA total energy calculations.

SAOP: the statistical average of orbital potentials [244,17]. It can be used for all electron calculations only. It will be expensive for large molecules, but requires no further parameter input.

IP: should be supplied only if GRAC is specified.

HartreeFock

Specifies that the Hartree-Fock exchange should be used during the SCF.

OEP

Defines the optimized effective potential expanded into a set of the fit functions. The subkeyword fitmethod can be any of the following: BARTLETT [248], SCUSERIA [249]. In the case of SCUSERIA one of the following approximations needs to be specified: CEDA, KLI or SLATER. An application of OEP in ADF can be found in Ref.[250].

HYBRID

Specifies that a hybrid functional should be used during the SCF. The hybrid can be one of the following:
B3LYP: ADF uses VWN5 in B3LYP. functional (20% HF exchange) by Stephens-Devlin-Chablowski-Frisch [176].
B3LYP*: modified B3LYP functional (15% HF exchange) by Reiher-Salomon-Hess [177].
B1LYP: functional (25% HF exchange) by Adamo-Barone [178].
KMLYP: functional (55.7% HF exchange) by Kang-Musgrave [179].
O3LYP: functional (12% HF exchange) by Cohen-Handy [180].
X3LYP: functional (21.8% HF exchange) by Xu-Goddard [172].
BHandH: 50% HF exchange, 50% LDA exchange, and 100% LYP correlation.
BHandHLYP: 50% HF exchange, 50% LDA exchange, 50% Becke88 exchange, and 100% LYP correlation.
B1PW91: functional by (25% HF exchange) Adamo-Barone [178].
mPW1PW: functional (42.8% HF exchange) by Adamo-Barone [25].
mPW1K: functional (25% HF exchange) by Lynch-Fast-Harris-Truhlar [181].

`PBE0`: functional (25% HF exchange) by Ernzerhof-Scuseria [211] and by Adamo-Barone [212], hybrid form of PBE.
`OPBE0`: functional (25% HF exchange) by Swart-Ehlers-Lammertsma [175], hybrid form of OPBE.
`S12H`: dispersion corrected (Grimme-D3) functional (25% HF exchange) by Swart [367].

`HFpart`

Specifies the amount of HF exchange that should be used in the functional, instead of the default HF exchange percentage for the given hybrid. Example HF=0.25 means 25% HF exchange.

`MetaHYBRID`

Specifies that a meta-hybrid functional should be used during the SCF. The meta-hybrid can be one of the following:
`M06`: functional (27% HF exchange) by Yan-Truhlar [223,224]
`M06-2X`: functional (54% HF exchange) by Yan-Truhlar [223,224]
`M06-HF`: functional (100% HF exchange) by Yan-Truhlar [223,224]
`TPSSH`: functional (10% HF exchange) by Tao-Perdew-Staroverov-Scuseria [246,247]

`XCFUN`

XCFun is a library of approximate exchange-correlation functionals, see Ref. [354], for which functional derivatives can be calculated automatically. For example, with XCFUN the full (non-ALDA) kernel can be evaluated and this has been implemented in the calculation of TDDFT excitations. The Full kernel can not be used in combination with symmetry, singlet-triplet excitation calculations, or excited state geometry optimizations. The following functionals can be evaluated with XCFUN at the present time:

- LDA: VWN5, X-ALPHA PW92
- GGA exchange: Becke88, PBEX, OPTX, PW91X, mPW, revPBEX
- GGA correlation: LYP, Perdew86, PBEC
- MetaGGA: TPSS, M06L, B95
- MetaHybrids: M06, M05, M062X, M06HF, PW6B95, PWB6K
- Hybrids: PBE0, B3LYP, BHandH, B1LYP, B3LYP*, PBEFALFX, CAMY-B3LYP

`RANGESEP {GAMMA=X} {ALPHA=a} {BETA=b}`

If RANGESEP is included, by default a long-range corrected (LC) functional is created with range separation parameter GAMMA of 0.75. As switching function in ADF the Yukawa potential is utilized, see Ref. [355]. Range separated functionals require XCFUN and are limited to GGA, meta-GGA, and CAMY-B3LYP. The CAMY-B3LYP functional is not the same as the CAM-B3LYP functional, since a different switching function is used. No other hybrids or meta-hybrids are supported. The special CAMYB3LYP functional is defined by three parameters, ALPHA, BETA and the attenuation parameter GAMMA. For CAMYB3LYP by default ALPHA is 0.19, BETA is 0.46, and GAMMA is 0.34. Singlet-triplet excitation calculations or excited geometry optimizations are not possible for functionals that require XCFUN.

`DISPERSION Grimme3 BJDAMP`

If `DISPERSION Grimme3 BJDAMP` is present a dispersion correction (DFT-D3-BJ) by Grimme [436] will be added to the total bonding energy, gradient and second derivatives, where applicable. This is the latest dispersion correction in the DFTB-D family. Parametrizations are implemented for the functionals: B3LYP, TPSS, BP86, BLYP, PBE, PBEsol, and RPBE. It has four parameter. One can overide this using *PAR1=.. PAR2=..*, etc. In the table the relation is shown between the parameters and the real parameters in the Dispersion correction.

variable	variable on Bonn website

PAR1	s6
PAR2	a1
PAR3	s8
PAR4	a2

`DISPERSION Grimme3`

If `DISPERSION Grimme3` is present a dispersion correction (DFT-D3) by Grimme [292] will be added to the total bonding energy, gradient and second derivatives, where applicable. Parametrizations are available for: B3LYP, TPSS, BP86, BLYP, revPBE, PBE, PBEsol, and RPBE, and will be automatically set if one of these functionals is used. There are also parameters directly recognized for S12g and S12h. For all other functionals, PBE-D3 parameters are used as default. You can explicitly specify the three parameters

variable	**variable on Bonn website**
PAR1	s6
PAR2	sr,6
PAR3	s8

`DISPERSION {s6scaling]} {RSCALE=r0scaling}`

If the `DISPERSION` keyword is present (without the argument Grimme3) a dispersion correction (DFT-D) by Grimme [226] will be added to the total bonding energy, gradient and second derivatives, where applicable. The global scaling factor with which the correction is added depends on the exchange-correlation functional used at SCF but it can be modified using the *s6scaling* parameter. The following scaling factors are used (with the XC functional in parantheses): 1.20 (BLYP), 1.05 (BP), 0.75 (PBE), 1.05 (B3LYP). In all other cases a factor 1.0 is used unless modified via the s6scaling parameter. The SSB-D functional includes the dispersion correction (factor 0.847455) by default.
Unlike the MMDISPERSION keyword, the van der Waals radii used in this implementation are hardcoded in ADF. However, it is possible to modify the global scaling parameter for them using the *RSCALE=r0scaling* argument. The default value is 1.1 as proposed by Grimme [226]. Please also see additional documentation for more information about this topic.

`DISPERSION dDsC`

The `DISPERSION dDsC` key invokes the density dependent dispersion correction [316], which has been parametrized for the functionals BLYP, PBE, BP, revPBE, B3LYP, PBE0 and BHANDHLYP.

`DISPERSION UFF`

The `DISPERSION UFF` key invokes the universal correction of density functional theory to include London dispersion (DFT-ulg) [366], which has been parametrized for all elements up to Lr (Z=103), and for the functional PBE, PW91, and B3LYP. For other functionals the PBE pareameters will be used.

`DISPERSION MBD`

The `DISPERSION MBD` key invokes the MBD@rsSCS method [377], which is designed to accurately describe long-range correlation (and thus dispersion) in finite-gap systems, including at the same time a description of the short-range interactions from the underlying DFT computation of the electronic structure.

Defaults, special cases, simple input

If the XC key is not used, the program will apply only the Local Density Approximation (no GGA terms). The chosen LDA form is then VWN.

If only a GGA part is specified, omitting the *LDA* sub key, the LDA part defaults to VWN, except when the LYP correlation correction is used: in that case the LDA default is Xonly: pure exchange.

The reason for this is that the LYP formulas assume the pure-exchange LDA form, while for instance the Perdew-86 correlation correction is a correction to a *correlated* LDA form. The precise form of this correlated LDA form assumed in the Perdew-86 correlation correction is not available as an option in ADF but the VWN formulas are fairly close to it.

Be aware that typing only the sub key *LDA*, without an argument, will activate the VWN form (also if LYP is specified in the GGA part).

PBE functionals

Starting from ADF2009.01 the default PBE functional uses LDA PW92, instead of LDA VWN, and uses for the GGA part routines provided by Burke.

Before this bug-fix the 'correct' PBE functional could be obtained with explicit specifying the LDA PW92 functional and the correct PBEc correlation functional:

```
XC
 LDA PW92
 GGA PBE USEBURKEROUTINES
End
```

Now this is default if one uses

```
XC
 GGA PBE
End
```

The old ('incorrect') defaults can be calculated with

```
XC
 LDA VWN
 GGA PBE USESPROUTINES
End
```

SSB-D functional

Currently (ADF2009.01), there are some numerical issues with the GGA implementation in ADF of SSB-D (Ref. [286,287]) for some systems. Because of this, the GGA SSB-D option is only available for single-points (and NMR). Geometry optimizations (etc.) are still possible by using instead:

```
XC
  METAGGA SSB-D
END
```

This METAGGA implementation is only possible with all-electron basis sets. Use GGA SSB-D for NMR calculations.

The SSB-D functional by definition already includes a dispersion correction by Grimme (factor 0.847455).

Meta-GGA potentials

Starting from ADF2009.01 the meta-GGA's M06-L and TPSS can be used during the SCF. Also starting from ADF2009.01 the meta-GGA's can be used in combination with geometry optimization, TS, IRC, LT, numerical frequencies, and excitation energies (ALDA kernel used). All electron basis sets should be used.

The M06-L functional needs high integration accuracy (at least BeckeGrid quality good) for reasonable gradients. For TPSS moderate integration accuracy for reasonable gradients is sufficient. For heavier elements (Z>36) and if one uses the M06-L functional it is also necessary to include the following keyword

```
FragMetaGGAToten
```

Using this key FRAGMETAGGATOTEN the difference in the meta-hybrid or meta-GGA exchange-correlation energies between the molecule and its fragments will be calculated using the molecular integration grid, which is more accurate than the default, but is much more time consuming. Default is to calculate the meta-GGA exchange-correlation energies for the fragments in the numerical integration grid of the fragments.

Model potentials

The LB94, GRAC, and SAOP functionals have only a SCF (=Potential) implementation, but no Energy counterpart. Therefore, they must not be used together with the Energy specification for Apply. If LB94 or GRAC is used for the Potential (SCF), the gga energy expression defaults to Becke (exchange part) + Perdew (correlation). For SAOP, the energy functional is PW91. This can be overruled by selecting another choice in the 'gga Energy ...' specification. However, it is recommendable to use a GGA for the XC potential if the main interest is in energies.

The LB94, GRAC, and SAOP forms are density functionals specifically designed to get the correct asymptotic behavior. This yields much better energies for the highest occupied molecular orbital (HOMO) and better excitation energies in a calculation of response properties (Time Dependent DFT). Energies for lower lying orbitals (sub-valence) should improve as well (in case of GRAC and SAOP, but not LB94). The energy expression underlying the LB94 functional is very inaccurate. This does not affect the response properties but it does imply that the energy and its derivatives (gradients) should not be used because lb94-optimized geometries will be wrong, see for instance [34]. The application of the LB94 functional in a runtype that involves the computation of energy gradients is disabled in ADF. You can override this internal check with the key ALLOW.

In case of a GRAC calculation, the user should be aware that the potential in the outer region is shifted up with respect to the usual level. In other words, the XC potential does not tend to zero in the outer region in this case. The size of the shift is the difference between the HOMO orbital energy and the IP given as input. In order to compare to regular GGA orbital energies, it is advisable to subtract this amount from all orbital energies. Of course, orbital energy differences, which enter excitation energies, are not affected by this shift in the potential.

The LB94, SAOP, and GRAC potentials cannot be used in a Create run (due to an implementation limitation in the code). If you need the energy difference of a molecule with respect to LB94-atoms, you have to run

the single-atom calculations with LB94 separately, using the same non-LB94 Create atoms as fragments as you did for the whole molecule. This will give you the required energy corrections.

Hartree-Fock and (meta-)hybrid potentials

Starting from ADF2009.01 the meta-hybrids M06, M06-2X, M06-HF, and TPSSH can be used during the SCF. Also starting from ADF2009.01 Hartree-Fock and the (meta-)hybrid potentials can be used in combination with geometry optimization, TS, IRC, LT, and numerical frequencies; hybrids can be used in calculating NMR chemical shift; PBE0 can be used in calculating NMR spin-spin coupling; Hartree-Fock and (meta-)hybrid can be used in calculating excitation energies, in which the kernel consists of the Hartree-Fock percentage times the Hartree-Fock kernel plus one minus the Hartree-Fock percentage times the ALDA kernel (thus no (meta-)GGA kernel). Hartree-Fock and the (meta-)hybrid potentials still can not or should not be used in combination with analytical frequencies, the (AO)RESPONSE key, EPR/ESR g-tensor, and frozen cores. Starting from ADF2010 it is possible to use Hartree-Fock and hybrids to calculate CD spectra.

In ADF one can do unrestricted Hartree-Fock (or hybrid or meta-hybrid) calculations, as long as one has integer occupation numbers (ROHF is not implemented in ADF, only UHF).

Starting from ADF2009.01 it is possible to change the amount of HF exchange in the input for hybrids (not for meta-hybrids and Hartree-Fock). For many hybrid functionals the sum of the amount of Hartree-Fock exchange and the amount of LDA exchange (or GGA exchange) is one. If that is the case, then if one changes the amount of Hartree-Fock exchange in the input the amount of LDA exchange (or GGA exchange) will also be changed, such that the sum remains one. Example:

```
XC
  Hybrid B3LYP HF=0.25
END
```

In this case the amount of Hartree-Fock for the B3LYP functional will be changed to 25% (instead of 20%), and the amount of LDA exchange to 75% (instead of 80%). The LDA correlation and GGA exchange and correlation part will be left unaltered.

Memory usage

Calculation of the Hartree-Fock exchange may require a lot of memory. Starting from ADF2012, shared memory is used to buffer the necessary data used by all processes on a multi-processor node. By default, ADF will use 30% of the total physical memory of the computer for this buffer, which may be more than is desirable or possible. This amount, in megabytes, can be set using a HFMAXMEMORY input keyword.

```
HFMaxMemory mbytes
```

The amount of memory used is related how many atoms can be done in a single pass. Thus, another way to limit the amount of memory used by ADF is to limit the number of atoms done per pass. The latter can be done using the HFATOMSPERPASS keyword. The safest, but also the slowest, is to set HFATOMSPERPASS to 1.

```
HFAtomsPerPass AtomsPerPass
```

If both HFMAXMEMORY and HFATOMSPERPASS are present, the value specified in HFATOMSPERPASS takes precedence. To debug memory usage in the Hartree-Fock routine, one can use a PRINT HFEXCHANGE keyword.

Numerical issues

Numerical problems have been found with the present implementation of Hartree-Fock or (meta-)hybrids during the SCF, especially if the molecule has symmetry NOSYM and a basis set TZP or larger is used. Workaround is to use always the DEPENDENCY key with rather strict criteria for the basis set dependence,

namely bas=4e-3. In ADF2010 these numerical problems have been reduced. Starting from the ADF2006.01 the DEPENDENCY key is automatically switched on in the case of a Hartree-Fock or a (meta-)hybrid potential. The result of the DEPENDENCY key is that linear dependence of the basis set is reduced by removing linear combinations that correspond with eigenvalues in the virtual SFOs overlap matrix, which are smaller than, in this case, 4e-3. Note that this is a rather large value, such that it will have an effect on the bonding energy. For DZ and DZP basis sets this value will normally not result in reduction of the virtual space. However, for TZP, TZ2P, QZ4P and larger this will often result in reduction of the basis set, which will have an effect on the accuracy of the bonding energy. In these cases one could try a smaller value than 4e-3, but be aware that numerical problems may occur. If the molecule has symmetry the numerical problems are reduced.

The origin of this problem is that for an accurate description of Hartree-Fock exchange one needs more (diffuse) fit functions in the fit procedure which is used in ADF, which uses only fit fuctions on the two centers of the two STOs. One can get more diffuse fit functions if one adds in the Create run of an atom the key:

```
AddDiffuseFit
```

If the BASIS key is used one can also add this key in the molecular calculation (the scripts in ADF will then automatically add this in the Create runs of the atoms). If one adds this key preliminary results indicate that one can lower the value for the dependency key to bas=1e-4. Such a low value for the dependency key normally means that the basis set is not reduced for basis sets of TZP or TZ2P quality.

A different way to add fit functions that is useful for standard basis setsl DZ, DZP, TZP, and TZ2P is to add the subkey 'Fittype QZ4P' in the BASIS key, thus:

```
BASIS
  Type ...
  Core ...
  FitType QZ4P
End
```

In ADF2013 use is made of distance cut-offs for the calculation of the HF exchange integrals. This can reduce computation time and memory usage, especially for large molecules, however, this can also reduce the precision, which can lead to numerical problems. It is possible to set the distance cut-off threshold (starting from ADF2013.01b) for the calculation of the HF exchange integrals, A value of 99 for hffit virtually excludes the possibility that something will be neglected:

```
LINEARSCALING
  HF_FIT hffit
END
```

An accuracy issue is relevant for some of the meta-GGA functionals, in particular the M06 functionals. These need high integration accuracy (at least BeckeGrid quality good) for reasonable gradients. For TPSSH moderate integration accuracy for reasonable gradients is sufficient. For heavier elements (Z>36) and if one uses one of the M06 functionals it is also necessary to include the following keyword

```
FragMetaGGAToten
```

Using this key FRAGMETAGGATOTEN the difference in the metahybrid or metagga exchange-correlation energies between the molecule and its fragments will be calculated using the molecular integration grid, which is more accurate than the default, but is much more time consuming. Default is to calculate the meta-hybrid or meta-GGA exchange-correlation energies for the fragments in the numerical integration grid of the fragments.

For benchmark calculations one would like to use a large basis set, like the QZ4P basis set. In such cases it is recommended to use a good numerical quality. Thus for accurate hybrid calculations of **small** molecules one could use:

```
basis
 type QZ4P
end
AddDiffuseFit
Dependency bas=1e-4
NumericalQuality good
```

However, for larger molecules and in case the molecule contains heavy elements (Z>36) one still should use rather strict criteria for the basis set dependence, such as bas=4e-3.

SCF problems

In case of SCF problems that might be related to numerical issues one can try one or more of the following, that were discussed before

```
LINEARSCALING
  HF_FIT 99
END
Basis
 FitType QZ4P
End
AddDiffuseFit
Dependency bas=5e-3
```

Note that HF_FIT can be used starting from ADF2013.01b.

Range-separated functionals

Range-separated functionals make use of a modified form of the Coulomb operator that is split into pieces for exact exchange and DFT. As switching function in ADF the Yukawa potential is utilized, see Ref. [355]. Global hybrids can be thought of as a special case of a range-separated functional where the split is independent of the interelectronic distance and is a simple X% exact and 1-X% DFT in all space.

In a general RS-functional the split depends on the interelectronic distance. How the split is achieved depends on the functional in question but it is achieved using functions that smoothly go from 1 to 0. In ADF an exponential function is used (the error function is common in Gaussian based codes). In a range-separated function the potential goes from a Coulomb interaction to a sum of Coulomb functions attenuated by an exponential function.

In practical terms, this means that a range-separated functional looks very much like a hybrid (or meta-hybrid) functional but with additional integrals over the attenuated interaction with fit functions on the exact exchange side and a modified functional on the DFT side.

DFT part of RS-functionals

Using Hirao's approach for creating RS-functionals, the RS form of a given exchange functional is created by multiplying the standard energy density by a factor that depends on the energy density. The factor is the same for all functionals and the only difference is introduced by the individual energy densities.

Exact exchange part of RS-functionals

The range-separation comes in at the level of the integrals over the operator with fit functions. They are very similar to the standard Coulomb integrals.

RS-functionals

An RS-functional is described by a series of regions describing each of the pieces of the Coulomb operator. The total function is built up by looping over the regions and adding up all the pieces. Currently, simple LC functionals can be defined where the exact exchange goes from 0 to 1 as the interelectronic distance increases and the DFT part does the reverse. In addition CAMY-B3LYP type functionals can be defined. More general functionals are not possible yet.

Functionality/Limitations

RS functionals require XCFUN and are limited to the GGA and meta-GGA functionals and one hybrid CAMY-B3LYP. The following functionals can be evaluated with range-separation at the present time:

- LDA: VWN5, X-ALPHA PW92
- GGA exchange: Becke88, PBEX, OPTX, PW91X, mPW, revPBEX
- GGA correlation: LYP, Perdew86, PBEC
- MetaGGA: TPSS, M06L, B95
- Hybrids: CAMY-B3LYP

The following functionality has been tested: XC potential, energy, ground state geometry, TDDFT. Singlet-triplet excitation calculations or excited geometry optimizations are not possible at thim time.

Numerical stability

The range-separated implementation requires that the range-separation parameter is not too close to the exponent of a fit function. In practice this means that values of the separation parameter between 1.0 and 50 can cause numerical problems. Typical useful values are in the range 0.2 to 0.9 so this should not be too serious a limitation.

Input

```
XC
 ...
 XCFUN
 RANGESEP {GAMMA=X} {ALPHA=a} {BETA=b}
END
```

Range separation is activated by putting RANGESEP in the XC block. Inclusion of XCFUN is required, see the XCFUN description. By default a long-range corrected (LC) functional is created with range separation parameter of 0.75. The parameter can be changed by modifying X in GAMMA=X in the RANGESEP card. Range separation typically will be used in combination with a GGA or METAGGA functional.

Range separation can not be included with a hybrid or meta-hybrid, with one exception, the special RS functional: CAMY-B3LYP. This is entered as HYBRID CAMY-B3LYP and must be used in combination with XCFUN (see XCFUN description) and RANGESEP. The CAMY-B3LYP functional is defined by three parameters, alpha, beta and the attenuation parameter gamma. The gamma parameter can be modified as for the LC functionals. For CAMY-B3LYP it defaults to 0.34. The alpha and beta parameters can be modified through ALPHA=a and BETA=b in the RANGESEP card. They default to 0.19 and 0.46 respectively.

```
XC
 HYBRID CAMY-B3LYP
 XCFUN
 RANGESEP GAMMA=0.34 ALPHA=0.19 BETA=0.46
END
```

Simple XC potential input

```
XC
    functionalspecification
```

```
    {dispersion [Grimme3]}
end
```

`functionalspecification`

The simplest use of the XC kewyord is to use one of the following lines for `functionalspecification`. More details and other options of the XC keyword can be found in the previous sections.

```
LDA VWN
LDA Xalpha
LDA PW92
GGA Becke Perdew
GGA BLYP
GGA PW91
GGA mPW
GGA PBE
GGA RPBE
GGA revPBE
GGA mPBE
GGA OLYP
GGA OPBE
GGA BP86
GGA BLYP
GGA PBE
GGA PBEsol
GGA HTBS
GGA KT1
GGA KT2
GGA SSB-D
GGA S12g
Model LB94
Model SAOP
HartreeFock
Hybrid B3LYP
Hybrid B3LYP*
Hybrid B1LYP
Hybrid KMLYP
Hybrid O3LYP
Hybrid X3LYP
Hybrid BHandH
Hybrid BHandHLYP
Hybrid B1PW91
Hybrid MPW1PW
Hybrid MPW1K
Hybrid OPBE0
Hybrid PBE0
Hybrid S12h
MetaGGA M06L
MetaGGA TPSS
MetaHybrid M06
MetaHybrid M06-2X
MetaHybrid M06-HF
MetaHybrid TPSSH
```

`dispersion {Grimme3 BJDAMP} {dDsC} {UFF} {MBD}`

If DISPERSION Grimme3 BJDAMP is present a dispersion correction DFT-D3-BJ is added. DISPERSION dDsC invokes the density dependent dispersion correction. The DISPERSION UFF key invokes the universal correction of density functional theory to include London dispersion (DFT-ulg). The DISPERSION MBD key invokes the MBD@rsSCS method.

Range separated functionals

```
XC
    functionalspecification
    XCFUN
    RANGESEP
end
```

`functionalspecification`

List of the most important functionals, for which one can use range separation:

```
LDA VWN
GGA BLYP
GGA BP86
GGA PBE
HYBRID CAMY-B3LYP
```

Post-SCF energy functionals

GGA energy functionals

In principle you may specify different functionals to be used for the *potential,* which determines the self-consistent charge density, and for the *energy* expression that is used to evaluate the (XC part of the) energy of the charge density. To be consistent, one should generally apply the same functional to evaluate the potential and energy respectively. Two reasons, however, may lead one to do otherwise:

- The evaluation of the GGA part in the *potential* is more time-consuming than LDA. The effect of the GGA term in the potential on the self-consistent charge density is often not very large. From the point of view of computational efficiency it may, therefore, be attractive to solve the SCF equations at the LDA level (i.e. not including GGA terms in the potential), and to apply the full expression, including GGA terms, to the energy evaluation *a posteriori*: post-SCF.
- A particular XC functional may have only an implementation for the potential, but not for the energy (or vice versa). This is a rather special case, intended primarily for fundamental research of Density Functional Theory, rather than for run-of-the-mill production runs.

One possibility is to calculate a whole list of post-SCF energy functionals using the METAGGA keyword, see next section. For some functionals the next possibility is enough. One has to specify different functionals for potential and energy evaluations respectively, using:

```
XC
 {LDA {Apply} LDA {Stoll}}
 {GGA {Apply} GGA}
end
```

`Apply`

States whether the functional defined on the pertaining line will be used self-consistently (in the SCF-potential), or only post-SCF, i.e. to evaluate the XC energy corresponding to the charge density. The value of apply must be SCF or Energy. A value postSCF will also be accepted and is equivalent to Energy. A value Potential will also be accepted and is equivalent to SCF. For each record separately

the default (if no Apply value is given in that record) is SCF. For each of the two terms (LDA, GGA) in the functional: if no record with Energy specification is found in the data block, the evaluation of the XC energy will use the same functional as is applied for the potential.

`LDA, GGA`

See the XC potential section for all possible values.

Meta-GGA and hybrid energy functionals

Starting from the ADF2004.01 version, several GGA [26-28,35-39], meta-GGA [40-46], hybrid GGA (for example B3LYP), and hybrid meta-GGA energy XC functionals have been implemented that have been shown to give good results for energies. This implementation enables only energies to be calculated with most of these functionals. In more recent versions of ADF some of the (meta-)GGA and (meta-)hybrid functionals can be used during the SCF, for optimizations and in property calculations, see the separate sections on meta-GGA potentials and (meta-)hybrid potentials.

The post SCF energy calculation is an easy and cheap way to get a reasonable guess for the bond energies for different XC functionals at the same time. Note that post-SCF energy calculations for a certain XC functional will not be so accurate if the functional form of the XC functional used in the SCF is very different from the XC functional used post SCF. The relative accuracy of post-SCF energies may not be so high if one looks at small energy differences. For accurate energy calculations it is recommended to use the same XC functional during the SCF as for the energy.

The implementation in ADF of the calculation of exact exchange (Hartree Fock exchange), which is needed for the hybrid functionals, is based on work by Watson et al., Ref. [138]. The difference with their method is the way in which ADF the orbital densities are fitted.

The calculation of a large, prespecified list of LDA, GGA, and meta-GGA energy functionals is invoked by specifying

```
METAGGA
```

as a separate keyword. The following (incomplete) list gives an idea of the (meta-)GGA density functionals that will then be calculated:

BP, PW91, mPW, BLYP, PBE, RPBE, revPBE, mPBE, OLYP, OPBE, KCIS, PKZB, VS98, FT97, BLAP3, HCTH, tau-HCTH, BmTau1, BOP, OLAP3, TPSS, KT1, KT2, B97, M06-L.

The hybrid GGA and hybrid meta-GGA energy functionals are calculated if in addition to the METAGGA key, the key

```
HARTREEFOCK
```

is included. The following (incomplete) list gives an idea of the extra hybrid (meta-)GGA density functionals that will then be calculated:

B3LYP, B3LYP*, B1LYP, KMLYP, O3LYP, X3LYP, BHandH, BHandHLYP, B1PW91, MPW1PW, MPW1K, PBE0, OPBE0, TPSSh, tau-HCTH-hybrid, B97, M05, M05-2X, M06, M06-2X.

In ADF2007.01 the Zhao-Truhlar M05 (Ref.[221,222]) and M06 (Ref.[223,224]) class xc energy functionals have been implemented.

The keys METAGGA and HARTREEFOCK can be used in combination with any XC potential. Note that at the moment hybrid functionals can not be used in combination with frozen cores. Also most METAGGA functionals will give wrong results if used in combination with frozen cores. Thus it is best to use an all electron basis set if one of the keywords METAGGA or HARTREEFOCK is used. One should include the

HARTREEFOCK keyword also in the create runs of the atoms. In ADF the hybrid energies only make sense if the calculation is performed with completely filled orbitals (ROHF is not implemented in ADF, only UHF).

For comparison with previous results instead of the key HARTREEFOCK one can also use the key HFEXCHANGE:

```
HFEXCHANGE
```

This key is now obsolete. The difference with the HARTREEFOCK key is the way in which the orbital densities are fitted. The key HFEXCHANGE can not be used in combination with frozen cores or spin-orbit coupling.

The Examples document describes an application to the OH molecule for the METAGGA option. More output, on the total XC energy of the system, can be obtained by specifying

```
PRINT METAGGA
```

This latter option is intended for debugging purposes mainly and is not recommended for general use.

The implementation calculates the total XC energy for a system and writes it to a file. This is always done in Create runs. If the basic fragments are atoms, the keyword

```
ENERGYFRAG
ATOM [filename]
ATOM [filename]
 ... ...
END
```

specifies that different atomic fragment files are to be used in the meta-GGA energy analysis than the regular atomic fragment files from the create runs. This keyword cannot be used for molecular fragment files. In order to compare meta-GGA energy differences between molecular fragments and the total molecule, results from the various calculations need to be combined by hand.

In such situations, it is advisable to use a somewhat higher integration accuracy than one would normally do, at least for the smaller fragments, as there is no error cancellation as in a regular ADF bond energy analysis.

A general comment is that some functionals show a more stable behavior than others (at least in our current implementation). In general, the functionals which are dependent on the Laplacian of the density may display a large variation with respect to basis set changes or different numerical integration accuracy. For this reason we currently recommend FT97 in favor of FT98. Similarly, the results with the BmTau1 functional should still be carefully checked. In our test calculations on the G2 set of molecules, the VS98 showed best performance, both for the average error and for the maximum error. The G2 set consists only of small molecules with elements up to Cl. The relative performance for transition metals and heavy elements is unknown and may well be very different from the ordering for the G2 set.

Post Hartree-Fock energy functionals

Next text is mostly taken from text by the authors of Ref. [381]. In the early days of DFT, non-self-consistent Kohn-Sham energy was often evaluated upon Hartree-Fock (HF) densities as a way to test new approximations. This method was called HF-DFT. It has been discovered that in some cases, HF-DFT actually gave more accurate answers when compared to self-consistent DFT calculations. In Ref. [381], it was found that DFT calculations can be categorized into two different types of calculations. The error of an approximate functional can be decomposed into two parts: error from the functional (functional error), and error from the density (density-driven error). For most calculations, functional error is dominant, and here self-consistent DFT is usually better than non-self consistent DFT on more accurate densities (called density corrected DFT (DC-DFT)). Unlike these 'normal' calculations, there is a class of calculations where the density-driven error is much larger, so DC-DFT give better a result than self-consistent DFT. These

calculations can be classified as 'abnormal'. HF-DFT is a simple implementation of DC-DFT and a small HOMO-LUMO gap is an indicator of an 'abnormal' calculation, thus, HF-DFT would perform better in such cases.

In ADF one can do HF-DFT with:

```
XC
  HartreeFock
END
MetaGGA
```

This will produce a large, prespecified list of LDA, GGA, meta-GGA, hybrid, and metahybrid energy functionals.

Self-Interaction Correction

In the ADF2004.01 version the Perdew-Zunger (PZ) self-interaction energy correction (SIC) with the Krieger-Li-Iafrate (KLI) approximation to the self-interaction corrected optimized effective potential (OEP) is implemented[47-49]. The block key SICOEP should be used.

Note: The existing auxiliary fitting sets, employed in ADF, were optimized for the calculation of the Coulomb potential of the total electron density. These standard fitting sets are quite flexible in the valence region, but do not include functions of high angular momentum in the inner regions. It was found (Ref [47]) that self-consistent SIC calculations with the standard fitting sets result in very poor SCF convergence, and large fit incompleteness corrections, particularly for the orbitals with substantial p and d contributions. Once the auxiliary fit sets are augmented with additional functions of high angular momentum, and reoptimized, the SCF convergence problems largely disappear. The reason is that the fit set must be able to approximate densities of each individual localized orbital, both in direction, and shell structure.

The key

```
SINGULARFIT FAST
```

activates fast(er) treatment of linearly dependent fits. It precomputes and stores singular value decomposition of fit overlap integrals, making fitting itself faster. Because SIC needs to fit a lot of densities on each iteration, it can improve performance by an order of magnitude. Use of 'SINGULARFIT FRUGAL' requests the old treatment of linearly-dependent fits, recomputing SVD parameters on each iteration. This is always slower than 'SINGULARFIT FAST', but uses less disk space.

The localization transformation is essential for obtaining realistic total and atomization energies. For example, without localization, SIC-Vosko-Wilk-Nusair (VWN) predicts molecular oxygen to be unstable with respect to the dissociation to two neutral oxygen atoms. At the same time, when using the localized orbitals, SIC-VWN compares favorably to both VWN itself, and to sophisticated gradient-corrected functionals. In ADF the Foster-Boys localization procedure is used.

Remarks: Fractional occupation are treated right. GGA functionals are supported, even for open-shell. SIC works for Linear Transit calculations. Frozen core SIC potentials are computed (GGA). The code is not well suitable for treating d and f electrons within the valence shell. Ds and Fs are fine within the frozen core. The SIC code does not allow parallel calculations. Geometry optimizations are not yet possible. SIC in combination with spin-orbit coupling is not possible yet.

Usage of the block key SICOEP:

```
SICOEP
 {IPRINT n}
 {NOLOCALIZE}
 {LOCALIZE {thrs}}
```

```
  {LDA name {Xa}}
  {GGA name}
  {SELFCONSISTENT n}
  {POSTSCF}
  {SKIPCYCLES every {start}}
  {STABLE frac}
  {NHOMO Nalpha {Nbeta}}
  {CORE cor}
  {DENSITY mode}
  {SHIPV filename}
  {READV filename}
  {READMOS filename {FREEZE}}
  {WRITEMOS filename}
  {NAILCANONICALS {eps}}
  {NPFITS n}
 end
```

`IPRINT n`

-2 = no printing except for fatal errors
-1 = condensed printing
0 = normal printing
1 = verbose printing
5 = debug printing
10 = exorcism printing

`NOLOCALIZE`

Calculate SIC energies for canonical MOs. Except for atoms and certain small molecules, this is guaranteed to produce non-optimal SIC energies, and/or lead to severe convergence problems.

`LOCALIZE {thrs}`

Calculate SIC energies for localized MOs. In this version, Boys-Foster procedure is used to obtain localized orbitals. By default, all occupied orbitals will participate in the Boys-Foster procedure. This can be changed by supplying thrs parameter, which will only allow mixing of canonical orbitals within thrs eV of each other. Calculate SIC energies for canonical MOs. Except for atoms and certain small molecules, this is guaranteed.

`LDA name {Xa}`

Requests a specific local density functional in Perdew-Zunger energy correction and KLI contribution to the XC potential (If name is XALPHA, the default alpha of 0.7 can be changed by specifying the additional argument). Normally, SIC code will use LDA functional requested (explicitly or implicitly) in the XC keyblock. Specifying 'NONE' with suppress LDA contribution in SIC energy and potential. Because the Perdew-Zunger energy correction, and the corresponding term in the XC potential are trying to remove a non-physical self-interaction contribution from the Kohn-Sham E(XC)/v(XC) contribution, this key should only be used for debugging.

`GGA name`

Requests a specific gradient-corrected functional in the SIC part, overriding the default choice (same as in 'XC' keyblock). Specifying 'NONE' will suppress GGA contribution to the SIC energy and potential. Note that very little input checking is done for this key. It is possible to specify a non-existent functional here, which will lead to unpredictable results. Use this key only for debugging.

`SELFCONSISTENT {n}`

Includes KLI contribution in the XC potential. This is the default. In case of convergence difficulties, v(KLI) will be recomputed until n-th SCF cycle (90 cycles by default). All subsequent SCF cycles will use v(KLI) from the n-th cycle. If a SIC calculation runs into convergence difficulties, it is important to make sure that the SIC energy does not change significantly between the last cycle where potential was computed, and the final SCF cycle.

`POSTSCF`

Calculates Perdew-Zunger energy correction, using orbitals from standard Kohn-Sham calculation (possibly after localizing them - see LOCALIZE/NOLOCALIZE). Except in a few special cases, such as atoms or the hydrogen molecule, post-SCF SIC corrections are not reliable (see S. Patchkovskii and T. Ziegler, JCP 116, 7806, Ref.[48]).

`SKIPCYCLES every {start}`

Requests that v(KLI) is recomputed after a given number of SCF iterations. By default, v(KLI) is recomputed on each SCF iteration (every=1). Because v(KLI) evaluation is expensive, using 2 or 3 here may lead to a reduction in computation time. If an optional second parameter is included, v(KLI) is omitted during the first few SCF cycles. This may reduce calculation time if starting guess is not very accurate. Because SIC energy expression is defined as a functional of orbitals, v(KLI) is always omitted in the first SCF cycle, regardless of the "start" setting.

`STABLE frac`

Specifies mixing coefficient for v(KLI) contribution to exchange-correlation potential. On each iteration where v(KLI) is recomputed, the it is updated as: v(n) = frac*v(KLI) + (1-frac)*v(n-1). The default for frac is 1.0 (i.e. no mixing). Supplying a value smaller than 1 may improve SCF convergence.

`NHOMO Nalpha {Nbeta}`

Chooses treatment of the free parameter in the KLI approximation (see S. Patchkovskii, J. Autschbach, and T. Ziegler, JCP 115, 26, Ref.[47]). The default is to choose the free parameter such that per-orbital potential shifts are non-negative. If this key is supplied, the per-orbital shifts of Nalpha/Nbeta highest occupied orbitals will be set to zero instead. Both choices are usually identical for converged solutions, but the default exhibits much better convergence behavior.

`CORE cor`

cor can be one of:
`IGNORE`: Ignores contributions of frozen core orbitals to SIC energy and potential. Only valence orbitals will appear in KLI potential mixing expression (eq. 14 of [49]).
`RHO`: Includes core density in the total density of eq.14, but ignore core orbitals otherwise.
`POTENTIAL`: Includes SIC potential of the core orbitals in v(KLI), but sets corresponding potential shifts to zero.
`FULL`: Full treatment of the frozen core - frozen core orbitals participate in KLI potential equilibration (eq. 16) on the equal footing with the valence orbitals. This is the default.
For POTENTIAL and FULL, tesseral harmonics are used for the angular part of the core orbitals.

`DENSITY mode`

This keyword controls evaluation of per-orbital densities in a SIC calculation. mode can be one of:
`EXACT`: Evaluate densities from molecular orbitals. This is the default.
`FIT`: Evaluate densities from auxiliary fits. Using this option is not recommended - it is both slower, and less accurate than EXACT.

`SHIPV filename`

Write v(KLI) contribution to XC potential on grid, to an external file. When using this option, it is a good idea to write out the numerical integration grid as well, by adding "SAVE TAPE10" to the ADF input file.

```
READV filename
```

v(KLI) is taken from an external file, and add it to the XC potential. Current integration grid must match the grid used to calculate the potential. The only certain way to guarantee this is to save the integration grid on TAPE10, and pass it around, together with the v(KLI) potential. Specifying this keyword will deactivate add other processing in SIC code, including calculation of Perdew-Zunger energies and SIC potential updates.

```
READMOS filename {FREEZE}
```

For the first evaluation of the v(KLI) potential, loads localized orbitals from the specified file, instead of localizing canonical Kohn-Sham MOs. Unless FREEZE is specified, subsequent SCF cycles will use localized canonical MOs.

```
WRITEMOS filename
```

Stores localized orbitals to an external file. Reading these orbitals back with READMOS provides a rudimentary restart capability.

```
NAILCANONICALS {eps}
```

Stabilize orientation of degenerate canonical MOs prior to localization. The default is not to stabilize. eps is the degeneracy criterion in eV (0.001 by default). Supplying this key may improve convergence when high local symmetry is present in a molecule.

```
NPFITS n
```

Number of fit coefficient sets to be computed in a single pass. Large values will improve performance, but need more memory. The default of 15 is usually adequate.

General remarks

- The phrase non-local in the discussion of density functionals does not mean that non-local potentials are involved. The potentials are perfectly local, but when you go beyond LDA and include gradient corrections, the value of the density functional potential in a point r is evaluated not only from the local value of the charge density, but also from the gradient of the charge density.
- The Stoll formula is considered to be a *correlation* correction to the *Local* Density Approximation. It is conceptually not correct to use the Stoll correction *and* apply non-local gradient (GGA) corrections to the correlation. It is the user's responsibility, in general and also here, to avoid using options that are not solidly justified theoretically.
- It is questionable to apply gradient corrections to the *correlation*, while not doing so at the same time for the exchange. Therefore the program will check this and stop with an error message. This check can be overruled with the key ALLOW.
- The issue of the 'best' density functional is a subject of extensive and widespread research. It is generally recognized that applying gradient corrections to the simplest Local Density Approximation usually gives better results for comparison with experimental data, especially as regards bond energies and the spectra computed from one-electron energies.
- The incorporation of gradient corrections during the SCF significantly increases the computing effort. In this respect it makes no difference which specific GGA formula is applied. The Energy (PostSCF) feature is therefore an alternative worthwhile considering: it saves a lot of time and the effects of this approximation are often small as regards the SCF solution, so the non-self-consistent aspect hardly shows up in the computed bond energy. In Geometry Optimizations, however, the Post-SCF option implies that the energy gradients are computed from the LDA

energy expression and hence the resulting optimized geometry corresponds to the LDA functional. In such a case, including the GGA term may make a substantial difference to the computed equilibrium geometry.

Dispersion corrected functionals

DFT-D3 functionals

In ADF2012 Stefan Grimme's latest dispersion correction is implemented. Grimme and his coworkers at the Universität Münster outlined the parameterization of this new correction, dubbed DFT-D3, in Ref. [292]. A slightly improved version with a more moderate BJ damping function appeared later, and was called DFTB-D3-BJ. [436] Here they list the advantages of the new method as the following:

- It is less empirical, i.e., the most important parameters are computed from first principles by standard Kohn-Sham (KS)-(TD)DFT.
- The approach is asymptotically correct with all DFs for finite systems (molecules) or nonmetallic infinite systems. It gives the almost exact dispersion energy for a gas of weakly interacting neutral atoms and smoothly interpolates to molecular (bulk) regions.
- It provides a consistent description of all chemically relevant elements of the periodic system (nuclear charge Z = 1-94).
- Atom pair-specific dispersion coefficients and cutoff radii are explicitly computed.
- Coordination number (geometry) dependent dispersion coefficients are used that do not rely on atom connectivity information (differentiable energy expression).
- It provides similar or better accuracy for "light" molecules and a strongly improved description of metallic and "heavier" systems.

DFT-D3-BJ is invoked with the XC block, for example

```
XC
  GGA BLYP
  Dispersion Grimme3 BJDAMP
END
```

Parametrizations are available for: B3LYP, TPSS, BP86, BLYP, revPBE, PBE, PBEsol, and RPBE (not for BJ damping), and will be automatically set if one of these functionals is used. Otherwise PBE parameters will be used. The parameters can be set manually, see the XC key block.

DFT-D functionals

An implementation for dispersion corrections based, called DFT-D is available starting from ADF2008. Like DFT-D3 this implementation is easy to use and is also supported by the GUI.

This DFT-D implementation is based on the paper by Grimme [226] and is extremely easy to use. The correction is switched on by specifying *DISPERSION*, possibly with parameters, in the XC input block. See description of the XC input block for details about the DISPERSION keyword.

Energies calculated Post-SCF using different DFT-D or GGA-D functionals are also present in table printed when METAGGA keyword is specified. These include: BLYP-D, PBE-D, BP86-D, TPSS-D, B3LYP-D, and B97-D. NOTE: this option does not require specifying a DISPERSION keyword in the XC block and thus there is **no correction added to the energy gradient** in this case. Please also note that although the original B97 functional includes HF exchange (and is thus a hybrid functional), the B97-D is a pure GGA.

B3LYP-D is, however, a hybrid functional. The following functional-dependent global scaling factors s_6 are used: 1.2 (BLYP-D), 0.75 (PBE-D), 1.05 (BP86-D), 1.0 (TPSS-D), 1.05 (B3LYP-D), and 1.25 (B97-D). These are fixed and cannot be changed.

Regarding performance of different functionals, testing has shown that BLYP-D gives good results for both energies and gradients involving VdW interactions. Post-SCF energy-only calculations at fixed geometries showed that also B97-D gives good binding energies compared to high-level reference data. Thorough comparison of different DFT-D functionals can be found in ref. [227]

Note: The original paper by Grimme included parameters for elements H throughout Xe. In ADF2009.01 values for dispersion parameters for DFT-D functionals for heavier elements (Cs-Rn) have been added. These new values have not been tested extensively. Thus, in this implementation, no dispersion correction is added for interactions involving atoms heavier than Radon.

DFT-D is invoked with the XC block, for example

```
XC
  GGA BLYP
  Dispersion
END
```

MM dispersion (old implementation)

The idea to get an accurate description of van der Waals complexes by density functional theory by including empirical corrections by Grimme [211] was implemented in ADF by J.M. Ducere from the group of Prof. L. Cavallo [215]. Please contact this group for more details on this functionality.

This is an expert option. As input one needs certain atomic parameters and (for a given basis set and functional optimized) parameters for a damping function. At the moment only for a few atoms atomic parameters can be found in the file $ADFHOME/atomicdata/MMDispersion/disp-param. Only for the PBE functional with a DZP or TZP basis set parameters are optimized for the damping function. This optimization was done with respect to MP2 theoretical data. The parameters from Grimme's paper can also be used.

```
MMDispersion
  {FILE_NAME  filename}
  {DAMPING    damping}
  {DAMP_PARAM damp_param {a b c}}
  {COMBI      combi}
  {DISPALL}
  {NODEFAULT}
  {ATOMTYPE
    attype  c6  pol  rad
  SUBEND}
End
```

`FILE_NAME filename`

Optional. The filename (full path) from which are read the C6 parameters, polarizabilities and radii. The file is expected to have the following structure:

```
 attype  c6  pol  rad
```

The attype must exactly match the atom-type name present in the ATOMS key-block (case-sensitive), for being recognized; c6, pol, and rad are in atomic units (hartree and bohr). A "---" sequence indicates the end of the read part. Even if the sqrt option is chosen for COMBI, a polarizability is needed. If the

environment variable ADFRESOURCES ($ADFHOME/atomicdata) is set, the default value for filename is $ADFRESOURCES/MMDispersion/disp-param.

`DAMPING damping`

Optional. Defines the kind of damping function to be used, damping can be one of:
sigm: sigmoid (default)
fermi: Fermi-like function (Grimme [211])

`DAMP_PARAM damp_param {a b c}`

Optional. Defines which parameters of the damping function should be used, damp_param can be one of:
tz: parameters optimized for PBE/TZP (default)
dz: parameters optimized for PBE/DZP
grimme: parameters from Grimme paper
cust a b c: parameters are a, b and c

`COMBI combi`

Optional. Defines the kind of combination rule to be used, combi can be one of:
s-k: Slater-Kirkwood combination rule (default)[214]
sqrt: square-root combination rule

`DISPALL`

Optional. If present, all atom-pairs are considered, else, only contributions from different fragments (different indexes, see below) are considered. DISPALL is NOT the default.

`NODEFAULT`

Optional. By default, if there is no match for a given atom-type, ADF looks in the parameter file specified in FILENAME for atomic default parameters (Grimme's ones). NODEFAULT switches off this check. Example: suppose the atom-type is H.text. By default if there is no match for H.text, but there is a match for H, parameters for H will be used. If NODEFAULT is set, and there is no match for H.text an error message is printed and ADF will stop.

`ATOMTYPE`

Optional. For input supplied c_6, polarizability and radius paramaters of atom-types; attype must exactly match an atom-type name present in the ATOMS block for being recognized; c_6, pol and rad are in a.u.

Atom-types and fragment-indexes are specified in the ATOMS keyblock:

```
ATOMS
  atom-type  x  y  z  FD=n
  ...
END
```

FD is the index of the fragment. FD=0 switch off the calculation for the atom. If DISPALL is present in the input, non-zero values of FD only have an analytical role. If DISPALL is not present, the contributions are calculated between atoms of different non-zero values of FD. By default, FD=1 for all atoms.

dDsC: density dependent dispersion correction

The `DISPERSION dDsC` key invokes the density dependent dispersion correction [316], which has been parametrized for the functionals BLYP, PBE, BP, revPBE, B3LYP, PBE0 and BHANDHLYP.

```
XC
  GGA BLYP
  Dispersion dDsC
END
```

DFT-ulg

The `DISPERSION UFF` key invokes the universal correction of density functional theory to include London dispersion (DFT-ulg) [366], which has been parametrized for all elements up to Lr (Z=103), and for the functional PBE, PW91, and B3LYP. For other functionals the PBE pareameters will be used. Example:

```
XC
  GGA PBE
  Dispersion UFF
END
```

DFT-MBD functionals

The `DISPERSION MBD` key invokes the MBD@rsSCS method [377], which is designed to accurately describe long-range correlation (and thus dispersion) in finite-gap systems, including at the same time a description of the short-range interactions from the underlying DFT computation of the electronic structure. The MBD (many-body dispersion) method [376] obtains an accurate description of van der Waals (vdW) interactions that includes both screening effects and treatment of the many-body vdW energy to infinite order. The revised MBD@rsSCS method [377] employs a range-separation (rs) of the self-consistent screening (SCS) of polarizabilities and the calculation of the long-range correlation energy. It has been parametrized for the elements H-Ba, Hf-Rn, and for the functional PBE and PBE0. Note that the MBD@rsSCS method depends on Hirshfeld charges. In calculating forces the dependence of the Hirshfeld charges on the actual geometry is neglected. The MBD method is implemented in case the BeckeGrid is used for the numerical integration. Example for PBE MBD@rsSCS:

```
XC
  GGA PBE
  Dispersion MBD
END
```

One can use user defined values with:

```
XC
  Dispersion MBD {RSSCS|TS} {BETA=beta}
END
```

`MBD {RSSCS|TS} {BETA=beta}`

The default method for MBD is MBD@rsSCS. Optionally one can use MBD@TS or change the used parameter β with setting beta.

Relativistic effects

```
RELATIVISTIC {level} {formalism} {potential}
```

`Level`

May be None (this suppresses the key, and is equivalent to not using the key at all), Scalar (default: scalar relativistic effects), or SpinOrbit (using double group symmetry).

`Formalism`

Pauli (default) or ZORA (ZORA is recommended!)

`Potential`

SAPA (default) or Full. The Full option is obsolete. It is here mainly for historical reasons. The SAPA method is described in Ref.[50] for the BAND program. The same potential is used in the ADF program. One may think that the Full option gives extra accuracy. However, this is not the case, it only leads to extra CPU time and extra DISK space usage.

The key RELATIVISTIC instructs ADF to take relativistic effects into account. By default (omission of the key) this is suppressed. Recommendation use: Relativistic Scalar ZORA or Relativistic SpinOrbit ZORA.

Pauli

Specification of the Pauli formalism means that the first order relativistic corrections (the Pauli Hamiltonian) will be used [51-60]. In a *scalar* relativistic run ADF employs the single point group symmetry and only the so-called *scalar* relativistic corrections, Darwin and Mass-Velocity. The treatment is not strictly first-order, but is *quasi*-relativistic, in the sense that the first-order scalar relativistic Pauli Hamiltonian is diagonalized in the space of the non-relativistic solutions, i.e. in the non-relativistic basis set.

The quasi-relativistic approach improves results considerably over a first-order treatment. There are, however, theoretical deficiencies due to the singular behavior of the Pauli Hamiltonian at the nucleus. This would become manifest in a complete basis set but results are reasonable with the normally employed basis sets. However, this aspect implies that it is not recommended to apply this approach with an all-electron basis set for the heavy atoms, and for very heavy elements even a frozen core basis set often fails to give acceptable results. The problems with the quasi relativistic approach of the Pauli Hamiltonian are discussed for example in Ref.[61].

ZORA

The ZORA approach gives generally better results than the Pauli formalism. For all-electron calculations, and in fact also for calculations on very heavy elements (Actinides), the Pauli method is absolutely unreliable. Therefore, with its formal introduction in ADF1999, the ZORA method is the recommended approach for relativistic calculations with ADF.

ZORA refers to the Zero Order Regular Approximation [61-65]. This formalism requires special basis sets, primarily to include much steeper core-like functions; applying the ZORA method with other, not-adapted basis sets, gives unreliable results. The ZORA basis sets can be found in the ADF database, in subdirectories under the $ADFHOME/atomicdata/ZORA directory.

The ZORA formalism can also be used in Geometry Optimizations. However, there is a slight mismatch between the energy expression and the potential in the ZORA approach, which has the effect that the geometry where the gradients are zero does not exactly coincide with the point of lowest energy. The differences are very small, but not completely negligible, order of magnitude: less than 0.001 Angstrom. In ADF2010 this difference is reduced to order 0.0001 Angstrom. In ADF2010 also the calculation of analytical frequencies has improved in case of QZ4P basis sets and heavy elements, like uranium.

Spin-Orbit coupling

The Spin-Orbit option uses double-group symmetry. The symmetry-adapted orbitals are labeled by the quantum number *J* rather than *L* and any references in input to subspecies, such as a specification of occupation numbers, must refer to the double group labels.

Create runs must *not* use the Spin-Orbit formalism. The SFO analysis of Molecular Orbitals for a Spin-Orbit calculation is only implemented in the case of a scalar relativistic fragment file, which is the whole molecule. Starting from the ADF2007.01 version gradient calculations for the Spin Orbit formalism have been implemented. Therefore, you may now calculate harmonic frequencies (numerical) and do geometry optimizations including spin-orbit coupling.

In a Spin-Orbit run each level can allocate 2 electrons (times the dimension of the irreducible representation) as in a normal restricted calculation. However, contrary to the normal case these two electrons are not directly associated with spin-α and spin-β, but rather with the more general Kramer's symmetry. Using the unrestricted feature in order to assign different numbers of electrons to a and b spin respectively cannot be applied as such. However, one can use the unrestricted option in combination with the collinear or noncollinear approximation. In that case one should use symmetry NOSYM, and each level can allocate 1 electron.

Relativistic core potentials

In all relativistic calculations - scalar as well as spin-orbit - the relativistic atomic core densities and the relativistic atomic potentials *must* be made available to ADF on a file specified with the key COREPOTENTIALS. Starting from the ADF2006.01 release this is necessary only in the 'create' run of the atoms. In the molecular calculation this key is not required anymore. If supplied then the file must contain data for *all* atom types in the molecule, even for those atoms where relativistic aspects are expected to be negligible or that may not have a frozen core at all (such as Hydrogen). Excepted are any Ghost atoms (for instance for a BSSE calculation): these can not have any core potentials. This is tested by the program, internally, by looking at the nuclear charge and at the number of electrons belonging to an atom: if both numbers are zero, no (relativistic or other) core potential is allowed. Also the potential used in the ZORA kinetic energy operator in the SAPA (sum of neutral atomic potential approximation) method should be present on this file (which will be the case if the program DIRAC is used to generate this file).

Relativistic potentials can and should be generated with the auxiliary program dirac, see the next section, and the examples.

As of ADF2003, the recommended way to generate atomic fragments and relativistic potentials is by using the BASIS keyword.

Dirac program: Relativistic Potentials

The auxiliary program DIRAC, which is installed together with ADF, serves to compute relativistic frozen core potentials (and densities), necessary to apply the (scalar) relativistic option in ADF. The database *atomicdata* has a subdirectory *Dirac*, which contains input files for DIRAC for all atoms of the periodic table of elements. The *names* of the input files indicate the frozen core level: *Ag.3d* for instance is the input file for a calculation on a Silver atom with a frozen core up to and including the 3d shell (i.e.: 1s, 2s, 2p, 3s, 3p, and 3d). The frozen core level used in the DIRAC calculation defines the core data computed and should therefore match the frozen core level in the ADF Create run for the atom that it will be used for.

A DIRAC run with the inputs provided in the database involves a fully relativistic calculation on the atom (spin-orbit coupling, double group symmetries). It generates a file TAPE12 with the corresponding core potential and density (a table of values for a sequence of radial distance values). Other files produced by

DIRAC should be removed after the DIRAC run; they are needed for other applications of the program but play no role here.

If you run DIRAC while a file TAPE12 already exists the computed core data will be written at the end of it, after the existing data. The program will assume, however, that the existing data on the file are also core-data from DIRAC runs, and may abort otherwise.

Starting from ADF2006.01 it is not necessary anymore to make one big TAPE12 which contains data for all atoms involved in the molecular calculation. Instead only in the ADF Create run for each atom one needs a TAPE12 which contains data for the atom that is created. The corresponding core data is written to the TAPE21 of this atomic fragment. In the molecular ADF run one then should not include the CorePotentials key, such that ADF will read core data on the TAPE21's of the atomic fragments. One can still use the CorePotentials key, but then one should proceed as in previous releases.

In previous releases (ADF2005 and older), if a CorePotentials file was needed for an *adf* calculation with the (scalar) relativistic option, the simplest approach was to subsequently run DIRAC for each of the involved atoms types. This builds up the TAPE12 file for this particular molecule. Then, specify in the ADF input which sections correspond to the distinct atom types. Alternatively, which we do not recommend, if you frequently perform relativistic ADF runs, with many different types of atoms, you might, once and for all, construct one big TAPE12 file, containing the core potentials of all atoms that you may ever need, and use that file again and again. Of course you need then to remember which section numbers correspond to which atoms.

Implied options

The DIRAC calculations imply the *local* Density Functional in its simple X-alpha approximation without any gradient corrections. Not the *scalar* relativistic but the *fully* relativistic Hamiltonian is used, including spin-orbit coupling. In ADF you may use the *scalar* relativistic Hamiltonian and most users will employ a more sophisticated lda than X-alpha, such as the default vwn (Vosko, Wilk, Nusair) formulas, and may in addition routinely apply gradient corrections. The core potential may not exactly match the Fock operator applied in the molecular calculation. The effect is very small and one can neglect the discrepancy.

Input

The ascii input files for DIRAC, as available in the database directory *$ADFHOME/atomicdata/Dirac* (point nucleus) and *$ADFHOME/atomicdata/Dirac* (finite nucleus), have a structure as described below. With this information you should be able to construct alternative input files, with other frozen cores for instance.

1 Title (60 characters at most). Plays no role

2 Ngrid, Nshell, rmin, rmax, Z, Xion, Anuc
Ngrid=number of radial grid points, in which the core potentials are computed.
Nshell=number of atomic orbital shells
rmin, rmax=minimum and maximum radial grid values
Z=nuclear charge
Xion=net charge of the 'atom'
Anuc=atomic weight

3 Pinit, Pfinal, eps, del, delrv
Pinit, Pfinal= initial and final density iteration averaging factors. Each iteration cycle changes the actual averaging factor by taking the average of the previous and the final one, starting with the 'initial' one.
eps=Exp(-sqrt(eps)) is set to zero, so eps determines the exponential underflow control.
del=absolute convergence criterion for orbital eigenenergies.
delrv=convergence criterion on the potential (multiplied by the radial distance *r*).

4 Idirc, Nmax, Ndebu, Nprin, Ipun, Ircor, Iwcor
Idirc=zero for non-relativistic, otherwise one.
Nmax=maximum number of iterations allowed to reach convergence.

Ndebu=non-zero for additional output (for debugging purposes mainly)
Nprin=print parameter. Use 2 or larger to get the orbitals printed.
Ipun=punched output is produced if Ipun is non-zero. (out-of-date)
Ircor=number of core orbitals from the fully relativistic run, to be kept frozen in the subsequent (if any) first-order perturbation calculation.
Iwcor=number of core orbitals used to construct the core density and core potential, that are output on TAPE12. So, here you specify the relativistic core.

5 Xalph, Xlatt, Rnuc
Xalph= Exchange parameter in the Xalpha formalism.
Xlatt=Coulomb tail parameter
Rnuc=size of nuclear radius, in bohr. If set to 1.0 or larger, it is recomputed as 0.0000208*Anuc**(1/3)

6 For each orbital shell:
N, L, J, E, Z, D
N, L, J = The usual orbital quantum numbers. J is used only for relativistic runs.
E = Initial estimate of orbital energy, in atomic units.
Z = Number of electrons on the shell
D = Initial estimate of the error in the orbital energy

7 Icorp, Npcl, Demp, Peps
Icorp = If 1 (one): Do a first order perturbation calculation after the fully relativistic run. This option plays no role in the current application for ADF.
Npcl = Maximum number of cycles in the perturbation calculation.
Demp = Damping factor in the perturbation iterations
Peps = Convergence criterion in the perturbation iterations.

Solvents and other environments

COSMO: Conductor like Screening Model

You can study chemistry in solution, as contrasted to the gas phase, with the implementation in ADF [66] of the Conductor like Screening Model (COSMO) of solvation [67-69]. The energy derivatives can also be calculated, so geometry optimization, harmonic frequencies, et cetera are available within this model.

The COSMO model is a dielectric model in which the solute molecule is embedded in a molecule-shaped cavity surrounded by a dielectric medium with given dielectric constant ε. Energy-related terms are computed for a conductor first, then scaled by the function

$$f(\varepsilon) = (\varepsilon-1)/(\varepsilon+x) \quad (2.1.1)$$

The empirical scaling factor x is specified in the input data block for the SOLVATION key. The block key SOLVATION turns the solvation calculation on. In most cases default values are available for the involved parameters.

It is also possible to include a linear parameterization of non-electrostatic terms as a function of surface area. To include such term can be specified in the input data block for the SOLVATION key. Starting from ADF2012 the default is to include only the part of this term that is proportional to the surface area (default CAV0=0.0, CAV1=0.0067639):

$$E_{\text{non-elst}} = f(\varepsilon) \times (\text{CAV0} + \text{CAV1}\times area) \quad (2.1.2)$$

The COSMO routine has never been tested with non-gas phase fragments. It is designed to correct the "total" energy (wrt fragments gas phase) for solvation. The COSMO energy of the fragments is never taken

into account, not even for the atoms, and no information is passed in regarding the COSMO energy when the fragments are defined.

If a calculations was done on a fragment, then the wavefunction obtained would be optimal for the fragment in solution, but not optimal for gas phase. The energies with the gas phase Hamiltonian would be higher, and the apparent solvation contribution to bonding would also be higher. The net point is that normally the COSMO procedure reports the energy of Esolv(AB), but to get the solvation energy, you need to subtract the E(AB, solv) from E(AB,gas) because the wavefunction changes (unless you are doing it post-SCF.) For this you need to have the same reference fragments in each case A(g) and B(g).

```
SOLVATION
 {SURF Esurf {NOKEEP}}
 {SOLV {Name=solvent} {Eps=78.4} {Del=1.4} {Rad=1.4}
       {Neql=1.9}{Emp=0.0}{Cav0=0.0}{Cav1=0.0067639} }
 {DIV {Ndiv=3} {NFdiv=1} {Min=0.5} {OFAC=0.8}
      {leb1=23} {leb2=29}{rleb=1.5} }
 {NOASS}
 {RADII
 name1=value1
 name2=value2
  ...
 subend }
 {CHARGED {Method=meth} {Conv=1e-8} {Omega=1.0} (Iter=1000} {Corr} }
 {C-MAT How {SCF} tol=1e-10 }
 {DISC {SC=0.01} {LEG=4} {TOL=0.1} }
 {SCF {When} {How}
 {NOCSMRSP}
 {LPRT}
End
```

Presence of the SOLVATION key block triggers the solvent calculation and does not require additional data. With subkeys you can customize various aspects of the model, for instance to specify the type of solute. None of the subkeys is obligatory. Follows a description of the subkeys

`SURF Esurf {NOKEEP}`

Esurf must be Wsurf, Asurf, Esurf, Klamt, or Delley. Five different cavity types are available. In ADF2010 the numerical stability of the COSMO surface has been improved, by merging close lying COSMO surface points, and removing COSMO surface points with a small surface area. The Wsurf, Asurf, and Esurf surfaces are constructed with the GEPOL93 algorithm [70]

`Wsurf`

Wsurf triggers the Van der Waals surface (VdW), which consists of the union of all atomic spheres. Not recommended to be used.

`Asurf`

Asurf gives the Solvent-Accessible-Surface (SAS). This is similar to VdW but consists of the path traced by the center of a spherical solvent molecule rolling about the VdW surface or, equivalently, a VdW surface created by atomic spheres to which the solvent radius has been added. These two surface types contain cusps at the intersection of spheres. Not recommended to be used.

`Esurf`

Esurf (the default) gives the Solvent-Excluding-Surface (SES), which consists of the path traced by the *surface* of a spherical solvent molecule rolling about the VdW surface. Primarily, this consists of

the VdW surface but in the regions where the spheres would intersect, the concave part of the solvent sphere replaces the cusp. This SES surface is the default in ADF.

`Klamt`

The fourth surface option is Klamt as described in [67]. It excludes the cusp regions also. Note that this surface might give an incomplete COSMO surface in case of more complicated molecules. Not recommended to be used.

`Delley`

The fifth surface is the so called Delley surface, see also Ref. [245]. This Delley type of cavity construction is recommended to be used in COSMO calculations, which results are used as input for COSMO-RS calculation, see the corresponding manual for COSMO-RS.

`NOKEEP`

The optional parameter NOKEEP controls surface creation during calculation of frequencies by numerical differentiation. By default, the surface is constructed only once at the central geometry and is used for the rest of the calculation. If the NOKEEP is specified then ADF will construct a new surface at each displaced geometry. The NOKEEP option was the default in ADF2005 and earlier versions but it was found to cause problems. Since ADF2006 one needs to specify SURF NOKEEP to get the same behavior.

`DIV`

The actual construction of the surface involves a few technical parameters controlled with the subkey DIV

`Ndiv, NFdiv`

Ndiv controls how fine the spheres that in fact describe the surface are partitioned in small surface triangles, each containing one point charge to represent the polarization of the cavity surface. Default division level for triangles Ndiv=3. Default final dision level for triangles NFdiv=1 (NFdiv≤Ndiv). Not used in the Delley surface.

`Min`

Min specifies the size, in angstrom, of the smallest sphere that may be constructed by the SES surface. For VdW and SAS surfaces it has no meaning. Default Min=0.5. Not used in the Delley surface.

`Ofac`

Ofac is a maximum allowed overlap of new created spheres, in the construction procedure. Default Ofac=0.8. Not used in the Delley surface.

`leb1, leb2, rleb`

Only used in the Delley surface. For the Delley type of construction one needs to set the variables `leb1` (default value 23), `leb2` (default value 29), and `rleb` (default value 1.5 Angstrom) to set the number of surface points. If the cavity radius of an atom is lower than `rleb` use `leb1`, otherwise use `leb2`. These values can be changed: using a higher value for `leb1` and `leb2` gives more surface points (maximal value `leb1`, `leb2` is 29). A value of 23 means 194 surface points in case of a single atom, and 29 means 302 surface points in case of a single atom Typically one could use `leb1` for the surface point of H, and `leb2` for the surface points of other elements.

`NOASS`

By default all new spheres that are created in the surface-construction are assigned to atoms, for the purpose of gradient computations (geometry optimization). Specifying the noass subkey turns this off. It has no argument.

SOLV

Solvent details.

Eps, Rad

Eps specifies the dielectric constant (the default relates to water). In ADF an infinite value for Eps is chosen if Eps is specified to be lower than 1.0. Rad specifies the radius of the (rigid sphere) solvent molecules, in angstrom. Instead of specifying Eps and Rad one can specify a solvent name or formula after 'name='. The following table lists names and formulas that are recognized with the corresponding values for Eps and Rad. The Rad in this table are calculated from the density, the molar mass, and a spherical approximation for the solvent. The names and formulas are case-insensitive.

Name	Formula	Eps	Rad
AceticAcid	CH3COOH	6.19	2.83
Acetone	CH3COCH3	20.7	3.08
Acetonitrile	CH3CN	37.5	2.76
Ammonia	NH3	16.9	2.24
Aniline	C6H5NH2	6.8	3.31
Benzene	C6H6	2.3	3.28
BenzylAlcohol	C6H5CH2OH	13.1	3.45
Bromoform	CHBr3	4.3	3.26
Butanol	C4H9OH	17.5	3.31
isoButanol	(CH3)2CHCH2OH	17.9	3.33
tertButanol	(CH3)3COH	12.4	3.35
CarbonDisulfide	CS2	2.6	2.88
CarbonTetrachloride	CCl4	2.2	3.37
Chloroform	CHCl3	4.8	3.17
Cyclohexane	C6H12	2	3.5
Cyclohexanone	C6H10O	15	3.46
Dichlorobenzene	C6H4Cl2	9.8	3.54
DiethylEther	(CH3CH2)2O	4.34	3.46
Dioxane	C4H8O2	2.2	3.24
DMFA	(CH3)2NCHO	37	3.13
DMSO	(CH3)2SO	46.7	3.04
Ethanol	CH3CH2OH	24.55	2.85
EthylAcetate	CH3COOCH2CH3	6.02	3.39
Dichloroethane	ClCH2CH2Cl	10.66	3.15
EthyleneGlycol	HOCH2CH2OH	37.7	2.81
Formamide	HCONH2	109.5	2.51
FormicAcid	HCOOH	58.5	2.47
Glycerol	C3H8O3	42.5	3.07
HexamethylPhosphoramide	C6H18N3OP	43.3	4.1
Hexane	C6H14	1.88	3.74
Hydrazine	N2H4	51.7	2.33
Methanol	CH3OH	32.6	2.53
MethylEthylKetone	CH3CH2COCH3	18.5	3.3
Dichloromethane	CH2Cl2	8.9	2.94

Methylformamide	HCONHCH3	182.4	2.86
Methypyrrolidinone	C5H9NO	33	3.36
Nitrobenzene	C6H5NO2	34.8	3.44
Nitrogen	N2	1.45	2.36
Nitromethane	CH3NO2	35.87	2.77
PhosphorylChloride	POCl3	13.9	3.33
IsoPropanol	(CH3)2CHOH	19.9	3.12
Pyridine	C5H5N	12.4	3.18
Sulfolane	C4H8SO2	43.3	3.35
Tetrahydrofuran	C4H8O	7.58	3.18
Toluene	C6H5CH3	2.38	3.48
Triethylamine	(CH3CH2)3N	2.44	3.81
TrifluoroaceticAcid	CF3COOH	42.1	3.12
Water	H2O	78.39	1.93

`Del`

Del is the value of Klamt's delta_sol parameter, only relevant in case of Klamt surface.

`Neql`

If Neql=ε_{NEQL} is included a nonequilibrium solvation is used, i.e. that the dielectric constant ε_{NEQL} used in RESPONSE is different from the ground state dielectric constant ε. Only relevant in case of TDDFT calculations. Default $\varepsilon_{NEQL} = \varepsilon$. The reason for using two different dielectric constants is that the electronic transition can so fast that only the electronic component of the solvent dielectric can respond, i.e., one should use the optical part of the dielectric constant. This is typically referred to as non-equilibrium solvation. The optical dielectric constant can be obtaining from the (frequency dependent) refractive index n of the solvent as: $\varepsilon_{neql} = n^2$.

`Emp`

Emp addresses the empirical scaling factor *x* in the formula 2.1.1 above.

`Cav0, Cav1`

Other options specify a linear parameterization of non-electrostatic terms as a function of surface area, see the formula 2.1.2 above. Possible values for CAV0 and CAV1 are CAV0 = 1.321 and CAV1 = 0.0067639, see Ref. [299]), which were the default values for CAV0 and CAV1 in ADF2009. In ADF2010 the default values for CAV0 and CAV1 are CAV0 = 0.0 and CAV1 = 0.0. However, starting from ADF2012 the default values for CAV0 and CAV1 are CAV0 = 0.0 and CAV1 = 0.0067639, If CAV0 is not zero, Esolv(AB) is not the same as Esolv(A) + Esolv(B) if A and B are far apart. This is the reason why CAV0 is set to zero, by default. By default CAV1 is not set to zero, thus by default there is a solvation energy term that does depend on the size of the cavity (surface area).

`COSMO Radii`

In order to construct the surface you have to specify the atomic ('Van der Waals') radii. There are three ways of doing this. In the first method you append 'R=value' to the atomic coordinates record, in the ATOMS key block. This would look like, for instance

```
C 1 2 3 CC CCO CCOH f=C.dz R=2.0
```

It assigns a radius of 2.0 to the Carbon atom.
In the second method you apply the same format, but specify a symbol (identifier) rather than a value

```
C 1 2 3 CC CCO CCOH f=C.dz R=C-sp3
```

The identifiers must be defined in the (optional) RADII subkey block in the Solvation data block (see next).
In the third method, you don't modify the Atoms block at all. In this case, the RADII subkey must be used and the 'identifiers' in it must be exactly the atom type names in the Atoms block.

RADII

This subkey is block type. Its data block (if the subkey is used) must terminate with a record subend. In the Radii data block you give a list of identifiers and values

```
SOLVATION
  ...
 Radii
 name1=value1
 name2=value2
  ...
 Subend
  ...
End
```

The values are the radii of the atomic spheres, in the same units of length as used in the Atoms block (angstrom or bohr). The names specify to which atoms these values apply. As discussed for the Solv subkey this depends on the Atoms block. If in the specification of atomic coordinates you have used the 'R=' construct to assign radii, with identifiers rather than values for the R-value, these identifiers must be defined in the Radii sub block. If no 'R=' construct was applied in the Atoms block, you must use the atom type names as they occurred in the Atoms data block. Referring to the example given in the Solv subkey discussion, you might have

```
...
 Radii
  C-sp3=2.0
  ...
 Subend
...
```

A simple atom type reference might look like

```
...
 Radii
  C=2.0
  ...
 Subend
...
```

When no radius specified a default value is used. The default value for an atom is the corresponding Van der Waals radius from the MM3 method by Allinger (Ref. [290]) divided by 1.2.

This concludes the discussion of the Radii subkey.

CHARGED

This addresses the determination of the (point) charges that model the cavity surface polarization. In COSMO calculations you compute the surface point charges q by solving the equation $A\mathbf{q}=-\mathbf{f}$, where $\mathbf{f}$ is the molecular potential at the location of the surface charges q and A is the self-interaction matrix of the charges. The number of charges can be substantial and the matrix A hence very large. A direct method, i.e. inversion of A, may be very cumbersome or even impossible due to memory limitations, in which

case you have to resort to an iterative method.
Meth specifies the equation-solving algorithm. Meth=INVER requests direct inversion. Meth=GAUS calls for the Gauss-Seidel iterative method. Meth=Jacobi activates another standard iterative procedure. The latter two methods require a positive-definite matrix (which may fail to be the case in an actual calculation) and can be used with a relaxation technique, controlled by the relaxation parameter OMEGA (1.0=no relaxation).
Meth=CONJ (default) uses the preconditioned biconjugate gradient method. This is guaranteed to converge and does not require huge amounts of memory.
CONV and ITER are the convergence criterion and the maximum number of iterations for the iterative methods.

Some of the molecular electronic charge distribution may be located outside the cavity. This affects the assumptions underlying the COSMO equations. Specifying the CORR option to the CHARGED subkey constrains the computed solvent surface charges to add up to the negative of the molecular charge.

C-MATRIX

- How: For the potential *f* we need the Coulomb interaction between the charges *q* and the molecular electronic density (and nuclei). Three methods are available, specified by the first option to the C-Matrix subkey.
a) EXACT: compute the straightforward Coulomb potential due to the charge *q* in each point of the molecular numerical integration grid and integrate against the electronic charge density. This is, in principle, exact but may have inaccuracies when the numerical integration points are very close to the positions of a charge *q*. To remedy this, the point charges *q* can be 'smeared out' and represented by a disc, see the next subkey DISC.
b) FIT: same as EXACT, but the *q*-potentials are now integrated not against the exact electronic charge density, but against the (much cheaper-to-compute) fitted density. The same DISC considerations apply.
c) POT: evaluate the molecular potential at the position of the charge *q* and multiply against the *q*-strength. Since the molecular Coulomb potential is computed from the fit density, any difference in results between the FIT and the POT approach should be attributed to the DISC issue.
POT is the default, because it is faster, and is only inadequate if the fit density is very inaccurate, which would be a problem anyway.
- SCF: If you specify this option, the computation of the Coulomb interaction matrix (between electrons and surface charges) is carried out during the SCF procedure, but this turns out to hamper the SCF convergence behavior. Therefore: not recommended. *IF* you use it, the program will switch to one of the other 3 methods, as given by the 'How' option, as soon as the SCF convergence error drops below TOL: (applies only to the SCF option, which is not recommended).

DISC

Applies only when the C-matrix method is EXACT or FIT. Note, however, that the default for the C-matrix method is POT, in which case the DISC subkey has no meaning. The DISC key lets the program replace the point charges *q* by a solid uniformly charged spherical surface disc whenever the numerical integration accuracy requires so, i.e. for those charges that are close to numerical integration points.
Options:
SC defines a shrinking factor, by which the actual disc radius used is reduced from its 'normal' value: an inscribed disc in the triangular surface partitions that define the distribution of surface charges, see the subkey DIV.
LEG gives the polynomial expansion order of the disc potentials. The Legendre expansion converges rapidly and the default should be adequate.

TOL is a tolerance parameter to control the accuracy of the disc potential evaluations.

SCF

In COSMO calculations you can include the surface charges in the Fock operator self-consistently, i.e. by recomputing the charges *q* at every SCF cycle and include them in the equations, or in a

perturbational manner, i.e. post-SCF. This is controlled with the first option. The When option must be either VAR or PERT, for variational and perturbational, respectively. Default is VAR.
The second (HOW) option applies only to the WHEN=VAR case and may affect the speed of SCF convergence. The COSMO calculation implies a considerable increase in CPU time! Values for HOW:
- ALL: This includes it in all SCF cycles (except for the first SCF cycle, which is gas-phase)
- LAST: This lets the program first converge the SCF completely without any solvent effects. Thereafter, the COSMO is turned on, hopefully converging in fewer cycles now, to compensate for the 'double' SCF effort.
- TOL=0.1 (or another value) is an in-between approach: converge the gas-phase SCF until the SCF error is below TOL, then turn on COSMO.

`NOCSMRSP`

Relevant only in combination with the time-dependent DFT (TDDFT) applications: the EXCITATION, the RESPONSE, or the AORESPONSE key. If this subkey NOCSMRSP is included the induced electronic charges which are present in the TDDFT calculations, will not influence the COSMO surface charges. No dielectric constant in the response might be closer to the optical dielectric constant than using the full dielectric constant, see also subargument NEQL of the subkey SOLV of the key SOLVATION. By default, in absence of this subkey NOCSMRSP, the induced electronic will influence the COSMO surface charges. If one does geometry optimization of the excited state this makes sense, since then the solvent dielectric has time to fully respond. Note that inclusion of the key ALLPOINTS is needed in case of TDDFT COSMO calculations.

`LPRT`

This is a debug switch and triggers a lot more output related to the cavity construction etc.

Warning about frequencies with COSMO model

Numerical frequencies calculated with COSMO should be checked for stability with respect to the *disrad*, the numerical differentiation step size. The problem is that the COSMO surface changes slightly when a nucleus is moved from its equilibrium position. The change is usually smal but in some cases it may result is creation or annihilation of surface points, which will lead to discontinuities in the potential energy surface and may result in inaccurate frequencies.

Thus, when calculating vibrational frequencies numerically with COSMO, one should try decreasing the *disrad* value until no changes in frequencies are observed. However, the value should not be too small because then the total numerical noise may become too large compared to the generated forces. A general recommendation would be to try to decrease disrad by a factor of 2 at a time. Of course, this procedure may be very expensive for a large molecule. If this the case, one should use the SCANFREQ keyword and recalculate only a small number of frequencies. It should be noted that generally frequencies that have a small force constant are more sensitive to the numerical noise.

QM/MM: Quantum mechanical and Molecular Mechanics model

ADF supports the QM/MM method to handle large systems or environment effects by treating only part of the atoms quantum-mechanically and the other ones by molecular mechanics. Use of this feature is invoked by the QM/MM keyword (block type). The functionality and all details of the keyword, involving quite a few options and aspects, are described in the separate QM/MM manual.

See also the pdb2adf utility, described in detail in the ADF QM/MM document, which transforms a PDB file into an ADF input file, for use with QM/MM.

Quild: Quantum-regions Interconnected by Local Descriptions

The QUILD (Quantum-regions Interconnected by Local Description) program has been developed for enabling calculations through multi-level approaches, in which different computational treatments are used for different regions of the system under study, see the separate Quild manual.

DIM/QM: Discrete Interaction Model/Quantum Mechanics

The Discrete Interaction Model/Quantum Mechanics (DIM/QM) method facilitates calculating the optical properties of molecules under the influence of a discrete solvent or a metal nanoparticle, see for example Ref. [362]. DIM/QM relies on one of three descriptions of the system: Discrete Reaction Field (DRF), where the atoms interact via induced dipoles and static charge, Capacitance Polarizability Interaction Model (CPIM), where the atoms interact via induced dipoles and induced charges, and Polarizability Interaction Model (PIM), where the atoms interact via induced dipoles only. DRF is best for solvents, CPIM is best for small metal nanoparticles, and PIM is best for large metal nanoparticles.

The DRF module is now a part of DIM/QM, so DRF is now a submodule of DIM/QM.

To perform a DIM/QM calculation, two block keys are required. The first is the DIMQM block which activates the DIM/QM module. The parameters for each DIM atom must also be given, and they can be given with either the DIMPAR block which is most convenient for metal nanoparticles, or with the EXTERNALS block which is designed for DRF.

DRF: Discrete Solvent Reaction Field Model

The Discrete Solvent Reaction Field (DRF) model is a hybrid Quantum mechanical and Molecular Mechanics (QM/MM) model for studying solvation effects on (time-dependent) molecular properties such as dipole moments, excitation energies and (hyper)polarizabilities[141-145]. The classical solvent molecules are represented using distributed atomic charges and polarizabilities.

Within the Discrete Solvent Reaction Field model the QM/MM operator is

$$H_{QM/MM} = \sum_i v^{DRF}(r_i,\omega) = \sum_i v^{el}(r_i) + v^{pol}(r_i,\omega)$$

where the first term, v^{el}, is the electrostatic operator and describes the Coulombic interaction between the QM system and the permanent charge distribution of the solvent molecules. The second term, v^{pol}, is the polarization operator and describes the many-body polarization of the solvent molecules, i.e. the change in the charge distribution of the solvent molecules due to interaction with the QM part and other solvent molecules. The charge distribution of the solvent is represented by atomic point charges and the many-body polarization by induced atomic dipoles at the solvent molecules. The induced atomic dipole at site *s* is found by solving a set of linear equations

$$\mu^{ind}_{s,\alpha}(\omega) = \alpha_{s,\alpha\beta}\,[F^{init}_{s,\beta}(\omega) + \sum_{t\neq s} T^{(2)}_{st,\beta\gamma}\,\mu^{ind}_{t,\gamma}(\omega)],$$

where $\alpha_{s,\alpha\beta}$ is a component of the atomic polarizability tensor at site *s*. The screened dipole interaction tensor is given by

$$T^{(2)}_{st,\beta\gamma} = 3f^{T}_{st}R_{st,\alpha}R_{st,\beta}/R^{5}_{st} - f^{E}_{st}\delta_{\alpha\beta}/R^{3}_{st}$$

where the damping functions f^{T}_{st} and f^{E}_{st} have been introduced, see also [146]. A smeared-out point charge model [147] is used for short-range damping of the QM/MM operator

$1/R_{st} \rightarrow 1/S_{st} = erf(R_{st})/R_{st}$

The scaled distance, S_{st}, then replaces the normal distance, R_{st}, in the QM/MM operator.

In order to perform a DRF calculation two types of parameters (model atomic charges and atomic polarizabilities) for each type of atom in the MM part are required. The point charges should represent at least the permanent molecular dipole moment, and the distributed atomic polarizabilities the full molecular polarizability tensor. The atomic charges can straightforward be obtained using e.g. Multipole Derived Charges (MDC) [See section on MDC] and the distributed polarizabilities by adopting standard parameters or refitting them to match the calculated polarizability tensor [146,147]. This allows for a simple procedure to obtain the solvent model parameters which subsequently can be used in the DRF calculation.

DIMQM block key

```
DIMQM
  <DRF|PIM|CPIM>
  NOPOL
  NOCHAR
  NOCROSS
  DDA
  FULLGRID
  LOCALFIELD
  EFIELD x y z
  SCREEN <ERF|EXP|ESP|NONE> {length}
  :: FREQUENCY-DEPENDENT PARAMETERS
  FREQUENCY
  OM_H value
  OM_C value
  OM_O value
  OM_N value
  :: CONTROL OVER SOLVER
  ALGORITHM <BEST|DIRECT|BRUTE|SINGLE|MULTI>
  TOLERANCE tol
  NITER iterations
  VOLUME vol_in_nm^3
  MULTIPLIER aterm bterm cterm
  GRID <COARSE|MEDUIM|FINE>
  :: GRADIENT OPTIONS
  FORCEFIELD
  CHEMBOND qmindex dimindex
  COORDDEPEND
  CHEMISORPTION
  COORDPAR atomtype e0 e1 r0 r1 CNmax Rmax Rmin
  CHEMPAR atomtype e0 e1 r0 r1 cutoff
  PROJECTIONMATRIXPOINTS <ALL|CUTOFF radius|OFF>
  :: OUTPUT CONTROL
  PRINTATOMICDIPOLES
  PRINTLJPAR
  DEBUG
END
```

`<DRF|CPIM|PIM>`

DIM/QM relies on one of three descriptions of the system: Discrete Reaction Field (DRF), where the atoms interact via induced dipoles and static charge, Capacitance Polarizability Interaction Model (CPIM), where the atoms interact via induced dipoles and induced charges, and Polarizability Interaction Model (PIM), where the atoms interact via induced dipoles only. DRF is best for solvents,

CPIM is best for small metal nanoparticles, and PIM is best for large metal nanoparticles. One and only one of these three keys must be included in every DIM/QM calculation.

NOPOL

The NOPOL key turns off the polarization terms, and thus all induced dipoles are zero. This key is only valid for DRF or CPIM calculations.

NOCHAR

The NOCHAR key turns off the charge terms, and thus all induced or static charges are zero. This key is only valid for DRF or CPIM calculations.

NOCHOSS

The NOCROSS key turns off the charge-dipole interactions. This key is only valid for CPIM calculations.

DDA

By default, the dipole-dipole, charge-dipole and charge-charge interactions are screened to take into account that atoms are not point charges. The DDA key will turn off this screening so that the results can be compared directly to the discrete dipole approximation (DDA).

FULLGRID

This is used in conjunction with the frozen density approximation.

LOCALFIELD

When the molecule interacts with a (for example) metal nanoparticle, there are two types of interactions: the image field and the local field. The image field is caused by the dipoles induced into the nanoparticle by the molecule's electron density. This is always taken into account in a DIM/QM calculation. The local field arises by direct interactions of the nanoparticle with an external field. Addition of the LOCALFIELD key causes the DIM/QM calculation to include this effect, but by default this is not included in a DIM/QM calculation.

EFIELD x y z

The EFIELD key is used to include an external static electric field in the vector x y z. Internally, the static charges used in a DRF calculation use ADF's EFIELD block, and therefore use of the EFIELD block is not allowed with the DIMQM block. This key is included so the user may include an electric field that would normally be included by the EFIELD block.

<SCREEN ERF|EXP|ESP|NONE> {length}

The SCREEN key indicates what functional form is used to screen the interactions between each DIM atom and the QM density. The choices are ERF (error function), EXP (exponential), ESP (error function for potential operator only), or NONE. For CPIM and PIM the default is EXP; for DRF, the default is ESP. In all cases, the default screening length is 1.0, but this may be changed with the optional length parameter.

FREQUENCY

The FREQUENCY key turns on frequency-dependent parameters.

OM_[HCON] value

The OM_H, OM_C, OM_O, and OM_N keys provide the resonance frequency (in atomic units) for the elements H, C, O and N, respectively. These keys are only for use with DRF and only when the older EXTERNALS block is used.

```
<ALGORITHM BEST|DIRECT|BRUTE|SINGLE|MULTI>
```

DIM/QM can choose between several solver algorithms. The DIRECT method solves the linear system of equations directly with a LAPACK routine; this should be considered the most robust method, but scales poorly with the number of atoms ($O(N^3)$ where N is the number of atoms in the system). The other three methods use an iterative technique, The BRUTE method (brute force) takes into account all atoms in the matrix-vector multiply step, and scales as $O(N^2)$. The SINGLE method uses the single-level cell-multipole-method (CMM), wherein dipoles that are spatially similar are collected into a multipole moment which effectively reduces the system size. This also scales as $O(N^2)$ but with a lower prefactor than BRUTE. The MULTI method uses the multi-level cell-multipole-method, which uses larger and large multipole the farther apart the dipoles are. This is the fastest method and scales as O(N log N).

Due to technical limitations, CPIM can only use DIRECT. Further, depending on the system size DIRECT or SINGLE may be more efficient than MULTI. To simplify choosing the solving algorithm, there is a BEST option that chooses the best algorithm for the particular system. BEST is the default option for algorithm.

```
TOLERANCE tol
```

The TOLERANCE key allows the user to specify a tolerance for the iterative solver. By default the tolerance is based on ADF's INTEGRATION key. This has no effect with the DIRECT solver.

```
NITER iterations
```

The NITER key allows the user to specify the maximum number of iterations for the iterative solver. By default this is MAX(N/100, 200) where N is the number of DIM atoms.

```
VOLUME vol_in_nm3
```

The VOLUME key is used to specify the DIM system volume. The volume is used to determine how to partition the system for the cell multipole method (ALGORITHM options SINGLE or MULTI), and is also use to determine the scattering efficiencies for frequency-dependent polarizability calculations. The volume does not need to be supplied; if it is missing, it will be calculated based on each atom's radius and the MULTIPLIER key.

```
MULTIPLIER aterm bterm cterm
```

An efficient way to get a approximation for the volume of the system is to sum the volume of each atom in the system, modified to account for the space between the atoms. This is done by modifying the atomic radius by a formula that takes into account the number of DIM atoms so that the effective radius changes with surface-to-bulk ratio. This formula is given by

$$r_{eff} = -ar/N^b + c$$

where r is the atomic radius, r_{eff} is the effective radius, N is the number of atoms, and the a, b, and c terms are the three parameters defined by the MULTIPLIER key. If the MULTIPLIER key is missing, the default values are 0.7, 0.5, and 1.13, respectively.

```
GRID <COARSE|MEDIUM|FINE>
```

In the cell multipole method (ALGORITHM options SINGLE or MUTLI), a certain number of the closest atom interactions must be calculated explicitly. The GRID key controls how many atoms must be

calculated this way, with COARSE being the least and FINE being the most. COARSE will be the fastest to calculate but may be numerically unstable. FINE is slowest to calculate but is the most stable. If the GRID key is missing, the default is MEDIUM.

`FORCEFIELD`

The FORCEFIELD key indicates that the DIM/QM calculation will include the DIM/QM force field. Currently the only maintained potential is the Lennard-Jones 12-6 potential (see Ref. [362]). This key is required to perform a DIM/QM geometry optimization and vibrational frequencies. By default, the DIM/QM force field is not included into the calculation.

Currently, DIM/QM geometry optimizations must be done in Cartesian coordinates which is specified in the GEOMETRY block. The user should be aware that ADF's default convergence criterion for a geometry optimization are relatively low, thus it is strongly suggested for a DIM/QM calculation to set the numerical integration quality (BeckeGrid) to good and change the convergence criterion of the max gradient to 1E-4.

`COORDDEPEND`

The COORDDEPEND key indicates that the DIM/QM force field will be coordination dependent. This only effects the Lennard-Jones parameters for DIM atoms (see Ref. [362]). By default, the DIM/QM force field is not coordination dependent.

`CHEMISORPTION`

The CHEMISORPTION key will include chemisorption corrections for all atoms that have chemisorption parameters within a given cutoff radius. By default, the DIM/QM force field does not included chemisorption corrections.

`COORDPAR atomtype e0 e1 r0 r1 CNmax Rmax Rmin`

The COORDPAR key allows the user to add additional coordination dependent parameter for a selected element type. atomtype specifies the element type (i.e., Ag for silver) for the given parameter set. e0, e1, r0, and r1 are the coordination dependent Lennard-Jones parameters; see Ref. [362]) for more details. The coordination numbers of the DIM atoms are computed as an effective coordination number. This scheme requires a maximum and minimum cutoff distances, Rmax and Rmin respectively, and a maximum coordination number, CNmax. All parameters need to be in atomic units.

`CHEMPAR atomtype e0 e1 r0 r1 CUTOFF`

The CHEMPAR key allows the user to add additional chemisorption dependent parameter for a selected element type. atomtype specifies the element type (i.e., N for nitrogen) for the given parameter set. e0, e1, r0, and r1 are the coordination dependent Lennard-Jones parameters; see Ref. [362]) for more details. The code determines if there is a chemical bond by a cutoff distance parameter, CUTOFF. If the QM-DIM bond is within the cutoff, the code uses the chemisorption parameter; otherwise, the code uses the standard parameter set. All parameters need to be in atomic units.

`CHEMBOND qmindex dimindex`

The CHEMBOND key indicates that there will a chemisorption correction used for the bond between the specified QM and DIM atoms. The user may repeat the CHEMBOD key up to 50 times to specify up to 50 different chemical bonds for the force field. The qmindex is an integer based on the order of atoms in the ATOMS block; i.e. the fifth QM atom in the ATOMS block would have qmindex = 5. The dimindex is the same but corresponds to the DIM atom involved in the bond. This key should only be used when the CHEMISORPTION key is also specified. When using CHEMBOND, the cutoff distance parameter for chemisorption correction parameter key will be ignored. It is suggested to use CHEMBOND if the user is generating a potential energy surface with a chemisorbed QM system.

`PROJECTIONMATRIXPOINTS <ALL|CUTOFF> <radius|OFF>`

The PROJECTIONMATRIXPOINTS key specifies what DIM atoms to include for the projection matrix when removing rigid motions out of the gradient. The methods available are ALL, CUTOFF, or OFF. The ALL option causes PROJECTIONMATRIXPOINTS to include all DIM atoms. OFF will turn off the removal of rigid motions. CUTOFF includes any DIM atom points within a cutoff radius from the center of mass of the QM system to the DIM atom points and requires a cutoff radius to be given in Angstrom. This key only applies to a geometry optimization. If the PROJECTIONMATRIXPOINTS key is not given, the option CUTOFF with a cutoff radius of 25.4 Angstrom is assumed.

`PRINTATOMICDIPOLES`

The PRINTATOMICDIPOLES key causes all the induced dipole moments of each DIM atom to be printed at the conclusion of each SCF cycle and each RESPONSE or AORESPONSE polarizability calculation. Because DIM/QM is typically used with many thousands of atoms, this can result in a large output file, but they may be useful for debugging purposes or to calculate electric fields. By default these are not printed.

`PRINTLJPAR`

The PRINTLJPAR key specifies that all Lennard-Jones parameters used for the calculation will be printed in the output file. The QM atoms' Lennard-Jones parameters are also printed with the DEBUG key.

`DEBUG`

The DEBUG key will print out extra information in the process of the calculation.

EXTERNALS block key

The `EXTERNALS` block key controls the input data for the MM atoms. The EXTERNALS block is designed for DRF calculations. For each MM atom the following data are required:

```
EXTERNALS
 atm num grp-nam grp-num, char, x, y, x, pol
 ...
 GROUP
 {...}
end
```

`atm`

Type of atom, i.e., H, O, ...

`num`

number of atoms (optional)

`grp-nam`

Name of the group to which the atom belongs

`grp-num`

Number of the group to which the atom belong

`char`

atomic charge (in atomic units)

`x`

x-coordinate

`y`

y-coordinate

`z`

z-coordinate

`pol`

atomic polarizability (in atomic units)

`GROUP`

Indicates the end of group

The separation of molecules into `GROUP`'s are important. Since in the many-body polarization operator only inter-molecular interactions, i.e. only interaction between sites which do not belong to the some group, are included. Therefore, it is important that the combined string (`grp-nam` + `grp-num`) is unique for each `GROUP`.

An example of a `EXTERNALS` block for two water molecules:

```
EXTERNALS
 O 4 water 2, -0.6690, -11.380487, -11.810553, -4.515226, 9.3005
 H 5 water 2,  0.3345, -13.104751, -11.837669, -3.969549, 0.0690
 H 6 water 2,  0.3345, -10.510898, -12.853311, -3.320199, 0.0690
 GROUP
 O 7 water 3, -0.6690,  -1.116350,   9.119186, -3.230948, 9.3005
 H 8 water 3,  0.3345,  -2.822714,   9.717033, -3.180632, 0.0690
 H 9 water 3,  0.3345,  -0.123788,  10.538199, -2.708607, 0.0690
 GROUP
 {...}
end
```

DIMPAR block key

In this block, the parameters for the DIM atoms are defined.

```
DIMPAR
  Element
    RAD val
    POL val
    CAP val
    CHAR val
    OM val
    OM1 val
    OM2 val
    GM1 val
    GM2 val
    SIZE val
    BOUND val
    EXP /path/to/experimental/dielectric/file
    DRUDE plasma damping {EV}
    FERMI val
    <LRTZ|LRTZ1> osc res damp {EV}
```

```
    LRTZ2 osc pls res damp {EV}
    LRTZ3 pls res damp {EV}
  SUBEND
  XYZ
    {/absolute/path/to/coordinates.xyz}
    {natoms
     elem x.xxx y.yyy z.zzz
     elem x.xxx y.yyy z.zzz
     elem x.xxx y.yyy z.zzz
     ...}
  SUBEND
END
```

Element

Within the DIMPAR block, you will need a sub-block that defines the parameters for each element that is in your DIM system. You will need to replace 'Element' with the element you are assigning parameters to, as in

```
  Ag
    ...
  SUBEND
```

if you are assigning parameters to Ag. Note that the first letter MUST be capitalized and the second MUST be lowercase.

RAD val

RAD specifies the atomic radius in the unit defined by the input file. RAD is required for all PIM calculations, all calculations with ALOGORITHM options SINGLE or DIRECT, and all frequency-dependent calculations where the AORESPONSE LIFETIME key is given.

POL val

POL specifies the polarizability parameter (in a.u.) used in DRF or CPIM.

CAP val

CAP specifies the capacitance parameter (in a.u.) used in CPIM.

CHAR val

CHAR specifies the atomic charge (in a.u) used in DRF.

OM val

OM specifies the resonance frequency (in a.u) used in DRF. This replaces the OM_[HCON] key in the DIMQM block.

OM1 val

OM1 specifies the ω_1 parameter (in a.u) used in CPIM.

OM2 val

OM2 specifies the ω_2 parameter (in a.u) used in CPIM.

GM1 val

GM1 specifies the γ_1 parameter (in a.u) used in CPIM.

```
GM2 val
```

GM2 specifies the γ_2 parameter (in a.u) used in CPIM.

```
SIZE val
```

SIZE specifies the size-dependent parameter used in CPIM.

```
EXP /absolute/path/to/experimental/dielectric/file
```

In PIM, the atomic polarizabilities are calculated from the dielectric constant. If you have access to the experimental dielectric constant, this may be supplied directly to DIM/QM. The values will be splined, so it is not necessary to ensure that each frequency at which you are calculating be in the file. DIM/QM expects the file to be formatted with the wavelength (in nm) in the first column, the real part of the dielectric in the second column, and the imaginary part of the dielectric in the third column. All other columns that may exist will be ignored, as well as lines beginning with the hash (#) symbol.

```
BOUND val
```

A Drude function is typically written as

$\varepsilon_\infty - \omega_p^2/(\omega(\omega+i\gamma))$

with the second term being the Drude function, and the first term accounting for bound electrons. For a conductor with no bound electrons, $\varepsilon_\infty = 1$ which is the default value for BOUND. To account for bound electrons you may set BOUND to a value greater than 1. This key only affects PIM.

```
DRUDE plasma damping {EV}
```

The formula for a Drude function is

$\varepsilon_\infty - \omega_p^2/(\omega(\omega+i\gamma))$

where ε_∞represents the bound electrons (as discussed for BOUND), ω_p (plasma) is the plasma frequency, and γ (damping) is the damping parameter (decay rate). Optionally, EV may be added to specify the values be read in units of electron volts, otherwise they are read in units of a.u. This key only affects PIM.

```
FERMI val
```

The FERMI key is used to specify a Fermi velocity (in m/s) so that the Drude function may be size-corrected using a modified Drude function:

$\varepsilon_\infty - \omega_p^2/(\omega(\omega+i(\gamma+v_{fermi}/R_{eff})))$

where $+v_{fermi}$ is the Fermi velocity and R_{eff} is the effective nanoparticle radius. This can also be used in conjunction with EXP and DRUDE to size-correct experimental dielectric parameters. This key only affects PIM.

```
<LRTZ|LRTZ1> osc res damp {EV}
LRTZ2 osc pls res damp {EV}
LRTZ3 pls res damp {EV}
```

There are three forms of the Lorentzian function seen in the literature:

$$\sum_n f_n \Omega_{0,n}^2/(\Omega_{0,n}^2 - \omega^2 - i \Gamma_n \omega)$$

$$\sum_n f_n \omega_p^2/(\Omega_{0,n}^2 - \omega^2 - i \Gamma_n \omega)$$

$$\sum_n \Omega_{p,n}^2/(\Omega_{0,n}^2 - \omega^2 - i \Gamma_n \omega)$$

where $\Omega_{0,n}$ (res) is a bound electron resonance frequency, f_n (osc) is a bound electron oscillator strength, Γ_n (damp) is a bound electron excited state decay rate (or damping parameter), ω_p (pls) is the free electron plasma frequency, and $\Omega_{p,n}$ (pls) is the bound electron plasma frequency. You may choose the Lorentzian for against which your parameters were parametrized. The top form is LRTZ1, the middle is LRTZ2, and the bottom is LRTZ3. Because LRTZ1 is the most common, it is also aliased as LRTZ. Optionally, EV may be added to specify the values be read in units of electron volts, otherwise they are read in units of a.u. This key only affects PIM. You may give any of the form of the LRTZ key up to 50 times supply up to 50 Lortenzian functions to make a Drude-Lorentz function.

`XYZ`

The XYZ sub-block is where the DIM atom coordinates are given. Two methods of supplying coordinates are allowed.

`In-File Coordinates`

As an example of how to supply coordinates in-file, imagine you wish to calculate a Au dimer system on the Z-axis. You might define your coordinates as:

```
XYZ
 2
 Au 0.0 0.0 0.0
 Au 0.0 0.0 3.0
SUBEND
```

The first line gives the number of atoms to follow. Every line after that contains the element in the first column (first letter MUST be capitalized, second MUST be lowercase), then the x-component, then the y-component, then the z-component. You may not number the atoms.

`External File Coordinates`

When the DIM system size becomes large, it is often more convenient to keep the DIM coordinates in a separate file. The XYZ block would then look like:

```
XYZ
 /absolute/path/to/coordinates.xyz
SUBEND
```

Note that you MUST include the absolute path to your file.

The .xyz file is set up identically to the in-file table, except that there is a space between the number of atoms and the first coordinate in case a comment need be added. The .xyz file for our dimer system would be:

```
2
A gold dimer (this line will be ignored)
```

```
Au 0.0 0.0 0.0
Au 0.0 0.0 3.0
```

FDE: Frozen Density Embedding

The Frozen-Density-Embedding (FDE) option invokes calculation of the effective embedding potential introduced by Wesolowski and Warshel [184] in order to take into account the effect of the environment on the electronic structure of an embedded system. The embedding potential (Eq. 3 in Ref. [240]) depends explicitly on electron densities corresponding to the embedded subsystem (e.g. a solvated molecule) and its environment (e.g. solvent). For a detailed review, see Ref. [205]. The ADF implementation of the method is described in detail in Ref. [185,217]. The implementation of FDE in ADF2007 has been completely revised and improved. Therefore, the input format has been changed with respect to ADF2006.

A time-dependent linear-response generalization of this embedding scheme was derived in Ref. [186]. Its implementation in an approximate form, which assumes a localized response of the embedded system only (uncoupled FDE), is described in the supplementary material to Ref. [187]. For possible drawbacks and pitfalls in connection with this approximation, see Refs. [185,190,193].

The theory of coupled excited states for subsystems is described in Refs. [296,297], and extended for general response properties in Ref. [298]. This theory (subsystem TDDFT, coupled FDE) allows to treat the mutual response of several subsystems, including the ones that are considered environment.

A generalization of the FDE scheme to the calculation of NMR shieldings has been given in Ref. [218], where also the approximations involved and possible problems are discussed.

With the exception of interaction energies, the current implementation in ADF only allows the calculation of molecular properties that only depend on the electron density and of response properties using TDDFT. For an application to the calculation of several molecular properties in solution and a comparison to the DRF model also available in ADF, see Ref. [190]. For further applications of the ADF implementation, see Ref. [189] (weakly interacting complexes), Refs. [185,190-192] (solvent effects), and Refs. [206-207] (other environment effects).

FDE Input

To invoke a frozen-density embedding calculation, two additional specifications in the input are required. First, one or more frozen fragments have to be included in the FRAGMENTS block, and second, the block key FDE has to be included. In the simplest case, this input should look like this:

```
FRAGMENTS
  ...
  FragType FragFile type=FDE
  ...
END

FDE
  PW91K
end
```

In the FRAGMENTS block, for any fragment it is possible to specify the option type=FDE to indicate that the density of this fragment is kept frozen. This density is imported from the file FragFile. The frozen fragments have to be included in addition to the usual, nonfrozen fragments. The atoms of the frozen fragments have to be included in the ATOMS block. As with normal fragments, the fragment found in the file will be rotated and translated to its position specified in the ATOMS block. For more details on specifying fragments, see the section 'fragment files'. In the FDE input block, the recommended PW91k (also known as GGA97) approximant is recommended for the non-additive kinetic energy (the default is the local density approximant). A recommended alternative is NDSD. For all other options the defaults will be used.

Please note that throughout the FDE part of the documentation, the word "approximant" is used instead of the more usual "functional" to emphasize that the exact functional is not known, also in the case of the kinetic energy functional. In the literature one may encounter both words used interchangeably.

By including more than one frozen fragment, it is possible to use a frozen fragment that is a superposition of the densities of isolated molecules (this was possible in the previous version of ADF using the DENSPREP option). For a discussion and tests of the use of such approximate environment densities, see Ref. [185].

There is no restriction on the use of symmetry in FDE calculations, and usually the correct symmetry will be detected automatically. However, in the preparation of frozen fragments that will be rotated and/or translated in the FDE calculation, one has to include the keyword NOSYMFIT for technical reasons.

In the current implementation, only the electron density of the embedded (nonfrozen) system is calculated. Therefore, with the exception of interaction energies, only properties that depend directly on the electron density (e.g. dipole moments) are available. In particular, the calculation of energy gradients is not implemented yet. All quantities given in the output refer (unless explicitly specified otherwise) to the nonfrozen system only.

The TDDFT extension of the FDE formalism allows the calculation of electronic excitation energies and polarizabilities. This extension is automatically activated if FDE is used in combination with the EXCITATIONS or the RESPONSE key. To allow the mutual response of several subsystems, see the secion on [subsystem TDDFT].

To employ the extension of FDE to the calculation of NMR shieldings, the file TAPE10 has to be used in the FDE calculation (by including the option SAVE TAPE10), and subsequently the NMR shielding has to be calculated using the program NMR (not with EPR).

Fragment-specific FDE options

For each frozen fragment, several additional options can be applied. To do this, the fragment specification is used as a subblock key by appending a & sign. The subblock is terminated with SubEnd. This subblock key looks, in the most general form, as follows:

```
FRAGMENTS
  ...
  FragType FragFile type=FDE &
     {FDEOPTIONS [USEBASIS] [RELAX or FREEZEANDTHAW]}
     {FDEDENSTYPE [SCF | SCFexact | SCFfitted ]}
     {RELAXCYCLES n or FREEZEANDTHAWCYCLES n}
     {XC [LDA | GGA ggapotx ggapotc | MODEL SAOP]}
  SubEnd
  ...
END
```

FDEOPTIONS

FDEOPTIONS USEBASIS

If the USEBASIS option is specified, the basis functions of this frozen fragment will be included in the calculation of the embedded subsystem. This allows to expand the density of the embedded subsystem using not only atom-centered basis sets localized in the embedded subsystem but also the ones in the environment Ref. [238]. In large-scale simulations using the embedding potential, this option is recommended to be used in the preparation stage to investigate the basis set dependence of the results (chapter 5.3 in Ref. [205]). This option is also an indispensable element in the procedure introduced in Ref. [238] to test approximants to the kinetic-energy component of the embedding potential introduced by Wesolowski and Warshel.

FDEOPTIONS RELAX or FREEZEANDTHAW

If the RELAX option (or equivalent FREEZEANDTHAW option) is specified, the density of this frozen fragment will be relaxed in freeze-and-thaw cycles [Ref. 240], i.e., the embedded subsystem is frozen, while this fragment is thawed. This is repeated, until convergence is reached or until the maximum number of iterations has been performed. By relaxing frozen fragments, it is possible to improve a given approximate environment density by including the polarization of the environment due to the embedded system.
This option is recommended to be used in the preparation stage of a large-scale numerical simulation. The freeze-and-thaw calculations lead to a pair of electron densities (embedded system and environment) that minimizes the total energy. As a consequence, the electron density of the environment derived from the freeze-and-thaw calculations can be used as a reference to verify the adequacy of the assumed electron density for the environment in a large-scale simulation. Due to technical restrictions, freeze-and-thaw is not possible if an open-shell (unrestricted) fragment is present.

```
FDEOPTIONS USEBASIS RELAX or FDEOPTIONS USEBASIS FREEZEANDTHAW
```

It is further possible to combine USEBASIS and RELAX or FREEZEANDTHAW. In this case, the basis functions of the nonfrozen fragment will be included when the density of the fragment is relaxed. This allows fully relaxed calculations with supermolecular expansion of the electron density of each subsystem. This option is to be used to test approximants to the kinetic-energy component of the embedding potential introduced by Wesolowski and Warshel by means of the procedure introduced in [Ref. 238].

```
FDEDENSTYPE
```

The FDEDENSTYPE option can be used to specify which density is read from the fragment file. The possible options are:

```
FDEDENSTYPE SCF (or FDEDENSTYPE SCFexact)
```

The exact density (not calculated using the fit functions) is used. This is the default.

```
FDEDENSTYPE SCFfitted
```

The fitted density is used. This is less accurate but can be significantly faster.

```
RELAXCYCLES n or FREEZEANDTHAWCYCLES n
```

This gives the maximum number of freeze-and-thaw cycles that are performed for this fragment. If the maximum number given in the FDE block is smaller, or if convergence is reached earlier, then fewer cycles are performed. For historical reasons, two equivalent keywords are available.

```
XC
```

The XC option can be used to select the exchange-correlation potential that is used for this fragment when it is relaxed. By default, the same potential as for the nonfrozen system is used, but in some cases it might be preferable to use another approximation for certain fragments. An example is given in Ref. [189].

```
XC LDA
```

This option selects LDA as exchange-correlation potential for relaxing this fragment.

```
XC GGA ggapotx ggapotc
```

This selects a GGA potential for relaxing this fragment. The GGA potential is specified by giving the name of the exchange potential, followed by the name of the correlation potential. The available potentials are listed in the documentation for the XC key.

`XC MODEL SAOP`

This selects the model potential SAOP for relaxing this fragment.

Kinetic energy approximants

The approximants to the kinetic energy dependent component of the embedding potential are described here.

```
FDE
  {approximants to the kinetic energy dependent
                component of the embedding potential}
  {CJCORR [rho_cutoff]}
  {GGAPOTXFD exchange approximant}
  {GGAPOTCFD correlation approximant}
end
```

`approximants to the kinetic energy dependent component of the embedding potential`

Several approximants to the kinetic-energy-dependent component of the effective potential given in Eq. (21) of [Ref. 184] are available. None of them is applicable if the embedded system is covalently bound to its environment. The user is recommended to look at the numerical value of the TSNAD(LDA) parameter which is given in the units of energy and can be considered as a measure of the overlap. The following rule of thumb should be applied: if this parameter is smaller than the estimated interaction energy between the embedded subsystem and the environment, then the available approximants are most probably adequate. If it exceeds this limit, the results can be less reliable. Printing TSNAD(LDA) is not done by default, as it can be quite time-consuming. Its printing is switched on by including "EXTPRINTENERGY", and "PRINTRHO2", and "FULLGRID" in the FDE input block.
If no kinetic energy approximant is specified, by default the local-density approximation (Thomas-Fermi approximant) is used. For an assessment of approximants for weakly overlapping pairs of densities see Refs. [238, 239, 188, 241]. Based on these studies, the use of PW91k (= GGA97) is recommended.

APPROXIMANTS TO BE USED IN NORMAL APPLICATIONS

`THOMASFERMI (default)`

Local-density-approximation form of vt[rhoA,rhoB] [237] derived from Thomas-Fermi expression for Ts[rho] [194, 195].

`GGA97 (or PW91K)`

Generalized-gradient-approximation form of vt[rhoA,rhoB] [239] derived from the Lembarki-Chermette [197] approximant to Ts[rho]. This approximant is currently the recommended one based on the numerical analysis of its accuracy [188,239] and the fact that the used enhancement factor disappears at large reduced density gradients, i.e. where the second-order gradient-expansion approximation fails [238, 241].

`NDSD`

Similarly to GGA97, the NDSD approximant is constructed by taking into account the asymptotic behavior of the functional vt[rhoA,rhoB] at small density gradients. In the construction of NDSD, the exact property of vt[rhoA,rhoB] at rho_A → 0$ and for ∫ rhoB = 2 given in Eq. A6 of Ref. [279] is also taken into account. The analysis of the accuracy of this potential [279] shows that NDSD is of the same or superior quality as GGA97. NDSD is, therefore, recommended as the successor of GGA97 to be used anywhere where the quality of the results depends directly on the accuracy of the potential vt[rhoA,rhoB], i.e., for obtaining electronic-structure-dependent properties. The analytical form of the corresponding approximant to the functional Ts^{nad} [\rho_A,\rho_B]$ exists

(Eq. 23 in Ref. [279]). It is not possible, however, to obtain the analytical form of the corresponding parent functional for the kinetic energy Ts[rho]. To reflect this and the fact that, similarly to the GGA approximants to vt[rhoA,rhoB], the numerical values of only first- and second derivatives of density are needed, the label NDSD (Non-Decomposable Second Derivatives) is used.

OBSOLETE APPROXIMANTS (can be used but GGA97 leads usually to a better embedding potential see [238,239])

`LLP91`

Generalized-gradient-approximation form of vt[rhoA,rhoB] [238] derived from Lee-Lee-Parr [Ref. 198] approximant to Ts[rho].

`PW86k`

Generalized-gradient-approximation form of vt[rhoA,rhoB] [238] derived from the Fuentealba-Reyes approximant to Ts[rho] [242].

`THAKKAR92`

Generalized-gradient-approximation form of vt[rhoA,rhoB] [239] derived from the Thakkar approximant to Ts[rho] [201].

APPROXIMANTS WHICH MIGHT BE USEFUL ONLY FOR THEORY DEVELOPMENT

The accuracy of **some** of these approximants was investigated in detail [239, 238, 188, 241]. Each of them was shown to lead to a qualitatively incorrect embedding potential. They shouldn't be used in practical applications.

`COULOMB`

Neglecting completely vt[rhoA,rhoB] (vt[rhoA,rhoB] equals zero) together with the exchange-correlation component of the embedding potential introduced by Wesolowski and Warshel.

`TF9W`

The approximant to vt[rhoA,rhoB] [184] derived from the second-order gradient expansion [242]] for Ts[rho].

`WEIZ`

The approximant to vt[rhoA,rhoB] [241] derived from the von Weizsäcker approximant to Ts[rho] [186].

`OL91A`

Generalized-gradient-approximation form of vt[rhoA,rhoB] [238] derived from the first Ou-Yang and Levy approximant to Ts[rho] [200].

`OL91B`

Generalized-gradient-approximation form of vt[rhoA,rhoB] [239] derived from the second Ou-Yang and Levy approximant to Ts[rho] [200].

`E00`

Generalized-gradient-approximation form of vt[rhoA,rhoB] [263] derived from a kinetic energy functional by Ernzerhof [264] which represents the gradient expansion approximation up to the fourth order.

P92

Generalized-gradient-approximation form of vt[rhoA,rhoB] [263] derived from a kinetic energy functional by Perdew [265] which represents the gradient expansion approximation up to the sixth order.

LONG DISTANCE CORRECTIONS TO THE EFFECTIVE POTENTIAL

CJCORR

Option to switch on a long-distance correction. By default this option is not used. As was shown in Ref. [220], with the available approximate kinetic-energy approximants, the embedding potential has the wrong form in the limit of a large separation of the subsystems. In particular, it was shown that this can have serious consequences in the case of "supermolecular expansion of electron density of each subsystem" calculations (USEBASIS option). In Ref. [220], a correction is proposed that enforces the correct long-distance limit. (See also this reference for limitations of this correction.)

CJCORR [rho_cutoff]

This option switches on the long-distance correction. This option has to be used in combination with one of the above kinetic-energy approximants. By default, a density cut-off of 0.1 is employed.

NONADDITIVE EXCHANGE-CORRELATION APPROXIMANT

GGAPOTXFD
GGAPOTCFD

Option to specify the nonadditive exchange-correlation approximant. By default, in the construction of the effective embedding potential the exchange-correlation approximant that was specified in the XC block is used. It is possible to specify a different approximant with the GGAPOTXFD and GGAPOTCFD options. This is particularly useful in combination with the use of model potentials like SAOP, that can not be used in the embedding potential because of their orbital dependence. (For a discussion, see Ref. [189].)

GGAPOTXFD exchange approximant

The exchange approximant is used in the construction of the embedding potential. The same exchange approximants as in the XC key are available.

GGAPOTCFD correlation approximant

The correlation approximant is used in the construction of the embedding potential. The same correlation approximants as in the XC key are available.

General FDE options

In addition to the fragment-specific options and the kinetic energy approximants, there are also a number of options available in FDE calculations that will be described in the following.

```
FDE
  {FULLGRID}
  {RELAXCYCLES n or FREEZEANDTHAWCYCLES n}
  {RELAXPOSTSCF or FREEZEANDTHAWPOSTSCF}
  {EXTPRINTENERGY}
  {PRINTRHO2}
  {ENERGY}
end
```

FULLGRID

By default, FULLGRID is not used, and in FDE calculations the integration grid is generated as described in Ref. [185] by including only atoms of the frozen subsystem that are close to the embedded subsystem in the generation of the integration grid. The distance cutoff used is chosen automatically, based on the extent of the basis functions of the embedded subsystem. (It can also be chosen manually, see the option qpnear in the INTEGRATION key) This scheme results in an efficient and accurate integration grid. However, it is possible that the default integration scheme is not accurate enough. This can be the case for weakly interacting systems and when the distance between the frozen and the embedded system is large. It is therefore recommended to check the quality of the default integration grid by comparing to results obtained using the full supermolecular grid (FULLGRID option).

If the subkey FULLGRID is included, all atoms of the frozen system are included in the generation of the integration grid. This results in the same grid that would be used in a supermolecular calculation of the combined frozen and embedded system. The integration grid generated by this option might be much larger than the default grid. This option should be used to check the quality of the default integration grid.

`RELAXCYCLES n or FREEZEANDTHAWCYCLES n`

Specifies the maximum number n of freeze-and-thaw iterations [Ref. 240] that are performed (for frozen fragments with the RELAX) option. If a smaller number of iterations is specified as a fragment-specific option, for this fragment this smaller number is used. Furthermore, if convergence is reached earlier, no more iterations will be performed.

`RELAXPOSTSCF or FREEZEANDTHAWPOSTSCF`

If this option is included, several post-SCF properties will be calculated after each freeze-and-thaw cycle [Ref. 240]. These are otherwise only calculated in the last cycle.

```
EXTPRINTENERGY
PRINTRHO2
```

If the options EXTPRINTENERGY and PRINTRHO2 are included (both are needed and should be listed on separate lines), several additional quantities will be printed, including TSNAD(LDA). In order to obtain meaningful numbers, also the FULLGRID keyword (see above) has to be used.

`ENERGY`

Option to switch on the calculation of the FDE energy as the sum of the energy E[rhoA] of the active, embedded system and the interaction energy Eint[rhoA,rhoB] of the embedded system with the frozen environment. This relies on the calculation of the total energy for the embedded system and all caveats and restrictions for total energy evaluations apply (see keyword TOTALENERGY). All energy contributions are evaluated on the grid of the active subsystem. Some contributions to the interaction energy Eint[rhoA,rhoB] require an accurate integration grid in the region of the environment. Thus, in pure embedding calculations (without fragment-specific option RELAX), an accurate calculation of the FDE energy requires a full supermolecular integration grid (FULLGRID option). Details on the implementation and the performance of kinetic energy functionals for interaction energies are documented in Ref. [263]

The calculation of the full, variationally minimized subsystem DFT energy, that is, the sum of the energy of two subsystems E[rhoA] and E[rhoB] and their interaction energy Eint[rhoA,rhoB] in the framework of FDE, is invoked if then the fragment densities are relaxed in freeze-and-thaw cycles (option RELAXCYCLES and fragment-specific FDE option RELAX). In this case the supermolecular integration grid is not required. Instead, in each step of the freeze-and-thaw cycle, the critical energy terms are taken from the previous freeze-and-thaw step of the presently frozen fragment. The convergence of the energy contributions with the number of freeze-and-thaw iterations should be carefully monitored. Due to conceptual problems for the evaluation of the non-additive kinetic energy contribution, only two subsystems, that is, one frozen fragment, is supported for FDE energy calculations with freeze-and-thaw.

Subsystem TDDFT, coupled FDE

The linear-response subsystem TDDFT code implements the theory of coupled excited states for subsystems as described in Refs. [296,297]. This theory is based on the FDE extension to excited states [186], which is implemented in ADF in a local response approximation, i.e., neglecting the dynamic response of the environment [187].

The subsystem TDDFT code allows to treat the mutual response of several subsystems, including the ones that are considered environment. A more typical situation would be a system composed of several equivalent chromophores treated as individual subsystems. In this case, the local response approximation leads to uncoupled excited states of the subsystems (hence the acronym FDEu is employed often), while the subsystem TDDFT code couples the monomer excitations to obtain the excited states of the total system (often denoted as coupled frozen density embedding, FDEc). This can be related to excitonic couplings between the monomers [297].

The current implementation is restricted to NOSYM calculations and Singlet-Singlet excitations without frozen core approximation. It makes use of the ALDA kernel (including a Thomas-Fermi part for the contribution arising from the non-additive kinetic energy) for consistency with the uncoupled FDE implementation for excited states. Some features have not or not extensively been tested and should be used with great care, e.g., linear dependencies in the basis set. Details on the calculation of transition moments, oscillator and rotational strengths are described in Ref. [298].

Subsystem TDDFT (FDEc) calculations can be invoked with the SUBEXCI key.

SUBEXCI input

FDEc calculations on coupled excited states first require that an uncoupled FDE-TDDFT calculation has been performed for every subsystem that should be included in the coupled calculation, and that the corresponding TAPE21 files, in which the considered subsystems are "active", have been saved (see the separate FDE input description). This means that it is not possible to use the information on frozen/inactive fragments from a TAPE21 file of a previous uncoupled FDE calculation, which contains all subsystems.

Although it is technically possible to use TAPE21 files from non-FDE calculations on the separate subsystems, this would lead to results that are inconsistent with the subsystem TDDFT methodology from Ref. [296]. In any case, a previous TDDFT calculation for each subsystem that should be included in the coupling procedure is necessary. If that is not the case, the subsystem will still be considered in the calculation of the total electron density (needed in the setup of the exchange-correlation kernel), but will not be included in the coupling procedure.

The first subsystem should always be one of the coupled subsystems. The input will then look like the corresponding input for an uncoupled FDE-TDDFT calculation, but in addition should contain the following block:

```
SUBEXCI
  {LOWEST nlowest}
  {OPTSTATES list_of_optstates}
  {CTHRES coupling_threshold}
  {SFTHRES solutionfactor_threshold}
  {COUPLBLOCK}
  {TDA}
  {CICOUPL}
END
```

`LOWEST nlowest`

The selection of the excited states to be coupled consists of two steps. First, a number of reference states are selected. As a default, the `nlowest` (default: 10) lowest excited states present on the fragment file for the first subsystem are considered.

`OPTSTATES list_of_optstates`

If the keyword OPTSTATES is given, only those excited states of the first subsystem are considered as reference states that are given in the `list_of_optstates` (numbers of states separated by blanks).

`CTHRES coupling_threshold`

Second, all excitations of all subsystems (present on the fragment TAPE21 files) with an excitation energy that differs by less than `coupling_threshold` (to be given in units of eV; default: 30 eV) from one of the reference states are selected to be included in the coupling. Note that additional excited states of system 1 may be included here.

`COUPLBLOCK`

If COUPLBLOCK is specified in the input, all couplings between all of these local excited states are included. Otherwise (default), the `coupling_threshold` is also applied to select pairs of states for which couplings are calculated. I.e., couplings are not calculated if the two particular states to be coupled differ in energy by more than `coupling_threshold`.

`SFTHRES solutionfactor_threshold`

To reduce the computational effort, it is possible to ignore the effect of orbital pairs with coefficients less than `solutionfactor_threshold` in the solution factors (TDDFT eigenvectors) of the underlying uncoupled calculation in the construction of the exact trial densities during the calculation of the coupling matrix elements. These orbital pair contributions are not ignored in the subsequent calculation of transition moments, oscillator, and rotational strengths. The default value of 0.00001 typically leads to a precision of the coupled excitation energies of about 0.0001 eV.

`TDA`

TDA specifies the use of the Tamm-Dancoff-Approximation (section on TDA) in the underlying uncoupled FDE-TDDFT calculations (Ref. [364]). Contrary to the full SUBEXCI-TDDFT variant, SUBEXCI-TDA allows for the usage of hybrid functionals in the underlying uncoupled FDE-TDDFT calculations.

`CICOUPL`

Within the Tamm-Dancoff Approximation, the couplings between localized excited states on different subsystems correspond directly to so-called exciton couplings (see Ref. [364]). The CICOUPL keyword, in conjunction with TDA, prints these exciton couplings. It is also possible to use CICOUPL with full FDEc-TDDFT. In that case, the excitonic couplings between monomers are reconstructed from an effective 2x2 CIS-like eigenvalue problem, as e.g. done in Ref. [297].

In addition, the input file may contain either an EXCITATION block or the keyword DIFFUSE. Both options lead to a slight adaption of the integration grid. Apart from this, the EXCITATION block will be ignored.

The key `ALLOW PARTIALSUPERFRAGS` is currently necessary to be able to use subsystem information for only one subsystem from a TAPE21 file of a previous FDE calculation:

```
ALLOW PARTIALSUPERFRAGS
```

Restrictions and pitfalls

In the current implementation, only the electron density of the embedded system is calculated. Therefore, with the exception of interaction energies, only properties that depend directly on the electron density (e.g., dipole moments) are available. In addition, the TDDFT extension allows the calculation of electronic excitation energies and polarizabilities, and NMR shieldings can be calculated.

EVERYTHING ELSE IS NOT YET IMPLEMENTED. THE RESULTS OBTAINED FOR OTHER PROPERTIES MIGHT BE MEANINGLESS.

In particular, energy gradients are not yet available.

Kinetic energy approximant:
Although the effective embedding potential is derived from first principles using universal density approximants, the ADF implementation relies on approximations. Currently, two implemented approximations are recommended [188]: PW91k (also known as GGA97) which uses electron densities and the corresponding gradients to express the nonadditive kinetic energy component of the embedding potential, or TF (Thomas-Fermi LDA approximant), which does not use gradients at all. Either approximation is applicable only in cases where the overlap between electron densities of the corresponding interactions is small. Note: so far, no approximation has been developed for the strong-overlap case - two subsystem linked by covalent bonds for instance.

SCRF: Self-Consistent Reaction Field

The following sections describe the SCRF method and explain the related ADF input options. SCRF may not be available in the standard distribution on all platforms, contact SCM (support@scm.com) to request SCRF on your platform.

Introduction

by Donald Bashford, St. Jude Children's Research Hospital Memphis, updated October 26, 2011.

SCRF (Self-Consistent Reaction Field) is a method of accounting for the effect of a polarizable solvent (and optionally, a classical macromolecular system) on the quantum system. Consider first the case with only a polarizable solvent. The solvent is modeled as a dielectric continuum with a dielectric constant, EPSSOL, that fills the space outside the quantum system. The boundary between the interior (where the dielectric constant is unity) and the higher-dielectric exterior is the molecular surface, as defined by Connolly [251]. The charges of the quantum system cause polarization of this continuum, giving rise to a reaction field which then acts back on the quantum system potentially altering its charge distribution. The SCRF algorithm calculates the reaction field through solutions to the Poisson or Poisson--Boltzmann equation, and iteratively obtains self-consistency between the reaction field and charge distribution of the quantum system. These ideas have their roots in Onsager's consideration of a dipole and a molecule with point-dipole polarizabilities inside a sphere [252], and have also been developed in the Polarizable Continuum Model of Tomasi and co-workers [253, 254]. The first use of these ideas in DFT calculations using ADF was by Chen et al. [255] and Mouesca et al. [256]. At around the same time the Tomasi group also used the PCM model with DFT [257]. In the past, SCRF calculations with ADF were done in the research groups of Noodleman, Bashford and Case using custom modifications of ADF. Now the method is available in a standardized form.

More recently, the ability to include a classical representation of a macromolecular environment, such as a protein around an enzyme active site, has been added to ADF. In what follows, we refer to the macromolecule surrounding the QM system as 'the protein' for historical reasons. It should be understood that the QM region and the protein region are distinct. For example, if the QM region includes some protein sidechain moieties, these are part of the QM system, and **not** the protein.

The protein exerts its effect in two ways: (1) The protein region can be assigned a dielectric constant (EPSPROT) that is different from either the QM region (1.0) or the solvent region (EPSSOL). This changes the reaction field that arises from the polarization of the environment by the charges of the QM atoms. (2) The protein has charges of it's own, which interact with the QM region in a way that is mediated by the complex dielectric environment described above. Calculations of this kind were first done by Li et al. [343].

The specific algorithm is:

1. Solve for the electronic structure in vacuum by the usual QM methods.

2. Nucleus-centered partial charges are derived from the electronic structure using a least-squares potential-fitting SVD algorithm [256], which is somewhat similar to the CHELPG algorithm [258].
3. *Only in calculations with a surrounding protein:* Calculate the electrostatic potential due to the partial charges of the protein in the three-dielectric environment defined by the protein and QM-region coordinates and radii. We refer to this field as the 'protein field' φ_{prot}.
4. Using the nucleus-centered partial charges:
    - Solve the Poisson equation (or, if an ionic strength is specified, the Poisson-Boltzmann equation) for φ_{sol}, the electrostatic potential in the presence of the above-described solvent dielectric environment.
    - Solve the Poisson equation for φ_{vac}, the electrostatic potential in a uniform vacuum environment.
    - The reaction field potential φ_{rf} is the difference between these two potentials: $\varphi_{rf} = \varphi_{sol} - \varphi_{vac}$.
5. Recalculate the electronic structure in the presence of the reaction field, and the protein field (if a protein field is present). This is done by adding the potentials φ_{rf} (and possibly φ_{prot}) to the total potential evaluated on the numerical integration grid when calculating Fock matrix elements.
6. Check changes in energy for convergence, and if not converged, return to step 4. Note that because φ_{prot} does not depend on electronic structure in the QM region, there is no need to return to step 3.

The algorithm also includes a correction for the small difference between the fit and true electron densities.

At this time, the SCRF procedures are only available for single-point calculations. No geometry optimization or frequency calculations can be done together with SCRF.

Several user-settable options can affect the SCRF procedure.

It is important the the QM calculation be done without use of molecular symmetry, as the finite-difference Poisson solver involves approximations that do not respect symmetry. Therefore, one should specify SYMMETRY NOSYM in the general input.

The choice of atomic radii and the probe radius determines the location of the dielectric boundary. This molecular surface comprises spherical patches of the contact surface generated by a solvent-sized probe rolling over the atomic radii, the toroidal surfaces swept out as the probe roles in grooves between pairs of atoms, and the inverse spherical patches generated when the probe simultaneously touches three or more [251, 259]. In volumetric terms, the In volumetric terms, the molecular surface divides the space of all points that are accessible to any part of a probe sphere that cannot penetrate into any of the atomic spheres, for the space that is not accessible [260]. The solvent accessible volume is assigned the solvent dielectric constant, while the inaccessible volume is assigned a dielectric constant of unity (the vacuum dielectric constant). Smaller radii move the dielectric surface closer to the atomic nuclei which typically leads to stronger calculated solvent effects.

The routine for calculating atomic partial charges chooses charges that best reproduce the potential outside the molecule that is generated by the nuclei and the electron density. It sets up a grid of potential-sampling points in a region outside the molecule, calculates the potential on this grid due to the electron density and nuclei, and finds the set of nucleus-centered charges that provides the best fit, in a least-squares sense, to the potential on the sampling points. The charge optimization is done using a singular value decomposition (SVD) method described by Mouesca et al. [256]. These calculations can be affected by user options concerning constraints on total charge and dipole, charge-fitting grid spacing and SVs to be deleted.

The solution of the Poisson or Poisson--Boltzmann equation utilizes libraries from Donald Bashford's MEAD programming suite. These use a finite-difference method that involves setting up cubic lattices around the molecule. Finer grids can be nested inside coarser ones to help manage trade-offs between accuracy and computational cost. The finest grid should cover the entire quantum system (that is, regions of significant electron density), and for good accuracy of φ_{rf} should be no coarser than about 0.15 Å. The outermost grid should extend 10 to 15 Å into the space beyond the model so that boundary conditions are accurate. A

reasonable scheme for grid selection based on atomic coordinates is implemented as the default, but the MEAD grids are also user-adjustable.

The SCRF method should not be used in the same ADF run with other solvation modeling methods such as COSMO or DRF. However, it is often useful to do geometry optimization with COSMO to get the single-point geometry for an SCRF run.

Additional considerations with a protein environment

The special considerations that apply when a protein environment is specified are similar to those in QM/MM calculations. The classical macromolecular region is specified by a .pqr file that provides the Cartesian coordinates, partial charges and radii of all the atoms of the region, as well as atom identifiers such as residue and atom numbers, in a format similar to PDB format. The QM atoms are specified in the ADF input file in the usual way. In contrast to the QM/MM situation, the ATOM block in the ADF input should specify only QM atoms, while the .pqr file should specify only the atoms of the protein. It is essential that no symmetry be used in the QM calculation (i.e. SYMMETRY NOSYM) even if the QM system possesses symmetries. The protein surroundings, as well as the numerical asymmetries of the finite-difference Poisson solver, will break any symmetry that might exist, but the QM part of the calculation has no way to know in advance about this broken symmetry.

Classical--QM bonds

If there is no covalent linkage between atoms in the QM region and the protein, the situation is straightforward. If there is such a linkage, some of the same considerations regarding capping atoms and link atoms as in QM/MM calculations apply [ref. Sec. 1.2 of ADF QM/MM manual]. One must typically add capping hydrogen atoms to the QM system to satisfy valence requirements. Because the interaction between the regions is purely electrostatic, we do not encounter the classical-side valency issues or van der Waals interaction issues that one finds in the QM/MM case. This is because this model involves no concept of bonds among the protein atoms. However, the link atom will typically be only a few tenths of an Angstrom from the capping hydrogen atom, so if the link atom has a significant partial charge, it could perturb the QM system in strange ways. Therefore we typically choose the link atom to be one that would normally only have a small partial charge (such as an aliphatic carbon) and then set its charge to zero in the .pqr file, perhaps with some minor adjustments of nearby partial charges to maintain the correct total charge.

Generation of the .pqr file

Typically one has at hand a PDB-format file of the macromolecule that one would like to include as the 'protein' part of the environment, and one needs to generate a suitable pqr file from this. PDB files usually lack hydrogen atoms, and often have missing heavy atoms at chain termini and in flexible loops. These problems need to be addressed in much the same way as for standard molecular mechanics calculations, although the requirements for a pqr file are less rigorous given the lack of classical bonds in the model. Although there are no specific tools for generating PQR files in the ADF distribution, a number of tools are available in the molecular modeling community. Amber users can use the usual LEaP tool to prepare prmtop and coordinate files, and then the utility, ambpdb can be used with the -pqr flag to generate a pqr file. Users of CHARMM can generate a structure file (PSF file) for the protein and a CHARMM-generation PDB file (containing hydrogens and using CHARMM naming conventions) and then use a perl script provided by the MEAD suite (available from http://stjuderesearch.org/site/lab/bashford) to generate a pqr file from the CHARMM-generated files. Another option is to use pdb2pqr either as a downloaded program or as web server. Information about pdb2pqr can be found at http://www.poissonboltzmann.org/pdb2pqr.

Once the initial version of the pqr file is generated, it is likely that some hand editing will be needed, particularly if there are covalent linkages between the protein and QM atoms. In cases with no linkages, such as a (QM) drug in a (classical) protein binding site, the most straightforward procedure is to leave the drug out during generation of the pqr file. Then no pqr-editing needs to be done, but it is necessary that the QM coordinates and the coordinates in the pqr file be compatible, e.g., that the drug is correctly positioned in the binding site. If there are covalent links between the QM and protein regions, for example if some sidechains in an enzyme active site are part of the QM region, then the pqr file will probably have been prepared with these sidechains included. *Any atoms that will be part of the QM region must be removed*

from the pqr file. If this is not done, the QM region will ``feel'' protein partial charges that sit on top of the QM-region nuclei; this will generate unphysical results or outright program crashes. There is no need to worry about dangling valences in the pqr file, but it may be necessary to edit link-atom charges, as noted above. The QM atom set may need to include capping atoms, (typically hydrogen) to satisfy valence requirements where bonds cross the QM--protein boundary.

Format of the pqr file

A pqr file is similar to a PDB (Protein Data Bank) file but with atomic charge and radii in the occupancy and B-factor columns, respectively. More specifically, lines beginning with either 'ATOM' or 'HETATM' (no leading spaces) are interpreted as a set of tokens separated by one or more spaces or TAB characters. Other lines are ignored. The tokens (including the leading ATOM or HETATM are interpreted as follows:

ignored ignored atName resName resNum x y z charge radius chainID

The first two (ignored) tokens must be present, or the line will not be parsed correctly. The chainID token is optional, and any tokens beyond that are ignored. Tokens can be of arbitrary length, but must not contain spaces or tabs. Lines that do not begin with "ATOM" or "HETATM" are ignored. The programs make no distinction between ATOMs and HETATMs. *No atname-resnum-chainid combination can occur more than once.*

Note that the .pqr format does not support some PDB-isms such as a altLoc fields, and a one-character chainID between resName and resSeq. Doing so would break the whitespace separated tokens convention that allows for easy processing with perl scripts, etc. Instead we put 'chainID' in a position more or less analogous with the PDB segID. (Note that the pdb2pqr program differs on this point, and pqr files with chainIDs between resName and resSeq may need to be modified.

SCRF Input

The SCRF input is contained in an SCRF input block as shown below, optional keywords being surrounded by curly brackets.

```
SCRF
  MEADGRID string integer real
  RADIUS string real
  {CYCLES integer}
  {TOLERANCE real}
  {ATOM_MAXR real}
  {CHGFIT_CONSTRAIN string}
  {DELATOM integer}
  {GRID_SPACING real}
  {SVD_CONDITION real}
  {SV_DELETE integer}
  {EPSSOL real}
  {IONIC_STR real}
  {SOLRAD real}
  {PROTEIN string}
  {EPSPROT real}
  {SAVESTATE string}
END
```

The SCRF block contains two mandatory keys: MEADGRID and RADIUS. All other keys are optional.

`MEADGRID string integer real`

Specifies the centering style, dimension and spacing for the MEAD grid. Recognized centering styles are "ON_ORIGIN" and "ON_GEOM_CENT". The grid dimension specifies the number of points on one edge of a cubic grid. The grid spacing is given in Angstroms. The edge length of the grid is the product

of the dimension minus 1 and the spacing. Multiple records may be used to specify sequentially finer grid levels, but finer grids must fit within the coarsest grid.

`RADIUS string real`

Specifies the radius in Angstroms for an atom type. Used in fitting the ADF electronic structure to partial atomic charges and for defining the boundary between regions of low and high dielectric in MEAD. The atom types should be the same as those used in the ATOMS input block. There must be one RADIUS record for each atom type in the ATOMS input block.

`CYCLES integer`

Specifies the maximum number of cycles of SCRF to perform. Whether or not the SCRF run has converged, it will terminate when the number of cycles exceeds the value specified by CYCLES.

`TOLERANCE real`

Specifies convergence criterion in kcal/mol for SCRF. For each cycle of SCRF the sum of ADF energy, reaction energy, the energy correction and nuclear reaction energy is calculated. If the difference in subsequent sums is less than the TOLERANCE value, SCRF is considered to have converged. Defaults to 0.01.

`ATOM_MAXR real`

Specifies the outer atomic radius in Angstroms for the system. For each atom, grid points that lie between the atomic radius specified by the RADIUS keyword and the outer atomic radius specified here are included in charge fitting. Defaults to 5.0.

`CHGFIT_CONSTRAIN string`

Specifies the type of constraints to be used in charge fitting. Recognized constraints are "NONE", "CHARGE" or "DIPOLE". NONE specifies no constraints will be applied, CHARGE specifies that only the molecular charge will be constrained and DIPOLE specifies that both the molecular charge and dipole will be constrained. Default is DIPOLE.

`DELATOM integer`

Specifies which atoms should be excluded from the charge fitting procedure. The input order in the ATOMS input block is used to identify the excluded atom. Multiple DELATOM records may be declared and will be applied cumulatively. The default is to include all atoms.

`GRID_SPACING real`

Specifies the grid spacing in Angstroms for the charge fitting grid. The default is 0.2.

`SVD_CONDITION real`

Specifies a condition number threshold for inclusion of singular values (SV) in singular value decomposition (SVD) during charge fitting. The default is 0.000001.

`SV_DELETE integer`

Instead of using a condition number threshold for deciding which SV to include in charge fitting, SV_DELETE may be used to specify how many SV should be excluded. The smallest SV are excluded first. The default is to use a condition number threshold. If both SV_DELETE and SVD_CONDITION are specified, the SV_DELETE value will take precedence.

`EPSSOL real`

Specifies the solvent dielectric for MEAD. Defaults to the dielectric of water: 80.0.

`IONIC_STR real`

Specifies an ionic strength in mol/L for the solvent in MEAD. Defaults to 0.0.

`SOLRAD real`

Specifies the radius in Angstroms of a solvent-sized probe that rolls along the surface of the molecular system to define the dielectric boundary. Defaults to a water-sized probe size of 1.4.

`PROTEIN string`

Use of this keyword turns on the SCRF solinprot option by specifying the prefix of a pqr file containing the protein definition for MEAD solinprot. The filename suffix must be pqr. Pqr format contains one line per atom and begins with the ATOM keyword followed by 10 fields separated by white space and in the order: atom number, atom name, residue name, residue number, x, y and z coordinates, partial atomic charge and atomic radius. The atoms of the quantum mechanical system should NOT be included in the pqr file.

`EPSPROT real`

Specifies the protein dielectric for MEAD. Defaults to 4.0.

`SAVESTATE string`

Specifies a filename in which ground state data for a subsequent VSCRF calculation will be saved. The data is saved in a binary KF file.

VSCRF: Vertical Excitation Self-Consistent Reaction Field

Introduction

by Donald Bashford, St. Jude Children's Research Hospital Memphis, October 26, 2011.

VSCRF is a method of accounting for the effect of the solvent environment (and possibly a macromolecular environment) on the energy of absorption/emission transitions between electronic states. It is built on the *Delta*SCF method of calculating electronic transition energies and a continuum dielectric model of the environment in a way that is consistent with the Franck--Condon principle. The original development and application of the procedure can be found in the papers of Liu et al. [344] and Han et al. [345].

Suppose one is calculating the energy for a transition between an initial state *i* with total density (including nuclei) ρ_i and a final state *f* with total density ρ_f. State *i* has existed long enough that the dielectric environment surrounding it has had time to fully respond, including changes in the distributions of nuclear coordinates and solvent-molecule dipole moments. The density ρ_i has therefore given rise to a reaction field $\varphi^{(r)}_{i,eq}$ and is self-consistent with it according to the ideas described in the SCRF part of the documentation, therefore the energy in solvent of the initial system is

$$G^i(\rho_i) = E_o(\rho_i) + 1/2 \int \rho_i(x)\, \varphi^{(r)}_{i,eq}(x)\, d^3x$$

where E_o is the vacuum density functional. During the excitation process, the change in density $\Delta \rho_{fi} = \rho_f - \rho_i$ must work against the non-optical component of the pre-existing equilibrium reaction field as well as the optical reaction field arising in response to $\Delta \rho$ itself. The full theoretical development is given in the paper of Liu et al. [344] and the result is that the excitation energy is:

$$\Delta G^{if}_{ex} = E_o(\rho_i) - E_o(\rho_f)\ 1/2 \int \Delta \rho_i(x) [2 \varphi^{(r)}_{i,eq}(x) + \Delta \varphi^{(r)}_{op}(x)]\ d^3x$$

where $\Delta \varphi^{(r)}_{op}$ is the reaction field that arises from the *optical* dielectric constant which is related to the index of refraction: $\varepsilon_{op} = n^2 \approx 2$ for organic liquids.

In the following description of the VSCRF algorithm, we adopt the language of an optical absorption calculation. The initial state is therefore referred to as the ground state, and the final state as the excited state. For emission calculations, these roles would be reversed. If there is a macromolecular environment surrounding the protein that modeled by macroscopic dielectric theory as described in the SCRF documentation in this manual, this is referred to as the 'protein'. With these definitions in mind, the VSCRF algorithm is:

1. An SCRF calculation for the ground state, either with or without a surrounding protein, is a prerequisite for a VSCRF calculation.
2. The electronic structure of the excited state (with occupations initially selected by the user) is calculated in vacuum.
3. The ground-state reaction field, the protein field (if any) and electron density are read in from the results of the previous ground-state calculation. It is essential that the excited-state and ground state calculations use exactly the same geometry, ADF grids, basis functions, XC functionals, MEAD grids and so on.
4. The excited-state electron density calculated in step 1 is subtracted from the stored ground-state density to obtain $\Delta\rho$.
5. Nucleus-centered partial charges are derived from $\delta\rho$ using a least-squares potential-fitting SVD algorithm [256] which is somewhat similar to the CHELPG algorithm [258].
6.
   - Set the exterior dielectric environment to the optical dielectric constant of the solvent, and the optical dielectric constant of the protein. These are equal to the square of the index of refraction of these materials, and are typically ≈ 2 (the default value).
   - Using the nucleus-centered partial charges for $\Delta\rho$, solve the Poisson equation using the finite-difference method in the above dielectric environment to obtain $\Delta\varphi_{sol}$.
   - Set the exterior dielectric constant to 1 everywhere and solve the Poisson equation by the finite-difference method to obtain $\Delta\varphi_{vac}$
   - The optical reaction field to the excitation, $\Delta\varphi^{(r)}_{op}$ is the difference between $\Delta\varphi_{sol}$ and $\Delta\varphi_{vac}$. In this subtraction, artifacts related to the finite difference method cancel out.
7. Recalculate the electronic structure in the presence of the optical reaction field, as well as the previously stored ground-state reaction field and the protein field (if any). This is done by placing these potentials on the ADF numerical integration grid when calculating Fock matrix elements.
8. Check changes in energy for convergence, and if not converged, return to step 6.

It should be noted that some additional calculations may be needed to obtain a proper excited-state energy. For example, if the first excited singlet state is prepared by promoting one of the α electrons from the HOMO to the LUMO, the resulting (single-determinant) wavefunction is not a pure singlet state but actually a mixture of singlet and triplet states [323]. Calling this mixed state S'_1, the resolution of the problem is to also calculate the energy of a corresponding triplet state, T_1 (prepared, for example, by adding one α electron to the LUMO and removing one β electron from the HOMO) and using the formula [323]:

$$E_{S_1} = 2_{S'_1} - E_{T_1}$$

See Liu et al. [344] or Han et al. [345]. for specific examples of this in the ADF/VSCRF context.

VSCRF Input

The VSCRF input is contained in a VSCRF input block as shown below, optional keywords being surrounded by curly brackets.

```
VSCRF
  MEADGRID string integer real
  RADIUS string real
  INITIAL_STATE string
  {CYCLES integer}
  {TOLERANCE real}
  {ATOM_MAXR real}
  {CHGFIT_CONSTRAIN string}
  {DELATOM integer}
  {GRID_SPACING real}
  {SVD_CONDITION real}
  {SV_DELETE integer}
  {EPS_OPT_SOL real}
  {SOLRAD real}
  {PROTEIN string}
  {EPS_OPT_PROT real}
END
```

The VSCRF block contains three mandatory keys: MEADGRID, RADIUS and INITIAL_STATE. All other keys are optional. It is highly recommended that the parameters specifying the MEAD grid and the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations. These procedures involve some numerical error, and keeping parameters the same will promote cancellation of these errors as energy differences are taken as the final results.

`MEADGRID string integer real`

Specifies the centering style, dimension and spacing for the MEAD grid. Recognized centering styles are "ON_ORIGIN" and "ON_GEOM_CENT". The grid dimension specifies the number of points on one edge of a cubic grid. The grid spacing is given in Angstroms. The edge length of the grid is the product of the dimension minus 1 and the spacing. Multiple records may be used to specify sequentially finer grid levels, but finer grids must fit within the coarsest grid. It is highly recommended that the parameters specifying the MEAD grid be the same for both the SCRF initial state and the VSCRF final state calculations.

`RADIUS string real`

Specifies the radius in Angstroms for an atom type. Used in fitting the ADF electronic structure to partial atomic charges and for defining the boundary between regions of low and high dielectric in MEAD. The atom types should be the same as those used in the ATOMS input block. There must be one RADIUS record for each atom type in the ATOMS input block.

`INITIAL_STATE string`

Specifies the filename of a binary KF file containing the ground state data for VSCRF. The KF file is created by a preliminary SCRF calculation using the SAVESTATE option. It is highly recommended that the parameters specifying the MEAD grid and the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations.

`CYCLES integer`

Specifies the maximum number of cycles of VSCRF to perform. Whether or not the VSCRF run has converged, it will terminate when the number of cycles exceeds the value specified by CYCLES.

`TOLERANCE real`

Specifies convergence criterion in kcal/mol for VSCRF. For each cycle of VSCRF the sum of ADF energy, Potential Term, and Response Term is calculated. If the difference in subsequent sums is less than the TOLERANCE value, VSCRF is considered to have converged. Defaults to 0.01.

ATOM_MAXR real

Specifies the outer atomic radius in Angstroms for the system. For each atom, grid points that lie between the atomic radius specified by the RADIUS keyword and the outer atomic radius specified here are included in charge fitting. It is highly recommended that the parameters specifying the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations. Defaults to 5.0.

CHGFIT_CONSTRAIN string

Specifies the type of constraints to be used in charge fitting. Recognized constraints are "NONE", "CHARGE" or "DIPOLE". NONE specifies no constraints will be applied, CHARGE specifies that only the molecular charge will be constrained and DIPOLE specifies that both the molecular charge and dipole will be constrained. It is highly recommended that the parameters specifying the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations. Default is DIPOLE.

DELATOM integer

Specifies which atoms should be excluded from the charge fitting procedure. The input order in the ATOMS input block is used to identify the excluded atom. Multiple DELATOM records may be declared and will be applied cumulatively. It is highly recommended that the parameters specifying the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations. The default is to include all atoms.

GRID_SPACING real

Specifies the grid spacing in Angstroms for the charge fitting grid. It is highly recommended that the parameters specifying the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations. The default is 0.2.

SVD_CONDITION real

Specifies a condition number threshold for inclusion of singular values (SV) in singular value decomposition (SVD) during charge fitting. It is highly recommended that the parameters specifying the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations. The default is 0.000001.

SV_DELETE integer

Instead of using a condition number threshold for deciding which SV to include in charge fitting, SV_DELETE may be used to specify how many SV should be excluded. The smallest SV are excluded first. The default is to use a condition number threshold. If both SV_DELETE and SVD_CONDITION are specified, the SV_DELETE value will take precedence. It is highly recommended that the parameters specifying the details of the charge-fitting SVD procedure be the same for both the SCRF initial state and the VSCRF final state calculations.

EPS_OPT_SOL real

Specifies the dielectric constant of the solvent at optical frequencies for MEAD. This value is equal to the square of the index of refraction of the solvent. Defaults to the value for water: 2.0.

SOLRAD real

Specifies the radius in Angstroms of a solvent-sized probe that rolls along the surface of the molecular system to define the dielectric boundary. Defaults to a water-sized probe size of 1.4.

PROTEIN string

Use of this keyword turns on the VSCRF solinprot option by specifying the prefix of a pqr file containing the protein definition for MEAD solinprot. This option can only be used if the ground state SCRF calculation also used a protein. The filename suffix must be pqr. Pqr format contains one line per atom and begins with the ATOM keyword followed by 10 fields separated by white space and in the order: atom number, atom name, residue name, residue number, x, y and z coordinates, partial atomic charge and atomic radius. The atoms of the quantum mechanical system should NOT be included in the pqr file.

`EPS_OPT_PROT real`

Specifies the dielectric constant of the protein at optical frequencies for MEAD. This value is equal to the square of the index of refraction of the protein. Defaults to the value for organic liquids: 2.0.

3D-RISM: 3D Reference Interaction Site Model

Introduction

The three-dimensional reference interaction site model with the closure relation by Kovalenko and Hirata (3D-RISM-KH) provides the solvent structure in the form of a 3D site distribution function, $g_{\gamma}^{UV}(r)$, for each solvent site, γ. Note that the use of 3D-RISM as implemented in ADF is an export option.

It enables, at modest computational cost, the calculations of thermodynamics, electronic properties and molecular solvation structure of a solute molecule in a given molecular liquid or mixture. Using 3D-RISM, one can study chemical reactions, including reaction coordinates and transition state search, with the molecular solvation described from the first principles. The method yields all of the features available by using other solvation approaches. The 3D-RISM part of ADF has not been parallellized.

Details on the implementation of 3D-RISM-KH in ADF can be found in Ref. [300], with applications in Ref. [301]. The theory of 3D-RISM-KH in combination with DFT can be found in Refs. [302-304,349]. A combination of 3D-RISM-KH with FDE (frozen-density embedding) can be found in Ref. [305].

Similar to explicit solvent simulations, 3D-RISM properly accounts for chemical peculiarities of both solute and solvent molecules, such as hydrogen bonding and hydrophobic forces, by yielding the 3D site density distributions of the solvent. Moreover, it readily provides, via analytical expressions, all of the solvation thermodynamics, including the solvation free energy potential, its energetic and entropic decomposition, and partial molar volume and compressibility. The expression for the solvation free energy (and its derivatives) in terms of integrals of the correlation functions follows from a particular approximation for the so-called closure relation used to complete the integral equation for the direct and total correlation functions.

Solvation free energy

The solvation free energy can be calculated as:

Solvation Free Energy

= [(Total Bonding Energy) + (Internal Energy) - T * (Entropy) + (Excess Chemical Potential)]_3D-RISM

- [(Total Bonding Energy) + (Internal Energy) - T * (Entropy)]_Gas-Phase

If one assumes that the internal energy and the vibrational and rotational entropy of the solute molecule is the same in solution and in gas phase, then this simplifies to:

Solvation Free Energy

= [(Total Bonding Energy) + (Excess Chemical Potential)]_3D-RISM

- [(Total Bonding Energy)]_Gas-Phase

However, a formally accurate calculation should include the difference between the thermal corrections from frequency calculations produced by ADF in the SCF calculation with 3D-RISM-KH solvation and in gas phase.

Input

When performing 3D-RISM simulations, each atom in the **ATOMS** block must have two parameters specified, SigU and EpsU, for example:

```
ATOMS
  C     0.00    0.00    0.00     SigU=3.50    EpsU=0.066
  ...
END
```

The SigU and EpsU parameters have the same meaning as Sigma_alpha and Eps_alpha for atoms of the solvent in the SOLVENT sub-block below. They can be obtained from a Lennard-Jones force-field parameter sets.

All 3D-RISM-related input keys are contained in a **RISM** input block. Below, only the mandatory keywords are shown. Optional keywords are described in the next section.

```
RISM title
  RISM1D
    FLUIDPARAM temper=298. DielConst=78.497 UTotDens=1/A3 0.03333
  SUBEND
  SOLVENT ArbitrarySolventName
    UNITS uWeight=g/mol  ULJsize=A  ULJenergy=kcal/mol Ucoord=A Udens=1/A3
    PARAMETERS Weight NAtomTypes
      N1    Z_alpha1   Sigma_alpha1   Eps_alpha1   X1_1 Y1_1 Z1_1
                                                  X1_2 Y1_2 Z1_2
                                                  ...  ...  ...
      N2    Z_alpha2   Sigma_alpha2   Eps_alpha2   X2_1 Y2_1 Z2_1
                                                  X2_2 Y2_2 Z2_2
                                                  ...  ...  ...
      ...
      DENSPE=density
  SUBEND
  SOLUTE ArbitrarySoluteName
    BOXSIZE   sizeX  sizeY  sizeZ
    BOXGRID    npX    npY    npZ
  SUBEND
END
```

The **RISM1D** sub-block contains general parameters for the 1D-RISM calculation of the solvent(s). Even though all RISM1D sub-keys have reasonable defaults, the FLUIDPARAM sub-key deserves a special attention because its default values are only applicable if the solvent is water. Thus, you may need to change some of these values when modeling a different solvent, at least the dielectric constant and the density. Note that even when using all default values from the RISM1D sub-block the sub-block itself must be specified, even if empty. See below for complete description of the RISM1D sub-block.

The **SOLVENT** sub-block can be repeated if the solvent is a mixture. Each SOLVENT sub-block contains parameters for one solvent. First, each solvent has a name, which is specified on the SOLVENT keyword's line and is arbitrary. The first line in the SOLVENT sub-block must contain the UNITS key. You should leave it at the default values. Then follow the actual solvent parameters. In principle, each solvent consists of multiple atoms and functional groups. For simplicity, we will call each of them an atom. For example, in 3D-RISM therms, methanol consists of 3 "atoms": CH_3, O, and H. Each such atom has a set of three parameters, shared between all atoms of the same type, and the coordinates. These parameters follow the PARAMETERS keyword. The line with the PARAMETERS keyword itself must specify the molecular weight

of the solvent and the number of atom types that follow. The first line for each atom type contains, in this order: number of atoms of this type, z_α, σ_α, ε_α, three coordinates for the first atom of this type. If there is more than one atom of this type then the coordinates for the 2nd and other atoms follow on subsequent lines. The SOLVENT sub-block is concluded by the specific density of this solvent, by default, in molecules per cubic angstrom. This number should be equal to the total density for mono-component solvents.

For example, the SOLVENT block for water would typically look as:

```
  SOLVENT water
    UNITS      uWeight=g/mol  ULJsize=A  ULJenergy=kcal/mol Ucoord=A
Udens=1/A3
    PARAMETERS Weight=18.015   nAtoms=2
     1    -0.8476   3.166    0.1554       0.000000  0.00000  0.000000
     2     0.4238   1.000    0.0460      -0.816490  0.00000  0.577359
                                          0.816490  0.00000  0.577359
    DENSPE=0.03333
  SUBEND
```

The **SOLUTE** sub-block specifies 3D-RISM parameters for your molecule. The BOXSIZE and BOXGRID sub-keys specify dimensions of the simulation box, in Angstrom, and the number of points of grid in each direction. The box should be twice as large as the molecule and the BOXGRID values must be a power of 2. The size/np ratio defines the grid spacing in each direction and this should be not larger than 0.5 angstrom.

The **optional** 3D-RISM keys for the RISM1D and SOLUTE sub-blocks are listed below together with their defaults.

```
...
  RISM1D Theory=RISM Closure=KH
!   optional RISM1D subkeys with their default values:
    FLUIDPARAM temper=298. DielConst=78.497 UTotDens=1/A3 0.03333
    OUTPUT PrintLev=5 File=solvent OutList=GXT
    GRID   Nr=8192  dR=0.025 MaxRout=100.0 MaxKout=0.0
    MDIIS  N=20 Step=0.5 Tolerance=1.e-12
    ELSTAT LRsmear=1.0 Adbcor=0.5
    ITER  Ksave=-1 Kshow=1 Max=10000
  SUBEND
...
```

OUTPUT key:
PrintLev - print level;
File - common base name for output files;
OutList - which information will be written to output files: G - distribution function, C - direct correlation function, H - total correlation function, T - thermodynamics.

GRID key:
Nr - the size of the 1D-RISM grid, must be a power of 2;
dR - mesh size in Angstrom;
MaxRout - plot range in direct space;
MaxKOut - plot range in reciprocal space.

MDIIS key:
N - number of vectors in the DIIS space;
Step - step size;
Tolerance - convergence criterion.

ELSTAT key:
LRsmear - smearing parameters for coulomb potential;

ITER key:
Ksave - save the current solution every Ksave steps (0 - do not save);
Kshow - print convergence progress every Kshow steps;
Max - maximum number of iterations.

FLUIDPARAM key:
Temper - temperature;
DielConst - dielectric constant;
UTotDens - units for total density: g/cm3, kg/m3, 1/A3 are valid units followed by the density value.

```
...
  SOLUTE ArbitrarySoluteName
    outlist=MH  closure=KH  xvvfile=solvent.xvv outfile=rism3d
    Nis=10  DELOZ=0.5 TOLOZ=2.0e-6
    Ksave=-1  Kshow=1  Maxste=10000
    Output=4
    CHRGLVL=MDCq
  SUBEND
...
```

Outlist - output requested: G - distribution function, C - direct correlation function, H - total correlation function
Closure - closure for the 3D problem: KH – Kovalenko-Hirata, HNC - hypernetted chain approximations, PY - Perkus-Yevik
Xvvfile - name of the file with the results of the 1D-RISM calculation specified in the RISM1D keyword above, with .xvv appended to it;
Outfile - name of the output text file;
Output - print level;
CHRGLVL - which charges computed by ADF to use. This can be MDCq (default), MDCd, MDCm, or EXACT.
The Nis, DELOZ, and TOLOZ have the same meaning for 3D-RISM as parameters of the MDIIS keyword of the RISM1D block. Likewise, Ksave, Kshow, and Maxste are analogous to the parameters of the ITER key in RISM1D.

Parameters for some solvents

Atom	z_α / a.u.	σ_α / Å	ε_α / kcal/mol	X / Å	Y / Å	Z / Å
Water Weight=18.015 nAtoms=3						
O	-0.8476	3.166	0.1554	-0.0646	0.0	0.0
H	0.4238	0.7	0.046	0.5127	0.8165	0.0
H	0.4238	0.7	0.046	0.5127	-0.8165	0.0
Methanol Weight=32.042 nAtoms=3						
CH_3	0.265	3.775	0.207	-1.43	0.0	0.0
O	-0.7	3.07	0.17	0.0	0.0	0.0
H	0.435	0.7	0.046	0.2998	0.8961	0.0
Ethanol						
CH_3	0.0	3.775	0.207	-1.9028	-1.4551	0.0
CH_2	0.265	3.905	0.118	-1.43	0.0	0.0
O	-0.7	3.07	0.17	0.0	0.0	0.0
H	0.435	0.7	0.046	0.2998	0.8961	0.0
Acetonitrile Weight=41.0520 nAtoms=3						
CH_3	0.269	3.6	0.38	1.46	0.0	0.0
C	0.129	3.4	0.099	0.0	0.0	0.0
N	-0.398	3.3	0.099	-1.17	0.0	0.0

Dimethylsulfoxide Weight=78.135 nAtoms=4						
CH_3	0.160	3.81	0.16	0.0	1.7167	-0.5413
CH_3	0.160	3.81	0.16	-1.6653	-0.4168	-0.5413
S	0.139	3.56	0.395	0.0	0.0	0.0
O	-0.459	2.93	0.28	0.0	0.0	1.53
Isopropanol Weight=60.09502 nAtoms=5						
O	-0.700	3.0700	0.1700	1.32393	0.14660	-0.01452
H	0.435	0.7000	0.0460	1.84501	0.01837	-0.79238
CH	0.265	3.8500	0.0800	0.84520	1.49406	-0.00612
CH3	0.000	3.7750	0.2070	1.38816	2.18026	1.24895
CH3	0.000	3.7750	0.2070	-0.68415	1.45191	-0.02015
Benzene						
C	-0.115	3.55	0.07	1.4	0.0	0.0
C	-0.115	3.55	0.07	0.7	-1.2124	0.0
C	-0.115	3.55	0.07	-0.7	-1.2124	0.0
C	-0.115	3.55	0.07	-1.4	0.0	0.0
C	-0.115	3.55	0.07	-0.7	1.2124	0.0
C	-0.115	3.55	0.07	0.7	1.2124	0.0
H	0.115	2.42	0.03	2.48	0.0	0.0
H	0.115	2.42	0.03	1.24	-2.1477	0.0
H	0.115	2.42	0.03	-1.24	-2.1477	0.0
H	0.115	2.42	0.03	-2.48	0.0	0.0
H	0.115	2.42	0.03	-1.24	2.1477	0.0
H	0.115	2.42	0.03	1.24	2.1477	0.0

OPLS-AA Parameters for common atoms and atom group

The table below contains sigma and epsilon parameters for some of the most common solvent groups collected from [433, 434, 435]. These parameters are kindly provided by Leonardo Costa.

Atom	Example	σ_α / Å	ε_α / kcal/mol	Reference
All atoms model				
C (SP^3)	methane	3.73	0.294	433
C (SP^3)	neopentane	3.80	0.050	433
C (SP^2)	isobutene	3.75	0.105	433
C (SP)	acetonitrile	3.40	0.099	434
C (arom)	benzene	3.55	0.07	434
C	chloroform	4.10	0.05	435
H	O---H	0.7	0.046	434
H	hydrocarbons	2.42	0.03	434
O	water	3.166	0.1554	434
O	alcohol	3.07	0.17	434
O	sulfoxide	2.93	0.28	434
N (SP)	acetonitrile	3.3	0.099	434
S	sulfoxide	3.56	0.395	434
Cl	chloroform	3.40	0.300	435
United atom model				
CH (SP^3)	isobutane	3.850	0.080	433
CH (SP^2)	2-butene	3.800	0.115	433

CH (arom)	benzene	3.750	0.110	433
CH_2 (SP^3)	n-butane	3.905	0.118	433
CH_2 (SP^2)	1-butene	3.850	0.140	433
CH_3	hydrocarbon	3.775	0.207	433
CH_3	acetonitrile	3.6	0.38	434
CH_3	sulfoxide	3.81	0.16	434
NH_2	amine	3.3	0.17	434

Electric Field: Homogeneous, Point Charges, Polarizability

A homogeneous external electric field and/or the field due to point charges can be included in the Fock operator. Either can be applied in both a Single-Point calculation (or a Create run) or in geometry optimization. When applied in geometry optimization, it will allow for the molecule to rotate with respect to point charges or the field vector but not translate. Rigid translation is explicitly disabled to avoid drifting of the molecule in the external electric field.

```
EFIELD {ex ey ez}
 {x y z q
 x y z q
  ...
end }
```

`EFIELD`

This *general* key can be used as a simple key or as a block key. The block form applies if no argument (ex, ey, ez) is given or when the argument is followed by the *continuation symbol* (&).

`ex,ey,ez`

Define a homogeneous electric field in atomic units: hartree/(e bohr) = 27.211 V/bohr; the relation to SI units is: 1 hartree/(e bohr) = 5.14 ... $*10^{11}$V/m.
The units applied by ADF for the interpretation of homogeneous field values are not affected by any units used for specifying atomic coordinates. By default no homogeneous E-field is included.

`x, y, z, q`

The Cartesian coordinates and strength of a point charge (in elementary charge units, +1 for a proton). Each point charge must be specified on a separate line in the data block. The Cartesian coordinates are in the units of length that was set by units (for interpreting atomic coordinates input). By default no point charges are included.

Orientation of the fields

When the atomic coordinates are input in z-matrix format, the direction of the homogeneous field and the location of the point charges as specified in input are interpreted as referring to the standard Cartesian frame associated with z-matrix input. The standard frame means: the first atom at the origin, the second on the positive x-axis, the third in the xy-plane with positive y.

If the program rotates (and translates, as the case might be) the atoms from the input frame - or the auto-generated frame in case of z-matrix input - to some other frame, for instance to accommodate the internal ADF symmetry orientation requirements, the fields are transformed along with the atoms.

Symmetry

The homogeneous electric field and the point charge fields may polarize the electronic charge density. This must be accounted for in the point group symmetry. If symmetry is not specified in input, the program computes the symmetry from the nuclear frame and the fields.

Bonding energy

The bonding energy is computed as: the energy of the molecule in the field minus the energy of the constituent fragments in the same field. Of course, the fragments may not be polarized and hence not be self-consistent in this field. This depends on how the fragments themselves were computed.

Polarizability and hyperpolarizability

ADF supports a direct calculation of the (hyper) polarizability (see section on Spectroscopic Propereties). The static (hyper) polarizabilities can also be computed by applying a small homogeneous field and comparing the results with the field-free data.

2.5 Structure and Reactivity

See also
ADF-GUI tutorial: geometry optimization, vibrational frequencies, HCN isomerization reaction, spin coupling in Fe_4S_4
GUI manual: structure and reactivity
Examples: geometry optimization, reactivity

Run Types

The different *run types* are characterized by how the geometry is manipulated:

`SinglePoint`

The SCF solution is computed for the input geometry.

`GeometryOptimization`

The atomic coordinates are varied in an attempt to find a (local) energy minimum. One may let all coordinates free or only a subset, keeping the others frozen at their initial values.

`TransitionState`

Search for a saddle point. Similar to a GeometryOptimization, but now the Hessian at the stationary point presumably has one negative eigenvalue.

`LinearTransit`

The geometry is modified step by step from an initial to a final configuration. All of the coordinates or only a subset of them may be involved in the transit. The coordinates to be modified are the LinearTransit *parameters*. For each of the LinearTransit points (geometries) the computation may be a Single Point SCF calculation or a GeometryOptimization. In the latter case only those coordinates (or a subset of them) can be optimized that are not LinearTransit parameters. The LinearTransit feature can be used for instance to sketch an approximate reaction path in order to obtain a reasonable guess for a transition state, from where a true TransitionState search can be started.

`IRC or IntrinsicReactionCoordinate`

Tracing a reaction path from a transition state to reactants and/or products. A fair approximation of the transition state must be input. The end-point(s) - reactants / products - are determined automatically.

`Frequencies`

Computation of force constants and from these the normal vibrational modes and harmonic frequencies. The force constants can be calculated by numerical differentiation of the energy gradients at the equilibrium geometry and the slightly deviating geometries (making small displacements of the atoms). There is however also a possibility with the post-ADF program SD, to calculate the second derivatives analytically. This should usually be faster and in some cases more robust (no SCF convergence problems in displaced geometries, as only a single SCF is required). Note that this program still has some limitations. Most importantly, it can only handle X$alpha and VWN (LDA), but not GGA, calculations at this moment.

`CINEB`

Calculation of the reaction path and transition state search using Climbing-Image Nudged Elastic Band method. This method is further referrer to as NEB or CI-NEB. Using this method one can find a transition state between two known states, further referred to as initial and final states. The choice which state is initial and which is final is arbitrary. During calculation with this method, a number of replicas, or images, of the system is calculated. These images can be considered as being linked by an elastic band. Each image is optimized in such a way that on each step the forces parallel to the reaction path are removed and spring forces are added that keep distances to this image's neighbors equal. At the end of the optimization the images are evenly distributed along the reaction path, the image highest in energy being the transition state (if the climbing-image option is on, the default).

For all features that involve changes in geometry, i.e. all run types except the SinglePoint, it is imperative that you use single-atom fragments. Larger molecular fragments can only be applied in SinglePoint calculations.

Four keys are involved in the specification of the geometry and its manipulation:

`atoms`

sets the atomic (starting) positions.

`geometry`

Controls the run type and strategy parameters, such as convergence thresholds and the maximum number of geometry steps to carry out.

`atoms and geometry`

These two keys together are sufficient for a straightforward Optimization, TransitionState search, IRC run or a Frequencies computation. (Of course, you also need to specify the Fragments or BASIS key.

`constraints`

May be used to impose constraints for geometry optimizations, LT, and TS, in the new branch for optimization, which is the default for these optimizations. This key can not be used for IRC, NEB, or for the old branch of optimizations.

`geovar`

May be used to impose constraints, for instance when only a subset of all coordinates should be optimized. This key should be used for IRC, NEB, or for the old branch of optimizations, to impose constraints. GeoVar may also be used in a LinearTransit or NEB run to define the LinearTransit or NEB parameters, respectively, and their initial and final values.

Constraints and LinearTransit parameters in the old branch of optimizations may also be controlled within the atoms block if a MOPAC-style input format is used, see below.

Runtype control and strategy parameters

With the block key GEOMETRY you define the runtype and strategy parameters.

```
GEOMETRY {RunType { RunTypeData}}
 RunType {RunTypeData}
End
```

`RunType`

Can be:
- SinglePoint or SP
- GeometryOptimization or GeoOpt or GO
- TransitionState or TS
- IntrinsicReactionCoordinate or IRC
- LinearTransit or LT
- Frequencies or FREQ
- CINEB

If omitted the run type is GeometryOptimization.
If the key GEOMETRY is not used at all the run type is SinglePoint.
The run type specification can be given as argument to the geometry key, or in the data block, but not both. For some run types additional data may be given after the run type specification.

`RunTypeData`

(Optional) further specifications, depending on the run type. See the sections below.

Omission of the GEOMETRY key altogether effectuates a SinglePoint calculation. A straightforward optimization, with all features that can be set with geometry at their default values, is activated by supplying the key with an empty block:

```
GEOMETRY
End
```

More subkeys are available in the geometry block than just the run type specification. They are used to control strategy parameters such as convergence criteria. All subkeys are optional: default values take effect for those omitted. Some of the subkeys are only meaningful for certain run types. They will be ignored for other run types.

The initial approximation of the Hessian matrix may affect the number of optimization steps that are carried out to reach convergence. See the section Initial Hessian.

Geometry Optimization

Geometry Optimizations in ADF is based on a quasi Newton approach [6-8], using the Hessian for computing changes in the geometry so as to make the gradients vanish. The Hessian itself is initialized (for instance based on a force field) and updated in the process of optimization. The optimization in delocalized coordinates is mainly based on the works by Marcel Swart [225]. The initial implementation relied heavily on [QUILD]. In ADF2010 and 2012, however, many algorithms have been reimplemented, including generation of delocalized coordinates, transformation of the computed step from delocalized to Cartesian coordinates, etc..

Several subkeys in the geometry block can be used for control of the Geometry Optimization procedure and related strategy parameters.

```
GEOMETRY
 Optim {Delocal/ Cartesian / Internal } {All / Selected}
 Branch {New / Old}
 Iterations Niter {Niter2}
 Hessupd HessUpdate
 Converge {E=TolE} {Grad=tolG} {Rad=TolR} {Angle=tolA}
 Step {Rad=MaxRadStep} {Angle=MaxAngleStep}
      {TrustRadius=MaxRadius}
 DIIS {N=NVect} {CYC=Ncyc}
 Externprogram externprog.exe coord=coords.inp
               energy=energy.out grad=grads.out
 Inithess inithessian.file
End
```

Optim

```
GEOMETRY
 Optim {Delocal/ Cartesian / Internal } {All / Selected}
..
END
```

Optim

`Delocal`

Optimization in delocalized coordinates (Delocal) is the default in geometry optimizations, transition state searches, and (linear) transits. Starting from ADF2007 with delocalized coordinates you can use constraints (see also CONSTRAINTS key) and restraints, and you can supply the atoms in Z-matrix coordinates. You can not use dummy atoms, ghost atoms, or alternative elements. Starting from ADF2007 the adapted delocalized coordinates are used as described in Ref. [225].

`Cartesian / Internal`

Cartesian or Zmatrix (equivalently: internal) specifies the type of coordinates in which the minimization is carried out. By default the coordinate type is applied that was used in the ATOMS key for the input of the (initial) atomic positions. (Cartesian if atoms were input in zcart format). Note that Zmatrix optimizations can only be done with the old branch of optimizations.

Cartesian optimization is allowed if the atoms were input in Z-matrix format, but no constraints (see the key GEOVAR) can then be used: *all* coordinates are optimized. An attempt to explicitly freeze variables may result in an error abort. Optimization in Z-matrix coordinates is not allowed if only Cartesian were supplied in atoms: the program does not construct a Z-matrix by itself. One should then use the zcart format: give Cartesian coordinates and supply the structure of the Z-matrix. Again, in this case you cannot use constraints.

`Selected`

For use with old branch of optimizations. Only those coordinates are optimized that are defined with the key GEOVAR.

`All`

(The default value) means that in principle all atomic coordinates will be varied. With the new branch of optimization the key CONSTRAINTS one may set constraints on the optimization. With the old branch of optimization one can use the key GEOVAR to set constraints.

Branch

```
GEOMETRY
 Branch {New / Old}
..
END
```

Branch

`Old / New`

Expert option. Specifies which branch of the code to use for making steps. Default the branch of code used depends on the optimization used. Optimization in delocalized coordinates can only be done with the new branch. Optimization in Z-matrix coordinates can only be done with the old branch. In case of Cartesian optimization default the new branch is used, but the old branch can also be used. The new branch can be used (and is default) in geometry optimizations, transition state searches, and in LT. The old branch is default in IRC and NEB, for which one can not use the new branch.

Iterations

```
GEOMETRY
 Iterations Niter {Niter2}
..
END
```

Iterations

`Niter`

The maximum number of geometry iterations allowed to locate the desired structure. The default is max(30,2*Nfree), where Nfree is the number of free variables, which is typically close to 3*N(atoms).
This is a fairly large number. If the geometry has not converged (at least to a reasonable extent) within that many iterations you should sit down and consider the underlying cause rather than simply increase the allowed number of cycles and try again.

`Niter2`

An optional second parameter that plays only a role in a LinearTransit run, see the LT section. It must not be used in other runtypes.

Hessupd

```
GEOMETRY
 Hessupd HessUpdate
..
END
```

Hessupd

`HessUpdate`

Specifies how the Hessian matrix is updated, using the gradient values of the current and the previous geometry. The methods available depend on the optimization branch being used. For both the old and new branches, the following options are available:
(i) BFGS : Broyden-Fletcher-Goldfarb-Shanno
(ii) MS : Murtagh-Sargent
(iii) FARKAS : Farkas-Schlegel, Eq. (15) and (16) of Ref. [139]
(iv) FARKAS-BOFILL : Farkas-Schlegel-Bofill, Eq. (15) and (14) of Ref. [139]
In the old branch, the following extra options are available:
(i) DFP : Davidon-Fletcher-Powell
(ii) FS : Fletcher switch
(iii) HOSHINO : Hoshino
In the new branch, the following extra option is available:
(i) BAKKEN-HELGAKER: Bakken-Helgaker, see Ref. [219]
The default is BFGS for geometry optimizations.

Converge

```
GEOMETRY
 Converge {E=TolE} {Grad=tolG} {Rad=TolR} {Angle=tolA}
..
END
```

Converge

Convergence is monitored for three items: the energy, the Cartesian gradients and the estimated uncertainty in the (chosen type of optimization) coordinates. For the latter, lengths (Cartesian coordinates, bond-lengths) and angles (bond-, dihedral-) are considered separately.
Convergence criteria can be specified separately for each of these items:

`TolE`

The criterion for changes in the energy, in Hartrees. Default: 1e-3.

`TolG`

Applies to gradients, in Hartree/angstrom. Default: 1e-3. Note, the default value has been changed in ADF2008.01, before it was 1e-2.

`TolR`

Refers to changes in the Cartesian coordinates or bond lengths, depending on in what coordinates you optimize, in angstrom. Default: 1e-2. Note that if the gradient is 10 times less than the convergence criterion for the gradient, then this convergence criterion for the step size is ignored.

`TolA`

Refers to changes in bond- and dihedral angles, in degrees. This is only meaningful if optimization takes place in Z-matrix coordinates. Default: 0.5 degree.
If only a numerical value is supplied as argument for *converge*, rather than a specification by name, it is considered to apply to the gradients (only). The other aspects (energy and coordinates) retain their default settings then.

Remarks:

1. Molecules may differ very much in the stiffness around the energy minimum. Application of standard convergence thresholds without second thought is therefore not recommended. Strict criteria may require a large number of steps, a loose threshold may yield geometries that are far from the minimum as regards atom-atom distances, bond-angles etc. even when the total energy of the molecule might be

very close to the value at the minimum. It is good practice to consider first what the objectives of the calculation are. The default settings in ADF are intended to be reasonable for most applications but inevitably situations may arise where they are inadequate.

2. The technical numerical accuracy of the calculation (e.g. numerical integration) should somehow match the required level of convergence in gradients (strict convergency criteria may require high numerical accuracy). A simple way of changing the numerical accuracy is through the NumericalQuality keyword.

3. The convergence threshold for the coordinates (TolL, TolA) is *not a* reliable measure for the precision of the final coordinates. Usually it yields a reasonable estimate (order of magnitude), but to get accurate results one should tighten the criterion on the gradients, rather than on the steps (coordinates). The reason for this is that the program-estimated uncertainty in the coordinates is related to the used Hessian, which is updated during the optimization. Quite often it stays rather far from an accurate representation of the true Hessian. This does usually not prevent the program from converging nicely, but it does imply a possibly incorrect calculation of the uncertainty in the coordinates.

Step

```
GEOMETRY
 Step {Rad=MaxRadStep} {Angle=MaxAngleStep} {TrustRadius=MaxRadius}
..
END
```

Step

Controls that changes in geometry from one cycle to another are not too large:

`MaxRadStep`

Can only be used in combination with the old branch. An upper bound on changes in Cartesian coordinates or bond lengths, as the case may be. Default: 0.3 angstrom when optimization is carried out in internal coordinates, 0.15 angstrom for Cartesian optimizations.

`MaxAngleStep`

Can only be used in combination with the old branch. Similarly this option limits changes in bond angles and dihedral angles. Default: 10 degrees.
Input for MaxRadStep, MaxAngleStep is in angstrom and degrees respectively, independently of the units used for atomic coordinates input.

Note: Optimization of ring structures carried out in internal (z-matrix) coordinates is sometimes tricky due to the ill-defined last segment of the ring. When problems arise, try Cartesian optimization or consider using smaller limits on the steps (in particular the angles) so as to prevent the program from breaking the ring beyond repair.

`MaxRadius`

Can only be used in combination with the new branch. By default, the trust radius is set to 0.01 Bohr times the number of atoms with a minimum of 0.2 Bohr. Using the key, the user can override this, setting a constant value. A conservative value is 0.2. A large system (eg 100 atoms) typically needs a larger trust radius (eg 0.8).

DIIS

```
GEOMETRY
 DIIS {N=NVect} {CYC=Ncyc}
```

```
..
END
```

DIIS `N=NVect CYC=Ncyc`

Can only be used in combination with the new branch. NVect is the number of vectors used by the DIIS interpolation method. NCYC is the number of geometry cycles run before the DIIS starts to modify the geometry steps. Default DIIS is used and default N=5 and CYC=0.

Externprogram

```
GEOMETRY
 Externprogram externprog.exe coord=coords.inp
               energy=energy.out grad=grads.out
..
END
```

Externprogram

Expert option, can only be used in combination with the new branch. Subkey EXTERNPROGRAM has been added to to allow energies and gradients to be calculated by an external program for use in a geometry optimization.

Note that you need to supply information about atomic fragments, such as the basis set, even though these are not actually used in the calculations.

When ADF is ready to perform an energy and gradient calculation, it writes the current cartesian coordinates to the file name given in the input. The format is similar to the ATOMS block in the ADF input file: it has one atom per line, with the element symbol given, followed by the x, y, and z coordinates.

ADF will then run the executable program, and then read in the energy and gradients from the file names given in the input file. The external program is thus responsible for reading the coordinates (in atomic units) written by ADF from file, generating the corresponding energy and gradients (in atomic units), and writing these to the appropriate files. ADF will then take another geometry step, and the process will repeat.

Inithess

```
GEOMETRY
 Inithess inithessian.file
..
END
```

Inithess

Can only be used in combination with the new branch. With this INITHESS subkey it is possible to read a hessian from a text files. The only argument is the name of the file containing the initial hessian. The hessian must be given in full, in non-mass-weighted cartesian coordinates, and in atomic units (hartree/bohr**2).

Transition State

A transition state (TS) search is very much like a minimization: the purpose is to find a stationary point on the energy surface, primarily by monitoring the energy gradients, which should vanish. The difference

between a transition state and a (local) minimum is that at the transition state the Hessian has a negative eigenvalue.

Because of the similarities between a minimization and a TS search most subkeys in geometry are applicable in both cases, see the Geometry Optimization section. However, practice shows that transition states are much harder to compute than a minimum. For a large part this is due to the much stronger anharmonicities that usually occur near the ts, which threaten to invalidate the quasi-Newton methods to find the stationary point. For this reason it is good advice to be more cautious in the optimization strategy when approaching a Transition State and for some subkeys the default settings are indeed different from those for a simple optimization. In addition, certain additional aspects have to be addressed.

```
GEOMETRY
 TransitionState {Mode=Mode} {NegHess=NegHess}
end
```

`NegHess`

The number of negative eigenvalues that the Hessian should have at the saddle point. In the current release it is a rather meaningless key, which should retain its default value (1).

`Mode`

Controls the first step from the starting geometry towards the saddle point: it specifies in which direction the energy is to be *maximized* while the optimization coordinates will otherwise be varied so as to *minimize* the energy. A positive value means that the eigenvector #mode of the (initial) Hessian will be taken for the maximization direction. This means: put all Hessian eigenvalues in ascending order, ignoring those that correspond to impossible movements (rigid rotations and translations, symmetry breaking) and then take the eigenvector of #mode in the remaining list.

Default: mode=1. Generally the program performs best with this default: it will simply concentrates on the mode with the lowest eigenvalue, which should of course finally be the path over the transition state (negative eigenvalue).

After the first geometry step, the subsequent steps will attempt to maximize along the eigenvector that resembles most (by overlap) the previous maximization direction.

As mentioned before, the other subkeys have the same functionality as for minimizations, but different defaults or options may apply:

Hessupd

`HessianUpdate`

Different (fewer) options apply now. The methods available depend on the optimization branch being used. For both the old and new branches, the following options are available:
(i) Powell: Powell
(ii) BFGS: Broyden-Fletcher-Goldfarb-Shanno
(iii) BOFILL : Bofill, Eq. (13) and (14) of Ref. [139]
(iv) MS : Murtagh-Sargent
In the old branch, the following extra option is available:
(i) DFP : Davidon-Fletcher-Powell
The default is Powell for transition state searches.

Step

`MaxRadStep`

Default: 0.2 angstrom for Z-matrix optimization, 0.1 angstrom for Cartesian optimization. Not used in the new branch.

`MaxAngleStep`

Default: 5 degrees. Not used in the new branch.

Note: in Transition State searches precision is often much more critical than in minimizations. One might consider using NumericalQuality Good or better.

Transition State Reaction Coordinate (TSRC)

Starting from ADF2010 it is possible to specify a reaction coordinate for transition state search via a TSRC input block, similar to QUILD. This feature is especially useful when an accurate Hessian is not available. In such a case ADF uses an approximate Hessian that can be poor when weak interactions and/or transition metals are involved. What then happens is that the mode with the lowest Hessian eigenvalue does not correspond to the reaction coordinate along which transition state is sought for, thus leading optimization in the wrong direction.

This problem can now be solved by specifying a reaction coordinate along which the transition state is sought for. Such a reaction coordinate can consist of one or more distance, valence or dihedral angle, or just a combination of vectors on certain atoms.

```
TSRC
  {ATOM i x y z {fac}}
  {DIST i j {fac}}
  {ANGLE i j k {fac}}
  {DIHED i j k l {fac}}
end
```

`i, j, k, l, x, y, z, fac`

Here, *i*, *j*, *k*, and *l* are atom indices, *x*, *y*, and *z* are corresponding components of a TSRC vector for atom *i*, and *fac* is the factor (and thus also the sign) of a particular coordinate in the TSRC. By default *fac* = 1.0.

Restrictions and notes:

The TSRC feature does **not** work in combination with the old optimization branch. In general, the old branch is no longer developed so all new features related to geometry optimization work with the new branch only.

The DIST, ANGLE and DIHED specifications should be used in combination with optimization in delocalized coordinates only (i.e. not with Cartesian).

Only one type of the keyword is allowed in a TSRC block. That is, the keys must be either all ATOM or all DIST, etc. Thus, mixing of different keywords is not allowed.

One should be careful when specifying more than one bond or angle as a TSRC. For example, suppose atom 2 is located between atoms 1 and 3. Then the following TSRC block:

```
TSRC
  DIST 1 2  1.0
  DIST 3 2 -1.0
END
```

means that the TSRC consists of two distances: R(1-2) and R(2-3). The positive direction of the TSRC is defined as increase of the distance R(1-2) and decrease of the distance R(2-3), which, in turn, means that this TSRC corresponds to atom 2 is moving along the R(1-3) axis.

Linear Transit

In a Linear Transit (LT) run you define a number of atomic coordinates (at least one) to be the LT *parameters*: these get an initial and a final value. The LT is defined as the simultaneous linear change of these parameters from their initial to their final values. This is carried out in a number of equidistant steps. In a non-linear transit calculation these may be not equidistant steps. The total number of LT *points* is specified on input. At each LT point the remaining atomic coordinates - those that are not LT parameters - may or may not be optimized: the (final) structure and energy at each LT point are computed. A Linear Transit (LT) run is therefore just a sequence of (related) constrained Geometry Optimizations.

The LT scan may be used for instance to sketch an approximate path over the transition states between reactants and products. From this a reasonable guess for the Transition State can be obtained which may serve as starting point for a true transition state search for instance.

Whenever a geometry subkey is applicable in a Geometry Optimization, it will apply in a Linear Transit run in each of the optimizations that are carried out at the distinct Linear Transit points, and the same default values apply.

Linear Transit (new branch)

A transit calculation option has been added in the new optimization branch. This is capable of performing both linear transits, and non-linear transits, and is the default when the LINEARTRANSIT or TRANSIT subkey is included in the GEOMETRY block key.

The new transit code works differently to the old: the transit is represented as a sequence of constrained optimizations. The CONSTRAINTS block key should be used to delineate the constraints applied at each stage of the transit.

To perform a linear transit, start and end values are supplied.

```
Constraints
  angle 2 1 3 start=100.0 end=120.0
End

Geometry
  Transit 4
  Optim Deloc
End
```

In the example above, 4 stages are required; ADF will interpolate the start and end values supplied for the angle between atoms 2, 1, and 3. Note that TRANSIT can now be used in place of LINEARTRANSIT, due to the more general nature of the new transit calculations.

Non-linear transits are possible, and can even be combined with linear transits in other coordinates. To perform a non-linear transit in a particular coordinate, explicit values must be given.

```
Constraints
  dist 1 2 0.8 0.9 1.1 1.15
  angle 2 1 3 start=100.0 end=120.0
End
```

```
Geometry
  Transit 4
  Optim Deloc
End
```

In the example above, 4 values are given for the distance between atoms 1 and 2. This distance constraint will be applied simultaneously with the linear transit constraints for the angle, with other degrees of freedom optimized at each stage of the transit.

It is worth noting that fixed constraints can also be used in a transit.

```
Constraints
  dist 1 2 0.8 0.9 1.1 1.15
  angle 2 1 3 100.0
End

Geometry
  Transit 4
  Optim Deloc
End
```

In this example, the angle between atoms 2, 1, and 3 will be fixed at 100.0 degrees at all stages of the transit.

Finally, it should be pointed out that fully converged constraints are used by default in the transit calculations. They do not have to be fully met in the input, but if the input geometry is far from meeting the constraints, a large, erratic first geometry step may result.

You can avoid fully enforcing constraints, by adding a CONSTRAINTS subkey to the geometry block key:

```
Geometry
  ..
  CONSTRAINTS partialconverge
End
```

In this case the constraints are not required to be fully met at each intermediate geometry, but are fully met at the converged geometries,

Linear Transit (old branch)

The LINEARTRANSIT runtype has to be specified. Additional specifications are optional.

```
GEOMETRY
 Branch Old
 LinearTransit {NPoints}
end
```

`NPoints`

The number of LT points for which an optimization will be carried out
If no value is supplied the default takes effect: 5.

There are a few obvious differences between a single optimization and a LT run. Most important is that the coordinate(s) that describe the LT path, the LT *parameters*, cannot be optimized: at each of the LT points they are frozen. This implies that technically speaking at each LT point a *constrained* optimization is carried out. One of the consequences is that the atoms coordinate type - Cartesian or Z-matrix - must also be the

optimization coordinate type. The LT parameters themselves must be defined with the key GEOVAR, see below.

It is possible to freeze *all* coordinates so that the LT run is similar to a sequence of Single Point runs. However, energy gradients will be computed at each step, so that more CPU time is spent at each LT point than for just a Single Point calculation.

The number of LT points by which the path is traced is defined by the npoints argument to the subkey LinearTransit. It is possible to execute only a subset of these points, usually with the purpose to complete the calculation by using the restart facility of ADF. In this way you can break down a very large calculation into several smaller ones, or have the opportunity to check how things have been going for the first few LT points before deciding whether a continuation is useful. This may be achieved of course by simply defining different start- and end-values for the LT parameters in a related series of calculations, but it is more comfortable to specify the complete path once and just execute parts of it at a time. This is accomplished by giving a *second* value to the *iterations* subkey in the geometry block.

```
...
iterations Niter Niter2
...
```

`Niter`

The first argument of the subkey iterations in the GEOMETRY block, controlling the maximum number of iterations allowed to reach convergence, applies now for each LT point separately.

`Niter2`

The second argument specifies the maximum number of LT *points* to calculate in this run. If omitted (default) the whole LT scan is completed. Doing only part of the scan may be combined with the restart feature, so that the remainder can be done in a continuation run. See the restart key.
A too large value of LT points is automatically adjusted: no more LT points are computed than required to complete the LT path as defined by the lineartransit subkey. A negative or zero value is not accepted and internally reset to one (1).

WARNING: if you use the QMMM functionality in combination with a Linear Transit, then only the coordinates of the true QM atoms can be used as LT parameters, no MM atoms must be involved in the LT parameter set.

Symmetry in a Linear Transit

In a Linear Transit run it is imperative that the complete Linear Transit path as defined by the parameters conforms to the specified symmetry. If such is not the case, an error will occur or possibly the program will continue but not produce correct results. Note that when no symmetry is specified in input, the initial geometry defines the specified symmetry.

Intrinsic Reaction Coordinate

The path of a chemical reaction can be traced from the Transition State to the products and/or to the reactants, using the Intrinsic Reaction Coordinate method (IRC) [9,10]. The starting coordinates should be a fair approximation of the Transition State. The final values at the endpoint(s) - reactants, products - are computed. The IRC path is defined as the steepest-descent path from the Transition State down to the local energy minimum. The energy profile is obtained as well as length and curvature properties of the path, providing the basic quantities for an analysis of the reaction path. Additional properties along the path (dipole moment, atomic charges) are computed.

Technically speaking the path is computed by taking small steps along the path meanwhile optimizing all atomic coordinates *orthogonal to it* so that, like in a Linear Transit run, a sequence of constrained optimizations is carried out. The total number of steps along the path is not known in advance. The maximum number of such steps can be set in input. If the path is not completed in the run, a Restart can be used to finish it. Each of the constrained optimizations in the run is treated as it would be in a Linear Transit run: convergence thresholds, maximum numbers of optimization iterations et cetera are set with subkeys in the geometry block.

You can set the IRC runtype by typing it in the geometry block

```
GEOMETRY
 IRC {Forward} {Backward} {Points=Points} {Step=Step}
     {StepMax=StepMax} {StepMin=StepMin} {Start=Start}
End
```

`IRC`

The runtype IntrinsicReactionCoordinate would also be recognized.

`Forward, Backward`

Specifies execution of the two possible paths from the Transition State to the adjacent local minima. By default both are computed. If Forward is specified only, the other path is turned off and similarly for Backward. For the definition of which of the two directions down from the Transition State to an adjacent minimum is 'forward' see below.

`Points`

The maximum number of IRC points computed in the run, for both paths together and including the initial (TS) central point (as far as applicable). Default 100.

`Step`

The (initial) step length when proceeding from one IRC point to another along the path. The difference between two geometries, to which the step quantity applies, is measured in mass-weighted coordinates.

The default value for step is 0.2 $(\text{amu})^{1/2}$ bohr. Larger steps reduce, in principle, the required number of IRC points from the transition state to the minimum, but usually at the expense of more optimization steps at each of the points so the net gain in computation time may not be very large, or even negative. The default size is rather conservative and in many cases you may increase it to save a few steps. However, to some extent you can leave that to the program. When going from one point to the next, the program will increase or decrease the stepsize depending on whether or not the previous point to a large number of geometry cycles to converge. The adjusting algorithm also tends to be more cautious when the successive IRC points show more drastic changes in the atomic geometrical configuration. In all cases the IRC step sizes remain between pre-set maximum and minimum values, see the next items.

`StepMax`

The maximum step length that the program will select in the step-adjusting algorithm. Default: 1.0 or 10 times the initial step length, whichever is larger.

`StepMin`

The minimum step length that the program will select in the step-adjusting algorithm. Default: 0.03 or 0.3 times the initial step length, whichever is smaller.

`Start`

Defines how the initial direction of the path is chosen to move away from the Transition State. It does *not* imply whether the first step along this direction is taken positively or negatively. See for this aspect the section about Forward/Backward IRC paths.
The admissible values for start are:
Grad: compute the gradient and take that direction right from the start. Obviously, if we start perfectly at the Transition State this will be meaningless since the gradient vanishes there completely.
Read : the initial path direction is read in with the key IRCstart, see the section IRC Start Direction.
Hess n : the initial path coincides with the n-th Hessian eigenvector (ordered by ascending eigenvalues); *n* must be an integer in the appropriate range.
The default (omission of any start specification at all) is the first Hessian eigenvector, presumably corresponding to the path over the Transition State (negative Hessian eigenvalue!).

IRC start direction

As mentioned above, the IRC path is initialized by a first step away from the Transition State. If perfect information is available this should be along the unique Hessian eigenvector with a negative eigenvalue. Therefore, it is preferable to supply (with a restart file) a good approximation of the Hessian at the Transition State. This can be computed in a Frequencies run. In many cases the automatic internally generated (force field based) Hessian will not severely disturb the procedure and may only require a few more initial search steps for the right direction to take, while saving a potentially expensive Frequencies calculation.

If you decide to use a precalculated Hessian, then usually the approximate Hessian resulting from a Transition State run will be good enough. The latter approach is more attractive of course since the TS run will usually be done anyway, as a preliminary to the IRC run, while an additional Frequencies run would be very demanding. At the other hand, Transition State runs often require a preceding Frequencies run. In such case, the Frequencies result file may be used both for the TS run and for the IRC run. The fact that the Frequencies run may have been performed not at the exact TS may affect slightly the adequacy for using it as a start-up for the IRC run, but this is likely not significant.

In some case you may want to specify the initial direction of the IRC path explicitly. This is done as follows:

```
IRCSTART
 data
 data
 ...
end
```

`IRCstart`

A block-type key. The data in the data block are values for *all* atomic coordinates (Cartesian or Z-matrix, as the case may be) that are not frozen and not (by geovar) explicitly instructed to remain equal. All such coordinate data together define a direction vector in the space of all (free) coordinates, which then serves as the initial segment of the IRC path.

Note that only a direction vector is defined here: the *size* of the total vector plays no role.

Furthermore, the initial step may be in the positive or negative direction along the so-defined initial path, see the section Forward / Backward IRC paths.

Forward / Backward IRC paths

Obviously there are two IRC paths down the transition state: Forward and Backward. We would have liked to chemically define forward and backward by determining (in advance!) which of the endpoints is reactants and which products. This is not well doable in practice. Therefore we define the directions in terms of the initial path vector: select simply the atomic coordinate with the largest (absolute) change in the initial vector

and define *Forward* as the direction in which this coordinate *increases* and *Backward* as the direction in which it *decreases*.

Climbing-Image Nudged Elastic Band

The reaction path can be found by simultaneous optimization of a number of replicas of the system in question starting from some rough approximation [159]. In the simplest case, implemented in ADF, the initial approximation is just a polynomial interpolation between initial and final states (see keyword geovar). The images are optimized not independently of each other but, in fact, forces on each image depend on its neighbors. At each step the forces parallel to the reaction path are eliminated and a so-called spring force is added which keeps the image in the middle between its neighbors. This does not let images slide to the initial or final reaction state and ensures that they are evenly distributed along the reaction path. There are also options to distribute images more densly near the transition state (energy-dependent spring force).

Below is the list of NEB options:

```
GEOMETRY
  CINEB {NumImages}
  {NEBSPRING  Nspring  Spring Spring2 Spower}
  {NEBOPT OptMethod}
  {NEBECONO}
  {NOCLIMB}
  {NONEBOPTENDS}
End
```

`CINEB`

The runtype. Nudged will also be recognized.

`NumImages`

The number of NEB images excluding initial and final stated. The default is 8.

`NEBSPRING Nspring Spring Spring2 Spower`

Nspring determines the type of spring used, which, in turn, determines which of the spring parameters are used:

1: constant spring, spring=Spring
2: exponential scaling, spring = Spring+Spring2*exp((dE-dEmax)**Spower)
3: power scaling, spring = Spring+Spring2*(dE/dEmax)**Spower
4: another exponential with different meaning of Spower, spring = Spring+Spring2*exp((dE-dEmax)*Spower)
5: another exponential scaling very close to #4, spring = Spring+Spring2*(2**(dE-dEmax))**Spower)

Units for Spring and Spring2 are Hartree/bohr. Default values, when NEBSPRING is not present in the input, are 1 for Nspring and 0.1 for Spring. If NEBSPRING is specified with Nspring 1 then the Spring parameter is required. If Nspring>1 is specified then also Spring2 and Spower are required.

`NEBOPT OptMethod`

Specifies the optimization procedure.
Since NEB is conceptually different from simple optimization, not all or not always options used in simple geometry optimization applicable to NEB. There are two optimization modes available for NEB: global (covering all images simultaneously) and local (that is, local to each image). Each method has its pro's and con's. The global method usually converges in fewer steps than local because its Hessian

takes into account all degrees of freedom at once. On the other hand, the size of the matrix may become too large for moderate-size system, which might lead to problems (one dimension of the Hessian matrix may be as large as N(atoms)*3*N(images)).
There are two geometry update methods available for both global and local optimization: Quasi-Newton and Conjugate-Gradient. Quasi-Newton is the preferred method at all times. NEB optimization in Z-matrix coordinates is not available at this time.

OptMethod can take any of the following values:

`GLOBALQN`: global Quasi-Newton.
`QN`: Local Quasi-Newton. The preferred (and default) method.

`NEBECONO`

(local optimization only) Requests that when at some point an image's geometry converges this image will not be recalculated in subsequent steps. This option can be used to speed up calculation in the end when some images have already converged. Please note though that even if an image has converged at one point it may become "un-converged" at a subsequent point due to increase in spring force (which is determined by position of the image with respect to its neighbors). This option is irrelevant in case of the global optimization because then the convergence state of a single image is not determined.

`NOCLIMB`

Switches off the climbing-image feature. This option is generally not recommended and exists for debugging and troubleshooting purposes.

`NONEBOPTENDS`

Do not optimize geometries the initial and final reaction states during NEB optimization.

Recommendations concerning the NEB method.

Preparing input
Please pay attention to the following points when preparing input for a NEB calculation:

- If an approximated transition state is known then try to use this information when preparing the input: do not just specify the initial and final state coordinates in the GEOVAR input section but also add some values in between to take advantage of the higher-order interpolation of the initial reaction path approximation.
- Try to optimize geometries of the initial and final reaction states (the end-points) as good as possible. ADF will by default optimize them too but doing this in advance can save quite a bit of time.
- If you do not want to optimize the end-points during NEB optimization you can use the NoNEBOptEnds input keyword and you know that the end-point geometries do **not** correspond to local minima, then you **must** make sure that they lie on the reaction path. If one (or both) of the end-points lie off the path then this may result in the images next to them sliding downhill behind the end-points, which will inevitably break the optimization.
- **Choosing an optimization method.** There are several optimization methods used in NEB but you should probably use one of the two Quasi-Newton methods: local (default) or global (*NEBOPT GlobalQN* input option). The main difference in functionality between the wto methods is that the Climbing image technique is possible only with the GlobalQN method. In the default (local) method, the images are optimized in such a way that the cartesian distance between them is always constant. This means that the image with the highest energy is not necessarily the transition state. With the GlobalQN method, you are guaranteed that the image with the highest energy is at the point of the maximum energy on the reaction path.

Problems during optimization
Many problems may be avoided if you follow the recommendations above. If, however, you did follow the recommendations and still have problems then please read below. Here follows a list of common problems with possible solutions:
Optimization stops with a message that the angle has become too small.

- Provided that the end-points **are** local minima, this may still happen if the initial guess for the reaction path was too rough an approximation. This usually result in very large forces on some of the images, which may result in very large steps. This is not a problem in itself but a problem may be that neighboring images get significantly different steps. If this happens, the NEB chain becomes jagged. This may get quickly out of hands if the Cartesian distances between images are comparable with the steps taken during optimization. The cure in this case is to either reduce the number of images (to increase distances between them) or to decrease the max step size (the *STEP RAD*= parameter).
- Another reason for the angle becoming too sharp is that the reaction path is very complex. In this case, it may help to use more images to "smoothen" it.

In all other cases it is recommended to contact the SCM support and specify exactly what went wrong and send along the input and output files. It is recommended to use the *DEBUG NEB* input keyword, which produces extra debugging information. Doing so will speed things up a lot because we won't have to repeat your calculation.

Special Features

Initial Hessian

In a Geometry Optimization (or Transition State search) the Hessian matrix - second derivatives of the energy with respect to changes in coordinates - is updated while the program steps around in an attempt to find the (local) energy minimum. The quality of the initial Hessian may have a considerable impact on the required number of steps to reach geometric convergence.

By default the initial Hessian is read from a restart file (see the key RESTART) if a restart file is present. Otherwise it is constructed from a force field [11] that is implemented in the program. When using the new branch, the initial Hessian is calculated by the same method as in QUILD [[225]]. In case of the old branch of optimization one can modify this with the key HESSDIAG. With the new branch of optimization it is possible to read an Hessian from an input file, see subkey INITHESS of the key GEOMETRY.

Constrained optimizations, LT (new branch)

The key CONSTRAINTS can only be used in case of the New branch for optimization of coordinates. The input for this key is very similar to that of the RESTRAINT keyword. The key CONSTRAINTS can, however, also be used to constrain Cartesian coordinates. Note that the key RESTRAINT and freezing of coordinates with the GEOVAR key can also be used in the New branch for optimization of coordinates. Currently, the New branch for optimization can only be used in geometry optimizations and transition state searches.

The CONSTRAINTS keyword allows geometry optimizations with constraints for the distance between two atoms, an angle defined by three atoms, a dihedral angle defined by four atoms, or to freeze a block of atoms:

```
CONSTRAINTS
 ATOM Ia1 {Xa1 Ya1 Za1}
 COORD Ia1 Icoord {valcoord}
 DIST Ia1 Ia2 Ra
 ANGLE Ib1 Ib2 Ib3 Rb
```

```
 DIHED Ic1 Ic2 Ic3 Ic4 Rc
 BLOCK bname
end
```

The ATOM, COORD, DIST, ANGLE, and DIHED constraints do not have to be satisfied at the start of the geometry optimization. The ATOM and COORD constraints can only be used in Cartesian optimizations while all other constraints may only be used with delocalized coordinates.

Important note about constraints and symmetry: if the system has some symmetry, which is going to be maintained, that is if a non-NOSYM symmetry is specified on input or if the initial geometry has non-NOSYM symmetry and "Symmetry NOSYM" is not specified, then the constraints specified here must also be symmetric. This means that if, for example, two distances are equivalent due to symmetry then and are going to be constrained then both must be specified in the CONSTRAINTS input block. It is incorrect to specify only one of them and hope that ADF will take care of the other automatically.

ATOM

When *ATOM* is specified, the Cartesian coordinates of atom Ia1 are constrained to: Xa1 Ya1 Za1. The atom number should be given in *Input order*; the value for the coordinates in Angstrom. Optionally one can give the three Cartesian coordinates a value. This key can only be used in Cartesian optimizations.

COORD

When *COORD* is specified, the Icoord Cartesian coordinate of atom Ia1 is constrained to: valcoord The atom number should be given in *Input order*; the value for the coordinate in Angstrom. Icoord is 1 means the x-coordinate. Icoord is 2 means the y-coordinate. Icoord is 3 means the z-coordinate. Optionally one can give the Cartesian coordinate a value. This key can only be used in Cartesian optimizations.

DIST

When *DIST* is specified, the distance between atoms Ia1 and Ia2 is constrained to the value Ra in Angstrom.

ANGLE

When *ANGLE* is specified, the angle between atoms Ib1, Ib2 and Ib3 (Ib1-Ib2-Ib3) is constrained to the value Rb in degrees.

DIHED

When *DIHED* is specified, the dihedral angle between atoms Ic1, Ic2, Ic3 and Ic4 (Ic1-Ic2-Ic3-Ic4) is restrained to the value Rc in degrees. The dihedral angle Ic1-Ic2-Ic3-Ic4 is defined in the same way as for the Z-matrix in ADF. The dihedral angle is projected onto the [0,2π] interval, so there should be no difference between specifying -30° or 330°.

BLOCK

Block constraints allow the internal degrees of freedom of a block of atoms to be frozen, so that the block moves as a whole. To apply block constraints, you add block labels to atoms in the Atoms block, and then add the block constraint in the Constraints input block:

```
ATOMS
  1.C        -0.004115   -0.000021    0.000023 b=b1
  2.C         1.535711    0.000022    0.000008 b=b2
  3.H        -0.399693    1.027812   -0.000082 b=b1
  4.H        -0.399745   -0.513934    0.890139 b=b1
  5.H        -0.399612   -0.513952   -0.890156 b=b1
  6.H         1.931188    0.514066    0.890140 b=b2
```

```
  7.H          1.931432    0.513819   -0.890121 b=b2
  8.H          1.931281   -1.027824    0.000244 b=b2
END

CONSTRAINTS
  BLOCK b1
  BLOCK b2
END
```

Note: The following restrictions apply to optimization with block constraints:

- block constraints may only be used with delocalized coordinates;
- there should be no other constrained coordinates used together with block constraints although this may work in many situation;
- the user should absolutely avoid specifying other constraints that include atoms of a frozen block.

Constrained optimizations, IRC, NEB, LT (old branch)

GEOVAR

The block key GeoVar is used

- In case of the old branch of optimizations
- To put restrictions on the number of coordinates that are varied and
- To define Linear Transit or NEB parameters and assign them initial and final (and in case of NEB - also intermediate) values.

geovar can also be used to assign (initial) values to coordinates without other implications, but this feature is accidental.

In the input section of atomic coordinates (key ATOMS) identifiers (names) may be used rather than numerical values wherever coordinate values are expected: x, y, z in case of Cartesian coordinate input; r, q, f in case of internal coordinates. All such identifiers must then be specified under geovar and assigned a value.

```
GEOVAR
 Name Data
 ...
end
```

`Name`

An identifier that can be used in place of a numerical value for one or more of the atomic coordinate values under atoms.

`Data`

Either of the following three formats:

1 A single value simply assigns the value to the corresponding atomic coordinate(s).

2 Two or more values (separated by a delimiter) imply that the corresponding atomic coordinate is a Linear Transit or a Nudged Elastic Band parameter. For Linear Transit, only two values are allowed in which case they specify initial and final values of the LT path, respectively. In case of a NEB calculation one can provide more than just initial and final values to get a better initial approximation of the reaction path. It is generally recommended (and in some cases necessary) to use more values. Intermediate images will be obtained by polynomial interpolation of degree N-1, where N is the number of values.

3 A single value followed by a letter F assigns the value to the corresponding atomic coordinates *and* specifies that these coordinates are frozen: they will not be optimized.

As regards the optimization of coordinates other than the frozen ones and the LT or NEB parameters, the meaning and effect of the input under geovar depends on the subkey *optim* in the geometry block:

If selected has been set, optimizations are carried out only for the coordinates that are referred to under geovar (and that are not Linear Transit parameters or Frozen). All coordinates that were input as simple numerical data under atoms are kept frozen then.

Alternatively, if selected has *not* been set (: all, the default) all atomic coordinates are optimized (except the Linear Transit parameters and the explicitly frozen coordinates). In that case, each assignments under geovar other than to freeze the coordinate or to define it as a Linear Transit parameter simply assigns an initial value to the pertaining coordinates. In this respect it is not different from typing the numerical value directly in the atoms block, except for the next aspect. Please note that whereas during a linear transit run the LT parameters are *never* optimized, the NEB parameters specified in the *geovar* section are always optimized.

The same identifier may be used for two or more coordinates in atoms. If they are varied (i.e. if they are not frozen) they will forcibly be kept equal throughout the optimization so that they constitute only one degree of freedom. Don't use the same geovar variable for coordinates that belong to atoms of different chemical types or to different *types* of coordinates (an angle and a bond length for instance). It is not sensible to do so and it will very probably lead to an error abort or to stupid results.

It is allowed to put as atomic coordinate under atoms *minus* a geovar variable name, i.e. the name preceded directly by a minus sign (without a blank in between!). The coordinate will then be kept equal, but with opposite sign, to coordinates that are defined by the same variable without the minus sign. The initial (and final, in case of a LT or NEB run) value for that coordinate is the negative of the geovar value.

Coordinate types

Restricted optimizations are performed by freezing certain coordinates, by explicitly referring to one and the same geovar identifier for different coordinates, or by using the selected option. In addition, they are implicit in each Linear Transit or IRC run. All restricted optimizations demand that the *type* of optimization variables (Cartesian or Z-matrix) equal the type of coordinates used in atoms. zcart input under atoms is considered to be *Cartesian* in this respect.

If this is violated in a Linear Transit calculation the program will abort. If you apply the Frozen option under geovar, while not using the same coordinate type for atoms as for optimization, an error will occur. If you refer to the same geovar identifier for distinct coordinates while the atoms and the optimization types of variables do not match, the program will continue and assume that you only have assigned the same *starting* values to the pertaining coordinates. No equality constraints will be in effect then during the optimization.

Linear combinations of constraint

It is often desirable to carry out a geometry optimization in internal coordinates where two or more of the coordinates are required to maintain constant values relative to each other. The most simple case, where two internal coordinates a kept equal can be achieved by referencing both coordinates to a single variable in the GEOVAR block. Ensuring a different relationship, such as forcing one bond length in a molecule to be 0.5 Angstrom longer than another is more difficult to achieve. These kind of constraints can often be be managed through the creative use of dummy atoms but this is generally laborious and not always possible at all.

This key can not be used in case of optimization in delocalized coordinates.

The LINEARCONSTRAINTS keyword allows geometry optimizations with constraints defined by arbitrary linear combinations of internal coordinates to be performed quite straightforwardly. The keyword allows the

linear combination to be constrained or used as part of a linear transit calculation with the constrained value being stepped as would a variable from the GEOVAR block.

```
LINEARCONSTRAINTS
 Name1 Data1
 VAR11 Coef11
 VAR12 Coef12
 ...
 SUBEND
 Name2 Data2
 VAR21 Coef21
 VAR22 Coef22
 ...
 SUBEND
 ....
end
```

`Namex`

Identifier of the xth linear constraint.

`Datax`

Either of two formats
1) A single number, giving the value of the xth constraint
2) Two numbers, the first as in 1) and the second the final value in a linear transit calculation. The LINEARTRANSIT keyword must be present in the GEOMETRY block.

`Varxy`

Name of the yth variable in that is part of the xth constraint. Varxy must be defined in the GEOVAR block.

`Coeffxy`

Coefficient of Varxy in the linear combination defining the constraint. Thus:
`Coeff11*Var11 + Coeff12*Var12 ...= Data1`
The summation must be consistent with the initial values of the Varxy.

This procedure is only possible when the geometry is defined in terms of internal coordinates. Although the program will not complain, it makes no sense to have linear combinations containing both bonds and angles of course.

The number of linear constraints must be less than or equal to the number of entries in the GEOVAR block. Only internal coordinates involving QM atoms can be included at this stage.

As a geometry optimization is run, the force acting on the linear constraints will be printed immediately after the forces on the internal coordinates. The constraint forces may be useful in the search for a transition state for instance.

Z-matrix and symmetry

If the structure of the Z-matrix does not reflect the symmetry of the molecule *and* constraints are applied the program may encounter algorithmic problems to match all demands. As a result some of the *frozen* coordinates may be found to change. Usually these changes are very small. To cure this: build the Z-matrix in a symmetric way.

Summary of geovar, optim, and atoms

For *unconstrained* optimization: don't use geovar, apply *optim* if *Cartesian* optimization is required while the data in the atoms block was in z-matrix format or when *z-matrix* optimization is required while the atoms input was in zcart format. Provide the atomic coordinates (atoms) directly as numerical data.

For optimizations where only very few coordinates are frozen: use geovar to set a few coordinates to frozen and/or to enforce equality of optimization coordinates whose values should remain equal. Don't use *optim*: the type of optimization coordinates - Cartesian or internal - must be identical to what is used in the atoms input part because you're using constraints now. In the atoms section, use identifiers for the frozen coordinates and for those that should satisfy equality conditions; use numerical input for all other (optimization) coordinates.

For very limited optimization: turn on the selected option with *optim* and assign with geovar initial values to the coordinates that you want to optimize. In the atoms input use identifiers for these coordinates. The numerical input coordinates are kept frozen automatically now.

Initial Hessian

By default the initial Hessian is read from a restart file - see the key RESTART - or constructed from a force field [11] that is implemented in the program. In the latter case the user can modify the so-generated initial Hessian in four ways:

1. By setting all diagonal elements to some constant.
2. By defining *three* constants, one for distances (or Cartesian displacements, as the case may be), one for bond angles, and one for dihedral angles. All diagonal elements of the Hessian are adapted accordingly.
3. By supplying a list of diagonal values.
4. By giving diagonal-Hessian values for one or more specific coordinates.

For each element *i* for which a diagonal Hessian value Hii is supplied the off-diagonal elements Hij, (all j i≠j) are set to zero.

A combination of the above options is possible. The rules of how combinations are interpreted by the program are:

- The program first initializes the Hessian using the force field (or restart data).
- If a single constant (*1*) or three constants (*2*) are supplied, all diagonal elements are adjusted (and all off diagonal elements are set to zero).
- If a list of diagonal values is supplied (*3*), this overrides the first so many values of the diagonal. Such a list is not required to cover *all* diagonal elements. If the list is shorter than the dimension of the Hessian, i.e. the number of atomic coordinates, only the first so many elements will be adjusted.
- If any individual elements are supplied specifically (*4*), their values are replaced in the diagonal defined thus far.

All input values of the Hessian are in units of Hartree/bohr2 for Cartesian coordinates and bond lengths. Hartree/radian2 for bond angles and dihedral angles.

The first 3 options are controlled by the key HESSDIAG:

```
HESSDIAG {General}
{ List
end }
```

`HESSDIAG`

A *general* key: it has either an argument (General), or a data block (List). It is also possible to supply the argument *and* the data block, but this requires that the continuation symbol (&) is given after the argument, separated from the argument by at least one blank.

`General`

Must be either a single numerical value, or one or more named specifications of options, in the format optionname=value.

If a single numerical value is given, this value is assigned to all the options that are available. If the named-option format is applied, any named options that are not found get the value 1.0.
The options are: rad=radvalue to assign a value to all Hessian diagonal elements that refer to distance coordinates (bond length in case of z-matrix coordinates, Cartesian coordinates otherwise),

ang=angvalue to assign a value to all elements that refer to bond angles, and finally dih=dihvalue for dihedral angles.
ang and dih are not significant in Cartesian optimizations.

`List`

A list of numerical values, which may expand over any number of lines. If *n* numbers are supplied, they are assigned to the first *n* diagonal elements of the Hessian. The remaining diagonal elements, if any, are not effected. The maximum number of Hessian diagonal elements equals the number of atomic coordinates.

The force field derived initial Hessian can be printed for inspection. Type in input:

```
HESSTEST
```

ADF will construct and print the initial Hessian and then abort.

Hessian values for selected coordinates

The diagonal elements for selected free coordinates can be given if these free variables are named in the geovar block.

```
GEOVAR
 Varname Data H=HessValue
end
```

`Varname, Data`

The name of the variable and any data as discussed in the sections above: assignment of initial value, final value (in case of a Linear Transit run), or a Frozen specification.

`HessValue`

The value for the diagonal element of the Hessian associated with that variable. All atomic coordinates that are defined by this variable will get the HessValue as diagonal element in the initial force field. Specification of a HessValue for a frozen coordinate or a Linear Transit parameter is meaningless.

Restrained optimizations

With the old branch of optimizations, the only way to constrain distances, angles or dihedral angles within a geometry optimization is by using a Z-matrix, and freezing that particular coordinate. With the key RESTRAINT it is possible to select **any** coordinate (distance, angle, dihedral), irrespective of the coordinates used, and restrain this coordinate.

Note the difference between *constrained* and *restrained* coordinates. At every step in the geometry optimization, the value of a *constrained* coordinate should match exactly a predefined fixed value. On the other hand, with *restraints*, a potential is added to the potential energy in order to satisfy the *restraint*, which

means that the *restraint* does not have to be satisfied exactly. For example, one can start with a geometry in a geometry optimization run in which the restraint is not satisfied.

The RESTRAINT keyword allows geometry optimizations with restraints for the distance between two atoms, an angle defined by three atoms, a dihedral angle defined by four atoms, and (or) a distance difference defined by four atoms:

```
RESTRAINT
 DIST Ia1 Ia2 Ra {[Aa] [Ba]}
 ANGLE Ib1 Ib2 Ib3 Rb {[Ab] [Bb]}
 DIHED Ic1 Ic2 Ic3 Ic4 Rc {[Ac] [Bc]}
 DD Id1 Id2 Id3 Id4 R0 [{Ad} {Bd}]
end
```

`DIST`

When *DIST* is specified, the distance between atoms Ia1 and Ia2 is restrained to the value Ra. The atom numbers should be given in *Input order*; the value for the distance in Angstrom. The Aa and Ba values are mere technical values, that don't have to be specified (in fact, recommended not to change these values); the default values of 2.0 resp. 0.1 have been chosen on sensible grounds.

`ANGLE`

When *ANGLE* is specified, the angle between atoms Ib1, Ib2 and Ib3 (Ib1-Ib2-Ib3) is restrained to the value Rb. The atom numbers should be given in *Input order*; the value for the angle in degrees. The Aa and Ba values are mere technical values, that don't have to be specified (in fact, recommended not to change these values); the default values of 1.0 resp. 0.1 have been chosen on sensible grounds.

`DIHED`

When *DIHED* is specified, the dihedral angle between atoms Ic1, Ic2, Ic3 and Ic4 (Ic1-Ic2-Ic3-Ic4) is restrained to the value Rc. The atom numbers should be given in *Input order*; the value for the angle in degrees. The Aa and Ba values are mere technical values, that don't have to be specified (in fact, recommended not to change these values); the default values of 0.5 resp. 0.1 have been chosen on sensible grounds. The dihedral angle Ic1-Ic2-Ic3-Ic4 is defined in the same way as for the Z-matrix in ADF. The dihedral angle is projected onto the [0,2π] interval, so there should be no difference between specifying -30° or 330°.

`DD`

When *DD* is specified, it is possible to restrain a difference between two distances: R0 = (r1 - r2) - (r3 - r4), where r1..r4 are positions of four atoms Id1..Id4 and R0 is the final restrained property. The functional form and meaning of Ad and Bd is the same as with plain distance restraints. Atoms Id1..Id4 need not all to be different.

Symmetry versus constraints

The symmetry of the atomic system defined by the input Schönfliess symbol is preserved during optimization. If the input information (which coordinates are kept frozen and which are optimized) conflicts with the symmetry, the result is unpredictable. For example, if two distances are equivalent due to symmetry then and are going to be constrained then both must be specified in the CONSTRAINTS input block. It is incorrect to specify only one of them and hope that ADF will take care of the other automatically.

Input specifications that are in conflict with the point group symmetry may lead to an error or a non-converging optimization.

Frequencies

Harmonic frequencies can be computed in ADF either numerically or analytically. The frequencies are computed numerically by differentiation of energy gradients in slightly displaced geometries [12, 13]. The analytical second derivatives implementation in ADF is based on [208, 209, 210].

Analytical Frequencies

The frequencies are calculated analytically by specifying the **AnalyticalFreq** block keyword (see below for more details). The analytical frequencies are as accurate as the numerical frequencies for the same integration accuracy, but can be up to 3 to 5 times quicker to compute, depending on the molecule, integration grid parameters, and choice of basis set. The analytical frequencies are fully parallelized and linearly scaled.

Note: The analytical calculation of frequencies in case of ZORA and frozen cores contains a bug in all version up to and including ADF2006.01b. In this case the numerical frequencies are more reliable. The analytical calculation of frequencies in case of ZORA and all electron basis sets does not give problems. In ADF2010 the accuracy of the calculation of analytical frequencies in case of ZORA using the large QZ4P basis sets for heavy elements, like uranium, has been improved a lot compared to ADF2009 and before. The calculated numerical frequencies were already reliable in these cases.

Calculating the analytical frequencies requires the solution of the Coupled Perturbed Kohn-Sham (CPKS) equations, which is an iterative process. This part of the process is of order 3 x *number of atoms*, and is generally the main bottle neck in calculating the frequencies. (The immediate result of the solution of the CPKS equations is the U1 matrix, the components of which are closely related to the derivatives of the MO coefficients. One of the adjustable parameters in the input of an analytical frequencies calculation can be used to control the accuracy of the U1 matrix components.)

One disadvantage in calculating analytical frequencies is that the range of exchange-correlation functionals is limited. (This is because derivative formulas have to be derived for each exchange-correlation functional in ADF, which is not an straight forward task). Here are the currently available functionals:

LDA: XONLY, VWN, STOLL, PW92
Exchange GGA: Becke88, OPTx, PBEx, rPBEx, revPBEx
Correlation GGA: LYP, Perdew86, PBEc
XC GGA shortcuts: BP86, PBE, RPBE, revPBE, BLYP, OLYP, OPBE

Any functional not mentioned above is not implemented, including PW91 and Hartree-Fock.

To calculate the frequencies analytically, include the block key word ANALYTICALFREQ. Subkeys are available, but in general, to calculate the frequencies, no subkeys are required, and including the following in your run file is sufficient:

```
AnalyticalFreq
End
```

Unlike the numerical frequencies, the analytical frequencies can be computed immediately after a geometry optimization by including both block keywords in the same input file.

```
Geometry
 ... Geometry optimization options here ...
End

AnalyticalFreq
End
```

A note of caution: For accurate frequencies it is especially important to also have an accurately optimized geometry. During a geometry optimization the integration accuracy is set by default to "Normal", and so the resulting frequencies will also have this level of integration accuracy while it may be desirable to have frequencies computed with a higher accuracy. One might consider using Good NumericalQuality (or BeckeGrid quality) and set the convergence criteria for the geometry optimization tighter.

The format for the AnalyticalFreq block key is:

```
AnalyticalFreq
    PRINT {eigs} {u1} {parts} {raw_freq}
    DEBUG {fit} {hessian} {b1} {densities} {numbers}
          {symmetry} {all}
    MAX_CPKS_ITERATIONS Niter
    CHECK_CPKS_FROM_ITERATION  N
    U1_ACCURACY  x
    NUC N1 N2 ... Nk
End
```

An explanation of the subkeys follow.

`PRINT`

This is primarly for debugging purposes. Choosing EIGS results in the print out of the MO eigenvectors, while U1 results in the print out of the U1 matrices. Except for small molecules this will result in a lot of data being output, and so they are not recommended. Choosing PARTS results in the print out of various sub-hessians that add up to give the final analytical hessian. RAW_FREQ gives the eigenvalues of the initial force matrix, which are essentially the frequencies before rotational and translational degrees of freedom have been removed from the force matrix.

`DEBUG`

This is for debugging purposes. The choice FIT results in the print out of information related to the calculation of the fit coefficients and their derivatives. HESSIAN results in the printing out of many of the sub-Hessians that add ups to give the final Hessian. Many more Hessians are printed out with this option that with the print parts subkey option (mentioned above). Choosing B1 gives data related to the frozen core orthogonalization coefficients and their derivatives. DENSITIES gives the integrals and moments of various densities computed during the calculation of the frequencies. Including NUMBERS results in the print out of numbers of basis functions, fit functions etc, as well as various integer arrays that are crucial to the calculation of the analytical second derivatives. SYMMETRY results in symmetry information and symmetry matrices being printed out. ALL can be used to print out all debug information.

`MAX_CPKS_ITERATIONS Niter`

Calculating the analytical frequencies requires the solution of the Coupled Perturbed Kohn-Sham (CPKS) equations, which is an iterative process. For most systems tested so far, convergence to the required accuracy in the U1 matrix is achieved within Niter=20 iterations, which is the default. If convergence is not achieved (a warning will be printed in the output if this is the case) then this subkey can be used to increase the number of iterations, although convergence is not guaranteed. The user required accuracy of the U1 matrix, as well as the ADF integration accuracy, can effect the rates of convergence.

`CHECK_CPKS_FROM_ITERATION N`

Solution of the CPKS equations is an iterative process, and convergence is achieved if the difference between U1 matrix of successive iterations falls below a certain threshold. This key can be used to determine at which iteration the checking should start taking place. The default is 1.

`U1_ACCURACY x`

Solution of the CPKS equations is an iterative process, and convergence is achieved if the difference between U1 matrix of successive iterations falls below a certain threshold. This subkey can be used to set the threshold. The accuracy of the U1 will be 10**(-x). So, the higher the number the more accurate the U1 will be. The default is 4. While this parameter effects the accuracy of the frequencies, other factors also effect the accuracy of the frequencies, especially the ADF integration accuracy.

`NUC N1 N2 ... Nk`

By default, when calculating the frequencies analytically, the derivatives of the energy with respect to all nuclei are calculated. This gives a complete Hessian (second derivative) matrix, from which the vibrational frequencies of the molecule can be calculated. However, there may be certain cases where only derivatives with respect to a subset of all the nuclei are required. In this case it is a considerable saving in time if only a partial Hessian is calculated. With this subkey, a list of the nuclei for which the derivatives are required can be specified. However, the frequencies in this case are not the vibrational frequencies of the molecule, but may be used in guiding certain transition state searches.

Restarting Analytical Frequency jobs

Analytical frequency jobs can be restarted if a previous job did not finish. A restart can be done if the file TAPE21 has been saved without any corruption to the file. Upon doing a restart for analytical frequencies, the ADF program will check for the presence of an incomplete Hessian and/or for the presence of an incomplete U1 matrix, then attempt to figure out what more needs to be done.

The restart option may also be useful in combination with the atom selection option (i.e. by specifying a list of atoms following the keyword nuc in the analyticalfreq block key, see previous notes in this section). So for instance, you could calculate the partial Hessian for a subset of atoms of a molecule, and at a later time add another subset of atoms, or the rest of the atoms of the molecule to complete the Hessian.

Numerical Frequencies

Input options

Calculation of the numerical frequencies is specified using the FREQUENCIES keyword in the GEOMETRY block. Most of the subkeys in the geometry block are meaningless for the calculation of frequencies. Indeed, a Frequencies calculation is not a variation on optimization, but rather a sequence of Single Point runs for the equilibrium geometry and a series of slightly different geometries. By comparison of the computed gradients the force constants and hence the frequencies are computed (in the harmonic approximation of the energy surface).

```
GEOMETRY
 Frequencies {Symm} {Allowed} {Numdif=Numdif} {Disrad=drad}
             {Disang=dang} {SCANALL} {NOSCAN}
 iterations Niter
end
```

`Symm`

This switch requests that frequencies are calculated in symmetric displacements. During such a calculation first symmetric atomic displacements are constructed. The number of such displacements in each irreducible representation corresponds to the number of frequencies with the corresponding symmetry. All displaced geometries within one representation have the same symmetry, which enables us to use it to speed up the computation significantly. This is a new option and for now it only works with geometries specified in Cartesian coordinates. This option does not work correctly with a restart file. **This option does not work correctly when symmetry is explicitly specified in the input file.**

`Allowed`

Another advantage of the symmetric displacements is that only a subset of frequencies can be calculated. The ALLOWED option requests computation of only IR-visible frequencies. This option is only useful for symmetric molecules where it can be a big time-saver.

`Numdif`

Must have the value between 1 and 4 and specifies the type of numerical differentiation that is applied to compute the force constants from gradients in slightly displaced geometries: 1-, 2-, 3-, or 4-point numerical differentiation. In the case of 1-point differentiation the gradients of the displaced geometry are compared with the gradients at the input (equilibrium) geometry. In 2-point case both a negative and a positive displacement are applied, yielding much more accurate results but at the expense of more computations. This option is the default.
In certain cases the 3-point differentiation method gives better result than 2-point because it also takes gradients in the middle point into account. This is the case when geometry has not completely converged and the residual gradients are not quite close to zero. In this method, a formula is used that interpolates the second derivative (i.e. force constant) at the zero-force point. This way, the error due to a small deviation from the minimum geometry is decreased. The requirement is that the residual forces are small enough, more precisely, less that forces at displaced geometries (that is, using numdif=3 for arbitrary geometries is a bad idea).
When Numdif=4 is specified, force constants matrix will be computed by making two displacements in each direction, the standard (see drad, dang below) and twice as short. The force constant is then computed using the Romberg formula that reduces the higher-order and noise components: H(tot)=(4*H(dx/2)-H(dx))/3. Although this method requires twice as many single-point evaluations, one can probably get reliable results using lower integration accuracy, which might be faster than the default.

`dang and drad`

The displacements of the coordinates that will be varied. Dang applies to angles (bond and dihedral) in degrees and drad applies to Cartesian (x, y, z) coordinates and to bond lengths, in angstrom. Defaults: 1 degree and 0.01 angstrom.

`Niter`

In a calculation of frequencies it is the total number of (displaced) geometries for which gradients are computed. By default this is internally determined such that the calculation of frequencies can be completed. If you reduce it, the run will only partially build the matrix of force constants and a restart is required to complete the computation.

WARNING: you cannot combine a Frequencies calculation with the QM/MM feature.

`SCANALL and NOSCAN`

ADF can scan some or all normal modes after a frequency calculation to verify the corresponding frequencies. By default, the normal modes corresponding all found imaginary frequencies are scanned. This can be switched off by specifying NOSCAN. Specifying SCANALL will tell ADF to scan along **all** normal modes. These options apply only to calculations in atomic displacements, that is when SYMM is not specified. These two options are mutually exclusive, the one specified last taking precedence.

Cartesian versus Z-matrix displacements

Cartesian displacements yield usually a higher accuracy than *Z-matrix* displacements because in the former case cancellation of numerical integration errors between the different geometries is (almost always) larger.

If Z-matrix coordinates are used as the displacement variables, then make sure that no bond angles of 180 (or zero) degrees are among them. They will very probably be treated incorrectly. If your molecule has such bond angles, use dummies to redefine the coordinates or use Cartesian displacements.

Frequencies and GEOVAR keyword

The use of the GEOVAR keyword in combination with a Frequencies run implies that constraints may be applied to the displacements, even if no coordinates are explicitly frozen. If different coordinates are connected to the same variable in the GEOVAR block, only combined displacements of the atoms will be allowed that correspond to a small change in the GEOVAR variable. For this reason, the combination of the GEOVAR and Frequencies keywords is to be handled with extreme caution. If no constraints are intended, it is recommended not to use the GEOVAR keyword, but to use DEFINE instead, or to specify the coordinates explicitly

Mobile Block Hessian (MBH)

The mobile block Hessian (MBH) method [282, 283] is useful when calculating vibrational frequencies of a small part of a very large system (molecule or cluster). Calculation of the full spectrum of such a system may be inefficient and is unnecessary if one is interested in one particular part. Besides, it may be difficult to extract normal modes related to the interesting sub-system out of the whole spectrum. Using MBH it is possible to treat parts of the system as rigid blocks. Each block will usually have only six frequencies related to its rigid motions compared to 3*N for when each atom of the block is treated separately.

The calculation of frequencies using mobile blocks is invoked by specifying FREQUENCIES and MBH keywords at the same time in the GEOMETRY input block:

```
GEOMETRY
  FREQUENCIES
  MBH blockname1 blockname2 ...
End
```

The names of blocks must correspond to the ones specified in the *b*= parameter in the ATOMS input block.

The second derivatives with respect to block motions are calculated by numerical differentiation. Since the number of degrees of freedom is reduced, the number of second derivatives is reduced as well. Therefore the MBH can realize a speed-up in the calculation of the Hessian compared to a full numerical frequency calculation.

MBH for partially optimized structures

MBH is suitable to calculate frequencies in partially optimized structures. Assume a geometry optimization is performed with the BLOCK subkey in the CONSTRAINTS section [see constrained geometry optimizations]. During the geometry optimization, the shape of the block is not changed. The internal geometry of the block is kept fixed, but the block as a whole can still translate or rotate.

At the end of such a partial geometry optimization, the position and orientation of the block are optimized. The total force on the block is zero. However, there might be still some residual forces within a block, since those degrees of freedom were not optimized.

A traditional frequency calculation performed on this partially optimized structure might result in non-physical imaginary frequencies without a clear interpretation. Therefore one should use an adapted formulation of normal mode analysis: the Mobile Block Hessian method. The MBH does not consider the internal degrees of freedom of the block (on which residual forces) apply, but instead uses the position/orientation of the block as coordinates. In the resulting normal mode eigenvectors, all atoms within the same block move collectively.

Of course, MBH can also be applied on a fully optimized structure.

Accuracy

The second derivatives with respect to Cartesian displacements (3 translations) of the free atoms (atoms not belonging to any block) and those with respect to block motions (3 block translation/3 block rotations) are calculated by numerical differentiation of the gradient. The accuracy of the second derivatives is mainly influenced by the accuracy of the gradient evaluation (e.g. accuracy numerical integration) and the step size in the numerical differentiation. The parameters DISRAD and DISANG can be specified to set the step size for Cartesian displacements (translations) and block rotations respectively. The step size for angles is automatically scaled with the block size.

```
FREQUENCIES {DISRAD=drad} {DISANG=dang}
```

The default values for drad and dang are the same as in the case of a standard numerical frequency run.

MBH Notes

Blocks consisting of 1 or 2 atoms are not supported and are ignored. This means that each atom of such a block is treated separately.

At this moment, it is not possible to calculate IR intensities with the MBH method.

The printed output of the MBH normal mode analysis lists all frequencies, including the ones corresponding to rotations and translation of the molecule as the whole. Note that while frequencies corresponding to translations should always be close to zero, the magnitude of the ones corresponding to rotations depends on how well the geometry is optimized.

Debug information can be obtained if one includes DEBUG MBHNormalModes.

Thermodynamics

At the end of a completed Frequencies calculation, a survey is given of thermodynamic properties: Heat Capacity, Internal Energy, Entropy. The computed results assume an ideal gas, and electronic contributions are ignored. The latter is a serious omission if the electronic configuration is (almost) degenerate, but the effect is small whenever the energy difference with the next state is large compared to the vibrational frequencies.

```
THERMO {P=pressure} {T=temp1 {temp2}} {nT=nT}
```

`pressure`

The Pressure in atmospheres. Default value: 1.0. A zero or negative pressure is adjusted by the program to a (very) small number 1e-4

`temp1,temp2`

The endpoints of the Temperature range (in K), for which the results are computed. By default only room temperature is considered (298.15 K).
If the option T= is used and only one value supplied (temp1), then temp2 is assumed to be equal to temp1.
A zero or negative temparture is adjusted by the program to a (very) small number 1e-4

`nT`

The number of steps by which the temperature interval is scanned. By default it is computed by the program from the temperature range (temp1, temp2), such that the step size is as close as possible to 10 K. Note that the number of temperatures for which the calculations are done is one more than the number of temperature *steps*.

The thermal analysis is based on the temperature dependent partition function. The energy of a (non-linear) molecule is (if the energy is measured from the zero-point energy)

$$E/NkT = 3/2 + 3/2 + \sum_j^{3N-6}[\, h\nu_j/(2kT) + h\nu_j/(kT(e^{h\nu_j/(kT)}-1))\,] - D/kT \quad (5.1.2)$$

The summation is over all harmonic ν_j, h is Planck's constant and D is the dissociation energy

$$D = D_0 + \sum_j h\nu_j/2 \quad (5.1.3)$$

Contributions from low (less than 20 1/cm) frequencies to entropy, heat capacity and internal energy have always been excluded from the total values for thermodynamical quantities listed before. In ADF2013 they are also listed separately so the user can add them if they wish.

Gibbs free energy change for a gas phase reaction

Here an example is given how to calculate the free energy change for a reaction. In the ADF output of a frequency calculation you can find the electronic bonding energy and nuclear kinetic energies, at room temperature. Example part of the ADF output of a nonlinear molecule:

```
                                                hartree              eV         kcal/mol
kJ/mol
                                      ---------------------     -----------       ----------
-----------
  Total Bonding Energy:                 -0.744356253597793        -20.2550          -467.091
-1954.31

Zero-Point Energy :      0.033172 a.u.
====================     0.902646 eV
     Temp                                                          Transl     Rotat    Vibrat
Total
     ----                                                          ------     -----    ------
-----
     298.15   Entropy (cal/mole-K):                               34.441    11.494     0.126
46.061
              Internal Energy (Kcal/mole):                         0.889     0.889    20.847
22.624
              Constant Volume Heat Capacity (cal/mole-K):          2.981     2.981     0.533
6.495
```

The nuclear internal energy = zero point energy + 3 kT + small correction term = 22.624 kcal/mol.
3 kT = 3/2 kT for rotation, and 3/2 kT for translation (i.e. 1/2 kT for each degree of freedom). The small correction term is a term due to the vibration partition function, depending on the temperature not only the ground state vibrational levels are occupied, see also Eq. (5.1.2).

The electronic internal energy = -467.091 kcal/mol.
In ADF the electronic internal energy is normally calculated with respect to (artificial) spherical averaged neutral atoms.

The electronic + nuclear internal energy U = -467.091 + 22.624 = -444.467 kcal/mol
Gas phase pV/n = RT = 8.314472 * 298.15 / 4184 = 0.592 kcal/mol
Enthalpy H = U + pV = -444.467 + 0.592 = -443.875 kcal/mol
Gibbs free energy G = H - TS = -443.875 - 298.15*46.061/1000 = -457.608 kcal/mol

For a calculation of the free energy change for reaction (Δ G), you will of course have to to this for the reactant and product molecules, and add and subtract these energies, for each molecule proportional to the number of molecules that take place in the reaction. Application of ADF for obtaining enthalpy, entropy and Gibbs free energy can for instance be obtained in Refs. [347,348]

Accuracy

Accuracy is a crucial aspect in the computation of frequencies, in particular for modes with low frequencies: the gradients at the geometries displaced along that mode will hardly change - analytically - from their equilibrium values, so numerical integration noise may easily affect the reliability of the computed *differences* in gradients. It is worthwhile to consider carefully the *size* of the displacements. At one hand they should be small in order to suppress the effect of higher order (anharmonic) terms in the energy surface around the minimum, at the other hand they should be large enough to get significant differences in gradients so that these are computed reliably.

High precision calculations where low frequency modes are involved may require high integration settings [14]. It is sometimes advisable to increase the numerical accuracy of the calculation:

```
NumericalQuality {Good|VeryGood}
```

Using 2-point differentiation rather than 1-point differentiation implies two-sided displacements of the atoms. This doubles the computational effort but in the so-computed force constants all anharmonic terms of *odd* order are eliminated. Since in general the lowest anharmonicity is third order this eliminates the first anharmonicity. Again, this is a feature directed primarily at obtaining highly accurate and reliable results.

The 3- and 4-point methods are intended to assist in special cases and as an extra check when the results obtained with the 2-point formula are not satisfactory. The 3-point formula should be used when residual forces after geometry optimization are between 0.01 and 0.0001 a.u./angstrom. In this case frequencies obtained with the 3-point formula are much closer to those that would be computed at the exact optimum geometry.

If a Frequencies calculation is carried out only to construct a good start-up Hessian for a TS search (see the restart key), accurate results are not crucial. The most important thing in such a situation is to get a fair guess for the negative eigenvalue and its associated mode, and to avoid spurious additional negative eigenvalues. We recommend to avoid the rather time-consuming standard Hessian-computing preparation run for a TS search and to lower the precision of the Frequencies run. A reasonable value should be 4.0.

Isotope Shifts of Vibrational Frequencies

To calculate isotopic shifts using ADF do the following:

- Calculate frequencies and save TAPE21 with a different name, say result.t21
- Modify the input file as follows:
    - add "RESTART result.t21" anywhere in the input file
    - create new fragment file with different mass
    - specify the fragment file in FRAGMENTS section
- Run ADF with the new input

Alternatively one can use the the key ATOMPROPS to change the masses of the atoms.

Please note that if you change the fragment file for an atom that has symmetry-equivalent ones then the new fragment file will be applied to all of the atoms.
Example: first calculate the NH3 frequencies in the C(3v) symmetry and then change H to D. This will mean that one calculates the frequencies of a ND3 molecule and not of NH2D as one might want to do. If one wants to calculate the frequencies of NH2D one first has to do a calculation with lower symmetry, say C(s), to be able to change isotope of only one of the hydrogens.

Scanning a Range of Frequencies

In ADF2006 it was already possible to request a full scan of all frequencies obtained by finite differences. This was originally done to help identify spurious imaginary frequencies that sometimes appear where one would expect a very low (nearly zero) frequency. Most frequently this happens when there is a barrier-free rotation of, for example, methyl groups.

Starting from ADF2007.01, it is possible to scan any range of frequencies calculated in the same run or found in a restart file. Note that one should not use the key SCANFREQ in combination with the key SYMMETRY. The input keyword used to request the scan is as follows:

```
SCANFREQ low high {NUM=num DISRAD=disrad}
```

`low, high`

Two values defining an interval of frequencies to scan. Frequencies that fall within the interval will be recalculated by numerical differentiation of the gradient along the respective frequency's normal mode. This means that 2*N single-point calculations with gradients will be performed, where N is the number of frequencies within the range. Imaginary frequencies are specified using negative values, which is consistent with the notation adopted within ADF.

`num`

Num is an integer number specifying how many points are to be used for numerical differentiation: 2, 4, or 6. The default value is 2.

`disrad`

Disrad specifies the step size (in Angstrom) to be made in each direction for 2-point differentiation. For the 4-point differentiation the maximum deviation from the equilibrium geometry will be twice as large as for the 2-point one and so on. The default value for disrad is the same as used for numerical frequencies.

The main advantage of this method is that single-point calculations used to obtain a force constant are performed within the same symmetry and, usually, with the same numerical integration grid, which significantly reduces the level of numerical noise and thus increases accuracy of the calculated frequency.

Moments of inertia

In case frequencies are calculated in ADF, ADF also reports the moments of inertia of the molecule in units of amu $bohr^2$ (amu = atomic mass unit).

Excited state (geometry) optimizations

See the key EXCITEDGO.

2.6 Spectroscopic properties

See also
ADF-GUI tutorial: excitation energies, vibrational frequencies
GUI manual: spectroscopic properties
Examples: IR spectra, excitation energies, response properties, NMR, ESR, EFG

IR spectra, (resonance) Raman, VROA, VCD

In ADF infrared and Raman spectroscopy can be studied for molecular vibrations. In the Born-Oppenheimer and harmonic approximations the vibrational frequencies are determined by the normal modes corresponding to the molecular electronic ground state potential energy surface. In resonance Raman spectroscopy the molecule is excited to near one of its electronic excited states, to improve the sensitivity compared to traditional Raman spectroscopy. Vibrational circular dichroism (VCD) is the differential absorption of left and right circularly polarized infrared light by vibrating molecules.

IR spectra

The IR frequencies can be calculated with the FREQUENCIES subkey of the key GEOMETRY (numerical frequencies),

```
GEOMETRY
  FREQUENCIES
END
```

or with the block key ANALYICALFREQ (analytical frequencies),

```
ANALYICALFREQ
END
```

These keys are described more extensively here: IR Frequencies.

Raman scattering

Raman scattering intensities and depolarization ratios for all molecular vibrations at a certain laser frequency can be calculated in a single run. The run type must be Frequencies, which is arranged with the FREQUENCIES subkey of the key GEOMETRY (numerical frequencies), or with the block key ANALYICALFREQ (analytical frequencies), see IR Frequencies.

The RESPONSE key is used to specify that Raman intensities are computed. The frequency dependent Raman scattering can be calculated for 1 laser frequency at the time.

```
RESPONSE
  RAMAN
  Nfreq 1
  FrqBeg Laserfreq
  [Optional Frequency/Energy Unit]
END
```

`Frequencies or wavelengths`

The number of frequencies Nfreq should be 1. With subkey Frqbeg the value of the Laser frequency value (Laserfreq) can be given. Default frequency unit is eV. This can be changed into Hartree units (a.u.) or in wavelengths (angstroms) by typing HARTREE or ANGSTROM on a separate line within the RESPONSE block, instead of [Optional Frequency/Energy Unit].

For static Raman scattering (ω = 0) use:

```
RESPONSE
 RAMAN
END
```

The Raman scattering calculation is very similar to an IR intensity calculation. In fact, all IR output is automatically generated as well. At all distorted geometries the dipole polarizability tensor is calculated. This is very time-consuming and is only feasible for small molecules. More details on the RESPONSE key can be found here.

There are a few caveats:
- Numerical integration (BeckeGrid) accuracy must be high
- A calculation in which only a subset of the atoms is displaced is not possible for Raman calculations.
- For good results, a well converged (with the same basis and functional) equilibrium geometry must be used.

Because of this last point, it is wise to always start the RAMAN calculation with a TAPE13 restart file from a previous geometry optimization with the same basis, accuracy parameters, and density functional.

Atomic coordinate displacements in a RAMAN calculation must be Cartesian, not Z-matrix. Furthermore, the current implementation does not yet support constrained displacements, i.e. you must use *all* atomic coordinate displacements. However, one can calculate Raman for selected frequencies, see next section.

The alternative Raman implementation with the AORESPONSE offers some unique features like lifetime options.

```
AORESPONSE
 RAMAN
END
```

Raman Intensities for Selected Frequencies

The RAMANRANGE keyword can be used to calculate Raman intensities for a range of frequencies only. Recommended to be used in the case one use a t21 as a restart file, which has frequencies on them. Using this option is a fast alternative for the existing method of calculating Raman intensities. Note that one should not use the key RAMANRANGE in combination with the key SYMMETRY. The input keyword is as follows:

```
RAMANRANGE low high {NUM=num DISRAD=disrad}
```

`low, high`

Two values defining an interval of frequencies to calculate the Raman intensities for. The Raman intensities are calculated by numerical differentiation of the polatizability tensor. Only frequencies within the interval that are known to be Raman-active will be included. This means that 2*N single-point TDDFT calculations will be performed, where N is the number of Raman-active frequencies within the range. Imaginary frequencies are specified using negative values, which is consistent with the notation adopted within ADF.

`num`

Num is an integer number specifying how many points are to be used for numerical differentiation: 2, 4, or 6. The default value is 2.

`disrad`

Disrad specifies the step size (in Angstrom) to be made in each direction for 2-point differentiation. For the 4-point differentiation the maximum deviation from the equilibrium geometry will be twice as large as for the 2-point one and so on. The default value for disrad is the same as used for numerical frequencies.

Main advantages of the method:

- Full symmetry at each displaced geometry is used, which does not only speeds the calculation up but also makes it more accurate.
- Only Raman-active modes are included in the calculation, which may save a lot of time for molecules with symmetry.
- There is no need to recalculate frequencies if you already have a t21 file with them as it can be used as a restart file.

For static Raman scattering (ω = 0) one does not need to add the RESPONSE block key. However, for the calculation of the frequency dependent Raman scattering the following RESPONSE block key is needed in the input:

```
RESPONSE
 RAMAN
 Nfreq 1
 FrqBeg Laserfreq
 [Optional Frequency/Energy Unit]
END
```

`Frequencies or wavelengths`

The number of frequencies Nfreq should be 1. With subkey Frqbeg the value of the Laser frequency value (Laserfreq) can be given. Default frequency unit is eV. This can be changed into Hartree units (a.u.) or in wavelengths (angstroms) by typing HARTREE or ANGSTROM on a separate line within the RESPONSE block, instead of [Optional Frequency/Energy Unit].

Resonance Raman: excited-state finite lifetime

In this method (Ref.[266]) the resonance Raman-scattering (RRS) spectra is calculated from the geometrical derivatives of the frequency-dependent polarizability. The polarizability derivatives are calculated from resonance polarizabilities by including a finite lifetime (phenomenological parameter) of the electronic excited states using time-dependent density-functional theory.

It is similar to the simple excited-state gradient approximation method (see next section) if only one electronic excited state is important, however, it is not restricted to only one electronic excited state. In the limit that there is only one possible state in resonance the two methods should give more or less the same results. However, for many states and high-energy states and to get resonance Raman profiles (i.e., Raman intensities as a function of the energy of the incident light beam) this approach might be more suitable. The resonance Raman profiles in this approach are averaged profiles since vibronic coupling effects are not accounted for. At the moment this method needs numerically calculated frequencies in Cartesian coordinates. The method described in the next section can use a basis of normal coordinates rather than Cartesian coordinates, so that in that method the calculation can be restricted to a couple of modes.

```
GEOMETRY
 FREQUENCIES
END
ALLPOINTS
AORESPONSE
 RAMAN
 FREQUENCY  1 freq1 units
 LIFETIME width
END
```

This method needs ALLPOINTS, because of the AORESPONSE key, and numerically calculated frequencies. The RRS is not calculated if one uses symmetric displacements in the numerically calculated frequencies or if one uses analytically calculated frequencies. The documentation of the AORESPONSE key explains in more detail the meaning of the subkeywords in the block key AORESPONSE, which are required

to calculate RRS. Similarly to the normal Raman module, the AOREPONSE-Raman only works with one frequency.

Resonance Raman: excited-state gradient

According to a the time-dependent picture of resonance-Raman (RR) scattering the relative intensities of RR scattering cross sections are, under certain assumptions, proportional to the square of the excited-state energy gradients projected onto the ground-state normal modes of the molecule (see Ref. [202]). For an alternative implementation of RR scattering using a finite lifetime of the excited states, and a discussion of some of the differences, see the previous section.

The excited-state gradients which are needed in this method can be computed numerically by ADF's VIBRON module, which is invoked by selecting the VIBRON runtype in the GEOMETRY block key, the use of the VIBRON block key and the EXCITATION block key:

```
GEOMETRY
  VIBRON
END
```

```
VIBRON
 NMTAPE filename
 RESRAMAN
 {...
 ..}
END
```

```
EXCITATIONS
  LOWEST nlowest
END
```

The VIBRON module always requires an EXCITATIONS input block, in which the total number of excited states to be calculated must be specified. NMTAPE is the only obligatory keyword for the VIBRON module. It specifies the name of a TAPE21 file from a previous frequency calculation. This TAPE21 file is needed to read the normal modes w.r.t. which the derivatives are computed. I.e., a separate frequency calculation must be carried out first. The second subkeyword RESRAMAN invokes the resonance Raman calculation. Note that the VIBRON module is not suited for open-shell TDDFT.

Resonance Raman for several excited states

The numerical evaluation of resonance Raman intensities has the advantages that

- Relative Intensities can be computed for several excited states at a time, since all excitation energies are determined simultaneously.
- Intensities can be computed for a selected no. of modes.

The intensities calculated for two different states cannot directly be compared, since the excited-state gradients only provide relative intensities for each excited state in resonance. For RR intensities from different excited states, also other quantities play a role. The most important one is the transition dipole moment to the excited state in resonance, which enters the intensity expression with to the fourth power.

Restrictions: (avoided) crossings between excited-states

The numerical calculation of excited-state gradients has a number of advantages, but also a possible problem: If the step size is chosen too large, or if there are close-lying excited-states, then the order of the excited states can change. For such cases, the excited-state gradient method to estimate relative RR intensities is not reliable: If states with different electronic character (but of the same symmetry) are close in energy, this will cause an avoided crossing. If the numerical derivatives are, in this case, computed w.r.t. the

adiabatic states, they will probably not reflect the true situation. Especially if the coupling matrix elements between the two excited states is small, the spectroscopic properties often behave as if there is no avoided crossing, i.e., according to the diabatic states. Such cases should be handled with extreme care, since it is often not possible in advance to see whether the adiabatic or the diabatic picture should be invoked.

Because of the possible (avoided) crossings the user must make sure that always enough excited states are calculated to include the state(s) of interest. E.g., if the resonance Raman intensities are required for the first excited state, also some higher excited states have to be included in the excitation calculation, as the first excited state at the ground-state equilibrium might be higher in energy for displaced structures.

Restrictions: results not trustworthy for higher excited states

Users should be aware of another technical point: Excited states are usually calculated from a Davidson diagonalization procedure, i.e., only a small number of eigenvalues and eigenvectors describing the lowest excitations are obtained. During finite displacements, some of the higher calculated states might leave the calculated energy window, while others enter it. Hence, the character of some of the higher calculated states can change. In such a case, the numerical differentiation based on a (simple) diabatic pictures will fail for the higher states, since no mapping between the excited states for reference (equilibrium) and displaced structure can be carried out.

The solution is rather simple: Users should always ask for more excited states than they are actually interested in, and discard the data for higher states, in particular for those which could not successfully be mapped for displaced structures (look for messages in the output like 'State No. X cannot be expressed in terms of reference states').

For advanced users it should be mentioned that it is possible to set an energy window within the range of states calculated, and only the states within this energy window will be taken into account in the evaluation. See the subkey ELTHRESH and EUTHRESH of the block key VIBRON.

Furthermore, it is possible to pick out certain states from this energy window, and only perform the mapping (and diabatization, if requested) and differentiation for them. See the subkey SELSTATE of the block key VIBRON.

Advanced Restarts

In some cases, it only becomes obvious which states have to be included in a (simple) diabatization after excitation energies for all displaced structures are calculated. Therefore, the selected states and the energy window settings can also be adjusted in a restart (with the usual restart key) after all single-point calculations are done. However, this is only possible if all raw data are saved to TAPE21, which might be an enormous amount of data.

Therefore, the user has to specify the subkeyword SAVERAWDAT of the key VIBRON in the production run, and the subkeyword USERAWDAT of the key VIBRON in the (evaluation) restart. Such a restart does not invoke any new SCF and will, therefore, typically only take a couple of seconds or minutes. If SAVERAWDAT is not specified, restarts are still possible, but the energy window cannot be adjusted differently, and no new state selection can be performed.

Resonance Raman Input options

A number of options are available for the VIBRON module, most of which are for special applications. All the options mentioned below have to appear in the VIBRON block. The VIBRON module always requires an EXCITATION input block, in which the total number of excited states to be calculated must be specified.

```
VIBRON
 NMTAPE filename
 RESRAMAN
 {DISPTYPE disptype}
 {STPSIZE stpsize}
```

```
  {ONLYSYM}
  {NOTONLYSYM}
  {DOMODES list}
  {DONTMODES list}
  {DSCHEME dscheme}
  {EUTHRES euthres}
  {ELTHRES elthres}
  {SELSTATE list}
  {SAVERAWDAT}
  {USERAWDAT}
 END
```

The most important ones in connection with RR calculations are:

`NMTAPE filename`

NMTAPE is the only obligatory keyword for the VIBRON module. It specifies the name of a TAPE21 file from a previous frequency calculation. This TAPE21 file is needed to read the normal modes w.r.t. which the derivatives are computed. I.e., a separate frequency calculation must be carried out first.

`RESRAMAN`

The second subkeyword RESRAMAN invokes the resonance Raman calculation.

`DISPTYPE disptype`

Select type of displacement steps; possible values are:
- `MASSWE`: steps in terms of mass-weighted normal mode vectors [default]
- `CARTES`: steps in terms of cartesian normal mode vectors
- `REDUCE`: steps in terms of reduced normal modes
- `ENERGY`: steps in terms of expected energy change (according to harmonic approximation)
- `ZPELEV`: like energy, but energy is expressed in ZPE units

`STPSIZE stpsize`

Sets the step size for the numerical differentiation in the default unit for the given DISPTYPE

`ONLYSYM`

Calculate derivatives only for totally symmetric modes (this is useful since this RR estimate only holds for Franck-Condon type Raman scattering, which is zero for non-symmetric modes). This option is ON per default in RR calculations, it can be switched off with the key NOTONLYSYM.

`DOMODES list`

Calculate derivatives only for the normal modes with numbers mentioned in list.

`DONTMODES list`

Calculate derivatives for all normal modes except the ones with numbers mentioned in list.

`DSCHEME dscheme`

The type of differentiation to be used can be set. Three different values for dscheme are available:

`ELCHAR`

A simple diabatic picture in which adiabatic states are mapped to the adiabatic states for the references structure based on a maximum transition density overlap criterion [default].

EIGVEC

A diabatic picture in which a diabatization is carried out as explained in Ref. [203]. I.e. the energies used here are the diagonal elements of the potential energy matrix for the nuclear Schrödinger Equation.

ADIABS

The adiabatic picture; can only be used if the symmetry of the excited states is supplied from a separate calculation, since the VIBRON module cannot check which states are allowed to cross as no symmetry is used in the excitation calculations. Consult the ADF-VIBRON manual [204] for information.

```
ELTHRESH elthresh
EUTHRESH euthresh
```

For advanced users it is possible to set an energy window within the range of states calculated, and only the states within this energy window will be taken into account in the evaluation.
- `elthresh`: lower bound in eV, default 0 eV.
- `euthresh`: upper bound in eV, default 1.0E10 eV.

```
SELSTATE list
```

For advanced users it is furthermore possible to pick out certain states from the energy window, and only perform the mapping (and diabatization, if requested) and differentiation for them. Here list includes the number (in ascending excitation energy) of the excited state at the reference (equilibrium) structure.

```
SAVERAWDAT
```

All raw data are saved to TAPE21, which might be an enormous amount of data. The selected states and the energy window settings can then be adjusted in a restart (with the usual restart key) and the inclusion of the subkey USERAWDATA after all single-point calculations are done.

```
USERAWDAT
```

All raw data are read from a previous calculation. The selected states and the energy window settings can now be adjusted. You need to invoke the usual restart key. Such a restart does not invoke any new SCF and will, therefore, typically only take a couple of seconds or minutes. If SAVERAWDAT is not specified in the previuous calculation, restarts are still possible, but the energy window cannot be adjusted differently, and no new state selection can be performed.

VROA: (Resonance) vibrational Raman optical activity

In ADF2010 a method is implemented to calculate both on- and off-resonance vibrational Raman optical activities (VROAs) of molecules using time-dependent density functional theory, see Ref. [306]. This is an extension of a method to calculate the normal VROA by including a finite lifetime of the electronic excited states in all calculated properties. The method is based on a short-time approximation to Raman scattering and is, in the off-resonance case, identical to the standard theory of Placzek. The normal and resonance VROA spectra are calculated from geometric derivatives of the different generalized polarizabilites obtained using linear response theory which includes a damping term to account for the finite lifetime. Gauge-origin independent results for normal VROA have been ensured using either the modified-velocity gauge or gauge-included atomic orbitals.

For the normal VROA use numerical frequencies, and the subkey VROA of the key AORESPONSE. Example input:

```
GEOMETRY
 frequencies
END
AORESPONSE
 NEWPOLCODE
 VROA
 scf  converge 1d-6 iterations 100
 frequency 1 5145 Angstrom
 ALDA
 FitAOderiv
 EL_DIPOLE_EL_DIPOLE VELOCITY
 EL_DIPOLE_EL_QUADRUPOLE VELOCITY
 EL_DIPOLE_MAG_DIPOLE VELOCITY
END
```

For the resonance VROA use numerical frequencies, and the subkey VROA and LIFETIME of the key AORESPONSE. Example input:

```
GEOMETRY
 frequencies
END
AORESPONSE
 NEWPOLCODE
 VROA
 scf  converge 1d-6 iterations 100
 frequency 1 5.15462 eV
 lifetime 0.0037
 ALDA
 FitAOderiv
 EL_DIPOLE_EL_DIPOLE VELOCITY
 EL_DIPOLE_EL_QUADRUPOLE VELOCITY
 EL_DIPOLE_MAG_DIPOLE VELOCITY
END
```

Vibrational Circular Dichroism (VCD) spectra.

Starting from ADF2007.01 it has become possible to calculate VCD spectra. The following keyword enables calculation of rotational strength during an analytical frequencies calculation:

```
VCD
```

It is important to note that the VCD keyword only works in combination `AnalyticalFreq` and `symmetry NOSYM`.

```
AnalyticalFreq
End

SYMMETRY NOSYM
```

The VCD intensities are calculated using Stephens' equations for VCD. For the calculation of the atomic axial tensors (AATs), analytical derivatives techniques and London atomic orbitals (the so called GIAO) are employed. As a result the calculated rotational strengths are origin independent, and therefore the common origin gauge is used [216].

Calculation of the AATs requires an analytical frequencies calculation. This limits the choice of functionals that can be used for VCD calculations. See the `ANALYTICALFREQ` keyword for a complete list of the

available functionals. The VCD calculations can be done immediately after a geometry optimization, just like analytical frequencies calculations.

The accuracy of the vibrational rotational strengths are determined by the accuracy of the harmonic force field, atomic polar tensors (APTs) and AATs. The most critical parameter being the harmonic force field. Thus, for a fair comparison with experimental data, accurate geometries and functionals that yield accurate force fields (e.g. BP86, OLYP, etc) should be used. Our tests showed that the BP86 functional in combination with TZP basis sets is always a safe choice. For a comparison of VCD spectra calculated with various functionals (e.g BP86, OLYP, BLYP, B3PW91 and B3LYP) see [216]. Regarding the geometries, we recommend the following strict settings, 10^{-4} for the geometry convergence of the gradients, and BeckeGrid quality good. The default settings should be used for the calculation of the frequencies.

The current VCD implementation does not support symmetry and therefore `symmetry NOSYM` should be specified in the input file. The frozen core approximation and open shell systems are also not supported.

By default, only the vibrational rotational strengths are printed in the ADF output file. The AATs can also be printed by specifying the keyword:

```
PRINT VCD
```

Vibrationally resolved electronic spectra

See the section on vibrationally resolved electronic spectra.

Time-dependent DFT

Excitation energies, frequency-dependent (hyper) polarizabilities, Van der Waals dispersion coefficients, higher multipole polarizabilities, Raman scattering intensities and depolarization ratios of closed-shell molecules are all available in ADF [71,72] as applications of time-dependent DFT (TDDFT) ; see [73] for a review.

New in ADF2004.01 is the calculation of circular dichroism (CD) spectra, and the calculation of the optical rotation (dispersion).

Starting from the ADF2005.01 version it is possible to calculate excitation energies for open-shell systems with TDDFT, including spin-flip excitation energies. New in ADF2005.01 is the possibility to use time-dependent current-density functional theory (TDCDFT).

New in ADF2006.01 is the possibility to calculate excitation energies for closed-shell molecules including spin-orbit coupling.

New in ADF2008.01 is the possibility to calculate lifetime effects in (dynamic) polarizabilities (AORESPONSE key).

The input description for these properties is split in three parts: (a) general advice and remarks, (b) excitation energies, and (c) frequency-dependent (hyper) polarizabilities (two alternative implementation: RESPONSE key and AORESPONSE key) and related properties.

General remarks on the Response and Excitation functionality

`Symmetry`

> As in calculations without TDDFT the symmetry is automatically detected from the input atomic coordinates and need not be specified, except in the following case: infinite symmetries cannot be

handled in the current release (ATOM, C(lin), D(lin)). For such symmetries a subgroup with finite symmetry must be specified in the input. The usual orientation requirements apply.
If higher multipole polarizabilities are required, it may also be necessary to use a lower subgroup (the program will stop with an error message otherwise). For verification of results one can always compare to a NOSYM calculation.

`Closed-shell`

The current implementation often supports only closed-shell molecules. If occupation numbers other than 0 or 2 are used the program will detect this, (but only at a later stage of the calculation) and abort. All 'RESPONSE' calculations must be spin-restricted.

`Open-shell`

Excitation energies can be obtained for open-shell systems in a spin-unrestricted TDDFT calculation. Spin-flip excitation energies can only be obtained in a spin-unrestricted TDDFT calculation.

`Atomic coordinates in a RAMAN calculation`

Atomic coordinate displacements in a RAMAN calculation must be Cartesian, not Z-matrix. Furthermore, the current implementation does not yet support constrained displacements, i.e. you must use *all* atomic coordinate displacements.

`Use of diffuse functions`

The properties described here may require diffuse functions to be added to the basis (and fit) sets. Poor results will be obtained if the user is unaware of this. As a general rule, diffuse functions are more important for smaller than for larger molecules, more important for hyperpolarizabilities than for normal polarizabilities, more important for high-lying excitation energies (Rydberg states) than for low-lying excitations, more important for higher multipole polarizabilities than for dipole polarizabilities. The user should know when diffuse functions are required and when they are not: the program will not check anything in this respect. For example, in a study on low-lying excitation energies of a large molecule, diffuse functions will usually have little effect, whereas a hyperpolarizability calculation on a small molecule is pointless unless diffuse functions are included. Diffuse even tempered basis sets are included in the ET/ directory of the database), for the elements H-Kr. Somewhat older basis sets can be found in the Special/Vdiff directory in the database. For other atoms, the user will have to add diffuse basis and fit functions to the existing data base sets. It is not necessary to start from basis V as was done for the basis sets in Special/Vdiff. For example, for heavier elements it may be a good idea to start from the ZORA/QZ4P basis sets. It may be expected that even more extensive basis sets will come available in the future, when usage and experience increase.

`Linear dependency in basis`

If large diffuse basis sets are used, or if diffuse functions are used for atoms that are not far apart the calculation may suffer from numerical problems because of (near-) linear dependencies in the basis set. The user should be aware of this danger and use the DEPENDENCY key to check and solve this.

`The LINEARSCALING input keyword`

For reasons of numerical robustness and safety rather strict defaults apply for the neglect of tails of basis and fit functions (see the key LINEARSCALING) in a Response or Excitation calculation. This may result in longer CPU times than needed for non-TDDFT runs, in particular for larger molecules. Possibly this precaution is not necessary, but we have not yet tested this sufficiently to relax the tightened defaults.

`Relativistic effects`

The Response and Excitations options can be combined with scalar relativistic options (ZORA or Pauli). The one-electron relativistic orbitals and orbital energies are then used as input for the property

calculation. Spin-orbit effects have been incorporated only in this part of the code (excitation energies). In case of a ZORA calculation, the so-called 'scaled' orbital energies are used as default.

Choice of XC potential

For properties that depend strongly on the outer region of the molecule (high-lying excitation energies, (hyper) polarizabilities), it may be important to use a XC potential with correct asymptotic behavior (approaching -1/r as r tends to infinity). Finally, several asymptotically correct XC potentials have been implemented in ADF, like the LB94 potential [15] and the statistical average of orbital potentials SAOP [244,17]. SAOP is recommended. Because of the correct asymptotic behavior the SAOP potential (and the LB94 potential) can describe Rydberg states correctly. The potentials based upon the so-called GRadient regulated seamless connection of model potentials (GRAC) for the inner and the outer region [16,18] have similar performance to SAOP, but have the disadvantage that the ionization energy of the molecule has to be used as input.

With the SAOP and GRAC functionals for the potential (as well as for LB94), the XC potential is computed from the exact charge density for reasons of stability and robustness (whereas for other functions the (cheaper) fit density is used). This implies that computation times may be longer. Another 'side effect' is that, since there is no energy expression corresponding to these potentials, the final (bonding) energy of such calculations uses another GGA and hence the energy result is not (exactly) consistent with the SCF procedure. Note, finally, that these potentials have been found to be not suitable for geometry optimizations because they maybe are not sufficiently accurate in the bonding region, see the discussion of the XC input key. Applications with SAOP to (hyper)polarizabilities and excitation energies, also for Rydberg transitions, can be found in [17] and with SAOP and Becke-Perdew-GRAC in [16,18]. Applications with the old LB94 potential to response calculations can be found in [74] (polarizabilities), [75-77] (hyperpolarizabilities), [78] (high-lying excitation energies), [79] (multipole polarizabilities and dispersion coefficients).

XC kernel

If most cases the adiabatic local density approximated (ALDA) kernel is used in the TDDFT functionality. In case of calculating excitation energies with a hybrid functionals the Hartree-Fock percentage times the Hartree-Fock kernel plus one minus the Hartree-Fock percentage times the ALDA kernel is used.

For excitation eenrgy calculations in some cases the full (non-ALDA) kernel can be evaluated, see the full XC kernel description. The Full kernel can not be used in combination with symmetry or excited state geometry optimizations.

COSMO

The COSMO model for solvation is a cheap method to include solvation effects in the TDDFT applications, see the SOLVATION key. Note that inclusion of the key ALLPOINTS is needed in case of TDDFT COSMO calculations. Note that in TDDFT calculations one may have to use two dielectric constants. The reason is that the electronic transition is so fast that only the electronic component of the solvent dielectric can respond, i.e., one should use the optical part of the dielectric constant. This is typically referred to as non-equilibrium solvation. The optical dielectric constant can be obtaining from the (frequency dependent) refractive index n of the solvent as: $\varepsilon_{opt} = n^2$. One can use the argumen NEQL in the subkey SOLV of the key SOLVATION to set a value for the optical dielectric constant. No dielectric constant in the response might be closer to the optical dielectric constant than using the full dielectric constant, This can be achieved if one includes the subkey NOCSMRSP in the block key SOLVATION, such that the induced electronic charges do not influence the COSMO surface charges. However, if one does geometry optimization then the full dielectric constant should be used in the TDDFT simulations since the solvent dielectric now has time to fully respond. The default is that the full dielectric constant is used in the TDDFT calculations.

Accuracy check list

As mentioned before, the TDDFT module is relatively new and not extensively tested for a wide range of applications. Therefore, we strongly recommend the user to build experience about aspects that may affect the accuracy of TDDFT results. In particular we advise to 'experiment' with

- Varying integration accuracy

- Varying the SCF convergence

- Varying the ORTHONORMALITY and TOLERANCE values in an Excitation calculation

- Varying the linearscaling parameters

- Using diffuse functions

- Using the Dependency key

- Applying the ZORA relativistic corrections for molecules containing heavy nuclei

- Using an asymptotically correct XC potential such as SAOP

Analysis options for TDDFT (excitation energies and polarizabilities)

Several options are available to obtain more detailed results than a few bare numbers for excitation energies, oscillator strengths, transition dipole moments, and (hyper)polarizabilities. For a zero-order understanding of which occupied and virtual orbitals play an important in the polarizability or intensity of an absorption peak,

it may be useful to know the values of the dipole matrix elements between (ground-state) occupied and virtual Kohn-Sham orbitals. If these dipole matrix elements are large for a particular occupied-virtual orbital pair, then this pair is almost certainly of great importance for the whole spectrum or polarizability. This information can be obtained by specifying somewhere in the input file (but NOT inside the RESPONSE or EXCITATION block keys):

```
PRINT DIPOLEMAT
```

Time-dependent Current DFT

The time-dependent current-density-functional (TDCDFT) implementation is built entirely upon the normal TDDFT implementation. Therefore all general remarks that are made for the TDDFT part of the program are also valid for TDCDFT. Only the polarizability and excitation energies of closed shell molecules can be calculated with TDCDFT in the present implementation.

If TDCDFT is used together with the ALDA functional (NOVK option) it will give the same results for the polarizability and excitation energies as TDDFT in a complete basis set. TDCDFT in ADF by default uses the VK functional [160,161], since this is the only current dependent functional that is known presently. Many aspects of the functional are still unknown and the functional should therefore be used with caution. The user is referred to the references for more information on when the VK functional gives good results and when not.

For more information on the implementation and applications of the TDCDFT and the VK functional please read the references: [162-165]. For more details on the theory and implementation in ADF see: [166].

To activate TDCDFT and the VK functional one should add the following block key to the input file:

```
CURRENTRESPONSE
END
```

To calculate the polarizability the keyword Response can be used with the following options:

```
RESPONSE
 ALLCOMPONENTS
 Nfreq Nfreq
 FrqBeg FirstFreq
 FrqEnd LastFreq
 [Optional Frequency/Energy Unit]
END
```

The block key EXCITATION can be used with all of its options.

In ADF2012 the block keyword AORESPONSE can also be used with the Vignale-Kohn functional. The current-density is generated on the fly but otherwise the computation is based on the time dependent density response.

In default the VK functional will be applied where the NCT parameterization [167] is chosen for the transverse exchange-correlation kernel for the polarizability and singlet excitation energies (giving the best results for the systems studied so far). For triplet excitation energies the only available parameterization will be used [168]. This option is not tested much and the results are in general much worse than ALDA [166]. It is therefore suggested that VK is not used to calculate triplet excitation energies.

In the output the polarizability tensor (in case of an ALLCOMPONENTS calculation) has a different shape, the results are printed in the more intuitive order x, y, z, instead of y, z, x that the TDDFT implementation uses.

The following subkeys are available within the datablock of CURRENTRESPONSE

```
CURRENTRESPONSE
QIANVIGNALE
NOVK
END
```

`QIANVIGNALE`

The QV parameterization [168] will be used for the transverse exchange-correlation kernel instead of NCT.

`NOVK`

TDCDFT will be applied with the ALDA functional instead of the VK functional. In a complete basis this will give the same results as a TDDFT calculation.

Excitation energies: UV/Vis spectra, X-ray absorption, CD, MCD

Ultraviolet-visible (UV/Vis) spectroscopy studies electronic excitations of valence electrons, whereas X-ray spectroscopy studies electronic excitations of core electrons. Excitation energies and oscillator strengths are all available in ADF as applications of time-dependent DFT (TDDFT). Excitation energies can be calculated for closed-shell as well as for open-shell molecules. It is also possible to include spin-orbit coupling and to calculate core excitations (X-ray absorption spectra). Circular dichroism (CD) is the differential absorption of left- and right-handed circularly polarized light.

Excitation energies, UV/Vis spectra

You can perform a calculation of singlet-singlet and singlet-triplet excitation energies of a closed-shell molecule by supplying in the input file the block key EXCITATION. See the next sections for settings of

technical parameters, the calculation of excitation energies for open shell molecules, inclusion of spin-orbit coupling, and the calculation of CD spectra.

```
EXCITATIONS
 EXACT &
   IRREP1 N1
   IRREP2 N2
 SUBEND
 DAVIDSON &
   IRREP3 N3
   IRREP4 N4
 SUBEND
 ALLOWED
 ONLYSING
 ONLYTRIP
 LOWEST nlowest
End
```

Several options can be addressed with subkeys in the data block. This functionality is based on TDDFT and consequently has a different theoretical foundation than the SCF techniques described elsewhere in this User's Guide. Two possible ways are available to solve the eigenvalue equation from which the excitation energies and oscillator strengths are obtained, of which the iterative Davidson procedure is the default. In this case, the program needs to know how many excitation energies are needed per irrep, what accuracy is required, and what type of excitation energies are required (singlet-singlet or singlet-triplet). Suitable defaults have been defined for all of these. Each of these points is discussed below.

`Exact diagonalization vs. iterative Davidson procedure`

The most straightforward procedure is a direct diagonalization of the matrix from which the excitation energies and oscillator strengths are obtained. Since the matrix may become very large, this option is possible only for very small molecules. It can be activated by specifying the word EXACT as one of the subkeys in the Excitations data block. The default is the iterative Davidson method. A few of the lowest excitation energies and oscillator strengths are then found within an error tolerance. An advantage of the EXACT option is that additional information is produced, such as the Cauchy coefficients that determine the average dipole polarizability. The EXACT option can not be used in unrestricted calculations.

`Singlet versus triplet`

By default, the singlet-singlet and singlet-triplet excitation energies are both calculated. The singlets are handled first, then the corresponding triplet excitation energies. One can skip one of these two parts of the calculation by specifying either ONLYSING or ONLYTRIP as a subkey in the data block.

In case of a calculation including spin-orbit coupling one can not separate the singlet-singlet and singlet-triplet excitations. The subkeys ONLYSING and ONLYTRIP are misused in this case to do a spin-restricted calculation, or a spin-polarized calculation, respectively. One should in fact only use the results of the spin-polarized calculation.

`Dipole-allowed versus general excitations.`

If you are interested in the optical absorption spectrum, you may not want to compute singlet-triplet excitation energies, nor singlet-singlet excitation energies which, by symmetry, have zero oscillator strengths. This subkey should not be used in case of spin-orbit coupling. The subkey ALLOWED tells ADF to treat only those irreducible representations for which the oscillator strengths will be nonzero. Of course, the oscillator strengths may still be negligibly small. The ALLOWED subkey automatically implies ONLYSING. The simplest, fastest, and recommended way to obtain information about the ten lowest dipole-allowed excitation energies would be:

```
EXCITATIONS
 ALLOWED
 LOWEST 10
END
```

Which excitation energies and how many?

The user can specify how many excitation energies per irrep should be calculated. If no pertaining input is available the program determines these numbers from the smallest differences between occupied and virtual Kohn-Sham orbital energies. By default it looks at the 10 lowest orbital energy differences. This number can be modified, by specifying inside the Excitation block key, for example:

```
LOWEST 30
```

One should be aware that this procedure does not guarantee that the lowest 10 (or 30) excitation energies will actually be found, since the orbital energy difference approximation to the excitation energy is rather crude. However, if the program decides on the basis of this procedure to calculate 4 excitation energies in a certain irreducible representation, these 4 excitation energies are certainly the lowest in that particular irrep.

The user has more control when the number of excitations per irrep is explicitly specified within the EXCITATION block key by the Davidson subkey:

```
DAVIDSON &
 E'' 5
 T1.u 2
SUBEND
```

The DAVIDSON sub key is a general (simple or block type) subkey. For usage as block type it must, be followed by the continuation code (&). Its data block may contain any number of records and must end with a record SUBEND. In the subkey data block a list of irreps, followed by the number of requested excitation energies is specified. Note that the irrep name may not be identical to the usual ADF name. For example E'' is called EEE in ADF. The Excitation code will skip an irrep if the label is not recognized. For multidimensional irreps, only the first column is treated, because the other would produce identical output. This implies that the oscillator strengths for E-irreps have to be multiplied by 2 and the oscillator strengths for T-irreps by 3. The ALLOWED subkey should not be used if irreps are specified with the Davidson block subkey, however, the subkey ONLYSING (or ONLYTRIP) can be used in this case.

The EXACT subkey, mentioned already above, can also be used as a block type subkey to treat only a few irreps instead of all. The number of excitation energies does not have to be specified then.

Tamm-Dancoff approximation

Excitation energies can be calculated using the Tamm-Dancoff approximation (TDA) [158] if one includes, besides the EXCITATION block key, the key TDA:

```
TDA
```

Full XC kernel

With XCFUN the full (non-ALDA) kernel can be evaluated, see the XCFUN description. To use the non-ALDA kernel the keyword FULLKERNEL should be put in the EXCITATIONS block. FULLKERNEL can be used with GGAs (including hybrids and RS functionals) but not meta-GGAs or meta-hybrids. FULLKERNEL

can not be used in combination with symmetry, excited state geometry optimizations or other response properties.

```
SYMMETRY NOSYM
XC
 ...
 XCFUN
End
EXCITATIONS
 ...
 FullKernel
END
```

Singlet-triplet excitations are not possible with FULLKERNEL. Thus for closed shell systems, one needs to include ONLYSING.

```
SYMMETRY NOSYM
XC
 ...
 XCFUN
End
EXCITATIONS
 ...
 ONLYSING
 FullKernel
END
```

Excitations as orbital energy differences

Instead of the relative expensive TDDFT calculation of excitation energies, sometimes just calculating Kohn-Sham orbital energy differences may already be useful. The keyword KSSPECTRUM, in combination with the block key EXCITATIONS, will calculate excitation energies as Kohn-Sham orbital energy differences. For a given excitation from an occupied orbital to a virtual orbital the oscillator strength is calculated from the the dipole transition moment between this occupied orbital and this virtual orbital. Especially useful for core excitation energy calculations. If KSSPECTRUM is used, it is possible to use fractional occupation numbers in the SCF, like is used in the DFT transition state (DFT-TS) scheme, see, for example, Ref. [359]. Note: for fractional occupation numbers, typically an orbital is treated in the excitation calculation as if it is fully occupied if the occupation number is 1.5 or more, and it is treated as if it is fully unoccupied if the occupation number is 0.5 or less.

```
KSSPECTRUM
EXCITATIONS
 Lowest 20
 KFWRITE 0
END
```

`KSSPECTRUM`

keyword to use only orbital energy differences

`KFWRITE kfwrite`

Subkeyword in EXCITATIONS block key. If kfwrite is 0 then do not write contributions, transition densities, and restart vectors to TAPE21, since this can lead to a huge TAPE21, especially if many excitations are calculated. Default value kfwrite is 3, which means that contributions, transition densities, and restart vectors are written to TAPE21.

Accuracy and other technical parameters

A summary of technical parameters with their defaults is:

```
EXCITATIONS
 VECTORS 40
 TOLERANCE 1e-6
 ORTHONORMALITY 1e-8
 ITERATIONS 200
 KFWRITE 3
END
```

`VECTORS vectors`

The maximum number of trial vectors in the Davidson algorithm for which space is allocated. If this number is small less memory will be needed, but the trial vector space is smaller and has to be collapsed more often, at the expense of CPU time. The default if usually adequate.

`TOLERANCE tolerance`

Specifies the error tolerance in *the square* of the excitation energies in hartree units. The default is probably acceptable but we recommend that you verify the results against a stricter default (e.g. 1e-8) for at least a few cases.

`ORTHONORMALITY orthonormality`

The Davidson algorithm orthonormalizes its trial vectors. Increasing the default orthonormality criterion increases the CPU time somewhat, but is another useful check on the reliability of the results.

`ITERATIONS iterations`

The maximum number of attempts within which the Davidson algorithm has to converge. The default appears to be adequate in most cases.

`KFWRITE kfwrite`

If kfwrite is 0 then do not write contributions, transition densities, and restart vectors to TAPE21, since this can lead to a huge TAPE21, especially if many excitations are calculated. Default value kfwrite is 3, which means that contributions, transition densities, and restart vectors are written to TAPE21.

Excitation energies for open-shell systems

Excitation energies can be obtained for open-shell systems in a spin-unrestricted TDDFT calculation [154]. To perform an open-shell TDDFT calculation one just needs to do an unrestricted SCF calculation and use the EXCITATION keyword. Presently the excitation energies can only be found with Davidson's procedure. In case of spin-orbit coupling, see the section on approximate spin-orbit coupled excitation energies open shell molecule.

The printed symmetry in the output in TDDFT calculations is actually the symmetry of transition density. For closed-shell systems, the symmetry of the excited state is the same as the symmetry of the transition density, while for open-shell systems, the symmetry of the excited states is the direct product between the symmetry of the transition density and the ground state symmetry. Note that the ground state symmetry of an open shell molecule is not necessarily A1.

For degenerate representations such as the 2-dimensional E-representations or the 3-dimensional T-representations, the occupation should be either fully occupied or zero. For example, for an orbital in an E-

representation the α and β occupation number should be either 2 or 0. The α occupation number can of course be different from the β occupation number.

As for the spin-state, the general rule is that if the excited state mainly results from transitions from the singly occupied orbitals to virtual orbitals or from fully occupied orbitals to the singly occupied orbitals, the spin state of the excited state should roughly be the same as that of the ground state. However, if the excited state mainly comes from transitions from fully occupied orbitals to virtual orbitals, the spin state of the excited state are usually a mixture since TDDFT can only deal with single excitations within adiabatic approximation for the XC kernel [155]. Sometimes we just suppose the spin state of this kind of excited states to be the same as that of ground state [154]. In the MO → MO transitions part for the excitations of the output file, the spin of each molecular orbitals are also specified to help assign the spin state of the excited states. The transitions are always from α spin-orbital to α spin-orbital or from β spin-orbital to β spin-orbital.

Spin-flip excitation energies

Spin-flip excitation energies [156,157] can only be obtained in a spin-unrestricted TDDFT calculation. This can not be used in case of spin-orbit coupling. At present, the spin-flip excitation energies can only be calculated with Tamm-Dancoff approximation (TDA) [158] and Davidson's method.

To calculate spin-flip excitation energies, one must specify two keys:

```
SFTDDFT
```

and

```
TDA
```

anywhere in the input file in addition to the `EXCITATION` block keyword.

In spin-flip TDDFT, the XC kernel can be calculated directly from the XC potential. To use the LDA potential for the XC kernel, which roughly corresponds to the ALDA in ordinary TDDFT, one must specify the key

```
FORCEALDA
```

anywhere in the input file. Only calculations using the LDA potential in the SCF are fully tested. Using other GGA potentials in the SCF and using the `FORCEALDA` key at the same time may introduce unreasonable results, while using LB94 or SAOP potential in the SCF without the `FORCEALDA` key may give unstable results. Unstable results have been reported for the PW91 functional.

For open-shell molecules, spin-flip transition can result in transition to the ground state with a different S_z value, while the symmetry of the transition density is A1. The excitation energy of this transition should be zero and this can be used to test the reliability of spin-flip TDDFT.

The symmetry of the excited states can be determined in the same way as that in spin-unrestricted TDDFT calculations. As for the spin state, similar to that in the spin-unrestricted TDDFT calculations, some states may be more or less pure spin states, others may just be mixtures. The users can interpret the excited state through the transitions that contribute to this state. Note that the transitions are always from α spin-orbital to β spin-orbital in spin-flip calculations, or from β spin-orbital to α spin-orbital.

Select excitation energies, Core Excitation energies, X-ray absorption

Two methods can be used to reduce the computational costs of, for example, core excitation energies, or some other high lying excitation energy. In the state selective method scheme a guess vector for the orbital transition has to be provided. An overlap criterion is used to follow the wanted eigenvector. In this scheme

the one-electron excited state configuration space remains complete, see Ref [346]. In the second scheme, the range of excitations that are calculated is modified, which means that the one-electron excited state configuration space is reduced to the interesting part, see Ref. [169]. The calculated excited states are more accurate with the state selective method if convergence is reached, however, the second scheme is more robust, and it is easier to find convergence.

These selection methods can also be used in case one calculates excitation energies as Kohn-Sham orbital energy differences, see key KSSPECTRUM.

State selective optimization excitation energies

The state selective method (key SELECTEXCITATION) can be used to reduce the computational costs of, for example, core excitation energies. In this scheme a guess vector for the orbital transition has to be provided. It should be used in combination with the Davidson method to calculate excitation energies. An overlap criterion is used to follow the wanted eigenvector. This method for state selective optimization of excitation energies is based on the method by Kovyrshin and Neugebauer, see Ref. [346]. This key can also be used in case of spin-orbit coupling. The use of the key SELECTEXCITATION is similar as the use of the key MODIFYEXCITATION. However, the key SELECTEXCITATION can not be used in combination with the key MODIFYEXCITATION. In the state selective method (key SELECTEXCITATION) the one-electron excited state configuration space remains complete, whereas it is reduced in case the scheme with the MODIFYEXCITATION key.

The starting guess vector(s) for the excitation energies can be selected, for example by selecting 1 occupied orbital and 1 virtual orbital.

```
SELECTEXCITATION
   OscStrength oscstrength
   UseOccVirtRange elowoccvirt ehighoccvirt
   UseOccVirtNumbers nrlowoccvirt nrhighoccvirt
   UseOccRange elowocc ehighocc
   UseVirtRange elowvirt ehighvirt
   UseOccupied
      irrep orbitalnumbers
      irrep orbitalnumbers
      ...
   SubEnd
   UseVirtuaL
      irrep orbitalnumbers
      irrep orbitalnumbers
      ...
   SubEnd
   UseScaledZORA
end
```

`OscStrength oscstrength`

Use only pairs of an occupied and virtual orbital as guess vectors, for which the oscillator strength of the single-orbital transition is larger than `oscstrength`.

`UseOccVirtRange elowoccvirt ehighoccvirt`

Use only pairs of an occupied and virtual orbital as guess vectors, for which the orbital energy difference is between `elowoccvirt` and `ehighoccvirt` (in hartree).

`UseOccVirtNumbers nrlowoccvirt nrhighoccvirt`

Use only pairs of an occupied and virtual orbital as guess vectors, for which in the sorted list of the orbital energy differences, the number of the single-orbital transition is between `nrlowoccvirt` and `nrhighoccvirt`.

`UseOccRange elowocc ehighocc`

Use only occupied orbitals in the guess vectors which have orbital energies between `elowocc` and `ehighocc` (in hartree).

`UseVirtRange elowvirt ehighvirt`

Use only virtual orbitals in the guess vectors which have orbital energies between `elowvirt` and `ehighvirt` (in hartree).

`UseOccupied`

Use only the occupied orbitals in the guess vectors which are specified.

`UseVirtual`

Use only the virtual orbitals in the guess vectors which are specified.

`irrep`

The name of one of the irreducible representations (not a subspecies) of the point group of the system. See the Appendix for the irrep names as they are used in ADF.

`orbitalnumbers`

A series of one or more numbers: include all numbers of the orbitals in the guess vectors that are to be used. In an unrestricted calculation the same numbers are used for the spin-α orbitals and the spin-β orbitals.

Modify range of excitation energies

The key MODIFYEXCITATION can be used to reduce the computational costs of, for example, core excitation energies. This key can also be used in case of spin-orbit coupling. The use of the key MODIFYEXCITATION is similar as the use of the key SELECTEXCITATION. However, the key MODIFYEXCITATION can not be used in combination with the key SELECTEXCITATION. In the state selective method (key SELECTEXCITATION) the one-electron excited state configuration space remains complete, whereas it is (effectively) reduced in case the scheme with the MODIFYEXCITATION key.

One possibility is to allow only selected occupied orbitals and or selected virtual orbitals in the TDDFT calculations. In this scheme the complete one-electron excited state configuration space is reduced to the subspace where only the core electrons are excited, see Stener et al. [169]. In the actual implementation this is done by artificially changing the orbital energies of the uninteresting occupied orbitals to a large negative value (default -1d6 hartree), and by by artificially changing the orbital energies of the uninteresting virtual orbitals to a large positive value (default 1d6).

In ADF2010 an extra possibility is added with the new subkey UseOccVirtRange, which restricts the space of excitation energies, by allowing only pairs of occupied and virtual orbitals, for which the difference in orbital energy is between a certain range.

```
MODIFYEXCITATION
   OscStrength oscstrength
   UseOccVirtRange elowoccvirt ehighoccvirt
   UseOccVirtNumbers nrlowoccvirt nrhighoccvirt
```

```
   UseOccRange elowocc ehighocc
   UseVirtRange elowvirt ehighvirt
   UseOccupied
      irrep orbitalnumbers
      irrep orbitalnumbers
      ...
   SubEnd
   UseVirtuaL
      irrep orbitalnumbers
      irrep orbitalnumbers
      ...
   SubEnd
   SetOccEnergy esetocc
   SetLargeEnergy epsbig
   UseScaledZORA
end
```

`OscStrength oscstrength`

Use only pairs of an occupied and virtual orbital as guess vectors, for which the oscillator strength of the single-orbital transition is larger than `oscstrength`.

`UseOccVirtRange elowoccvirt ehighoccvirt`

Use only pairs of an occupied and virtual orbital, for which the orbital energy difference is between `elowoccvirt` and `ehighoccvirt` (in hartree).

`UseOccVirtNumbers nrlowoccvirt nrhighoccvirt`

Use only pairs of an occupied and virtual orbital as guess vectors, for which in the sorted list of the orbital energy differences, the number of the single-orbital transition is between `nrlowoccvirt` and `nrhighoccvirt`.

`UseOccRange elowocc ehighocc`

Use only occupied orbitals which have orbital energies between `elowocc` and `ehighocc` (in hartree).

`UseVirtRange elowvirt ehighvirt`

Use only virtual orbitals which have orbital energies between `elowvirt` and `ehighvirt` (in hartree).

`UseOccupied`

Use only the occupied orbitals which are specified.

`UseVirtual`

Use only the virtual orbitals which are specified.

`irrep`

The name of one of the irreducible representations (not a subspecies) of the point group of the system. See the Appendix for the irrep names as they are used in ADF.

`orbitalnumbers`

A series of one or more numbers: include all numbers of the orbitals that are to be used. In an unrestricted calculation the same numbers are used for the spin-α orbitals and the spin-β orbitals.

`SetOccEnergy esetocc`

All occupied orbitals that have to be used will change their orbital energy to `esetocc`. In practice only useful if one has selected one occupied orbital energy, and one want to change this to another value. Default: the orbital energies of the occupied orbitals that are used are not changed.

`SetLargeEnergy epsbig`

The orbital energies of the uninteresting occupied orbitals are changed to `-epsbig` hartree, and the orbital energies of the uninteresting virtual orbitals are changed to `epsbig` hartree (Default: `epsbig` = 1d6 hartree).

`UseScaledZORA`

Use everywhere the scaled ZORA orbital energies instead of the ZORA orbital energies in the TDDFT equations. This can improve deep core excitation energies. Only valid if ZORA is used. Default: use the unscaled ZORA orbital energies.

Excitation energies and Spin-Orbit coupling

Spin-orbit coupling can be included in the TDDFT calculation of excitation energies for closed-shell molecules. Two methods can be used in ADF. The first one includes spin-orbit coupling as a perturbation to a scalar relativistic calculation of excitation energies. The second one includes spin-orbit coupling self-consistently in the ground state calculation. If spin-orbit coupling is large, the second one is more accurate, but is also more time-consuming.

The results of these spin-orbit coupled TDDFT calculations include the calculation of the zero field splitting (ZFS) of triplet excited states and the calculation of radiative rate constants, which could be used to calculate radiative phosphorescence lifetimes.

Perturbative inclusion of spin-orbit coupling

```
SOPERT {NCALC=ncalc} {ESHIFT=eshift}
RELATIVISTIC SCALAR ZORA
EXCITATIONS
END
```

The perturbative method, which is described in Ref.[280], is an approximate time-dependent density-functional theory (TDDFT) formalism to deal with the influence of spin-orbit coupling effect on the excitation energies for closed-shell systems. In this formalism scalar relativistic TDDFT calculations are first performed to determine the lowest single-group excited states and the spin-orbit coupling operator is applied to these single-group excited states to obtain the excitation energies with spin-orbit coupling effects included. The computational effort of the present method is much smaller than that of the two-component TDDFT formalism. The compositions of the double-group excited states in terms of single-group singlet and triplet excited states are obtained automatically from the calculations. In Ref.[280] it was shown that the calculated excitation energies based on the present formalism affords reasonable excitation energies for transitions not involving 5p and 6p orbitals. For transitions involving 5p orbitals, one can still obtain acceptable results for excitations with a small truncation error, while the formalism will fail for transitions involving 6p orbitals, especially 6p1/2 spinors.

Although this method is not completely correctly implemented for (meta-)hybrids or Hartree-Fock, it still gives reasonable excitation energies, and can thus be useful also in that case. Note that SYMMETRY C(2H) is not implemented for spin-orbit coupled excitations, use SYMMETRY C(S), C(I) or NOSYM, instead.

`NCALC=ncalc`

Number of spin-orbit coupled excitation energies to be calculated. Default (and maximum) value: 4 times the number of scalar relativistic singlet-singlet excitations.

`ESHIFT=eshift`

The actually calculated eigenvalues are calculated up to the maximum singlet-singlet or singlet-triplet scalar relativistic excitation energy plus eshift (in hartree). Default value: 0.2 hartree.

Some extra information about the spin-obit matrix is written to the output if one includes:

```
SOPERT {NCALC=ncalc} {ESHIFT=eshift}
PRINT SOMATRIX
```

If one includes PRINT SOMATRIX the spin-orbit matrix on basis of singlet and triplet excited states will be printed. On the diagonal the singlet or triplet energies is added. The spin-orbit matrix has a real and imaginary part. This spin-orbit matrix is the one that is diagonalized to get the spin-orbit coupled excitation energies.

```
GSCORR
```

If one includes GSCORR the singlet ground state is included, which means that spin-orbit cupling can also have some effect on energy of the ground state. The spin-orbit matrix in this case is on basis of the ground state and the singlet and triplet excited states.

Self-consistent spin-orbit coupling

```
RELATIVISTIC SPINORBIT ZORA
EXCITATIONS
  {ALSORESTRICTED}
END
```

Starting from the ADF2006.01 version in ADF the relativistic TDDFT formalism, including spin-orbit coupling, is implemented for closed-shell molecules with full use of double-group symmetry [182]. This relativistic time-dependent density-functional theory (TDDFT) is based on the two-component zeroth-order regular approximation (ZORA) and a noncollinear exchange-correlation (XC) functional. This two-component TDDFT formalism has the correct nonrelativistic limit and affords the correct threefold degeneracy of triplet excitations.

In case of a calculation including spin-orbit coupling one can not separate the singlet-singlet and singlet-triplet excitations. By default the spin-polarized excitation energies are calculated (the noncollinear scheme is used for the spin-dependent exchange-correlation kernel). The subkeys ALSORESTRICTED can be used to include also excitation energies in which a spin-restricted exchange-correlation kernel is used. One should in fact only use the results of the spin-polarized calculation, which is based on the noncollinear exchange-correlation (XC) functional. For the same reason, the ALLOWED subkey should not be used if spin-orbit coupling is included. Note that SYMMETRY C(2H) is not implemented for spin-orbit coupled excitations, use SYMMETRY C(S), C(I) or NOSYM, instead.

To perform a spin-orbit coupled TDDFT calculation one just needs to do a spin-orbit coupled SCF calculation and use the EXCITATION keyword. The molecule needs to be closed shell, and should be calculated spin-restricted. Thus do not use the UNRESTRICTED, COLLINEAR, or NONCOLLINEAR keyword. See, however, also next section.

The contribution to the double group excited states in terms of singlet and triplet single group excited states can be estimated through the inner product of the transition density matrix obtained from two-component and scalar relativistic TDDFT calculations to better understand the double group excited states [183]. In order to get this analysis one needs to perform a scalar relativistic TDDFT calculation of excitation energies

on the closed shell molecule first, and use the resulting TAPE21 as a fragment in the spin-orbit coupled TDDFT calculation of excitation energies, including the keyword STCONTRIB (Singlet and Triplet CONTRIButions):

```
STCONTRIB
```

This STCONTRIB analysis is not performed for (meta-)hybrids, unless one uses the Tamm-Dancoff approximation (TDA) approximation, but then it may also fail. If one wants this STCONTRIB analysis for (meta-)hybrids one may consider to the perturbative inclusion of spin-orbit coupling in the calculation of excitation energies.

Note that if hybrids are used, the dependency key is automatically set, and this may effectively reduce the number of excitations, which may give problems in the STCONTRIB analysis. A workaround for these problems is to first calculate the scalar relativistic fragment without the EXCITATIONS keyword. Use the TAPE21 of this calculation as fragment in a scalar relativistic calculation with the EXCITATIONS keyword. Use the TAPE21 of the second calculation as fragment in the spin-orbit coupled calculation, inclding the STCONTRIB keyword.

Highly approximate spin-orbit coupled excitation energies open shell molecule

Excitation energies can be obtained for open-shell systems in a spin-unrestricted TDDFT calculation including spin-orbit coupling. This approximate method uses a single determinant for the open shell ground state. The Tamm-Dancoff approximation (TDA) is needed and symmetry NOSYM should be used. Best is to use the noncollinear approximation. For analysis it is advised to calculate the molecule also with the scalar relativistic spin-restricted method and use it as fragment in the spin-orbit coupled calculation. This will make it easier to identify the excitations.

```
Relativistic spinorbit ZORA
Unrestricted
NonCollinear
Symmetry Nosym
TDA
Excitations
End
```

Note that this approximate method for open shell molecules is not able to show the subtle effects of spin-orbit coupling. Some of the reasons are the approximate nature of the XC functionals for open shell molecules, the single determinant that is used for the open shell ground state, and that only single excitations are included in the excitation. If one does not include spin-orbit coupling the spin-unrestricted TDDFT approach introduces spin-contamination such that the result does not represent transitions between pure spin states. Inclusion of spin-orbit coupling will not simplify this. However, if spin-orbit coupling is large, then this method may help to identify excitations.

Note that the approximations made in this approximate method are much worse than for spin-orbit coupled TDDFT for closed shell systems. In that case one can get a reasonable description of the subtle effects of spin-orbit coupling, for example, for the zero-field splitting of a triplet excited state.

CD spectra

Circular dichroism (CD) is the differential absorption of left- and right-handed circularly polarized light. Starting from ADF2010 Hartree-Fock and hybrids can also be used to calculate CD spectra.

```
EXCITATIONS
 CDSPECTRUM
 ANALYTIC
```

```
 VELOCITY
End
```

CDSPECTRUM

If the subkey *CDSPECTRUM* is included in the key EXCITATIONS the rotatory strengths for the calculated excitations are calculated, in order to simulate Circular Dichroism (CD) spectra [80,81]. Interesting for chiral molecules. This subkey should not be used in case of spin-orbit coupling. For accuracy reasons you should also use the subkey *ANALYTIC* in the block key EXCITATIONS, otherwise the results may be nonsense.

ANALYTIC

If the subkey *ANALYTIC* is included the required integrals for the CD spectrum are calculated analytically, instead of numerically. Only used in case of CD spectrum.

Velocity

If the subkey *VELOCITY* is included ADF calculates the dipole-velocity representation of the oscillator strength. If applicable (use of subkey *CDSPECTRUM*) the dipole-velocity representation of the rotatory strength is calculated. Default the dipole-length representation of the oscillator strength and rotatory strength is calculated.

MCD

MCD or magnetic circular dichroism is the differential absorption of left and right circularly polarized light in the presence of a magnetic field. MCD intensity is usually described in terms of different contributions called A, B and C terms, see Refs. [273,274]. A further parameter D is often discussed in MCD studies. D is proportional to the intensity of an absorption band and is closely related to the oscillator strength. A and B terms for closed and open-shell molecules and C terms of open-shell molecules induced by spin-orbit coupling can be calculated. Starting from ADF2010 C terms related to spatially degenerate states, i.e. breaking of degeneracies can be calculated.

For MCD calculations for molecules that have C(2) or D(2) symmetry use SYMMETRY NOSYM.

Input options

```
EXCITATIONS
 MCD {options}
 ONLYSINGLET
 {SELECT transition number}
 {DTENSOR {Dxx Dxy Dyy Dxz Dyz|D E/D}}
End
ALLPOINTS
{RELATIVISTIC ZORA}
{SOMCD}
{ZFS}
```

MCD

If the subkey MCD is included in the key EXCITATIONS the MCD parameters of some or all of the excitations considered in the TDDFT procedure are calculated [275-278]. This subkey should not be used with spin-orbit coupling (but, see below). Several other keywords could be important.
ALLPOINTS: required for an MCD calculation.
ONLYSINGLET: this keyword should be used in combination with a MCD calculation.
RELATIVISTIC ZORA: required for a calculation of temperature-dependent C terms. In this case the keyword SOMCD must also be added as a key by itself. If only A and B terms are calculated then

ZORA is not needed but can be included if desired.
ZFS: If the ZFS keyword and MCD with SOMCD are also included then the influence of the calculated zero-field splitting (ZFS) on the temperature-dependent MCD is evaluated. The MCD in the presence of ZFS is described as anisotropic in the output because the Zeeman splitting becomes orientation dependent in the presence of ZFS.

In ADF2010 the temperature-dependent MCD due to the breaking of degeneracies of excited states by spin-orbit coupling can be calculated. Although all temperature-dependent MCD is typically called "C terms", the parameters associated with the MCD are labeled "CE" to distinguish them from the MCD due to mixing between states caused by spin-orbit coupling that is labeled "C". The CE terms have a derivative shape like A terms. They have the same temperature-dependence as normal C terms. If they are present, CE terms are calculated automatically along with C terms if the keyword SOMCD is included in the input.

```
MCD {options}
```

Options include NMCDTERM, NMIX, DCUTOFF, MCDOUT, CGOUT, NANAL, NANAL2, FULLOMEGA, NOAB, NODIRECT, NOCG, CONVCG, ITERCG, ITER2CG, BMIN, BMAX, TMIN, TMAX and NTEMP.

`NMCDTERM=nmcdterm`

Number of excitations for which MCD parameters are to be calculated. The nmcdterm lowest energy excitations are treated. The default is the number of transitions considered in the TDDFT calculation.

`NMIX=nmix`

Number of transitions allowed to mix in a SOS calculation. Default is the number of transitions considered in the TDDFT calculation.

`DCUTOFF=dcutoff`

MCD parameters will only be calculated for transitions with sufficient intensity. Each cartesian component of each transition is considered separately. If the dipole strength D of that component is below dcutoff then the MCD is not calculated. The default is 1.0e-6.

`MCDOUT=mcdout`

Number that determines the amount of output to be printed about the MCD calculation. Higher means more output. Possible values are 0, (orientationally averaged and cartesian components of MCD parameters only) 1 (as for 0 but with the addition of a short analysis) or 2 (as for 1 but with the addition of a lengthy analysis). Theoretical analyses of MCD parameters are presented in several places including Refs. [273-274,276-278]. The default for MCDOUT is 0.

`CGOUT=cgout`

The perturbed transition densities used to evaluate the B and C term parameters can be obtained through an iterative conjugate-gradient procedure. Convergence information of the conjugate-gradient algorithm is printed every cgout iterations. Default is 10.

`NANAL=nanal, NANAL2=nanal2`

If MCDOUT is set to 2, a detailed analysis of the B and/or C term parameters in terms of which states mix and how much MCD each mixing causes, is presented. The parameters NANAL and NANAL2 determine how many contributions are included in the analyses. Defaults are 10 for NANAL and 5 for NANAL2.

`FULLOMEGA`

A standard TDDFT calculation involves the solution of an eigenvalue equation to obtain the excitation energies and transition densities of interest. ADF can solve this eigenvalue equation two ways: through diagonalization of the full Omega matrix or through the Davidson procedure where Omega is never explicitly constructed. Construction of the complete Omega matrix is generally only feasible for smaller problems. The matrix Omega appears again in the equations solved to obtain MCD. Here again Omega can be built or only the products of Omega with a vector can be used as is the case in the Davidson procedure. The default is to not construct Omega. If the keyword FULLOMEGA is included then Omega is constructed. Note that the choice of FULLOMEGA is completely independent of whether EXACT or DAVIDSON is chosen in the earlier TDDFT calculation.

`NOAB`

If this keyword is included then A and B terms are not calculated. NOAB only makes sense if SOMCD is included in the input otherwise no MCD will be calculated at all.

`NODIRECT`

The perturbed transition density needed to evaluate B and C term parameters is obtained through the solution of a large system of equations. This system of equations is solved in two ways: through a sum-over-states (SOS) type approach where the solution is expanded in a known set of transition densities or through the direct solution of the system of equations by the conjugate gradient procedure. The SOS method is much faster but also less accurate, particularly for larger systems. By default MCD parameters are evaluated through both approaches. If the NODIRECT keyword is included then only the SOS calculation is performed.

`NOCG`

The conjugate gradient procedure is first used in combination with a preconditioner that generally speeds up convergence significantly. If no solution is found in a reasonable number of iterations then the procedure is restarted without the preconditioner. If the NOCG keyword in included then the preconditioner is never used.

`CONVCG`

Convergence criterion for the CG iterative methods. The default value of 0.01 is probably good enough for most applications. This choice seems to produce B and C terms that are converged to 3 significant figures. Except for small systems, it is not recommended that CONVCG be set to a much smaller number as this will probably cause a large number of convergence failures.

`ITERCG=itercg`

Number of iterations before failure in the first (preconditioned) CG solver. This solver either succeeds quickly or not at all so the default value is 30.

`ITER2CG=iterc2g`

Number of iterations before failure in the B of C term parameter calculation of the unconditioned CG solver. This solver is often slow so the default value is 200.

`BMIN=bmin, BMAX=bmax, NBFIELD=nbfield, TMIN=tmin, TMAX=tmax, NTEMP=ntemp`

Temperature dependent MCD intensity often varies nonlinearly with T and B when T is small and/or B is large. It may therefore be of interest to evaluate the MCD intensity over a range of temperatures and/or magnetic fields. This can be achieved through the use of the BMIN, BMAX, NBFIELD, TMIN, TMAX and NTEMP keywords. The MIN and MAX keywords give the maximum values of B or T. NBFIELD and NTEMp indicate how many values are to be considered. Note that magnetic fields are assumed to be given in Tesla and temperatures in Kelvin. For example,

BMIN=1, BMAX=5, NBFIELD=5 means that fields of 1,2,3,4 and 5 T will be considered. Defaults are BMIN=BMAX=1, TMIN=TMAX=5 and NBFIELD=NTEMP=1.

```
SELECT nselect1 nselect2 nselect3...
```

Rather than selecting the first nmcdterm transitions for consideration individual transitions can be selected through the SELECT keyword. The transitions of interest are listed after the SELECT keyword. Note that the numbering follows that given in the summary table at the end of the TDDFT calculation. To consider a degenerate transition only the first component need be included. Note that it makes no sense to use both the SELECT and NMCDTERM keywords together.

```
DTENSOR Dxx Dxy Dyy Dxz Dyz
DTENSOR D E/D
```

As noted earlier, if the ZFS keyword is included with MCD and SOMCD then the influence of zero-field splitting on temperature-dependent MCD will be evaluated. As an alternative to the ZFS keyword the D-tensor parameters can be entered directly through the DTENSOR keyword in the EXCITATIONS block. Two input formats are possible. Five real numbers Dxx Dxy Dyy Dxz Dyz can be entered. These five numbers are sufficient to define the traceless tensor D. Alternatively, the two parameters D and E/D can be entered. In this case the coordinate system chosen to define the molecular geometry must be the the principle axis system of the D-tensor. D, Dxx, Dxy, Dyy, Dxz and Dyz should be given in wavenumbers (cm-1).

Notes

If an MCD calculation is run, the transition densities obtained in the TDDFT calculation are saved to TAPE21. For large molecules this can result in a very large TAPE21 file.

An MCD calculation relies on the excitation energies and, in particular, the transition densities that result from the preceding TDDFT calculation. If the results of the TDDFT calculation are poor then it is likely that the results of the MCD calculation will be poor. It therefore should be kept in mind that most TDDFT calculations will make use of the Davidson method for finding the eigenvalues and eigenvectors of the TDDFT equation. The Davidson approach involves some approximations that can lead to some variation in results with the applied parameters. The most important example of this is the fact that the results vary depending on how many eigenvalue/eigenvector pairs are calculated, ie how many transitions are selected through the LOWEST keyword. The variation is small for the eigenvalues (excitation energies) but can be significant for the eigenvectors (transition densities). A variation in the transition densities leads to variation in the transition dipoles which can significantly impact calculated MCD parameters. The moral of this story is that when calculating MCD parameters it is best to choose one value of LOWEST and stick with it.

The most time-consuming part of an MCD calculation is the solution of the system of equations through the conjugate-gradient solver. The solver can fail so be aware of warnings concerning convergence in the output. A few hints to improve convergence are: a) choose a value of LOWEST that is at least double the number of transitions for which you desire MCD parameters. This helps to improve the SOS calculation which provides an initial guess for the conjugate gradient solver. The solver is sensitive to the initial guess so changing LOWEST by a small amount may help (or hinder) convergence significantly. Keep the previous note in mind when playing with LOWEST however. b) The preconditioned conjugate gradient solver is usually fast but does not converge monotonically to the correct answer. The unpreconditioned solver is much slower but tends to converge monotonically. If the preconditioned solver fails but leaves a fairly well converged result for the unpreconditioned solver the latter usually converges quickly. If the preconditioned solver does not leave a fairly well converged result it may be worth changing the number of iterations it uses since a few iterations earlier or later may provide a much better converged answer. c) The SELECT keyword can be used to work on the remaining transitions for which converged results have not been obtained.

All MCD parameters are presented in au. To convert A and C terms to the alternative unit D^2 (Debye squared) the value in au should be multiplied by 6.46044. To convert the B term to the alternative unit of D^2/cm-1 the value in au should be multiplied by 2.94359e-05.

The A, B and C terms are defined through the equation suggested by Stephens (equation 1 in [278] and also see [273-274,107]). This equation assumes that MCD intensity varies linearly with applied magnetic field and that the temperature-dependent component varies linearly with temperature as 1/T. For the most part, these assumptions are reasonable. An exception is that the temperature-dependent part varies from linearity when T is very small. To allow for this situation a temperature and magnetic field dependent multiplicative constant (chi(B,T)) is evaluated whenever temperature-dependent MCD parameters are considered. This constant includes all magnetic field and temperature dependence of the temperature-dependent MCD. Thus chi(B,T)*C can be used in place of B*C/kT in equation 1 of [278] when MCD spectra are to be simulated. Note that, since the g-factor for all states is here approximated by 2.0, chi applies to all transitions.

Applications of the Excitation feature in ADF

It may be useful to consult the following (early) applications of the Excitation feature in ADF:

1. For excitation energies based on *exact* XC potentials: [82]
2. Calculations on Free Base Porphin: [83]; calculations on metal-porphyrins: a series of papers by Rosa, Ricciardi, Baerends, e.g. [84,85].
3. Calculations on MnO_4-, $Ni(CO)_4$ and $Mn_2(CO)_{10}$: [86]
4. Calculations on $M(CO)_5$ (M=Cr, Mo, W), using the scalar ZORA relativistic approach: [87]
5. Excitation energies of open-shell molecules: [154,157]
6. Calculations on $[PtCl_4]^{2-}$, $[PtBr_4]^{2-}$, and $[Pt(CN)_4]^{2-}$, using the ZORA relativistic approach including spin-orbit coupling: [183]
7. For details regarding the (near linear scaling and parallelized) implementation, please check Refs.[71,88]

Excited state (geometry) optimizations

Starting from ADF2010 it is possible to do excited state geometry optimizations, see Ref. [350]. Note that not all aspects of such calculations have been tested thoroughly.

With the keyword EXCITEDGO the gradients of the TDDFT excitation energy can be calculated. Naturally, the EXCITATIONS block must also be included in the input. The excitation energy gradients will only be calculated if the ground state gradients are calculated. Thus, the GEOMETRY keyword is also required.

The gradients of the excitation energy are combined with the ground state gradients to give the gradients of the excited state. These gradients can be used in much the same way as ground state gradients are used. The type of calculation is chosen in the same way as for a ground state calculation. Possible run types are:

- Geometry optimization
- Frequency analysis with numerical second derivatives: (analytical second derivatives (ANALYTICALFREQ) are not possible).
- Linear transit
- Transition state search
- IRC calculations may be possible but this possibility has not been tested yet.

In general, an option that applies to a ground state geometry optimization will also apply to an excited state geometry optimization. For example, convergence criteria can be set and constraints can be used. These options are set through the GEOMETRY block as usual. A TDDFT geometry optimization will proceed in very much the same way as a ground state geometry optimization. The major difference will be that a TDDFT calculation will take place after the SCF and before the ground state gradients are evaluated. TDDFT gradients are calculated after the ground state gradients.

Gradients for closed-shell singlet-singlet, closed shell singlet-triplet, conventional open shell and spin-flip open-shell TDDFT calculations can be evaluated. The FORCEALDA option and TDA options should be used with spin-flip calculations.

Not all functionals can be used in combination with TDDFT gradients. The following should work:
LDA: VWN, XALPHA
GGA: Any allowed combination of the Perdew86, LYP and PBEc correlation functionals and the Becke88, revPBEx, RPBEx, PBEx and OPTx exchange functionals.
Hybrid: B1LYP, B3LYP, B3LYP*, BHANDHLYP, BHANDH, O3LYP, X3LYP, B1PW91, MPW1PW, PBE0, OPBE0

QM/MM TDDFT gradients can be calculated.

Scalar relativistic effects can be included with the ZORA or mass-velocity-Darwin Hamiltonians.

At this time, gradients involving frozen cores, spin-orbit TDDFT and solvation can not be calculated.

TDDFT gradients can take advantage of symmetry but if the point group of interest includes degenerate irreducible representations then all grid points are needed in integration (equivalent to the ALLPOINTS keyword). This situation is detected automatically. This use of the full grid may make it more efficient to use a point group with only one-dimensional irreducible representations where only the symmetry-unique slice is utilized.

Degenerate excitations can be optimized. However, since in reality such degeneracies will be split by a Jahn-Teller distortion it is recommended that the symmetry of the chosen point group be lowered so that the transition of interest is no longer labeled by a degenerate representation. A Jahn-Teller distortion will not occur when the degeneracy cannot be broken by nuclear motion, e.g. for a diatomic molecule.

The EXCITEDGO block key has the following form:

```
EXCITEDGO
 {STATE Irreplab nstate}
 {SINGLET/TRIPLET}
 {OUTPUT=n}
 {CPKS EPS=err PRECONITER=precon NOPRECONITER=noprecon ITEROUT=iter}
 {EIGENFOLLOW}
END
```

`STATE Irreplab nstate`

Choose the excitation for which the gradient is to be evaluated.

`Irreplab`

Irreplab is the label from the TDDFT calculation. NOTE: the TDDFT module uses a different notation for some representation names, for example, A′ is used instead of AA.

`nstate`

This value indicates that the nstate-th transition of symmetry Irreplab is to be evaluated. Default is the first fully symmetric transition.

Note that in a numerical FREQUENCIES calculation symmetry is turned off except to reduce the number of points calculated so irrespective of the specified point group Irreplab is A in this case. Care should be taken to ensure that nstate is correct in a frequencies calculation as this number can change when the point group is changed.

`SINGLET/TRIPLET`

SINGLET: A singlet-singlet excitation is considered. The default.
TRIPLET: A singlet-triplet excitation is considered.

`OUTPUT=n`

The amount of output printed. A higher value requests more detailed output. Default: Output=0

`CPKS EPS=err PRECONITER=precon NOPRECONITER=noprecon ITEROUT=iter`

Some control parameters for the CPKS(Z-vector) part of the TDDFT gradients calculation.

`EPS=err`

err is a real number that gives the convergence requirement of the CPKS. Default is 0.0001

`PRECONITER=precon`

precon is the maximum number of iterations allowed for the preconditioned solver. Default = 30.

`NOPRECONITER=noprecon`

noprecon is the maximum number of iterations allowed for the unpreconditioned solver. Default=200.

`ITEROUT=iter`

Details of the CPKS calculation are printed every iter iterations. Default is 5.

`EIGENFOLLOW`

This key tries to follow the eigenvector in excited state geometry optimizations. In the initial implementation the target state of an excited state geometry optimization was indicated by a number and a symmetry, e.g. A2g 3 or the 3rd state of A2g symmetry. This approach becomes problematic when states cross and the state you are interested in become the 4th A2g state for example. An eigenvector-following option has been added that attempts to alleviate this problem. This option is off by default. If the subkeyword EIGENFOLLOW in is included, the state of interest in the first iteration is the same as before. In the second and subsequent iterations the state for which gradients are determined is decided on the basis of the overlap between the transition density of the transition from the previous iteration and the transition densities available in the current iteration. The same symmetry is maintained. Note that this method is not full proof. It assumes that the transition density changes only because of the contributions from the various occupied-virtual orbital pairs change but that the orbitals remain unchanged. This is not necessarily the case. Secondly, the sign of the transition density components is not taken into account.

At each iteration of a TDDFT-gradients calculation the excited state electric dipole moment is also calculated. If the Output parameter is 1 or greater then the excited state dipole moment will be printed out.

Vibrationally resolved electronic spectra

To calculate vibrational effects on the electronic excitations (Uv/vis, X-ray), one needs to do a frequency calculation both at the ground state as well as the excited state of interest. Next Franck-Condon factors need to be calculated for the transition between the two electronic states, which can be done with the FCF program, described below. These Franck-Condon factor can then be used to predict the relative intensities of absorption or emission lines in the electronic spectra. Note that the Herzberg-Teller effect is not taken into account.

FCF program: Franck-Condon Factors

fcf is an auxiliary program which can be used to calculate Franck-Condon factors from two vibrational mode calculations [418].

fcf requires an ascii input file where the user specifies the TAPE21 files from two *adf* vibrational mode calculations, carried out for two different electronic, spin or charge states of the same molecule. These calculations can be either numerical or analytical. The number of vibrational quanta that have to be taken into account for both states in the evaluation of the Franck-Condon factors have to be specified.

fcf produces a (binary) KF file TAPE61, which can be inspected using the KF utilities. Furthermore, *fcf* writes the frequencies, vibrational displacements and electron-phonon couplings for both states too the standard output, including any error messages.

Introduction

Franck-Condon factors are the squares of the overlap integrals of vibrational wave functions. Given a transition between two electronic, spin or charge states, the Franck-Condon factors represent the probabilities for accompanying vibrational transitions. As such, they can be used to predict the relative intensities of absorption or emission lines in spectroscopy or excitation lines in transport measurements.

When a molecule makes a transition to another state, the equilibrium position of the nuclei changes, and this will give rise to vibrations. To determine which vibrational modes will be active, we first have to express the displacement of the nuclei in the normal modes:

$$\mathbf{k}=\mathbf{L}'^{T}\mathbf{m}^{1/2}(\mathbf{B}_0\mathbf{x}_0-\mathbf{x}'_0)$$

Here, **k** is the displacement vector, **L** is the normal mode matrix, **m** is a matrix with the mass of the nuclei on the diagonal, **B** is the zero-order axis-switching matrix and $\mathbf{x}_0$ is the equilibrium position of the nuclei. For free molecules, depending on symmetry constraints, the geometries of both states may have been translated and/or rotated with respect to each other. To remove the six translational and rotational degrees of freedom, we can center the equilibrium positions around the center of mass and rotate one of the states to provice maximum overlap. The latter is included with the zero-order axis-switching matrix **B**, implemented according to [419].

When we have obtained the displacement vector, it is trivial to calculate the dimensionless electron-phonon couplings. They are given by:

$$\boldsymbol{\lambda}=(\boldsymbol{\Gamma}/2)^{1/2}\mathbf{k}$$

Here, $\boldsymbol{\Gamma}=2\pi\boldsymbol{\omega}/h$ is a vector containing the reduced frequencies. [420]. The Huang-Rhys factor **g** is related to **λ** as:

$$\mathbf{g} = \boldsymbol{\lambda}^2$$

The reorganization energy per mode is

$$\mathbf{E} = (h/2\pi)\ \boldsymbol{\omega}\ \boldsymbol{\lambda}^2$$

When the displacement vector **k**, the reduced frequencies **Γ** and **Γ'**, and the Duschinsky rotation matrix $\mathbf{J}=\mathbf{L}'^{T}\mathbf{B}_0\mathbf{L}$ have been obtained, the Franck-Condon factors can be calculated using the two-dimensional array method of Ruhoff and Ratner[420].

There is one Franck-Condon factor for every permutation of the vibrational quanta over both states. Since they represent transition probabilities, all Franck-Condon factors of one state which respect to one vibrational state of the other state must sum to one. Since the total number of possible vibrational quanta,

and hence the total number of permutations, is infinite, in practice we will calculate the Franck-Condon factors until those sums are sufficiently close to one. Since the number of permutations rapidly increases with increasing number of vibrational quanta, it is generally possible to already stop after the sum is greater than about two thirds. The remaining one third will be distributed over so many Franck-Condon factors that their individual contributions are negligible.

In the limiting case of one vibrational mode, with the same frequency in both states, the expression for the Franck-Condon factors of transitions from the ground vibrational state to an excited vibrational state are given by the familiar expression:

$$|I_{0,n}|^2=e^{-\lambda^2}\lambda^{2n}/n!$$

Input

The input for *fcf* is keyword oriented and is read from the standard input. *fcf* recognizes several keywords, but only two have to be specified to perform the calculation. All input therefore contains at least two lines of the following form:

```
$ADFBIN/fcf << eor
STATES state1 state2
QUANTA l1 l2
eor
```

`STATES state1 state2`

The filenames of two TAPE21 files resulting from a numerical or analytical frequency calculation. The calculations must have been performed on the same molecule, i.e. the type, mass and order of occurrence of all the atoms (or fragments) has to be the same in both files.

`(optional) MODES first last`

The first and last mode to be taken into account in the calculation. If this option is omitted, all modes are taken into account. This option can be used to effectively specify and energy range for the Franck-Condon factors. When using this options, always check if the results (electron-phonon couplings, ground state to ground overlap integral, average sum of Franck-Condon factors, etc.) do not change too much.

`(optional) LAMBDA lambda`

The minimum value of the electron-phonon coupling for a mode to be taken into account in the calculation. The default value is zero. Together with the MODES option, this provides a way to significantly reduce the total number of Franck-Condon factors. As with the MODES option, always check if the results do not change too much.

`QUANTA l1 l2`

The maximum number of vibrational quanta to be taken into account for both states. Franck-Condon factors will be calculated for every permutation of up to and including l1/l2 quanta over the vibrational modes.

`(optional) TRANSLATE`

Move the center of mass of both geometries to the origin.

`(optional) ROTATE`

Rotate the geometries to maximize the overlap of the nuclear coordinates.

```
(optional) SPECTRUM freqmin freqmax nfreq
```

If SPECTRUM is included the vibrational spectrum is calculated. A histogram of the spectrum is calculated for the frequency range that is provided on input. The three parameters that define the frequency range are:

`freqmin`

minimum frequency for which the spectrum is calculated.

`freqmax`

maximum frequency for which the spectrum is calculated.

`nfreq`

number of frequencies for which the spectrum is calculated.

Only a few keys from the TAPE21 file are used for the calculation of the Franck-Condon factors. Disk space usage can be significantly reduced by extracting just these keys from the TAPE21 file before further analysis. The following shell script will extract the keys from the KF file specified by the first argument and store them in a new KF file specified by the second argument using the *cpkf* utility:

```
#!/bin/sh
cpkf $1 $2 "Geometry%nr of atoms" "Geometry%xyz" "Geometry%nr of
atomtypes" \
"Geometry%fragment and atomtype index" "Geometry%atomtype"
"Geometry%mass" \
"Freq%Frequencies" "Freq%Normalmodes"
```

Result: TAPE61

After a successful calculation, *fcf* produces a TAPE61 KF file. All results are stored in the Fcf section:

contents of TAPE61	comments
firstmode, lastmode	the first and last vibrational mode taken into account
lambda	the minimum value of the electron-phonon coupling
maxl1, maxl2	maximum level (or maximum number of vibrational quanta) in both states
translate, rotate	whether the TRANSLATE and ROTATE options were specified in the input
natoms	number of atoms in the molecule
mass	atomic mass vector (**m**)
xyz1, xyz2	equilibrium geometries of both states ($\mathbf{x}_0$ and $\mathbf{x}_0$')
b0	zero-order axis-switching matrix matrix ($\mathbf{B}_0$)
nmodes	number of vibrational modes with a non-zero frequency
gamma1, gamma2	reduced frequencies of both states (**Γ** and **Γ**')
lmat1, lmat2	mass-weighted normal modes of both states (**L** and **L**')
jmat	Duschinsky rotation matrix (**J**)
kvec1, kvec2	displacement vectors for both states (**k** and **k'**, kvec1 is used for the calculation of the Franck-Condon factors)
lambda1, lambda2	electron-phonon couplings for both states (**λ** and **λ**')
maxp1, maxp2	maximum number of permutations of maxl1/maxl2 quanta over the vibrational modes
i0	ground state to ground state overlap integral ($I_{0,0}$)

freq1, freq2	frequencies of every permutation of the vibrational quanta for both states
fcf	maxp1 by maxp2 Franck-Condon factor matrix
fcfsum1, fcfsum2	average sum of the Franck-Condon factors for both states

In addition to producing a binary TAPE61 file, *fcf* also writes the frequencies, displacement vectors and electron-phonon couplings for both states to the standard output.

Example absorption and fluorescence

In this example it is assumed that the molecule has a singlet ground state S_0, and the interesting excited state is the lowest singlet excited state S_1. First one needs to a the ground state geometry optimization, followed by a frequency calculation. The TAPE21 of the ground state frequency calculation will be called s0.t21. Next one needs to do an excited state geometry optimization. Here it is assumed that the lowest singlet excited state S_1 is of interest:

```
EXCITATION
   Onlysing
   Lowest 1
END
EXCITEDGO
   State A 1
   Singlet
END
GEOMETRY
END
```

To get the frequencies for this excited state, numerical frequencies need to be calculated, at the optimized geometry of the first excited state.

```
EXCITATION
ONLYSING
lowest 1
END
EXCITEDGO
    State A 1
    Singlet
END
GEOMETRY
    frequencies
END
...
mv TAPE21 s1.t21
```

Next for the absorption spectrum, we look at excitations from the lowest vibrational state of the electronic ground state to the vibrational levels of the first singlet excited state S_1 ($S_1 \leftarrow S_0$), using the FCF program, which calculates the Franck-Condon factors between the vibrational modes of the two electronic states, with input

```
STATES s0.t21 s1.t21
QUANTA 0 5
SPECTRUM 0 10000 1001
TRANSLATE
ROTATE
```

The number of vibrational quanta for the excited state should be larger in case of small molecules. See the description of FCF program for more details.

For the fluorescence spectrum, we look at excitations from the lowest vibrational state of the first singlet excited state S_1 to the vibrational levels of the singlet ground state state S_0 ($S_1 \rightarrow S_0$). Input for the FCF program is in this case:

```
STATES s0.t21 s1.t21
QUANTA 5 0
SPECTRUM -10000 0 1001
TRANSLATE
ROTATE
```

The number of vibrational quanta for the ground state should be larger in case of small molecules.

Note that the FCF program calculates the spectrum relative to the 0-0 transition. Thus one should add to spectrum calculated with FCF the difference in energy of the lowest vibrational state of the ground state S_0 and the lowest vibrational state of the electronically singlet excited state S_1.

Example phosphorescence

In this example it is assumed that the molecule has a singlet ground state S_0, and the interesting excited state is the lowest triplet excited state T_1. Emission from a triplet state to a singlet state is spin forbidden, however, due to spin-orbit coupling such transitions may occur. In the following we assume that the geometry of the triplet excited state is not influenced much by spin-orbit coupling.

First one needs to a the ground state geometry optimization, followed by a frequency calculation. The TAPE21 of the ground state frequency calculation will be called s0.t21. Next one needs to do an excited state geometry optimization of the lowest triplet excited state, followed by a frequency calculation.

```
CHARGE 0.0 2.0
UNRESTRICTED
GEOMETRY
END
AnalyticalFreq
End
...
mv TAPE21 t1.t21
```

For the phosphorescence spectrum, we look at excitations from the lowest vibrational state of the first triplet excited state T_1 to the vibrational levels of the singlet ground state state S_0 ($T_1 \rightarrow S_0$). Input for the FCF program is in this case:

```
STATES s0.t21 t1.t21
QUANTA 5 0
SPECTRUM -10000 0 1001
TRANSLATE
ROTATE
```

The number of vibrational quanta for the ground state should be larger in case of small molecules.

Note that the FCF program calculates the spectrum relative to the 0-0 transition. Thus one should add to spectrum calculated with FCF the difference in energy of the lowest vibrational state of the ground state S_0 and the lowest vibrational state of the electronically triplet excited state T_1.

Zero field splitting (ZFS) and the radiative rate constants (i.e. radiative phosphorescence lifetimes) could be calculated with spin-orbit coupled ZORA time-dependent density functional theory (ZORA-TDDFT). In ADF spin-orbit couping can be treated self-consistently (i.e. non perturbatively) during both the SCF and TDDFT parts of the computation, see the section on excitation energies and spin-orbit coupling.

An alternative to the use of the unrestricted formalism to calculate the lowest triplet excited state is to use the TDDFT formalism:

```
EXCITATION
   Onlytrip
   Lowest 1
END
EXCITEDGO
   State A 1
   Triplet
END
GEOMETRY
END
```

To get the frequencies for this excited state, numerical frequencies need to be calculated, at the optimized geometry of the first excited state.

```
EXCITATION
   Onlytrip
   Lowest 1
END
EXCITEDGO
   State A 1
   Triplet
END
GEOMETRY
   frequencies
END
...
mv TAPE21 t1.t21
```

(Hyper-)Polarizabilities, ORD, magnetizabilities, Verdet constants

A (frequency dependent) electric field induces a dipole moment in a molecule, which is proportional to the (frequency dependent) molecular polarizability. Van der Waals dispersion coefficients describe the long-range dispersion interaction between two molecules. Optical rotation or optical activity (ORD) is the rotation of linearly polarized light as it travels through certain materials. A (frequency dependent) magnetic field induces a magnetic moment in a molecule, which is proportional to the (frequency dependent) molecular magnetizability. The Faraday effects describes the rotation of the plane-polarized light due to a magnetic field, which is proportional to the intensity of the component of the magnetic field in the direction of the beam of light. The Verdet constant describes the strength of the Faraday effect for a particular molecule. All these properties are available in ADF as applications of time-dependent DFT (TDDFT).

Polarizabilities

The calculation of frequency-dependent (hyper)polarizabilities and related properties (Raman, ORD) is activated with the block key RESPONSE

```
RESPONSE
END
```

In this example only the *zz* component of the dipole polarizability tensor is calculated, at zero frequency. The orientation of the molecule is therefore crucial. Be aware that the program may modify the orientation of the molecule if the input coordinates do not agree with the symmetry conventions in ADF!
(This calculation could equivalently be done through a finite field method).

See also the alternative implementation with the AORESPONSE key that offers some unique features like magnetizability, and lifetime options.

The impact of various approximations on the quality of computed polarizabilities has been studied in, for instance, Refs. [74,82,89]. If you are new to this application field, we strongly recommend that you study a few general references first, in particular when you consider hyperpolarizability calculations. These have many pitfalls, technically (which basis sets to use, application of the DEPENDENCY key) and theoretically (how do theoretical tensor components relate to experimental quantities, different conventions used). Please, take a good look both at ADF-specific references [75-77,90] and at general references related to this subject: Refs. [91-93], the entire issues of Chem.Rev.**94**, the ACS Symposium Series #628, and further references in the ADF-specific references.

```
RESPONSE
 ALLCOMPONENTS
 Nfreq Nfreq
 FrqBeg FirstFreq
 FrqEnd LastFreq
 [Optional Frequency/Energy Unit]
 ALLTENSOR
 Quadrupole
 Octupole
END
```

`Entire tensor or only one component`

You specify the ALLCOMPONENTS subkey to get the entire polarizability tensor, instead of just the *zz* component.

`Frequencies or wavelengths`

Instead of performing the calculation at zero frequency (which results in the *static* polarizability), one can specify an even-spaced sequence of frequencies, using the subkeys Nfreq, FrqBeg, and FrqEnd with obvious meaning. The (first and last) frequency values are by default in eV. This can be changed into Hartree units (a.u.) or in wavelengths (angstroms) by typing HARTREE or ANGSTROM on a separate line within the RESPONSE block, instead of [Optional Frequency/Energy Unit].

`Higher multipole polarizabilities`

Instead of just calculating the dipole-dipole polarizability, one may address the dipole-quadrupole, quadrupole-quadrupole, dipole-octupole, quadrupole-octupole, and octupole-octupole polarizability tensors. These can all be calculated in a single run, using the subkey ALLTENSOR. If only quadrupole-quadrupole or octupole-octupole tensors are needed, the subkey quadrupole or octupole should be used.

Accuracy and convergence, RESPONSE key

```
RESPONSE
 erralf 1e-6
 erabsx 1e-6
 errtmx 1e-6
```

```
 ncycmx 30
END
```

`erralf, erabsx, errtmx`

The subkeys erralf, erabsx, errtmx determine the convergence criteria for a polarizability calculation. The strict defaults are shown. It is rarely necessary to change the defaults, as these are rather strict but do not lead to many iterations.

`ncycmx`

The maximum number of attempts within which the algorithm has to converge. The default appears to be adequate in most cases.

Hyperpolarizabilities

`Hyperpolarizabilities`

```
RESPONSE
 HYPERPOL LaserFreq
END
```

The first hyperpolarizability tensor b is calculated (in atomic units in the 'theoreticians convention', i.e. convention T=AB in Ref. [92]) if the subkey HYPERPOL is present with a specification of the laser frequency (in hartree units). If also the subkey ALLCOMPONENTS is specified, all components of the hyperpolarizability tensor will be obtained.

As mentioned before, by default only the static dipole hyperpolarizability tensor is computed. If one is interested in the frequency-dependent hyperpolarizability, the input could look like:

```
RESPONSE
 ALLCOMPONENTS
 HYPERPOL 0.01
 DYNAHYP
END
```

The subkey DYNAHYP has to be added and the main frequency ω has to be specified in Hartrees after the subkey hyperpol. In the output all nonzero components of the tensors governing the static first hyperpolarizability, second harmonic generation, electro-optic pockels effect, and optical rectification are printed.

Note: Second hyperpolarizabilities are currently not available analytically. Some can however be obtained by calculating the first hyperpolarizability in a finite field.
The effect of using different DFT functionals (LDA, LB94, BLYP) on hyperpolarizabilities in small molecules has been investigated in [77].

Van der Waals dispersion coefficients

```
RESPONSE
 ALLCOMPONENTS
 VANDERWAALS NVanderWaals
 {ALLTENSOR}
END
```

`Dispersion coefficients`

Simple dispersion coefficients (the dipole-dipole interaction between two identical molecules, the C_6 coefficient) are calculated in a single ADF calculation. General dispersion coefficients are obtained with the auxiliary program DISPER, which uses two output files (file named TENSOR) of two separate ADF runs as input. See the Properties and the Examples documents.
To get the dispersion coefficients one has to calculate polarizabilities at imaginary frequencies between 0 and infinity. The ADF program chooses the frequencies itself. The user has to specify the number of frequencies, which in a sense defines the level of accuracy, as an argument to the subkey VanDerWaals.

`NVanderWaals`

One can specify the number of frequencies with NVanderWaals. Ten frequencies is reasonable. Without the key ALLTENSOR only dipole-dipole interactions are considered. If ALLTENSOR is specified, higher dispersion coefficients are also calculated. This ADF calculation generates a file with name TENSOR, which contains the results of multipole polarizabilities at imaginary frequencies. This TENSOR file has to be saved. Similarly, the TENSOR file for the second monomer has to be saved. The files have to be renamed to files 'tensorA' and 'tensorB' (case sensitive) respectively. Then the program DISPER has to be called in the same directory where the 'tensorA' and 'tensorB' files are located. DISPER needs no further input.

DISPER program: Dispersion Coefficients

The DISPER program was originally written by V.Osinga [79]. The original documentation was written by S.J.A. van Gisbergen.

Van der Waals dispersion coefficients

The program DISPER computes Van der Waals dispersion coefficients up to C_{10} for two arbitrary closed-shell molecules. ADF itself can already compute some C_6 and C_8 coefficients between two identical closed-shell molecules. These coefficients describe the long-range dispersion interaction between two molecules. It requires previous ADF-TDDFT calculations for the polarizability tensors at imaginary frequencies for the two interacting molecules. Each such ADF calculation produces a file TENSOR (if suitable input for ADF is given). The TENSOR files must be renamed tensorA and tensorB, respectively and must be present as local files for DISPER. The DISPER program takes no other input and prints a list of dispersion coefficients.

A schematic example, taken from the set of sample runs, for the usage of DISPER is the following:

Step1: run ADF for, say, the HF molecule. In the input file you specify the RESPONSE data block:

```
RESPONSE
 MaxWaals 8     ! Compute dispersion coefficients up to C8
 ALLTENSOR      ! This option must be specified in the ADF calc for a
                ! subsequent DISPER run
 ALLCOMPONENTS ! Must also be specified for DISPER
End
```

At the end of the run, copy the local file 'TENSOR' to a file 'tensorA'. For simplicity, we will now compute the dispersion coefficients between two HF molecules. Therefore, copy 'tensorA' to 'tensorB'.

Now run DISPER (without any other input). It will look for the local files 'tensorA' and 'tensorB' and compute corresponding dispersion coefficients to print them on standard output.

```
$ADFBIN/disper -n1 << eor
eor
```

The output might look something like this:

```
DISPER 2000.02 RunTime: Apr04-2001 14:14:13
********** C-COEFFICIENTS **********
 n  LA  KA  LB  KB  L  coefficient(Y)    coefficient(P)
 6   0   0   0   0  0   28.29432373        28.29432373
 6   2   0   0   0  2    7.487547697        3.348533127
 8   0   0   0   0  0  416.1888455        416.1888455
 8   0   0   2   0  2    0.4323024202E-05   0.1933315197E-05
 8   2   0   0   0  2  402.3556946        179.9389368
 8   2   0   2   0  4    0.4238960180E-05
 8   4   0   0   0  4  -36.67895539       -12.22631846
 8   4   0   2   0  6   -0.2000286301E-05
```

The n-value in the first column refers to the long-range radial interaction. The case n=6 refers to the usual dipole-dipole type interaction related to a $1/R^6$ dependence in the dispersion energy. The n=7 case relates to a dipole-quadrupole polarizability on one system and a dipole-dipole polarizability on the other (this is not symmetric!). The n=8 term may contain contributions from a quadrupole-quadrupole polarizability on one system in combination with a dipole-dipole polarizability on the other as well as contributions from two dipole-quadrupole polarizabilities.

Terms which are zero by symmetry are not printed. In the example above, this is the case for all n=7 terms, because the systems (apparently) are too symmetric to have a nonzero dipole-quadrupole polarizability. The best known and most important coefficients are the isotropic ones, determining the purely radial dependence of the dispersion energy. They are characterized by the quantum numbers: 6 0 0 0 0 0 (or 8 0 0 0 0 0 etc.) Other combinations of quantum numbers refer to different types of angular dependence. The complete set determines the dispersion energy for arbitrary orientations between the two subsystems A and B.

The complete expressions are rather involved and lengthy. We refer the interested reader to the paper [79] which contains a complete description of the meaning of the various parts of the output, as well as references to the earlier literature which contain the mathematical derivations. In particular, a useful review, which was at the basis of the ADF implementation, is given in [414]. Of particular significance is Eq.(8) of the JCP paper mentioned above, as it defines the meaning of the calculated coefficients $C_n^{(L_A,K_A,L_B,K_B,L)}$ as printed above.

For highly symmetric systems, a different convention is sometimes employed. It is based on Legendre polynomials (hence the 'P' in the final column) instead of on the spherical harmonics (the 'Y' in the column before the last). The 'P' coefficients are defined only for those coefficients that are nonzero in highly symmetric systems and never contain additional information with respect to the 'Y' coefficients. They are defined [Eq. (14) in the mentioned J. Chem. Phys. paper] in terms of the 'Y' coefficients by:

$$C_n^L = (-1)^L C_n^{L,0,0,0,L}/\sqrt{(2L+1)}$$

Because the quality of the dispersion coefficients is determined by the quality of the polarizabilities that are the input for DISPER, it is important to get good polarizabilities from ADF. For that it is important, in the case of small systems, to use an asymptotically correct XC potential (several choices are available in ADF, such as SAOP or GRAC) and a basis set containing diffuse functions. We refer to the ADF User's Guide for details.

Optical rotation dispersion (ORD)

```
RESPONSE
 OPTICALROTATION
END
```

```
OPTICALROTATION
```

With the subkey OPTICALROTATION the (frequency dependent) optical rotation [80,94] will be calculated. For correct calculations one should calculate the entire tensor (see also the subkey ALLCOMPONENTS), which is done automatically.

An alternative implementation uses the AORESPONSE key, in which life time effects can be included.

AORESPONSE: Lifetime effects, polarizabilities, ORD, magnetizabilities, Verdet constants

The AORESPONSE key offers some unique features compared to the RESPONSE key, namely lifetime effects (polarizabilities at resonance), polarizabilities in case of spin-orbit coupling, the calculation of (dynamic) magnetizabilities, Verdet constants, the Faraday B terms, and an alternative way to calculate (resonance) Raman scattering factors. Note that the RESPONSE key also has many unique features, like the use of symmetry during the calculation.

AORESPONSE key

If the block key AORESPONSE is used, by default, the polarizability is calculated. This can be modified using one of the keys below. Note that if the molecule has symmetry the key ALLPOINTS should be included.

```
AORESPONSE
 OPTICALROTATION
 VELOCITYORD
 MAGNETICPERT
 MAGOPTROT
 RAMAN
 FREQUENCY      Nfreq freq1 freq2 ... freqN units
 FREQRANGE      freq1 freqN TotFreq units
 LIFETIME width
 ALDA|XALPHA
END
ALLPOINTS
```

`OPTICALROTION`

Specify OPTICALROTION to calculate optical rotatory dispersion spectrum instead of polarizabilities.

`VELOCITYORD`

This option should be used instead of OPTICALROT with GIAO if the finite lifetime effects need to be taken into account (LIFETIME option).

`MAGNETICPERT`

Calculate static or time-dependent magnetizability, see also Ref. [230].

`MAGOPTROT`

Specify MAGOPTROT to calculate the Verdet constant instead of polarizability, see for the details of the implementation Ref. [307]. When it is specified together with the LIFETIME key the real and imaginary part of the damped Verdet constant will be calculated. Combination of three keys MAGOPTROT, LIFETIME and FREQRANGE yields the magnetic optical rotatory dispersion and magnetic circular dichroism spectrum (Faraday A and B terms) calculated simultaneously in the range from freq1 to freqN. It is also possible to combine MAGOPTROT, LIFETIME and FREQUENCY. In order to obtain the

Faraday B terms from the Verdet constant calculations it is necessary to perform several steps, involving a fit of the imaginary Verdet data to the MCD spectrum. You can request SCM for details on the fitting procedure. For details of the method, see Ref. [272].

`RAMAN`

Calculates the Raman scattering factors. The AOREPONSE-Raman only works with one frequency. If one frequency is specified the Raman scattering factors are calculated at that frequency. The Raman option is compatible with the lifetime option so that resonance Raman scattering can be calculated. For details of this method, see Ref. [266]. To get Raman intensities with AORESPONSE, numerical frequencies need to be calculated using a FREQUENCIES key in the GEOMETRY input block. Non-resonance Raman intensities can also be obtained using the RESPONSE key or, alternatively, using RAMANRANGE in combination with analytically or numerically pre-calculated frequencies.

`FREQUENCY Nfreq freq1 freq2 ... freqN units`

To calculate time-dependent properties, one needs to specify frequency of perturbation field. Here Nfreq specifies the number of frequencies that follow. The last item on the line specifies the units and is one of EV, HARTREE, ANGSTROM.

`FREQRANGE freq1 freqN TotFreq units`

This key is useful when it is necessary to specify more than 20 equally spaced frequencies for the response calculations. The first frequency is freq1 and the last one is freqN. The total number of frequencies including the first and the last one is TotFreq. The last item specifies the units: EV, HARTREE or ANGSTROM.

`LIFETIME width`

Specify the resonance peak width (damping) in Hartree units. Typically the lifetime of the excited states is approximated with a common phenomenological damping parameter. Values are best obtained by fitting absorption data for the molecule, however, the values do not vary a lot between similar molecules, so it is not hard to estimate values. A value of 0.004 Hartree was used in Ref. [266].

`ALDA|XALPHA`

If ALDA is specified the VWN kernel is used. This option is the default. If ALPHA is specified the Xα kernel is used instead of the default VWN one.

The spin-orbit ZORA polarizability code (Ref. [311]) is automatically selected if the AORESPONSE keyword is given in a spin-orbit coupled calculation. In this case a spin-restricted calculation is required, but, unlike the rest of AORESPONSE, also SYMMETRY NOSYM. Spin-polarization terms in the XC response kernel are neglected. In Ref. [311] the imaginary polarizability dispersion curves (spin-restricted) match well the broadened spin-orbit TDDFT data from Ref. [182]. Thus the corrections from the spin-polarization terms appear to be rather minor. No picture change corrections were applied in the ZORA formalism.

Technical parameters and expert options

```
AORESPONSE
 ...
 SCF        {NOCYC} {NOACCEL} {CONV=conv} {ITER=niter}
 GIAO
 FITAODERIV
END
```

`SCF {NOCYC} {NOACCEL} {CONV=conv} {ITER=niter}`

Specify CPKS parameters such as the degree of convergence and the maximum number of interations:
NOCYC - disable self-consistence altogether
NOACCEL - disable convergence acceleration

CONV - convergence criterion for CPKS. The default value is 10^{-6}.
The value is relative to the uncoupled result (i.e. to the value without self-consistence).
ITER - maximum number of CPKS iterations, 50 by default.
Specifying ITER=0 has the same effect as specifying NOCYC.

`GIAO`

Include the Gauge-Independent Atomic Orbitals (GIAO). This option should not be used with damping (LIFETIME keyword) and the VELOCITYORD option should be used instead.

`FITAODERIV`

Use fitted AO Derivatives. This will improve the density fitting, can only be used in cae of STO fitting. In case of ZlmFit one can improve the fitting with the ZLMFIT block key.

`COMPONENTS {XX} {XY} {XZ} {YX} {YY} {YZ} {ZX} {ZY} {ZZ}`

Limit the tensor components to the specified ones. Using this option may save the computation time.

Applications of AORESPONSE

It may be useful to consult the following applications of the AORESPONSE key in ADF:

1. Calculation of static and dynamic linear magnetic response in approximate time-dependent density functional theory [230]
2. Calculation of CD spectra from optical rotatory dispersion, and vice versa, as complementary tools for theoretical studies of optical activity using time-dependent density functional theory [231]
3. Calculation of origin independent optical rotation tensor components for chiral oriented systems in approximate time-dependent density functional theory [232]
4. Time-dependent density functional calculations of optical rotatory dispersion including resonance wavelengths as a potentially useful tool for determining absolute configurations of chiral molecules [233]
5. Calculation of optical rotation with time-periodic magnetic field-dependent basis functions in approximate time-dependent density functional theory [234]
6. A Quantum Chemical Approach to the Design of Chiral Negative Index Materials [235]
7. Calculation of Verdet constants with time-dependent density functional theory. Implementation and results for small molecules [236]
8. Calculations of resonance Raman [266,267]
9. Calculations of surface-enhanced Raman scattering (SERS) [268,269]
10. Calculation of magnetic circular dichroism spectra from damped Verdet constants [272]
11. Calculation of the polarizability in case of spin-orbit coupling [311]

NMR

NMR chemical shifts and NMR spin-spin couplings can be calculated. Effects due to spin-orbit coupling can be included. All electron basis sets can be used.

The separate program EPR/NMR ($ADFBIN/epr) program is no longer documented, since most of its capabilities are implemented in newer modules. See for the old documentation the ADF2010 EPR/NMR module documentation.

NMR Chemical Shifts

NMR Chemical shifts have been implemented [113-117] in a separate property program NMR. It requires the TAPE21 result file from an ADF calculation. The NMR module can be combined with the ZORA treatment for relativistic effects and with Spin-Orbit effects, making it suitable for treatment of heavy elements. See also the general review on relativistic computations of NMR parameters [331].

Important notes

TAPE21 and TAPE10

NMR requires an ASCII input file and TAPE21 and TAPE10 result files from an ADF calculation on the molecule to be analyzed. The ADF result files TAPE21 and TAPE10 must be present with names TAPE21 and TAPE10 in the directory where you execute NMR. Use the keywords SAVE TAPE10 in the adf calculation in order to obtain a TAPE10 result file.

Reaclculation of TAPE10 by NMR

Warning: the NMR property program will not always give the correct result for every SCF potential in the ADF calculation, like for example the SAOP potential, or if one uses COSMO in the ADF calculation, if one lets the NMR recalculate TAPE10. This is due to the GIAO method used in this program, which requires the calculation of the SCF potential, which is not done correctly for potentials, other than the standard LDA and GGA potentials. To obtain correct results one should, in addition to the use of TAPE21, also use TAPE10 that ADF generates, using the keywords SAVE TAPE10, and use it as input for the NMR property program. On TAPE10 the SCF potential is written, which is read in by the NMR program.

Atomic calculation

NMR calculations on 1 atom must have symmetry NOSYM.

Spin-orbit coupling

NMR calculations on systems computed by ADF with Spin Orbit relativistic effects included must have used NOSYM symmetry in the ADF calculation. NMR can also be combined with ADF ZORA calculations. The NMR program reads from TAPE21 the relativistic option that is used in the ADF calculation, and will use the same relativistic option in the NMR calculations.

Bug spin-orbit part ADF2008 - ADF2013

In the ADF2008.01 a bug was introduced in the spin-orbit part of the calculated chemical shielding, which caused the calculated chemical shielding to be gauge dependent. This bug is relevant for spin-orbit coupled calculations for ADF versions ADF2008-ADF2013. In ADF2014 this bug has been fixed.

Unscaled ZORA default ADF2014

There is gauge dependence if the scaled ZORA method is used in the calculation of NMR chemical shieldings. Therefore the default method for NMR chemical shielding calculations is changed in ADF2014 to use the unscaled ZORA method.

SAOP

The use of the model SAOP potential leads to isotropic chemical shifts which are substantially improved over both LDA and GGA functionals, and of similar accuracy as results with a self-interaction-corrected functional (SIC), see [421]. SAOP is computationally expedient and routinely applicable to all systems, requiring virtually the same computational effort as LDA and GGA calculations.

NICS

The Nucleus-Independent Chemical Shift (NICS) can be calculated at any point in the molecule.

Hybrids

Starting from ADF2009.01 Hartree-Fock and the hybrid potentials can used in combination with NMR chemical shielding calculations. see Refs. [422,423]. Use SAVE TAPE10 in the ADF calculation. The use of frozen cores and hybrids gives gauge dependent results for the NMR chemical shieldings, therefore the NMR program will stop in this case.

Meta-GGA's and meta-hybrids

Meta-GGA's and meta-hybrids should not be used in combination with NMR chemical shielding calculations. The results are wrong due to an incorrect inclusion of GIAO terms.

Bug fix ADF2005.01 off-diagonal part shielding tensor

Bug fix off-diagonal part shielding tensor: In the ADF2005.01 the bugs in the NMR module are fixed that gave problems in the ADF2004.01 and older versions.

Input options

The input file for NMR uses the block key NMR, with several (optional) sub keys, each having a series of options. For analysis a separate block key can be used.

```
$ADFBIN/nmr << eor
ZSOAO2007
RECALCULATETAPE10
NMR
   ...
End
Analysis
   ...
End
eor
```

ZSOAO2007 keyword

In the ADF2008.01 a bug was introduced in the spin-orbit part of the calculated chemical shielding, which caused the calculated chemical shielding to be gauge dependent. This bug is relevant for spin-orbit coupled calculations for ADF versions ADF2008-ADF2013. Workaround in ADF versions ADF2008-ADF2013 is to include the keyword ZSOAO2007 in the NMR part of the input, which causes a one-center approximation to be used. The bug has been fixed in ADF2014 by introducing an extra gauge-correction term for the spin-orbit coupled part. One can still get the (slightly) incorrect results in ADF2014 by using the keyword WRONGSOGAUGE, and not including ZSOAO2007.

REACALCULATETAPE10 keyword

If there is no TAPE10 present the NMR program will stop. One can use the key REACALCULATETAPE10 such that TAPE10 will be recalculated by the NMR module. Not recommendended to be used. Better use 'SAVE TAPE10' in the ADF calculation, and use this TAPE10 as input for NMR.

NMR block key

```
NMR
   Out  OutOptions
```

```
    Calc CalcOptions
    Use UseOptions
    U1K  U1KOptions
    Nuc  NucOptions
    Atoms  AtomsOptions
    Ghosts  GhostsOptions
    Analysis  AnalysisOptions
End
```

`Out OutOptions`

The sub key Out controls printed output. Its options specify the details by their (optional) presence. The following OutOptions are recognized (Default ISO):

`All`

Implies all the other options except for 'ISO', which may be specified in addition.

`ISO`

Isotropic shielding constants

`Tens`

Shielding tensors

`Eig`

Eigenvectors

`U1`

The U1 matrix

`F1`

The first order change in the Fock matrix

`S1`

The first order change in the Overlap matrix

`AOP`

The paramagnetic AO matrix (= the matrix in the representation of elementary atomic basis functions)

`AOD`

The diamagnetic AO matrix

`AOF`

The Fermi-contact AO matrix

`REFS`

Literature references

`INFO`

General information

`Calc CalcOptions`

The sub key Calc controls what is actually calculated. The following options are available (Default ALL):

`All`

Implies all of the other options to this key

`Para`

The paramagnetic part

`Dia`

The diamagnetic part

`FC`

The Fermi-contact part in case of the Pauli Hamiltonian

`SO`

The Fermi-contact part in case of the ZORA Hamiltonian

`Use UseOptions`

The sub key Use controls some optional options (default none)

`SCALED`

Implies the scaled ZORA method, which gives (slightly) gauge dependence results. Note that in case of the ZORA Hamiltonian default the unscaled ZORA method is used. For chemical shifts, only compare results with the same options.

`SO1C`

Before ADF2008.01 in the the spin-orbit term a 1-center approximation was used, which does not suffer from gauge dependence. This 1-center approximation can be used with USE SO1C.

`U1K U1KOptions`

The sub key U1K determines which terms are included in the calculation of the U1 matrix (first order changes in MO coefficients). Options (Default none):

`Best`

The best (recommended) options for each relativistic option are included for this sub key.

`All`

Implies all the other options to this key.

`MV`

The mass-velocity term

`Dar`

The Darwin term

`ZMAN`

The Spin-Zeeman term.

`ESCL`

Scaled ZORA orbital energies in U1 matrix

Note: for chemical shifts, only compare results with the same options. If the sub key U1K is used with the option ALL in the ZORA calculation, then the scaled ZORA orbital energies are used in the making of the U1 matrix, which is not recommended. Recommended is to use 'U1K Best' in all cases, which uses plain ZORA orbital energies in the making of the U1 matrix.

`NUC NucOptions`

The (sub) key Nuc determines for which nuclei the chemical shifts are computed. If this (sub) key is omitted from the NMR block, the calculations are carried out for all nuclei. Else you may use this options by simply typing Nuc in the NMR block (without any further data); this means: for no nuclei at all. Alternatively you may type the index of the atom(s) you want to see analyzed. Default all nuclei are calculated, i.e. as for omitting this sub key.

Example:

```
NUC 2 1
```

The numbers refer to the internal numbering of the nuclei as it appears somewhere early in the general ADF output. This internal numbering is also the internal NMR numbering, but it is not necessarily the same as the input ordering. Use the subkey ATOMS to specify the nuclei according to this input ordering in the ADF calculation.

Note that the number of nuclei has a significant consequence for the total CPU time.

`Atoms AtomsOptions`

This subkey ATOMS specifies for which nuclei the NMR shielding is calculated. Default all nuclei are calculated, i.e. as for omitting this sub key.

Example:

```
ATOMS 2 1
```

The numbers refer to the input ordering in the ADF calculation. Use the subkey NUC to specify the nuclei according to the internal NMR numbers of the atoms.

`GHOSTS`

The subkey GHOSTS is a block type subkey. The format is:

```
Ghosts
   xx1 yy1 zz1
   xx2 yy2 zz2
   ......
SubEnd
```

With this key, the user can specify ANY point(s) within the molecule at which the shielding is to be calculated (whatever the physical meaning of this shielding is). One can think of those points as neutrons within the molecule. There is a publication by P. Schleyer et al. using a similar feature (J. Am.

Chem. Soc. **118**, 6317, 1996). They call it NICS, Nucleus-Independent Chemical Shift. Note that the NICS value is minus 1 times the isotropic part of the shielding tensor that is calculated at these points.

```
xx1 yy1 zz1
```

real numbers that specify the Cartesian coordinates of 'ghost' 1, etc.

The coordinates have to be specified in the same units as any other input (ADF subkey Units). That is, you use Angstrom for the ghosts if you did so for the atomic coordinates, or bohr otherwise. The same set of coordinates has to be specified as 'point charges with charge zero' using the key EFIELD. This is necessary in order to allow the appropriate distribution of integration points around the ghosts.

E.g., if you want to have two 'ghosts' with the coordinates xx1 yy1 zz1 and xx2 yy2 zz2 then you **must** also have in the input the key EFIELD as follows

```
EFIELD
 xx1 yy1 zz1   0.0
 xx2 yy2 zz2   0.0
END
```

(the last number is the charge at these coordinates - zero).

Eventually, this step should be programmed internally but for now the procedure outlined above works. No check is done to verify whether those 'point charges' are taken care of or not, but their omission leads to unpredictable results.

Only Cartesian coordinates are possible for ghosts, even if the atoms were originally specified using internal coordinates. This shouldn't be a problem, though (e.g., one could start an ADF run of the molecule of interest, and get very soon the Cartesian coordinates of the atoms in the output. This run would then be aborted, and restarted with the ghosts specified as desired.) The ghosts are numbered in the output as NNUC+1, NNUC+2 ... where NNUC is the total number of nuclei in this molecule. Default: no ghosts.

`Analysis AnalysisOptions`

The sub key Analysis controls the MO analysis. After the word (sub key) Analysis you type an integer, which then specifies that the first so many MOs are to be analyzed. Default no Analysis. The value of this analysis subkey in the block key NMR is somewhat limited. The separate ANALYSIS block key can give more analysis of the NMR chemical shielding.

Analysis block key

The NMR shielding tensor, can be analyzed in detail, see Refs. [329-331]. For the analysis option with the ANALYSIS block key there are some restrictions. In the ADF calculation all electron basis sets should have been used, and SYMMETRY NOSYM. Can not be used in case of non-relativistic calculations. The ADF calculation should use relativistic scalar ZORA or relativistic spinorbit ZORA. In case of scalar relativistic ZORA the keyword FAKESO should be added to the NMR input (outside of the NMR or Analysis block keys). The analysis utilizes the ZORA spin-orbit branch of the NMR code. For scalar ZORA, the NMR analysis contributions will appear in equivalent pairs for spin-orbitals even if the ADF calculation is closed-shell spin restricted. The MO numbering then also reflects this doubling of MOs. In the analysis, canonical MOs number 1 and 2 are the alpha and beta spin ADF MO 1, canonical MOs number 3 and 4 correspond to the alpha and beta spin ADF MO 2, and so on. In case of spinorbit relativistic ZORA the keyword FAKESO should not be included. For an NBO analysis of NMR, see the section on NBO analysis.

```
Analysis
   print threshold
   canonical
   {components}
```

```
End
{FakeSO}
```

`print threshold`

The print keyword selects printout of contributions relative to the total diamagnetic, paramagnetic. For example in case of 'print 0.01' only contributions greater than 1% are printed. Set to zero to print ALL contributions.

`canonical`

It enables an analysis of the shielding in terms of the canonical MOs.

`components`

The components keyword is optional and enables an analysis not only of the isotropic shielding but also of the diagonal cartesian components of the tensor XX, YY, and ZZ). In order to analyze the principal shielding tensor components with canonical MOs you can calculate the shielding tensor first with the NMR code, rotate the molecule such that the principal axes system aligns with the Cartesian coordinate system, and then repeat the NMR calculation with the analysis features switched on.

Paramagnetic NMR Chemical Shifts

Knowledge of the g-tensor and hyperfine A-tensor can be used in the prediction and analysis of paramagnetic NMR shifts, see Refs. [340-342]. Because of the dependence on the g-tensor, prediction of paramagnetic NMR (pNMR) shifts is not straightforward, as the dependence of the pNMR shift on excess α or β electron spin density at a nucleus is to be combined with the sign and magnitude of the isotropic g-value. Ref. [340] describes in detail how to calculate pNMR contact chemical shifts and pNMR pseudo-contact chemical shifts, using the ADF program.

Of course, like for NMR chemical shielding of closed shell molecules, one also needs to calculate of the so called orbital dependent part of the NMR chemical shielding of the open shell molecule. For open shell molecules in ADF this can only be calculated without taken into account spin-orbit coupling:

```
$ADFBIN/adf << eor
CHARGE charge spinpolarization
unrestricted
Relativistic scalar ZORA
Basis
 Core None
End
SAVE TAPE10
...
eor
$ADFBIN/nmr << eor
ALLINONE
nmr
 u1k best
 calc all
 out iso tens
end
end input
eor
```

Note that one can only do this at the scalar relativistic ZORA level, one needs to use all electron basis sets, and one needs to include the key ALLINONE in the input for NMR,

NMR spin-spin coupling constants

The NMR spin-spin coupling constants [118, 119] have been implemented in a separate program CPL. It can be combined with ZORA and Spin-Orbit treatment of relativistic effects to study heavy elements. The original version of this part of the User's Guide was written by Jochen Autschbach, primary author of the CPL code.

Introduction

The CPL code of the Amsterdam Density Functional program system allows the user to calculate *Nuclear Spin-spin Coupling Constants* (NSSCCs) [118,119]. NSSCCs are usually observed in NMR (Nuclear Magnetic Resonance) spectroscopy and give rise to the splitting of the signals of the NMR spectrum in multiplets. They contain a wealth of information about the geometric and electronic structure of the compound being investigated.

The calculation needs a standard TAPE21 ADF output file. CPL reads also an input key and optional settings from stdin (usually from an input file). Technical parameters such as the maximum memory usage can be set here as well.

One of the key features of the program is its ability to treat heavy nuclei with the ZORA relativistic formalism. We refer the reader to the literature for details about our implementation [118,119], and the general review on relativistic computations of NMR parameters [331]. Please use the information printed in the output header of the CPL program in order to provide references of this work in scientific publications.

The development of the CPL program started in 2000. CPL provides the main functionality in order to evaluate NSSCCs based on DFT, as well as a number of additional features in order to provide an analysis of the results. Several analysis features for the coupling constant have been added, see the CONTRIBUTIONS sub key. Please report bugs or suggestions to SCM at support@scm.com.

Theoretical and technical aspects

Within the non-relativistic theory of nuclear spin-spin coupling, there are four terms contributing to the NSSCC between two nuclei A and B: the paramagnetic and diamagnetic orbital terms (OP and OD, respectively), and the electron-spin dependent Fermi-contact (FC) and spin-dipole term (SD). In the literature, the OP and OD terms are often named PSO and DSO (for paramagnetic and diamagnetic spin-orbital). In the more general ZORA formulation, very similar operators are responsible for the NSSCC, therefore we use the same terminology for the individual contributions. In general, the interpretation of the results for a heavy atom system is basically equivalent to a non-relativistic situation.

In most cases, the FC term yields the most important contribution to the NSSCC. However, many exceptions are known for which one or each of the other terms can be non-negligible or even dominant. We therefore suggest that you always check, at least for a smaller but similar model system, or by using a smaller basis set, which of the four terms are negligible and which are dominant.

By default, the CPL program computes the FC coupling between the first and all other nuclei of the molecule, respectively. Other couplings or the computation of the OP, OD and SD terms can be requested by input switches (see the 'Running CPL' section of this document for details).

All contributions to the NSSCC are evaluated with the help of the numerical integration scheme implemented into ADF. In general, the computation of the OD term is computationally very cheap, since only integrals involving the electron density have to be evaluated. The next expensive term is the OP term. For this contribution, the first-order perturbed MOs have to be computed. With the available density functionals in ADF, the OP term does not cause a change in the Kohn-Sham potential, and the first-order MOs can be computed directly (i.e. without an iterative procedure). This is equivalent to the approach that has been implemented in the NMR code for ADF.

Both the FC and the SD terms induce electron spin-density to first-order as a perturbation. Equivalent to the iterative solution of the unperturbed Kohn-Sham equations, the first-order MOs depend on that first-order spin-density, which in turn depends on the first-order MOs. Therefore, in order to evaluate the FC and SD NSSCC contributions, the CPL program carries out a SCF cycle. In the scalar or non-relativistic case, the computational cost for the FC term is comparable to an ADF single point calculation with a local density functional. The evaluation of the SD term is more expensive. The current implementation utilizes the CPL spin-orbit code to compute the combined FC+SD contribution and therefore leaves some room for future speed-ups. In most cases, the SD term yields a negligible NSSCC and the much faster code for the scalar- or non-relativistic FC term can be used. However, it is very important to include the SD term in the computation if coupling anisotropies are to be evaluated.

In the case where the NSSCC computation is based on spin-orbit coupled relativistic two component ZORA MOs, the SD term causes only a marginal increase in computational time as compared to the FC term alone. Generally, in this case the computational cost for the FC term is already approximately one order of magnitude higher than in the scalar or non-relativistic case, since the 3 (x, y, z) components of the spin-density with respect to 3 components of the perturbation, respectively, have to be determined self-consistently. The additional presence of the SD term only shows up in a somewhat more costly evaluation of the matrix elements of the perturbation operator. However, CPL spends most of its computational time in the SCF cycle. Therefore, in spin-orbit computations the computation of the FC+SD terms is the default. The OP term has to be evaluated self-consistently, too, in this case and is added as a perturbation in the SCF cycle upon request.

We use the terminology 'perturbing' and 'responding nucleus' within the CPL output. The 'perturbing' nucleus is the one, for which the first-order MOs have to be computed (self-consistently), while the NSSCC is then determined by these first-order MOs and the FC, SD, and OP matrix elements of the second, 'responding' nucleus. For the OD term, this distinction makes no sense but is used in the output for reasons of consistency.

Experimental NSSCCs between two nuclei A and B are usually reported as J(A,B) in Hertz. From a computational point of view, the so-called reduced NSSCCs K(A,B) are more convenient for comparisons. CPL outputs both. The J's are set to zero in case the nuclear magneto-gyric ratio of one of the nuclei A or B is not available at run time.

Further technical aspects and current limitations

In order to facilitate the future computation of rather large molecules, all matrix elements of the perturbation operators FC, SD, and OP are evaluated in the Slater AO basis that is specified as input in the CREATE runs of ADF. The AO matrix elements are further transformed to the basis of MOs and the calculation proceeds within the MO basis. This allows for a convenient analysis of the results in terms of contributions from individual occupied and virtual MOs. Such an analysis can be requested by input.

The matrix elements themselves as well as the first-order contributions to the potential are evaluated by numerical integration. The CPL code, which is parallelized, can use multiple processors for these steps of the computation. The accuracy setting for the numerical integration is of high importance to obtain accurate matrix elements. Furthermore, the basis set being employed needs to be flexible enough to describe the perturbation correctly. This means usually that modified basis sets have to be used in particular for heavy element calculations.

The first-order potential is currently approximated by the VWN functional. The Xα potential is available as an alternative but usually leads to less accurate results. In ADF2009.01 the first order potential of the PBE family of GGA functionals and the hybrid PBE0 functional can be used.

Currently, only spin-restricted computations for systems with an even number of electrons are supported. Further, the calculation does not make use of symmetry and must be based on an ADF run with input SYMMETRY=NOSYM. Non-Aufbau configurations are not supported. The atom input list must not contain dummy atoms.

With the present version of CPL, the SD term and the FC/SD cross term cannot be evaluated separately. Either, the sum of FC + SD + cross terms, or the FC term individually, are computed.

CPL is restartable after various time-consuming steps of the computation.

In ADF2009.01 the hybrid PBE0 functional can be used in combination with NMR spin-spin coupling calculations, see the documentation for the extra keys that are needed. However, other hybrid functionals and Hartree-Fock can not or should not be used in combination with NMR spin-spin coupling calculations.

In ADF2009.01 the effects of a finite size of a nucleus on the spin-spin couplings can be calculated. A finite size of the nucleus can be set with the NUCLEARMODEL key in the input for the ADF calculation.

Bug fix in case more than 1 perturbing atom and DSO or PSO

In the ADF2006.01b version a bug in the CPL module is fixed that gave problems in ADF2006.01 and older versions. The problem in ADF2006.01 and older versions is: In case there is more than 1 perturbing atom and the DSO or PSO term is calculated, only the results of the spin-spin couplings for the first perturbing atom are correct, but the results of the other spin-spin couplings may be incorrect.

Input file for CPL: TAPE21

In order to run the CPL code, you need the general ADF output file TAPE21 being present in the directory where CPL is running. Most of the computation's specific settings will be taken from TAPE21, such as the integration accuracy, the basis set, the density functional being employed, nuclear coordinates, and so on. *That also means that nearly all of the aspects that affect the quality of CPL's results are already determined in the input for the ADF run*. Five aspects are of particular importance here:

1. The numerical integration accuracy: the perturbation operators are large in the vicinity of the nuclei. Therefore, you have to make sure that the integration grid is fine enough in the atomic core regions.

2. The basis set: NSSCCs are sensitive chemical probes, and therefore flexible basis sets have to be employed in order to yield a valid description of the MOs that determine the NSSCCs. We have found that it is imperative to use at least basis set TZ2P (V) from the ADF basis set database. Additional polarization functions in the valence shell may be necessary. Furthermore, the FC perturbation usually requires additional steep 1s functions (i.e. with exponents much higher than the nuclear charge) for a proper description. In the relativistic heavy element case, the use of additional steep basis functions as compared to the ZORA/TZ2P basis is mandatory. The use of steep functions is only of high importance for those nuclei, for which the NSSCC is to be evaluated.
In ADF2009.01 some basis sets suitable for NSSCCs has been added to the ADF basis set directory, in the directory $ADFHOME/atomicdata/ZORA/jcpl. For elements not available in this directory we suggest to use basis TZ2P as a starting point and to add some 1s basis functions (and appropriate fit functions) with higher exponents in order to improve the accuracy of the FC term. This is especially important for the heavy NMR nuclei.
For the nuclei for which NSSCCs are to be evaluated, it is necessary to use all-electron basis sets. This is not a restriction due to the implementation, but we have found that, with the available frozen core basis sets, the flexibility of the basis in the vicinity of the nuclei is not sufficient. It is possible to use frozen core basis sets if you add enough basis functions in the core region such that the basis approaches the flexibility of at least a double-zeta all-electron basis there [118]. In that sense, the savings in computational time due to usage of a frozen core basis are not as pronounced as in standard ADF computations. Unless reliable frozen-core basis sets for the NSSCC computation are available we strongly discourage the use of frozen core basis sets with the CPL program!

3. The finite size of a nucleus: typically, the isotropic J-couplings are reduced in magnitude by about 10 to 15 % for couplings between one of the heaviest NMR nuclei and a light atomic ligand, and even more so for couplings between two heavy atoms, see Ref. [270]. However, one should have really large basis sets with tight basis functions to observe this effect in calculations, see the previous point about basis

sets. The basis sets in the directory $ADFHOME/atomicdata/ZORA/jcpl are suitable for finite nucleus calculations. A finite size of the molecule can be set in the ADF program with the key NUCLEARMODEL:

```
NuclearModel Gaussian
```

4. The density functional: the results of the CPL code depend mostly on the shape of the MOs that have been determined by ADF, and their orbital energies. Both, in turn, depend on the density functional or Kohn-Sham potential that has been chosen for the ADF run (and the basis set quality). It is difficult to give a general advice here concerning the NSSCCs. So far we have found that the use of GGAs improves the NSSCCs with respect to experiment in most cases in comparison to LDA. Different GGAs often yield very similar results. Further, in particular for those cases for which the OP term is large or even dominant, both standard LDAs and GGAs sometimes do not provide an accurate enough description of the orbitals, and deviation of the CPL results as compared to experiment can be substantial. Future developments of density functionals might be able to cure these problems. For the time being, we recommend that you base the CPL run on different choices of density functionals in the ADF run, and investigate the convergence of the result with respect to basis set and integration accuracy. Note that CPL itself uses the VWN functional by default to determine the first-order perturbed MOs. There are enough indications to believe that this is a reasonable approximation for NMR purposes. In ADF2009.01 the first order potential of the PBE family of GGA functionals and the first order potential of the hybrid PBE0 functional can be used. See Refs. [425,426] for applications of such first order potentials. However, other hybrid functionals and Hartree-Fock can not or should not be used in combination with NMR spin-spin coupling calculations.

5. Modeling the experimental setup: computing such sensitive numbers as NMR chemical shifts and in particular NSSCCs can result in substantial deviations from experimental data. The simple reason might be that the isolated system that has been computed at zero temperature is not at all a good approximation to the system that has been studied experimentally. We [403,404] and other authors have found that in particular solvent effects can contribute very substantially to the NSSCC. In case you are comparing CPL results to experimental data obtained in strongly coordinating solvents we suggest that you consider solvent effects as a major influence. We have found that even weakly coordinating solvents can cause sizeable effects on the NSSCCs for coordinatively unsaturated metal complexes. Other sources of errors can be the neglect of vibrational corrections to the NSSCCs (usually in the range of a few percent).

If the parameters of the underlying ADF computation are carefully chosen and the density functional is able to provide an accurate description of the molecule under investigation, it is possible to compute NSSCCs by means of DFT with very satisfactory accuracy (please note that for properties as sensitive as NSSCCs, agreement with experimental results within about 10% error can be regarded as quite good). Further, chemical trends will be correctly reproduced for a related series of molecules in most cases. However, due to the inherent approximate character of the density functionals currently available with ADF, and necessary basis set limitations, great care should be taken that the results are reliable. *CPL assumes Aufbau configurations. Please make sure that there are no empty orbitals with energies below the highest occupied MO (HOMO). In addition, the SYMMETRY NOSYM key has to be used in the ADF computation. It is currently not possible to use dummy atoms in the ADF input if the* TAPE21 *is intended to be used for a subsequent CPL computation.*

Running CPL

Main input switches

With the ADF output TAPE21 present in the current working directory, the CPL code is invoked by:

```
$ADFBIN/cpl < input_file
```

where input_file contains the input for CPL. We have tried to ensure some backward compatibility with older ADF versions, such as ADF 1999 and ADF 2.3. Normally, you will use the ADF suite that contains the CPL code of the same version. CPL tries to detect if the TAPE21 belongs to an older version of ADF and exits with an error message in case it is not able to process this file. For ADF 2.3, you have to supply also the TAPE10 of ADF 2.3 in addition to TAPE21 (specify SAVEFILE TAPE10 in the ADF input file). We provide this option for testing purposes, however this functionality is not supported and we do not recommend to run CPL on top of the output of an older ADF version.

input_file must contain at least one block-type input key in order to start the CPL run. The input key is

```
$ADFBIN/cpl << eor
NMRCOUPLING
END
eor
```

This represents a minimal input file for CPL. The NMRCOUPLING key hosts all optional keys that are relevant for the NSSCCs themselves. In addition to the mandatory NMRCOUPLING key, CPL recognizes the following input switches:

```
$ADFBIN/cpl << eor
GGA
..
eor
```

See the separate section for this key, which influences the first order potential that is used.

```
$ADFBIN/cpl << eor
RESTART restart_file
..
eor
```

restart the computation from file restart_file. This is the TAPE13 produced during a CPL run. By default, TAPE13 is deleted after a successful completion of CPL. As with ADF restarts, you can not use the name TAPE13 for restart_file but you have to rename it, e.g. to tape13.restart.)

```
$ADFBIN/cpl << eor
SAVEFILE TAPE13
..
eor
```

keep the restart file even after a successful completion of CPL. TAPE13 is currently the only file that is meaningful as a parameter to SAVEFILE

NMRCOUPLING subkeys

The available switches within a NMRCOUPLING/END block control the computation of the NSSCCs. By default, the program will evaluate the FC coupling contribution for the first nucleus being the perturbing nucleus and all remaining nuclei responding.

Please **note** that the ordering of atoms in CPL is generally different from the ADF input. The ordering of atoms is the one being stored in TAPE21 and it is grouped by fragment types. In case you are in doubt about the ordering of atoms, you can run CPL for a few seconds. It will print a list of atoms with their coordinates. The ordering is currently the same as required the NMR program in the ADF program system. On the other hand, note that for the subkeys ATOMPERT and ATOMRESP the number of the atoms refer to the input ordering in the ADF calculation.

Available subkeys are:

```
$ADFBIN/cpl << eor
 NMRCOUPLING
   NUCLEI {npert nresp1 nresp2}
   ATOMPERT {npert1 npert2 npert3}
   ATOMRESP {nresp1 nresp2 nresp3}
   GAMMA {nnuc gamma}
   DSO
   PSO
   SD
   FC
   SCF {ITERATIONS=25 | NOCYCLE | CONVERGE=1e-4 }
   XALPHA
   CONTRIBUTIONS {1E19} {LMO, SFO, LMO2, SFO2}
END
..
eor
```

`NUCLEI {npert nresp1 nresp2}`

Use nucleus no. npert as the perturbing nucleus, and nuclei nresp1, nresp2, etc as responding nuclei. You can supply more than one NUCLEI keys, in which case CPL evaluates the first-order MOs for each perturbing nucleus that is specified and computes the NSSCCs between all specified responding nuclei. For each NUCLEI line in the input, CPL has to perform an SCF cycle. Note: for the numbers of the atoms the internal CPL numbering should be used.

`ATOMPERT {npert1 npert2 npert3}`
`ATOMRESP {nresp1 nresp2 nresp3}`

ATOMPERT: use nucleus no. npert1, npert2, etc. as the perturbing nuclei. ATOMRESP: use nucleus no. nresp1, nresp2, etc. as the responding nuclei. You can supply more than one ATOMPERT and (or) ATOMRESP key. CPL computes the NSSCCs for all pairs of combinations of perturbing atoms and responding atoms. For each perturbing atom CPL has to perform an SCF cycle, which is the expensive part in the calculation. Note: the numbers refer to the input ordering in the ADF calculation. Use the subkey NUCLEI to specify the nuclei according to the internal CPL numbers of the atoms.

`GAMMA {nnuc gamma}`

Input a non-default magneto-gyric ratio of g = gamma for nucleus no. nnuc, in units of rad/(T s). Note that one should include the the typical 10^7 factor. CPL normally uses the g value of the most abundant NMR active isotope for a nucleus of a given charge by default. With the GAMMA keyword you can override this value or supply a value if CPL does not know about it. A list of g's that is used in the computation is printed in the output. You have to provide the GAMMA key for each nucleus you want to specify.

`DSO`

Compute the diamagnetic orbital term for each NSSCC that is requested (not default)

`PSO`

Compute the paramagnetic orbital term for each NSSCC (not default)

`SD`

Compute the SD term for each NSSCC. This is only default for spin-orbit ADF runs. The output will contain the sum of the FC and SD contributions. *Please note that requesting this option results in a greatly increased computational cost in scalar or non-relativistic runs.* The option NOSD will turn the SD computation off in spin-orbit runs and has no effect otherwise.

`FC`

Compute the FC contribution to the NSSCCs. This is the default option. Please note that it is currently not possible to compute the SD term without the FC term. Consult the 'practical aspects' section for instructions how to estimate the FC/SD cross term. The option NOFC will disable both the FC and SD computation.

`SCF {ITERATIONS=25 | NOCYCLE | CONVERGE=1e-4 }`

Settings related to the SCF cycle that is carried out by CPL. Valid options are (with default values if applicable):

`ITERATIONS 25`

maximum number of iterations

`NOCYCLE`

perform no cycle, equivalent to ITERATIONS 0

`CONVERGE 1e-4`

convergence criterion, an input of e corresponds approximately to a convergence of log(-e) digits, i.e. the results will be converged to about four significant digits by default. The measurement for the convergence is based on the sum *S* of the magnitudes of all occupied-virtual matrix elements of the induced first-order exchange potential. Note that the actual convergence criterion being used in the computation is e times *S* of the first cycle, i.e. the convergence criterion is set relative to the initial value of *S*.

`XALPHA`

Use first-order Xalpha potential instead of VWN potential (default). This will usually decrease the accuracy for couplings involving hydrogen, and does not have a large effect for couplings between heavier nuclei (not default). The key is mainly intended to ensure compatibility with our previously published results.

`CONTRIBUTIONS {1e19} {LMO, SFO, LMO2, SFO2}`

Print contributions from individual orbitals to the FC and OP term of the NSSCCs that are larger in magnitude than a certain threshold. The threshold refers to the reduced coupling constant K in SI units (not default). Additionally, an analysis in terms of Boys localized MOs (see User's Guide and SFOs. At present, either each key LMO, SFO, LMO2, SFO2 can be used individually, or grouped as {LMO, SFO2} or {SFO2, LMO}. If you need all analyses or different combinations, it is recommended to restart the CPL calculation from TAPE13, and to specify 0 iterations in the SCF. This way, the only additional computational cost should be the analysis itself.

The equation and an application for the analyses due to the LMO and SFO keys is described in Refs. [368,369]. The other analysis is based on the same equation as in Ref. [370]. For an NBO analysis of the spin-spin couplings, see the section on NBO analysis.

In order for the LMO-based analyses to work, the MO → LMO transformation matrix needs to be stored on TAPE21. In the ADF input, you can achieve this with the option "STORE" to the LOCORB key, i.e.

```
LOCORB STORE
 ... options
END
```

GGA key

```
$ADFBIN/cpl << eor
GGA
..
eor
```

GGA

Use first-order GGA potential instead of the first-order VWN potential. Should only be used for the PBE family of GGA exchange-correlation functionals and for the hybrid functional PBE0. See Refs. [425,426] for applications of calculating spin-spin couplings with PBE0. However, other hybrid functionals and Hartree-Fock can not or should not be used in combination with this key GGA. For consistency reasons of the first-order potential one should use the keyword USESPCODE in the ADF calculation. An example input for ADF for the hybrid PBE0 would then contain:

```
USESPCODE
XC
 hybrid PBE0
End
```

Practical Aspects

Minimal input

The default settings for CPL are invoked by the simple minimal content of the input file:

```
$ADFBIN/cpl << eor
NMRCOUPLING
END
eor
```

This is equivalent to

```
$ADFBIN/cpl << eor
NMRCOUPLING
 NUCLEI 1 2 3 4 5 6 7 8 ..(up to number of atoms)
 SCF CONVERGE 1e-4 ITERATIONS 25
 FC
END
eor
```

Restarts

CPL is restartable after the computation of each the complete set of FC or FC/SD and OP matrix elements, and after their transformation to the MO basis. Further, in spin-orbit runs or in scalar- or non-relativistic computations involving the SD term, CPL is restartable after each SCF cycle. As with ADF restarts, you need to supply a proper input file for a restarted computation, and the restart file TAPE13 (which needs to be renamed). Changing the input of a calculation for a restart is not supported. In restarted runs, the program will automatically continue at the latest possible point before the execution stopped, and changing the input between restarts can cause inconsistencies that may lead to a crash.

Unless you are computing a very large molecule, the most likely need for a restart will probably occur during a computation of the FC/SD SCF cycle. We have already mentioned that this is a very time consuming part of the computation, and for this reason CPL can be restarted after each completed SCF cycle. The convergence of the results should not be affected by a restart. You can, e.g., use this in order to complete a lengthy CPL computation in case you have tight time limits in your queuing system, or after a power loss.

How to avoid the unnecessary computation of many SCF cycles

As already mentioned, once the first-order MOs with respect to the perturbation by one of the nuclear spins have been determined, the NSSCC between this and all other nuclei can be computed rather quickly. For each nucleus that participates in at least one of the coupling constants to be determined, the matrix elements of the FC, SD, and OP operators have to be evaluated once (unless the computation of the respective terms is disabled).

You can use this information in order to minimize the number of nuclei for which an SCF cycle has to be performed. This can lead to a great speedup of the computation. The final result, the NSSCC between A and B, does not depend on which nucleus has been chosen as the 'perturbing' one, and which as the 'responding' one (convergence has to be good enough, though). Suppose you want to compute the NSSCCs in the water molecule, with O being nucleus no. 1. In that case,

```
NUCLEI 1 2 3
NUCLEI 2 3
```

yields the same O-H and H-H coupling constants as the input

```
NUCLEI 2 1
NUCLEI 1 3
NUCLEI 3 2
```

but with less computational effort due to the fact that only 2 instead of 3 SCF cycles will be performed. The example chosen here is trivial, but in other cases it can be worthwhile to consider different sequences of computations.

Alternatively you can use the ATOMPERT and ATOMRESP subkeys:

```
ATOMPERT 1 2
ATOMRESP 2 3
```

which will calculate the spin-spin coupling of the nuclei 1-2, 1-3, and 2-3 (skips 2-2, since the nuclei are the same), which are the same O-H and H-H couplings as before.

Note: the numbers of the nuclei for the subkeys ATOMPERT and ATOMRESP refer to the input ordering in the ADF calculation, whereas the numbers of the nuclei for the subkey NUCLEI refer to the internal CPL numbers of the atoms.

Computing individual terms in the coupling tensor

As we have mentioned before, the FC, OP and OD terms can be calculated individually, but not the SD term. In case the SD input option is given, the FC+SD contribution is evaluated instead. This is NOT equal to the sum of the individual FC and SD contributions since there is a cross term between these two. Due to computational simplicity and efficiency, CPL evaluates either the matrix elements for the FC operator, or the combined ones for FC+SD. The final result therefore contains either FC only, or FC, SD plus the cross terms. Only the latter, in addition to the OP and OD contributions, should be compared to experimental results. We will implement the computation of the individual SD term in a future version of CPL in order to assist the analysis of the CPL results.

Likewise, in a spin-orbit based relativistic computation, there exists a cross term between the spin-dependent FC and SD terms, and the OP term. In the scalar- or non-relativistic limit, this contribution is always zero. With the PSO option present, CPL computes the FC, SD and OP terms including all cross contributions. Even though the output suggests that the individual OP and FC+SD terms are printed, they contain additional cross terms if spin-orbit coupling is large. You can run CPL with the options

```
$ADFBIN/cpl << eor
NMRCOUPLING
 NOFC
 NOSD
```

```
 PSO
END
..
eor
```

in order to evaluate the individual OP contribution(s). In a second run, you can then compute just the FC+SD contributions. The differences between these two CPL runs and a third one with all three terms present yields the relativistic (FC+SD)/OP cross term.

Two-bond and more-bond couplings

CPL does not discriminate between one-bond and two-bond couplings etc. in any technical sense. Even though we [118,119,403,404] have validated the code mostly for one-bond NSSCCs, the coupling between any pair of nuclei in the molecule can be computed. See Ref. [404] for an example.

Principal axis system, the whole coupling tensor

CPL evaluates the complete 3x3 coupling tensor with respect to the Cartesian input coordinate system. Depending on the orientation of the molecule, and the local symmetry, the coupling tensor has in fact often only a small number of independent components. CPL evaluates the 'principal components' by the following procedure: the 3x3 matrix is transformed into the basis of the eigenvectors of its symmetric part. This diagonalizes the symmetric part of the coupling tensor. A set of eigenvectors (= 'principal axis system') is also printed.

References

[118] J. Autschbach, T. Ziegler, J. Chem. Phys. 2000, **113**, 936.
[119] J. Autschbach, T. Ziegler, J. Chem. Phys. 2000, **113**, 9410.
[403] J. Autschbach, T. Ziegler, J. Am Chem. Soc. 2001, **123**, 3341.
[404] J. Autschbach, T. Ziegler, J. Am Chem. Soc. 2001, **123**, 5320.
[270] J. Autschbach, ChemPhysChem, 2009, **10**, 2274.
[425] J. Autschbach, J. Chem. Phys. 2008, **129**, 094105, J. Chem. Phys. 2009, **130**, 209901.
[426] D.L. Bryce and J. Autschbach, Can. J. Chem. 2009, **87**, 927.
[331] J. Autschbach and S. Zheng, Ann. Rep. NMR Spectr. 2009, **67**, 1.

See also:
[371] N. F. Ramsey, Phys. Rev. **91**, 303 (1953).
[372] Dickson, R.M.; Ziegler, T. J. Phys. Chem. 1996, **100**, 5286.
[370] Khandogin, J.; Ziegler, T. Spectr Acta A 1999, **55**, 607.
[373] D. L. Bryce, R. Wasylishen, J. Am. Chem. Soc. **122**, 3197 (2000).
this ADF User's manual, SCM, Vrije Universiteit, Amsterdam, The Netherlands.
[374] G. Schreckenbach, S. K. Wolff, T. Ziegler, Modeling NMR chemical shifts, ACS Symposium Series, Washington DC (1999).

ESR/EPR

The EPR (ESR) g-tensor, hyperfine interaction (A-tensor), nuclear quadrupole interaction (Q-tensor), and zero-field splitting (ZFS, D-tensor) can be calculated. Effects due to spin-orbit coupling can be included. All electron basis sets can be used.

The separate program EPR/NMR ($ADFBIN/epr) program is no longer documented, since most of its capabilities are implemented in newer modules. See for the old documentation the ADF2010 EPR/NMR module documentation, and Refs. [120,121,416,417].

ESR/EPR g-tensor and A-tensor

A-tensor, no spin-orbit coupling

```
$ADFBIN/adf << eor
ESR
END
CHARGE charge spinpolarization
unrestricted
{Relativistic scalar ZORA}
{NUCLEARMODEL gaussian}
eor
```

If spin-orbit coupling is neglected, the spin in the effective spin Hamiltonian, which is commonly used for the interpretation of ESR experiments, is the real electronic spin of the paramagnetic molecule. In the spin-unrestricted DFT calculations one then uses eigenfunctions of S_z. The A-tensor can then simply be calculated as expectation value of the corresponding operator, see [96].

The A-tensor will be calculated for all nuclei. Terms due to the spin-polarization density at the nucleus are included in the evaluation of the A-tensor. For an accurate evaluation of the spin-polarization density at the nucleus it is important to use an all-electron basis set for the nuclei that one is interested in, avoiding the frozen core approximation. For heavy elements the incorporation of a Gaussian finite nucleus model can be important. However, one should have really large basis sets with tight basis functions to observe this effect in calculations. One possibility is to use the $ADFHOME/atomicdata/ZORA/QZ4P basis set, although even this large basis set is not large enough sometimes. The basis sets in the directory $ADFHOME/atomicdata/ZORA/jcpl (not available for all elements) are suitable for finite nucleus calculations.

In case one uses a finite nuclear model for the charge distribution, starting from ADF2013 ADF also uses a finite distribution of the nuclear magnetic dipole moment for the calculation of the A-tensor.

A-tensor, perturbative inclusion spin-orbit coupling

```
$ADFBIN/adf << eor
CHARGE charge spinpolarization
unrestricted
Relativistic scalar ZORA
Symmetry NOSYM
...
eor
$ADFBIN/cpl << eor
hyperfine
 atoms 1 2 :: calculates A-tensor for atom 1 and 2, input order
 SCF Converge=1e-7 {Iterations=25}
end
...
eor
```

The calculation of A-tensors is implemented in the CPL program as a second derivative property (spin-orbit coupling and nuclear magnetic field as perturbation) within the two-component relativistic zeroth-order regular approximation (ZORA), see Ref. [340]. This implementation allows for hybrid (only PBE0) DFT calculations, but not metaGGA's and not metahybrids.

Note that the CPKS convergence in CPL has to be set tightly (1e-7 or 1e-8) to get converged PSOSO terms for the A-tensor. For hyperfine calculations the default value is 1e-7.

g-tensor, perturbative inclusion spin-orbit coupling

```
$ADFBIN/adf << eor
CHARGE charge spinpolarization
unrestricted
Relativistic scalar ZORA
Symmetry NOSYM
...
eor
$ADFBIN/nmr << eor
nmr
 gfactors
 u1k best
 calc all
 out iso tens
end
end input
eor
```

The calculation of g-tensors is implemented in the NMR program as a second derivative property (spin-orbit coupling and external magnetic field as perturbation) within the two-component relativistic zeroth-order regular approximation (ZORA), see Ref. [339]. This implementation allows for hybrid (B3LYP, PBE0, etc) DFT calculations, but not metaGGA's and not metahybrids. This implementation requires the use of all electron basis sets.

For an older implementation of this method, see the ADF2010 EPR/NMR module documentation, and Refs. [120,121,416,417].

g-tensor and A-tensor, self consistent spin-orbit coupling

```
$ADFBIN/adf << eor
ESR
END
CHARGE charge
unrestricted
Relativistic spinorbit ZORA
Collinear
Symmetry NOSYM
eor
```

In a spin-orbit coupled spin unrestricted relativistic ZORA calculation and the ESR block key, the g-tensor and the nuclear magnetic dipole hyperfine interaction (A-tensor) will be calculated, see also Refs. [95,96]. In such a calculation degenerate perturbation theory is used with the external magnetic field or nuclear magnetic field as perturbation. The calculation must use the collinear approximation, and symmetry must be NOSYM. This implementation does allow for metaGGA, and (meta-)hybrid DFT calculations, but then GIAO's are not used. There may be more than one unpaired electron. Terms due to the spin-polarization density at the nucleus are included in the evaluation of the A-tensor. However, one can not set the number of unpaired electrons, the 'spinpolarization' argument of the key CHARGE will be ignored.

Note: in a spin-orbit coupled spin restricted relativistic ZORA calculation and the ESR block key, ADF will also calculate and print the nuclear magnetic dipole hyperfine interaction, but the terms due to the spin-polarization density at the nucleus are absent. Furthermore, if there is more than one unpaired electron, the computed results will simply be incorrect, without any warning from the program. On the other hand, in case of one unpaired electron, and very large effects of spin-orbit coupling, the spin-restricted calculation may be of interest, since it uses Kramer's symmetry exact.

ESR/EPR Q-tensor

For the calculation of the ESR Q-tensor see the key QTENS.

ESR/EPR Zero-field splitting (D-tensor)

With the keyword ZFS the zero-field splitting (ZFS) of the ground state can be calculated.

```
ZFS
RELATIVISTIC SCALAR ZORA
```

Zero-field splitting is the breaking of degeneracies of the ground state that is not described by a standard nonrelativistic Hamiltonian. ZFS as calculated by ADF is that exhibitted by molecules whose ground state has spin S>1/2 and no spatial degeneracy. This type of ZFS has two contributions, second-order spin-orbit coupling and spin-spin coupling. In the present implementation only the spin-orbit coupling term is included. The calculation of ZFS with DFT is described in [312-315]

ZFS can be calculated in combination with LDA and GGAs but not hybrid or meta-GGA functionals. In order to calculate ZFS the RELATIVISTIC SCALAR ZORA option must be included.

Just the simple keyword ZFS is needed in order to calculate zero-field splitting. Several optional additional keywords can also be included. The complete list is:

```
ZFS {PEDERSON|NEESE} {ANALYSIS|FULLANALYSIS}
```

`PEDERSON|NEESE`

PEDERSON: The available approaches for calculating ZFS with DFT each differ subtly from the others. We believe that the method proposed by van Wüllen and coworkers [314,315] is the most theoretically complete but it may be that for certain systems the other approaches are more accurate. The van Wüllen formulation is the default but if the PEDERSON keyword in included then the equation proposed by Pederson and Khanna [312] is used.
NEESE: If the NEESE keyword is included then the equation for ZFS proposed by Neese [313] is used.

`ANALYSIS|FULLANALYSIS`

ANALYSIS: Neese has presented some interesting analyses of ZFS [313]. If the ANALYSIS keyword is invoked then the contributions to the ZFS is divided into terms from alpha-beta, alpha-alpha, beta-beta and beta-alpha one-electron excitations.
FULLANALYSIS: The output requested by the ANALYSIS keyword is further extended to analyze each of the alpha-beta, alpha-alpha, beta-beta and beta-alpha contributions in terms of the individual one-electron excitations.

Nuclear Quadrupole Interaction (EFG)

```
QTENS
```

This key activates the computation of the Nuclear Electric Quadrupole Hyperfine interaction. It can be applied to open-shell and to closed-shell systems. QTENS gives you the Nuclear Electric Quadrupole Hyperfine interaction (Q-tensor) [97]. The latter is directly related to the Electric Field Gradient (EFG) . The Q-tensor elements (in MHz) equal the the electric field gradient tensor elements (in a.u.) times 234.9647 times the nuclear quadrupole moment (NQM in barn units, 1 barn = 10-28m^2 = 10-24cm^2) and divided by 2I(2I-1), where I is the nuclear spin. The Nuclear Quadrupole Coupling Constant (NQCC) (in MHz) is the

largest value of the principal values of the EFG (in a.u.) times 234.9647 times the nuclear quadrupole moment (in barn units). The electric field gradient tensor is printed next to the Q-tensor.

In the case of ZORA the program will also calculate the EFG in the so called ZORA-4 approximation, which includes a small component density ("picture-change correction"), see [97]. If one includes spin-orbit coupling the EFG in the ZORA-4 approximation is only calculated if the symmetry in the calculation is NOSYM.

In case QTENS is used for ^{57}Fe, ^{119}Sn, ^{125}Te, ^{193}Ir, and ^{197}Au, quadrupole splittings are written in units of mm/s, used in Mössbauer spectroscopy.

Analysis of the EFG

With the EFG keyword in AOResponse a Mulliken type analysis of the EFG principal components, and an analysis in terms of canonical MOs, can be performed. Required is symmetry NOSYM. This not implemented in case of spin-orbit coupling. For an NBO analysis of the EFG, see the section on NBO analysis. For an explanation of the output and a general usage tutorial, see [327]. Further references and recommended citations, see [328].

```
Symmetry NOSYM
Aoresponse
 efg NUC
end
```

`efg NUC`

Here NUC is the number of the nucleus at which the EFG is to be computed (ADF internal atom ordering). Avaliable for one nucleus at the time.

Mössbauer spectroscopy

Isomer shifts

By default the electron density at the nuclei is calculated, no input key is required. In the implementation in ADF, the electron density is not calculated exactly at the center of the nucleus, however, at points on a small sphere around the center of a nucleus. The printed electron density in the output of ADF is the average electron density on these points. The radius of the sphere is an approximated finite nuclear radius. The electron density at the nuclei could be used for the interpretation of isomer shifts in Mössbauer spectroscopy. Typically one needs to perform a fit of the experimentally measured isomer shifts and calculated electron densities, like, for example, is done in Ref. [281].

One should use all electron basis sets for the Mössbauer active elements. Important is to use the same basis set, same exchange correlation functional, same integration accuracy, and same nuclear model (see key NUCLEARMODEL), if electron densities at nuclei in different molecules are compared. Note that the absolute electron density at a nucleus heavily depends on the accuracy of the basis set in the core region of this nucleus, especially if relativistic effects are included.

Quadrupole splittings

For the calculation of Mössbauer quadrupole splittings see the key QTENS.

Nuclear resonance vibrational spectroscopy (NRVS)

The nuclear resonance vibrational spectroscopy (NRVS) experiment can be thought of as Mösbauer spectroscopy with vibrational sidebands. The NRVS experiment provides the complete set of bands corresponding to modes that involve motion of the Mössbauer active atoms. In order to calculate this with the ADF program a partial vibrational Density-Of-States (PVDOS) has been implemented. A PVDOS factor for a given atom is the ratio of this atom nucleus kinetic (vibrational) energy to the total vibrational energy of all nuclei, for a given mode. PVDOS factors for every atom and every mode are written to TAPE21 if IR Frequencies are calculated. To visualize the calculated PVDOS use the ADFspectrum program: select the PVDOS spectrum type. Next select one or more atoms to get the PVDOS spectrum generated by the selected atoms. This is useful for analysis of NRVS spectra in bioinorganic chemistry for NRVS-active nuclei.

2.7 Transport properties

See also
Examples: transport properties

Charge transfer integrals (transport properties)

ADF can provide input parameters, such as charge transfer integrals, that are needed in approximate methods that model charge transport properties. ADF has the unique feature that it can (also) calculate such transfer integrals based on the direct method by the use of its unique fragment approach.

In theoretical models of charge transport in organic materials, see Refs. [293-295], the whole system is divided into fragments, in which an electron or hole is localized on a fragment, and can hop from one fragment to another. In the tight-binding approximation that is used in these models the electron or hole is approximated with a sinlge orbital, and it is assumed that only the nearest neighboring fragments can couple. The models require accurate values of electronic couplings for charge transfer (also referred to as charge transfer integrals or hopping matrix elements) and site energies (energy of a charge when it is localized at a particular molecule) as a function of the geometric conformation of adjacent molecules. Charge transfer integrals for hole transport can be calculated from the energetic splitting of the two highest-occupied molecular orbitals (HOMO and HOMO-1) in a system consisting of two adjacent molecules, also called "energy splitting in dimer" (ESID) method. For electron transport these can be calculated from the two lowest-unoccupied orbitals (LUMO and LUMO+1) in this ESID method. ADF can also calculate transfer integrals based on the direct method by the use of its unique fragment approach. see Refs. [294,295]. ADF allows one to use molecular orbitals on individual molecules as a basis set in calculations on a system composed of two or more molecules. The charge transfer integrals obtained in this way differ significantly from values estimated from the energy splitting between the highest occupied molecular orbitals in a dimer. The difference is due to the nonzero spatial overlap between the molecular orbitals on adjacent molecules. Also, ADF's methods are applicable in cases where an orbital on one molecule couples with two or more orbitals on another molecule.

Charge transfer integrals with the TRANSFERINTEGRALS key

In this method the matrix elements of the molecular Kohn-Sham Hamiltonian in the basis of fragment orbitals is used to calculate site energies and charge transfer integrals. Likewise the overlap integrals between fragment orbitals are calculated. No explicit electrons are removed or added in this method. For electron mobility calculations the fragment LUMO's are considered. For hole mobility calculations the fragment HOMO's are considered.

To calculate the charge transfer integrals, spatial overlap integrals and site energies, include the key TRANSFERINTEGRALS in the input for ADF. Symmetry NOSYM should be used. The molecular system typically should be build from 2 fragments. In the fragment calculation full symmetry can be used.

```
TRANSFERINTEGRALS
Symmetry NOSYM
Fragments
 frag1 frag1.t21
 frag2 frag2.t21
End
```

By default, integrals are calculated only for the HOMO (LUMO) of the fragments, and possibly HOMO-1, HOMO-2 (LUMO+1, LUMO+2) if the energy of those fragment orbitals are close to the HOMO (LUMO) of that fragment. To calculate the matrix elements and overlap integrals based on all fragment orbitals one can use the key:

```
PRINT FMATSFO
```

The method described here to calculate charge transfer integrals is more approximate than the next method that uses FDE. The major difference is how effects of a localized charge are included.

Charge transfer integrals with FDE

Overview

The ELECTRONTRANSFER keyblock invokes the calculation of Hamiltonian (site energies and couplings) and overlap matrix elements with FDE-derived localized states. Two FDE calculations are (not strictly) needed before running the ELECTRONTRANSFER calculation. The calculated matrix elements are theoretically similar to the ones obtained with the TRANSFERINTEGRALS keyword.

Relation with the TRANSFERINTEGRALS key

The key difference between TRANSFERINTEGRALS and ELECTRONTRANSFER is that the latter allows to include

- Effects of orbitals relaxation due to localized charges, Refs.[352,353]
- Effects of polarization due to molecules in the environment, Ref.[353]
- ELECTRONTRANSFER is linear scaling in the number of fragments when the system is composed by more than one fragment

Limitations

- Hybrid functionals are not yet supported
- The code cannot be used in conjunction with the BASIS key
- The code is not parallelized yet and must run on a SINGLE processor
- The code can only tackle hole and excess electron transfer, i.e. charge separation or triplet energy transfer are not yet implemented

The first three limitations do not apply to the method with the TRANSFERINTEGRALS key.

ELECTRONTRANSFER input

The minimum input for the ELECTRONTRANSFER key is:

```
ELECTRONTRANSFER
FRAGMENTS
 frag1 FragFile1
 ...
 fragN FragFileN
END
ELECTRONTRANSFER
```

```
 NumFrag N
END
```

where frag1 ... fragN are the labels of the fragments in the calculation, and FragFile1 ... FragFileN are the TAPE21 files of spin RESTRICTED calculations of the isolated fragments, N is the total number of fragments employed in the calculation.

Files and file names

The fragment files to be used in the ELECTRONTRANSFER calculation are generally different from the TAPE21 files used in the FRAGMENTS key block. Two types of fragment TAPE21 files are needed by the calculation:

1. The isolated closed-shell TAPE files for the FRAGMENTS keyblock
2. The TAPE21 files of the charge or spin localized states (which can be obtained with an FDE calculation as done in the example below)

There are 2 charge localized states. They are labelled with A and B. The respective TAPE21 files must be names as follows:

```
fragA1.t21, fragA2.t21, ... , fragAN.t21 (for state A)
fragB1.t21, fragB2.t21, ... , fragBN.t21 (for state B)
```

The above files should be copied to the working folder of the ADF calculation prior to executing ADF.

Options

```
ELECTRONTRANSFER
 NumFrag N
 {Joint|Disjoint}
 {Debug}
 {Print EIGS|SAB}
 {FDE}
 {INVTHR threshold}
END
```

`Joint|Disjoint`

The default is "Joint". Joint is always recommended. The "Disjoint" formalism is described in Ref.[353] and is much faster than the "Joint" formalism when more than 2 fragments are considered. Joint and Disjoint are equivalent for systems composed of only 2 fragments. Disjoint should only be used if the fragment files are obtained in an FDE calculation (see FDE below).

`Debug`

The code performs additional checks (determinants, diagonalizations, inversions, traces, etc.). Substantial increase in the output should be expected.

`Print`

If EIGS, it will print the (unformatted) matrix of the MO coefficient in the AO representation. If SAB, it will print the (unformatted) matrix of the diagonal and transition overlap matrix in the MO representation.

`FDE`

An FDE calculation including more than 2 fragments must include the following key block

```
ELECTRONTRANSFER
  FDE
END
```

and the numericalintegration precision in the *last* FDE calculation for every subsystem should be set to no less than:

```
BeckeGrid
   Quality Good
End
```

if in the subsequent ELECTRONTRANSFER calculation the DISJOINT subkey is used.

`Invthr threshold`

Default 1.0e-3, is a threshold for the Penrose inversion of the transition overlap matrix. If warnings about density fitting are printed, invthr may be increased up to 1.0e-2. Larger invthr might affect the quality of the calculated couplings and excitation energies.

Output

The output of the example in $ADFHOME/examples/adf/ElectronTransfer_FDE_H2O is discussed here. This example involves the calculation of electronic coupling, site energies and charge-transfer excitation energy for the hole transfer in a water dimer.

```
  ============  Electron Transfer RESULTS ===================

  Electronic Coupling =         0.000000 eV
  Electronic Coupling =        -0.003569 cm-1
  H11-H22             =        -1.396546 eV
  Excitation Energy   =         1.396546 eV
  Overlap             =         0.000000
  H11 H22 H12 =  -152.443000816341  -152.391678701092  -151.743979368040
Eh
  S11 S22 S12 =     0.981795415192     0.981006454450    -0.000000023700

  =========== END Electron Transfer RESULTS ================
```

Due to symmetry, the overlap is almost diagonal (Overlap = 0.00), thus the transition density is evaluated with one less electron as explained in Ref. [353].

The electronic coupling between the state with a positive charge localized on one water molecule and another with the charge localized on the other water molecule is given by "Electronic Coupling" and is reported in eV and cm^-1.

"H11-H22" is the difference of the site energies in eV. Values of the site energies are given by the first two values of "H11 H22 H12" in atomic units.

"Excitation Energy" reports the value of the transfer excitation energy as calculated by diagonalization of the 2X2 generalized eigenvalue problem in the basis of the charge-localized states, see Refs. [352,353].

"S11 S22 S12" are the values of the non-normalized overlaps.

GREEN: Non-self-consistent Green's function calculation

green is an auxiliary program which can be used to calculate the density of states (DOS) and transmission of molecules connected to semi-infinite contacts. The transmission is the electron transmission through a

molecule connected to semi-infinite contacts. The calculation is based on the non-self-consistent Green's function method. The details of this method can be found in chapter 2 and appendix C of the PhD thesis of Jos Seldenthuis (2011). See Ref. [365] for more details on the applicability of the wide-band limit approximation.

Introduction

The utility program *green* calculates the density of states (DOS) and zero-bias transmission of molecule connected to two semi-infinite contacts. A typical calculation consists of two parts. The first is the calculation of the effect of the semi-infinite contacts, contained in the so-called self-energy matrix. The second is the calculation of the desired properties of the molecule with the self-energies.

Self-energy

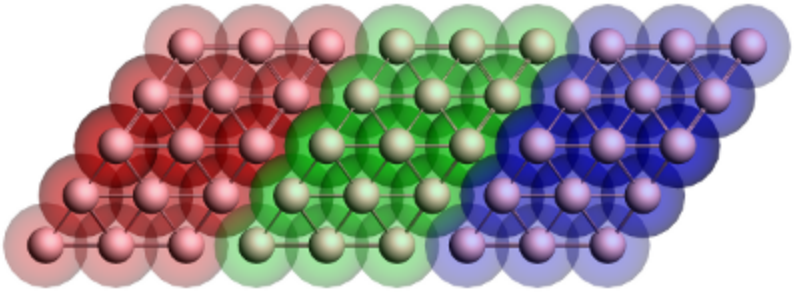


Figure 1: Geometry of the gold contact used in the calculation of the self-energy. The lead consists of two surface layers, left (red) and right (blue), and a bulk layer (green). Each principal layer in turn consists of three atomic layers. This should be sufficient to ensure that the Hamiltonian of the central (green) layer is a bulk Hamiltonian.

Since the contacts are semi-infinite, the calculation of their self-energy is effectively a bulk calculation. Since ADF only works with systems of finite size, approximations have to be made. Fig. 1 shows the typical geometry used in the calculation of a gold contact. The geometry consists of three parts, the so-called principal layers. These layers should be large enough that the atoms on one side are not influenced by whatever is attached to the other side. Three atomic layers usually suffice. The green region is the bulk layer. The red and blue regions are the surface layers. Note that the blue region corresponds to the left contact of a molecule and the red region to the right contact.

To calculate the self-energy, we first need to do a single-point calculation of a principal layer. This layer is then used as a fragment in the following calculations. Note that all ADF calculations have to be performed with SYMMETRY NOSYM. We then build up the contact geometry from three copies of the layer fragment as in Fig. 1 and perform another single-point calculation. This results in a Hamiltonian describing the three contact layers and the coupling between them.

From the TAPE21 file *green* can now calculate the self-energy matrices with the SURFACE key. This has to be done once for every energy for which we want to calculate the DOS or transmission. For the left contact of the molecule, *green* needs the blue and green fragments. The self-energy is calculated by taking the (blue) surface layer and iteratively adding more (green) bulk layers until matrices converge to the semi-infinite result. The self-energy of the right contact is similarly calculated from the red and green fragments. Since the self-energy described the effect of an infinite chain of (green) bulk regions on a (red or blue) surface layer, this calculation does not depend on whatever is attached to the contacts. The self-energy matrices can therefore be reused for different molecules.

DOS and transmission

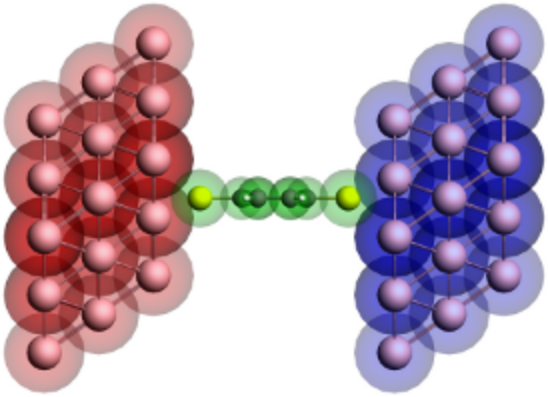


Figure 2: Geometry of the extended molecule used in the calculation of a benzenedithiol junction. The molecule is shown in green, while the left and right contact regions are shown in red and blue, respectively. Note that the red region corresponds to the blue surface layer in Figure 1 and vice versa.

Once the self-energy matrices have been calculated for the desired energies, we can compute the DOS and transmission of a molecule. However, since the self-energy matrices couple to the surface layers of the contacts, we need to include those surface layers in the calculation of the molecule (see Fig. 2). We therefore first perform a single-point calculation with ADF of the isolated molecule. The result is then used as a fragment and combined with the fragments of the surface layers to construct the so-called extended molecule. We then perform another single-point calculation of the final geometry.

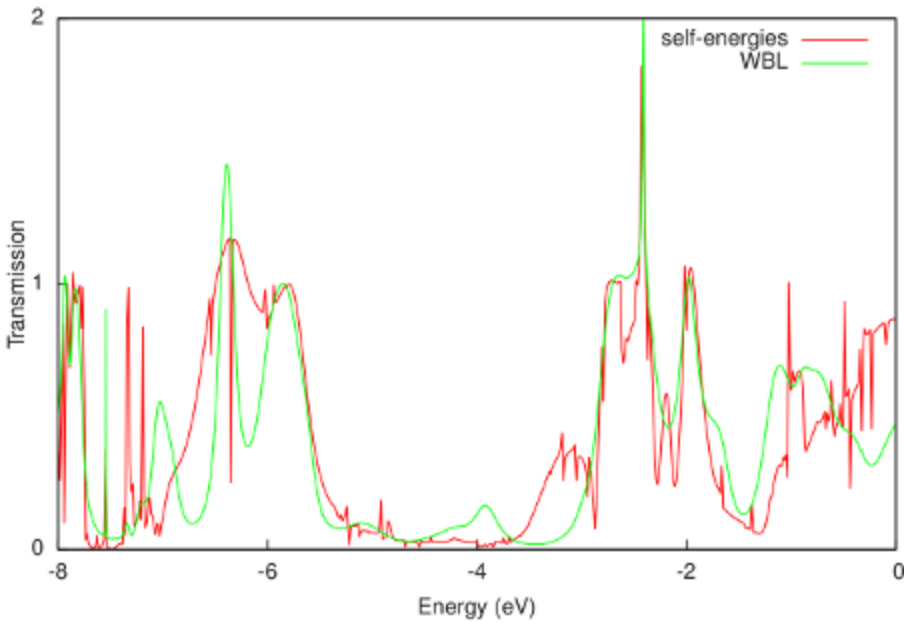
From the self-energies of the contacts and the TAPE21 file of the extended molecule, *green* can now compute the DOS and transmission. This calculation is non-self-consistent since the ADF calculations are all performed on finite instead of semi-infinite systems. This will result in certain artifacts in the DOS and transmission spectra, but those can be made arbitrarily small by choosing the principal layers large enough.

Wide-band-limit

In the wide-band limit (WBL) the coupling to the leads is assumed to be independent of energy. Therefore one does not need to calculate any self-energies. This also means that the eigenspace of the Green's function is independent of energy. It can therefore be diagonalized in advance, greatly speeding up the calculation of the DOS and the transmission. See Ref. [365] for more details on the applicability of the wide-band limit in DFT-based molecular transport calculations.

In the example $ADFHOME/examples/adf/green_Al/green_WBL.run of *green*, the transmission of benzenedithiol junction in the wide-band limit (WBL) is calculated. In order to model the molecule-metal interface, we do need to include a few gold layers in the calculation. However, unlike before, only a single atomic layer as the principal layer is used. Because a single atomic layer is an unnatural configuration for gold, a minor amount of smearing is necessary to make the calculation converge. The molecule is sandwiched in between the electrodes just like before (see Fig. 2 in the example for benzenedithiol). However, this time each atomic layer of gold gets its own fragment. The reason for this configuration is that if the WBL is used on the entire gold contact the result is an an unphysical coupling to the leads; even the gold atoms contacting the molecule would have a direct coupling to the environment. A much better result can be obtained by only using the WBL on the back-most atomic layer and letting the electrons propagate naturally through the rest of the contact. Because the WBL is computationally so inexpensive, we can easily calculate the DOS and transmission for 10,000 points instead of 1000.

A comparison of the resulting transmission with the calculation with self-energies is shown in the following figure:



The WBL shows good agreement with the non-WBL transmission around the Fermi energy (-0.195 Hartree or -5.306 eV). Note that the quality of the WBL depends on the choice of the coupling (ETA). For this particular contact geometry we obtain good agreement for ETA = 0.02 Hartree, but a better value may be found for other electrodes. Finally, the WBL can be incrementally improved by adding more gold layers to the extended molecule. For many layers it converges to the calculation with full self-energies.

Input options

The input for *green* is keyword oriented and is read from the standard input. *green* is typically first used to calculate the self-energy matrices of the left and right contacts (with the SURFACE key), and then to calculate the density of states (DOS) and transmission (with the DOS and TRANS keys, respectively), using those self-energies. The only keyword required to be present in all calculations is the EPS keyword, which specifies the energy range.

```
$ADFBIN/green << eor
 EPS mineps maxeps numeps
 {ETA eta}
 {SO sh sl {moc}}
 {SURFACE filename
     FRAGMENTS f1 f2
  END}
 {DOS filename}
 {TRANS
  LEFT filename
     FRAGMENT fragment
     ETA eta
  END
  RIGHT filename
     FRAGMENT fragment
     ETA eta
  END}
eor
```

`EPS mineps maxeps numeps`

The energy range for which either the self-energy matrices or the DOS and transmission have to be calculated. The range consists of numeps (≥ 1) points running from mineps to maxeps inclusive.

`(optional) ETA eta`

The imaginary energy, or the distance from the real axis, in the calculation of the Green's function. The value needs to be a small positive number to prevent singularities in the calculation. The default value (10^{-6} Hartree) is sufficient for most calculations.

`(optional) SO sh sl {moc}`

The shifts for the scissors operator. All occupied orbitals (HOMO and below) are shifted by sh, while the unoccupied orbitals (LUMO and above) are shifted by sl. Orbitals are considered occupied if their (possibly fractional) occupation is larger than moc (default 0). The scissor operator can partially remedy the underestimation of the HOMO-LUMO gap in DFT. The sh and sl shifts generally have the same magnitude, but opposite sign (with sh usually being negative and sl positive). A good estimate for the magnitude is the sum of the ionization potential and the energy of the HOMO of the *free* molecule. This can be improved by including image charge effects. For more details, see Ref. [351]. By default, sh=sl=0.

`(optional) SURFACE`

```
SURFACE filename
    FRAGMENTS f1 f2
END
```

The SURFACE block key enables the calculation of the self-energy matrices. The filename specifies the TAPE21 file resulting from an ADF calculation of the contacts. This calculation has to be performed with SYMMETRY NOSYM. The FRAGMENTS key is used to specify the two principal layers between which the surface is defined. The resulting self-energy matrices (one for every energy point given by EPS) is stored in a binary KF file named SURFACE.

`(optional) DOS filename`

The DOS key enables the calculation of the density of states. The filename specifies the TAPE21 file containing the result of an ADF calculation of the extended molecule (performed with SYMMETRY NOSYM). Two text files will be generated: DOS_A and DOS_B, containing, respectively, the DOS of the spin-A and spin-B electrons. In the case of a spin-unrestricted calculation, DOS_A and DOS_B might differ. If only the DOS of the spin-A electrons is required, the calculation can be sped up by specifying NOSAVE DOS_B. The DOS key requires the presence of the LEFT and RIGHT keys.

`(optional) TRANS`

The TRANS key enables the calculation of the transmission. The filename specifies the TAPE21 file containing the result of an ADF calculation of the extended molecule (performed with SYMMETRY NOSYM). Two text files will be generated: TRANS_A and TRANS_B, containing, respectively, the transmission of the spin-A and spin-B electrons. In the case of a spin-unrestricted calculation, TRANS_A and TRANS_B might differ. If only the transmission of the spin-A electrons is required, the calculation can be sped up by specifying NOSAVE TRANS_B. The TRANS key requires the presence of the LEFT and RIGHT keys.

`LEFT/RIGHT`

```
LEFT filename
    FRAGMENT fragment
    ETA eta
END
```

The LEFT and RIGHT block keys specify the left and right self-energies used in a calculation of the DOS and transmission. If a filename is specified, the self-energy matrices are read from that file. The energy range of the self-energies has to be consistent with the range specified by the EPS keyword. The FRAGMENT key is used to denote the fragment in the extended molecule (given by the argument to the DOS or TRANS key) to which the self-energy couples. If no filename is specified, the wide-band limit is used. The ETA key can then be used to specify the magnitude of the coupling (10^{-3} Hartree by default).

Output

After a successful calculation of the self-energy matrices, *green* produces a binary KF file named SURFACE containing two sections. The Surface section contains the energy range:

contents of Surface	comments
mineps	start of the energy range
maxeps	end of the energy range
numeps	number of points

The Sigma section contains the real and imaginary parts of the self-energy matrices:

contents of Sigma	comments
nfo	number of fragment orbitals (dimension of the self-energy matrices)
Re(Sigma_%d)	the real part of the %d self-energy matrix (numbered from 1 up to numeps)
Im(Sigma_%d)	the imaginary part of the %d self-energy matrix (numbered from 1 up to numeps)

A successful calculation of the density of states (DOS) or transmission results in the text files DOS_A and DOS_B, and TRANS_A and TRANS_B, respectively. The suffixes _A and _B denote the different spins. The text files the DOS and transmission for every energy point and can be plotted with, for example, gnuplot.

2.8 Analysis

See also
ADF-GUI tutorial: all ADF tutorials, fragment analysis
GUI manual: analysis
Examples: analysis

Molecules built from fragments

ADF analyzes the results in terms of user-specified subsystems from which the total system is built. The program tells you how the 'Fragment orbitals' (FO's) of the chemically meaningful sub-units mix with FO's on other fragments to combine to the final molecular orbitals.

ADF builds a molecule from user-defined fragments, which may be single atoms or larger moieties, for example, ligands, functional groups, or complete molecules in a donor-acceptor complex. In practice, this means that the results of the ADF calculation on a fragment are saved on a file and that the fragment files are then used in setting up the calculation on the overall system. The fragment orbitals (FOs), i.e., the MOs from the calculations on the fragments, are employed as basis functions in the new calculation. This does not imply a basis set truncation or contraction because the virtual FOs are included: the FOs constitute only a transformation of the basis set. If there are symmetry-equivalent fragments, for example, the six CO

molecules in octahedral $Cr(CO)_6$, the program generates symmetry combinations of the FOs and uses the symmetrized fragment orbitals (SFOs) as basis functions. The SFOs transform as the irreducible representations (irreps) of the molecule, allowing a symmetry-driven analysis of the results. In absence of any symmetry the SFOs are identical to the FOs.

The fragment approach offers considerable advantages. It enhances the interpretative power of ADF as it leads to a more transparent picture of bonding, which reduces from a complicated mixing of many primitive basis functions (possessing little physical relevance) to a few key interactions between meaningful fragment (frontier) orbitals. The fragment approach also improves the numerical precision. In ADF, energies are calculated directly, with respect to the fragments, by one single numerical integral of the difference energy density $\varepsilon[\rho,\mathbf{r}]-\sum_A\varepsilon_A[\rho_A,\mathbf{r}]$ between the overall molecule and the constituting fragments.

$$\Delta E[\rho] = \int d\mathbf{r}\ (\varepsilon[\rho,\mathbf{r}]-\sum_A\varepsilon_A[\rho_A,\mathbf{r}]\ )$$

In other words, we evaluate the energy of the overall molecule, $E[\rho] = \int d\mathbf{r}\ \varepsilon[\rho,\mathbf{r}]$, and the energies of each of the fragments, say the atoms that constitute the overall molecule, $E_A= \int d\mathbf{r}\ \varepsilon_A[\rho_A,\mathbf{r}]$, in the same numerical integration grid. This provides more accurate relative energies than subtracting total energies from separate calculations, because the same relative numerical integration error applies to a much smaller quantity, yielding, in turn, a much smaller absolute error.

Note that the user has the freedom to make his own choice of fragments. This is, however, not a matter of plain arbitrariness, and it does not make the analysis tools less meaningful. On the contrary, this freedom simply reflects the many perspectives from which a particular chemical phenomenon can be viewed.

In practice, many calculations are performed using as fragments the so-called basic atoms, which are the smallest possible building blocks in ADF. The basic atoms are not necessarily physically realistic objects - indeed, usually they are not, as they must be spin-restricted and spherically symmetric. The computed (bonding) energy w.r.t. basic atoms, then, does not yield quantities that can be compared to experimental data directly. Rather, one must correct for the true ground state of the isolated single atoms.

Text is mostly taken from: *Chemistry with ADF*, G. te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler J. Comp. Chem. **22** (2001) 931.

Link: How to make molecular fragments Tutorial: ADF fragment analysis
Examples: analysis options

Bond energy analysis

No special input keys are required, except if one has open shell fragments, see key FRAGOCCUPATIONS.

ADF calculates various chemically meaningful terms that add up to the bond energy, with an adaptation of Morokuma's bond energy decomposition to the Kohn-Sham MO method. The individual terms are chemically intuitive quantities such as electrostatic energy, Pauli repulsion, and orbital interactions. The latter are symmetry decomposed according to the Ziegler transition state method. For a discussion of bonding energy decompositions and applications see e.g. [3, 110, 112, 130-136]

In ADF2012 the calculation of the Pauli repulsion for metaGGA's and metahybrids is implemented. Note that for hybrids this was already implemented before in case of closed shell fragments. In ADF2012 for hybrids, metaGGA's, and metahybrids the calculation of the Pauli repulsion is also implemented if one is simulating an unrestricted fragment with the key FRAGOCCUPATIONS.

In ADF2012 for hybrids the exact exchange contribution to the Pauli term is isolated and the contributions to the orbital term are divided amongst orbital symmetries.

Bond energy details

In the framework of Kohn-Sham MO theory and in conjunction with the fragment approach, one can decompose the bond energy between the fragments of a molecular system - say, a base and a substrate for E2 elimination - into contributions associated with the various orbital and electrostatic interactions. In ADF, we follow a Morokuma-type energy decomposition method. The overall bond energy ΔE is divided into two major components. In the first place, the preparation energy ΔE_{prep} corresponding to the amount of energy required to deform the separated fragments, A and B say, from their equilibrium structure to the geometry they acquire in the overall molecule ($\Delta E_{prep,geo}$), and to excite them to their valence electronic configuration ($\Delta E_{prep,el}$). In the second place, the interaction energy ΔE_{int} between the prepared fragments.

$$\Delta E = \Delta E_{prep} + \Delta E_{int} = \Delta E_{prep,geo} + \Delta E_{prep,el} + \Delta E_{int}$$

In the following step, the interaction energy ΔE_{int} is further decomposed into three physically meaningful terms, which are printed in the ADF output file.

$$\Delta E_{int} = \Delta V_{elst} + \Delta E_{Pauli} + \Delta E_{oi} = \Delta E^0 + \Delta E_{oi}$$

The term ΔV_{elst} corresponds to the classical electrostatic interaction between the unperturbed charge distributions of the prepared fragments as they are brought together at their final positions, giving rise to an overall density that is simply a superposition of fragment densities $\rho_A+\rho_B$. (Note that we use the convention that energy terms containing potential energy only, kinetic energy only, or both kinetic and potential energy are indicated by V, T, and E, respectively.) For neutral fragments, ΔV_{elst} is usually attractive. The Pauli repulsion ΔE_{Pauli} arises as the energy change associated with going from $\rho_A+\rho_B$ the wave function $\Psi^0=NA[\Psi_A\Psi_B]$ that properly obeys the Pauli principle through explicit antisymmetrization (*A* operator) and renormalization (*N* constant) of the product of fragment wave functions. It comprises the destabilizing interactions between occupied orbitals, and is responsible for any steric repulsion. In case of neutral fragments, it can be useful to combine ΔV_{elst} and ΔE_{Pauli} in a term ΔE^0 which, in the past, has been conceived as the steric interaction. However, we prefer to reserve the designation steric interaction or repulsion for ΔE_{Pauli} because that is, as already mentioned, the only source of net repulsive interactions between molecular fragments. Finally, the wavefunction is allowed to relax from Ψ^0 to the fully converged wave function Ψ. The associated orbital interaction energy ΔE_{oi} accounts for electron pair bonding, charge transfer (e.g., HOMO-LUMO interactions) and polarization (empty/occupied orbital mixing on one fragment due to the presence of another fragment). This can be further decomposed into the contributions from the distinct irreducible representations Γ of the interacting system using the extended transition state method. In systems with a clear σ/π separation, this symmetry partitioning proves to be very informative.

$$\Delta E_{oi} = \sum_\Gamma \Delta E_{oi,\Gamma}$$

An extensive discussion of the physical meaning of all the terms in the energy decomposition is given in
F.M. Bickelhaupt and E.J. Baerends,
Kohn-Sham Density Functional Theory: Predicting and Understanding Chemistry,
In: Rev. Comput. Chem.; Lipkowitz, K. B. and Boyd, D. B., Eds.; Wiley-VCH: New York, 2000, Vol. 15, 1-86.

Text is mostly taken from: *Chemistry with ADF*, G. te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler J. Comp. Chem. **22** (2001) 931.

Total energy evaluation

ADF normally does not calculate the total energy of a system (the energy wrt bare nuclei and free electrons). However, ADF calculates the energy of the system with respect to fragment energies. By default, these

fragments are the spherical spin-restricted neutral atoms, but one can also use larger fragments. For this reason total energies from other programs could not be compared to ADF directly. Note, however, that only energy difference comparisons are meaningful. These are the only energies that play a role in chemistry of course, and for this one does not need total energies.

If you really want to calculate the total energies, there are two options in ADF

Total energy by adding the binding energy of the atoms

There is a work-around to calculate the total energy of a system: calculate the total energies of the atomic fragments and add them to the bonding energy. Because total energy of an atom is, by definition, the energy difference between the atom and the (nucleus+free electrons) system one can calculate it by calculating a single atom with the charge equal to the number of electrons. 'Bonding energy' of such an 'atom' will then be equal to negative of the total energy of the atomic fragment. Care should be taken to apply this procedure to frozen-core fragments. In this case, it only makes sense to remove the valent electrons and leave the frozen core.

TOTALENERGY keyword

The total energies have not been tested extensively and should therefore be used with caution. In addition to bond energies it is now possible to compute total energies with ADF by including the keyword TOTALENERGY in the input. This work is in progress.

```
TOTALENERGY
```

The total energy will be computed for the chosen XC functional (LDA, GGA, hybrid functionals, or Hartree-Fock). MetaGGA functionals, (ZORA) scalar relativistic and relativistic spin-orbit calculations, electric fields and QM/MM are not supported yet.

In particular the requirements to the integration accuracy are somewhat higher than for bond energies. It is recommended to use an integration grid (BeckeGrid) of quality "Good". If in doubt, a convergence test with respect to the integration accuracy is recommended.

Symmetry

Together with the point group symmetry, a tolerance parameter can be supplied.

```
SYMMETRY {symbol} {tol=tolerance}
```

`symbol`

The Schönfliess symmetry symbol. A complete list of allowed values for this argument is given in Appendix 5.3.

`tolerance`

The tolerance (absolute deviation in the Cartesian coordinates) for atomic positions being symmetry equivalent. The same tolerance applies to check the mapping of fragments on attached fragment files with the actual fragments.
If the tolerance is specified it is interpreted in the chosen unit of length (units). The default tolerance is 0.001 Angstrom and the maximum is 0.1 a.u.

Input atomic coordinates that are slightly (within the tolerance) off from their correct positions are adjusted by the program.

Localized Molecular Orbitals

ADF provides the Boys-Foster method for localization of Molecular Orbitals [122-124]. This implies a unitary transformation of the occupied molecular orbitals as computed in the SCF procedure, with the objective to obtain a (transformed) set of orbitals that represent exactly the same charge density but with molecular orbitals that are more localized in space than the original MOs.

The goal of orbital-localization lies in analysis: the localized orbitals provide an easier-to-interpret picture.

Orbital localization procedures require a measure of the localization of the orbitals which can then be optimized in the space of the allowed unitary transformations. Methods advocated in the literature differ in the definition of this measure. The Boys-Foster method minimizes the mean extension of the occupied orbitals around their center of gravity; see the literature for details.

Occasionally it is useful to apply the localization only to a subset of the MOs, with the objective to expose certain features better. This is accomplished by performing the localization in a number of distinct steps, where at each step the localization is restricted by keeping a subset of the MOs frozen. A case is worked out in the Examples document.

The computation of localized orbitals is controlled with the block-type key. By default (if the key is not supplied in input) no orbital localization is carried out.

```
LOCORB {nopop store}
  Spintype FrozenMOs
  Spintype FrozenMOs
  ...
end
```

`nopop`

Specifies that no SFO population analysis is to be carried out on the localized MOs. By default this population analysis will be printed in the output file.

`store`

Specifies that the transformation from MOs to localized MOs is stored on TAPE21.

`Spintype`

Must be either alfa or beta (not case sensitive) and refers to spin-A and spin-B orbitals respectively. In a spin-restricted run beta records are meaningless and must not be used.

`FrozenMOs`

A list (possibly empty) of integers, referring to a list of MOs from the SCF, and/or labels of irreducible representations. The integers and/or labels may be given in any order.

Each record Spintype FrozenMOs in the data block defines a localization *cycle* in which the localization procedure is carried out on all orbitals (of the indicated spin), except those indicated by the FrozenMOs.

For either spin at least one localization cycle is carried out. If no data record for that spin is found in the data block, a full localization is performed, without any MOs excluded.

The data block may be completely empty (but the record end must be supplied since the key is block-type) and would be equivalent with specifying two records, one for either spin, without any FrozenMOs:

```
LOCORB {nopop}
end
```

is equivalent with

```
LOCORB {nopop}
  alfa
  beta
end
```

The integers in FrozenMOs refer to an overall list of SCF MOs consisting of all valence MOs in each symmetry representation up to and including the highest non-empty one. So, when for instance in the first irrep MO #4 is the highest non-empty one and in the second irrep mo #2 is the highest non-empty one, then in the overall list the first 4 are the orbitals of the first irrep, the no.s 5 and 6 are from the second irrep, et cetera.

Each symmetry label in FrozenMOs collectively denotes in one stroke all molecular orbitals of that representation up to and including the highest occupied one (in that symmetry). The label may be the name of an irreducible representation or of a subspecies. In the former case all partner representations are denoted collectively. In an atom symmetry for instance, specifying P would be equivalent to P:x P:y P:z.

Note that if the final SCF has in any symmetry representation empty orbitals *below* the highest non-empty orbital in that symmetry - violating the Aufbau principle - then these empty orbitals are included in the above-defined overall list and hence a FrozenMOs specification is necessary, namely to avoid mixing MOs with different occupation numbers in the localization.

Note:

It is imperative that in a particular localization cycle only MOs from the SCF are combined that have identical occupation numbers. If this is violated the program will carry out the localization without error message, but the results are incorrect in the sense that the density defined by the localized orbitals is *not* the same anymore as the SCF density.

So, if any of the MOs in the overall list defined above is not *fully* occupied (open shell, excited state, ...) you need to define precisely the localization cycles - localizing in each cycle only MOs with identical occupations and freezing all others - in order to obtain sensible results.

In the output file the localized MOs are printed as expansions in SFOs and (optionally) a population analysis is given, again in terms of the SFOs. Furthermore, each localized MO has associated with it an energy value and an occupation number. The energy is the expectation value of the Fock operator for the orbital. The occupation number is obtained as a weighted sum from the SCF MOs that were combined into the localized orbital. As mentioned before one should combine only SCF MOs with identical occupations into a localized orbital, in which case its occupation number will be the same. The printout of the occupation number of the localized orbital allows therefore a verification that a correct localization procedure has been carried out.

Advanced charge density and bond order analysis

In addition to Mulliken charge analysis, ADF calculates several atomic charges that do not share the flaws of Mulliken (strong basis set dependence). The multipole-derived charge analysis exactly reproduces dipole and higher multipole moments of the molecule. Other charge analysis methods ('Voronoy deformation density' and 'Hirshfeld' provide atomic charges that agree well with chemical intuition. Nalewajski bond orders can be calculated and show good agreement with experimental trends and chemical intuition, even for transition metal compounds.

Note that the amount of data can be regulated with the keys PRINT, NoPrint, EPrint and Debug.

Charges, Populations, Bond orders

Mulliken populations

See the input key EPRINT. See also the section on Mulliken populations.

Hirshfeld charges, Voronoi deformation density

No special input key required. See also the section on Hirshfeld charges, Voronoi deformation density.

Multipole derived charges

No special input key required. See also the section on MDC.

Charge model 5 (CM5)

Charges calculated with CM5 activated by keyword

```
CM5
```

See also the section on CM5.

Bond orders

See also the section on bond order analysis.

Nalewajski-Mrozek Bond orders

Bond order analysis in ADF is activated by keyword

```
BONDORDER {tol=xxx} {printall}
```

By default bond order indices calculated by the Nalewajski-Mrozek [148-152] method are calculated. There exist three alternative definitions of the valence and bond order indices within the Nalewajski-Mrozek approach. By default the values obtained from partitioning of Tr(PΔP) are calculated and printed in the output. For more information on alternative Nalewajski-Mrozek bond order indices see Results/Properties section (4.1).

`tol=xxx`

The tol=xxx option specifies the threshold value for bond orders to be printed in the output (default=0.2).

`printall`

The values calculated from all three versions of the Nalewajski-Mrozek approach are printed when the option printall is present; in addition the Gopinathan-Jug [153] and Mayer [140] bond order indices are calculated for comparison.

Present bond order analysis is based on SFOs. Symmetry used in the calculation should be NOSYM. For this reason the analysis may be used only if the symmetry in the calculation is NOSYM. The analysis may be used also for multi-atomic fragments, the fragment-fragment bond orders are printed in such a case. Note that in the present implementation all fragment types should be different.

Mayer Bond orders

The Mayer bond orders are calculated and printed if the keyword

```
EXTENDEDPOPAN
```

is included in the input. Next to the Mayer bond orders Mulliken atom-atom populations per l-value will be calculated and printed if this keyword is included in the input. Note that this keyword is not a subkey.

ETS-NOCV: Natural Orbitals for Chemical Valence

With the ETS-NOCV charge and energy decomposition scheme the deformation density is partitioned into the different components (σ, π, δ) of the chemical bond. The energy contributions to the total bond energy is calculated for each specific orbital interactions between fragments, giving insight in the orbital interactions also for non-symmetric molecules. The ETS-NOCV analysis offers a compact quantitative picture of the chemical bond, which is also qualitatively attractive to chemists.

Theory

The Natural Orbitals for Chemical Valence (NOCV) approach has been derived from the Nalewajski-Mrozek valence theory [148, 150]. From the mathematical point of view, each NOCV ψ_i is defined as an eigenvector of the deformation density matrix in the basis of fragment orbitals.

$$\Delta P\, \psi_i = v_i\, \psi_i$$

Thus, the deformation density $\Delta\rho$ can be expressed in the NOCV representation as a sum of pairs of complimentary eigenfunctions (ψ_{-k},ψ_k) corresponding to eigenvalues $-v_k$ and v_k with the same absolute value but opposite signs:

$$\Delta\rho(r) = \Sigma\, \Delta\rho_k(r) = \Sigma\, v_k[-\psi^2_{-k}(r)+\psi^2_k(r)]$$

here, k goes over the pairs of NOCV's.

In the combined ETS-NOCV scheme the orbital interaction term ΔE_{orb} is expressed in terms of NOCV's as [261, 262]:

$$\Delta E_{orb} = \Sigma\, \Delta E_k^{orb} = \Sigma\, v_k[-F^{TS}_{-k}+F^{TS}_k]$$

here, F^{TS}_{-k} and F^{TS}_k are diagonal transition-state Kohn-Sham matrix elements corresponding to NOCV's with eigenvalues $-v_k$ and v_k, respectively. The advantage of this expression is that usually only a few complimentary NOCV pairs significantly contribute to the total ΔE_{orb}. Another advantage of this approach is that not only can each $\Delta\rho_k(r)$ be visualized but there is also a well defined bonding energy contribution ΔE_k^{orb} corresponding to it.

Remarks

The ETS-NOCV analysis is often not very useful when atomic fragments are used. No symmetry must be used in the final calculation, thus, use a Symmetry NOSYM keyword if your molecule is symmetric. The analysis is not completely implemented for meta-GGA's and meta-hybrids.

Improvements in ADF2012 to both the ETS and NOCV analysis with hybrids. ETS: Now the exact exchange contribution to the Pauli term is isolated and the contributions to the orbital term are divided amongst orbital symmetries. NOCV: The exact exchange contribution to the Fock operator is included when calculating energy contributions. These changes do not apply to meta-hybrids.

Usage

In order to perform the ETS-NOCV analysis, the following two keywords must be specified at the same time:

```
ETSNOCV RHOKMIN=rhokmin EKMIN=ekmin ENOCV=enocv
PRINT {ETSLOWDIN | ETSLOWDIN-Unrestricted}
```

ETSNOCV

The ETSNOCV keyword specifies thresholds for printing of NOCV-related information. All three arguments are optional and when all three are omitted only the NOCV's corresponding to eigenvalues abs(v_k) ≥ 0.05 are included in the analysis.

RHOKMIN

The threshold for population analysis of each deformation density contribution in terms of individual SFO's.

EKMIN

The threshold for orbital interaction energy contributions corresponding to deformation density components originating from each NOCV-pairs.

ENOCV

The threshold for NOCV-eigenvalues.

PRINT {ETSLOWDIN | ETSLOWDIN-Unrestricted}

Only one of the two PRINT options is supposed to be used to activate printing of ETS-NOCV results. The choice depends on the bonding situation.

ETSLOWDIN

If one is interested in a description of bonding between closed-shell molecular fragments, then 'PRINT ETSLOWDIN' keyword must be used. In such a case one set of NOCV's originating from the total deformation density matrix $\Delta P=(\Delta P_\alpha+\Delta P_\beta)$ will be printed out. See the example of carbene bonding between closed shell CH2 and Cr(CO)5.

ETSLOWDIN-Unrestricted

If, however, one is interested in a description of bonding between open-shell molecular fragments then the 'PRINT ETSLOWDIN-Unrestricted' keyword must be used. In this case two sets of NOCV's originating from ΔP_α and ΔP_β will be printed out. See the example of CH3-CH3 bonding between two CH3 radicals with opposite spins. This option must also be used when one wants to analyze bonding in a molecule with unpaired electrons.

Adfnbo, gennbo: NBO analysis

Dr. Autschbach, SCM, and Prof. Weinhold have collaborated to prepare a simple in put file generator, called adfnbo, for the GENNBO program of Prof. Weinholds Natural Bond Orbital (NBO) package. In ADF2013 the NBO 6.0 version is supported http://nbo6.chem.wisc.edu.

The GENNBO executable is included in the ADF distribution and can be enabled via the license file for all those who buy an NBO manual from either the NBO authors or from SCM (info@scm.com). An extensive documentation of GENNBO is part of the NBO manual. The application of ADFNBO to frozen-core basis sets needs to be further tested. Usage can be found below and in the Examples Document.

Next a brief summary of the capabilities of GENNBO is given (by Prof. Weinhold).
GENNBO implements most capabilities of the full NBO 6.0 program suite as described on the NBO website: http://nbo6.chem.wisc.edu

These include determination of natural atomic orbitals (NAOs), bond orbitals (NBOs), and localized MOs (NLMOs), as well as the associated NPA (atomic charges and orbital populations) and NRT (resonance structures, weightings, bond orders) valence descriptors, for a wide variety of uncorrelated and correlated (variational, perturbative, or density functional) theoretical levels. GENNBO-supported options include all keywords except those explicitly requiring interactive communication with the host electronic structure system (viz., $DEL deletions, NEDA, NCS, NJC). The GENNBO program typically sits conveniently on the PC desktop, ready to analyze (or re-analyze at will, with altered options) the final results of a complex ADF calculation performed on a remote cluster.

GENNBO "communicates" with the original ADF calculation through an archive file (JOB.47 file, preserving all necessary details of the final density) that is initially generated by ADF and subsequently becomes the input file for GENNBO. The .47 file contains a standard $NBO ... $END keylist that can be edited with a standard word processor or text editor to include chosen NBO keyword options, just as though they might have appeared in the original input stream of an interactive ADFNBO run. The stand-alone GENNBO program therefore allows many alternative NBO analysis options to be explored at leisure, without costly re-calculation of the wave function.

Using the GENNBO executable is possible only if NBO6 is enabled in your license file by SCM. In that case you will get access to an NBO 6.0 manual in electronic form that explains in detail how GENNBO can be used and how the output should be interpreted.

NBO analysis of EFG, NMR chemical shifts, NMR spin-spin coupling

For certain molecular properties it is possible to perform detailed analyses in terms of Natural Bond Orbitals (NBOs) and Natural Localized Molecular Orbitals (NLMOs). These features generally require a sequence of ADF and/or property code runs. An initial nonrelativistic or scalar relativistic ADF run, followed by the generation of NBO and NLMO data, is required, and the resulting data files need to be present in subsequent property calculations, along with a keyword indicating that the NBO analysis is requested in the property module.

We have noted in the past some slight loss of numerical accuracy of the results after going through the various orbital transformations in the NBO - NLMO sequence. It is important that the user verifies in each case that the total contributions from the analysis are in agreement with the total calculated property, within the numerical integration accuracy limits. In order to assist the user with this, the analysis program always print the total analysis contributions, including small nonprinted values.

Moreover, there appears to be a problem with the analysis of the Fock matrix in the NBO program in conjunction with ADF calculations. Therefore please do NOT use the Fock matrix second order perturbation theory analysis in NBO at this time. We will remove this disclaimer once the issue has been fixed. Applications of the NBO-NLMO property analysis codes have so far given no indication that the Fock matrix issue interferes with the analysis.

Important note: If properties are analyzed from within spin-orbit relativistic computations, the NBO/NLMO analysis is performed in terms of scalar (spin-free) relativistic orbitals, as detailed in the technical references. The results from these analyses are exact in the sense that they fully reproduce the final spin-orbit property result, and they allow to dissect the property in terms of more intuitive one component real scalar relativistic localized orbitals. Typically, the property analysis in a spin-orbit calculation involves contributions from unoccupied scalar NLMOs, whereas there are no such contributions if a nonrelativistic or scalar relativistic property is analyzed.

Available properties for NBO analysis: EFG, NMR chemical shifts and NMR spin-spin coupling.

NBO analysis of EFG

EFGs: nonrelativistic and scalar ZORA, in ADF/AORresponse. Requires initial ADF run with

```
AOresponse
 donothing
End
```

in order to generate orbitals that re equivalent to those generated in the subsequent ADF run where the EFG is calculated. Alternatively, simply calculate the EFG twice, once before the NBO generation step, and once afterward.

The next step (see below) is to create the NBOs and the required data files for the analysis. Afterward, in the second ADF run, use

```
Aoresponse
 efg NUC nbo
end
```

`efg NUC nbo`

> Here NUC is the number of the nucleus at which the EFG is to be computed (ADF internal atom ordering). Example: efg 1 nbo.

In addition to the optional NBO analysis, the EFG program in AOResponse prints a Mulliken type analysis of the EFG principal components, and an analysis in terms of canonical MOs.

WARNING: the ordering of the principal components is lowest to highest including the sign. That is, we have V11 ≤ V22 ≤ V33. This does not conform to the usual convention of |V11| ≤ |V22| ≤ |V33|. Please make sure you select the right component for your analysis.

Example job: $ADFHOME/examples/adf/AlCl3_efgnbo.
For an explanation of the output and a general usage tutorial, see [327]. Further references and recommended citations, see [328].

NBO analysis of NMR Chemical shift

An implementation is currently available for spin-orbit ZORA computations. If scalar ZORA calculations are to be analyzed, provide the input keyword

```
FAKESO
```

in the NMR input (outside of the 'nmr' keyword). A native scalar ZORA implementation that does not require this keyword is under development. If this feature is requested one should restrict the calculation to a single shielding tensor per NMR run. It would be good practice to check the results against regular NMR calculations where the analysis feature is not requested. No ZORA scaling is applied in the analysis results. The data should be equivalent to a regular computation in the NMR input with

```
NMR
 u1k best
 calc all
END
```

Depending on whether scalar or spin-orbit calculations are to be analyzed, the sequence of calculations is different:

scalar:
1. ADF, scalar ZORA
2. generate NBOs and required data files for analysis
3. NMR with FAKESO and analysis keywords, use TAPE21, TAPE10 from step 1.

spin-orbit:
1. ADF, scalar ZORA

2. generate NBOs and required data files for analysis
3. delete TAPE21, TAPE10, TAPE15
4. ADF, spin-orbit ZORA
5. NMR with analysis keywords, using TAPE21, TAPE10 from step 4

In the NMR run, in addition to the NMR keyword, provide the following

```
analysis
 print 0.01
 canonical
 nbo
 components
end
```

The optional canonical keyword can be used independently from the NBO analysis features. It enables an analysis of the shielding in terms of the canonical MOs. The components keyword is optional and enables an analysis not only of the isotropic shielding but also of each principal component of the tensor. The print keyword selects printout of contributions relative to the total diamagnetic, paramagnetic. In the example, only contributions greater than 1% are printed. Set to zero to print ALL contributions.

Example job: $ADFHOME/examples/adf/CH4_nmrnbo.
References [329-331].

NBO analysis of NMR spin-spin coupling (J-coupling)

Nonrelativistic, scalar ZORA, spin-orbit ZORA

The sequence of jobs is similar to those in the NMR section.

scalar or **nonrelativistic**:
1. ADF, scalar ZORA or nonrel.
2. generate NBOs and required data files for analysis
3. CPL with analysis keyword, use TAPE21, TAPE10 from step 1.

spin-orbit:
1. ADF, scalar ZORA
2. generate NBOs and required data files for analysis
3. delete TAPE21, TAPE10, TAPE15
4. ADF, spin-orbit ZORA
5. CPL with analysis keyword, using TAPE21, TAPE10 from step 4

In the CPL run provide the following 'contributions' keyword to enable the analysis

```
nmrcoupling
 ... other options
 contributions 1E19 nbo
end
```

The numerical value selects a print threshold in SI units of T**2/J for the analysis. Increase the value to obtain less detail in the analysis. By default, 'contributions' triggers an analysis of the J-coupling in terms of canonical MOs. The nbo keyword enables in addition the NBO-NLMO analysis.

Please note that due to the history of how the program was developed the output from the scalar/nrel. analysis and from the spin-orbit calculations differs somewhat. The qualitative content is the same.

In scalar ZORA or nonrelativistic CPL calculations without the SD term an orbital based analysis is only performed for the Fermi-contact mechanism. If you also need an analysis for the PSO and SD mechanisms but do not want to run a spin-orbit calculation with ADF please use the SD or NOSD keywords which will cause the spin-orbit branch of the CPL code to be used. In ZORA spin-orbit calculations the FC, SD, PSO,

and cross terms are analyzed together by default. You can selectively switch them on or off in order to get individual mechanism analyses. The DSO mechanism is often negligible. An analysis tool for this mechanism has therefore not yet been developed.

Example job: $ADFHOME/examples/adf/CPL_CH3OH_NBO.
References NMR spin-spin couplings with NBO analysis [331-334]:

Generation of NBOs

How to generate the NBOs, NLMOs, and the data files needed for these calculations (step 2 below is step 2 in the examples above):

1. run ADF with scalar ZORA or nonrelativistic options, and keep TAPE21 and TAPE15.

2.

```
# run adfnbo in WRITE  mode to create the gennbo input file FILE47
# and one of the required property analysis files, adfnbo.kf

$ADFBIN/adfnbo << eor
write
spherical
end input
eor

rm -f adfnbo.37 adfnbo.39 adfnbo.49 adfnbo.48
$ADFBIN/gennbo6 FILE47

# run adfnbo in COPY mode to create the second property analysis
# file, adfnbo2.kf

$ADFBIN/adfnbo << eor
 spherical
 copy
end input
eor

# run adfnbo in READ mode: prepare locorb on TAPE21

$ADFBIN/adfnbo << eor
 spherical
 read
end input
eor

rm -f adfnbo.37 adfnbo.39 adfnbo.49 adfnbo.48

# keep the TAPE21 after this sequence in order to
# be able to plot the NBOs and NLMOs with adfview:

mv TAPE21 nbonlmo.t21

# clean up, keep adfnbo*.kf for any NBO property analyses.
```

QTAIM: Atoms in Molecules

Starting from the ADF2008.01 version in ADF one can calculate Bader atomic charges using a grid based method. This is based on the quantum theory of atoms in molecules (QTAIM). In ADF2012 one can also calculate critical points of the density and bond paths. Another possibility for Bader's analysis is to use the *adf2aim* utility such that a third party program Xaim can be used.

Bader atomic properties (grid based method)

A fast Bader atomic property calculation (see Refs. [228,229]]) is performed by specifying the BADER keyword in the input file. This calculation can be used for relatively big systems (hundreds of atoms). The default calculation produces atomic electron density populations, charges, density Laplacian, dipole moments and quadrupole moments. In ADF2012 the default calculation also includes the calculation of critical points of the density and bond paths (see Ref. [363]]), which can be visualized with the ADF-GUI.

```
BADER {Energy} {Spacing=value}
```

`Spacing`

Specifies spacing (distance between neighboring points) of the initial grid when searching for critical points, in Angstrom. The default spacing value is 0.5 Bohr (or approximately 0.26 Angstrom). It may be useful to specify a smaller value if the default results in some critical points being missed, which will result in a more accurate but slower calculation.

`Energy`

If the argument Energy is included, also calculate Bader atomic energies. Note that this option to calculate Bader atomic energies is only valid in cases, where the key TOTALENERGY can be used, and that this option further implies that the key EXACTDENSITY is set.

The accuracy of this this calculation [228,229] can be determined by the standard method: the integration of the Laplacian of the electron density must vanish over the Bader atomic basins. The default output produces these integrations. The accuracy of the method can be improved by using larger integration grids. Usually the default grid suffices an average atomic integration accuracy of 10^{-3} a.u. (differences of milliHartree in the energies energies). The convergence of the integration of the electron density Laplacian is not monotone but sinusoidal. So the integration of the Laplacian is not always closer to zero as a larger grid is used. So this type of Bader atomic property calculation might be considered as an approach where computational efficiency is critical and moderate accuracy is sufficient [228,229].

ADF2AIM

The ADF utility adf2aim (original name rdt21) developed by Xavi López, Engelber Sans and Carles Bo converts an ADF TAPE21 to WFN format (for Bader analysis)

The program rdt21 is now called adf2aim and is part of the ADF package, starting from ADF2004.01.

The WFN file is an input file for the third party program Xaim (see http://www.quimica.urv.es/XAIM for details), which is a graphical user interface to programs that can perform the Bader analysis. Usage of adf2aim can be found in the Examples Document.

Printed Output

The amount of printed output is regulated with the keys Print, NoPrint, EPrint and Debug. (No)Print and Debug are simple keys, EPrint is a block type key.

Many print options pertain to debugging situations and are included here only for completeness. This section is intended to give a survey of all possibilities. Some items may be mentioned again in other sections where the subject of a particular print switch is discussed.

Print / NoPrint

```
PRINT Argumentlist
Print Argumentlist
NoPrint Argumentlist
```

`Argumentlist`

A sequence of names separated by blanks or commas.
The keys Print and NoPrint may occur any number of times in the input file. The names in the argument list may refer to various items. For some of them printing is normally on, and you can turn them off with NoPrint. For others the default is not printing; use Print to override that.

Follows a list of the recognized items that are applicable in the argument lists, with a short explanation and defaults. Item names must be used exactly as given in the table - abbreviated or elongated forms will not be recognized - but they are not case sensitive.

Item	*Default*	*Explanation*
$ALL	No	Turns on *all* print options. This will not be affected by any additional noprint instructions. Be careful: this generates a large amount of output. To be used only for debugging purposes.
Atdist	No	Inter-atomic distance matrix at each new geometry (in an optimization)
Bas	Yes	General control of output related to elementary basis functions (bas).
BlockCheck	No	Intermediate data during the determination of the block length. (see Blocks)
Blocks	No	Numerical integrals, consisting of loops over large numbers of points, are split up in loops over blocks of points. The block length is determined by the available amount of workspace. Given this amount, the maximum block lengths, according to memory usage in a few relevant routines, are computed (and printed with this print option) and used to impose upper bounds on the block length actually use.
Character-Table	No	Table of characters for the irreducible representations of the pointgroup symmetry.
Computation	Yes	Reports progress of the computation, with (concise) info about each SCF cycle and each Geometry update in an optimization.
Core	No	Description of the frozen core: frozen core expansion functions (corbas) and the expansion coefficients for the frozen orbitals. This printing can only be activated if Functions is also *on*, otherwise it is ignored.
CoreOrt	No	The valence basis set contains auxiliary Core Functions. They are not degrees of freedom but are used solely to ensure orthogonalization of the valence set to the frozen Core Orbitals. The orthogonalization coefficients and some related overlap matrices are printed.
CoreTable	No	Internally the charge density and potential of the atomic frozen cores are processed as tables with values for a sequence of radial distances. A few initial and a few final values from these tables are printed, along with the (radial) integral of the core density, which should yield the number of core electrons.
EKin	No	At the end of SCF: Kinetic energy of each occupied MO.
EndOf	No	Flags the exit from a few major routines, with cpu times used in these modules. Primarily a debug tool.

EPauli	Yes	The repulsive Pauli term in the bonding energy (also called exchange repulsion) with its decomposition in density functional (lda and nl) and Coulomb terms.
Fit	Yes	General control of output related to the density fitting.
Fmat	No	Fock matrix computed at each cycle of the SCF.
FmatSFO	No	Fock matrix (and overlap matrix) in the basis of symmetrized fragment orbitals (SFOs). This option requires the FULLFOCK and ALLPOINTS keyword to be present in the input. The matrix is printed only at the last SCF cycle. Use 1 iteration in the SCF for the Fock matrix at the first SCF cycle.
ForceConstants	Yes	Force constants matrix (Frequencies run only)
FreqHess	No	matrix of force constants (Frequencies run) after each applicable step in its processing: transformation from/to Cartesian and Z-matrix coordinates, symmetrizations, ...
Frag	No	General control of output related to build-molecule-from-fragments.
Functions	Yes	List of employed Slater-type exponential basis functions and fit functions.
Gradients	No	detailed info of computed energy gradients (in optimization runs)
Group-Operators	No	3*3 matrices of pointgroup symmetry operators, with the axis and angle of rotation
HessEig	No	Eigenvalues of the Hessian in each cycle of a Geometry Optimization. The print-out in the intermediate cycles is suppressed if output of updated coordinates etc. is turned off (see the eprint subkey Repeat (option GeoStep).
Idfree	No	List of free atomic coordinates with indication whether they are optimization coordinates (this info is also contained in the output of new atomic coordinates at each step of an optimization)
Inertia	No	Warning message in the log file in case of zero product of moments of inertia (this may correctly be the case for certain molecules)
Inputkeys	No	List of keys that were specified in input, together with some of the associated data. The list is printed directly after the echo of the Input File, before the header with ADF program information. A few special keys will not be echoed: (No)Print,(No)Skip, Allow.
Irrep-Matrices	No	Irreducible representation matrices
Logfile	Yes	At the end of the calculation a copy of the log file is appended to standard output
low	No	Construction of the LOW basis from the elementary BAS functions and from the SFOs: combination coefficients
lowMO	No	MOs are printed in the LOW (Lowdin) representation, in the RESULTS section
OvlBAS	No	Overlap matrices processed during the construction of the LOW basis
Parser	No	Most input records are echoed twice (at the very beginning of output). First the original version, then the parsed version in which expressions have been replaced, redundant blanks removed, etc. The parsed version is what the program really uses as input. Comment blocks, and function definitions (in define blocks) are not parsed, and are not affected by this switch. If the print switch is off only the original, non-parsed input record is echoed in output. This print switch affects only the part of input after its occurrence.
Pmat	No	The density matrix (in Lowdin representation) in each cycle of the SCF.
QMpot	Yes	At the end of the SCF for each atom the electrostatic potential at its nucleus (excluding its own contribution of course).
	No	Redundant Coordinates used in the construction of the initial - force field derived - Hessian
RedCrdBonds	No	atom-atom bonds determined for the construction of the initial Hessian

RedCrdH	No	Hessian in the redundant coordinates representation
SCF	Yes	Controls the information about progress of the SCF procedure. Applies only if the print switch computation is on.
sdiis	No	Expansion coefficients applied by the DIIS procedure during the SCF.
sdiismat	No	Turns on sdiis(see above) *and* prints the *error vector* constructed by the DIIS routine (this is the commutator of the Fock matrix and the Density matrix). This is used to determine the DIIS expansion coefficients and to assess convergence.
SFO	Yes	General control of SFO-related output. If turned off, (almost) all such output is suppressed. If on, as is the case by default, such printing is controlled by the eprint subkey SFO.
Smat	No	Overlap matrix of BAS functions.
Smearq	No	Smear parameter - if and when applied - used in the determination of electronic occupation numbers for the MOs, with details of how it works out at every cycle of the SCF. For debugging purposes.
SpinOrbit	No	detailed information about how double-group symmetry representations are related to the single group representations
Tails	No	In each block of integration points (see Blocks) the evaluation of (Slater-type) exponential functions (basis, fit) is skipped when the function has become negligible for all points in that block due to the distance of those points from the atom where the function is centered. The relative savings due to this distance screening is printed at the first geometry cycle (use debug for printing at all cycles).
TechPar	Yes	Technical parameters such as maximum vector length in vectorized numerical integration loops, SCF and Geometry Optimization strategy parameters.
Timing	No	Print out of more timing info (in particular with respect to performance of the parallel version of ADF) than is provided by the standard Timing Statistics tables at the end of each output.
TimingDetail	No	Similar, but more details.
TimingTooMuchDetail	No	Similar, but even worse.
Workspace	No	Statistics of calls to the Workspace Manager (memory management).
Arguments for the keys print and noprint.		

For print switches that start with Frag.., Fit.., Freq..,Geostep..., Numint.., Repeat..., SCF..., TF..., see the key EPRINT below.

Debug

The key DEBUG is used to generate extensive output that is usually only relevant for debugging purposes. It operates exactly like the PRINT key but there is no converse: nodebug is not recognized; it would be irrelevant anyway because by default all debug print switches are off.

A list of the possible items for the DEBUG key is given below.

All items of the print list can also be used with the debug key. If they are not mentioned in table III, the meaning is the same as for the print key, but the corresponding output may be generated more often, for instance at every SCF cycle rather than at the last one only.

Item	*Explanation*
Basis	Construction of the orthonormal LOW basis from elementary (BAS) and fragment (FO) basis.
Core	Core Orthogonalization procedure

Ekin	Kinetic energy matrices. (compare the print switches EKIN)
Fit	Construction of the symmetry adapted fit functions
Fitint	Construction of integrals used in the Fit procedure.
Freq	Force matrices processed in the computation of frequencies: Cartesian and internal representation, before and after symmetrization, etc. (as far as applicable).
GeoStep	Geometry optimization procedure. All relevant items.
Hess	Complete eigensystem of the Hessian during geometry optimizations.
NumInt	Numerical integration. Very extensive output (including the coordinates and weights of all generated points).
Pmat	P-matrix (density matrix) during SCF and in the ETS analysis program in the BAS representation.
Rhofih	Computation of fit coefficients during the SCF.
SCF	Extensive output during the SCF procedure about many different items. See also EPRINT, subkey SCF.
SDIIS	All data concerning the DIIS as used during the SCF. See ERPRINT, subkey SDIIS.
TransitionField	The Transition State procedure to compute and analyze certain terms in the bonding energy. The distinct components, the involved transition field Fock matrices, etc.
Table III. Arguments for the print key DEBUG. All debug switches are by default off.	

Eprint

The key EPRINT is an extended version of the (no)print key, employed for print switches that require more specification than just off or on.

Contrary to what is the case for the keys print and noprint, the key EPRINT must occur only once in the input file; any subsequent occurrences are incorrect and ignored or lead to abort.

```
EPRINT
  subkey
  subkey
  ...
end
```

`subkey`

A subkey-type structure: it consists of a keyword followed by data, so that it functions as a simple (sub)key, *or* it is a keyword followed by a data *block* which must then end with the word subend.

The subkeys used in the eprint data block are called Eprint keys. A complete list of them is given below. All available eprint keys are discussed in the schemes below. The enclosing records eprint and end are omitted in these schemes.

EPRINT subkeys	*Subject*
AtomPop	Mulliken population analysis on a per-atom basis
BASPop	Mulliken population analysis on a per-bas-function basis
Eigval	One-electron orbital energies
Fit	Fit functions and fit coefficients
Frag	Building of the molecule from fragments.
FragPop	Mulliken population analysis on a per fragment basis
Freq	Intermediate results in the computation of frequencies (see debug: freq).
GeoStep	Geometry updates (Optimization, Transition State, ...)
NumInt	Numerical Integration

OrbPop	(Mulliken type) population analysis for individual MOs, both on a per-SFO basis and on a per-bas function basis. In a SpinOrbit calculation no SFO-type analysis is available (not yet implemented).
OrbPopEr	Energy Range (ER) in hartree units for the OrbPop subkey
Repeat	repetition of output in Geometry iterations (SCF, optimization, ...)
SCF	Self Consistent Field procedure
SFO	Information related to the Symmetrized Fragment Orbitals and the analysis (populations and MO coefficients) in this representation.
TF	Transition Field method for the evaluation and analysis of certain bonding energy terms.
Table IV. List of eprint subkeys.	

Eprint subkeys vs. Print switches

Several eprint subkeys are merely shortcuts for normal (no)print switches. All such simple subkeys are used in the following way:

```
EPRINT
  ESUBKEY argumentlist
END
```

`Esubkey`

One of the following eprint subkeys: Fit, Frag, GeoStep, NumInt, Repeat, SCF, sdiis, SFO, TF, Time.

`argumentlist`

A sequence of names, separated by delimiters. Each of these names will be concatenated with the esubkey and the combination will be stored as a normal print switch.
Example:
Frag rot, SFO
will be concatenated to fragrot and fragsfo and both will be stored as print switches. All such combinations can also be specified directly with the key PRINT. The example is therefore exactly equivalent with the input specification:
print FragRot, Fragsfo

If any of the names starts with the two characters no, the *remainder* of the name will be concatenated with the eprint, but now the result will be stored and treated as a noprint switch. Items that are on by default can in this way be turned off. Example:

```
EPRINT
  FRAG noRot Eig
END
```

This turns Rot *off* and Eig *on* for the eprint subkey Frag. Equivalent would be:

```
NOPRINT FragRot
Print FragEig
```

Follows a description of all simple EPrint subkeys:

Fit

The subkey fit controls output of how the elementary fit functions are combined into the symmetric (A1) fit functions. It controls also printing of the initial (start-up) and the final (SCF) fit coefficients.

```
EPRINT
  FIT list
END
```

`list`

A list of items, separated by blanks or commas. The following items are recognized: Charge, Coef, Comb.

`Charge`

The amount of electronic charge contained in the fit (start-up), total and per fragment.

`Coef`

The fit coefficients that give the expansion of the charge density in the elementary fit functions.

`Comb`

The construction of the totally symmetric (A1) fit function combinations from the elementary fit functions.

By default all options are off.

Frag

The subkey frag controls output of how the molecule is built up from its fragments.

```
EPRINT
  FRAG list
END
```

`list`

A list of items, separated by blanks or commas. The following items are recognized: Eig, Fit, Rot, SFO.

`Eig`

The expansion coefficients in elementary functions (bas) of the fragment Molecular Orbitals as they are on the fragment file.

`Rot`

The rotation (and translation) required to map the master fragment (i.e. the geometrical data on the fragment file) onto the actual fragment which is part of the current molecule.
N.B.: if eig and rot are both *on*, the rotated fragment orbitals are printed also.

`Fit`

The fit coefficients that describe the fitted charge density of the fragments after the rotation from the *master* fragment on file to the actual fragment. These are the molecular fit coefficients that are used (by default) to construct the total molecular start-up (fitted) charge density and hence the initial Coulomb and XC potential derived from it.

`SFO`

The Symmetry-adapted combinations of Fragment Orbitals that are used in the current calculation. This feature ensures that the definition of the SFOs is printed. This will happen anyway whenever the eprint subkey SFO itself is activated.
By default all options are off.

Remark: SFO analysis in a Spin-Orbit relativistic calculation is implemented only in the case there is one scalar relativistic fragment., which is the whole molecule.

Freq

Controls printing of Force matrices and a few more data that are intermediate results in the computation of frequencies after all coordinate displacements have been carried out.

```
EPRINT
  FREQ list
END
```

`list`

contains any of the items SymCoord, DMuRot, Hess.

`SymCoord`

print the Symmetry Coordinates, both as they are generated, and some related info in their processing later on. The Symmetry Coordinates are symmetry-adapted combinations of cartesian displacements, with the pure translations and rotations projected out.

`dmuRot`

info about generating Dipole-Derivative information in transformation between cartesian and internal coordinate representation as regards the rotational aspects.

`Hess`

processing of the completed matrix of force constants, symmetrization, transformation to other coordinates.

By default all options are off.

GeoStep

Controls output concerning the geometry update method, parameters, energy gradients, etc. It plays no role in a SinglePoint calculation.

```
EPRINT
  GEOSTEP list
END
```

`list`

A list of items, separated by blanks or commas. The following items are recognized: Energy, GradientTerms, Gradients, Upd.

`Energy`

summary of the (bonding) energy and its components as computed in the geometry update procedure.

`Gradients`

Energy gradients on the free variables. These may be all or some of the cartesian or the Z-matrix coordinates, depending on the case.

`GradientTerms`

The decomposition of the gradients in computed terms, as described in the thesis of L.Versluis [7].

`Upd`

parameters used and adapted in the geometry update procedure.

By default Gradients, Upd are on, the other items off.

NumInt

Output related to the numerical integration procedure: parameters, generated points, tests on the accuracy of the generated scheme, etc.

```
EPRINT
  NUMINT list
END
```

`list`

A list of items, separated by blanks or commas. The following items are recognized: All, Geo, Ovl, Par, Pnt, Res, Sym, Test.

`All`

includes all other options and prints in addition the coordinates and weights of all generated points. *This can be a lot of output!*

`Geo`

geometric data such as boundary planes around the molecule, as they are computed and used in the program section where the point grid is generated.

`Ovl`

numerically integrated are the auto-overlaps of symmetry-adapted combinations of elementary basis functions SBAS. The deviations from the analytically computed values is printed. The test option, see below, yields a summary of these data: the maximum error and the root-mean-square error.

`Par`

employed precision parameters, atomic spheres radii etc.

`Pnt`

the generated *numbers* of points in each of the subregions processed in the point-generating procedure.

`Res`

results as regards the total number of points, the sum-of-weights and the partitioning of the points in blocks (for segmented vectorization).

`Sym`

the symmetry operators that are computed directly from the coordinates (irrespective of the input Schönfliess symbol) and that are used to construct thec numerical integration grid in a symmetric fashion.

`Test`

a few external tests are performed after the grid has been generated, such as the numerical integration of the sum-of-fragment densities. See also the norms option.

By default Res and Test are on, the other options off.

OrbPop

Specifies that (Mulliken type) population analysis should be printed for individual MOs, both on a per-SFO basis and on a per-bas function basis. The format of the subkey is as follows:

```
EPRINT
  ORBPOP TOL=X Nocc Nunocc
  SUBEND
END
```

X is the threshold for the SFO coefficient value to include in the listing for the per-SFO analysis. Nocc is the number of the highest occupied and Nunocc is the number of the lowest unoccupied orbitals to analyze.

OrbPopER

Specifies the energy range for the MOs to which the OrbPop key applies. The default range is from -0.7 below the HOMO to 0.2 hartree above the LUMO. Usage:

```
EPRINT
  OrbPopER minEn maxEn
END
```

where minEn and maxEn are both in hartree, and have the defaults just specified. In order to get information on many more orbitals, simply specify a large negative value for minen and a large positive value to maxen.

Repeat

Control the repetition of output in Geometry iterations: optimization, computation of frequencies, transition state search.

```
EPRINT
  Repeat list
END
```

`list`

contains one or more of the following items: NumInt, SCF.

`NumInt`

Output from the numerical integration procedure, like parameters, numbers of points generated, test data is controlled by the *numint* subkey (see below). The *repeat* subkey controls whether the output is repeated for all geometries (if the flag is on) or only for the first (if the flag is off). Some concise info is produced (repeatedly) anyway if the print switch computation is on.

`SCF`

Controls similarly the SCF output, like population analysis and orbital eigenvalues. If the flag is on, these items are printed at the last SCF cycle in every geometry, otherwise only at the last (in case of an optimization, not in case of a Frequencies calculation).

By default both options are off.

SCF

Output during the SCF procedure.

```
EPRINT
  SCF list
END
```

`list`

is a list of items, separated by blanks or commas. The following items are recognized: Eigval, Eigvec, Err, Fmat, Keeporb, MOPop, Occ, Pmat, Pop, Start.

`Eigval`

Eigenvalues of the one-electron orbitals at the last SCF cycle. In a run with multiple SCF runs (Geometry Optimization,..) this printing occurs only for the last SCF procedure. See also the eigval subkey of EPRINT. (Use debug or the *repeat* subkey of EPRINT to get output on *all* cycles).

`Eigvec`

MO eigenvector coefficients in the BAS representation. Only printed on the last SCF cycle.

`Err`

SCF error data which are checked for convergence. By default this takes effect after cycle 25 of the SCF. If the key is set it takes effect at the first cycle. Optionally one may type ErrN,where n is an integer (written directly after Err without a blank in between), in which case the key takes effect at cycle n.

`Fmat`

Fock matrix in the low representation.

`Keeporb`

If the KeepOrbitals option is activated (see the key SCF), output is generated whenever this option actually results in a change of occupation numbers as regards the energy ordering.

`Occ`

concise output of SCF occupation numbers on last SCF cycle if no eigenvalues are printed (see: Eigval).

`moPop`

Mulliken populations in terms of the elementary basis functions (bas), per MO, for input-specified MOs (see the eprint subkey *orbpop*)

`Pmat`

Density matrix

`Pop`

General control of bas Mulliken populations. This supervises all printing (whether populations are printed or not) according to the eprint subkeys *atompop, fragpop, orbpop* (the latter only as regards the bas population analysis at the end of the SCF procedure).

`Start`

Data pertaining to the *first* SCF cycle (of the *first* SCF procedure, in case of an optimization; use *repeat* to get this for *all* SCFs).

By default Eigval, Keeporb, Occ, and Pop are on, the others off.

SFO

Information pertaining to the use of Symmetrized Fragment Orbitals (for analysis purposes).

```
EPRINT
  SFO list
END
```

`list`

A list of items, separated by blanks or commas. The following items are recognized: eig, eigcf, orbpop, grosspop, fragpop, ovl.

`Eig`

The MO coefficients in terms of the SFOs.

`Eigcf`

idem, but now also containing the coefficients pertaining to the CoreFunctions.

`OrbPop`

population analysis of individual orbitals. The orbitals analyzed are set with the eprint subkey *orbpop*.

`GrossPop`

Gross populations of the SFOs, split out in symmetry representations. GrossPop is automatically turned on when OrbPop is activated.

`FragPop`

Population analysis on a per-FragmentType basis. This analysis does in fact not depend on the SFOs (ie, the result does not depend on how the SFOs are defined), but the computation of these populations takes place in the SFO-analysis module, which is why it is controlled by the SFO print option. FragPop output is given per orbital when OrbPop is activated, per symmetry representation when GrossPop is activated, and as a sum-over-all-orbitals-in-all-irreps otherwise (if FragPop is active).

`Ovl`

Overlap matrix of the SFO basis, separately for each symmetry representation.

By default orbpop is on, the other options off. Note that before ADF2008.01 eig en ovl were on by default.

In a Spin-Orbit calculation the SFO analysis is not yet implemented completely.

Remark: the options eig, eigcf replace the pand revious (now disabled) simple print options eigsfo and eigsfo.

Note that the simple print key SFO controls whether or not the eprint subkey *sfo* is effective at all.

TransitionField

Part of the bonding energy is computed and analyzed by the so-called Transition State procedure [3, 110]. This has nothing to do with physical transition states, but is related to the Fock operator defined by an average charge density, where the average is taken of the initial (sum-of-orthogonalized-fragments) and the final (SCF) charge density. There is also an analogous term where the average is taken of the sum-of-

fragments and the sum-of-orthogonalized-fragments. Various terms, Fock operators and Density Matrices used in this approach may be printed. To avoid confusion with real Transition States (saddle points in the molecular Energy surface) the phrase TransitionField is used here.

```
EPRINT
  TF list
END
```

`List`

A list of items, separated by blanks or commas. The following items are recognized: Energy, Fmat, DiagFmat, FragPmat, DiagFragPmat, F*dPmat, DiagF*dPmat, OrbE.

`Energy`

Energy terms computed from the TransitionField.

`Fmat`

TransitionField Fock matrices.

`DiagFmat`

Idem, but only the diagonal elements.

`FragPmat`

The molecular P-matrix constructed from the sum-of-fragments.

`DiagFragPmat`

idem, but only the diagonal elements.

`F*dPmat`

The TransitionField energy term can be expressed as a Fock operator times the difference between two P-matrices (initial and final density).

`DiagF*dPmat`

only diagonal elements

`OrbE`

Orbital energies in the TransitionField.

By default all options are off.

Other Eprint subkeys

We discuss now the remaining eprint sub keys that are not simple shortcuts for print switches.

Orbital Energies

```
EPRINT
  Eigval noccup {nvirtual}
END
```

This specifies the *number* of one-electron orbitals for which in the SCF procedure energies and occupation numbers are printed whenever such data is output: the highest noccup occupied orbitals and the lowest nvirtual empty orbitals. Default values are noccup=10, nvirtual=10. If only one integer is specified it is taken as the noccup value and nvirtual is assumed to retain its standard value (10). Printing can be turned off completely with the eprint sub key SCF, see above.

Mulliken Population Analysis

All population subkeys of eprint refer to *Mulliken* type populations.

```
EPRINT
  ATOMPOP level
END
```

Populations accumulated per atom.

level must be none, gross or matrix. none completely suppresses printing of the populations; gross yields the gross populations; matrix produces the complete matrix of net and overlap populations. Default value: matrix.

```
EPRINT
  BASPop level
END
```

Populations are printed per elementary (bas) basis function. The level options are none, short, gross, matrix. none, gross and matrix are as for atompop.

short yields a summary of BAS gross populations accumulated per angular momentum (*l*) value and per atom.

Default value: gross.

```
EPRINT
  FragPop level
END
```

Completely similar to the atompop case, but now the populations per *fragment*. Of course in the case of single-atom fragments this is the same as atompop and only one of them is printed. Default: matrix.

For all three population keys atompop, fragpop and baspop, specification of a higher level implies that the lower-level data, which are in general summaries of the more detailed higher level options, are also printed.

Printing of any populations at the end of the SCF procedure is controlled with the eprint sub key *SCF* (pop).

Population Analysis per MO

A very detailed population analysis tool is available: the populations *per orbital* (MO). The printed values are independent of the occupation numbers of the MOs, so they are not populations in a strict sense. The actual populations are obtained by multiplying the results with the orbital occupations.

The analysis is given in terms of the SFOs and provides a very useful characterization of the MOs at the end of the calculation, after any geometry optimization has finished. This feature is now also available in a Spin-Orbit coupled relativistic calculation, in the case there is one scalar relativistic fragment, which is the whole molecule.

The same analysis is optionally (see EPRINT subkey *SCF*, option mopop also provided in terms of the elementary basis functions (bas).

```
EPRINT
  OrbPop {noccup {nvirtual}} {tol=tol}
    subspecies orbitals
    subspecies orbitals
    ...
  subend
END
```

`noccup`

Determines how many of the highest occupied orbitals are analyzed in each irrep. Default noccup=10.

`nvirtual`

Determines in similar fashion how many of the lowest virtual orbitals are analyzed in each irrep. Default nvirtual=4.

`tol`

Tolerance parameter. Output of SFO contributions smaller than this tolerance may be suppressed. Default: 1e-2.

`subspecies`

One of the subspecies of the molecular symmetry group. Can not be used (yet) in a Spin-Orbit coupled calculation.

`orbitals`

A list of integers denoting the valence orbitals (in energy ordering) in this subspecies that you want to analyze. This overrules the noccup,nvirtual specification for that symmetry representation. In an unrestricted calculation two sequences of integers must be supplied, separated by a double slash (//).

Any subset of the subspecies can be specified; it is not necessary to use all of them. No subspecies must occur more than once in the data block. This can not be used in a Spin-Orbit coupled equation (yet).

A total SFO gross populations analysis (from a summation over the occupied MOs) and an SFO population analysis per fragment type are preformed unless *all* MO SFO-populations are suppressed.

Reduction of output

One of the strong points of ADF is the analysis in terms of fragments and fragment orbitals (SFOs) that the program provides. This aspect causes a lot of output to be produced, in particular as regards information that pertains to the SFOs.

Furthermore, during the SCF and, if applicable, geometry optimizations, quite a bit of output is produced that has relevance merely to check progress of the computation and to understand the causes for failure when such might happen.

If you dislike the standard amount of output you may benefit from the following suggestions:

If you are not interested in info about progress of the computation:

```
NOPRINT Computation
```

If you'd like to suppress only the SCF-related part of the computational report and make the GeometryUpdates related part more concise:

```
NOPRINT SCF, GEO
```

(Keep computation on, so you get at least some info about the GeometryUpdates)

If you don't want to see any SFO stuff:

```
NOPRINT SFO
```

To keep the SFO *definitions* (in an early part of output) but suppress the SFO-mo coefficients and the SFO overlap matrix:

```
EPRINT
 SFO noeig, noovl
END
```

Note: the SFO-overlap matrix is relevant only when you have the SFO-MO coefficients: the overlap info is needed then to interpret the bonding/anti-bonding nature of the various SFO components in an MO.

If you are not interested in the SFO *populations*:

```
EPRINT
 SFO noorbpop
END
```

2.9 Accuracy and Efficiency

See also
GUI manual: accuracy
Examples: accuracy

Precision and Self-Consistency

The precision of a calculation is determined by

- The function sets (basis sets, levels of frozen core approximation, and fit sets for the computation of the Coulomb potential)
- Numerical integration settings in real space
- The accuracy of the density fitting procedure
- Convergence criteria (for the SCF procedure and the geometry optimization)
- A few more items that are rather technical and usually irrelevant (these are not discussed here).

The fragments you attach determine, through the fragment files, the function sets. Since each fragment traces back to one or more Create runs, the employed data base files in the Create runs determine the finally employed function sets.

For convergence of the geometry optimization see the key GEOMETRY.

In this part we examine numerical integration, density fitting and the SCF procedure.

With the key NUMERICALQUALITY one can set the density fitting quality (ZlmFit) and the numerical integration quality (BeckeGrid) simultaneously

```
NUMERICALQUALITY {basic|normal|good|verygood|excellent}
```

Numerical Integration

Becke grid for numerical integration

Many integrals in ADF are evaluated by numerical integration: Fock matrix elements, several terms in the (bonding) energy, gradients in geometry optimization, and so on.

The default numerical integration method in ADF2013 is a refined version of the fuzzy cells integration scheme developed by Becke [361]. This implementation in ADF is described in Ref. [375].

Note that in ADF2012 and previous versions the default integration scheme was the cellular Voronoi quadrature scheme, implemented by te Velde and Baerends. Thanks to a smoother behavior of the relative integration error as a function of the nuclear coordinates, the Becke grid is better suited for geometry optimization and TS search compared to the Voronoi scheme.

The default quality of the Becke grid is normal. It can be changed with the block key BECKEGRID:

```
BECKEGRID
 Quality {basic|normal|good|verygood|excellent}
 AtomDepQuality
    Ia1 {basic|normal|good|verygood|excellent}
    Ia2 {basic|normal|good|verygood|excellent}
    ...
 SubEnd
 {qpnear qpnear}
End
```

`Quality`

For a description of the various "qualities" and the associated numerical accuracy, see Ref. [375]

`AtomDepQuality`

One can define a different grid quality for each atom, with input numbers Ia1, Ia2, etc. If an atom is not present in the AtomDepQuality seciton, the quality defined in the Quality key will be used.

`qpnear`

Only relevant if you have specified point charges in the input file. ADF generates grids only about those point charges that are close to any real atoms. The criterion, input with the qpnear subkey, is the closest distance between the point charge at hand and any real atom. Default 4.0 Angstrom. Any input value is interpreted in the unit-of-length specified with the Units key.

Notes:

- If either the (block) key INTEGRATION or the key NOBECKEGRID are used in the input, the Voronoi grid is used.

  ```
  NOBECKEGRID
  ```

- A Becke grid of normal quality is roughly equivalent (in both absolute accuracy and computation time) to INTEGRATION 4 (Voronoi scheme), and a Becke grid of good quality is roughly equivalent to INTEGRATION 6 (Voronoi scheme).
- The Becke grid is not very well suited to calculate Voronoi deformation density (VDD) charges. For accurate calculation of VDD charges the Voronoi integration scheme is recommended.

- A new atomic partition function, called MPV, is used in ADF2014. To use the same partition function that was used in ADF2013, include the following key in the block key BECKEGRID:

```
PartitionFunction YukawaLike
```

Voronoi grid

A sophisticated numerical integration procedure is the Voronoi integration method [104, 105]. It requires only one input parameter which determines the precision of numerical integrals and derives from that the number of integration points. Starting from ADF2013 this method is no longer the default sheme for integration. The so called Becke grid is the default, see the key BECKEGRID. If the key INTEGRATION is used, the Voronoi scheme will be used.

```
INTEGRATION accint
```

`accint`

A positive real number: the numerical integration scheme generates points and weights such that a large number of representative test integrals are evaluated with an accuracy of accint significant digits. The default for accint depends on the runtype: 4.0 for Single Point runs and simple Geometry Optimizations, including Linear Transits; 5.0 for Transition State searches; 6.0 for the computation of Frequencies; 10.0 in Create runs.

The *number* of integration points varies strongly with accint, and this determines to a large extent the computational effort. Decreasing accint from 4.0 to 3.0 for instance roughly halves the number of points (this depends somewhat on the molecule).

The defaults should yield good precision for the very large majority of applications. Lower values (3.0 or even 2.0) can be used if precision is not crucial and the purpose is to get an impression. We recommend that you experiment for yourself to get a feel for how results may vary in quality and computing time.

The default in Create mode is very large: 10.0. This is computationally no problem thanks to the simplicity of the single atom case, in particular due to the high symmetry. There is no reason to override the default integration settings when creating basic atoms.

We've now only explained the normal, simple application of the Integration key, which we hope and expect is adequate for all your computations. Next additional details will be discussed. The distribution of points over space is internally regulated by quite a few parameters. Each of these parameters can be controlled in input. By default they depend on one another, and all of them depend on the main parameter accint. Advanced users may wish to experiment and override the default relations between the parameters.

You may also have rather non-standard applications where the default relations are less adequate. A thorough understanding of the underlying method is required to make a sensible choice for all parameters [105, 109].

The key INTEGRATION has been introduced in its simple form before.

```
INTEGRATION accint
```

accint is a real number. The key is used as a simple key here.

Alternatively you can use it as a block key. This is activated if you give no argument. In the data block you specify which of several integration methods you want to use, and you give values for the involved parameters. Consult the literature for detailed information about the various schemes.

```
INTEGRATION
  data
```

```
  data
  ...
end
```

The block form is used to override default relations between various parameters that are applied in the generation of the integration grid in the polyhedron method [105]. All these parameters are accessible with subkeys in the data block of Integration. Most of the subkeys are simple keys with one single value as argument; a few subkeys are block-type (sub) keys themselves and hence require the usual format of a data block closed by subend.

`accint`

The main precision parameter
Its value defines the number of significant digits by which an internal set of standard integrals must be evaluated. The number and distribution of integration points is tuned accordingly. For normal applications this should yield a nearly optimal (given the underlying method) generation of points and weights. The default depends on the run type.

`accsph`

The polyhedron method of generating integration points partitions space in atomic polyhedrons, partitioned in pyramids with their tops at the atom in the center of the polyhedron. A core like atomic sphere is constructed around the atom; this truncates the tops of the pyramids. accsph specifies the test precision for the generation of points within the spheres. By default accsph=accint.

`accpyr`

Similarly this subkey sets the test level for the parts of the pyramids outside the atomic sphere. Default: accpyr=accint.

`accpyu, accpyv, accpyw`

The truncated pyramids are mathematically transformed into unit cubes. A product Gauss integration formula is applied to the cubes, with three (test precision) parameters for the three dimensions. Accpyw controls the direction that is essentially the radial integration from the surface of the atomic sphere to the base of the pyramid. The other two control the orthogonal directions (angular). By default all three equal accpyr.

`accout`

The region of space further away from the atoms, outside the polyhedrons, has its own precision parameter. By default accout=accint.

`nouter`

This outer region is treated by a product formula: outwards times parallel. The latter involves two dimensions: the surface of the molecule say. The outward integration is performed with Gauss-Legendre quadrature, in a few separate steps. The lengths of the steps are not equal, they increase by constant factors. The total length is fixed. The number of steps is controlled with this subkey; default: 2.

`outrad`

The parameter that defines the number of Gauss-Legendre integration points for each outward step. The precise relation between the actual number of points and this subkey, and the default relation between outrad and accout can be found in the implementation.

`outpar`

Similarly the integration in the directions parallel to the surface of the atomic system is controlled by a parameter. See the implementation for details.

`dishul`

Sets the distance between the outermost nuclei of the molecule and the boundary planes that define the boundary between the polyhedrons and the outer region. By default dishul=2.3*R, where Ris the radius of the largest atomic sphere in the molecule.

`frange`

The outward range of the *outer region*: integration is not performed to infinity but to a distance frange from the outermost atoms, where all functions can be assumed to be essentially zero. By default frange is derived both from accint, the general precision parameter, and from the present chemical elements: heavier atoms have longer-range functions than hydrogen say. The precise relations can be found in the implementation.

`linrot`

This parameter is significant only for symmetries with an axis of infinite rotational symmetry: Cand D
It is the highest rotational quantum number around this axis that occurs among the integrands. This depends on the employed basis functions and fit functions. By default the program finds this out for itself.

`qpnear`

If you specify point charges in the input file, there are two considerations implied for the numerical integration grid.
First, since the point charges create a Coulomb singularity. The integrands (of for instance the basis function products against the Coulomb potential) can only be evaluated with high precision if the grid around the point charges has spherical symmetry and uses local spherical coordinates, exactly as is done for the atomic nuclei. Second, the point charges do not carry fit or basis functions, hence they play only a role in the more diffuse tails of the actual functions involved in integrals. Therefore, a relative low precision of the integral part close to the point charge may have little effect on the total integration accuracy.
Since additional 'spherical centers' with their own surrounding grids increase the total number of points significantly, typically a few thousands *per Coulomb center*, this may result in high computational effort. Therefore, the program generates spherical grids only about those point charges that are close to the other atoms. The criterion, input with the qpnear subkey, is the closest distance between the point charge at hand and any real atom. Default 4.0 Angstrom. Any input value is interpreted in the unit-of-length specified with the Units key.

Next come the subkeys that require a list of data. The subkey must be placed on one line, the data on the next. This somewhat peculiar structure suggests that the subkeys are block keys; however their data blocks have no end code (subend) as for normal block type subkeys.

The list of data for such a subkey contains one value for each atom type. The data must be in the order in which the atom types were defined under atoms, implicitly or explicitly: remember that atoms belonging to different fragment types automatically have different atom types, even if their atom type *names* have been specified as identical under atoms.

`rspher`

gives the radii of the atomic spheres, one value for each atom type. By default, the radii are derived from the chemical element (heavier atoms get larger spheres) and from the environment: the sphere must not be too large for the atomic cell (polyhedron).

`linteg`

The maximum angular momentum quantum number of integrands centered on an atom of that type (one value for each atom type). This depends on the basis functions and on the fit functions. By default the program checks the function sets and sets the linteg values accordingly. This subkey is applied for the generation of grid points in the atomic spheres.

Items that relate to geometric lengths (dishul, frange, rspher) must be given in bohr (=atomic units), irrespective of the unit of length defined with units.

Atomic radial grid

For each atom the charge densities and the coulomb potentials of frozen core and valence electrons are computed in a radial grid and stored on TAPE21. The values in the points of the molecular numerical integration grid are then evaluated by interpolation from the table of radial values.

The radial grid consists of a sequence of r-values, defined by a smallest value, a constant multiplication factor to obtain each successive r-value, and the total number of points. Equivalently it can be characterized by the smallest r-value, the largest r-value, and the number of points; from these data the program computes then the constant multiplication factor. The characteristics are set with

```
RADIALCOREGRID {nrad=points} {rmin=rmin} {rmax=rmax}
```

`points`

The number of radial grid points; default: 5000.

`rmin`

The shortest distance used in the radial grid; default 1e-6 Angstrom

`rmax`

The largest distance in the radial grid; default: 100 Angstrom.

rmin and rmax, when specified, are interpreted as specified in units of length defined by units.

The keyword name radialcoregrid has historical reasons: in earlier releases the radial grid was used only for the frozen core density and potential.

SCF

The SCF procedure is regulated with keys that set the maximum number of iterations, the convergence criterion, and various items that control the iterative update method. Molecules may display wildly different SCF-iteration behavior, ranging from easy and rapid convergence to troublesome oscillations. We expect that the default settings take care of most cases, but one should realize that this is a difficult and tricky subject. The user has a few (main) options to adapt the procedure to the situation at hand: simple damping or the DIIS procedure (Direct Inversion in the Iterative Subspace). Either of them can be combined with Level-Shifting.

Please refer to the SCF troubleshooting section in case of SCF convergence problems.

At each cycle the density is computed as a sum of occupied orbitals squared; the new density defines the potential from which the orbitals are re-computed. The cycle is repeated until convergence is reached. To speed-up convergence and to avoid non-convergent oscillatory behavior the values at the next iteration are constructed as a mixture of the computed new data and those used at the cycles before. This may involve only the previouss cycle and is then called *damping*. Alternatively the DIIS procedure can be invoked, which is a generalization of damping to include more previouss iterations.

In ADF2009.01 two methods, called ARH and Energy-DIIS, have been implemented that can solve problematic cases for SCF convergence. These methods will be discussed separately, after the main options of the SCF key. Both methods require the total energy to be calculated at each step, which makes them much more expensive compared to the standard SCF procedure, and not applicable in all cases. Therefore, these methods should only be used when the standard SCF procedure fails.

In ADF2010.01 ADIIS was implemented, which performs similar to the Energy-DIIS scheme but it does not require calculation of the total energy. Therefore it is much faster than E-DIIS.

In ADF2012, LISTi has been implemented.

Main options

Subkeys in the data block of the master key SCF control the aspects mentioned above. Each subkey is optional. Omission means the application of default values. Omission of the SCF key altogether implies defaults for all subkeys.

```
SCF
 Iterations Niter
 Converge SCFcnv { sconv2 }
 Mixing mix
 Diis {N=n} {OK=ok} {CX=cx} {CXX=cxx} {BFAC=bfac} {cyc=cyc}
 Lshift vshift {Err=shift_err} {Cyc=shift_cyc}
End
```

`Iterations Niter`

`Niter`

The maximum number of SCF cycles allowed. In case of Geometry Optimizations it applies separately to each of the SCF procedures that are executed. Default is 300. The program executes at least one cycle, even if you specify a non-positive number.

`Converge SCFcnv { sconv2 }`

`SCFcnv`

The criterion to stop the SCF updates. The tested error is the commutator of the Fock matrix and the P-matrix (=density matrix in the representation of the basis functions) from which the F-matrix was obtained. This commutator is zero when absolute self-consistency is reached. Convergence is considered reached when the maximum element falls below SCFcnv and the norm of the matrix below 10*SCFcnv. The default is 1e-6 (in Create mode: 1e-8).

`sconv2`

A second criterion which plays a role when the SCF procedure has difficulty converging.
When in any SCF procedure the currently applicable criterion does not seem to be achievable, the program stops the SCF. When the secondary criterion (sconv2) has been met, only a warning is issued and the program continues normally. If the secondary criterion was not met either, the program terminates any further geometry optimizations, frequency steps, etc. You can prevent the program from terminating in such a case with the key ALLOW.
The default for sconv2 is 1e-3.

`Mixing mix`

`mix`

The relative weight of the new potential, as computed from the occupied orbitals, to be mixed with the potential that was used in the previous cycle, to define the potential for the next. Mixing is used only if, and as long as, the DIIS procedure (see below) is not operational. Default: 0.2.
For problematic systems that require strong damping, one should *decrease* the mix-parameter.

`Diis {N=n} {OK=ok} {CX=cx} {CXX=cxx} {BFAC=bfac} {cyc=cyc}`

The DIIS subkey specification(s) can be given to control the DIIS procedure. Each of these specifications is optional. Simple damping will be used during the first few cycles, until the DIIS procedure becomes operational. Two conditions must be satisfied for this: 1) at least two iterations must have been done anyway (to build up sufficient information for the DIIS to work at all) and 2) the error must be small enough; see however the cyc option below.
There have been claims in the literature that the DIIS should not be used until fair convergence has been reached. Our experience thus far does not indicate that this should be taken too seriously, except in special situations. To allow the user complete control, the start-up criteria can be set in input.

`N`

The number of expansion vectors used in the DIIS. The number of previous cycles taken into the linear combination is then n-1 (the new computed potential is also involved in the linear combination.
Default n=10. An input value smaller than 2 disables the DIIS. Note that this number applies not only to Pulay DIIS scheme but also to other DIIS-like methods, such as A-DIIS, Energy-DIIS, and LISTi.

`OK`

The DIIS starting criterion. The DIIS procedure is not invoked until a) the maximum commutator element is smaller than OK (default: 0.5) or *b*) a certain number of SCF cycles has been executed.

`Cyc`

The SCF cycle no. at which the DIIS will start irrespective of the OK value above. Default: 5.

`Cx`

An upper bound on linear combination coefficients as applied in the DIIS. As soon as any coefficient exceeds cx, all information about older cycles but the last two is discarded and the DIIS starts again to accumulate info from the current cycle on. The computed linear combination, with the large coefficient(s), is used for the next iteration, however. Default=5.0

`Cxx`

A second upper bound on the coefficients (should in principle be bigger than cx). When a coefficient exceeds cxx, the computed linear combination is not used for the next cycle, but simple damping is applied.

`Bfac`

A factor to bias the DIIS combination vector in favor of the new computed potential. Default=0 (no bias). A sensible alternative value is 0.2

`Lshift vshift {Err=shift_err} {Cyc=shift_cyc}`

`VShift`

The level shifting parameter. The diagonal elements of the Fock matrix, in the representation of the orbitals of the previous iteration, are raised by vshift hartree energy units for the virtual orbitals. This may help to solve convergence problems when during the SCF iterations charge is sloshing

back and forth between different orbitals that are close in energy and all located around the Fermi level. Level shifting is not supported in the case of Spin-Orbit coupling. **At the moment properties that use virtuals, like excitation energies, response properties, NMR calculations, will give incorrect results if level shifting is applied.**

`Shift_err`

Specifies that level shifting will be turned off by the program as soon as the SCF error drops below a threshold; default value: 0. (Note that the default value has changed in ADF2004.01, previous value: 1e-2).

`Shift_cyc`

Specifies that level shifting is not turned on before the given SCF cycle number (for the start-up geometry); default value: 1.

Note1: very strong damping, i.e. a very small value of mix such as 1e-3, may not combine very well with the DIIS procedure. The reason is that with strong damping successive SCF cycles tend to be very similar and the vectors building up the DIIS space become linearly dependent. We recommend in difficult cases either to use a not too strong damping (mix=0.03) *or* to use strong damping while the DIIS is disabled (by setting n=0 for the DIIS subkey and include the NOADIIS subkey of SCF) during a limited number of SCF iterations, and then *restart* with DIIS activated and less stringent damping.

Note2: Another feature, *electron smearing*, may be used to overcome convergence difficulties. The idea is to distribute electron occupations fractionally over a few states around the Fermi level, by a pseudo-thermal distribution function. This aspect is controlled with the Smear option to the Occupations key. One should be aware that the applied distribution of occupations is not really an approximation to the finite-temperature case. In fact, the results are unphysical and one should not use the results as a meaningful outcome. The smearing trick is *only* to be used to overcome convergence difficulties. Having reached convergence with it, one should typically do a follow-up restart calculation without smearing, using the converged outcomes to hopefully get the thing to converge properly. A typical 'allowed' application is the usage of smearing during geometry optimizations, because the intermediate geometries are not relevant anyway and only a step towards the final results. By default, the program does *not* apply any smearing unless during a geometry optimization. See the Occupations key for more details.

Energy-DIIS

```
SCF
 EDIIS
 ...
End
```

IN ADF2009 Energy-DIIS is implemented following the paper by Kudin, Scuseria, and Cances [289]. The method is invoked by specifying an EDIIS keyword in the SCF block. Please note that similar to ARH and unlike the standard SCF procedure in ADF this method requires energy evaluation at each SCF cycle, which makes it significantly slower compared to energy-free SCF. You might need a higher integration accuracy to get an accurate total energy. The same restrictions apply as for the key TOTALENERGY. The EDIIS method will start at the 2nd SCF cycle, and the size of the DIIS space will be the same as for the normal DIIS. This subkey EDIIS can be used in addition to the other subkeys of the block key SCF.

ADIIS

In ADF2010 ADIIS is implemented following the paper by Hu and Wang [291]. The method is invoked by specifying an ADIIS keyword in the SCF block. According to some test, ADIIS performs similar to the

Energy-DIIS scheme but it does not require calculation of the total energy. Therefore it is much faster than E-DIIS.

```
SCF
 ADIIS {THRESH1=a1 THRESH2=a2}
 ...
End
```

a1, a2

Here, *a1* and *a2* (*a1* > *a2*) correspond to values of the maximum element of the [F,P] commutator matrix, ErrMax. If ErrMax ≥ *a1*, only A-DIIS coefficients are used to determine the next Fock matrix. If ErrMax < *a2* then only SDIIS coefficients are used. For ErrMax between *a2* and *a1* the total DIIS coefficients are calculated from SDIIS and A-DIIS values weighted proportionally according to the ErrMax value. Thus, the weight of A-DIIS coefficients decreases with the ErrMax value. The default values for a1 and a2 are 0.01 and 0.0001, respectively.

ADIIS is automatically switched on when there is no SCF convergence after 15 iterations. To disable this behavior, specify NoADIIS in the SCF block.

Note: A-DIIS is not compatible with enforced non-aufbau electronic configurations it should be disabled in such a case. A non-aufbau electronic configuration may be enforced using a block form of the Occupations key, but it may also result from the KeepOrbitals (a.k.a. orbital tracking) feature. In both cases A-DIIS should not be used.

LISTi

The LISTi method has been implemented in ADF2012 following the paper by Y.K. Chen and Y.A. Wang [432]. This method is invoked by specifying a single LISTi keyword in the SCF block.

```
SCF
 LISTi
 ...
End
```

Even though the keyword does not take any arguments, the number of vectors to store specified in the DIIS keyword also applies to LISTi. This number is a very important parameter and it is worthwhile increasing or decreasing it in case of SCF convergence problems. A word of caution: do not just blindly increase the number for every system. Testing showed that a large number breaks convergence for some, mainly small, systems.

Our tests show that LISTi performance is similar to A-DIIS although sometimes it may require fewer or more steps to converge. A nice feature of LISTi is that the algorithm used in it scales better in parallel because it does not require evaluation of the [F,P] commutator.

Augmented Roothaan-Hall (ARH)

ARH Introduction

The Augmented Roothaan-Hall method has been developed by T. Helgaker and coworkers and is extensively discussed in Ref. [271]. The basic idea of the method is that the density matrix is optimized directly to minimize total energy. At each step, the new density matrix is parametrized in terms of matrix exponent:

$P_{new} = \exp(-X)\, P_{old} \exp(X)$,

here, X is an anti-symmetric step matrix subject to the following conditions:

$X = \text{argmin}\{E(P(X))\}$ - X minimizes the energy

$|X| < h$ - length of X is smaller than or equal to some trust radius

The optimal X is found using a Conjugate Gradient method, possibly with pre-conditioning. The trust radius is updated based on how well the energy change is predicted.

ARHOPTIONS Input

It is possible to specify ARH options without turning on the method itself unconditionally. The ARHOPTIONS subkey has the same arguments as the ARH subkey (see below). It should be used in combination with the ALLOW ARH keyword. The ARH procedure will then be invoked automatically if SCF has trouble converging.

```
SCF
  ARHOPTIONS ....
  ...
End
SYMMETRY NOSYM
ALLOW ARH
```

ARH Input

The ARH procedure is invoked using an ARH keyword in the SCF input block. This subkey ARH can be used in addition to the other subkeys of the block key SCF.

```
SCF
  ARH {CONV=conv} {ITER=iter} {NSAVED=nsaved} {START=start}
      {FINAL} ...
  ...
End
SYMMETRY NOSYM
```

All parameters in the ARH keyword are optional. The following arguments determine the main parameters of the ARH procedure.

`CONV=conv`

ARH convergence criterion. When the RMS gradient and its maximum components are both lower than the criterion, the ARH procedure will be considered converged. The default value is 10^{-4}.

`ITER=iter`

Maximum number of ARH iteration to perform. Please note that in difficult cases a huge number of iterations may be required for complete SCF convergence. The default value is 500.

`FINAL`

Determines whether SCF is continued after ARH has completed. If this option is set, one Fock matrix diagonalization will be performed to get orbitals and the SCF procedure will be halted. By default this option is OFF.

`START=start`

Sets the SCF cycle number on which the ARH method is invoked. The default value is 2. Using a larger value may provide a better starting guess for the ARH minimization.

`NSAVED=nsaved`

Sets the number of saved density and Fock matrices used for augmentation of the electronic Hessian. The default value is 8. A larger nsaved value should be used in difficult cases when the number of orbitals very close to the Fermi level is large.

The default minimization method is Untransformed Pre-conditioned Conjugate Gradient. The following two parameters may be used to change this.

`NOPRECOND`

Disables pre-conditioning during the CG minimization. This option should not be used if atoms heavier than the second-row elements are present.

`TRANSPCG`

Specifying this option will enable the use of the Transformed Pre-conditioned CG method, which may result in better SCF convergence in some cases.

At each SCF step, the procedures begins by performing usual CG minimization keeping track of the total step length. If at some micro-iteration the step length exceeds the trust radius, the procedure switches to trust-radius optimization in the reduced space, which, in turn, is halted as soon as the level-shift parameter *mu* has converged. The final step is then calculated as a Newton step in the reduced space of all the trial vectors generated during CG minimization. The following options may be used to modify this behavior.

`NOSWITCHING`

Setting this option turns OFF the switching from the normal CG to a trust-radius minimization in reduced space. Using this option helps to reduce the total number of SCF cycles is some cases.

`SHIFTED`

Setting this option will turn ON the trust-radius optimization from the first micro- iteration.

`CGITER=cgiter`

Sets the maximum number of micro-iterations.

The next two options determine the trust radius.

`TRUSTR=trustr`

Initial value for the trust radius. Default: 0.5

`MAXTRUSTR=maxtrustr`

The maximum trust radius value. This is set to 0.5 by default and should never be changed.

ARH Notes and Recommendations

Restriction: The method currently works for symmetry NOSYM calculations only. The NOSYM requirement comes from the fact that during direct optimization of the density matrix it may have a symmetry lower than that of the molecule.

The method requires the total energy to be calculated at each step, which makes it much more expensive compared to the standard SCF procedure that does not need or use the energy. Therefore, the method

should only be used when the standard SCF procedure fails. Another complication caused by the use of the total energy is that somewhat higher integration accuracy may be required to get stable SCF convergence, and that the method may not be applicable in all cases. It is also recommended to use the ADDDIFFUSEFIT keyword to increase accuracy of the total energy and, thus, improve convergence. Please refer to the TOTALENERGY keyword for more information.

Scalable SCF

The new (still experimental) ScalableSCF module has been added to ADF. This is an almost complete rewrite of the SCF procedure using distributed matrices. It is supposed to scale much better in parallel compared to the original SCF and it also writes less data to disk during SCF.

The ScalableSCF can be turned on by including ScalableSCF in the key SCF.

```
SCF
 ...
 ScalableSCF
End
```

ScalableSCF uses the A-DIIS+SDIIS convergence acceleration scheme automatically but the LISTi is also implemented.

Currently, the following features are not yet supported by ScalableSCF:

- Spin-orbit in combination with Hartree-Fock exchange, including hybrid functionals such as B3LYP;
- spin-unrestricted spin-orbit calculations;
- SOPERT, SOMCD, ZFS, STCONTRIB, RESTOCC, CONSTRUCTPOT keywords;
- some types of FDE calculations such as FDESubtract;
- some types of SCF convergence aids are not implemented (smearq, steep, vshift, E-DIIS).

In all cases listed above ScalableSCF will be automatically disabled and the calculation will proceed using the old SCF.

Density fitting

Zlm Fit: density fitting with radial spline functions and real spherical harmonics

Note: In ADF2013 and previous versions, a different density-fitting scheme (pair-fit) was used. Include the key STOFIT if you want to use the old fitting scheme.

The basic ideas behind the so-called Zlm Fit can be described as follows. The total electron-density is split into atomic densities (in a similar way as the volume is partitioned for the Becke grid). These atomic densities are then approximated by a combination of radial spline functions and real spherical harmonics (Zlm). The implementation in ADF is described in Ref. [379]. The algorithm that is used in ADF is related to the procedures proposed by Becke [380] and Delley [360]).

The Zlm Fit scheme (which is the default fitting scheme in ADF2014) offers certain advantages compared to the old pair-fit method, especially the possibility to calculate the Coulomb potential to very high precision.

```
ZLMFIT
 Quality {basic|normal|good|verygood|excellent}
 AtomDepQuality
    Ia1 {basic|normal|good|verygood|excellent}
    Ia2 {basic|normal|good|verygood|excellent}
```

```
    ...
  SubEnd
End
```

`Quality`

The default quality of the Zlm Fit is normal. It can be changed with the subkey Quality.

`AtomDepQuality`

One can define a different Zlm Fit quality for atoms, with input numbers Ia1, Ia2, etc.

The Zlm Fit method can be used for most features of the ADF program. However, it is not implemented for the calculation of Hartree-Fock exchange integrals.

Pair fit: symmetric density fit

The non-default density fitting procedure in ADF, called pair fit method, is carried out separately for each pair of atoms. To use it one needs to include the keyword STOFIT.

```
STOFIT
```

The implemented approach has several advantages in efficiency but it has a drawback in that it necessitates the use of all available fit functions rather than only the symmetric combinations although the final result of course needs only a symmetric fit because the total density is a symmetric (A1) function. For atoms far apart the density fitting is performed with only symmetric functions. Given the implemented algorithm this entails an approximation which can be tuned:

```
A1FIT atomicseparation
```

`atomicseparation`

is the threshold distance between atoms, *in Angstrom*. The symmetric fit approximation is applied only for atoms farther apart. Default is 10.0 Angstrom

Pair fit: fit integrals

```
STOFIT
```

For the computation of the Coulomb potential with the pair fit method the program uses a large number of so-called *fit integrals*: the overlap integrals of a fit function with a *product* of two basis functions, where at least two of the involved three functions are centered on the same atom. In fact these are ordinary overlap integrals of STOs because the fit and basis functions are all STOs and a product of STOs on a center is itself also an STO. To use this STO fitting method, which was previously the default, use the key STOFIT in the input of adf (also include it in the create mode of an atom, if that is explicitly used).

Obviously, when the two involved atoms are far enough apart, such overlap integrals become negligibly small. All fit integrals are ignored (and not computed) that are smaller - according to a rough but reasonable estimate - than a preset threshold.

The value of this treshold can be set via input, using the subkey CUTOFF_FIT of the LINEARSCALING block key word.

True density in XC potential

For the computation of the exchange-correlation potential (XC-potential) the program uses as default the fitted density. This is an approximation. For the XC potential the true density can be used if one includes the keyword `EXACTDENSITY`:

```
EXACTDENSITY
```

Using the EXACTDENSITY keyword makes the calculation more time-consuming but more accurate in the following cases:

- calculations that require accurate description of virtual orbitals, such as most of the TDDFT;
- when studying systems where weak interaction, such Van der Waals forces and hydrogen bonds, are important. For example, EXACTDENSITY should be switched on when performing geometry optimization of DNA pairs.

Dependency (basis set, fit set)

Conceivably the sizes of basis and/or fit sets may be so large that the function sets become almost linearly dependent. Numerical problems arise when this happens and results get seriously affected (a strong indication that something is wrong is if the core orbital energies are shifted significantly from their values in normal basis sets). Although for the fit set a few (incomplete) tests are carried out, the program will generally not check such aspects and carry on without noticing that results may be unreliable.

A new feature has been implemented to take care of this. For reasons of compatibility with previous versions and also because our experience with it is limited so far, we have chosen to make application of it not the default.

You have to activate it explicitly. Our experience so far suggests that real problems only arise in case of large basis sets with very diffuse functions (i.e.: not with the normal basis sets provided in the standard package).

Use of the key DEPENDENCY turns internal checks on and invokes countermeasures by the program when the situation is suspect. A few technical (threshold-type) parameters can be set as well, but this is not necessary, assuming that the defaults are adequate.

```
DEPENDENCY {bas=tolbas} {eig=BigEig} {fit=tolfit}
```

`tolbas`

A criterion applied to the overlap matrix of unoccupied normalized SFOs. Eigenvectors corresponding to smaller eigenvalues are eliminated from the valence space. Default value: 1e-4. Note: if you choose a very coarse value, you'll remove too many degrees of freedom in the basis set, while if you choose it too strict, the numerical problems may not be countered adequately.

`BigEig`

Merely a technical parameter. When the DEPENDENCY key is activated, any rejected basis functions (i.e.: linear combinations that correspond with small eigenvalues in the virtual SFOs overlap matrix) are normally processed until diagonalization of the Fock matrix takes place. At that point, all matrix elements corresponding to rejected functions are set to zero (off-diagonal) and BigEig (diagonal). Default: 1e8.

`tolfit`

Similar to tolbas. The criterion is now applied to the overlap matrix of fit functions. The fit *coefficients*, which give the approximate expansion of the charge density in terms of the fit functions (for the evaluation of the coulomb potential) are set to zero for fit functions (i.e.: combinations of) corresponding to small-eigenvalue eigenvectors of the fit overlap matrix. Default 1e-10.

Notes:

- Application / adjustment of tolfit is not recommended: it will seriously increase the cpu usage while the dependency problems with the fit set are usually not so serious anyway.
- Application of the dependency/tolbas feature should not be done in an automatic way: one should test
  and compare results obtained with different values: some systems look much more sensitive than others.
  We have, so far, not been able to understand an unambiguous pattern in these experiences. Of course,
  when things become clearer in this respect, we will implement the corresponding intelligence into the program.
- When the dependency key is used, the numbers of functions that are effectively deleted is printed in
  the output file, in the SCF part (cycle 1) of the computation section.
- The TAPE21 result file of a calculation that used the DEPENDENCY key contains information about the omitted
  functions and these will also be omitted from the fragment basis when the TAPE21 is used as a fragment file.

Basis Set Superposition Error (BSSE)

The Ghost Atom feature enables the calculation of Basis Set Superposition Errors (BSSE). The idea is as follows. In a normal calculation of the bonding energy of a molecule c, composed of fragments a and b, one compares the total energies of c vs. those of isolated a and isolated b added together. In ADF this can be done in one stroke by running c from fragments a and b.

The BSSE is determined as the bonding energies of a pseudo-molecule d composed of (1) a plus a ghost b and (2) b plus a ghost a. The ghost atoms in the calculations are at their normal positions in the true molecule c, and they have their normal basis (and fit) functions. However, they do not have a nuclear charge and no electrons to contribute to the molecule. To set such a calculation up one needs first to make the appropriate ghost database files: for each involved atom, copy the database file that was used for its creation and modify it so as to remove the frozen core. Next, Create the ghosts with zero mass and zero nuclear charge. Apply these ghost fragments in the BSSE runs.

An example is worked out in the Examples document.

Control of Program Flow

Limited execution

```
STOPAFTER programpart
```

`programpart`

Must be a predefined name associated with a (major) part of the program With this key you tell ADF to terminate the job after the named program part has been executed.

A survey of the recognized names with a brief explanation follows below. The program parts are listed in order of execution: by taking a name further down the list you execute a larger part of the program.

`init`

initialization procedure, input reading and printing of the output header with the job identification.

`input`

input-reading module.

`geomet`

geometry section: organization of atoms in types of atoms and fragments, checks of the actual fragments against information on the attached fragment files.

`config`

electronic configuration (if not determined only by the SCF procedure), printout of symmetry subspecies.

`mainsy`

generation of symmetry information, representation matrices, etc.

`symfit`

construction of symmetry adapted fit functions.

`cblock`

generation of integration points and the distribution of them in the blocks that control the internally used segmented vectorization loops.

`engrad`

Relevant only in an optimization calculation. Engrad calculates energy gradients. The geometry is not yet updated and no printing of convergence tests and new coordinates is carried out.

`geopt`

This routine evaluates energy gradients and updates the geometry accordingly; it also prints the convergence tests and the computed new coordinates. Compare 'stopafter engrad'.

`forcematrix`

in a Frequencies run, terminate the calculation when all displacements have been done and before any further processing of the computed hessian, such as the determination of normal modes, takes place.

Direct SCF: I/O vs. recalculation of data

The program's performance can be defined in terms of the amounts of time (cpu and i/o seconds) and disk space used in a calculation. Also important for the human user is the turn-around time. On multi-user machines cpu-cheap jobs may take a lot of real time to execute due to i/o scheduling.

Therefore it can be a good idea to recompute some items rather than store them on disk. This will increase the amount of cpu time but reduce disk access and it may also improve the turn-around. Another consideration is of course that storage of data on disk may exhaust the available disk space in case of big calculations so that recalculation rather than storage is unavoidable.

```
DISK {{no}fit} {{no}basis}
```

instructs ADF how to handle the values of the fit functions and basis functions in all integration points: calculate once and store on disk or recompute whenever needed. The (optional) arguments are fit or nofit, and basis or nobasis.

fit and basis tell ADF to store the corresponding data on file; the prefix no induces recalculation whenever the data is needed.

Defaults are nofit and nobasis: direct-SCF mode for both features (this can be modified at the installation of ADF, see the Installation Manual).

The key DISK has replaced in ADF 2.0 the key directSCF in ADF 1.x, and extended the applicability of the I/O versus recalculation choice from fit functions-only to basis functions as well.

Skipping

With the following key you can restrict which parts of the program are actually executed:

```
SKIP argumentlist
```

`argumentlist`

A sequence of names, separated by blanks or commas. skip may occur any number of times in input. The names in the argument list refer to various items that are associated with parts of the program. With this key you tell ADF to skip the named program part(s) and to continue execution thereafter. The program does not check any consequences and may even crash when variables have not been initialized or have attained incorrect values due to the skipping.

Use of this key should be contemplated only in debugging and testing sessions, in which you may skip the computation of certain data when before that data will be needed you'll halt the program to inspect something.

Recognized and operational arguments are for instance (possibly not complete due to frequent extensions in this respect): atpair, ets, fitint, orthon, qmpot

Ignore checks

ADF performs several checks during a calculation, and stops with an error message when intermediate results are suspicious, when input-specified instructions are incompatible, etc. These controlled aborts can in some cases be overruled. Of course, the checks have been inserted for good reasons and one should realize that ignoring them probably produces incorrect results and/or may lead to a program-crash.

```
ALLOW argumentlist
```

`argumentlist`

A sequence of names, separated by blanks or commas. allow may occur any number of times in input, see the list below for the names that can be used.

`BadCoreInt`

Numerical integration of the frozen core density should closely approximate the analytical value. If the deviation is large compared to the user-specified numerical integration precision the program aborts with an error message like 'BAD CORE INTEGRAL'. This control is overruled by using this ALLOW option.

`BadIntegrals`

Only applicable when the direct-SCF option is turned off for the basis functions. (This happens automatically for ZORA full-potential calculations). In that case, a sequence of elementary overlap integrals are evaluated with the numerical integration grid and the outcomes tested against the

analytical value. If the deviation is too large a warning is issued. Above a certain threshold the program will abort, unless you override the exit with this Allow option.

`BadSCF`

If the SCF procedure hasn't converged, any geometry manipulations (optimization, linear transit ...) will be aborted because the energy gradients are not reliably computed in a non-self-consistent field.

`CloseAtoms`

Atom-atom distances should not be less than 0.2 Bohr. This is checked in the program section where the numerical integration grid is generated.

`RelGeo`

Geometry manipulation (optimization, linear transit...) is not supported for all of the relativistic options. See Relativistic

`SmallBlocks`

The list of numerical integration points is partitioned in blocks, so as to fit data arrays (for instance values of all basis functions in the points of a block) in available memory. The program computes the maximum block length from available memory and size parameters such as numbers of basis functions. A small block size implies a severe reduction in CPU efficiency. Therefore, the program aborts (by default, to override by this ALLOW option) if the block length turns out to be very small (less than 10).

`xc`

Certain combinations of the Density Functional options or application of them with some other features are not allowed. See XC.

Parallel Communication Timings

With the key

```
COMMTIMING
```

in the input you instruct ADF to skip normal execution and perform only a test on the gather, broadcast and combine routines, used in a PVM version of ADF. Obviously, this is only meaningful if such an ADF version has been installed.

Technical Settings

GPU Acceleration

ADF2014 can use NVidia GPUs for accelerating ADF calculations that use GGA functionals.

To benefit from GPU acceleration in ADF, you will need:

- A 64 bit Linux OS
- The IntelMPI+CUDA version of ADF2014
- An NVidia GPU with good double precision performance (we also support multi-GPU)
- The official NVidia drivers version 319 or higher installed

What parts of ADF are accelerated?

Single Point, Geometry Optimization and Frequency calculations can be accelerated. Only GGA functionals are supported at the moment. The GPU acceleration in ADF uses a hybrid scheme in which the GPU works in tandem with the CPU, so both need to be beefy for fast calculations.

What GPU do I need for accelerating ADF?

You will need an NVidia GPU that has good double precision performance. This currently (september 2014) means it should be any of the following devices: Tesla C2050/C2075/M2050 /M2070/M2090, Tesla K20/K20x/K40, GeForce GTX Titan/Titan Black/Titan Z.

How do I get it to work?

- Make sure your Linux OS is 64bit: run the following command in a terminal and check that it says x86_64 or amd64:
  `uname -m`
- Make sure your NVidia device is recognized and your driver version is higher than 319: run the following command in a terminal and check the output for driver version and your GPU:
  `nvidia-smi`
- The user running the ADF calculation must have read and write access to the device. To make sure you can access the GPU, run the following command in a terminal and check the permissions:
  `ls -l /dev/nvidia*`
- Download and install the IntelMPI+CUDA version of ADF from the "Less frequently used platforms" section on the download page
- Make sure your calculation uses a GGA functional

What input options are available?

The IntelMPI+CUDA version of ADF2014 enables GPU acceleration by default for any ADF calculation, no changes to the input are needed for this. To disable GPU acceleration (while using the IntelMPI+CUDA version), set the SCM_GPUENABLED environment variable to "FALSE". You can do this by inserting the following two lines on the second line in the .run script of your ADF job:

```
SCM_GPUENABLED="FALSE"
export SCM_GPUENABLED
```

The ADF input recognizes the following input options in the GPUENABLED block:

```
GPUENABLED
  SETVALIDDEVS [0 1 2 ...]
  NOFOCKY
  NOGRADS
END
```

The SETVALIDDEVS keyword should be followed by a list of integers that correspondent to the CUDA id numbers of your GPUs (note: these numbers are not the same as the ones reported by nvidia-smi). This keyword can also be (ab)used for manual load balancing if you have multiple GPUs. For example if you want to distribute load 2:1 between device 0 and device 1, specify SETVALIDDEVS 0 0 1. If you simply want to make sure that only your Tesla card is used and not some other NVidia GPU that is in the system, have a look at the CUDA_VISIBLE_DEVICES environment value in the NVidia documentation.

The NOFOCKY keyword disables GPU acceleration in the SCF cycles that calculate the Fock matrix.

The NOGRADS keyword disables GPU acceleration in the Frequency calculation part.

Some technical details about the GPU acceleration in ADF:

ADF is parallelized using the MPI standard, and CUDA 5.5 is used on top op that.

The code decides per CPU process which GPU device to use from the SETVALIDDEVS list, or from all devices that are available in the system with CUDA compute capability 2.0 or higher. This usually means that multiple CPU cores will be sharing a GPU board for the acceleration, and the ratio of CPU cores per GPU should thus not be too high. As a rule of thumb it is best to have no more than 6 cores per Fermi Tesla GPU (the C2050/C2075/M2050 /M2070/M2090 devices), no more than 12 cores per Telsa K20 GPU, and no more than 16 cores per K40 GPU. A GeForce GTX Titan should perform similar to a Tesla K20 GPU. Note that other GeForce boards should not be used, as they in general have a very low double precision performance, which means they likely to slow down your calculation instead of speeding it up.

Only parts of the ADF calculations are GPU accelerated, so you will still need a fast CPU. The accelerated parts work in a hybrid scheme: the GPU performs a numerical integration while the CPU calculates the needed input values. The speed-up that can be obtained this way depends on your system, the basis set size and the grid quality settings. We have seen speed-ups of up to 1.9x for single-point calculations, and up to 2.2x for frequency calculations.

ADF can use multiple GPUs when running on a single machine and also when running on multiple nodes. Each CPU process uses one GPU that is local to the CPU it is running on. This means that ADF should run on at least as many cores as GPUs it is supposed to use. When running in a cluster environment, ADF will only detect and use GPUs that are inside the node the job is running on.

What should I do if I have problems?

Send an email to support@scm.com that mentions GPU acceleration and explains your problems. Attach the input and output of your job to this email, as well as the outputs of “nvidia-smi”, “ls -l /dev/nvidia*”, “uname -a” and “cat /proc/cpuinfo”

Memory usage

The amount of memory used by the program during a calculation is determined by three quantities:

- The size of the program itself (executable statements, static arrays).
  This quantity depends on the program version and is currently around 20 MB.
- Buffer space used by ADF for more efficient I/O handling. This quantity is set at installation.
  See the Installation Manual.
- Dynamically allocated arrays. The program allocates memory dynamically during the run conform the requirements of the actual calculation.

Starting from ADF2010 in case of parallel calculations some of the data arrays that are used within ADF will be shared by processes on the same node, provided the operating system allows shared memory. This will reduce the total amount of memory used by all ADF processes on each node because only one copy of certain large arrays per node will be present. Note that shared arrays is not the same as distributed arrays. To disable the use of shared memory one can specify the following keyword:

```
NoSharedArrays
```

Vector length

Numerical integration is applied in ADF to evaluate Fock matrix elements and many other quantities that are defined as integrals over basis functions, the charge density, the potential, etc. As a consequence a large part of the CPU time is spent in simple do-loops over the integration points. The total number of points depends on the required precision and on the number of atoms, the geometry and symmetry. All such numerical integration loops are segmented into loops over *blocks* of points, each block consisting of a certain number of points. This latter defines the most inner do-loop and hence determines vectorization aspects.

Depending on the computer, c.f. the compiler, vector operations may be executed more efficiently using longer vectors. Long vectors increase the demand on Central Memory however because the program may sometimes have to access large numbers of such vectors in combination (for instance all basis functions) so that they must be available in memory simultaneously. The optimum vector length depends therefore on the balance between vectorization efficiency and memory usage. The maximum vector length that you allow the program to use can be set via input.

```
VECTORLENGTH vectorlength
```

The default is set at the installation of ADF on your platform, see the Installation manual. For organizational reasons the true vector length actually used in the computation may be smaller than the value defined with this key, but will not exceed it (except in a Create run, but in that case performance and memory usage are no hot topics).

Tails and old gradients

The key TAILS is currently obsolescent because of the introduction of the LINEARSCALING keyword and may be removed in future versions. The key TAILS was used in older versions, ADF2004.01 and before, in the calculation of the gradients.

Each block of points (see above) covers (more or less) a certain region in space and can hence be assigned a distance value with respect to a particular atom. These distances are used to control whether or not to evaluate functions centered on that atom in that particular block of points.

```
TAILS {bas=tailbas} {fit=tailfit}
```

`tailbas, tailfit`

> Accuracy levels, similar to the integration parameter: a higher value implies higher precision: in this case, basis functions and fit functions respectively are assumed zero in blocks of points that are at a sufficiently large distance from the atom at which the function is centered. Sufficiently large is defined by comparing the integral of the (radial part of the) function beyond that distance with the total integral. By default tailbas and tailfit both depend on the numerical integration parameter

Note: in contrast with some of the older versions, supplying only the keyword without parameters does not switch off the use of function cutoffs. To effectively switch off the distance effects in gradients evaluation one should specify large values for the BAS and FIT parameters. The value of 100 should be more than enough, thus, for example:

```
TAILS bas=100 fit=100
```

Improved performance in geometry optimizations and frequency runs is achieved by a new implementation of the calculation of the gradients that now uses linear scaling techniques. This is now the default. One can still use the old implementation if one includes in the input:

```
OLDGRADIENTS
```

The key TAILS is not used in geometry optimizations anymore. For controlling the use of distance effects in normal SCF calculations, and for calculations with the RESPONSE or EXCITATIONS keywords, please check the LINEARSCALING keyword.

Linearscaling

The LINEARSCALING keyword has a very similar function to the TAILS keyword described above. In addition to defining the precision of operations related to operations in the numerical integration grid, it also defines the precision for the calculation of the overlap matrix, the fit integrals, and the density fit procedure.

Default values have been chosen which result in negligible differences in the results for our test calculations, so that these defaults can be considered safe. They have been chosen similarly to the defaults for the TAILS keyword.

However, it may be advisable to modify the settings for the linear scaling parameters in two cases. First, if a very accurate result is needed, and numerical noise is to be completely eliminated, strict values can be specified. Especially for small molecules, where timings are not so large anyway, this may be of interest. Second, for large molecules, in which the calculations are very time-consuming, one can experiment with less strict values for the LINEARSCALING block keyword. In such a case one should be aware of the reduced accuracy and preferably test the influence of the changes on the results.

In the simplest application of the LINEARSCALING keyword, only one parameter is provided. All the subkeys described below will then be given this value. A very large value implies a calculation where no distance cut-offs are used. A normal value (almost default situation) would be 8 for linscal, 6 gives a faster but somewhat sloppier result. Whether this is acceptable is strongly case-dependent. A value of 10 or 12 is already quite strict and, unless there are some sort of numerical problems, there should not be much influence on the results by choosing a stricter value than that. A value of 99 for linscal virtually excludes the possibility that something will be neglected.

```
LINEARSCALING linscal
```

More refined control is possible by using the full block key

```
LINEARSCALING
  CUTOFF_FIT epsfit
  OVERLAP_INT ovint
  PROGCONV progconv
  CUTOFF_COULOMB epsvc
  CUTOFF_MULTIPOLES epsmp
END
```

CUTOFF_FIT

determines how many atom pairs are taken into account in the calculation of the fit integrals and the density fit procedure. If the value is too low, charge will not be conserved and the density fitting procedure will become unreliable. This parameter is relevant for the timings of the FITINT and RHOFIH routines of ADF.

OVERLAP_INT

determines the overlap criterion for pairs of AO's in the calculation of the Fock-matrix in a block of points. Indirectly it determines what the cut-off radii for AO's should be. The value of ovint has a strong influence on the timing for the evaluation of the Fock matrix, which is very important for the overall timings. The default value for ovint is accint + 2 (typically 6). Again, a higher value implies a safer but slower calculation.

PROGCONV

determines how the overall accuracy changes during the SCF procedure ('progressive convergence'). The idea is that one might get away with a lower accuracy during the initial SCF cycles, as long as the last cycle(s) is/are sufficiently accurate. The current default is that progconv has the value 0, which means that the accuracy in the beginning of the SCF is the same as in the rest of the SCF. This keyword is currently still in the testing phase, so we do not recommend changing its default value. The value of progconv determines how much lower the other parameters in the LINEARSCALING input block are at the beginning of the SCF than at the end.

CUTOFF_COULOMB

determines the radii for the fit functions in the evaluation of the (short-range part of) the Coulomb potential. As the Coulomb potential may take a sizable amount of time, the value chosen for epsvc may influence the total ADF timing significantly as well. The default value for epsvc is accint + 4 (typically 8).

`CUTOFF_MULTIPOLES`

determines the cut-offs in the multipole (long-range) part of the Coulomb potential. This term scales quadratically with system size, but has a small prefactor. In most cases, change in the epsmp value will not affect the CPU time significantly. The default value for epsmp is accint + 4 (typically 8).

All Points

ADF makes use of symmetry in the numerical integrations. Points are generated for the *irreducible wedge,* a symmetry unique sub region of space. Optionally the symmetry equivalent points are also used. This is achieved by setting the key

```
ALLPOINTS
```

The key has no argument. The CPU time increases roughly by a factor equal to the number of symmetry operators, and the results should be the same. This key is available only as a debugging feature, to check the correctness of certain symmetry related algorithms.

Full Fock

At every cycle in the SCF procedure the Fock operator is computed in all integration points. By default the *difference* with the values of the previous cycle are used to compute *changes* in the Fock matrix elements. This leads in general to better computational efficiency in two ways: 1) when all such difference values in a block of integration points are very small such a block is skipped in the calculation. 2) if the values are not negligible but still rather small, the contribution from such a block to matrix elements between basis functions with small overlaps are neglected.

With the key

```
FULLFOCK
```

this is turned off, so that the complete matrix elements are computed, no blocks are skipped and the neglect of matrix elements between functions with small overlaps (see also the key TAILS) is controlled solely by the function characteristics and precision requirements, not by the development of the SCF.

Electrostatic interactions from Fit density

By default the program tries to evaluate the electrostatic Coulomb interaction energy between the fragments in a molecule using the exact fragment charge densities. The implemented algorithm requires that all fragments are spherically symmetric. This is checked by the program by verifying that all fragments have been computed in atom symmetry. It that is not the case, an alternative method is applied, using the fitted charge densities of the atoms; this is an approximation with a small, but not insignificant error. The following key forces the program to apply the fit density approach even in the case of spherically symmetric fragments. This aspect applies only to the final bonding energy analysis, not to energy computations and their gradients within the automatic geometry optimizer. The purpose of this option is to simulate a previously existing situation where the electrostatic term in the bonding energy was computed from the fit density regardless of the fragments and their internal symmetries.

```
FITELSTAT
```

presence of this key in the input file triggers using the fit density.

Save info

Several types of information, gathered during the run, are lost on exit. The SAVE key allows you to prevent the removal of such information.

```
SAVE info
```

`info`

A sequence of names separated by blanks or commas. save may occur any number of times in the input file.

save turns save-info options on. A lists of the available options, with their default status.

item	*default*	*explanation*
TAPE10	no	File with numerical integration data: points and weights, values of functions (depends on direct-SCF options) and core densities and potentials.
TAPE11	no	File with fit integrals (STOFIT).
TAPE13	no	Check point file. This file is lost (by default) only upon normal program exit, i.e. a program-controlled termination (including a program-detected error condition leading to controlled exit). In all such cases all info on TAPE13 is also present on TAPE21. tape13 exists when the program crashes into a core dump for instance, in which case it is uncertain what the contents of TAPE21 will be. The save feature allows you to specify that TAPE13 is kept *also* upon normal exit.
TAPE14	no	Scratch file with numerical integration data, mainly pertaining to individual fragments.
Timing	no	During an ADF calculation the program gathers a large amount of timing information about the performance of different program parts. It can be printed, at various levels of detail, on standard output (key PRINT). It can also be stored on TAPE21, for later inspection, in a section Timing.
Table VII. Arguments for the key save.		

2.10 Restarts

Restart files

When an ADF calculation terminates abnormally - not controlled by the program itself, for instance after a core dump due to some bug - there will usually be a file TAPE13, which serves as a checkpoint file. tape13 can be used to restart the calculation at a point not too far before the fatal condition occurred. It contains only data for the restart, but none of the special analysis data on TAPE21 that would be useful for analysis, to serve as fragment file, etc.

TAPE13 is upgraded during the calculation but discarded upon normal termination, namely when all relevant information has been saved on TAPE21. At that point all info that would have been on TAPE13 is present on TAPE21. If you wish to keep tape13 anyway - for instance because you plan a restart after normal termination and don't intend to keep the substantially bigger TAPE21 - you must use the save key.

Upon normal (i.e. program-controlled) termination of a calculation, the TAPE21 result file can be used for restart purposes. When a crash occurs, however, chances are that TAPE21 has not correctly been closed and that its data structure is inconsistent: during the calculation large portions of TAPE21 are kept in memory rather than on file, and only at the point of final termination, all data is flushed to file.

General remarks

In all restart calculations a normal input file must be supplied (you can, for instance, simply take the original one), with a specification of the restart file added: the restart file does *not replace* the input file. From the program's point of view, it first reads the 'normal' input file and then inspects whether a restart file is present to replace some of the information read from input.

The concept of restarts in ADF is rather simple and primarily directed at increasing computational efficiency by providing cost-expensive data. The continuation run is to a large extent independent from the one that generated the restart file. The runtype, the choice of density-functional and other features in the Hamiltonian, precision of numerical integration, thresholds on convergence, et cetera are all determined solely from the input file for the new run: no such data is read from the restart file. Most input items should, therefore, be supplied in the restart run again, even if it is a direct continuation of a previous calculation: omission implies using the standard defaults, which are not necessarily the settings of the calculation that generated the restart file.

Even the key ATOMS with the list of atomic coordinates must be supplied again: the program needs the information herein to deduce what fragments are used, which coordinates are free or frozen respectively in an optimization, etc. The coordinate *values* may be supplied with the restart file and these will then overwrite those specified in the input file.

Obviously, the two runs cannot be completely unrelated. To let the restart data make sense the runs should correspond to the same molecule (i.e. its general definition in terms of fragment building blocks). The program does not check all aspects related to this and certain abuses will therefore survive the internal tests, but will surely lead to some error later on: it is the user's responsibility to ensure that the restart data match the calculation one has in mind.

Interdependencies between data read from the restart file (rather than from input or fragment files) and other items imply that some input keys and some options to specific keys may be inaccessible when restart data are provided. In most cases supplying such inaccessible input options will simply be ignored; in some cases a warning is issued or an error abort occurs.

A restart file supplies data from a previous run that might be useful in the current one. The applications are (combinations are possible):

- Get a better start in the (first) SCF procedure by providing the electronic charge density (in the form of fit coefficients) from the preceding run,
- Continue an optimization by supplying the latest geometry (coordinates) from a previous run via the restart file (rather than typing them in),
- Get faster geometry convergence by supplying a Hessian,
- Breaking large jobs (Linear Transits, Frequencies) in smaller ones, each time doing a part and passing this on to the continuation run.

WARNING**.** The SCF and optimization procedures use *history* to improve convergence behavior. Most of such history information is not stored on a restart file. As a consequence, a restart may not continue exactly as the original run would have done if it hadn't terminated. In a SCF restart, for instance, the DIIS procedure has to rebuild the information. The same holds for geometry optimizations, although history plays usually not a very big role there.

The restart key

The name of the restart file must be provided with the key RESTART (see below). A list of data items is read from the file (if present on the file and only as far as significant for the new run) and used unless their usage is explicitly suppressed by the user.

Simple key:

```
RESTART restartfile
```

Block key:

```
RESTART restartfile &
  optionlist
  optionlist
  ...
end
```

`restart`

This *general* key can be used as a simple key - to supply the name of the restart file - or as a block key. In the latter case the continuation code (&) must be applied to tell the program that a data block follows.

`restartfile`

The name of a file with restart data. The path (absolute or relative) to the file must be included if the file is not local to the directory where the calculation executes. In most cases it will be a TAPE21 file from an ADF calculation, but this is not necessary. It may be any file - constructed by the user for instance - provided it has the right structure. It must be a kf file and the data to be used must be stored in sections and under variable names as defined below, which is exactly how such data are generated by a normal ADF run on TAPE21 or on the checkpoint file TAPE13.
Note: the filename must not be one of the standard filenames used internally by the program, such as TAPE21, TAPE13 etc. Generally: don't use a name like tapenn where nn is a two-digit number.

`optionlist`

A list of options, separated by blanks or commas. The following options are applicable:

`noSCF`

Do not use any fit coefficients from the restart file as a first approximation to the (fitted) SCF density for the new calculation. Instead, the sum-of-fragments density will be used, as in a non-restart run. Note, typically noSCF should be used in combination with noORB.

`noORB`

Do not use orbitals from the restart file.

`nogeo`

Do not use the geometry - Cartesian, Z-matrix, etc. coordinates - from the restart file.

`nohes`

Do not use any Hessian from the restart file.

`SPINFLIP atomnumbers`

See the separate section about the spin-flip method for converging broken-symmetry systems.

Note: in the continuation of a Linear Transit, IRC or Frequencies run, geometric data are read from the restart file and will be used: the option nogeo is ignored. In a continued Frequencies run the input coordinates (key ATOMS) must be correct (i.e. the equilibrium geometry). In a continued LT or IRC run, the input coordinate values from atoms are ignored (but they must be supplied to give the program a preliminary count of atoms and fragments involved).

Structure of the restart file

All data that may be retrieved from the restart file must be stored in a specific location on the restart file. If you're simply using a TAPE21 result file or a TAPE13 checkpoint file you don't need to bother about this: ADF has put all data in the right place; the following discussion is primarily for those who want to manipulate the restart file or even construct one themselves.

Since the restart file must be a kf file, the location of the data is of the form Section%Variable, specifying the section and the variable name. The section and variable names are case sensitive. See the utilities document for general information about kf files.

If the specified variable is not present in the specified section on the restart file - or if there is no such section at all - the data is not used, usually without an error message. In some cases a few global tests are carried out on the retrieved data; if they fail the tests the data are not used and a warning - in some cases an error abort - may be issued by the program.

KF files are binary files and so are the TAPE21 result file, the TAPE13 checkpoint file and generally any restart files. If you wish to edit and modify the contents, or just inspect them, the standard KF utilities can be used. Apply pkf to get a survey of the sections and variables on the file, dmpkf to get a complete ASCII version of the file and udmpkf to transform an ASCII version - presumably edited and modified - back into binary format. See Appendix 5.5.

Data on the restart file

Follows a survey of all data items that the program may search for on the restart file.

SCF data

`Fit%coef_SCF`

The fit-expansion of the charge density to be used as start-up for the next SCF. Without these restart fit data the first SCF will start from the (fitted) sum-of-fragments charge density.

`Fit%coef_FreqCenter`

Only in a Frequencies run: the fit-expansion of the SCF-converged equilibrium geometry. It usually helps to get a somewhat better start-up of the SCF in displaced geometries.

If the noSCF option is used to the restart key, any Fit%coef_? data on the restart file are ignored.

Coordinates

`Geometry%xyz`

Cartesian atomic coordinates. The option nogeo suppresses using such data. In a Frequencies or continued Linear Transit run, they may be read but will be ignored (i.e. replaced by other coordinates data from the restart file, see below).

In most applications, when coordinates are read (and used) from the restart file, only *Cartesian* coordinates are retrieved and the corresponding Z-matrix values are computed from them, using the Z-matrix *structure* defined in the atoms data block. This is one of the reasons why the ATOMS key must be used even when the atomic coordinates are supplied on the restart file.

Hessian

`GeoOpt%Hessian_CART`

`GeoOpt%Hessian inverted_CART`

`GeoOpt%Hessian_ZMAT`

`GeoOpt%Hessian inverted_ZMAT`

> All these four varieties are searched for if the new run searches for a restart Hessian matrix at all, that is: in an optimization, Linear Transit or Transition State search. As the names should suggest these variables stand for the Hessian, respectively the inverse of the Hessian in Cartesian or z-matrix coordinates.
> In all cases the full square matrix must be present, with dimension the number of atomic coordinates, 3 times the number of atoms. This holds also for z-matrix coordinates. The 6 dummy coordinates play no role, the corresponding matrix elements in the Hessian should be zero. The order of atoms is the same as in the input.

If a Hessian is searched for on the restart file, all four possibilities above are tried and the first one found is used, the other ones being ignored. The order in which they are tried is:

If the current run uses *Cartesian* coordinates as optimization variables, then first the two cart varieties are tried, and vice versa for *z-matrix* optimization.

In a minimization (simple optimization or Linear Transit) first the *inverted* variety is tried; in a Transition State search the normal (not inverted) Hessian is looked for first.

Note**:** If a z-matrix Hessian is retrieved from the restart file the program will use the underlying z-matrix *structure* to derive a Cartesian Hessian from it. In such case the restart file must also contain:

`GeoOpt%kmatrix`

> The z-matrix structure (references to the atoms in this matrix assume the ordering of atoms as used internally by the program).

Note: the kmatrix on the file need not be identical to the kmatrix used in the current calculation. In fact, the current calculation may not even have a z-matrix structure.

Transition State

In a continued TS run the program retrieves, apart from general geometry *optimization* data such as the Hessian - see above - only the latest TS search vector: the eigenvector of the (approximate) Hessian that points to the Transition State. All other TS-specific data are input-determined with corresponding defaults.

The TS search vector is stored in:

`TS%mode to follow`

> A list of atomic coordinates (Cartesian or Z-matrix, depending on the type of optimization variables used. The underlying list of atoms has the atoms not necessarily in the order in which they have been given in input: rather they are grouped together by atom type.

Linear Transit

In a continued Linear Transit (LT) calculation the continuation run proceeds from where the previous run stopped. The total number of points by which the transit is scanned, the current point (its index and the Cartesian coordinates), the accumulated results of completed points on the transit etc. are copied from the restart file. If the restart file contains a section LT, then all relevant data must be present on it and correct (i.e. matching those of the current run: same number of LT *parameters*, and of course the same molecule.

`LT%nr of points`

> The number of points by which the LT is scanned; this is identical to the Fortran variable ltimax in the code. The value on the restart file applies in the calculations and overwrites any input/default value (see the subkey *lineartransit* of the geometry block)

`lt%current point`

> Index of the current LT scan point. This is where the program will continue. In a non-restart LT run, this index initializes at 1.

`LT%Energies`

> An array with energy values, one for each LT point. When the LT run is completed, this array allows you to map out the energy along the LT path. The values for the completed LT points are stored on the restart file. This size of the array on the restart file must (at least) be the total number of points on the complete path.

`LT%Parameters`

> Initial and final values for the LT parameters, which describe roughly the path (all other coordinates may be optimized at each point, depending on other input keys). The values from the restart file overwrite input values. The input values should be supplied, however, as if it were a non-restart run.

`LT%atmcrd`

> zmat if a z-matrix structure is available for the molecule, cart otherwise. This is used to control printing of results. It does not define the type of optimization variables: see the next item.

`LT%geocrd`

> zmat or cart: the type of optimization variables. This defines in which type of coordinates the LT parameters are defined and any optimization of other coordinates takes place.

`LT%xyz`

> Cartesian coordinates for *all* LT points: 3*atoms*ltpoints. The size of the array must conform to this. Only the values of the completed LT points *and* those of the current point are relevant. Those of the current LT point are used as initial coordinates to start the current run.

`LT%zmatrix`

> Same for the Z-matrix coordinates. They should match the Cartesian coordinates for the completed LT points (this is not checked). Those for the current LT point will be recomputed from the current *Cartesian* coordinates.

IRC

In a continued Intrinsic Reaction Coordinate (IRC) calculation, the continuation run processes the path(s) as specified in input. Any info for such path(s) on the restart file will then be used to continue from there. If the

restart file contains the relevant IRC sections, see below, then all relevant data must be present on it and correct (i.e. matching those of the current run.

The sections on file pertaining to the IRC are:

IRC: this section contains information about the central (TS) point, which variables are optimized in each of the IRC points, the connection matrix defining the z-matrix structure, etc.

IRC_Forward and IRC_Backward: these sections contain the data of the two paths from the Transition State down to the two adjacent local energy minima: for each point the distance from the previous point and the local curvature and molecular properties such as energy, atomic charges and dipole moment.

`LT%nr of points`

The number of points by which the LT is scanned; this is identical to the Fortran variable ltimax in the code. The value on the restart file applies in the calculations and overwrites any input/default value (see the subkey *lineartransit* of the geometry block)

`LT%current point`

Index of the current LT scan point. This is where the program will continue. In a non-restart LT run, this index initializes at 1.

`lt%Energies`

An array with energy values, one for each LT point. When the LT run is completed, this array allows you to map out the energy along the LT path. The values for the completed LT points are stored on the restart file. This size of the array on the restart file must (at least) be the total nr of points on the complete path.

`lt%Parameters`

Initial and final values for the LT parameters, which describe roughly the path (all other coordinates may be optimized at each point, depending on other input keys). The values from the restart file overwrite input values. The input values should be supplied, however, as if it were a non-restart run.

`lt%atmcrd`

zmat if a z-matrix structure is available for the molecule, cart otherwise. This is used to control printing of results. It does not define the type of optimization variables: see the next item.

`lt%geocrd`

zmat or cart: the type of optimization variables. This defines in which type of coordinates the LT parameters are defined and any optimization of other coordinates takes place.

`lt%xyz`

Cartesian coordinates for *all* LT points: 3*atoms*ltpoints. The size of the array must conform to this. Only the values of the completed LT points *and* those of the current point are relevant. Those of the current LT point are used as initial coordinates to start the current run.

`LT%zmatrix`

Same for the Z-matrix coordinates. They should match the Cartesian coordinates for the completed LT points (this is not checked). Those for the current LT point will be recomputed from the current *Cartesian* coordinates.

Frequencies

In the continuation of a Frequencies calculation all Frequencies-related data are retrieved from the section Freq on the restart file. (SCF fit data are, as always, retrieved from the section Fit). A fairly large number of items will be read and must all be present (*if* a section Freq is present in a restart file supplied to a Frequencies run). Technical parameters such as the type of numerical differentiation, size of displacements etc. are read from the restart file. Any input specifications are ignored.

`Freq%kountf`

Counter of number of geometries completed. In a non-restart run this is initialized at zero; in a restart it is read from the file.

`Freq%nraman`

Flag for RAMAN calculations

`Freq%numdif`

1 or 2: defines numerical differentiation used to compute the force constants from the gradients in slightly displaced geometries (by 1-point or 2-point differentiation).

`Freq%disrad`

Size of displacement for cartesian or bond-length displacements.

`Freq%disang`

Size of angular (bond angle, dihedral angle) displacements.

`Freq%atmcrd`

zmat or cart: specifies whether a z-matrix structure is present. This does not define the type of displacement coordinates, see the next item.

`Freq%geocrd`

Type of coordinates in which the displacements are carried out: zmat or cart

`Freq%nfree`

Number of free and independent displacement variables.

`Freq%idfree`

References from the atomic coordinates (in internal order) to the independent displacement variables.

`Freq%all freedoms`

(logical) flags whether or not the complete energy surface is scanned around the equilibrium or only part of the internal degrees of freedom are used.

`Freq%xyz`

equilibrium coordinates (internal order of atoms).

`Freq%kmatrix`

Z-matrix structure. Pointers are indexed by and refer to atoms in the internally used order.

`Freq%zmatrix`

Z-matrix coordinates of the equilibrium geometry (internal ordering of atoms).

`Freq%rigids`

6 rigid motion vectors (one may be zero, in case of a linear molecule). Each vector has as many components as there are atomic coordinates. The values correspond to the internal ordering of atoms.

`Freq%xyz displaced`

Cartesian coordinates of displaced geometry to carry out now. In a non-restart run this would be the equilibrium geometry.

`Freq%zmatrix displaced`

Similar for the Z-matrix coordinates.

`Freq%Dipole previous`

dipole vector (3 components) for the last geometry handled.

`Freq%Dipole`

dipole at the equilibrium geometry.

`Freq%Dipole derivatives`

Derivatives of the dipole wrt atomic coordinate displacements.

`Freq%Gradients previous`

Energy gradients (derivatives wrt atomic coordinate displacements) in the last handled geometry.

`Freq%Force constants`

Matrix of force constants. This is, together with the *Dipole derivatives* the final quantity to compute. At each cycle of the Frequencies data are added to it. Upon completion of the Frequencies cycles the frequencies and normal modes are computed from it. Together with the dipole derivatives it then also yields the InfraRed intensities.

2.11 Examples

See the separate Examples document.

3 Recommendations, problems, Questions

3.1 Recommendations

Precision

The quality of the calculation, *given* the selected model Hamiltonian - density-functional, relativistic features, spin-restricted/unrestricted... - is determined to a large extent by several technical precision parameters.

The most significant ones are:

`Basis set`

Obviously, the quality of the basis set may have a large impact on the results. As a general rule, minimum and almost-minimum basis sets (types SZ and DZ) may be used for pilot calculations, but polarization functions should be included (DZP, TZP) for more reliable results.

`SCF convergence`

The self-consistent-field (SCF) and geometry optimization procedures terminate when convergence criteria are satisfied. If these are set sloppy the results may carry large *error bars*. The default SCF convergence tolerance is tight enough to trust the results from that aspect. However, when the SCF procedure encounters severe problems an earlier abort may occur, namely if a secondary (less stringent) criterion has been satisfied (see the key SCF). Although this still implies a reasonable convergence, one should be aware that for instance the energy may be off by a few milli hartree (order of magnitude, may depend quite a bit on the molecule). It is recommended that in such cases you try to overcome the SCF problems in a secondary calculation, by whatever methods and tricks you can come up with, rather than simply accept the first outcomes.
Note: in a geometry optimization the SCF convergence criteria are relaxed as long as the geometry optimization has not yet converged. This should generally not affect the final results: the SCF density and hence the energy gradients may be somewhat inaccurate at the intermediate geometries, but since these are not a goal in themselves the only concern is whether this might inhibit convergence to the correct final geometry. Our experiences so far indicate that the implemented procedure is reliable in this aspect.

`Geometry convergence`

This is a far more troublesome issue. Three different types of convergence criteria are monitored: energy, gradients and coordinates. The energy does not play a critical role. Usually the energy has converged well in advance of the other items. The coordinates are usually what one is interested in. However, the program-estimated uncertainty in the coordinates depends on the Hessian, which is not computed exactly but estimated from the gradients that are computed in the various trial geometries. Although this estimated Hessian is usually good enough to guide the optimization to the minimum - or transition state, as the case may be - it is by far not accurate enough to give a reasonable estimate of force constants, frequencies, and as a consequence, neither of the uncertainties in the coordinates. An aspect adding to the discrepancy between the Hessian-derived coordinate-errors and the true deviations of the coordinates from the minimum-energy geometry is that the true energy surface is not purely quadratic and using the Hessian neglects all higher order terms. The gradients provide a better criterion for convergence of the minimizer and therefore it is recommended to tighten the criterion on the gradients, rather than anything else, when stricter convergence than the default is required.
The default convergence criteria, in particular for the gradients, are usually more than adequate to get a fair estimate of the minimum energy. Tighter convergence should only be demanded to get more reliable coordinate values (and in particular, when the equilibrium geometry needs to be determined as a preliminary for a Frequencies run).

`Numerical integration accuracy`

The key BeckeGrid key block (or alternatively the old INTEGRATION key block) determines the numerical precision of integrals that are evaluated in ADF by numerical integration, primarily the Fock matrix elements and most of the terms in the gradients. In addition the integration settings also determine several other computational parameters. The demands on numerical integration precision depend quite a bit on the type of application. The SCF convergence seems to suffer hardly from limited integration precision, but geometry convergence does, especially when tight convergence is required and also in transition state searches, which are generally more sensitive to the quality of the computed energy gradients. An extreme case is the computation of frequencies, since they depend on differences in gradients of almost-equal geometries. Frequency calculations on molecules with sloppy modes suggest that a BeckeGrid of "good" quality may be required. Note: a large integration value implies that a lot more points will be used in the numerical integrals, thereby increasing the computational effort (roughly linear in the number of points).

Electronic Configuration

Not specifying occupation numbers in input will *not* automatically result in the computational of the ground state. It may even lead to non-convergence in the SCF and/or in the determination of minimum-energy geometries or transition states. Therefore: whenever possible, specify occupation numbers explicitly in input (key OCCUPATIONS)!

Misunderstanding results of a calculation may easily result from a lack of awareness of how ADF treats the electronic configuration, which orbitals are occupied and which are empty. Unless you specify occupation numbers in input they will be determined from the aufbau principle but only during the first few SCF cycles. Thereafter the distribution of electrons over the different symmetry representations is frozen (see the key OCCUPATIONS, options AUFBAU and aufbau2). If at that point the potential has not yet sufficiently relaxed to self-consistency the *final* situation may be non-aufbau.

A related aspect is that the *ground state* does not necessarily *have* an aufbau occupation scheme. In principle, different competing electronic states have to be evaluated to determine which has the lowest total (strongest bonding) energy.

Check output always carefully as to which orbitals are occupied. In general, whenever possible, supply occupation numbers in input. Be aware that the automatic choice by the program may in a Geometry Optimization result in different configurations in successive geometries: the automatic assessment by the program will be carried out anew in each SCF procedure. If competing configurations with comparable energies have different equilibrium geometries, the geometry optimization has a high failure probability. The gradients computed from the SCF solution of a particular configuration drive the atoms in a certain direction, but in the next geometry, when the program re-determines the occupations and finds a different configuration, the resulting gradients may drive the atoms in another direction.

See the keys CHARGE and OCCUPATIONS for user-control of occupation numbers.

Spin-unrestricted versus spin-restricted, Spin states

If your molecule has unpaired electrons, you should run an unrestricted calculation, in principle. However, if this exhibits convergence problems (or if you simply want to save time: an unrestricted calculation takes a factor 2 more CPU time and data storage), you may consider to do it in two steps. First, run a spin-*restricted* calculation. Then perform a spin-unrestricted calculation using the restricted TAPE21 as a restart file. In the follow-up calculation you should specify the precise occupation numbers for the state you're interested in, *and* use the SCF input key to specify *only one* SCF cycle (iterations=1). This prohibits convergence (so you keep the converged *restricted* orbitals) and gives you a fairly adequate approximation to a converged unrestricted result. See also the H2 example run for a discussion in the Examples document.

An unrestricted calculation does not necessarily yield the multiplet configuration (triple, doublet ...). This is a rather complicated matter, see the discussion on multiplet states, key SLATERDETERMINANTS.

Geometry Optimization

Bond angles of zero or 180 degrees

Avoid bond angles of 0 or 180 degrees. Use a dummy atom at a location orthogonal to the co-linear triple and define angles w.r.t. the dummy atom.

Be aware that bond angles can be explicit - these are easily recognized - but also implicit, in the definition of dihedral angles: it is absolutely imperative that such implicit bond angles are never 0 or 180 degrees: the dihedral angle will not be properly defined and an error will occur.

The program may in some cases be able to recover from 0/180 degree bond angles, but this is not a certainty. If it fails, the geometry update steps may go completely wild. Even worse: the steps may remain small but convergence is not reached, without a clear and explicit indication in the output about the cause.

Sloppy modes

Many molecules have sloppy modes, implying that geometric departures along these modes from the true minimum hardly change the energy and do not result in sizeable gradients. This usually shows up in slow convergence: energy and gradients appear to be converged but the computed step lengths, an assessment of the error in the geometry itself, do not disappear.

Starting from ADF2005.01 delocalized coordinates can be used in geometry optimizations and transition state searches. The use of delocalized coordinates often help in convergence of these problematic sloppy modes.

It depends then on the purpose of the run whether a continued search for the minimum is useful if one has slow convergence: not if you only want to know the energy at the minimum, but certainly so if you want to determine all geometric parameters to high precision. Depending on the case you may therefore want to relax the convergence criterion on the coordinate steps. In the case of Z-matrix optimization this has to be done primarily for the *angular* coordinates because the bond lengths are usually much stiffer and will therefore not suffer from sloppy mode problems. If you insist on strict convergence of sloppy modes you should use a fair integration precision (BeckeGrid quality: "good").

Step convergence

The criterion on convergence of the coordinates (steps) is often *not* a reliable measure for the precision of the final coordinates, although it does give a reasonable estimate (order of magnitude). To get accurate results you should tighten the criterion for the gradients, rather than for the steps.

What basis set should I use in ADF?

See the section on what basis set should I use in ADF? in this ADF Manual.

Frequencies

Outcomes of Frequencies calculations are usually quite sensitive to the geometry, so before computing the frequencies, one should make sure that the geometry is well converged *at the level of the subsequent Frequencies calculation:* the same model parameters and basis sets.

In all cases one should take care that the precision of Numerical Integration is adequate (this is good advice anyway for a sound Frequencies calculation).

Doing one-point, rather than two-point differentiation will roughly save you half of the time needed to complete the calculation. Increasing the integration precision will work the other way. To obtain high-precision results using one-point differentiation requires for one thing that you use very small displacements (smaller than the defaults) *and* high accuracy of numerical integration. Some studies [14, 107] suggest to use *a*) two-sided displacements, *b*)

This may not always be feasible due to the high CPU costs, but it should at least stress the importance of accuracy in the computation of frequencies.

A computation of frequencies runs over discrete displacements of atomic coordinates. When using Cartesian displacement coordinates, the program applies symmetry to skip symmetry-equivalent displacements and thereby save CPU time. In the output and logfile you'll find in such a case that the 'frequency displacement counter' skips one or more values: the counter counts all possible displacements, while only the symmetry-unique ones are actually carried out.

Starting from ADF2005.01 symmetric displacements can be used. This speeds up the computation significantly.

Relativistic methods

The ZORA relativistic approach is often superior and in other cases at least similar to the older Pauli method. In particular for all-electron calculations generally, and for very heavy elements even within the frozen core approach, the Pauli method may exhibit significant shortcomings. This is mostly due to the variational instability of the Pauli formalism in the deep-core region near the nucleus. The bigger the basis set and the smaller the frozen core, the more likely this will show up, while generally speaking you might be tempted to use smaller cores and bigger basis sets to *improve* your results. The ZORA approach does not suffer from these problems and is, therefore, highly recommended over the Pauli formalism.

3.2 Trouble Shooting

This chapter contains hints to help you solve some problems and comments on frequently asked questions.

License file corrupt

You may find that, after having installed the license file, the program still doesn't run and prints a message like 'your license file is corrupt'. To explain how this may come about, and how you overcome this, a few words on license files.

Each license file consists of pairs of lines. The first of each pair is text that states, in more or less readable format typical aspects such as an expiration date, the version number of the software and so on. The second line contains the same information in encrypted format: a (long) string of characters that seem to make little sense. The program reads the license file and checks, with its internal encrypting formulas, that

the two lines match. If not, it stops and prints the 'corrupt' message. So, there are two common reasons why it may happen to you:

- You are using a license file for another version of the software than your executables correspond to.
  Newer (major) releases may contain a different encrypting formula, so that the match in old license files
  is not recognized anymore.
  So, please verify that your license file and executable belong to the same major release.
- More likely: the license file as it has been created has been modified in some way.
  Sometimes, people inspect it and 'clean it up' a little bit, for instance by removing 'redundant' spaces,
  or by making some other 'improvements'.
  Unfortunately, every such modification will destroy the encryption match and lead to the 'corrupt' error.
  Most of the times, however, the reason lies in the mailing system, by which the license file has been sent
  to you. If the encrypted line is rather long, the mailer may have cut it in two shorter lines.
  To verify (and correct) this: edit the license file and see if it consists of pairs of lines as described above.
  If not, re-unify the broken lines and try again.
- Finally, the problem may lie in your O/S, which may have inserted additional hidden <CR> characters
  (Carriage-Return) into the license file. You can remove them with our fix_license utility
  (in $ADHFOME/Install), see the Installation manual.

Recover from Crash

A calculation may terminate in two ways: controlled or uncontrolled. Controlled termination includes cases where the program itself detects an error and decides that continuation of the calculation is impossible or pointless. In all such cases the standard exit routine is executed, resulting in an output section with some final information. This also ensures that the general result file TAPE21 is closed properly and all relevant information flushed to it.

Uncontrolled termination may occur, for instance when some bug causes the program to divide by zero, violate memory access restrictions, etc. Usually this leads to an immediate abort of the program by the Operating System and hence loss of control by the program. In such situations the information on TAPE21 may be incomplete because some of the data are kept in memory until the final termination of the program is carried out. It would be a terrible nuisance to see all time spent so far being lost. To remedy this ADF supports a check point file, named TAPE13, to help you recover at least some, if not most, of the results: not for analysis, but for continuation from a point not too long before the fatal condition occurred. TAPE13 can be used, just like TAPE21, as a normal restart file. See the restart key.

Memory Management

Problem: The program aborts with an error message "MEMORY ALLOCATION ERROR". This message is issued both in the logfile and in the output file.

Cause: Memory allocation may fail due to:

1. Insufficient virtual (i.e. total RAM + swap) memory
2. On Unix: too low values for per-process memory limits
3. Restrictions of the 32-bit architecture

Cure: Problem 1: add more physical RAM or increase the size of the swap space (page file).
Problem 2: add one or more ulimit commands to your run script setting relevant limits to "unlimited".
Problem 3: Perform your calculations on a 64-bit system. ADF version for the most common 64-bit operating systems are available so use them!

All the three problems above can be avoided by reducing the size of the calculation. The most important parameter defining the amount of used memory is the size of the basis set or, more precisely, the total number of Cartesian Slater functions, naos. Current value can always be found in the out file of the calculation, just search for the "naos" string. The amount of memory used by a particular calculation depends on the naos value and of the type of the calculation and, for large naos, it scales as $naos^2$. For example, a non-relativistic calculation during SCF can use up to 40 $naos^2$ bytes of memory. Using spin-orbit coupling may double this amount and using a hybrid or a meta-GGA XC functional will add extra on top of it. Also TDDFT calculations require additional memory.

What can be done to reduce memory usage? First of all, reducing the basis set size for non-critical parts of the molecule will reduce the memory requirement without reducing the quality of the results. Secondly, performing a calculation with a pure GGA instead of B3LYP will not only reduce the amount of memory used but also make the calculation faster. The latter especially applies to geometry optimizations because there B3LYP does not perform any better than some of the GGAs.

Note: If workspace problems occur for relatively small calculations, there might be a bug. Notify your ADF contact: send us the output file so that we can have a look and check things out.

SCF

SCF convergence problems can have various reasons. Thus, finding the reasons for a particular SCF behavior is half of solving the problem. You'll be surprised but the majority of SCF convergence problems are caused by an **unphysical calculation setup**, such as mistakes in the geometry or a too large negative charge. Thus, the first thing to do is to check if the geometry is really what it is meant to be. Check for too short interatomic distances, make sure the coordinates are specified in the right units. By default ADF expects atomic coordinates in Andgstrom so check that the coordinates are provided in these units. Also check that no atoms got "lost" when importing coordinates.

So, your calculation is set up correctly, but the SCF still does not converge. Before discussing other options, let's look at how the SCF process in ADF is organized. In a nutshell, it consists of the following steps:

1. A Fock matrix is constructed from the current density and the potential.
2. The Fock matrix is used in a DIIS procedure where it is mixed with some previous Fock matrices to construct a new one.
3. This new Fock matrix is diagonalized to obtain molecular orbitals (MOs).
4. MOs are populated by electrons following the *aufbau* principle or, if the KeepOrbitals feature is ON, by overlap with a previous density matrix.
5. A new density matrix is constructed from occupied MOs and the steps 1-5 are repeated.

Problems on any of the steps 1, 2, and 4 above can cause problems in the whole SCF process. Usually one can identify which step causes the problems by looking in the logfile and in the output file. In the logfile, two values are printed for each SCF cycle: ErrMat and MaxEl. Both values are related to commutator of the current Fock and density matrices, [F,P]. ErrMax is a sum of squares of the commutator matrix elements while MaxEl is its largest (by absolute value) element. Below, different SCF patterns will be discussed with suggestions on how to solve them.

By far the most common reason for non-converging SCF is a very small or absent HOMO-LUMO gap. This problem is most frequently observed for compounds containing d- and f-elements (transition and rare-earth metals). This causes different MOs to be populated in subsequent cycles at the step 4 above, which, in turn,

leads to large changes in the density and Fock matrices between cycles. In the logfile, the problem manifests itself by the ErrMat and MaxEl values remaining rather large (in the order of 0.1 to a few tens) sometimes going down to smaller values but then jumping back up. By looking at the MO population numbers in the output file one can sometimes see that the HOMO changes from cycle to cycle. There are a few ways to get the SCF converged in such a situation.

For open-shell electronic configurations, it is possible that a spin-unrestricted calculation will converge better than a spin-restricted one. Thus, if the molecule is not going to be used as a fragment (in which case it *must be* spin-restricted) then it is recommended to perform a spin-unrestricted calculation in a high-spin configuration. This is particularly useful for molecules with multiple radical centra, such bi- or multi-nuclear transition metal complexes. After a high-spin calculation of the complex has converged one can perform a broken-symmetry low-spin calculation using the high-spin results as a restart and a SpinFlip feature. See also a tutorial on spin coupling of an iron compound for an example of this approach.

If a spin-unrestricted calculation is not desirable or if it also has SCF convergence problems one may consider trying different DIIS methods in the step 2 above or try a completely different SCF method, preferrably in the order listed below:

- **A-DIIS** is a recently published alternative DIIS procedure that combines the strength of E-DIIS and ARH methods discussed below, but does not require (time-consuming) evaluation of the energy. By varying the two threshold parameters (see the link above) one can control the switching between A-DIIS and Pulay DIIS, which can get the SCF converged even in very difficult cases. To enable A-DIIS from the 1st SCF cycle, just add ADIIS to the SCF input block. A-DIIS is also automatically invoked when ADF detects SCF problems. In some rare cases, A-DIIS may fail while the Pulay DIIS converges, even if after many cycles. In this case adding NoADIIS to the SCF block will disable the automatic switching on the A-DIIS. Note: A-DIIS is not compatible with enforced non-aufbau electronic configurations it should be disabled in such a case. A non-aufbau electronic configuration may be enforced using a block form of the Occupations key, but it may also result from the KeepOrbitals (a.k.a. orbital tracking) feature. In both cases A-DIIS should not be used.
- **LISTi** is a method developed by Y.A. Wang and coworkers. LISTi frequently converges as well as A-DIIS, if not better. Like A-DIIS, this method also affects (replaces) the DIIS step only leaving other steps unchanged. In the LISTi method, the number of the expansion vectors (also known as the size of the DIIS space) is an important factor which may have to be adjusted to achieve convergence.
- **Energy-DIIS** by Scuseria and Kudin is a powerful method. It requires evaluation of the total energy, which is its strong point and a weakness at the same time. It is a strong point because it lets the algorithm converge the SCF to a configuration with the lowest energy and it is a weakness because energy evaluation is a computationally expensive procedure in ADF. To use Energy-DIIS just add an EDIIS keyword to the SCF input block.
- The **Augmented Roothaan-Hall (ARH)** method is an alternative SCF method that combines steps 2-4 above into a single step. Essentially, ARH performs a direct minimization of the energy as a function of the density matrix combining a preconditioned conjugate-gradient method with a trust-radius approach. This is probably the most powerful SCF method to date because it can converge even the most difficult cases. However, it also has its limitations and drawbacks discussed in the corresponding section of the ADF User's Guide. The most important drawback is that, like Energy-DIIS, it also requires evaluation of the total energy. Besides, for the method to work reliably, the energy must be accurate, which means that an accurate fit and high integration accuracy should be used.

As mentioned above, A-DIIS, LISTi and Energy-DIIS affect only the DIIS step of the whole SCF process. Thus, it is possible that the SCF still has trouble converging even with the best DIIS method, because different MOs are occupied in different SCF cycles, which induces large changes in the density. ADF has a built-in feature called KeepOrbitals that assigns electrons to MOs based on their overlap with occupied MOs from the previous SCF cycle. KeepOrbitals is usually enforced starting from the 25th SCF cycle. However, if there are SCF problems, switching on KeepOrbitals may not have the desired effect. For example, the system may be trapped in an excited state due to it. Changing the SCF cycle at which KeepOrbitals kicks in

may affect the final electronic configuration. Thus, playing with KeepOrbitals and trying different values for its parameter (the SCF cycle number) is encouraged.

For the DIIS-like methods listed above (A-DIIS and LISTi), the number of expansion vectors is an important parameter. We found that in some cases, for example, transition metal clusters, increasing this number to 20 or 25 solves most of the problems. Thus, the first thing to try when there are SCF convergence problems is to use the following input parameters:

```
SCF
  ADIIS
  DIIS N=20
  Iterations 300
END
Occupations KeepOrbitals=300
```

The input above turns A-DIIS on from the first SCF cycle, increases the number of expansion vectors to 20 and disables KeepOrbitals during the whole SCF process. With these settings, the previously non-converging transition metal clusters have converged in under 100 cycles.

Another trick that may help sometimes is allowing more cycles with simple damping before switching to DIIS. It is also recommended to use a smaller damping factor in this case. The following SCF settings are then recommended:

```
SCF
! The default value of 0.2 may be too high for some systems
! so we change it to 0.1
  Mixing 0.1
! The default value DIIS ok=0.5 might make DIIS kick in too soon
! thus making it unstable. The cycle starting from which DIIS is
! enforced may also be shifted forward.
  DIIS ok=0.01 cyc=20
! Set the max number of SCF cycles to 100 but you can use an
! even larger value.
  Iterations 100
END
```

Numerical noise in the exchange-correlation potential may also contribute to SCF convergence probems. If this is the case, the SCF process starts off converging quite well until some point after which the ErrMat and MaxEl values remain relatively small but do not decrease further. This behavior is typically observed for systems where weak (Van der Waals or hydrogen bonding) interactions are present. It is caused by a relatively low accuracy of the density fit in the chemically relevant region between weakly bonded fragments.

This problem can be resolved by adding the EXACTDENSITY keyword to the input file. When using ADFinput, the corresponding option called "Density used in XC-potential" found on the Accuracy tab should be set to "Exact" or "Exact MO-based", "Exact" being preferred.

Geometry Optimization

No convergence

First of all one should look how the energy changed during the latest ten or so iterations. If the energy is changing more or less in one direction (increasin or decreasing), possibly with occasional jumps, then there is probably nothing wrong with the optimization. This behavior is typical in the cases when the starting geometry was far away from the minimum and the optimization has a long way to go. Just increase the allowed number of iterations, restart from the latest geometry and see if the optimization converges.

If the **energy oscillates** around some value and the energy gradient hardly changes then you may need to look at the calculation setup. We give some recommendations below.

The success of geometry optimization depends on the accuracy of the calculated forces. The default accuracy settings are sufficient in most cases. There are, however, cases when one has to increase the accuracy in order to get geometry optimization converged. First of all, this may be necessary if you tighten the optimization convergence criteria. In some cases it may be necessary to increase the accuracy also for the default criteria. Here's what you can do to increase the accuracy of gradients:

- Increase the numerical quality to "good"
- Add an ExactDensity keyword or select "Exact" in the "Density used in XC-potential" list in the Details:Accuracy panel. This will make the calculation 2 or 3 times slower.
- Tighten the SCF convergence criteria, for example, to 1e-8.

Example input with some of these stricter settings using a TZ2P basis set. ExactDensity is not included here since it will make the calculation much slower.

```
NumericalQuality Good
Basis
  Type TZ2P
  Core None
End
SCF
 converge 1e-8
End
```

Small HOMO-LUMO gap: check the HOMO-LUMO gap at the last SCF cycle at a recent geometry. Is it comparable with the changes in the MO energies between geometries? If yes, then it is possible that the electronic structure changes between optimization steps, which may lead to non-convergence. This may indicate a fundamental problem with the calculation setup. First of all, check that you obtain a ground state in a single-point calculation. Is the spin-polarization value correct? Try calculating some of the high-spin states if possible and see if they have lower energy. If the MO repopulation that is taking place is between MOs of different symmetry, then you can try freezing the number of electrons per symmetry using an OCCUPATIONS block.

In the next two sections we've collected some troubleshooting tips that apply only to the new or only to the old branch. In the last section, we also address a very rare issue with very short bonds that may appear independently of the optimization branch.

New Branch

Are there **constraints** imposed on the optimization? If yes, then check that the constraints do not break the symmetry. ADF automatically preserves symmetry when the starting structure is symmetric and no *Symmetry NOSYM* has been specified. This symmetry preservation may conflict with constraints if the constraints are not specified according to symmetry. For example, if two interatomic distances are symmetry-equivalent and you freeze one of them, then you should also freeze the other explicitly.

Optimization in Cartesian coordinates usually needs more steps to converge compared to delocalized. If you haven't tried delocalized optimization for the system then you should do it.

Near 180-degree angles with optimization in delocalized coordinates. ADF normally does not have a problem with a near-180-degree valence angle if the initial value of the angle was larger than 175 degrees or if it is a terminal bond angle. If the initial angle was larger than 175 degrees then it gets a special treatment. For example, any torsion angle that contains the three atoms is removed or replaced with a torsion angle involving only the two end atoms of the near-180-degrees angle. If the angle defines a terminal bond then two angles in two coordinate planes are used to define the bond instead of a valence and a torsion angle. However, if the initial value of the angle was far from 180 degrees and has become close to it during

optimization then optimization may become unstable, especially if this is an angle connecting large fragments. In this case, it may be useful to restart geometry optimization from the latest geometry. As a last resort, you may want to constrain the angle to a value close, but not equal, to 180 degrees.

Old Branch

Spurious jumps

Problem: During geometry optimization in Z-Matrix coordinates, the atomic configuration makes a large unrealistic jump.

Possible cause 1: the triplet of atoms to which the current atom is related by the Z-matrix is (almost) co-linear. When, in a geometry step, the triplet passes through co-linearity, the dihedral angle for the current atom should make a discontinuous jump of 180 degrees. This is not checked in the program and the dihedral angle may not get corrected, resulting in a geometric jump of the atom (and hence of all atoms related to it by the Z-matrix).

Cure: Check the triplets of atoms, used in your Z-matrix to define the dihedral angles. If one of them is almost colinear, then redefine the Z-matrix or use *Cartesian* optimization.

Possible cause 2: the connectivity of the Z-matrix does not reflect some important bonds. Especially when the molecule contains (a large number of) rings, this badly affects the stability of the geometry update step. The reason is basically that computed *Cartesian* forces are transformed into changes of the curvilinear *internal* coordinates. The transformation between the two systems of coordinates is non-linear, but mathematically assumed to be linear. This is only a good approximation for small steps.

Cure: Redefine the Z-matrix and/or (if the geometry steps are very large) set a smaller upper bound on the maximum step (key GEOMETRY, subkey *step*).

Constraints are violated

Problem: constraints are violated: coordinates that were specified as frozen turn out to change during the optimization or coordinates that should remain the same start to differ after a few geometry update cycles.

Possible cause: there is an internal conflict between different demands, usually: symmetry versus constraints. The problem arises easily when a constrained optimization is requested for a molecule with some symmetry while the coordinates were defined with a Z-matrix structure that does not properly reflect the symmetry. Usually the deviations from the requested constraints are small. If they are really large, there might be a bug and you should contact ADF support.

Cure: redefine the Z-matrix and/or use Cartesian optimization (if the constraints are expressible in Cartesians).

Very short bonds

If the computed equilibrium geometry appears to exhibit unlikely values, typically significantly *too short* bond lengths, you may have run into a basis set problem, in particular (but not only) if the Pauli relativistic method is applied.

Problem: Optimized bond lengths are clearly too short. The energy may also look suspicious.

Possible cause 1: Basis set trouble: onset of Pauli variational collapse, if you have applied the Pauli relativistic option. Caused by small (or absent) frozen cores and/or relatively large basis sets, applied to heavy elements.

Possible cause 2: Basis set trouble also, but quite different from the previous potential cause: you have used relatively *large* frozen cores. When the atoms approach each other during the optimization and the frozen cores start to overlap, the energy computation and the computed energy gradients become more and more incorrect. This is a result of the inappropriateness of the frozen core approximation, which indeed assumes that frozen cores of neighboring atoms do not significantly overlap. Without going into a detailed explanation here, the net effect is that certain repulsive terms in the energy computation are missing and hence a spurious tendency to a 'core collapse' arises, yielding too short bond lengths.

Cure: Best is to abandon the Pauli method and use the ZORA approach instead for any relativistic calculation. If for whatever reason you insist on using the Pauli formalism, apply bigger frozen cores and, if that doesn't help, reduce the basis set (not by deleting polarization functions, but by reducing the flexibility of the occupied-atomic-orbitals space, in particular *s*- and *p*-functions). Note, however, that large frozen cores can be a cause for trouble by themselves, irrespective of any relativistic feature. If you have reason to believe that your frozen cores might be too *large*, given the resulting bond lengths in your calculation, you have to pick smaller cores (and hence be very wary of using the Pauli formalism for any relativity).

Frequencies

Imaginary Frequencies

Problem: totally unexpected significant imaginary frequencies are obtained (in a Frequencies run) where you are pretty convinced that all frequencies should be real.

Possible cause 1: problems with the electronic configuration. If there are competing configurations, the electronic *states* in the different displaced geometries may be different, resulting in energies and gradients belonging to different potential energy surfaces to be compared and combined into force constants (frequencies).

Check: orbital occupations and SCF convergence behavior: if the SCFs in the displaced geometries start with large errors and/or converge very slowly you are likely to have stumbled into different configurations, so that the results from the displaced geometries are incompatible.

Cure: This is a difficult situation that may require some experimenting and judicious manipulation of the various SCF options. The bottom line is that you should try anything you can to ensure that all involved geometries have the same electronic configuration. As long as you fail to achieve this, the results are meaningless.

Possible cause 2: flat potential energy surface (think about almost free rotation modes) coupled with relatively high noise level in gradients caused by numerical integration errors or not sufficiently converged geometry optimization.

Check: visualize the imaginary frequencies in ADFspectra and check that their respective normal modes correspond to movements that are expected to be have (nearly) flat energy profile.

Cure

:

- restart geometry optimization with more strict convergence criteria. The default criterion on gradients 0.001 Hartree/Angstrom may be not strict enough for some systems. In such cases a value of 0.0001 is recommended, and for accuracy reasons use "good" numerical quality, and EXACTDENSITY (important for GGA's).

Example input with strict settings using analytical frequencies, and a TZ2P basis set. Unlike the numerical frequencies, the analytical frequencies can be computed immediately after a geometry optimization by including both block keywords in the same input file.

```
Geometry
  converge grad=1e-4
End
AnalyticalFreq
End
NumericalQuality Good
Basis
  Type TZ2P
  Core None
End
```

Geometry-displacement numbers in the logfile are not contiguous

Problem: successive displaced geometries in the logfile are numbered, but in your case these numbers make sudden jumps, like '0, 1, 2, 5, 6, 13...'

Cause: you're using Cartesian displacements in a system that has some symmetry in its equilibrium geometry. The program skips the displacements of symmetry-equivalent atomic coordinates to save time. The displacement counts in the logfile do not run over the actually performed displacements but over all atomic coordinates that could be displaced if no use were made of symmetry properties.

Cure: there is no error, don't worry.

Input ignored

Problem: the program doesn't get past input and aborts with a message eof while reading (....). Or the program seems to ignore some parts of input and as a consequence goes wrong somewhere. Or it seems that part of the input has not been read correctly or not at all.

Cause 1. You have used tab characters in your input file. These are not normally visible when you edit your file, but they will affect the program's scanning of the input. When you use tab characters in the input, it is very likely that the program will do something wrong somewhere. Tabs may be ignored by the program, so that items that you believed were separate (by a tab!) are in fact read as contiguous.

Check: the input file on tab characters.

Cause 2: misusage of one of the block-type keys or general keys.

A case that relatively often shows up is typing a title as first line of the input file, *without preceding it by the keyword title*. The program does not understand this as the title, but rather tries to interpret the first word as a keyword. This leads to an error if the first word is recognized as one of the pre-defined block-type keys (possibly abbreviated).

Check: the input file on usage of block-type keys and on proper usage of a title.

Cause 3: incorrect processing of expressions or unintended replacement of names by numerical values. Various kinds of mis-typing or incorrect usage of variables may cause this.

Check: how the program sees input, *after parsing*. This can be done by rerunning the job, with as first line in input : print parser.

This will cause the program to copy each input line *twice* to output, the second time after having parsed it. You may use StopAfter Input or StopAfter Init to let the program quit early so you can inspect what is going on with the input reading.

SFO Populations

In the section that prints the SFO populations of (selected) MOs you may occasionally find, for some SFOs in some MOs, *negative* SFO contributions. This may seem unphysical and hence suspicious, but it is 'only' a result of the Mulliken-type analysis method that underlies the computation of the SFO contributions. See the section below that discusses the output file. Likewise for larger-than-100% contributions: don't worry too much, these numbers may be correct (mathematically, given the Mulliken population formulas).

Error Aborts

The program performs a large number of checks during the calculation and may stop when it detects and error. It is close to impossible to show here a complete list of all possible error messages. In a large number of cases, additional information is printed in the output file to provide a clue as to the cause of the error. It is always useful to carefully inspect the printed info and to try to understand the meaning of any error- or warning messages. If you can't find your way out, try to get help from your ADF provider. If that fails, contact us directly at support@scm.com

Warnings

The program attempts to detect bugs, instabilities, convergence problems, et cetera and may issue warnings when something looks suspicious. This is not necessarily fatal to your results, but you should be cautious and try to understand what the messages are about. Most warnings are printed in the logfile. Usually there is corresponding and more extensive information in the standard output file.

3.3 Questions

`Overlap matrix in BAS representation`

> How do I get the overlap (S) matrix in the BAS representation?
> It is stored on a scratch file TAPE15, which is normally deleted at the end of the calculation because it can be pretty big. To retain that file, insert 'SAVE TAPE15' in your input, see the Save key.
> TAPE15 is a KF file, which you can manipulate with the KF utilities. On TAPE15 the overlap matrix is stored as the variable 'smat' in the section 'Matrices', in reduced (triangular) format: (1,1), (1,2), (2,2), (1,3) et cetera

4 RESULTS

ADF produces two ASCII files: standard output and the log file. The latter is a very concise summary of the calculation's progress during the run. Furthermore, ADF produces and reads binary data files. Most of these files have the so-called KF format. KF stands for Keyed File: KF files are keyword oriented, which makes them easy to process by simple procedures. KF files are Direct Access binary files. Consult Appendix 5.5 for how to use some standard utilities for processing KF files.

4.1 Results on standard output

The (standard) output file contains information of the main characteristics of the run, the SCF and geometry optimization results, bonding energy and population analyzes. Major parts of output can be regulated with print switches, see the keys (NO)PRINT and EPRINT.

By default the program produces quite a bit of output, for a large part related to (Mulliken-type) population analyzes of the molecule in total, as well as of individual orbitals, both in terms of the elementary basis functions and in terms of the SFOs, the symmetry-adapted Fragment Orbitals.

The fragment-oriented approach of ADF is very suitable for a thorough chemical analysis of molecular orbital properties and a conceptual representation of results. New users are advised to spend time and get familiar with the SFO-type analysis. It is an extremely more powerful tool to understand the electronic structure of the molecule than the classical atomic orbital populations.

A summary of output is given below, assuming that default values apply for all print switches. Keep one of the Example outputs at hand when reading the description below.

Job Characteristics

Input Echo, Output Header

- Copy of the input file, except any InLine records: these are expanded and the contents of the inlinefile
  replaces the InLine command in the echo.
- Header with the program name, the release number and a copyright statement.
- Directly below the header are printed the job identification, title, and any comments that may have been
  supplied via input (key COMMENT).
  The job identification is comprised of the ADF release number and the date and time of the calculation.

Main Job Characteristics

- The Model Parameters such as the Density Functional and relativistic options.
- A list of attached files: restart data files and fragment files.
- The run type: Geometry Optimization, Frequencies...
- (Initial) geometric data: atomic positions, atom types, defined fragments, and the inter-atomic distance matrix.
- The point group symmetry, with a list of the irreducible representations and subspecies.
- The electronic configuration: occupation numbers (if specified), their distribution over spin-α and spin-β,
  and the net charge of the molecule.

Build Info: Fragments and Function Sets

See the print options eprint:*frag*, eprint:*sfo* and functions.

- Correspondence between fragments in the molecule and the corresponding *master* fragments on the pertaining fragment file. (This output is by default off)
- SFOs: the Symmetry combinations of Fragment Orbitals. The SFOs are the basic conceptual entities
  for the analysis of MOs and other results.
  Note: The FO *coefficients* that expand the SFOs are normalized in the sense that they add up (squared) to unity.
  The resulting SFO *function* is not necessarily a normalized function. The FOs are normalized, so it depends on the *overlap* between the FOs what the self-overlap and hence the norm of the SFO is.
  Also printed are, for each subspecies in each irrep separately, the indices of the elementary basis functions
  from which the FOs and hence the SFOs are built up. (The overlap matrix of SFOs is printed much later,
  in the SFO Populations section after everything (SCF, Geometry) has cycled to convergence).
- The elementary basis functions, fit functions, and the frozen-core levels of the atoms.
  First the lists of function *sets*, defined by radial behavior and the angular quantum number, are printed for all atom types on which the functions are centered. Thereafter follows the complete BAS
  list where the function sets have been expanded over all atoms (the *sets* are printed only for the atom *types*) and also over all Cartesian harmonics (6, not 5 *d*-functions, et cetera).
  In this printout the numbering can be found to which the SFO survey above refers.

Technical Parameters

See the PRINT key techpar.

- Parallelization and vectorization characteristics.
- *Direct* versus *Store-On-Disk* options.
- Update strategy parameters for Geometry updates (if applicable) and for the SCF procedure.
- General precision settings for numerical integration and neglect-of-small function values (in integral evaluations).

Computational Report

See the print switches computation, eprint:*numint*, eprint:*SCF*, eprint:*geo*.

`Numerical integration`

General grid-generating parameter(s) and the number of generated (symmetry unique) integration points, with their distribution over the distinct kinds of integration regions: the atomic (core-like) spheres, the remaining interstitial regions between the atoms (atomic polyhedra), and the outer region, i.e. the part of space around the molecule.

Partitioning of the points in blocks. In general there are too many integration points to have all pertaining data (values of basis functions in the points etc.) in memory. A segmentation in blocks of points is therefore applied, processing a block of data at a time after loading it from disk or recomputing it (depending on the *Direct* options). This also determines vector lengths and hence vectorization performance in numerical integral evaluations.

Integration Tests. The generation of the points involves an adaptive procedure to tune the point distribution such that a pre-set precision of several test integrals is achieved with a minimal number of points. The generated scheme is *a posteriori* tested by evaluating a few integrals in the actual molecule.

This does not result in any subsequent adaptation of the grid but only produces info for the user to verify that all goes well. If the results are suspicious a warning is issued and if the results are too bad, the program will abort.
The most important and significant test is the evaluation of the self-overlaps of all symmetry-adapted elementary basis functions. The maximum and root-mean-square (relative) errors are printed. This extensive testing is not carried out in Direct-SCF (bas) mode because in that case the necessary information is not available (basis functions are only computed when needed in the SCF).
A test that is always carried out is the numerical integration of the total frozen core density (summed over all atoms in the molecule). Also here a warning or even abort will occur when the result indicates that the integral has insufficient accuracy compared with the integration precision parameter.

`SCF procedure`

at each cycle: for each irreducible representation: the one-electron orbital energies and the occupation numbers for a contiguous sequence of orbitals. The indices of the lowest and highest MOs (in energy ordering) are printed directly after the irrep label. With this information you can check the electronic configuration. When convergence is problematic, more info appears at the higher iterations.
The involved orbitals are usually the highest few occupied and the lowest few unoccupied orbitals, see the eprint subkey eigval. During the SCF, as soon as the distribution of electrons over irreps is frozen, only the occupied orbital energies are computed and hence printed.
Also printed at each SCF cycle is the difference of the density matrix (P-matrix) with the previous cycle: the average and maximum difference in the diagonal elements.

At the end of the SCF: concise information about the density-fit precision: the error integral for the SCF density. The error integral is the integral of the difference between the exact density and the fit density, squared. Such values have very little to do with numerical integration, rather they show whether or not the employed set of fit functions are adequate to describe the SCF density. Error integral values that significantly exceed 1e-4 times the number of atoms are suspicious and may indicate some deficiency in the fit set for the actual calculation.
On the last geometry (in an optimization) the fit-error integrals are also printed (in the Results section, see below) for the initial (sum-of-fragments) density and the orthogonalized fragments (see Chapter 1.2)

- Gross atomic charges, computed from a Mulliken population analysis.
- Geometry Updates. The contents of this section depends on the RunType:
  Geometry Optimization, Frequencies.... It is absent in a Create run and in a SinglePoint calculation.
- Gradients on the atoms: derivatives of the energy w.r.t. changes in the nuclear coordinates.
- Summary of convergence issues. One of the items considered for convergence is the maximum Cartesian
  gradient. This value corresponds in principle to one of the Gradients on the Atoms.
  Differences may occur due to user-set and automatic constraints.
  The printed Gradients are the raw gradients, the maximum Cartesian gradient is the maximum over
  relevant gradients: this ignores gradients in frozen coordinates.
  Furthermore, gradients in coordinates that are forced to remain equal are averaged before the maximum
  is selected; finally the raw gradients are processed to eliminate spurious components such as gradients
  in rigid motions (translations and possibly rotations).
  In a Z-matrix optimization any user-set constraints apply to the Z-matrix coordinate-derivatives and the maximum Cartesian gradient is selected from the Cartesian gradients that are recomputed from the constrained z-matrix gradients.
- New coordinates: Cartesian and z-matrix if applicable. Optionally the new inter-atomic distance matrix
  is given (not by default).

The Computational info is repeated in all cycles (SCF and geometry) until the iterations have terminated.

Exit Procedure

normal termination or an error message.

A list of all files that are (still) open when the exit routine is called. The program closes such files at this point.

Information about buffered I/O processing during the calculation.

A check of workspace to see whether all dynamically allocated arrays have been cleaned-up. If so the program mentions All Arrays Delocated. Otherwise there is something wrong and the situation will be summarized. If the calculation seems to have completed normally, but nevertheless workspace has been found not-clean, we would appreciate to get the complete output file because it might signal a programming error. This does not apply when you have used the stopafter feature: the program will then abort before the standard termination and usually not all workspace will have been cleaned up then.

Timing Statistics: a survey of cpu, System (I/O) and Elapsed times spend in various sections of the program.

Logfile

At the end of the calculation the log file is copied (optionally, see print) to the tail of the standard output file. The log file contains a concise summary of the run.

Nuclear and Electronic Configuration

- The final atomic coordinates (only in an optimization run).
- One-electron orbital data: occupation numbers and energies, HOMO and LUMO energies and, if applicable, a list of partially occupied MOs.
- Orbital energies of the Core Orbitals

The direct results from the SCF are the orbital energies and occupation numbers. This defines the electronic configuration: the occupation numbers and HOMO and LUMO energies for instance show whether or not the aufbau principle is satisfied in the final situation.

The energies of the Core Orbitals can be used to interpret for instance XPS (X-ray Photoelectron Spectroscopy) data: from Koopman's theorem these core orbital energies are an approximation to the core ionization energies. This neglects the effect of relaxation upon the ionization so that absolute energy values may not be very good; relative values, however, should be fair and can therefore be used to study (relative) chemical shifts.

Structure and Reactivity

Summary of LT or IRC path(s)

At the very end of the results section, a completed LT or IRC calculation will show tables of a few key properties in each point of the scanned path: atomic coordinates, energy, dipole moment, atomic charges and a few others, depending on the case. This gives you a quick survey of the computed profile.

Frequencies Results

In a Frequencies calculation the computed harmonic frequencies are printed. If a complete variation of coordinates has taken place, the program will compute the frequencies and normal modes also in terms of Symmetry Coordinates, along with the representation in the coordinates that were specified in input.

The zero-point energy is printed, computed as sum over frequencies:

$E_0 = \sum v/2$ (4.2.1)

Any imaginary frequencies (printed in the output file as *negative* frequencies) are not included in the summation.

Thermodynamic properties (Heat Capacity, Entropy, Internal Energy) are printed, based on the ideal gas approximation. Electronic contributions are omitted. These are small when the energy gap with the next electronic configuration is large compared with the vibrational frequencies. For (near) degenerate configurations this assumption is incorrect.

Imaginary frequencies and very small frequencies are ignored in this calculation.

Spectroscopic Properties

The results for the spectroscopic properties that are printed are meant to be self-explanatory. See also the input options for each spectroscopic property.

Analysis

Mulliken populations

Mulliken populations are based on the elementary atomic basis functions (bas). The individual BAS populations are printed together with summaries of the populations in all basis functions with the same angular moment quantum number on the same atom.
A final summary is obtained by adding all functions on each atom, yielding the atom-atom populations. The atom-atom populations per l-value can be obtained if the key EXTENDEDPOPAN is included. The atomic gross charges are derived from the net and the overlap populations in the usual way.
In addition, a population analysis may be given of individual MOs (by default this is suppressed). See the EPrint keys SCF (option mopop) and orbpop.

Mulliken-type populations are computed and printed at various levels of refinement (ranging from *per-basis function* to *per-fragment type*, data for the whole molecule as well as for individual MOs), and in two different representations, one based on the elementary basis functions (bas), the other on SFOs (Symmetrized Fragment Orbitals). This is potentially a very large amount of data. Precisely what is printed by default, and how this can be modified so as to suppress output or, alternatively, to get more information, is regulated by the print keys (print, eprint).

Hirshfeld charges, Voronoi deformation density

Mulliken populations can be summarized to yield atomic charges. Alternative methods exist to deduce atom charges from the self-consistent results of a molecular calculation. Several of those alternatives are provided by ADF: Hirshfeld analysis, Voronoi analysis, multipole derived charges, and charge model 5.

Of the methods applied in ADF to compute charges (Mulliken, Hirshfeld, Voronoi) we recommend the Hirshfeld analysis [125, 126] and the analysis based on Voronoi *deformation* density (VDD) charges [109, 127], see below. The fragments to which the Hirshfeld charges apply are enumerated in the early geometry part of the output file, where for each fragment the numbers of the atoms are given that belong to the fragment. The sum of the Hirshfeld charges may not add up to the analytical net total charge of the molecule. Any deviation from this is caused by numerical integration precision (small effect) and the neglect

of long-distance terms that ADF uses to speed up the integral evaluations. This approximation does not affect very much the energy and molecular orbital properties, but it does show up in the sum-of-charges somewhat more. It does not indicate an error (unless the deviation is really large, say in the order of 1‰ of the total number of electrons).

The Hirshfeld analysis produces a charge value per fragment, computed as the integral of the SCF charge density over space, in each point weighted by the relative fraction of the (initial) density of that fragment in the total initial (sum-of-fragments) density:

$$Q^{\text{frag(i)}} = \int \rho^{\text{SCF}} \rho^{\text{initial frag(i)}} / (\sum_j \rho^{\text{initial frag(j)}}) \quad (5.1.1)$$

The VDD method is based on the *deformation* density and a rigorous partitioning of space into non-overlapping atomic areas, the so-called Voronoi cells [109, 127, 128]. The Voronoi cell of an atom *A* is the region in space closer to nucleus *A* than to any other nucleus (cf. Wigner-Seitz cells in crystals). The VDD charge of an atom *A* monitors the *flow* of charge into, or out of the atomic Voronoi cell as a result of 'turning on' the chemical interactions between the atoms. The VDD method summarizes the three-dimensional deformation density on a per-atom basis. It is conceptually simple and affords a transparent interpretation based on the plausible notion of charge redistribution due to chemical bonding, i.e. the gain or loss of charge in well-defined geometrical compartments of space. For the use of VDD in analyzes involving molecular fragments, see Ref. [129].

In the same fashion as for the Hirshfeld analysis, a summation over all atoms is given which should yield zero (for a neutral molecule). The deviation from zero is caused by numerical integration and by neglect-of-long-distance-terms; the same remarks apply as for the Hirshfeld analysis above.

The partitioning of space, using mid-way separation planes, is inappropriate to produce useful absolute numbers when neighboring atoms have very different sizes, for instance, Hydrogen and a heavy metal. However, *changes* in the density analyzed in this way do give a reasonable general insight in the effect of bonding on the location of charge densities, in particular because the Voronoi data per atom are split up in contributions within the atomic sphere and the rest of its Voronoi cell.

Hirshfeld and Voronoi charge analyzes are printed at the end of the SCF (of the last geometry, in case of an Optimization).

The Hirshfeld analysis in ADF produces charges *per fragment*, so that *atomic* charges are obtained only if single-atom fragments are used. This limitation does not apply to Voronoi charges (data per atom). Mulliken charges are given both per atom *and* per fragment.

In the printout of charges per fragment (as for the Hirshfeld analysis), you have to be aware of the *ordering* of fragments. A complete list of fragments is printed in the early GEOMETRY section of standard output, where you also find which atom(s) correspond(s) to which fragment. Note that even when you use single-atom fragments only, the order of fragments is usually quite different from the order of atoms in your input file. Typically (but not necessarily exactly in each case), when you use single-atom fragments: consider the first non-dummy atom in your ATOMS block. This defines the first atom *type*. Then browse the ATOMS list until you find an atom of a different type. This defines the second atom type, and so on. The single-atom fragment list will often be such that you first get *all* atoms of the first atom type, then all atoms of the second type, and so on. Check the printed list-of-fragments always, to avoid mistakes in assigning Hirshfeld charges to atoms (fragments).

Multipole derived charges

The multipole derived charges (MDC) analysis [170] uses the atomic multipoles (obtained from the fitted density) up to some level X, and reconstructs these multipoles exactly (up to level X) by distributing charges over all atoms. This is achieved by using Lagrange multipliers and a weight function to keep the multipoles local. Since the atomic multipoles are reconstructed up to level X, the molecular multipoles are represented

also up to level X. The recommended level is to reconstruct up to quadrupole: MDC-q charges. The SCF should have converged for a meaningful MDC analysis.

Charge model 5

The charge model 5 (CM5) [378] uses the Hirshfeld analysis in combination with a parametrization to yield atomic charges that can accurately reproduce dipole moments obtained from experimental results. For input, use the keyword CM5.

Bond order analysis

The Mayer bond order between two atoms is calculated from the density and the overlap matrices (key EXTENDEDPOPAN), see Ref. [140].

The bond order analysis with the key BONDORDER. produces the output in which the bond order values are printed for each pair of atoms for which the Nalewajski-Mrozek bond order value is larger than the threshold that can be specified with the keyword BONDORDER. For convenience the printed bond orders are accompanied by the corresponding inter-atomic distance. In the Nalewajski-Mrozek approach [148-153] the bond order indices b_{AB} are calculated based on the one- and two-center valence indices

$$b_{AB} = V_{AB} + w^{AB}_{A} V_A + w^{AB}_{B} V_B$$

with the weighting factors for one-center indices given by

$$w^{XY}_{X} = V^{cov}_{XY}/\sum_Z V^{cov}_{XZ}$$

Unlike other definitions of covalent bond orders, the Nalewajski-Mrozek valence indices comprise both, covalent and ionic contributions. There exist three alternative sets of the Nalewajski-Mrozek valence indices, [148-153, 140]. The bond order indices calculated from each set of the valence indices differ slightly due to arbitrariness in the way of splitting the one-center terms between bonds. More detailed description of alternative valence indices and their physical meaning is summarized in [148]; see also original papers [149-153]

By default the bond order indices based on the valence indices obtained from partitioning of Tr(PΔP) are printed in the ADF output. Note that in this version the covalent two-center part (also printed in the output) is equal to the Gopinathan-Jug [153] bond order. The default values are:

$$V_A = V^{ion}_{A} + V^{cov}_{A}$$

$$V^{ion}_{A} = \sum_{a\in A} \{P^{\alpha}_{aa}\, \Delta P^{\alpha}_{aa} + P^{\beta}_{aa}\, \Delta P^{\beta}_{aa}\}$$

$$V^{cov}_{A} = 2 \sum_{a\in A} \sum_{a'\in A, a<a'} \{P^{\alpha}_{aa'}\, \Delta P^{\alpha}_{a'a} + P^{\beta}_{aa'}\, \Delta P^{\beta}_{a'a}\}$$

$$V^{cov}_{AB} = 2 \sum_{a\in A} \sum_{b\in B} \{P^{\alpha}_{ab}\, \Delta P^{\alpha}_{ba} + P^{\beta}_{ab}\, \Delta P^{\beta}_{ba}\}$$

To produce the values from all alternative versions of Nalewajski-Mrozek valence indices, accompanied by the Gopinathan-Jug [153] and Mayer [140] bond orders, see the keyword BONDORDER.

The Mayer [140] bond orders can also be calculated using the keyword EXTENDEDPOPAN. The two implementations of calculating the Mayer bond orders differ slightly if one uses frozen cores. They should agree exactly in all electron calculations.

Dipole moment, Quadrupole moment, Electrostatic potential

Dipole moment. Note that in a ion the value of the dipole moment depends on the choice of the origin, as follows from elementary electrostatic theory.

Quadrupole moment. Note that the value of the quadrupole moment often depends on the choice of the origin, as follows from elementary electrostatic theory.

Electrostatic potential at the nuclei: the Coulomb potential of the molecule at the nuclear positions, where the contribution from the nucleus itself is omitted.

MO analysis

`MOs expanded in SFOs`

This gives a useful characterization of the character of the self-consistent molecular orbitals. Additional information is supplied by the SFO population analysis, see below.
The definition of the SFOs in terms of the Fragment MOs has been given in a earlier part of output (section build). The SFO occupation numbers that applied in the fragments are printed. This allows a determination of the orbital interactions represented in a MO.
Be aware that the bonding/antibonding nature of a SFO combination in a mo is determined by the relative signs of the coefficients *and* by the overlap of the SFOs. This overlap *may be negative!* Note also that SFOs are generally *not* normalized functions. The SFO overlap matrix is printed later, in the SFO-populations part below.

`SFO population analysis`

For each irrep:

- Overlap matrix of the SFOs. Diagonal elements are not equal to 1.0 if the SFO is a linear combination of two or more Fragment Orbitals. The Fragment Orbitals themselves are normalized so the diagonal elements of the SFO overlap matrix give information about the overlap of the Fragment Orbitals that were combined to build the SFO.
- Populations on a per-fragment basis for a selected set of MOs (see EPrint, subkey *OrbPop*). This part is by default *not* printed, see EPRINT subkey *SFO*.
- SFO contributions per MO: populations for each of the selected MOs. In these data the MO occupation numbers are not included, so that also useful information about the virtual MOs is obtained. The printout is in matrix form, with the MOs as columns. In each printed matrix a row (corresponding to a particular SFO) is omitted if all populations of that SFO are very small in all of the MOs that are represented in that matrix. See eprint, subkey *orbpop*.
Note that this method to define SFO populations (for orbitals) is very similar to the classical Mulliken type analysis, in particular regarding the aspect that *gross* populations are obtained as the diagonal (*net*) populations plus half of the related off-diagonal (overlap) populations. Occasionally this may result in negative (!) values for the population of certain SFOs, or in percentages higher than 100%. If you have such results and wonder if they can be right, work out one of the offending cases by hand, using the printed SFO overlap matrix and the printed expansion of the MOs in SFOs to compute 'by hand' the population matrix of the pertaining MO. To avoid doing large calculations it is usually sufficient to take only the few largest MO expansion coefficients; this should at least qualitatively give the correct outcomes.
- Total SFO gross populations in a symmetry representation: from a summation over all MOs (not only those analyzed in the previous section of output) in the symmetry representation under consideration. In

the gross populations the MO occupation numbers have been included.
- (Per spin): A full list of all MOs (combining all symmetry representations), ordered by energy, with their most significant SFO populations. Since there might be several significant SFO populations for a particular MO, and an SFO may actually be a linear combination of several (symmetry-related) Fragment Orbitals, this table could get quite extensive. In order to confine each SFO population specification to one line of output, the SFOs are indicated by the characteristics of the first term (Fragment Orbital) of its expansion in Fragment Orbitals. So, if you see the SFO given as the '2 P:x on the first Carbon fragment', it may actually refer to the symmetry combination of, for instance, 2P:x and 2P:y orbitals on the first, second and third Carbon fragments. A full definition of all SFOs in terms of the constituting Fragment Orbitals is given in an early part of the output.

Bond energy analysis

The bond energy and its decomposition in conceptually useful terms: Pauli (exchange) repulsion, total steric repulsion, orbital interactions (partitioned into the contributions from the distinct irreducible representations), and corrections for some approximations (fitting and Transition State analysis procedure).
For a discussion of bonding energy decompositions and applications see e.g. [3, 110, 112, 130-136]

The program prints the bonding energy (not in a Create or Frequencies run) and its decomposition in terms that are useful for chemical interpretation. The *total* energy is not computed. The bonding energy is defined relative to the fragments. When *basic atoms* are employed as fragments one should realize that these do not represent the atomic ground state since they are computed as spin-restricted and spherically symmetric objects, with possibly fractional occupation numbers. The correct multiplet state is not computed. To obtain the bonding energy with respect to isolated atoms you should therefore add atomic correction terms to account for spin polarization and the multiplet state. See also the SLATERDETERMINANTS key and the discussion on multiplet states.

The spin polarization energy can be computed by running the single atom unrestricted, using as fragment the corresponding (restricted) basic atom. The true multiplet state is not necessarily obtained in this way.

For the comparison of computed bonding energies with experimental data one should furthermore be aware of any aspects that are not represented in the computational formalism, such as zero-point motions and environment (solvent) effects.

In a Geometry Optimization or Transition State search, the program may print a bonding energy evaluation at each geometry (depending on print switches). A test-energy value is written in the log file. This is *not* the bonding energy, although the difference is usually small. The test-energy printed in the log file is the energy expression from which the energy gradients are computed. The true bonding energy contains in addition a few (small) correction terms that are mostly related to the fit incompleteness. These correction terms are usually very small.

If Electric Fields are used in the computation (homogeneous and/or point charges), the printed Bonding Energy is the energy of the molecule in the field minus the energy of the fragments in the same field. The energy terms due to the field are also printed separately so that one can subtract them from the total bonding energy to obtain the energy-change without field-terms.

4.2 Log file, TAPE21, TAPE13

Log file

The log file (logfile) is generated during the calculation and flushed after (almost) each message that is sent to it by the program. Consequently, the user can inspect it and see what is going on without being delayed by potentially large system I/O buffers. Each message contains date and time of the message plus additional info.

A major part of the messages simply states the name of a procedure. Such messages are sent when the procedure is entered. During the SCF procedure, the SCF errors, which are a measure for non-self-consistency, are written at every cycle. In calculations where the geometry is changing (optimization, frequencies...) each set of new coordinates is sent to the log file (Cartesian, in angstrom and also Z-matrix, if a Z-matrix structure was provided in the input file). Other messages should be self-explanatory.

Be alert on error messages. Take them seriously: inspect the standard output carefully and try to understand what has gone wrong. Be also alert to warnings. They are not necessarily fatal but you should understand what they are about before being satisfied with the results of the calculation. Do not ignore them just because the program has not aborted: in some cases the program may not be able to determine whether or not you really want to do what appears to be wrong or suspicious. If you believe that the program displays erratic behavior, then the standard output file may contain more detailed information. Therefore, in such case save the complete standard output file, together with the logfile, in case we need these files for further analysis.

TAPE21

TAPE21 is the general result file of an ADF calculation. It is a KF file: Direct-Access, binary, and keyword driven. It contains information about the calculation. You can use it as a fragment file in a subsequent calculation on a bigger molecule, where the current one may be a part, or in an analysis program. For more information on TAPE21, see Appendix 5.4.

TAPE13

TAPE13 is the checkpoint file for restarts after a crash. It is a concise version of TAPE21, containing only the data the program uses for restarting the calculation. See the RESTART keyword. Like TAPE21, TAPE13 is a binary, keyword driven KF file. For more information on TAPE21, see Appendix 5.4.

4.3 ADF-GUI

The graphical user interface ADF-GUI enables all users to set up complicated calculations with a few mouse clicks, and provides graphical representations of calculated data fields, see the ADF-GUI overview tutorials, and advanced ADF-GUI tutorials.

4.4 Densf: Volume Maps

densf is an auxiliary program to generate values of molecular orbitals, charge densities and potentials in a user-specified grid, to be used typically for plotting or graphical display. The TAPE41 result file can be used directly by the ADFview program to visualize these properties.

densf requires an ascii input file where the user specifies the grid and the items that he/she wishes to see calculated on the grid, plus the standard result file TAPE21 from an *adf* calculation. *densf* writes a summary of the items that have been requested to standard output, together with some general information.

densf produces a (binary) KF file TAPE41, see OUTPUTFILE keyword below. TAPE41 is a KF file and all KF utilities can be used to inspect and process its data.

Furthermore, TAPE41 can be processed by *cntrs* to generate contour plot data. The ADFview program can be used to view the data available in TAPE41 in a variety of ways. *Cntrs* is a separate utility program, see later sections in this documentation. For the ADFview program separate documentation is available.

Starting from ADF2007, *densf* can also read and write cube files. See the CUBINPUT and CUBOUTPUT input options for details.

Examples of using *densf* are contained in the set of sample runs; see the *Examples* document.

Input

The input for *densf* is keyword oriented. The keywords may be specified in any order with one exception: INPUTFILE, if present, must be specified before any other option. Reading input by *densf* ends when it encounters the record EndInput or the end-of-file, whichever comes first.

The current version of *densf* does have reasonable defaults for all input. That means that in many cases you probably will not need to specify any input at all.

Below follows a list of the allowed keywords with their description.

```
$ADFBIN/densf << eor
INPUTFILE {file}
OUTPUTFILE {file}
VTKFILE {file}
CUBINPUT {file}
CUBOUTPUT {file}
GRID ...
UNITS ...
Density ...
KinDens ...
Laplacian ...
DenGrad ...
DenHess ...
Potential ...
Orbitals ...
NOCV ...
NCI ...
SEDD
eor
```

Input/Output files

```
INPUTFILE {file}
```

INPUTFILE keyword specifies path to the TAPE21 file from which *densf* reads the input data. Absence of the keyword is treated as if **INPUTFILE TAPE21** has been specified.

```
OUTPUTFILE {file}
```

OUTPUTFILE keyword specifies path to the (possibly existing) TAPE41 file. If the file exists, *densf* will read grid specifications from it ignoring GRID keyword in the input. Computed quantities are saved in the file overwriting existing data with the same name, if any.

```
VTKFILE {file}
```

VTKFILE keyword specifies path to a file in the format readable by VTK directly. This option exists primarily for better integration with ADF-GUI and the user should not specify it.

```
CUBINPUT {file}
```

If the CUBINPUT keyword is present then the grid as specified in the **file** is used to calculate all requested quantities. Any volume data found in the cube file is also saved in the output file. NOTE: CUBINPUT option cannot be used with a pre-existing TAPE41 file because they both specify the grid, which may lead to a conflict.

```
CUBOUTPUT {file}
```

Presence of the CUBOUTPUT keyword tells densf to save all computed quantities as cube files using **file** as filename prefix. The prefix can also contain a complete path including directories. For example, specifying the following in the densf input

```
CUBOUTPUT /home/myhome/H2O
Density SCF
```

will result in a file /home/myhome/H2O%SCF%Density.cub being created containing volume data for the total SCF density. One file per requested quantity is created.

The OUTPUTFILE, CUBOUTPUT and VTKFILE options are mutually exclusive. Absence of any of these options is treated as if **OUTPUTFILE TAPE41** has been specified.

Grid

The Grid key is available either as simple key, or as block key.

The simple key options are as follows:

```
GRID {save} {coarse|medium|fine}
```

If the word save is specified, the program will store all grid points on TAPE41 (in addition to the specification of the grid that is always stored). The default is NOT to store all grid points.

Either coarse, medium or fine may be specified. This instructs the program to generate the grid automatically within a box enclosing all atoms of the molecule. The distance between grid points is 0.5, 0.2 or 0.1 bohr for respectively a coarse, medium or fine grid. Evidently the size of the result file TAPE41 depends strongly on this specification. The default value (used when the user does not specify the grid) is to generate a coarse grid.

If GRID is used as a block key it must be followed by the word end in a later record. The records until the end are the data for the Grid keyword:

```
Grid {save}
 x0, y0, z0
 n1, n2, n3
 v1x, v1y, v1z, length1
 v2x, v2y, v2z, length2
 v3x, v3y, v3z, length3
END
```

If the word save is specified, the program will store all grid points on TAPE41 (in addition to the specification of the grid that is always stored). The default is NOT to store all grid points.

The records in the data block must contain (*in the order specified below!*):

- 1 Three coordinates for the 'origin' (lower-left corner) of the grid.
- 2 Three integers: the numbers of points in three independent directions.
  If fewer integers are supplied the grid will accordingly be less-dimensional.
- 3 Three records each containing the coordinates for the direction of the independent vector (size irrelevant)
  and the total length of the grid in that direction.

If a lower-dimensional grid is requested (see item #2), then fewer such direction-records are read and
the redundant ones, if any, are ignored.
The unit of length in which the grid size is input is by default Angstrom.
The default can be overridden by using the input key UNITS, see below.

Notes:

- The second record ('three integers...') specifies the number of grid *points* in the different directions.
  The corresponding number of steps or intervals is one less!
- If the TAPE41 result file is to be used by the contour generating program *cntrs*,
  the grid used in the *densf* calculation must be *two*-dimensional.
- If the TAPE41 result file is to be used by ADFview, the grid used must be an three-dimensional orthogonal grid,
  with a single step size for all three dimensions.
- If the output TAPE41 file already exists and it contains valid grid data or if CUBINPUT is specified then the GRID input is ignored.
- The unit of length used in the input file has no relation to how the data are stored on the result file and how the program processes the data internally.
  Internal processing and storage on file is in bohr (atomic units).

Inline Grid

DENSF can now read grid as list of points. When specifying inline grid the GRID keyword should look as follows:

```
Grid Inline
  x1 y1 z1
  x2 y2 z2
  ...
  xN yN zN
End
```

Here, x#, y#, and z# are coordinates of points at which requested properties will be calculated. This feature may be used, for example, by external programs to calculate various properties at a number of points exactly and avoid interpolation with its inaccuracy. This feature should be used only when the output file has a TAPE41 format. NOTE: the coordinates must be specified in atomic units regardless of the value of units of length below.

Units

As in the ADF main program, the unit of length can be set with the block type key

```
UNITS
 Length unit_of_length
END
```

In *densf* the only item that can be specified in the UNITS block is the length, so it seems a bit pointless to make UNITS a *block* type key rather than a simple key. However, to make its usage identical to the application in the *adf* main program the block form has been chosen to apply also here. The unit-of-length will apply to the grid specification in the input file. Default is angstrom, except for inline grid where it is always Bohr and cannot be changed.

Density

Generates the charge density in the grid. It is a simple keyword (not block-type).

```
density {fit} {frag} {ortho} {scf} {trans}
```

Occurrence of the word fit specifies that all densities specified in this record will be computed from the fit functions (an approximation to the exact density), rather than from the occupied molecular orbitals.

frag, ortho, scf, and trans causes each of the corresponding densities to be computed. frag stands for the sum-of-fragments (i.e. the initial) density, scf for the final result of the *adf* calculation, ortho for the orthogonalized fragments (orthogonalization to account for the Pauli repulsion, see the ADF User's Guide), and trans for excitation transition density.

Transition density is a product of initial and final states of an excitation. In the simplest case when initial and final states consist of one molecular orbital each, in this case the corresponding transition density is a product of the two MOs. To otain transition densities one needs to perform an excitations calculation with ADF, see EXCITATIONS keyword in ADF User's Guide. Transition densities for all excitations found in the input TAPE21 file will be calculated. The transition densities are always fit-densities.

If both the exact and the fit-densities are required the density keyword must be repeated, once with and once without the fit option specified.

The default (when the DENSITY key does not occur in the input file) is to calculate the final SCF density and the sum-of-fragments density.

The frozen core density is calculated with:

```
density core
```

Kinetic Energy Density and Electron Localization Function (ELF)

```
KinDens {frag} {orth} {scf}
```

Generates the Kinetic energy density and electron localization function on the grid.

Occurrence of any of the words requests calculation of the two quantities (KinDens and ELF) based on the corresponding density: sum-of-fragments, orthogonalized fragments, or SCF, respectively. If none of the options is present, scf is assumed.

Laplacian of the Density

The Laplacian of the exact SCF density is calculated with:

```
Laplacian
```

The Laplacian of the fitted SCF density is calculated with:

```
Laplacian fit
```

The LAPLACIAN key can occur multiple times. The LAPLACIAN feature is also supported by ADFview.

Gradient of the Density

The gradient of the exact SCF density is calculated with:

```
DenGrad
```

The gradient of the fitted SCF density is calculated with:

```
DenGrad fit
```

The gradient of the frozen core density is calculated with:

```
DenGrad core
```

The DENGRAD key can occur multiple times. This feature should be used only when the output file has a TAPE41 format.

Hessian of the Density

The hessian of the exact SCF density is calculated with:

```
DenHess
```

The Hessian of the fitted SCF density is calculated with:

```
DenHess fit
```

The Hessian of the frozen core density is calculated with:

```
DenHess core
```

The DENHESS key can occur multiple times. This feature should be used only when the output file has a TAPE41 format.

Potential

Generates the coulomb and/or exchange-correlation potential in the grid.

```
potential {coul / XC} {frag} {ortho} {scf}
```

frag, ortho, and scf are as for the density: at least one must be specified.

coul and XC specify that the Coulomb potential, respectively the exchange-correlation potential must be computed. Precisely one of these options must be specified in the record. If both potential types are required, another input record with the potential key must be used.

In the present release the xc option is not yet operational.

The default (when the POTENTIAL key does not occur in the input) is to calculate the SCF Coulomb potential.

Orbitals

A block type key in which the required molecular orbitals are specified. The key can be repeated in input any number of times; all occurrences are read and applied.

```
Orbitals type
 (data)
END
```

The argument of the orbitals key (type) must be scf (for the scf orbitals) or frag (for the fragment orbitals) or loc (for the localized molecular orbitals, see the ADF User's Guide) or generic (see separate section).

The frag option is not operational in the present release.

In many data records in the ORBITALS block, as noted in the description of these data records, you may specify a HOMOLUMO range.

A HOMOLUMO range is the following:

```
   {HOMO{{-}n}} {LUMO{{+}n}}
```

HOMO: the highest occupied orbital
HOMO-n, with n an integer: the highest (n+1) occupied orbitals

LUMO: the lowest virtual orbital
LUMO+n, with n an integer: the lowest (n+1) virtual orbitals.

The HOMO part, or the LUMO part, or both must be specified. The integer n with sign is always optional, and the sign is always optional (and has no meaning, it is intended to enhance readability).

Thus, as an example,

```
    HOMO-1 LUMO+1
```

means a range of 4 orbitals: the two highest occupied ones, and the two lowest virtuals.

Each data record in the orbitals block must have either of the following formats:

1. the word alpha or beta.
This specifies that subsequent records refer to spin-alpha or spin-beta orbitals respectively. In a restricted calculation this has no meaning and beta must not be specified.
alpha and/or beta may occur any number of times in the orbitals block. All records until the first occurrence of alpha or beta are assumed to refer to spin-alpha orbitals.
2. label n1, n2, n3, ...
label is one of the subspecies of the point group symmetry used in the *adf* calculation and n1 etc. are indices of the molecular orbitals (in that subspecies) that are to be computed. This format is meaningless and must not be used for the loc orbitals type, because *localized* orbitals do not (necessarily) belong anymore to a particular symmetry representation.
3. label HOMOLUMO
label is one of the subspecies of the point group symmetry used in the calculation, the orbitals follow from the HOMOLUMO range.
4. label occ or **label virt**
occ specifies all orbitals (in that symmetry representation) up to and including the highest occupied one. virt specifies all orbitals above the highest occupied one. In this context *partially* occupied orbitals are considered occupied. Note carefully that if in a particular symmetry representation an empty orbital is computed below the highest occupied one in that same representation (excited state), that particular empty one is included in the list of occ.
Again, this format is meaningless and must therefore not be used for the loc type of orbitals.
5. all occ or **all virt** or **all HOMOLUMO**
Specifies *for each symmetry representation:*

- all orbitals up to and including the highest occupied one (in that symmetry), or
- all orbitals above the highest occupied one, or
- all orbitals defined by the HOMOLUMO range.

This form is not to be used for the LOC type of orbitals. However, using this for LOC will not result in an error but will be interpreted as identical to the following format.
6. all
This format must be used only for the LOC type of orbitals and simply means: all computed localized orbitals (irrespective of occupation numbers).
7. n1, n2, ...
a simple list of integer indices. This format must be used only for the loc type of orbitals since no reference is made to any symmetry representation. The indices refer of course to the list of localized orbitals as computed by *adf*, see the *User's Guide*.

The default value used when the ORBITALS key is not present is:

```
Orbitals SCF
 All HOMO-1 LUMO+1
End
```

NOCV

In ADF2009.01 it is possible to use DENSF to calculate $\varepsilon^*\varphi^2$ values of Natural Orbitals for Chemical Valence (NOCVs). Additional information on NOCVs is available in Ref. [335].

The relevant part of the DENSF input is as follows:

For spin-unrestricted:

```
NOCV
  Alpha
   N1α
   N2α
   ...
  Beta
   N1β
   N2β
   ...
END
```

For spin-restricted:

```
NOCV
   N1
   N2
   ...
END
```

N1, N2, etc. specify sequential numbers of the orbitals for which $\varepsilon^*\varphi^2$ is to be calculated.

Alpha and *Beta* specify that the numbers that follow refer to spin α and β, respectively. Both *Alpha* and *Beta* are optional, *Alpha* being assumed if omitted. The NOCV input block must be closed with "END".

Alternatively, one can specify to calculate all (alpha- or beta-) NOCV's:

For spin-unrestricted:

```
NOCV
  Alpha
   ALL
  Beta
   ALL
END
```

For both spin-restricted and spin-unrestricted:

```
NOCV
   ALL
END
```

The last and probably the most convenient form of the NOCV input blocks lets one to specify an NOCV eigenvalue threshold as a criterion for selecting orbitals:

For spin-unrestricted:

```
NOCV
  Alpha
   THRESH threshold
  Beta
```

```
   THRESH threshold
END
```

For both spin-restricted and spin-unrestricted:

```
NOCV
   THRESH threshold
END
```

When this form of the input is used, only those NOCVs will be included whose absolute eigenvalue is equal to or larger than the given *threshold*.

Generic orbitals

There is also a possibility to calculate any orbital as long as it is present in the t21 file in the BAS representation. The input syntax is as follows:

```
Orbitals GenBas
 section1%variable1
 section2%variable2
End
```

In the example above, each line contains the section and variable name of the orbital in the input t21 file. The length of the variable should be equal to the number of atomic functions (naos) and it is supposed to contain expansion coefficients of the orbital on the basis of atomic (primitive) functions.

The calculation results are stored in the output file in sections and variables with exactly the same names as specified in the input. The section and variable names may contain spaces although the leading and training spaces are discarded.

NCI

The areas of non-covalent interactions (NCI), see Refs.[356,357], can be recognized by the a low value of the electron density coupled with a low value of RDG (reduced density gradient $s = 1/2\ (3\pi^2)^{-1/3}\ |\nabla\rho|\ \rho^{-4/3}$) and a negative (or a small positive) value of the second eigenvalue of the Hessian of the electron density (λ_2). The regions of significant hydrogen bonding are recognized by strictly negative λ_2 while in the regions of VdW interactions it may be slightly positive. The relevant DENSF input keyword is:

```
NCI {BOTH|FIT} {RHOVDW=RhoVdW} {RDG=Rdg}
```

All arguments are optional

By default, the exact density is used to calculate the NCI properties. If FIT is specified then the fitted density is used to calculate the fields and their names are prepended with "Fit". If BOTH is specified then the NCI properties are calculated using both exact and the fittend density. Again names of the fields calculated from the fitted density start with "Fit"

The remaining arguments set relevant thresholds (all in atomic units):
RhoVdW: density threshold for detection of weak interaction regions (default 0.02);
Rdg: threshold on the reduced density gradient value *s* (default 0.5). A point is considered for NCI only if *s* value is smaller than Rdg.

DENSF creates three variables per density type (exact or fitted) when NCI is present in the input:
SCF%RDG (or *SCF%FitRDG*): the reduced density gradient value *s*;
SCF%DenSigned (or *SCF%FitDenSigned*): the sign(λ_2) ρ value for regions where *s* < *Rdg*;
SCF%NCI (or *SCF%FitNCI*): the NCI flag value, see below;

If the point is considered for NCI (that is if *s* < *Rdg*), the sign(λ_2)ρ value (or ρ) is tested against *RhoVdW*. If ρ < *RhoVdW* then the NCI value is set to 1 to flag a VdW interaction region. If sign(λ_2)ρ < -*RhoVdW* then the NCI value is set to -1 to flag a hydrogen bonding region. In all other cases the NCI value is zero.

SEDD

The single exponential decay detector (SEDD), see Ref. [358], extracts information about bonding and localization in atoms, molecules, or molecular assemblies. The practical evaluation of SEDD does not require any explicit information about the orbitals. The only quantity needed is the electron density (calculated or experimental) and its derivatives up to the second order. For the exact equation to be used, and pictures, see Ref. [358].

Result: TAPE41

Follows a description of the contents of TAPE41. We start with a brief discussion of the sections. At the end you can find an uncommented list of all variables and sections. Note that some data are only generated when certain keywords are provided.

Sections on TAPE41

Grid

This is a general section. It contains the grid data and some more general info.

The grid characteristics are stored as:

- The 'origin' of the grid.
- The numbers of points in three independent directions.
- Three vectors, called 'x-vector', 'y-vector' and 'z-vector'.
  They are the *steps* in the three independent directions that define the grid.

If the save option was used in input (key grid) also all grid coordinates are stored: for each point three coordinates (xyz), also if only a 2-dimensional or 1-dimensional grid has been generated (a 2D grid does not necessarily lie in the xy-plane).

Note that the grid values are now stored in a simpler manner than in previous (prior to 2004) versions of densf, because the 'x values', 'y values', and 'z values' now each have their own, separate sections.

The remaining (general) data in this section comprises:

- The number of subspecies ('symmetries') for which data such as Molecular Orbitals may be present.
- The names of the subspecies.
- A logical with the name 'unrestricted', which flags whether the data pertain to an unrestricted calculation.
- The total number. of grid points.

SumFrag

Contains grid data of the Sum-of-fragments (charge density, coulomb potential, kinetic energy density, ELF, etc.).

Ortho

Contains similar data for the orthogonalized-fragments.

SCF

Contains the (spin) density, potential, etc. of the final (scf) solution.

Core

Contains grid data of the frozen core (charge density, gradients, Hessian).

TransDens_L1_L2

Contains grid data for electron transition densities. L1 is either SS or ST, and L2 is a symmetry label for all transitions in the section. Here SS and ST stand for Singlet-Singlet and Singlet-Triplet, respectively. Variables in each section are Fitdensity_N and Coulpot_N for the density and Coulomb potential for excitation N within this spin and symmetry.

SCF_label

'Label' is one of the symmetry subspecies.

Each such section contains the total number of orbitals in that subspecies (as used in the *adf* calculation), with their occupation numbers and energy eigenvalues.

In addition it contains the grid-values of the (user-specified subset of) MOs in that subspecies. The variable name corresponding to an orbital is simply its index in the energy-ordered list of all orbitals (in that subspecies): '1', '2', etc.

LocOrb

Values of the localized orbitals.

NOCV

Values related to the NOCVs.

Geometry

Some general geometric information: the number of atoms (not counting any dummy atoms that may have been used in the *adf* calculation), their Cartesian coordinates (in bohr) and nuclear charges.

Note: the *order* of the atoms here is not necessarily identical to the input list of atoms: they are grouped by atom type.

Notes

- In an unrestricted calculation the section SCF_label is replaced by SCF_label_A and SCF_label_B for the spin-alpha and spin-beta data, respectively, and similarly for LocOrb: LocOrb_A and LocOrb_B.
- One or more subspecies may not have been used in the *adf* calculation.
  This happens when the basis set used in that calculation does not contain the necessary functions to span symmetry-adapted combinations of basis functions for that subspecies.
  In such a case the corresponding section on TAPE41 will not be created by *densf*.
- If you want to verify the contents of TAPE41, use the *pkf* utility to obtain a survey or *dmpkf* to get a complete ascii printout.

Contents of TAPE41

The information is presented in three columns. In the left-most column, section and variable names are printed, variable names being indented. In the middle column, variable's type and size is given. If the type is omitted, double precision floating point is assumed. The right-most column contains comments, if any.

Note that the name of a section of variable may consist of more than one word and that blanks in such names are significant. Furthermore, they are case-sensitive. Each line below contains the name of only one section or variable.

```
      NAME            length            Comment
Grid
  Start-point          (3)
  nr of points x       (one integer)
  nr of points y       (idem)
  nr of points z       (idem)
  total nr of points   (idem)
  x-vector             (3)
  y-vector             (3)
  z-vector             (3)
  nr of symmetries     (one integer)
  labels               (nr of symmetries*160 characters)
  unrestricted         (one logical)
SumFrag
  CoulPot              (total nr of points)
  XCPot_A              (idem)    spin-restricted: XCPot
  XCPOt_B              (idem)
  Density_A            (idem)    spin-restricted: Density
  Density_B            (idem)
  Fitdensity_A         (idem)    spin-restricted: Fitdensity
  Fitdensity_B         (idem)
  Kinetic Energy Density_A (idem)    spin-restricted:
                                      Kinetic Energy Density
  Kinetic Energy Density_B (idem)
  ELF_A                (idem)    spin-restricted: ELF
  ELF_B                (idem)
Ortho
  Same variables as in SumFrag
SCF
  Same variables as in SumFrag and Ortho, and:
  DensityLap_A         (idem)    spin-restricted: DensityLap
  DensityLap_B         (idem)
  DensityGradX_A       (idem)    spin-restricted: DensityGradX
  DensityGradX_B       (idem)
  DensityGradY_A       (idem)    spin-restricted: DensityGradY
  DensityGradY_B       (idem)
  DensityGradZ_A       (idem)    spin-restricted: DensityGradZ
  DensityGradZ_B       (idem)
  DensityHessXX_A      (idem)    spin-restricted: DensityHessXX
  DensityHessXX_B      (idem)
  DensityHessXY_A      (idem)    spin-restricted: DensityHessXY
  DensityHessXY_B      (idem)
  DensityHessXZ_A      (idem)    spin-restricted: DensityHessXZ
  DensityHessXZ_B      (idem)
  DensityHessYY_A      (idem)    spin-restricted: DensityHessYY
  DensityHessYY_B      (idem)
  DensityHessYZ_A      (idem)    spin-restricted: DensityHessYZ
  DensityHessYZ_B      (idem)
  DensityHessZZ_A      (idem)    spin-restricted: DensityHessZZ
  DensityHessZZ_B      (idem)
Core
```

```
  Density               (total nr. of points)
  DensityGradX          (idem)
  DensityGradY          (idem)
  DensityGradZ          (idem)
  DensityHessXX         (idem)
  DensityHessXY         (idem)
  DensityHessXZ         (idem)
  DensityHessYY         (idem)
  DensityHessYZ         (idem)
  DensityHessZZ         (idem)
TransDens_L1_L2                     L1: SS or ST; L2 is excitation's symmetry
  Fitdensity_1          (total nr. of points)
  Fitdensity_2          (idem)
  Fitdensity_3          (idem)
  Coulpot_1             (idem)
  Coulpot_2             (idem)
  Coulpot_3             (idem)
SCF_label_A
      (label is a symmetry subspecies.
       Spin-restricted: SCF_label)
  nr of orbitals        (one integer)
  Occupations           (nr of orbitals)
  Eigenvalues           (idem)
  1                     (total nr of points)
  2                     (idem)
  3                     (idem)
      (as many as there are Molecular Orbitals in that
       symmetry representation for the indicated spin)
SCF_label_B
      (only if spin-unrestricted same variable as
       in SCF_label_A)
LocOrb_A                            if unrestricted, otherwise
                                    LocOrb
  nr of orbitals        (one integer)
  1                     (total nr. of points)
  2                     (idem)
  (etc)
NOCV
  Dens_A number*(occupation number)     (total nr. of points)
  Dens_B number*(occupation number)     (idem)
  (etc)
Geometry
  nnuc                  (one integer)
      (nr of nuclei, omitting dummy atoms)
  xyznuc                (nnuc times 3)
      (the atoms are not in the same order as in the adf input
       file. Rather they are grouped by atomtype.)
  qtch                  (nnuc)        Atomic charges
x values
       x values         (total nr. of points)
y values
       y values         (idem)
z values
       z values         (idem)
```

4.5 Dos: Density of States

The auxiliary program *dos* computes various types of densities-of-states (DOS) for a user-specified energy interval.

dos requires an ascii input file where the user specifies the items to be calculated and computational details, plus the standard result file TAPE21 from an *adf* calculation. The latter file must be present as a local file with name TAPE21 in the directory where *dos* is executed.

dos produces as result one or more ascii files with the density-of-states values. Error messages and computational info (if any) are written to standard output.

Introduction

The program *dos* gives information on the number and character of one-electron levels (molecular orbitals) as a function of the (orbital) energy. The total density of states $N(E)$ is a well known concept in electronic structure theory of infinite systems (crystals). $N(E)dE$ denotes the number of one-electron levels (orbitals) in the infinitesimal energy interval dE. The total density of states (TDOS) at energy E is usually written as

$$N(E) = \sum_i \delta(E-\varepsilon_i) \quad (4.5.1)$$

where the ε_i denote the one-electron energies. So the integral of $N(E)$ over an energy interval E_1 to E_2 gives the number of one-electron states in that interval. Usually the δ-functions are broadened to make a graphical representation possible.

When the δ-functions are multiplied by a weight factor that describes some property of the one-electron state φ_i at energy ε_i various types of densities-of-states are obtained that provide a graphical representation of the state character (orbital character) as a function of one-electron energy.

In calculations on finite molecules the total density of states as a function of (orbital) energy may also be useful, but the main use of various types of densities-of-states is to provide a pictorial representation of Mulliken populations. The weight factors employed are related to the orbital character determined by means of a Mulliken population analysis *per orbital* (see below). The program *dos,* therefore, provides the same information as can be generated by the ADF program (a population analysis per orbital) but *dos* enables an easy graphical representation and is particularly useful when there are many one-electron levels, for instance in calculations on clusters. You can obtain a simple view of the character of the orbitals in a certain energy range. You can also find out in which orbitals (at which energies) certain basis functions or fragment orbitals give a large contribution, and whether such contributions are bonding, nonbonding or antibonding with respect to particular bonds. Such information is provided by *dos* in the form of (weighted) density of states values over a user-specified energy range, which can for instance be plotted by *gnuplot*.

The following options are available for computations by *dos*:

- TDOS: Total Density of States
- GPDOS: Gross Population Density of States
- OPDOS: Overlap Population Density of States
- PDOS: Projected Density of States

The total density of states (TDOS) has large values at energies where there are many states per energy interval.

The GPDOS (Gross Population based Density Of States) of a function χ_μ (or a sum of such functions) has large values at energies where this function (these functions) occur(s) in the molecular orbitals.

The PDOS of a function χ_μ provides similar information, but with the projection of χ_μ onto the orbital φ_i as weight factor for the importance of χ_μ in the orbital φ_i.

The OPDOS (Overlap Population based Density Of States) between χ_μ and χ_ν has large positive values at energies where the interaction between them is bonding, and negative values where the interaction is antibonding. An example of the use of these plots is provided in [326].

We review below the Mulliken population analysis, and then describe the forms of density of states analysis performed by DOS. Finally an input description of DOS is given.

Mulliken population analysis

The orbitals φ_i with energies ε_i are expanded in basis functions χ_μ, which leads to the definition of density matrices P_i describing orbital densities, from which the total density matrix can be constructed:

$$\varphi_i(\mathbf{r}) = \sum_\mu \chi_\mu(\mathbf{r})\, C_{\mu i}$$
$$\rho_i(\mathbf{r}) = \int |\varphi_i(\mathbf{r})|^2 = \sum_{\mu\nu} P_{i,\mu\nu}\, \chi_\mu(\mathbf{r})\, \chi_\nu(\mathbf{r}); \quad P_{i,\mu\nu} = C_{\mu i}\, C_{\nu i}$$
$$\rho(\mathbf{r}) = \sum_i n_i \rho_i(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu}\, \chi_\mu(\mathbf{r})\, \chi_\nu(\mathbf{r}); \quad P_{\mu\nu} = \sum_i n_i C_{\mu i}\, C_{\nu i} \quad (4.5.2)$$

Here μ and μ run over the basis functions, which may be either primitive functions, or combinations of primitive functions, for instance the SCF orbitals of atoms or larger fragments.

The Mulliken population analysis provides a partitioning of either the total charge density or an orbital density. The total density is written as

$$\rho(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu}\, \chi_\mu(\mathbf{r})\, \chi_\nu(\mathbf{r}) = \sum_{A \le B} \sum_{\mu \in A} \sum_{\nu \in B} P_{\mu\nu}\, \chi_\mu\, \chi_\nu = \sum_{A \le B} \rho_{AB} \quad (4.5.3a)$$
$$\rho_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} P_{\mu\nu}\, \chi_\mu\, \chi_\nu \quad (4.5.3b)$$

The total number of electrons, $N = \int \rho(\mathbf{r}) d(\mathbf{r})$, is now partitioned over the atoms by assigning an overlap population $P_{\mu\nu}S_{\mu\nu} + P_{\nu\mu}S_{\nu\mu}$ for one half to the atom A of χ_μ and one half to atom B of χ_ν,

$$N = \int \rho(\mathbf{r}) d(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu}\, S_{\mu\nu} = \sum_\mu GP_\mu \quad (4.5.4a)$$
$$GP_\mu = \sum_\nu P_{\mu\nu}\, S_{\mu\nu} \quad (4.5.4b)$$

GP_μ is the gross population of χ_μ. It contains the net population $P_{\mu\mu}$ and half of each total overlap population $P_{\mu\nu}S_{\mu\nu} + P_{\nu\mu}S_{\nu\mu}$ between χ_μ and χ_ν. Summing the gross populations over the functions $\mu \in A$ yields the total number of electrons assigned to atom A, or the gross population of atom A, GP_A, and hence the gross charge Q_A of atom A,

$$GP_A = \sum_{\mu \in A} GP_\mu \quad (4.5.5a)$$
$$Q_A = Z_A - GP_A \quad (4.5.5b)$$

The overlap population $OP_{\mu\nu}$ between two functions and the overlap population Q_{AB} between two atoms are defined in an analogous manner,

$$OP_{\mu\nu} = P_{\mu\nu}S_{\mu\nu} + P_{\nu\mu}S_{\nu\mu} \quad (4.5.6a)$$
$$Q_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} OP_{\mu\nu} \quad (4.5.6b)$$

These quantities can be evaluated for a single orbital density, $N = 1 = \int |\varphi_i(\mathbf{r})|^2 d\mathbf{r}$. The gross population $GP_{i,\mu}$ of a function in a specific orbital density $|\varphi_i(\mathbf{r})|^2$ is then associated with the fraction of the orbital density belonging to that function (or the percentage χ_μ character of orbital φ_i, and the overlap population $OP_{i,\mu\nu}$ gives an indication of the strength of bonding or antibonding between χ_μ and χ_ν in orbital φ_i,

$GP_{i\mu} = \sum_v P_{i,\mu v} S_{\mu v} = \sum_v C_{\mu i} C_{vi} S_{\mu v}$ (4.5.7a)
$OP_{i,\mu v} = P_{i,\mu v} S_{\mu v} + P_{i,v\mu} S_{v\mu} = 2 C_{\mu i} C_{vi} S_{\mu v}$ (4.5.7b)

Density of states analyses based on Mulliken population analysis

Total density of states

The total density of states TDOS at energy E is written as

TDOS: $N(E) = \sum_i \delta(E-\varepsilon_i)$ (4.5.8)

so the integral of $N(E)$ over an energy interval E_1 to E_2 gives the number of one-electron states in that interval. In practice the delta functions are approximated by Lorentzians,

TDOS: $N(E) = \sum_i L(E-\varepsilon_i) = \sum_i \{\sigma/\pi \cdot 1/[(E-\varepsilon_i)^2+\sigma^2]\}$ (4.5.9)

A plot of $N(E)$ *versus* E reveals energetic regions where many levels are located. The width parameter s determines of course the appearance of the plot. A typical value is 0.25 eV (used as default in *dos*).

Partial (gross population and projected) density of states

In order to find out if a given function χ_μ contributes strongly to one-electron levels at certain energies, one may weigh a one-electron level with the percentage χ_μ character. We usually determine the χ_μ character by the gross populations, obtaining the GPDOS form of the partial density of states,

GPDOS: $N_\mu(E) = \sum_i GP_{i,\mu} L(E-\varepsilon_i)$ (4.5.10)

If the weight factor is determined by projection of φ_i against χ_μ, we obtain the projected density of states PDOS,

PDOS: $N_\mu(E) = \sum_i |< \chi_\mu| \varphi_i >|^2 L(E-\varepsilon_i)$ (4.5.11)

One should not use the PDOS for d-type or f-type primitive basis functions ('BAS'). A d-type function consists of 6 Cartesian functions, while there can of course be only 5 true d-type functions among them: one (linear combination) of them is in fact an s-type function (x2+y2+z2). Similarly, there are 10 f-type Cartesian functions, 3 of which are in fact p-functions. The PDOS is calculated for the 6 d-type and 10 f-type Cartesian functions, which leads to undesired results. An PDOS for SFOs does not suffer from this problem.

Overlap population density of states (OPDOS)

If the delta function representing orbital φ_i is weighed with the overlap population between χ_μ and χ_v in φ_i, the overlap population density of states OPDOS is obtained,

OPDOS: $N_{\mu v}(E) = \sum_i OP_{i,\mu v} L(E-\varepsilon_i)$ (4.5.12)

If an orbital φ_i at energy ε_i is strongly bonding between χ_μ and χ_v the overlap population is strongly positive and OPDOS(e) will be large and positive around $E=\varepsilon_i$. Similarly, OPDOS(E) will be negative around energy ε_i when there is antibonding between χ_μ and χ_v in φ_i.

The OPDOS(E) has been used under the name coop (crystal orbital overlap population) in Extended-Hückel solid state calculations by Hoffmann and coworkers [2].
[2] R. Hoffmann, *A chemist's view of bonding in extended structures* (VCH Publishers, New York, 1988).

Generalizations of OPDOS, GPDOS, PDOS

As observed above, the basis functions in the above expressions may be primitive basis functions ('Slater type orbitals'), but of course the formulas are equally applicable for other types of MO expansions. In *dos* the user may select either the expansion in primitive basis functions ('BAS') or the expansion in SFOs (Symmetrized Fragment Orbitals) for the DOS analyses.

It is also possible in DOS to treat a *set* of basis functions simultaneously. For instance, the GPDOS for a set of basis functions μ_1, μ_2, ... is simply defined as the summation of the corresponding single-function GPDOS(E) values

$$N_{\mu\text{-set}}(E) = \sum_{\mu\in\mu\text{-set}} \sum_i GP_{i,\mu}\, L(E-\varepsilon_i) \quad (4.5.13)$$

In a similar fashion the OPDOS can be defined for *two sets* of basis functions μ_1, μ_2, ... and ν_1, ν_2, ... as

$$N_{\mu\text{-set},\nu\text{-set}}(E) = \sum_{\mu\in\mu\text{-set}} \sum_{\nu\in\nu\text{-set}} \sum_i OP_{i,\mu\nu}\, L(E-\varepsilon_i) \quad (4.5.14)$$

and finally for the PDOS we get in similar fashion

$$N_{\mu\text{-set}}(E) = \sum_{\mu\in\mu\text{-set}} \sum_i |\langle \chi_\mu | \varphi_i \rangle|^2\, L(E-\varepsilon_i) \quad (4.5.15)$$

Input

The (ASCII) input for *dos* is keyword oriented. Reading input by *dos* terminates whenever it finds a line END INPUT or the end-of-file, whichever comes first.

Follows a list of keywords with their meaning. Generally keys may occur more than once and *the order in which they appear is relevant in some cases*. For instance the key energyrange (which defines for what energy values to compute densities-of-states, see below) applies to all items that come after it in input until the next occurrence of energyrange.

```
$ADFBIN/dos << eor
ENERGYRANGE {Npoint=nr} {E-start=e1} {E-end=e2 / E-step=de}
LORENTZIAN width=width
FILE file
TDOS { title }
OPDOS ...
GPDOS ...
PDOS ...
eor
```

Energy scan values

```
ENERGYRANGE {Npoint=nr} {E-start=e1} {E-end=e2 / E-step=de}
```

This specifies for which energy values the densities-of-states are computed that are specified *after* it in the input file and *until* the next occurrence of ENERGYRANGE.

ENERGYRANGE specifies the lower bound, upper bound and number of equidistant energy values (including end-points). All items are optional with defaults applying for those omitted.

The E end and E-step values determine one another and must therefore not be specified both (or be consistent).

The initial defaults are:

```
nr=301
e1=-20
de=0.1
```

All energy data are in eV.

When values have been changed with the key ENERGYVALUE, the so-modified values are the defaults for the next occurrence of ENERGYVALUE.

Peak widening

The peaks in the DOS curves corresponding to the energies of the molecular orbitals are widened by a Lorentzian curve, the width of which can be adjusted.

```
LORENTZIAN width=width
```

Initial default width is 0.25 (eV).

As for ENERGYRANGE, the key LORENTZIAN may occur more than once and each occurrence sets the width for all items after it.

Result files

The computed densities-of-states are stored on one or more ascii files, which have to be specified in input.

```
FILE file
```

The key FILE may occur any number of times in input. Each time it occurs the specified file is opened by *dos*. The file must not yet exist and the new file will accumulate (ascii) the densities-of-states data of all DOS items subsequently specified, until the next occurrence of FILE. The first occurrence of the key FILE must be given before any DOS specification (by the keys TDOS, OPDOS, GPDOS, PDOS, see below).

The format of the result file is such that it can be fed directly into *gnuplot*.

Densities of States

Total density of states.

```
TDOS { title }
```

TDOS
instructs the program to compute the *total* density of states.
title (optional)
will appear as title to the section of corresponding Density-of-States data in the result file.

The other types of densities-of-states require block-type keyword input.

```
OPDOS { title }
Ftype numbers
Ftype numbers
...
SUBEND
Ftype numbers
Ftype numbers
...
END
```

Ftype
Specifies the type of basis functions to use in the MO expansions. If the primitive basis functions are to

be used Ftype must be bas. For the SFO representation Ftype must be one of the irreducible representations of the pointgroup symmetry. All Ftype values in the data block must be consistent: either all are bas or all are irrep labels. The scope of this consistency requirement is the data block of the current key: in a next OPDOS data block, for instance, a different choice may be made.

numbers

Must be a sequence of integers referring to the basis functions to be selected, i.e. the 'μ-set' and 'ν-set' in (4.5.13) etc.

If bas-type basis functions are selected the numbers refer to the overall list of all basis functions as printed in the output file of the *adf* run. If SFOs are selected the numbers refer to the SFO list of the pertaining symmetry representation *without the core functions*, see the *adf* output file.

SUBEND

Must be typed as such and separates the 'μ-set' and the 'ν-set': all records before subend specify together the 'μ-set' and all records below subend comprise the 'ν-set'. Each of these two sections may consist of any number of records.

The input for GPDOS and PDOS are similar, but simpler because only one set of functions ('μ-set') has to be specified, so there is no subend in the data blocks for these keys.

```
GPDOS { title }
Ftype numbers
Ftype numbers
...
END
```

```
PDOS { title }
Ftype numbers
Ftype numbers
...
END
```

The keys GPDOS, OPDOS, PDOS and (TDOS) may occur any number of times in input and in any order. Each time the DOS key occurs the current energyrange and lorentzian settings apply and the results are written to the current file.

4.6 Other plotting programs

Most of the functionality that the two following two programs can offer, can already be done with the ADF-GUI.

- *cntrs*: a program to generate contours for data computed by *densf*.
- *adfplt*: a program to graphically display orbitals, densities or potentials computed in a 2D or 3D grid.

The separate adfplt program is no longer documented (and is not part of the distribution). See for the old documentation the ADF2010 Analysis document. The ADFview module of the ADF-GUI, see the ADF-GUI tutorials is the suggested alternative for adfplt that is supported by SCM.

Cntrs: Contour Plots

cntrs is an auxiliary program to generate plot data from TAPE41 produced by *densf*. In the future *cntrs* will be superseded by the ADF-GUI, which will offer both two-dimensional and three-dimensional visualization possibilities.

cntrs requires an ascii input file where the user specifies which items should be plotted and what scan values are to be used, plus the standard result file TAPE41 from *densf*. TAPE41 must be present as a local

file in the directory where *contrs* executed. For usage by *contrs* TAPE41 must have been generated in a *densf* run using a *two*-dimensional grid.

cntrs produces as result one or more ascii files with plot data.

An example of using *cntrs* is contained in the set of sample runs (NO_2), see the *Examples* document.

Input

The (ascii) input for *cntrs* is keyword oriented. *The order of keywords in input is relevant for* cntrs.

Scan

```
scan
 scanvalues
END
```

With the scan key you read in values for which contours are to be generated for the items that are specified subsequently in input. Up to a maximum number of 20 scan values can be supplied. scan may occur any number of times in input. Each occurrence resets the scan values for the subsequent items. The initial values, which apply until the first occurrence, if at all, of scan are the eleven values 0, ±2e-2, ±5e-2, ±1e-1, ±2e-1, ±5e-1.

Dash

```
DASH length
```

contours corresponding to positive scan values are plotted as solid lines, the zero-contour is plotted by a dash-dot-dash line, and negative contours are dash-lines.

The dash key defines the length of a dash. Default: 0.2 bohr. dash may occur any number of times in input, each occurrence resets the dash-length for the items that follow.

Items to be plotted

The remaining part of input has the format:

```
FILE filename
item {factor}
item {factor}
...
FILE filename
item {factor}
...
(etc.)
END INPUT
```

Each FILE key requires the name of a file. This file must not yet exist and will be created by *cntrs* as an ascii file on which to write plot data. All 'items' until the next occurrence of dash will be combined into one quantity for the contours of which the plot data are generated.

Each item must be of the form Section%Variable and must in this way correspond to one of the variables on TAPE41 (*case-sensitive!*), see the description in the chapter about *densf*.. All items that belong to one file will be added up, each one multiplied by its factor (default: 1.0), to the quantity to be plotted. In this way you can generate contours for instance of density differences or a summation of densities.

Result

Each of the ascii result files, the names of which are defined in input (key file), consists of a sequence of data blocks.

Each block consists of a number of records that contain two values ('x' and 'y') and it ends with a blank line.

Each block defines a contour by plot instructions as follows:

- for the first {x,y}: start to plot at that point.
- for each next {x,y}: (continue to) draw the contour to that point.
- the blank line signals the end of the contour.

The last block does not correspond to a contour, but draws a rectangle around the whole picture.

5 APPENDICES

5.1 Database

The database contains standard basis sets (and fit sets, frozen core orbitals) for all chemical elements of the periodic table at different levels of accuracy. The database is partitioned in subdirectories. Some of these are special: for example, the subdirectory Dirac contains input files for the program *dirac* (computation of relativistic potentials and charge densities). Most subdirectories contain files for the create runs: for example, the subdirectories SZ through TZ2P. The section about the database of STO basis sets describes all subdirectories in more details. See also the section about the standard basis sets available in ADF.

The names of the files in the database consist of two parts: the standard symbol for the chemical element and the level of frozen core approximation. Mn.2p for instance is a data file for Manganese with a frozen core up to and including the 2p shell.

Polarization functions are provided for most elements. If you contemplate to compile more extended basis sets, by including one or more polarization functions, a good rule of thumb to choose the functional characteristics, is the following. Take the next higher *l*-value that does not yet occur in the function set (however, do not go beyond *f*-functions: the program cannot (yet) handle *g*-type basis functions), select the minimum value for the main quantum number *n* that is compatible with the *l*-value (i.e.: 2p, 3d, 4f), and determine the exponential decay factor ζ, such that the function attains its maximum value at somewhere between 1/3 and 1/2 times the bond length. The functional maximum for a Slater-type function is at $R=(n-1)/\zeta$. The maximum for r^2 times the square of a Slater-type function is at $R=n/\zeta$.

Many all-electron basis can be found in the data base, especially for the elements H-Kr. All electron basis sets for the heavier elements can be found in the ZORA subdirectory. Fit functions for the all-electron basis sets must include more, in particular more contracted functions than the standard fit sets that are provided in the frozen core database files. If you would combine a basis set with an inadequate fit set the results are unreliable and absolutely inadequate, in the same fashion as when you would have used a highly inadequate basis set.

Data File for Create

The data file supplied to ADF in Create mode contains the following sections:

```
Title
Basis Functions
Core Expansion Functions
Core Description
Fit Functions
Start-up Fit Coefficients
```

Each of these items is discussed below. The data file does *not* define the applied density functional, the electronic configuration, precision parameters (numerical integration, SCF convergence criterion...), etc etera. These items can be set in the normal input file if the default is not satisfactory.

Title

is the first record of the file. It may contain any text. Only the first 60 characters are actually used. This title is (by default) printed in the output; it is also used to stamp an identification on the result file (TAPE21). The file stamp will be printed whenever you use it as a fragment file in another calculation.

Basis functions

A list of Slater type basis function characteristics. This part has the following format (example):

```
BASIS
  1s 5.4
  2s 1.24
  ...
  (etc.)
  ...
end
```

The words basis and end signal the beginning and the end of this section in the data file. The records in-between list the basis functions; each record contains the main quantum number, the angular quantum number, and the exponential decay factor for a *set* of Slater type basis functions. A function description *3d 2.5* for instance represents the functions *re*Y, *m*=-2,...,2.

The order of specification of the basis functions is not free. First must come the Core Functions used for core-orthogonalization, see Chapter 1.2. The CFs must be in order: s-functions first, then p-functions, then d-functions, and finally f-functions (as far as applicable). In the valence basis set there must be exactly one core-orthogonalization function for each frozen core shell (1s, 2s, 2p, ...).

Here as well as in all other function definitions below, the unit of length, implicit in the exponential decay factor, is bohr (atomic units), irrespective of the unit of length used in input for geometric items such as atomic positions (see units).

Core expansion functions

This part has the form

```
CORE ns, np, nd, nf
  1s 7.68
  ...
  (etc.)
  ...
end
```

It looks very much like the *basis functions*: a list of Slater type function descriptions, closed by end. The header record however (core...) contains in addition four integers ns, np, nd, nf. They are the numbers respectively of *s-, p-, d-, and f-* frozen core shells in the atom. If you create for instance a Ruthenium atom with a frozen core up to the 4p shell, these numbers would be *4 3 1 0* : four frozen *s*-shells (1s,2s,3s,4s), three frozen *p*-shells (2p,3p,4p), one frozen *d*-shell (3d), and no frozen *f*-shells.

The core expansion sets defined in this section are used to describe the frozen core orbitals; they are not included in the valence basis set. In the list of core expansion sets all *s*-type functions must come first, then the *p*-type functions, then the *d*-functions, and then the *f*-functions (as far as applicable).

Core description

Describes the frozen core shells as linear combinations of the core expansion functions. This section has the form

```
COREDESCRIPTION
  coefficients for the first frozen s-shell
  for the second s-shell
  ...
  for the n-th shell
  coefficients for the first frozen p-shell
  for the second p-shell
```

```
  ...
  for the d-shells
  for the f-shells
  pseudopotential parameters
end
```

For each of the angular momentum quantum numbers *l*=*s, p, d, f* all n*l* frozen shells are described by giving expansion coefficients. There are as many coefficients as there are function *sets* with the pertaining *l*-value in the list of expansion functions. There are no separate coefficients for all *m*-values: all *m*-values are equivalent in a spherically symmetric model atom. See the Ca example below.

At the end of the (core) description section there is a record with pseudopotential parameters. The pseudopotential option, as an alternative to the frozen core approximation, is currently not supported, all values in this record must be zero, one for each frozen core shell. Equivalently you can put one zero, followed by a slash (/).

Fit functions

is again a list of Slater type functions. These are used for an expansion of the density. The Coulomb potential due to the electronic charge distribution is computed from this expansion, see Chapter 1.2.

The format of this section is similar to the *basis functions*:

```
FIT
  1s 10.8
  ...
  ...
  (etc.)
  ...
end
```

The program cannot handle fit functions with *l*-value higher than 4, i.e. not higher than *g*-type functions. Bear this in mind if you construct alternative fit sets.

In view of the next item, one is well advised to put the *s*-functions first.

Start-up fit coefficients

The initial (start-up for the SCF procedure) expansion of the atomic charge density in terms of the fit functions. Since the atom is spherically symmetric, only *s*-type functions should have non-zero coefficients. This is why the *s*-type fit functions should be listed first: the list of coefficients can then, after the *s*-set, be closed by a slash, rather than putting a long series of zeros.

The higher *l*-values (p, d...) in the fit set play no role in the creation of the basic atom, because it is spherically symmetric. They should not be omitted however as they will be needed when the atom is used as a fragment in a molecule: the charge density around the atom is then not spherically symmetric anymore.

The form of this section is simple:

```
FITCOEFFICIENTS
coefficients
end
```

Example: Calcium

An example may serve to illustrate the format of a Create data file for Ca (DZ, note that compared to the old basis II an extra 3D polarization function is added) (empty records inside and between the various sections are meaningless and ignored):

```
Calcium (II, 2p frozen)

BASIS
1S  15.8
2S   6.9
2P   8.1

3S   2.6
3S   3.9
3P   2.1
3P   3.4
4S   0.8
4S   1.35
4P   1.06

3D   2.000
END

CORE    2  1  0  0
 1S  24.40
 1S  18.25
 2S   7.40
 2S   4.85
 3S   4.00
 3S   2.55
 4S   0.70
 4S   1.05
 4S   1.65
 2P  10.85
 2P   6.45
 3P   1.85
 3P   2.70
 3P   4.00
END

DESCRIPTION
  0.2076143E+00  0.7975138E+00 -0.7426673E-04  0.1302616E-03 -0.6095738E-04
  0.1508446E-04  0.1549420E-06 -0.2503155E-07 -0.1843317E-05
  0.8487466E-01 -0.4505954E+00  0.1009184E+01  0.9627952E-01 -0.3093986E-01
  0.1678301E-01 -0.2381843E-02  0.6270439E-02 -0.8899688E-02
  0.3454503E+00  0.6922138E+00 -0.1610756E-02  0.5640782E-02 -0.5674517E-02

0/
END

FIT
 1S  31.80
 2S  29.37
 3S  25.15
 4S  21.06
```

```
 4S  13.99
 5S  11.64
 5S   8.05
 6S   6.69
 6S   4.76
 6S   3.39
 7S   2.82
 7S   2.06
 7S   1.50
 2P  24.10
 3P  14.78
 4P   9.29
 5P   5.98
 6P   3.94
 6P   2.24
 7P   1.50
 3D  16.20
 4D  10.47
 5D   6.91
 6D   4.65
 6D   2.70
 7D   1.85
 4F   7.00
 5F   4.00
  5G   3.50
END

FITCOEFFICIENTS
   .567497268648811470E+02  -.452377281899367176E+03   .326145159087736033E+03
   .337765644703942453E+05   .131300324467109522E+04  -.704903218559526340E+04
   .755210587728052587E+03   .281241738156731174E+03   .864928185630532020E+01
  -.230025056878739281E+00   .366639011114029689E-01   .905663001010961841E-03
   .160080832168547530E-04   .000000000000000000E+00   .000000000000000000E+00
/
END
```

5.2 Elements of the Periodic Table

A few characteristics are predefined in ADF for all elements of the periodic table, as shown below.

The electronic configuration defines the default occupation numbers in Create mode. Basis sets for the elements Rf-Uuo (Z=104-118) are only available in the ZORA atomic database.

	Nuclear Charge Z	*mass number of default isotope used for mass*	*electronic configuration*
H	1	1	$1s^1$
He	2	4	$1s^2$
Li	3	7	$2s^1$
Be	4	9	$2s^2$
B	5	11	$2s^2 2p^1$
C	6	12	$2s^2 2p^2$

N	7	14	$2s^2 2p^3$
O	8	16	$2s^2 2p^4$
F	9	19	$2s^2 2p^5$
Ne	10	20	$2s^2 2p^6$
Na	11	23	$3s^1$
Mg	12	24	$3s^2$
Al	13	27	$3s^2 3p^1$
Si	14	28	$3s^2 3p^2$
P	15	31	$3s^2 3p^3$
S	16	32	$3s^2 3p^4$
Cl	17	35	$3s^2 3p^5$
Ar	18	40	$3s^2 3p^6$
K	19	39	$4s^1$
Ca	20	40	$4s^2$
Sc	21	45	$3d^1 4s^2$
Ti	22	48	$3d^2 4s^2$
V	23	51	$3d^3 4s^2$
Cr	24	52	$3d^5 4s^1$
Mn	25	55	$3d^5 4s^2$
Fe	26	56	$3d^6 4s^2$
Co	27	59	$3d^7 4s^2$
Ni	28	58	$3d^9 4s^1$, $3d^8 4s^2$
Cu	29	63	$3d^{10} 4s^1$
Zn	30	64	$3d^{10} 4s^2$
Ga	31	69	$3d^{10} 4s^2 4p^1$
Ge	32	74	$3d^{10} 4s^2 4p^2$
As	33	75	$3d^{10} 4s^2 4p^3$
Se	34	80	$3d^{10} 4s^2 4p^4$
Br	35	79	$3d^{10} 4s^2 4p^5$
Kr	36	84	$3d^{10} 4s^2 4p^6$
Rb	37	85	$5s^1$
Sr	38	88	$5s^2$
Y	39	89	$4d^1 5s^2$
Zr	40	90	$4d^2 5s^2$
Nb	41	93	$4d^4 5s^1$
Mo	42	98	$4d^5 5s^1$
Tc	43	(98)	$4d^5 5s^2$
Ru	44	102	$4d^7 5s^1$
Rh	45	103	$4d^8 5s^1$

Pd	46	106	$4d^{10}$
Ag	47	107	$4d^{10}5s^{1}$
Cd	48	114	$4d^{10}5s^{2}$
In	49	115	$4d^{10}5s^{2}5p^{1}$
Sn	50	120	$4d^{10}5s^{2}5p^{2}$
Sb	51	121	$4d^{10}5s^{2}5p^{3}$
Te	52	130	$4d^{10}5s^{2}5p^{4}$
I	53	127	$4d^{10}5s^{2}5p^{5}$
Xe	54	132	$4d^{10}5s^{2}5p^{6}$
Cs	55	133	$6s^{1}$
Ba	56	138	$6s^{2}$
La	57	139	$5d^{1}6s^{2}$
Ce	58	140	$4f^{1}5d^{1}6s^{2}$
Pr	59	141	$4f^{3}6s^{2}$
Nd	60	142	$4f^{4}6s^{2}$
Pm	61	(145)	$4f^{5}6s^{2}$
Sm	62	152	$4f^{6}6s^{2}$
Eu	63	153	$4f^{7}6s^{2}$
Gd	64	158	$4f^{7}5d^{1}6s^{2}$
Tb	65	159	$4f^{9}6s^{2}$
Dy	66	164	$4f^{10}6s^{2}$
Ho	67	165	$4f^{11}6s^{2}$
Er	68	166	$4f^{12}6s^{2}$
Tm	69	169	$4f^{13}6s^{2}$
Yb	70	174	$4f^{14}6s^{2}$
Lu	71	175	$4f^{14}5d^{1}6s^{2}$
Hf	72	180	$4f^{14}5d^{2}6s^{2}$
Ta	73	181	$4f^{14}5d^{3}6s^{2}$
W	74	184	$4f^{14}5d^{4}6s^{2}$
Re	75	187	$4f^{14}5d^{5}6s^{2}$
Os	76	192	$4f^{14}5d^{6}6s^{2}$
Ir	77	193	$4f^{14}5d^{7}6s^{2}$
Pt	78	195	$4f^{14}5d^{9}6s^{1}$
Au	79	197	$4f^{14}5d^{10}6s^{1}$
Hg	80	202	$4f^{14}5d^{10}6s^{2}$
Tl	81	205	$4f^{14}5d^{10}6s^{2}6p^{1}$
Pb	82	208	$4f^{14}5d^{10}6s^{2}6p^{2}$
Bi	83	209	$4f^{14}5d^{10}6s^{2}6p^{3}$
Po	84	(209)	$4f^{14}5d^{10}6s^{2}6p^{4}$

At	85	(210)	$4f^{14}5d^{10}6s^{2}6p^{5}$
Rn	86	(222)	$4f^{14}5d^{10}6s^{2}6p^{6}$
Fr	87	(223)	$7s^{1}$
Ra	88	(226)	$7s^{2}$
Ac	89	(227)	$6d^{1}7s^{2}$
Th	90	232	$6d^{2}7s^{2}$
Pa	91	231	$5f^{2}6d^{1}7s^{2}$
U	92	238	$5f^{3}6d^{1}7s^{2}$
Np	93	(237)	$5f^{4}6d^{1}7s^{2}$
Pu	94	(244)	$5f^{6}7s^{2}$
Am	95	(243)	$5f^{7}7s^{2}$
Cm	96	(247)	$5f^{7}6d^{1}7s^{2}$
Bk	97	(247)	$5f^{9}7s^{2}$
Cf	98	(251)	$5f^{10}7s^{2}$
Es	99	(252)	$5f^{11}7s^{2}$
Fm	100	(257)	$5f^{12}7s^{2}$
Md	101	(258)	$5f^{13}7s^{2}$
No	102	(259)	$5f^{14}7s^{2}$
Lr	103	(260)	$5f^{14}6d^{1}7s^{2}$
Rf	104	(261)	$5f^{14}6d^{2}7s^{2}$
Db	105	(262)	$5f^{14}6d^{3}7s^{2}$
Sg	106	(263)	$5f^{14}6d^{4}7s^{2}$
Bh	107	(264)	$5f^{14}6d^{5}7s^{2}$
Hs	108	(265)	$5f^{14}6d^{6}7s^{2}$
Mt	109	(268)	$5f^{14}6d^{7}7s^{2}$
Ds	110	(269)	$5f^{14}6d^{8}7s^{2}$
Rg	111	(272)	$5f^{14}6d^{9}7s^{2}$
Cn	112	(277)	$5f^{14}6d^{10}7s^{2}$
Uut	113	(280)	$5f^{14}6d^{10}7s^{2}7p^{1}$
Fl	114	(280)	$5f^{14}6d^{10}7s^{2}7p^{2}$
Uup	115	(280)	$5f^{14}6d^{10}7s^{2}7p^{3}$
Lv	116	(280)	$5f^{14}6d^{10}7s^{2}7p^{4}$
Uus	117	(280)	$5f^{14}6d^{10}7s^{2}7p^{5}$
Uuo	118	(280)	$5f^{14}6d^{10}7s^{2}7p^{6}$
Default (most abundant) isotope, used to set atomic mass (nr. of brackets gives mass directly). Default electronic configurations used in Create mode.			

5.3 Symmetry

Schönfliess symbols and symmetry labels

A survey of all point groups that are recognized by ADF is given below. The table contains the Schönfliess symbols together with the names of the subspecies of the irreducible representations as they are used internally by ADF. The subspecies names depend on whether single-group or double-group symmetry is used. Double-group symmetry is used only in relativistic spin-orbit calculations.

Note that for some input of TDDFT (Response) calculations, other conventions apply for the subspecies. This is explicitly mentioned in the discussion of that application.

Point Group	*Schönfliess Symbol in ADF*	*Irreducible representations in single-group symmetry*	*Irreducible representations in double-group symmetry*
C_1	NOSYM	A	A1/2
R^3	ATOM	s p d f	s1/2 p1/2 p3/2 d3/2 d5/2 f5/2 f7/2
T_d	T(D)	A1 A2 E T1 T2	E1/2 U3/2 E5/2
O_h	O(H)	A1.g A2.g E.g T1.g T2.g A1.u A2.u E.u T1.u T2.u	E1/2.g U3/2.g E5/2.g E1/2.u U3/2.u E5/2.u
$C_{\infty v}$	C(LIN)	Sigma Pi Delta Phi	J1/2 J3/2 J5/2 J7/2
$D_{\infty h}$	D(LIN)	Sigma.g Sigma.u Pi.g Pi.u Delta.g Delta.u Phi.g Phi.u	J1/2.g J1/2.u J3/2.g J3/2.u J5/2.g J5/2.u J7/2.g J7/2.u
C_i	C(I)	A.g A.u	A1/2.g A1/2.u
C_s	C(S)	AA AAA	A1/2 A1/2*
C_n	C(N), n must be 2	A B E1 E2 ... odd n: without B	A1/2 A1/2*
C_{nh}	C(NH), n must be 2	even n: A.g B.g A.u B.u E1.g E1.u E2.g E2.u ... odd n: AA AAA EE1 EE2 ... EEE1 EEE2 ...	A1/2.g A1/2.g* A1/2.u A1/2.u*
C_{nv}	C(NV), n<9	A1 A2 B1 B2 E1 E2 E3 ... odd n: without B1 and B2	E1/2 E3/2 E5/2 ... for odd n also: An/2 An/2*
D_n	D(N), n<9	n=2: A B1 B2 B3 other: A1 A2 B1 B2 E1 E2 E3 ... odd n: without B1 B2	E1/2 E3/2 ... for odd n also: An/2 An/2*
D_{nh}	D(NH), n<9	n=2: A.g B1.g B2.g B3.g A.u B1.u B2.u B3.u even n (≠2): A1.g A2.g B1.g B2.g E1.g E2.g E3.g ... A1.u A2.u B1.u ... odd n: AA1 AA2 EE1 EE2 ... AAA1 AAA2 EEE1 EEE2	even n: E1/2.g E1/2.u E3/2.g E3/2.u ... odd n: E1/2 E3/2 E5/2 ...
D_{nd}	D(ND), n<9	even n: A1 A2 B1 B2 E1 ... odd n: A1.g A2.g E1.g E2.g ... E(n-1)/2.g A1.u A2.u E1.u E2.u ... E(n-1)/2.u	even n: E1/2 E3/2 ... odd n: E1/2.g E1/2.u E3/2.g E3/2.u An/2.g An/2.u An/2.g* An/2.u*
Schönfliess symbols and the labels of the irreducible representations.			

Most labels are easily associated with the notation usually encountered in literature. Exceptions are AA, AAA, EE1, EEE1, EE2, EEE2, etcetera. They stand for A', A'', E1', E1'', and so on. The AA, etc. notation is used in ADF to avoid using quotes in input files in case the subspecies names must be referred to.

The symmetry labeling of orbitals may depend on the choice of coordinate system. For instance, B1 and B2 representations in C_{nv} are interchanged when you rotate the system by 90 degrees around the z-axis so that x-axis becomes y-axis and vice-versa (apart from sign).

Labels of the symmetry subspecies are easily derived from those for the irreps. For one-dimensional representations they are identical, for more-dimensional representations a suffix is added, separated by a colon:
For the two- and three-dimensional E and T representations the subspecies labels are obtained by adding simply a counting index (1, 2, 3) to the name, with a colon in between; for instance, the EE1 irrep in the Dnh

pointgroup has EE1:1 and EE1:2 subspecies. The same holds for the two-dimensional representations of $C_{\infty v}$ and $D_{\infty h}$. For the R3 (atom) point group symmetry the subspecies are p:x, p:y, p:z, d:z2, d:x2-y2, etc.

All subspecies labels are listed in the Symmetry section, very early in the ADF output. To get this, perform a quick run of the molecule using the STOPAFTER key (for instance: stopafter config).

Molecular orientation requirements

ADF requires that the molecule has a specific orientation in space, as follows:

- The origin is a fixed point of the symmetry group.
- The z-axis is the main rotation axis, xy is the σ_h-plane (axial groups, C(s)).
- The x-axis is a C_2 axis (D symmetries).
- The xz-plane is a σ_v-plane (C_{nv} symmetries).
- In T_d and O_h the z-axis is a fourfold axis (S_4 and C_4, respectively) and the (111)-direction is a threefold axis.

If the user-specified symmetry equals the true symmetry of the nuclear frame (including electric field and point charges) the program will adapt the input coordinates to the above requirements, if necessary. If no symmetry has been specified at all ADF assumes you have specified the symmetry of the nuclear frame, accounting for any fields. If a subgroup has been specified for the molecular symmetry the input coordinates will be used as specified by the user. If a Z-matrix input is given this implies for the Cartesian coordinates: first atom in the origin, second atom on the positive x-axis, third atom in the xy-plane with positive y value.

5.4 Binary result Files, KF utilities

TAPE21

TAPE21 is the general result file of an ADF calculation. It is a KF file: Direct-Access, binary, and keyword driven. It contains information about the calculation. You can use it as a fragment file in a subsequent calculation on a bigger molecule, where the current one may be a part, or in an analysis program.

The contents of TAPE21 is keyword-accessible and you may use the KF utilities (see Appendix 5.5) for conversion of TAPE21 from binary to ASCII format and vice versa. This facility is also useful when you intend to use a TAPE21 result file produced on one type of hardware, for a continuation run on a quite different computer: Transform the binary file to ASCII format with the KF utilities on the first machine. Then transport the ASCII file to the other machine, and make a binary out of it again.

Another utility (*pkf*) can be used to obtain a summary of the contents of TAPE21. The output should be more or less self-documenting: all variables are listed by name, type (integer, real, ..) and size (number of array elements) and grouped in named sections.

The data on TAPE21 is organized in Sections which group together related data. Each section contains a number of variables. Each variable may be an array or a scalar and may be integer, real, logical or character type.

A complete dump of the contents of TAPE21 is obtained with *dmpkf*. The resulting ASCII file contains for all variables on the file:

- The name of the section it belongs to;
- The name of the variable itself;
- Three integers coding for the data of the variable:
    - The number of data elements reserved on the file for the variable;

    - The number of data elements actually used for the variable.
      In virtually all cases the number of *used* elements is equal to the number of *reserved* elements.
      The number of *used* elements is relevant for interpreting the data, the number of *reserved* elements
      has only relevance for the management of data on the file by kf-specific modules and utilities;
    - An integer code for the data type: 1=integer, 2=real, 3=character, 4=logical;
- The variable value(s).

A typical case of the contents of TAPE21 obtained by *dmpkf* operating on the binary TAPE21 file from an optimization run on H2O would be:

contents of TAPE21	***comment***
General	name of (first) section
file-ident	name of (first) variable in the current section (General)
6 6 3	characteristics of the data: 6 elements reserved on file for the variable, 6 data elements actually used, 3=integer code for the data type: character
TAPE21	Value of the variable fileident in the section General.
General	again: name of the section
title	name of the (second) variable
80 80 3	reserved and used number of data elements (both 80), and the data type code (3: character)
Water Geometry Optimization with Internal Coordinates	value
(etc.)	(etc.)

A description of the various utilities that can be used to process TAPE21 can be found in other parts of this ADF manual.

Contents of TAPE21

Follows a survey of the sections and variables on TAPE21. Details may differ between different run types (SinglePoint, Frequencies...). Most items should be self-explanatory. Some are only significant for internal proceedings of the program and will not be explained in detail. The sections are described in an order that corresponds to the order in which they are generated and hence printed by the KF utility programs. However, the order of generation depends somewhat on the type of application, so some difference may be found when comparing to your own TAPE21 printout.

Note that variable and section names may contain spaces: these are significant.

A special section is the 'SUPERINDEX' section, which is in fact a table-of-contents that lists all the sections in the file, with technical information regarding their position on the file, the amount of data corresponding to that section and similar items. The SUPERINDEX section is not discussed further here. See the KF documentation for more details.

Section General

General information about the calculation and the file

`fileident`

Name of the file. Here: TAPE21

`jobid`

ADF release number with date and time of the calculation

`title`

Title of the calculation. This may have been set in the input file, or be internally generated. In a create run it is picked up from the Create database file (if no input value for the title key has been given).

`runtype`

The type of calculation, for instance SinglePoint or Frequencies

`nspin`

1 for a spin-restricted calculation, 2 for spin-unrestricted

`nspinf`

Similar for the fragment occupation numbers as they are used in the calculation, See the key FRAGOCCUPATIONS

`ldapot`

An integer code for the applied LDA part of the XC potential functional used in the SCF. 1 for VWN, 2 for VWN+Stoll ...

`xcparv`

X-alpha parameter value. Only relevant for the X-alpha LDA potential, meaningless if another LDA potential functional has been selected.

`ldaen`

As for ldapot: integer code for the LDA part of the Density Functional, now however pertaining to the (post-SCF) energy evaluation. Usually ldaen and ldapot are identical. See the key XC for details.

`xcpare`

As xcparv, but now for the energy evaluation.

`ggapot`

Specification (string) of the GGA part of the XC potential used in the SCF, for instance 'Becke Perdew'. If no GGA potential is applied, the string ggapot is empty.

`ggaen`

Similar for the GGA part of the XC energy evaluation

`iopcor`

Code for usage of frozen core: 1=use frozen cores, 0=pseudopotentials. Pseudopotentials are not supported anymore in ADF, so this variable must always be 1.

`electrons`

The number of valence electrons
Note that this is not necessarily the same as what may consider, chemically, as the valence space.

Rather, it equals the total number of electrons in the calculation minus the electrons in the frozen core orbitals.

`unit of length`

Transformation factor between input-used geometrical units (for distances) and atomic units (bohr). If input of, say, the atomic coordinates is in Angstrom, the unit of length is approximately 1.89

`unit of angle`

Similar for angles. Internal units in the program are radians. Input (bond and dihedral angles) may be in degrees, in which case the unit of angle equals approximately 0.017

Section Geometry

Geometrical data such as number of atoms, coordinates, etc: Most variable names should be self-explanatory

`grouplabel`

Point group symmetry (string) used in the calculation, for instance O(H). This may be set in the input file.

`Geometric Symmetry`

Auto-determined ('true') symmetry (considering the nuclear frame and any external fields, but not taking into account any user-defined MO occupation numbers and hence the electronic charge distribution.

`symmetry tolerance`

Threshold for allowed deviation of input atomic coordinates from symmetry to be detected or verified.

`orient`

Affine transformation (3,4 matrix: rotation and translation) between the input coordinates and the frame in which the program processes the atoms. ADF has certain orientation requirements for all supported point group symmetries and may rotate and translate the input coordinates accordingly.

`oinver`

The inverse transformation of orient

`lrotat`

A logical flag to signal whether or not a rotation has been applied between the input frame and the internally used frame.

`nr of fragmenttypes`

The number of distinct types of fragments

`nr of dummy fragmenttypes`

Idem, but counting only dummy atom fragments. A dummy fragment, if it exists, must consist of one single (dummy) atom.

`fragmenttype`

Names (string) of the fragment types.

`fragment mass`

Sum of atomic masses in the fragment.

`fragment charge`

An array with 3 values per fragment type (nftypes,3): 1=sum of nuclear charges, 2=sum of effective nuclear charges (discounting for the frozen core shells), 3=nr of valence electrons

`fframe`

Signals whether or not special local coordinate frames are used for the atoms. Usually this is not so, in which case the variable has the value DEFAULT. fframe is an array that runs over the atoms. See the 'z=' option to the data in the ATOMS input key block.

`cum nr of fragments`

An array (0:nftyps) that gives the total number of fragments for the fragment types up to and including the indexed one. The ordering of fragments and fragment types is printed in the standard output file.

`nr of fragments`

The total number of fragments in the calculation
This equals the last element of the previous variable 'cum nr of fragments'

`nr of dummy fragments`

The total number of fragments that each consist of a single dummy atom.

`fragment mapping`

Affine transformation matrices (3,4: rotation and translation), one for each fragment in the molecule, that transform the fragment coordinates as they are on the fragment file(s), to the actual position of the fragments in the molecule.

`cum nr of atomtypes`

An array (0:fragmenttypes) that counts the number of atom types up to and including the indexed fragment type.

`nr of atomtypes`

Total number of atom types in the molecule. Must equal the last element of the 'cum nr of atomtypes' array

`nr of dummy atomtypes`

Similar, now counting only the atom types consisting of a dummy atom.

`atomtype`

Names (strings) of the atom types

`mass`

Atomic masses: array running over the atom types. Compare 'fragment mass'.

`charge`

Similar as for 'fragment charge', but now the values per atom type.

`cum nr of atoms`

An array (0:atomtypes) that counts the number of atoms up to and including the indexed atom type.

`nr of atoms`

Total number of atoms. Must equal the last element of the array 'cum nr of atoms'.

`nr of dummy atoms`

Total number of dummy atoms

`atmcrd`

Type of atomic coordinates in input: CART (Cartesian) or ZMAT (Internal).

`kmatrix InputOrder`

The connection matrix listing (and referencing) the atoms in the order as they were in the input file. This ordering aspect is significant because internally the program reorders the atoms and groups them together by atom type and fragment type. Hence it is relevant to know what ordering (input- or internal-) is assumed in data arrays.

`zaxis`

For each atom the direction of the local z-axis. Normally this is identical to the standard (0,0,1), but it may be different for analysis purposes. See the 'z=' option to the data records in the ATOMS block.

`fragment and atomtype index`

An integer array (natoms,2) that specifies for each atom the fragment and the atom type it belongs to.

`atom order index`

An integer array (natoms,2) that defines the re-ordering of atoms between the list in the input file and the internally used list (which is driven by fragment types, fragments, atom types; dummies come last). The internally used list can be derived from the printout of the fragments, early in the standard output.

`kmatrix`

The connection matrix using the internally applied ordering of atoms

`xyz`

Cartesian coordinates of the atoms, in the internally used ordering of atoms

`xyz Inputorder`

Similar, but now for the ordering of atoms as in the input file.

`zmatrix`

Internal (Z-matrix) atomic coordinates

`zmatrix Inputorder`

Internal coordinates in the input-order of atoms

`Atomic Distances`

Inter atomic distance matrix

`ntyp`

Number of atom types, not counting dummy atoms,

`nqptr`

A cumulative counting array, very similar to 'cum nr of atoms'
Differences: it runs only over 'ntyp' atom types (not including dummy atoms) and its indexing as well as its values are shifted by one: nqptr(k) is the total number of atoms plus one, counting the atom types up to and including #(k-1)

`nnuc`

Total number of non-dummy atoms

`qtch`

Nuclear charges of the non-dummy atoms

`qeff`

Effective nuclear charges (subtracting charge for the frozen core shells) of the non-dummy atoms

`nfragm`

Total number of non-dummy fragments

`nofrag_1`

Integer array specifying for each non-dummy atom the fragment it belongs to.

`nofrag_2`

Integer array specifying for each non-dummy atom the fragment type it belongs to

`nuclab`

Names of the non-dummy atom types.

Section Fragments

(To be completed)

`FragmentFile`

Names of all used fragment files

`FragRun Ident,Title`

Job identification and title of each fragment run that is used in the current molecule

Section AtomTypes

(To be completed)

Section Properties

`AtomCharge Mulliken`

Atomic charges derived from Mulliken population analysis.

`Dipole`

Dipole moment in atomic units.

`FragmentCharge Hirshfeld`

Fragment charges derived from Hirshfeld analysis

`AtomCharge_initial Voronoi`

Atomic charges derived from Voronoi analysis for the initial (sum-of-fragments) charge density

`AtomCharge_SCF Voronoi`

Similar as the previous item, but now for the SCF density

`Electrostatic Pot. at Nuclei`

Coulomb potentials at the positions of the atoms, not including the contribution from the nucleus itself

Section Basis

Information about the (valence) basis set

`nbset`

The total number of basis 'sets', where a 'set' here means a Cartesian function set (3 for a *p*-type function, 6 for a *d*-type function, and so on), given by an entry in the 'list-of-basis-functions' in the data base file.

`nbaspt`

Cumulative number of basis sets (see previous variable, for 'set'), on a per atom type basis. Only non-dummy atoms (type) are considered. nbaspt(k) is 1+nr-of-basis sets up to, but not including atom type #k

`nqbas`

Main quantum number of each basis set. A 1s function has nqbas()=1

`lqbas`

Angular momentum quantum number of each basis set. The current implementation of ADF supports only *s, p, d,* and *f* basis functions, so the allowed lqbas values are 0, 1, 2, and 3

`alfbas`

The exponential decay parameters of the STO functions in the basis set

`basnrm`

Normalization coefficients for the basis sets

`naos`

The total number of basis functions, counting all Cartesian polynomials and all copies of the functions on the atoms of the pertaining atom type

`nbos`

The total number of Cartesian basis functions, *not* counting the copies of the functions on the different atoms of the atom type: the functions are defined per atom type and are (for nbos) counted only once. The next few variables relate to lists of basis functions that run from 1 to nbos: all the Cartesian polynomials, but counting the function only once per atom type. Essentially, this means counting all functions with distinct characteristics (apart from their geometrical center).

`nbptr`

Index array of the nbos functions, where the entries are the cumulative numbers of functions (+1) up to, but not including the atom type. The size of the array is (ntyp+1): one plus the number of (non-dummy) atom types.

`kx`

Powers of x of the nbos Cartesian STO basis functions

`ky`

Powers of y of the nbos Cartesian STO basis functions

`kz`

Powers of z of the nbos Cartesian STO basis functions

`kr`

Powers of r of the nbos Cartesian STO basis functions

`alf`

Exponential decay factors of the nbos Cartesian STO basis functions

`bnorm`

Normalization factors for the nbos Cartesian STO basis functions

`nprta`

Consider a list of all (naos) Cartesian STO basis functions, including copies of the functions on all atoms of the same atom type. Build that list by first taking all true valence functions on all atoms (loop over atom types, inner loops over atoms, inner loop over basis sets of the atom type, inner loop over Cartesian polynomials for the function set), then all auxiliary core-orthogonalization functions (similar loop structure). nprta(i) gives the index in that list of function #*i*, where *i* corresponds to a similar list of all naos functions in which the core and valence subsets are not separated.

`norde`

An array that runs over the non-dummy atom types. Each element gives the maximum of the main quantum number for all STO basis and fit functions corresponding to that atom type.

`lorde`

As norde, but lorde applies to the angular momentum quantum numbers.

Section Core

Information about frozen core orbitals and the Slater-type exponential functions used to describe them.

`nrcset`

The number of STO function sets to describe the frozen core orbitals in the calculation. The array is sized (0:llqcor,1:ntyp). llqcor is the maximum l-value in core orbitals (3), ntyp is the number of non-dummy atom types.

`nrcorb`

An array (0:llqcor, 1:ntyp) specifying the number of frozen core orbitals per *l*-value and per non-dummy atom type.

`ncset`

The total number of core expansion STO function sets, not counting copies on all atoms, and not counting the Cartesian polynomials (1 value per *p*-set, et cetera)

`ncorpt`

Index array: 1 + cumulative number of core expansion sets up to, but not including, the indexed atom type. The array runs from 1 to ntyp+1

`nqcor`

Main quantum numbers for the core expansion sets

`lqcor`

Angular momentum quantum numbers for the core expansion sets.

`alfcor`

Exponential decay factors for the core expansion sets.

`cornrm`

Normalization factors for the core expansion sets.

`ncos`

Total number of core expansion functions, counting all copies on different atoms of each atom type, and counting all Cartesian polynomials.

`nccpt`

Index array: 1 + cumulative number of core orbitals, counting all copies on different atoms and all Cartesian (sub) functions.

`ncptr`

Similar, but applying to the STO core expansion functions.

`ccor`

All core expansion coefficients, which express the core orbitals in the core expansion functions. The array stores the expansion coefficient sequence for each core orbital shell (not for each Cartesian sub

function) and only one sequence per orbital per atom type (no duplication for the different atoms of the atom type).

`npos`

An index array. For each atom type: the index where its data are stored on the TAPE12 core data file. npos(k) may be zero if no data for atom type #k are available on TAPE12.

`kcos`

The total number of core expansion functions, like ncos, but now counting only the truly independent functions. For instance: 5 functions per *d*-set, while in ncos there are 6 functions per *d*-set. The *s*-type combination in the 6-membered *d*-set is in the calculation projected out and does not represent a degree of freedom.

`s`

The (kcos,kcos) overlap matrix of the core expansion functions. Note that, since the dimension is (kcos,kcos), the *s*-type combination has been eliminated, and likewise for the 3 *p*-type functions in each *f*-set.

`idfcor`

Integer that indicates whether or not the core set contains *d*- and/or *f*-type functions. 1=yes, 0=no

`nd`

Total number of *d*-type core orbital sets (not counting the Cartesian sub functions)

`nf`

Total number of *f*-type core orbital sets (not counting the Cartesian sub functions)

`ndorb`

An array running over the *d*-type core orbital sets (loop over atom types, loop over atoms, loop over core orbitals with *l*=2). It gives for each the index of the orbital (the first of the Cartesian subset) in the overall list of all core orbitals in the molecule (including the spurious *s*-type functions in the *d*-sets, and so on)

`nforb`

Similar as ndorb, but now for the *f*-type core orbitals.

`cmat`

Overlap matrix between core-orbitals (ncos, counting all Cartesian functions including the *s*-type function in each *d*-set, et cetera), and the basis functions. In the list of basis functions, all core functions (the auxiliary orthogonalization functions) come before all true valence basis functions, see array NPRTA.

Section Fit

This section stores information about the fit functions, which are used for the Coulomb potential evaluation.

`Unrestr.SumFrag`

A logical that flags whether or not the fit coefficients have been set and stored for the sum-of-fragments, but adjusted for the unrestricted fragments option (see the keys UnrestrictedFragments, ModifyStartPotential).

`coef_SumFrag`

Fit coefficients pertaining to the sum-of-fragments charge density.

`coef_SCF`

SCF fit coefficients.

`nfset`

Total number of fit function sets (not counting the Cartesian sub functions, not counting the copies of the functions on the atoms of an atom type)

`nfitpt`

Index array: 1+the total number of fit function sets up to, but not including, the indicated atom type.

`nqfit`

Main quantum numbers of the fit sets

`lqfit`

Angular momentum quantum numbers of the fit sets

`alffit`

Exponential decay factors of the STO fit sets.

`fitnmr`

Normalization factors for the STO fit sets.

`nfos`

Total number of Cartesian fit functions, not counting copies on all atoms of an atom type, but including all (for instance, 6 for a *d*-set) Cartesian sub functions.

`nfptr`

Index array: 1+ total number of Cartesian (see variable nfos) fit functions, up to but not including the indicated atom type.

`nprimf`

Total number of Cartesian ('primitive') fit functions, counting also the copies on all atoms of each atom type.

`nsfos`

The total number of fully symmetric (A1 symmetry) fit function combinations that represent the true dimension (variational freedom) of the space of fit functions in the calculation.

`na1ptr`

Index array, like nfptr, but applying to the nsfos symmetric function combinations.

`niskf`

This refers to an atom-limited symmetry combination of primitive fit functions, in the code and some documentation indicated as a 'g'. A 'g' is the specific part of a molecule-wide A1 fit function combination (see nsfos) that consists of all the terms that are centered on one particular atom. The number niskf gives the total number of such 'g' function combinations.
To clarify this, consider an A1 fit function combination in the molecule. Assume, that it consists of a specific linear combination the following functions: a p-x function on atom A, its partner p-y function, and the corresponding p-x and p-y functions on atom B. (Atoms A and B must be symmetry equivalent). In this example we have one A1 function (in the list of nsfos such functions) and two 'g"s. Each 'g' consists of a p-x and a p-y function combination on a specific atom.

`iskf`

Compound index array. It runs over the niskf 'g' fit function combinations and has 4 entries for each function (1:4,1:niskf). The meaning of the entries is as follows. #1=number of the fit set (not counting the copies of fit functions on different atoms of an atom type, and not counting the Cartesian sub functions) this 'g' belongs to. #2=index where the combination coefficients for this 'g' start in the arrays cofcom and numcom (see next). #3=number of terms in the expansion of this 'g'. #4=number of the molecular fit A1 function combination this 'g' belongs to.

`na1cof`

Length of the arrays numcom and cofcom, see next

`numcom`

Numcom (and cofcom) consists of a sequence of smaller sub arrays. Each sub array gives the expansion of a 'g' function in terms of the Cartesian functions in the pertaining fit function set. The elements of numcom specify the particular Cartesian sub functions that participate in the expansion. Its values are therefore limited to lie between 1 and (L+1)(L+2)/2, where L is the maximum *l*-value occurring in the fit function sets.

`cofcom`

Compare numcom: cofcom gives the actual expansion coefficients for the expression of a 'g' function in primitive Cartesian fit functions.

Section Num Int Params

Numerical integration parameters: the general precision parameter, but also more technical parameters used by the grid-generating modules.

`method`

Label of the method used to generate the grid. Usually: 'polyhedra'

`accint min`

Minimum integration precision parameter. It is the lower bound of the range in which the value of the actual numerical integration precision parameter may vary.

`accint max`

Maximum value of the precision general parameter

`accint`

Actual value of the precision parameter. This variable governs by default almost all other integration parameters.

`ldim`

In fact, this a geometric parameter: the number of dimensions in which the system is periodic. For molecules this is zero.

`PointChargeTypes`

The number of point charges types used in the calculation. Point charges belong to a different point charge type if, and only if, their strengths are not equal.

`accsph`

The precision parameter that determines the (radial) integration grid in the atomic spheres

`accpyr`

The precision parameter that determines the general precision level of the grid in the atomic polyhedra

`accout`

The precision parameter that determines the general precision level of the grid in the outer region

`accpyu`

The precision parameter that determines the 1D grid along the first direction in the quadrangles and triangles of the bases of the atomic pyramids

`accpyv`

The precision parameter that determines the 1D grid along the second direction in the quadrangles and triangles of the bases of the atomic pyramids

`accpyw`

The precision parameter that determines the 1D radial integration in the atomic pyramids, between the atomic sphere surfaces and the pyramid basis

`frange`

Estimated maximum range of functions, to determine how far the integration grid has to extend outwards, away from the molecule

`rspher`

An array with the radii of the atomic sphere (a value per atom type)

`rsph0`

The smallest sphere radius

`rsphx`

The largest sphere radius

`dishul`

The distance between the innermost boundary planes, which separate the atomic pyramids from the outer region, and the surfaces of the outermost atoms

`nouter`

The number of intervals in which the outward (radial) integration in the outer region is broken up

`outrad`

The precision parameter that determines the outward radial integration in the outer region

`outpar`

The precision parameter that determines the 2D integrals in the outer region parallel to the boundary planes

`linteg`

An array with maximum angular momentum quantum numbers (one value per atom type), to determine the angular integration grid in the atomic spheres

`lintgx`

Maximum of linteg()

`linrot`

Angular momentum quantum number to determine the rotational integration parameter around the molecular axis (in linear molecules only)

`ntyps`

The number of atom types as seen by the numerical integration grid generator. This means in practice: the number of non-dummy atom types plus the number of point charge types.

`nnucs`

The number of atoms as seen by the numerical integration grid generator. This means in practice: the number of non-dummy atoms plus the number of point charges.

`qatm`

Nuclear charges for all ntyps atom types

`nratst1`

The numerical integration grid generator automatically determines the symmetry of the nuclear (nnucs atoms!) frame and then puts the atoms in sets of symmetry equivalent ones. nratst1() is an array (0:ntyps) that contains the cumulative number of atoms in the symmetry sets. nratst1(k) is the total number of atoms in the sets up to and including set #k

`xyzatm`

Cartesian coordinates of the atoms.

`linteg all`

Similar to array linteg(), extended to include also the point charge types

`npowx`

Maximum power of the radial variable *r*, in the set of test functions that the grid generator uses to tune the grid

`alfas`

An array that stores the exponential decay factors of all test functions, ordered by atom type and by the power of the radial variable *r*.

Section Symmetry

Symmetry related data.

`nogr`

The number of symmetry operators in the point group used in the calculation. NB, for the special cases of infinite symmetries, only the operators corresponding to finite elements are counted. Therefore, ATOM has nogr=1 (only the unit operator); C(LIN) has nogr=1, D(LIN) has nogr=2.

`faith`

An array that stores all the (3,3) symmetry operator matrices in the real space representation

`nsetat`

The number of sets of symmetry equivalent atoms under the used symmetry

`napp`

An array that stores for each atom the number of the symmetry set it belongs to

`notyps`

An array that stores for each set of symmetry equivalent atoms, the atom type to which the set belongs

`noat`

Map between the normal list of atoms and the symmetry sets. When you loop over the symmetry sets and, inside, loop over the atoms in each set, you thereby run over the index of noat(). The value points to the position of that atom in the original (not set-ordered) list.

`ntr`

An array (nogr,nnuc) that stores for the each atom A and each symmetry operator R, the atom onto with A is mapped by R. The row index runs over all symmetry operators, the column index over the atoms.

`npeq`

The number of symmetry unique pairs of atoms

`jjsym`

An array that runs over the npeq sets of symmetry equivalent atom pairs. Its value gives for the indicated set the index of a (c.f. the first) atom pair in that set.

`jasym`

An array that runs over the npeq sets of equivalent atom pairs. Its value gives for the indicated the set the number of pairs in that set.

`ja1ok`

An array (1:npeq), with values 0 or 1. 1=the pair density can be fitted using A1 fit functions only. 0=all fit functions (on the involved atoms) are to be used. The value 1 may arise because of symmetry properties, or because the distance between the atoms is so large that the inaccuracy from using only A1 fit functions can be neglected.

`ntr_setat`

A condensed variety of array ntr: the columns are not the atoms, but the nsetat sets of symmetry equivalent atoms. The value is the index of the atom, onto which a representative (c.f. the first) atom of the indicated symmetry set is mapped by the given symmetry operator.

`igr`

A code that fixes, together with nogr and ngr, the point group symmetry. See the header of routine adf/maisya for a list

`ngr`

One of the code components that fix the symmetry group. See routine adf/maisya

`grouplabel`

Schönfliess symbol as used in ADF

`nsym`

The number of symmetry representation (including subspecies) used in the calculation.

`norb`

For each of the nsym representations the number of basis function combinations (SFOs) that belong to it.

`nfcn`

For each of the nsym representations the number of primitive atom centered basis functions that participate in the representation.

`ncbs`

For each of the nsym representations the number of core orthogonalization functions that participate in the representation.

`jsyml`

For each of the nsym representations: if it belongs to a one-dimensional irrep, the value is 1, otherwise: for the first subspecies in the irrep the value is the dimension of the irrep, for the other subspecies in the same irrep the value is 0

`symlab`

For each of the nsym representations the label (string) of the representation

`norboc`

An array (-2:2,nsym). The column runs over the symmetry representations. The positive row indices (1,2) specify for spin-A and spin-B (the latter only if the calculation is spin-unrestricted), the highest non-

empty orbital. The negative indices (-1,-2) specify for spin-A and spin-B (if the unrestricted fragment option is used) the total number of non-empty SFOs. The zero row index specifies the number of non-empty SFOs, before applying any fragment occupation changes.

Section Spin_orbit

(To be completed)

Section Energy

`XC energies`

16 elements of an array enxc(2,2,4): exchange-correlation energies of various charge densities:
first index: 1=exchange term, 2=correlation term
second index: 1=lda tern, 2=gga term
third index: 1=energy of fragments (summed over fragments), 2=energy of sum-of-fragments density, 3=energy of orthogonalized fragments, 4=SCF.

`Pauli TS Correction (LDA)`

Correction to the 'Transition State' method to compute terms in the bonding energy, in this case the Pauli exchange energy term. The Pauli TS Correction is not separately printed in the standard output file, but included in the Pauli interaction term.

`Pauli FitCorrection`

The first-order correction to the Pauli exchange interaction term, for the error in the Coulomb energy due to the fit incompleteness. This correction term is not printed in the output file but included in the Pauli interaction term

`Elstat Core terms`

An obsolete variable, not used in the energy computation

`Elstat Fitcorrection`

The first-order correction to the electrostatic interaction term (putting the fragments together, without any relaxation of Pauli orthogonalization), for the error in the Coulomb energy due to the fit incompleteness

`Orb.Int. FitCorrection`

The first-order correction to the electrostatic interaction term in the SCF relaxation energy (Orbital Interactions), for the error in the Coulomb energy due to the fit incompleteness. This term is not printed (anymore) separately, but incorporated in the symmetry-specific interaction terms.

`Orb.Int. TSCorrection (LDA)`

The difference between the representation-specific orbital interaction terms added, and a straightforward computation of the SCF relaxation energy is the result of the neglect of higher order terms in the Taylor expansion that underlies the 'Transition State' method. This difference, therefore, corrects exactly this neglect. It is not printed separately anymore in the output, but incorporated in (distributed over) the representation-specific orbital interaction terms.

`Ebond due to Efield`

Bond energy term due to any homogeneous electric field

`Corr. due to Orthogonalization`

For analysis purposes, the concept of 'orthogonalized fragments' has been introduced and the bonding energy is split in a part that describes the difference between the sum-of-fragments situation and the orthogonalized-fragments density at the one hand, and the SCF relaxation (from the orthogonalized fragments density) at the other. Both terms contain a first order fit correction term. The result of adding the two parts is not identical to computing the total bonding energy directly and applying the first order correction to that approach. The difference is given by this term, which therefore corrects for the additional second order fit errors caused by using the orthogonalized fragments split-up

`SumFragmentsSCF FitCorrection`

The 'true' first order fit correction for the complete bonding energy, resulting from a direct calculation that takes the sum-of-fragments as starting point and the SCF as final situation, without the intermediate step of orthogonalized fragments.

`Pauli Efield`

The contribution to the Pauli interaction energy due to any electric field

`Orb.Int. Efield`

The contribution to the SCF relaxation energy (orbital interactions) due to any electric field

`Electrostatic Interaction`

The electrostatic interaction energy including any first order fit correction (if computed from the fit density)

`Pauli Total`

The Pauli exchange (orbital orthogonalization) interaction energy

`Steric Electrostatic`

INCORRECT. Do not use. The electrostatic interaction energy including any first order fit correction (if computed from the fit density)

`Steric Total`

The total steric interaction energy, consisting of the electrostatic and the Pauli interactions

`Orb.Int.` *Irrep*

Irrep stands for one of the irreps of the point group symmetry. The value gives the orbital interaction (SCF relaxation) term for that symmetry representation

`Orb.Int. Total`

The total orbital interaction energy

`SCF Bond Energy`

Total bonding energy

`elstat`

INCORRECT. Do not use. Electrostatic interaction energy. Same as the 'Electrostatic Interaction' variable in this section

`Bond Energy`

Total bonding energy, same as the 'SCF Bond Energy' variable

`Pauli Kinetic`

Kinetic energy term in the Pauli exchange interaction energy

`Pauli Coulomb`

Coulomb energy term in the Pauli exchange interaction energy

`Pauli Kinetic+Coulomb`

Sum of the kinetic and Coulomb terms in the Pauli exchange interaction energy

Section Point_Charges

`NumberofPointCharges`

The total number of point charges used

`PointCharges`

The array with point charge values: (4,np), where np is the number of point charges and the 4 components are, respectively, the x y z components and the strength.

Section GeoOpt

Optimization data.

Where references are made to the list of atoms, the atoms are assumed to be in internal order. This may be different from the input-list of atoms.

`nfree`

number of independent optimization variables

`idfree`

indices (3,nr-of-atoms) for all atomic coordinates referring to the optimization variables (values 1..nfree) and/or LinearTransit parameters (values nfree+k, k being the k-th LT parameter). A zero value means that the coordinate is frozen.

`all freedoms`

A logical the flags whether or not all fundamental degrees of freedom in the system are allowed to vary. This is not the case when constraints are applied.

`Gradients`

The most recent values for the derivatives of the energy with respect to the atomic coordinates (cartesian or z-matrix, depending on the type of optimization variables).

`Displacements`

The most recent step executed for the atomic coordinates (optimization variables)

`kmatrix`

The connection matrix.

`Hessian_CART`

The Hessian matrix (second derivatives) as a n*n matrix, in the Cartesian coordinates representation. Note that the reduced storage mode (typically, Fortran upper-triangular) is not applied.

`Hessian_ZMAT`

Same, but now in the internal coordinates representation

`Hessian inverted_CART`

The *inverted* Hessian, in Cartesian coordinates

`Hessian inverted_ZMAT`

Likewise, in internal coordinates

Note: in most cases only one, or maybe two of the Hessian cases are present on TAPE13. They can be transformed into each other quite easily. The order of atoms is the same as in the input.

`xyz old`

cartesian coordinates at previous geometry cycle

`zmatrix old`

idem for internal coordinates

Section TS

Information about the Transition State search

`modtrc`

Defines the initial search direction. Positive value *n*: the n-th Hessian eigenvector (default:1). Negative value *n*: the Hessian eigenvector with the largest (absolute value) component in the n-th optimization variable

`itrace`

Index of the Hessian eigenvector that is being followed

`neghes`

The assumed number of negative eigenvalues of the Hessian at the Transition State. Should be 1: searches for higher-order transition states are not supported.

`mode to follow`

Direction vector in atomic coordinates (Cartesian or Z-matrix, depending on the variable geocrd) that corresponds to the current estimate of the unique Hessian eigenvector with negative eigenvalue

Section LT

Information about the Linear Transit calculation

`nr of points`

The total number of LT points to be computed.

`current point`

Index of point that is currently computed

`Energies`

Energy values in the LT points

`Dipole`

Dipole moments in the LT points

`Parameters`

LT parameters, initial and final values (along the path, the values are obtained by even-spaced linear interpolation)

`atmcrd`

ZMAT if a Z-matrix structure (connection matrix) is available. CART otherwise. Used for printing

`geocrd`

Type of coordinates to optimize and scan along the path (CART or ZMAT)

`xyz`

Cartesian coordinates in the LT points (3,natoms,nlt)

`zmatrix`

Internal coordinates in the LT points

`AtomCharge Mulliken`

Mulliken atomic charges in the LT points

`FragmentCharge Hirshfeld`

Hirshfeld fragment charges in the LT points

`AtomCharge_initial Voronoi`

Voronoi atomic charges corresponding to the sum-of-fragment densities in the LT points

`AtomCharge_SCF Voronoi`

Voronoi atomic charges corresponding to the SCF densities in the LT points

Section IRC

This section contains general information about the IRC (Intrinsic Reaction Coordinate) calculation. Details of the computed reaction path are in sections IRC_Forward and IRC_Backward.

`atmcrd`

ZMAT is a Z-matrix structure (connection matrix) is available. CART otherwise

`geocrd`

CART or ZMAT: the type of coordinates to change, optimize and trace

`PointStatus`

A string status variable of the current IRC point. Value can be 'DONE' (if the point has been computed), 'EXEC' if its computation has not yet finished.

`nfree`

Number of optimization coordinates that can be varied. See section GeoOpt

`idfree`

(3,natoms) pointers to the optimization variables for each of the atomic coordinates. A zero means: frozen by constraint

`xyz`

Cartesian coordinates

`kmatrix`

Connection matrix, if a Z-matrix structure is available

`zmatrix`

Internal coordinates

`Energies`

Energy at the Transition State

`Dipole`

Dipole moment at the Transition State

`Gradients`

Computed energy gradients at the (assumed) Transition State (should be very small)

`AtomCharge Mulliken`

Mulliken atomic charges, for the TS geometry

`AtomSpinDen Mulliken`

Atomic spin densities (Mulliken) at the TS

`AtomCharge_initial Voronoi`

Voronoi atomic charges at the TS, from the sum-of-fragments density

`AtomCharge_SCF Voronoi`

Similar, for the SCF density

`modtrc`

Defines the start direction for the IRC path. A positive value *n* selects the n-th eigenvector of the Hessian. A value -1 selects the gradient vector (which must then, of course, not be exactly zero). A value -2 specifies that the start direction is specified in the input file.

`step`

Step length (in mass-weighted metric) between successive points of the IRC path.

`stepMin`

The minimum value for the step

`stepMax`

The maximum value for the step

`Hessian inverted_ZMAT`

Inverse Hessian in internal coordinates.

`lfree`

The number of independent optimization step directions (for the restricted optimization orthogonal to the IRC path.

`vfree`

Direction vectors (3,natoms,lfree) for the independent optimization directions

`GradientVector`

The current gradient vector (during optimization)

Section IRC_Forward

Information about the 'forward' IRC path. The choice, which direction down from the Transition State is forward or backward is arbitrary. By definition, in ADF the forward direction is in the positive direction along the first Hessian eigenvector, for which the sign convention is that the largest coefficient is positive.

`PathStatus`

Status (string) variable for the 'forward' half of the IRC path. May be 'EXEC', or 'DONE', 'UNKNOWN', 'WAIT', 'OFF'

`PointStatus`

Status variable for the current point at the 'forward' path. May be 'DONE', 'EXEC'

`nset`

Size of arrays to store data in the IRC points along the path. Will be increased when too small

`pivot`

Coordinates of the current pivot point

`xyz`

Cartesian coordinates of the IRC points (3,natoms,nset)

`zmatrix`

Internal coordinates of the IRC points (3,natoms,nset)

`Path`

Lengths measures in mass-weighted metric along the path to the IRC points

`Curvature`

Local curvature values of the path at the IRC points

`Energies`

Energy values at the IRC points

`Gradients`

Energy gradients at the IRC points (one value: the gradient along the path. The orthogonal components are presumably zero)

`Dipole`

Dipole moments at the IRC points

`AtomCharge Mulliken`

Mulliken atom charges at the IRC points

`FragmentCharge Hishfeld`

Hirshfeld fragment charges at the IRC points

`AtomCharge_initial Voronoi`

Voronoi atomic charges at the IRC points, corresponding to sum-of-fragments densities

`AtomCharge_SCF Voronoi`

Voronoi atomic charges at the IRC points, corresponding to the SCF densities

`CurrentPoint`

Integer index of the current IRC point (in the set of nset points)

`step`

Current step length

Section IRC_Backward

All entries in this section match those in the section IRC_Forward. Of course, here they refer to the other half of the IRC path.

Section Freq

This section contains information about (progress) of the Frequencies calculation and results.

`kountf`

An integer counter that keeps track of how many geometric displacements have been carried out to scan the potential energy surface around the equilibrium

`nraman`

Integer to flag whether or not Raman intensities are computed

`numdif`

Integer to determine the type of numerical differentiation (of gradients, to get the second derivative): 1=one-sided, 2=two-sided displacements.

`disrad`

Size of displacements of Cartesian coordinates or bond lengths (in case of displacements in internal coordinates)

`disang`

Size of displacements of angular coordinates

`geocrd`

Type (string) of coordinates to displace: CART or ZMAT

`atmcrd`

ZMAT if a z-matrix structure is available. This determines printed output but does not affect the computation. Compare the variable geocrd

`nfree`

The number of degrees of freedom

`idfree`

An array (3,natoms) that stores for each atomic coordinates (Cartesian or internal, depending on geocrd) the number of the (1..nfree) variational freedom it corresponds to. If zero, the coordinate is frozen by constraint.

`xyz`

Cartesian coordinates of the equilibrium geometry

`kmatrix`

Connection matrix that defines a Z-matrix

`zmatrix`

The Z-matrix variable values of the equilibrium geometry

`all freedoms`

Logical: true if all atomic coordinates are allowed to be displaced, not restricted by constraints.

`nr of atoms`

The total number of atoms, including dummy's

`rigids`

Vectors of rigid motion directions, expressed in the atomic coordinates (3,natoms,6)

`Dipole previous`

The dipole moment of the previous geometry. This is used to compute dipole derivatives by numerical differentiation. The 'previous' geometry is the equilibrium geometry in case of one-sided displacements.

`Dipole`

The dipole moment corresponding to the current geometry

`Dipole derivatives`

The matrix of dipole derivatives with respect to atomic displacements

`Polbty previous`

The polarizability tensor (6 elements, triangular representation) of the 'previous' geometry. See the remarks about the dipole moment

`Polbty`

The polarizability tensor corresponding to the current geometry

`Polbty derivatives`

The matrix of derivatives (w.r.t. the atomic coordinates) of the polarizability tensor

`Gradients`

The energy gradients corresponding to the current geometry

`Gradients previous`

The energy gradients of the 'previous' geometry. See the remarks about 'previous' dipole moment

`Force constants`

The matrix of force constants (second derivatives), built up during the frequencies calculation.

`xyz displaced`

The Cartesian coordinates of the current (displaced) geometry

`zmatrix displaced`

Internal coordinates of the current (displaced) geometry

`Dipole derivatives_CART`

Dipole derivatives with respect to Cartesian coordinate changes

`Hessian_CART`

The Hessian matrix in Cartesian coordinates, computed at the end, when the construction of the 'Force constants' has been completed.

`Frequencies`

An array with harmonic frequencies.

Sections Ftyp n

n is an integer. All such sections give general information about fragment type #n, and more specifically about the ADF calculation that produced the corresponding fragment file.

`jobid`

Job identification of the fragment run

`title`

Title of that calculation

`nsym`

Number of symmetry representations (subspecies) used

`norb`

For each representation the size of the Fock matrix (variational degrees of freedom)

`bb`

Labels of the subspecies

`igr`

(Partial) code for the point group symmetry

`ngr`

(Partial) code for the point group symmetry

`grouplabel`

Schönfliess symbol of the point group symmetry (of the fragment calculation)

`nfcn`

An array over the representation: for each subspecies the number of primitive STO basis functions that participate in that subspecies

`jsyml`

An array (1:nsym). Value 1 means that the corresponding subspecies belongs to a 1D irrep. A value larger than 1 means a correspondingly higher dimensionality of the irrep *and* indicates that that subspecies is the first in that irrep. A value 0, finally, means that it is not the first subspecies in its irrep.

`nfrag`

Number of fragments used in that fragment calculation

`natom`

Number of atoms in the fragment

`naos`

Number of primitive atomic basis functions

`nrat 1`

Maps the atoms of this fragment (the '1' signals the first fragment of this type) onto the list of all atoms

`rotfrg`

Rotation matrix to map the fragment coordinates as they are on the fragment file onto their actual orientation in the molecule

`nsot`

Total number of MO degrees of freedom, summation over all subspecies

`nmis`

The number of symmetry representations that could not be spanned by the basis set

`mis`

Indices of the missing symmetry representations

Sections Ftyp n?

n stands for the n-th fragment type. The ? stands for one of the symmetry representations (of the point group symmetry used in the fragment calculation)

`froc`

MO occupation numbers for the MOs in this subspecies

`eps`

Orbital energies
When they result from a ZORA calculation, the non-scaled values are stored on file (the scaled values are printed in the standard output file).

`eigvf`

Fragment MO eigenvectors, expressed in all the primitive atomic orbitals of the fragment.

`nsos 1`

Total number of MOs in this subspecies: size of variational problem

`nbas 1`

Number of primitive atomic basis functions that participate in this subspecies

`npart 1`

Indices that give for each of the nbas functions, the number of the basis function in the list of all basis functions

`FO 1`

The fragment MOs (nbas*nsos coefficients)

`nocc 1`

Number of non-empty orbitals

Section Freq Symmetry

Information about the true (possibly input-specified) symmetry of the equilibrium geometry (in a frequencies calculation). The displaced geometries may loose symmetry. Therefore, the program uses NOSYM symmetry, internally, for a frequencies calculation. The 'true' symmetry of the system is used for analysis purposes.

`nr of operators`

Number of symmetry operators used

`operators`

(3,3) matrices of the operators

`nr of symmetries`

Number of subspecies

`symmetry labels`

Names of the subspecies

`atom indices`

List of indices to map the symmetry-ordered atoms (loop over symmetry sets, loop over atoms in each set) to the 'normal' list of all atoms

`nr of atomsets`

Number of sets of symmetry equivalent atoms

`atom mappings`

Integer array that provides mapping (back and forth) between the atom list in the input file and the internally used list, which is atom type driven

`atomset indices`

The number of atoms in each of the sets of symmetry equivalent atoms

`nr of displacements_X`

(X must be one of the symmetry representations.) The number of symmetry-combined atomic displacements that transform as X

`degeneracy_X`

Degeneracy of X

`displace_X`

The actual displacement direction vectors (3,natoms,N). N is the number of symmetry displacements for X.

`nr of rigids_X`

The number of rigid motion direction vectors that transform as symmetry representation X

`displ_InputOrder_X`

The displacement vectors, but now expressed in the atomic coordinates using the ordering of atoms in the input file

`NormalModes_X`

Harmonic frequency normal modes in representation X

`Frequencies_X`

The harmonic frequency values

`IR intensities_X`

The infrared intensities

Sections X

X stands here for the label of a subspecies of the point group symmetry, for instance A1. Depending on the point group symmetry, there may be many such sections, each corresponding to one of the subspecies. All such sections have an identical structure.

`nmo_A`

The number of MOs with spin-A, for which the coefficient vectors are calculated. During the SCF this may be severely reduced, at the end it is usually the complete basis in the pertaining symmetry representation.

`nmo_B`

Similar for spin b. This variable is not present in a restricted calculation.

`SFO`

The definition of the SFOs in the representation, consisting of expansion coefficients in terms of the primitive atomic STO basis functions

`frocf`

The occupation numbers of the SFOs in this representation

`npart`

A list of indices of the bas functions that are used in this symmetry representation

`froc_A`

The occupation numbers of the MOs in the representation, for spin-A

`froc_B`

Similar for spin-B, if a spin-unrestricted calculation is performed

`smx`

Overlap matrix between core functions and SFOs

`frocor`

SFO occupation numbers

`Orth-Bas`

The orthogonalized fragment orbitals in the BAS representation.

`Low-Bas`

The Lowdin orbitals in the BAS representation: the matrix to transform the MOs from Lowdin representation (orthonormalized SFOs) to the BAS representation

`Eigen_Bas_A`

mo expansion coefficients in the bas representation for all nmo_A orbitals. The coefficients run over all bas functions indicated by npart

`Eigen_Bas_B`

Similar for spin-B, if present

`eps_A`

The orbital energies for the nmo_A orbitals of spin-A
When they result from a ZORA relativistic calculations, the non-scaled values are stored on file. (The scaled energies are printed in standard output.

`eps_B`

Similar for spin-B, if present

`Eig-CoreSFO_A`

MOs expressed in SFOs, for spin-A MOs

`Eig-CoreSFO_B`

Similar for spin-B

Sections Atyp n X

Each such section contains the (core- and possibly also valence-) radial density and potential of one particular atom type. *X* is the atom type label and *n* is an index running over all atom types in the calculation. The list of all atom types is printed on standard output in the early geometry section.

The radial densities and potentials may be represented as simple tables - a sequence of values for *r*, the distance to the nucleus, and the corresponding density or potential - or as a piecewise expansion in Chebyshev polynomials over a sequence of intervals (r1,r2).

The core density and potential have been constructed from the Frozen Core orbitals, which are defined in the section Core. If a TAPE12 (corepotentials) file has been attached to the calculation the core data is read off from that TAPE12 and stored also.

`rx val`

Maximum r-value for which the valence density is non-negligible

`nrint val`

Number of intervals for piecewise expansion of the valence density in Chebyshev polynomials

`rup val`

Arrays (1..nrint) of upper bounds of the intervals. The lower bound of the first interval is zero

`ncheb val`

Array (1..nrint) with the number of expansion coefficients for each interval

`ccheb val`

Coefficients of the expansion. All coefficients, for all intervals, are stored contiguously in one linear array. The parts pertaining to a particular interval are determined by using the arrays ncheb()

`nrad`

Number of points used in the direct tabular representation of the atomic densities and potentials

`rmin`

The first r-value of the table: the radial grid is defined by a first value (rmin), a constant multiplication factor defining rk+1 w.r.t. rk (rfac, see next), and the total nr of points (nrad).

`rfac`

The multiplication factor of the radial grid

`valence den`

The valence density, in a table of nrad values.

`valence pot`

Similar for the Coulomb potential of the density, including a nuclear term Q/r, such that the long-range monopole term in the potential is zero

`qval`

The number of electrons contained in the valence density

`rx core`

Maximum r-value for which the core density is non-negligible

`nrint core`

Number of intervals for piecewise expansion of the core density in Chebyshev polynomials

`rup core`

Arrays (1..nrint) of upper bounds of the intervals. The lower bound of the first interval is zero

`ncheb core`

Array (1..nrint) with the number of expansion coefficients for each interval

`ccheb core`

Coefficients of the expansion. All coefficients, for all intervals, are stored contiguously in one linear array. The parts pertaining to a particular interval are determined by using the arrays ncheb()

`qcore`

The number of electrons contained in the core density

`core den`

The core density, in a table of nrad values.

`core pot`

Similar for the Coulomb potential of the density, including a nuclear term Q/r, such that the long-range monopole term in the potential is zero

Section LqbasxLqfitx_xyznuc

This section will be removed again in the future. Temporarily it serves to transfer data from the calling program to the grid generator.

`lqbasx`

An array with for each atom type the maximum angular moment quantum number in the basis functions for that type

`lqfitx`

An array with for each atom type the maximum angular moment quantum number in the fit functions for that type

`xyznuc`

Cartesian coordinates of the non-dummy atoms

Section GenptData

This section will be removed in the future. It serves, temporarily, to transfer data from the calling program to the numerical integration grid generator. Most of the entries here occur also in other sections but are packed together as replacement for previous common block structure.

`numint`

Integer code for the type of integration grid. Usual value: 2 (polyhedra method)

`iexcit`

Integer flag for excitations (response) calculation

`lpolar`

Integer flag for polarizability (response) calculation

`ldim`

Number of dimensions of periodicity

`mdim`

Dimensionality of the molecule, for instance a linear molecule has mdim=1

`r0mult`

A technical parameter that sets the radius outside which the multipole part of the fit coulomb potential functions is separated (from the exponentially decaying part), for separate treatment in the evaluation of the molecular coulomb potential.

`avec`

(3,3) matrix with lattice vectors. Only the (ldim,ldim) sub matrix is significant.

`bvec`

Inverse of avec (apart from a factor of 2 pi): lattice vectors in reciprocal space.

`ngimax`

Maximum number of geometry optimization iterations

`llbloc`

Block length determination parameter (maximum)

`ipnbl`

Number of integration blocks processed by the current process

`nbleqv`

The number of symmetry equivalent blocks to each symmetry unique block of points. This value is 1 if any equivalent blocks are not constructed and used.

`ngmax`

The number of integration points, accumulated over all parallel processes

`nblock`

The number of integration blocks

`lblock`

The block length

`lblx`

An upper bound of the block length applied during the computation of the block length

`nmax`

The number of integration points generated by this process

`twopi`

Value of the constant 2π

`fourpi`

Value of the constant 4π

Section Multipole matrix elements

Information in a response calculation

`dipole elements`

The matrix elements of the 3 dipole operator components between occupied and virtual orbitals: outer loop over the operators (in order: y, z, x), loop over virtual MOs, inner loop over occupied MOs

`quadrupole elements`

Similar as for dipole. Order of operators:
√(3)*xy
√(3)*yz
z2-(x2+y2)/2
√(3)*xz
√(3)*(x2-y2)/2

`octupole elements`

Similar as for dipole and quadrupole. Order of operators:
√(10)*y*(3*x2-y2)/4
√(15)*xyz
√(6)*y*(4*z2-x2-y2)/4
z*(z2-3(x2-y2)/2)
√(6)*x*(4*z2-x2-y2)/4
√(15)*z*(x2-y2)/2
√(10)*x*(x2-3y2)/4

`hexadecapole elements`

Similar as for dipole and quadrupole. Order of operators:
√(35)*xy*(x2-y2)/2
√(70)*z*(3x2y-y3)/4
√(5)*xy*(6z2-x2-y2)/2
√(10)*(4yz3-3yz*(x2+y2))/4
(8z4-24*z2*(x2+y2)+3(x4+2x2y2+y4))/8
√(10)*(4xz3-3xz*(x2+y2))/4
√(5)*(x2-y2)*(6z2-x2-y2)/4
√(70)*z*(x3-3xy2)/4
√(35)(x4-6x2y2+y4)/8

Section Irreducible matrix elements

Information in a response calculation

`irreducible dipole elements`

The dipole matrix elements between occupied and virtual MOs, as in the section Multipole matrix elements. Here, however, the matrix elements are ordered by symmetry representations and 'symmetry zeros' are omitted. The stored arrays, however, have the same size as in the previous section. See the implementation for details about the storage of this data. (Directory $ADFHOME/adf/response/)

`irreducible quadrupole elements`

Similar as for the dipole elements

`irreducible octupole elements`

Similar as for the dipole elements

`irreducible hexadecapole elements`

Similar as for the dipole elements

Section ETS

Technical data used in the ets procedure.

`nff`

Size of array ncspt (next)

`ncspt`

Pointer array to find, for each atom type, the first element corresponding to that atom type's section in the arrays ncsett, alfcst, and cfcset, see below

`ncs`

Size of the matrices ncsett, alfcst, and cfcset, see below

`ncsett`

Build a list of products of core orbital expansion functions, taking only the one-center products and looping over the atom types (not the atoms). ncsett stores the powers of the radial variable *r* for the products (from the main quantum numbers, one subtracted). A product of a 1s and a 2p yields ncsett()=1

`alfcst`

Similar as ncsett: the sum of the exponential decay factors of the factor functions

`cfcset`

The density matrix corresponding build from the frozen core orbitals (all atom types, but no copies for the distinct atoms of a type), in the representation of the core orbital expansion functions. Stored are, per atom and per *l*-value (0..3) the upper-triangles of the corresponding density matrices, one after the other, all in cfcset

`nnuc`

The number of (non-dummy) nuclei

`qcore`

For each atom the number of electrons summed over its core orbitals, resulting from analytical integration of the core orbital expansions in STO core expansion functions.

Using Data from TAPE21

An ASCII dump of TAPE21 (complete or partial) can be obtained with the kf utility dmpkf, see the utilities document. Alternatively you may build your own small program to extract any required information, using the KF library routines in the ADF package. Consult the KFS documentation for a description of this software.

Representation of functions and frozen cores

adf uses the cartesian representation for the spherical harmonics part in functions:

f(x,y,z)=xaybzcrde-ar

The angular momentum quantum number *l* is then given by *l*=a+b+c, and the main quantum number *n*=*l*+d+1.

There are (*l*+1)(*l*+2)/2 different combinations of (a, b, c) for a given *l*-value, rather than (2*l*+1). The excess is caused by the presence of spurious non-*l* Functions in the set; a Cartesian d-set for instance consists of six functions, five of which are true d-functions while one linear combination is in fact an s-type function (x2+y2+z2). Only the five true d-combinations are actually used as degrees of freedom in the basis set, but lists of primitive basis functions (bas) for instance run over all Cartesian functions including the improper ones.

A function *set* in ADF is characterized by the quantum numbers *l* and *n*, and by the exponential decay factor a. A set thus represents (*l*+1)(*l*+2)/2 Cartesian functions and (2*l*+1) degrees of freedom.

The atomic frozen core orbitals are described as expansions in Slater-type functions; these are not the functions of the normal basis set but another set of functions, defined on the data files you use in Create mode.

Orthogonality of the valence space to the frozen core states is enforced as follows: for each frozen core shell (characterized by the quantum numbers *l* and *n*: all orbitals with *m*=-*l*...+*l* are identical apart from rotation in space) the set of valence basis functions is augmented with a so-called core orthogonalization function set. You may conceptually interpret the core orthogonalization functions as single zeta expansions of the true frozen core states. Each of the normal valence basis functions is now transformed into a linear combination of that valence function with all core orthogonalization functions, where the coefficients are uniquely defined by the requirement that the resulting function is orthogonal to all true core functions.

So the list of all Cartesian basis functions is much larger than the degree of freedom of the basis: it contains the spurious *non-l* combinations and it contains also the core orthogonalization functions.

Evaluation of the charge density and molecular orbitals

TAPE21 contains all the information you need to evaluate the charge density or a Molecular Orbital (MO) in any point in space. Most of the information is located in section Basis:

A list of function characteristics (kx,ky,kz,kr,alf), including the core orthogonalization functions. This list does *not* run over *all* bas functions used in the molecule: if a particular function is used on the atoms of a given atomtype, the function occurs only once in the list, but in the molecule it occurs as many times as there are atoms of that type.

With array nbptr you locate the subsections in the function list that correspond to the different types of atoms: for atom type i the functions nbptr(i)...nbptr(i+1)-1.
The distinct atom types are listed in an early section of the standard output file.

Array nqptr gives the number of atoms for type i: nqptr(i+1)-nqptr(i). With this information you construct the complete list of all functions. Repeat the subsection of type i as many times as there are atoms of that type: the complete list can be considered to be constructed as a double loop, the outer being over the atom types, the inner over the atoms that belong to that type.
The total 'overall' list of functions you obtain in this way contains naos functions.
Note that in this way we have implicitly also defined a list of all atoms, where all atoms that belong to a particular atom type are contiguous. This list is the so-called 'internal' atom ordering, which may not be identical to the order in which atoms were specified in input, under atoms.

For a given symmetry representation (Sections S) the array npart gives the indices of the basis functions in the overall list that are used to describe orbitals in this representation.
In case of an unrestricted run the array npart applies for either spin: the same basis functions are used; the expansion coefficients for the molecular orbitals are different of course.

In the symmetry-representation sections Eigen_bas gives the expansion coefficients that describe the MOs. The expansion refer to the functions indicated by npart, and the function characteristics are given by the arrays kx,ky,kz,kr,alf, *and* bnorm, i.e. the expansion is in *normalized* functions.

The value of an MO is now obtained as a summation of values of primitive basis functions. For the evaluation of any such basis function you have to take into account that its characteristics are defined in the local coordinate system of its atom.

To obtain the charge density you sum all MOs, squared and multiplied by the respective occupation numbers (array froc in the appropriate irrep section).

Note that the auxiliary program densf, which is provided with the ADF package, generates orbital and density values on a user-specified grid. See the utilities document.

TAPE13

TAPE13 is the checkpoint file for restarts after a crash. It is a concise version of TAPE21, containing only the data the program uses for restarting the calculation. See the restart keyword.

Like TAPE21, TAPE13 is a binary, keyword driven KF file. You can manipulate it with the KF utilities, to get a print-out of its 'table of contents', or a complete ASCII dump of its full contents.

The calculation that produces TAPE13 determines which section are written on it. The following sections may occur (and if they occur, the listed variables are stored in them). The actual values of the variables should be identical to the corresponding variables on TAPE21. Also they should have the same names and be located in the same sections. In some cases, TAPE13 contains the complete corresponding section of TAPE21.

Contents of TAPE13

Section Fit

`coef_SCF`

SCF fit coefficients. Total number of them is nprimf, the number of primitive fit functions (counting all *Cartesian* spherical polynomials: 3 for a *p*-set, 6 for a *d*-set, and so on). If the calculation is spin-unrestricted, each spin has its own set of fit coefficients: first all coefficients of spin-A, then those of spin-B

`coef_FreqCenter`

Only in a Frequencies calculation: the fit coefficients that correspond to the equilibrium geometry. The variable coef_SCF corresponds always to the current geometry, or the previous one if the geometry has just been changed and the new SCF has yet to start.

Section Freq

This section is identical to the same section on TAPE21.

Section Geometry

This section is identical to the same section on TAPE21

Section GeoOpt

This section is identical to the same section on TAPE21

Section IRC

This section is identical to the same section on TAPE21

Section IRC_Forward

This section is identical to the same section on TAPE21

Section IRC_Backward

This section is identical to the same section on TAPE21

Section LT

This section is identical to the same section on TAPE21

Section TS

This section is identical to the same section on TAPE21

KF browser

With the GUI module kfbrowser one can browse through the raw data on KF files (like the .t21 result files).

```
$ADFBIN/kfbrowser file.t21
```

KF command line utilities

There are four utility programs for manipulating KF files from the command shell. Two of them convert kf files from binary to ASCII and vice versa. See the *pkf* and *dmpkf* utilities for a description of the ASCII format of a kf file. An ASCII version of a KF file can be useful to inspect its contents in detail.

In the versions of ADF prior to ADF2006, the conversion back and forth between binary and ASCII was necessary when a binary KF file generated on a particular platform was to be used on another platform that is not binary compatible. To do so one had to convert binary file to ASCII, transfer to the other platform and transfer back to binary. Although a bit tedious, it was occasionally the only way to avoid recomputing a TAPE21 result file only because you needed it on another machine.

As of ADF2006 this is no longer necessary. All programs from the package will convert a KF file to the format native to this platform if necessary. In such a case, the original file will be renamed to a file with tilde "~" appended to its name and a message will printed on the standard output.

The KF software (KF= Keyed File) has been developed at the Vrije Universiteit in Amsterdam as a general-purpose package to store data on files and retrieve it again by keyword-driven procedures. For more information about the KF package (usage, implementation) consult the SCM web site (http://www.scm.com) where information about the ADF software is available.

pkf

```
pkf file1 { file2 ... filen }
```

pkf prints a summary of the contents of the kf files file1... filen.

The output should be more or less self-documenting: all variables are listed by name, type (integer, real, character, logical) and size (number of array elements) and they are grouped together in named sections.

To put the results in an ASCII file for later inspection:

```
pkf file > ascii_result
```

Each section on the file contains an index of its variables and their associated values. All data are organized in blocks. Each section may have any number of index blocks and any number of data blocks (this depends simply on the amount of data to be stored in such a block). In addition there is one special section, the SuperIndex, which is an index of all sections on the file.

The output of *pkf* consists of:

- General information about the file (name of the file, internally used unit numbers during processing the file...)
- A summary of the SuperIndex: an index of blocks on the file and the sections they are associated with.
- A summary: total numbers of blocks associated with the different types of blocks.
- For each section a list of its variables with for each variable:
    - Its name.
    - Its length: the amount of space reserved on the file for the variable.
    - Its size ('used'): the amount of data associated with the variable; for reals, integers and logicals:
      the number of such elements; for strings: the number of characters.
    - The (logical) index of the data block it is stored on;
    - Off-set: its position within the data block in which it is stored;
    - Its value (for an array: the value of the first element);

Remark: 'length' and 'size' are usually the same, but not necessarily.

cpkf

```
cpkf file1  file2 {key1 .. keyn}
```

cpkf copies the sections and/or variables key1 .. keyn from file1 to file2.

If a referenced section or variable already exists on file2 it is overwritten, else it is created. Sections and variables which are already present on file2 but which are not referenced in the command are not affected.

If no sections and/or variables are explicitly mentioned at all the copy is carried out for *all* sections and variables on file1.

A side effect of the copy is that any 'holes' that may be present on the first file are not copied to the second file so that they no longer take up any space. The data copied to the second file is rearranged for optimum

storage efficiency. 'Holes' may be due to sections and variables that have existed in the past but that have formally been deleted later; one of the KF functionalities is to delete variables or sections. Such activity does not actually rearrange the KF file, but simply deletes the corresponding entries in index tables.

Alternative form:

```
cpkf file1 file2 -rm section1 ...
```

In this form, all sections will be copied except for the ones specified on the command line, thus effectively removing them from the file.

dmpkf

A utility to extract information from a KF file and make it available in ASCII format.

```
dmpkf file {key1 .. keyn}
```

dmpkf prints the sections and/or variables from the file file indicated by key1 .. keyn on standard output. If no sections and/or variables are explicitly mentioned the complete file is printed.

The format to be used for key1 et cetera is:

```
Sec%Var
```

where Var is the variable, which must exist in section Sec. If no Var is mentioned, the complete section Sec is dumped.

By redirecting the result to another file you get an ASCII version of file:

```
dmpkf file > ascii_result
```

The output contains for each printed variable:

- One line with the name of the section it belongs to;
- One line with the name of the variable itself;
- One line with three integers:
    - The amount of space reserved for the variable on the file (this aspect is relevant for the software that handles kf files and is mentioned here only for completeness);
    - The amount of data associated with the variable: for reals, integers, logicals: the number of such elements; for strings: the number of characters;
    - An integer code for the data type of the variable: 1=integer, 2=real, 3=character, 4=logical;
- The values of the variable (on as many lines as necessary): for scalar variables only one value, for arrays as many values as the array contains.

udmpkf

A utility to put information read from standard input into a KF file.

```
udmpkf file
```

udmpkf reads an ASCII file in the format created by *dmpkf* from standard input and creates the KF file file from this data if it does not already exist, otherwise it adds the sections and/or variables in the input file to the existing kf file. If sections or variables on the input file already exist in the target file, that data on the target file is overwritten. Other data on the target file are not affected.

The combination of *dmpkf* and *udmpkf* makes it easy to modify KF files with a normal text editor:

```
dmpkf TAPE21 > t21_ASCII
```

If necessary, edit and modify t21_ASCII, and then

```
udmpkf < t21_ASCII TAPE21_new
```

Note carefully that *dmpkf* and *udmpkf* do NOT have two arguments, but only one each. The ASCII 'argument' is standard input, respectively output, which you may of course redirect to a file.

5.5 Scripting with ADF

- **adfprep**: prepare an ADF job from a script (or command line).
- **adfreport**: get information (including images) from an ADF result file (for use in your script, or to generate an HTML or tab-separated report).
- **pkf**, **cpkf**, **dmpkf**, **udmpkf**: the KF utilities, which are command-line utilities to process KF files.

ADFprep: generate (multiple) ADF jobs

The module *adfprep* is intended to facilitate scripting: it makes it very easy to construct proper jobs, as saved by ADFinput, from within a script. This module can be used, for example, to run the same type of job on various molecules, or to change input settings such as basis set choice or numerical integration accuracy.

ADFprepare ($ADFBIN/adfprep) generates a job script from a .adf file (the template). The .adf file can be produced by ADFinput, or you can use one of the default templates included. These default templates are identical to those present in ADFinput.

Two examples are included to give you an idea what you can do with adfprep.

In $ADFHOME/examples/adf/BakersetSP you will find how to use adfprep to run a particular job (a single point calculation in this case) for all molecules in the Baker set. The molecules are simply xyz files and contain no ADF specific information. adfreport is used to collect the resulting bonding energies

In $ADFHOME/examples/adf/ConvergenceTestCH4 you will find how to use adfprep to test convergence of the bonding energy with respect to basis set and integration accuracy. adfreport is used to collect the resulting bonding energies

Another use of adfprep can be seen in the run script used to calculate something for a batch of conformers. Please try the conformer GUI tutorial, and study the run script to see how adfprep is used."

The most convenient way to see the options of adfprep is to run the adfprep command without arguments. You will get output very much alike the following description, but probably more up-to-date.

```
% adfprep -h
ADFprepare (adfprep) generates a job script from a .adf file (the template),
with user specified changes to input options / method / system.

Usage: adfprep -t template.adf [-m molecule.(adf|xyz|mol|t21)] [-z charge] [-s
spin]
                               [-i integration] [-b basis] [-c core] [-r
relativity]
```

```
                                [-x xcpotential] [-e xcenergy] [-bondsonly]
                                [-dftbmodel DFTB|SCC-DFTB|DFTB3]
[-dftbparameters dir]
                                [-logfile logfile] [-j jobname] [-a adffile]


Start with a job template, adjust it for this particular job, and write the
resulting job
to standard output. Values specified should match exactly the values as you
would specify
 using ADFinput, also for menu choices.

TEMPLATE
-t: the .adf file (saved by ADFinput) to be used as template, defining the
whole job
    All other options override values from this job

    Instead of a .adf file, you may also specify the name of one of the
standard templates
    as defined in ADFinput: "Single Point", Frequencies, "Geometry
Optimization", etc

    Some shortcuts: SP, GO, FREQ, optionally prefixed by
(ADF|BAND|DFTB|UFF|MOPAC)-
    For example: ADF-FREQ, BAND-SP, DFTB-GO

CHANGES TO TEMPLATE
-m: the molecule to use, element types and coordinates
    This can be taken from anything that ADFinput can import,
    for example .adf, .mol, xyz or .t21 files

    If you specify an .sdf file, you can select which frames to import:
        conformers.sdf#1-10         loop over the first 10 frames
        conformers.sdf#e2.0         loop over all frames with energy below 2.0
           (units as in the file, wrt the lowest energy of all frames in the
file,
           energies from comment lines)
        conformers.sdf#1-10e2.0     loop over the first 10 frames,
                                    and use only those  with energy below 2.0
        conformers.sdf              use the first frame of the sdf file

    If you specify a .t21 file, you can select which frames or range of frames
to import:
        ajob.t21#ircf3               3rd frame in the IRC forward path
        ajob.t21#ircb2               2nd frame in the IRC backward path
        ajob.t21#h7                  7th frame in the history
        ajob.t21#lt8                 8th frame in the LT path
        ajob.t21#ircf3-10            IRCForward frame 3, 4, ... 10
        ajob.t21#ircf                IRCForward all frames, starting at 1
        ajob.t21#ircf0-              IRCForward all frames, starting at 0
                                     (original geometry, before first step)

    If you specify a .cry file, the compound to import may be specified:
        $ADFHOME/atomicdata/Molecules/Crystals/Cubic/CsCl.cry#MgTl

    When looping, all resulting jobs will be joined together, the jobname and
adf files
```

```
    may get the frame sequence number appended after an _ when the
-addmolnumber
    flag is used

-addmolnumber: see -m when looping over molecules
-irc: when using IRC frames in the -m flag, revert the backwards order

-runtype: run type (SinglePoint,GeometryOptimization,Frequencies)
-z: charge (real number)
-s: spin (integer), if not zero this implies an unrestricted calculation
-i: integration (integer)
-i: Becke integration (Basic, Normal, Good, VeryGood or Excellent)
-i: teVelde integration (integer)
-b: basis type (SZ, DZ, DZP, TZ, TCP, TZ2P, QZ4P)
-c: core type (None, Small, Medium, Large)
-r: relativistic level (None, Scalar, Spin-Orbit), using ZORA
-x: XC potential during SCF, one from the options available in ADFinput:
        LDA,
        GGA:BP, GGA:BLYP, GGA:PW91, GGA:mPW, GGA:PBE, GGA:RPBE, GGA:revPBE,
GGA:mPBE,
        GGA:OLYP, GGA:OPBE,
        Model:SAOP, Model:LB94,
        Hartree-Fock,
        Hybrid:B3LYP, Hybrid:B3LYP*, Hybrid:B1LYP, Hybrid:KMLYP, Hybrid:O3LYP,
Hybrid:X3LYP,
        Hybrid:BHandH, Hybrid:BHandHLYP, Hybrid:B1PW91, Hybrid:MPW1PW,
Hybrid:MPW1K,
        Hybrid:PBE0, Hybrid:OPBE0
-e: XC energy after SCF (Default, LDA+GGA_METAGGA, LDA+GGA+METAGGA+HYBRIDS)
-k: replace any key, the single argument will be broken into:
     the key, the replacement value, and END for a block key
     all separated by spaces. To insert a return, add a |
     When the key is not found, it is added just before the ATOMS key
     The -k key may be repeated, and is applied at the end, replacing even
earlier changes

-dftbmodel DFTB|SCC-DFTB|DFTB3: select the DFTB modele
-dftbparameters dir: select the directory with DFTB parameters

OUTPUT
-bondsonly: only the bonds as generated by the GUI will be exported (the
GUIBONDS block)
-logfile: force the specified logfile to be used in the run script
-j: produce a fully runable job (as the .job files from ADFjobs),
    using the specified jobname.
    The job script produces files like jobname.out, jobname.t21 etc. Several
job scripts can simply
    be concatenated, the results will be stored in different files using th
jobname parameter
    the default is a simple run script (the .run file from ADFinput, files are
left as they are)
-a: save a .adf file that matches the run script, except for the -k arguments
    (they are listed in the user input field)
    adffile is the name of the adffile, including the .adf extension (required)

Example: calculate gradients for a molecule in file mymol.xyz
         adfprep -t GO -m mymol.xyz -k "stopafter ggrads"
```

```
Example: calculate gradients for a molecule in file mymol.xyz, using a good
Becke grid
         adfprep -t GO -i Good -m mymol.xyz -k "stopafter ggrads"

Example: calculate DFTB frequencies for a molecule in file mymol.xyz
         adfprep -t DFTB-FREQ -m mymol.xyz
```

ADFreport: generate report

The utility *adfreport* is intended to facilitate scripting: it makes it very easy to get results calculated by ADF, BAND, ReaxKF, DFTB, UFF or MOPAC in your own script. The results are available as an HTML file, a tab-separated file, or on standard output for use in other scripts. In this way, one can quickly compare calculate values across a set of calculations, as well as specified pictures, such as HOMO and LUMO orbitals.

ADFreport ($ADFBIN/adfreport) gets information (including images from molecules or fields) from a result file. For ADF this is the .t21 file (TAPE21). It can also be the .runkf file from BAND, the .rxkf file from ReaxKF or the .rkf file from DFTB, MOPAC or UFF.

The selected information is printed as concisely as possible on standard output. Alternatively, you can write the information to a tab separated file or to an HTML file. If the file does not exist, adfreport will also generate one line with headers to identify the information. Images are generated using the ADF-GUI.

Results on a .t21 file may depend on the ordering of atoms. adfreport will report such results in the input order when you use one of the predefined keys, which is the order that you will normally expect. If you ask for a dump of a particular KF variable you will get the data exactly as it is present in the file, and it is up to you to make sense of the data.

A simple example to get the bonding energy from a file:

```
adfreport job.t21 BondingEnergy
```

or to generate high-quality pictures of some orbitals:

```
adfreport job.t21 HOMO LUMO+1 -v "-grid Fine" -v "-antialias" -v "-bgcolor
#ffffff"
```

The most convenient way to see the options of adfreport is to run the adfreport command without arguments. This will print a description of all available options.

For up to date documenatation, please run adfreport without arguments:

```
% adfreport -h
ADFreport (adfreport) converts selected data from ADF and SCM result files
(.t21, .rkf,
.rxkf or .runkf) to either an HTML table or plain text. The plain text version
by default
includes tabs, headers and units for easy use in for example Excel.

To get a minimalist output (as it was before this release), add the -plain
flag or
set the SCM_ADFREPORT_PLAIN environment variable to yes. As the standard
output
already is properly formatted with tabs, the tsv option has been removed.
```

```
Results on a .t21 file may depend on the ordering of atoms.
adfreport will report such results in the input order when you use one of the
predefined keys, which is the order that you will normally expect.
If you ask for a dump of a particular KF variable you will get the data exactly
as it is present in the file, and it is up to you to make sense of the data.

ADF may also reorient your molecule, as required by the symmetry
(see the ADF User's Guide). All information on a .t21 file is with respect to
the
internal orientation as used by ADF.
The information reported by adfreport will also be in the INTERNAL orientation!

Usage: adfreport [-h] [-i] result.t21 [-o result.html] [[-r] result] [-v
adfview]

Successive calls with the same output file will update the information
contained in that file.

The following flags may be used:

-h: help
    Print this help text
    If also an input file is present, the command line flags for the
information on that
    file will be listed (like a table of contents)
-i: result.t21 is the name of the input .t21 (or .rkf etc) file with the ADF
results
    the -i is optional, if no -i flag found the first argument
    without flag will be used

-o: result.html
        is the name of the html formatted file (a simple html table) in which
the
        results will be saved
  : absent:
        the output will be written to standard output,
        with headers and units, nicely formatted with tabs

-plain: only the data, no labels or units
        alternatively, set SCM_ADFREPORT_PLAIN to yes in your environment.
-noplain: the default output with labels and units, tab separated.
          Only needed to undo the effect of SCM_ADFREPORT_PLAIN in you
environment.

-r: result is a particular result to be reported from the .t21 file.
    The -r flag is optional: all arguments without flags will be considered
    as result arguments, with exception of the first argument if the -i flag
    is also absent

    *** KF variables ***

    Any proper KF variable is allowed (section%variable, see the KF utilities
    documentation), After the variable you can specify details, all separated
with #

    - range: one or two numbers, separated with a :, array counts start at 1
    - format: TclTk format string for one number, like 8.3f or 12.6g
```

```
    - %nperline: (% with number) insert new line after nperline items

    For example: -r "Geometry%xyz#12.4f##3"      prints a nicely formatted
table
                                                   of the coordinates
                 -r "Geometry%xyz#1:9#12.4f##3"  similar, for the first two
atoms only
                 -r "Geometry%xyz#12.4f#1:9"     the coordinates of the first
two atoms,
                                                   all on one line
                 -r "Geometry%xyz#1"             print just the first
coordinate
                 -r "Energys%Bond Energy"        print the bond energy (a
scalar)

    *** Predefined Keys ***

    When a result without % is specified, it must be one of the pre-defined
keys
    If you use HTML output, the name of the keys are used as table headers.
    The keys are not cases-sensitive, and * matches any text.
    Note: LUMO is interpreted as HOMO+1, so it might not be the real LUMO

    *** For .t21 files (ADF result files): ***

       orient*  : affine transform (3x4) from input to internal ADF
orientation,
                  format after #
       iorient* : affine transform (3x4) from internal ADF to input
orientation,
                  format after #

       title        : title of the calculation
       type         : calculation type (single point, geometry optimization,
...)
       weight       : molecular weight
       symmetry     : molecular symmetry
       natoms       : number of atoms
       integration     : integration accuracy
       integration-min : minimum integration accuracy
       integration-max : maximum integration accuracy

       scfstatus    : status string (CONVERGED if no problems)
       charges      : shorthand for Voronoi, Hirshfeld and Mulliken charges
       voronoi      : voronoi deformation charges
       hirshfeld    : hirshfeld fragment charges, will only work with atomic
fragments
       mdc          : all available MDC atom charges
       mdc-m        : MDC-M charges
       mdc-d        : MDC-D charges
       mdc-q        : MDC-Q charges
       mulliken     : Mulliken charges
       bondorders   : Mayer bond orders
       nmr          : nmr shieldings
       nmr-shieldings: nmr shieldings
       nmr-shielding-tensor : nmr shielding tensor
       nmr-j-coupling-tensor : nmr j coupling tensor
```

```
       nmr-k-coupling-tensor : nmr k coupling tensor
       nmr-j-coupling-constant : nmr j coupling constant
       nmr-k-coupling-constant : nmr k coupling constant
       dipolev*     : dipole vector
       dipole       : dipole moment (length of dipole vector)
       quadrupole   : quadrupole tensor
       orbital-info : orbital info (energy, occupation and label),
                      format for energy after #,
                      range after # with HOMO or LUMO
                      for example:
                          orbital-info#HOMO, orbital-info#HOMO-1,
                          orbital-info#HOMO-2:LUMO+2, orbital-info#HOMO#12.8f
       orbital-e*   : orbital energies, format and range after # as in
orbital-info
       orbital-o*   : orbital occupations, format and range after # as in
orbital-info
       orbital-l*   : orbital labels, format and range after # as in
orbital-info
       homo-lumo-gap*: HOMO-LUMO gap, format after #
       atomlabels          : name of atoms with sequence number, starting at 0
       atomlabels-from0    : name of atoms with sequence number, starting at 0
       atomlabels-from1    : name of atoms with sequence number, starting at 1
       nstep        : number of steps in history / LT / IRC data,
                      type (h,lt,ircf,ircb) after #
       step         : use coordinates from history / LT / IRC data, step
number after #
                      with h for history, lt for LT, ircf/ircb for forward/
backward IRC
                      if no letter after #, history data will be used
                      (if not, last step will be used)
                      for example: step#23 (or step#h23), step#lt4, step#ircf3
       geometry,
       geometry-a*,
       geometry-b*  : geometry (element type and coordinates), in input order,
                      in angstrom or bohr (default bohr)
       distance     : distance between two atoms, in angstrom. Input separated
by #:
                      labels (optional): include atom labels in output
                      format (optional): format field
                      atom numbers, starting at 1, in input order
                      for example: distance#2#3, distance#labels#2#3,
distance#-8.3f#5#8,
                      distance#labels#8.4f#1#2, distance#2#3#4#5,
distance#labels#1#2#3#4
       angle        : angle between three atoms, in degrees.
                      Input see distance, but with three atoms per angle
       dihedral     : dihedral between four atoms, in degrees.
                      Input see distance, but with our atoms per dihedral
       hessian*     : hessian (from GeoOpt%Hessian_CART), fmt and nperline
options after #
       gradients*   : gradients (from GeoOpt%Gradients), fmt and nperline
options after #
       energies*    : all avaliable energies (bonding up to xc, with labels),
                      fmt option after #
       bonding      : total bonding energy
       pauli        : total pauli repulsion
       steric       : total steric interaction
```

```
       orbital      : total orbital interaction
       electrostatic   : electrostatic energy
       kinetic         : electrostatic energy
       coulomb         : elstat (steric +OrbInt) energy
       xc              : XC energy
       dispersion      : dispersion energy
       frequencies*    : IR Frequencies, format, nperline and range
                         (n, or n:n, start at 1) after #
       freqint*        : IR Intensities, format, nperline and range
                         (n, or n:n, start at 1) after #
       freqlabel*      : IR Frequencies label (symmetry), format, nperline and
range
                         (n, or n:n, start at 1) after #
       normalmode*     : Normal Modes (mass weighted), format, nperline and
range
                         (n, or n:n, start at 1) after #
       zeropoint*      : Zero-Point energy
       excitation*        : Excitation energies, format, nperline and range
                            (n, or n:n, start at 1) after #
       oscillatorstrength*: Oscillator strengths for the excitation energies
                            format, nperline and range (n, or n:n, start at 1)
after #
       excitlabel*        : Excitation labels (symmetry), format, nperline and
range
                            (n, or n:n, start at 1) after #

    *** For .rkf files (general SCM result file): ***

       natoms       : number of atoms
       step         : use coordinates from history, step number after #
                      (if not, last step will be used)
                      for example: step#23
       geometry,
       geometry-a*,
       geometry-b*  : geometry (element type and coordinates), in input order,
                      in angstrom or bohr (default bohr)
       distance     : distance between two atoms, in angstrom. Input separated
by #:
                      labels (optional): include atom labels in output
                      format (optional): format field
                      atom numbers, starting at 1, in input order
                      for example: distance#2#3, distance#labels#2#3,
distance#-8.3f#5#8,
                      distance#labels#8.4f#1#2, distance#2#3#4#5,
distance#labels#1#2#3#4
       angle        : angle between three atoms, in degrees.
                      Input see distance, but with three atoms per angle
       dihedral     : dihedral between four atoms, in degrees.
                      Input see distance, but with our atoms per dihedral
       hessian*     : hessian, fmt and nperline options after #
       gradients*   : gradients, fmt and nperline options after #
       energies     : all avaliable energies (with labels)

    *** For .rxkf files (SCM ReaxKF result file): ***

       natoms       : number of atoms
       geometry,
```

```
       geometry-a*,
       geometry-b*  : geometry (element type and coordinates), in input order,
                      in angstrom or bohr (default bohr)
       distance     : distance between two atoms, in angstrom. Input separated
by #:
                      labels (optional): include atom labels in output
                      format (optional): format field
                      atom numbers, starting at 1, in input order
                      for example: distance#2#3, distance#labels#2#3,
distance#-8.3f#5#8,
                      distance#labels#8.4f#1#2, distance#2#3#4#5,
distance#labels#1#2#3#4
       angle        : angle between three atoms, in degrees.
                      Input see distance, but with three atoms per angle
       dihedral     : dihedral between four atoms, in degrees.
                      Input see distance, but with our atoms per dihedral
       atomlabel          : name of atoms with sequence number, starting at 0
       atomlabel-from0    : name of atoms with sequence number, starting at 0
       atomlabel-from1    : name of atoms with sequence number, starting at 1
       rx-frame n options : information for a particular reaxff frame.
                            Note the spaces, you will need to quote this key.
                            n: frame number 0, 1, 2, ...
                               (thus this is not the ReaxFF step number)
                            options: a combination of the following
                                     (if omitted, all will be reported):
                                nframes  : total number of frames
                                step     : the ReaxFF step number for the
specified frame
                                nats     : number of atoms
                                xyz      : the xyz coordinates
                                names    : the element names (C, H etc) for
each atom
                                           in the same order as the XYZ
                                neighbors: bond information
                                cell     : cell information
                           for example: adfreport water.rxkf "rx-frame 20
step xyz cell"
       pdbtrajectory: the trajectory information (including molecule details)
                      as a sequence of PDB models
                      due to limitations of the PDB format no more then 99999
atoms,
                      and it will not be a standard conforming PDB file
       xmol         : the trajectory information (only element, xyz) in xmol
format
       gro          : trajectory as .gro file (xyz and velocities)
                      options after a - sign:
                      m : print list of molecule names and formulas only
                      x : allow xyz only frames (missing velocities)
                      f : add forces if available
                      tf : add the time step, f is a floating point number
                           that is the time per step in ps
                      examples: gro-x, gro-f, gro-xf,  gro-ft0.0001,
gro-xt0.001, etc

    *** For .runkf files (SCM BAND result file): ***

       natoms       : number of atoms
```

```
       geometry,
       geometry-a*,
       geometry-b*  : geometry (element type and coordinates), in input order,
                      in angstrom or bohr (default bohr)
       distance     : distance between two atoms, in angstrom. Input separated
by #:
                      labels (optional): include atom labels in output
                      format (optional): format field
                      atom numbers, starting at 1, in input order
                      for example: distance#2#3, distance#labels#2#3,
distance#-8.3f#5#8,
                      distance#labels#8.4f#1#2, distance#2#3#4#5,
distance#labels#1#2#3#4
       angle        : angle between three atoms, in degrees.
                      Input see distance, but with three atoms per angle
       dihedral     : dihedral between four atoms, in degrees.
                      Input see distance, but with our atoms per dihedral
       atomlabel          : name of atoms with sequence number, starting at 0
       atomlabel-from0    : name of atoms with sequence number, starting at 0
       atomlabel-from1    : name of atoms with sequence number, starting at 1


    The -r flag (or arguments without flag) may be repeated for multiple
results

-v: command line to pass to adfview (without filenames) to generate an image
    The image will be generated by ADFview, the image will be stored in a
    directory with a name based on the result file, and with extension .jpgs.
    The result file will contain a path to the image file (directly, or in an
IMG tag)
    After the -v the arguments must be listed, with proper quoting.
    Repeat the -v flag for multiple arguments.
    The individual -scmgeometry, -bgcolor, -zoom, -viewplane, -antialias and
-grid options
    will be collected and applied to all view options

    Some shortcuts have been defined (HOMO, HOMO+1, LUMO, Molecule, Density,
Potential)

    Some useful flags include -scmgeometry (default 200x200), -bgcolor
(default #220000),
                               -zoom (default 1.0), -viewplane (default {1 2
5}),
                               -antialias (off when not present, especially
useful
                                    with light bgcolors),
                               -grid (Coarse when not present, Medium when
specified,
                                    or value after flag if a value is present)

    Examples: Molecule HOMO LUMO
              HOMO-1 LUMO+1 -v "-viewplane {0 0 1}" -v "-grid Fine" -v
"-antialias"
```

6 References

1. E.J. Baerends, V. Branchadell and M. Sodupe, *Atomic reference energies for density functional calculations*, Chemical Physics Letters **265**,481 (1997)

2. F.M. Bickelhaupt and E.J. Baerends, *Kohn-Sham DFT: Predicting and Understanding Chemistry*, in *Reviews in Computational Chemistry*, D.B. Boyd and K.B. Lipkowitz, Editors. 2000, Wiley-VCH: New York. p. 1-86.

3. T. Ziegler and A. Rauk, *On the calculation of Bonding Energies by the Hartree Fock Slater method. I. The Transition State Method*, Theoretica Chimica Acta **46**, 1 (1977)

4. P.J. van den Hoek, A.W. Kleyn and E.J. Baerends, *What is the origin of the repulsive wall in atom-atom potentials*, Comments Atomic and Molecular Physics **23**, 93 (1989)

5. E.J. Baerends, *Pauli repulsion effects in scattering from and catalysis by surface*, in *Cluster models for surface and bulk phenomena*, ISBN13: 9780306441028, G. Puccchiori, P.S. Bagus and F. Parmigiani, Editors. 1992, Springer: New-York. p. 189-207.

6. L. Versluis and T. Ziegler, *The determination of Molecular Structure by Density Functional Theory*, Journal of Chemical Physics **88**, 322 (1988)

7. L. Versluis, *The determination of molecular structures by the HFS method*, 1989, University of Calgary

8. L. Fan and T. Ziegler, *Optimization of molecular structures by self consistent and non-local density functional theory*, Journal of Chemical Physics **95**, 7401 (1991)

9. L. Deng, T. Ziegler and L. Fan, *A combined density functional and intrinsic reaction coordinate study on the ground state energy surface of H_2CO*, Journal of Chemical Physics **99**, 3823 (1993)

10. L. Deng and T. Ziegler, *The determination of Intrinsic Reaction Coordinates by density functional theory*, International Journal of Quantum Chemistry **52**, 731 (1994)

11. T.H. Fischer and J. Almlöf, *General Methods for Geometry and Wave Function Optimization*, Journal of Physical Chemistry **96**, 9768 (1992)

12. L. Fan and T. Ziegler, *Application of density functional theory to infrared absorption intensity calculations on main group molecules*, Journal of Chemical Physics **96**, 9005 (1992)

13. L. Fan and T. Ziegler, *Nonlocal density functional theory as a practical tool in calculations on transition states and activation energies*, Journal of the American Chemical Society **114**, 10890 (1992)

14. A. Bérces, *Application of density functional theory to the vibrational characterization of transition metal complexes*, 1995, University of Calgary: Calgary

15. R. van Leeuwen and E.J. Baerends, *Exchange-correlation potential with correct asymptotic behavior*, Physical Review A **49**, 2421 (1994)

16. M. Grüning, O.V. Gritsenko, S.J.A. van Gisbergen and E.J. Baerends, *Shape corrections to exchange-correlation Kohn-Sham potentials by gradient-regulated seamless connection of model potentials for inner and outer region*, Journal of Chemical Physics **114**, 652 (2001)

17. P.R.T. Schipper, O.V. Gritsenko, S.J.A. van Gisbergen and E.J. Baerends, *Molecular calculations of excitation energies and (hyper)polarizabilities with a statistical average of orbital model exchange-correlation potentials*, Journal of Chemical Physics **112**, 1344 (2000)

18. M. Grüning, O.V. Gritsenko, S.J.A. van Gisbergen and E.J. Baerends, *On the required shape correction to the LDA and GGA Kohn Sham potentials for molecular response calculations of (hyper)polarizabilities and excitation energies*, Journal of Chemical Physics **116**, 9591 (2002)

19. D.P. Chong, O.V. Gritsenko and E.J. Baerends, *Interpretation of the Kohn-Sham orbital energies as approximate vertical ionization potentials*, Journal of Chemical Physics **116**, 1760 (2002)

20. S.H. Vosko, L. Wilk and M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, Canadian Journal of Physics **58** (8), 1200 (1980)

21. H. Stoll, C.M.E. Pavlidou, and H. Preuss, *On the calculation of correlation energies in the spin-density functional formalism*, Theoretica Chimica Acta **49**, 143 (1978)

22. A.D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Physical Review A **38**, 3098 (1988)

23. J.P. Perdew and Y. Wang, *Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation*, Physical Review B **33**, 8822 (1986)

24. J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Sing and C. Fiolhais, *Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation*, Physical Review B **46**, 6671 (1992)

25. C. Adamo and V. Barone, *Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models*, Journal of Chemical Physics **108**, 664 (1998)

26. J. P. Perdew, K. Burke and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Physical Review Letters **77**, 3865 (1996)

27. B. Hammer, L.B. Hansen, and J.K. Norskøv, *Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals*, Physical Review B **59**, 7413 (1999)

28. Y. Zhang and W. Yang, *Comment on "Generalized Gradient Approximation Made Simple"*, Physical Review Letters **80**, 890 (1998)

29. N.C. Handy and A.J. Cohen, *Left-right correlation energy*, Molecular Physics **99**, 403 (2001)

30. J.P. Perdew, *Density-functional approximation for the correlation energy of the inhomogeneous electron gas*, Physical Revied B **33**, 8822 (1986)
Erratum: J.P. Perdew, Physical Review B **34**, 7406 (1986)

31. C. Lee, W. Yang and R.G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Physical Review B **37**, 785 (1988)

32. B.G. Johnson, P.M.W. Gill and J.A. Pople, *The performance of a family of density functional methods*, Journal of Chemical Physics **98**, 5612 (1993)

33. T.V. Russo, R.L. Martin and P.J. Hay, *Density Functional calculations on first-row transition metals*, Journal of Chemical Physics **101**, 7729 (1994)

34. R. Neumann, R.H. Nobes and N.C. Handy, *Exchange functionals and potentials*, Molecular Physics **87**, 1 (1996)

35. J.P. Perdew, K. Burke and M. Ernzerhof, *ERRATA for "Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]"* Physical Review Letters **78**, 1396 (1997)

36. F.A. Hamprecht, A.J. Cohen, D.J. Tozer and N.C. Handy, *Development and assessment of new exchange-correlation functionals*, Journal of Chemical Physics **109**, 6264 (1988)

37. A.D. Boese, N.L. Doltsinis, N.C. Handy and M. Sprik, *New generalized gradient approximation functionals*, Journal of Chemical Physics **112**, 1670 (2000)

38. A.D. Boese and N.C. Handy, *A new parametrization of exchange-correlation generalized gradient approximation functionals*, Journal of Chemical Physics **114**, 5497 (2001)

39. T. Tsuneda, T. Suzumura and K. Hirao, *A new one-parameter progressive Colle.Salvetti-type correlation functional*, Journal of Chemical Physics **110**, 10664 (1999)

40. J.B. Krieger, J. Chen, G.J. Iafrate and A. Savin, in *Electron Correlations and Materials Properties*, ISBN13: 9780306462825, A. Gonis and N. Kioussis, Editors. 1999, Plenum: New York.

41. J.P. Perdew, S. Kurth, A. Zupan and P. Blaha, *Erratum: Accurate Density Functional with Correct Formal Properties: A Step Beyond the Generalized Gradient Approximation [Phys. Rev. Lett. 82, 2544 (1999)]*, Physical Review Letters **82**, 5179 (1999).

42. T. van Voorhis and G.E. Scuseria, *A novel form for the exchange-correlation energy functional*, Journal of Chemical Physics **109**, 400 (1998)

43. M. Filatov and W. Thiel, *A new gradient-corrected exchange-correlation density functional*, Molecular Physics **91**, 847 (1997)

44. M. Filatov and W. Thiel, *Exchange-correlation density functional beyond the gradient approximation*, Physical Review A **57**, 189 (1998)

45. E.I. Proynov, S. Sirois and D.R. Salahub, *Extension of the LAP functional to include parallel spin correlation*, International Journal of Quantum Chemistry **64**, 427 (1997)

46. E.I. Proynov, H. Chermette and D.R. Salahub, *New tau-dependent correlation functional combined with a modified Becke exchange*, Journal of Chemical Physics **113**, 10013 (2000)

47. S. Patchkovskii, J. Autschbach and T. Ziegler, *Curing difficult cases in magnetic properties prediction with self-interaction corrected density functional theory*, Journal of Chemical Physics **115**, 26 (2001)

48. S. Patchkovskii and T. Ziegler, *Improving "difficult" reaction barriers with self-interaction corrected density functional theory*, Journal of Chemical Physics **116**7806 (2002)

49. S. Patchkovskii and T. Ziegler, *Phosphorus NMR chemical shifts with self-interaction free, gradient-corrected DFT*, Journal of Physical Chemistry A **106**, 1088 (2002)

50. P.H.T. Philipsen, E. van Lenthe, J.G. Snijders and E.J. Baerends, *Relativistic calculations on the adsorption of CO on the (111) surfaces of Ni, Pd, and Pt within the zeroth-order regular approximation*, Physical Review B **56**, 13556 (1997)

51. J.G. Snijders and E.J. Baerends, *A perturbation theory approach to relativistic calculations. I. Atoms*, Molecular Physics **36**, 1789 (1978)

52. J.G. Snijders, E.J. Baerends and P. Ros, *A perturbation theory approach to relativistic calculations. II. Molecules*, Molecular Physics **38**, 1909 (1979)

53. T. Ziegler, J.G. Snijders and E.J. Baerends, *Relativistic effects on bonding*, Journal of Chemical Physics **74**, 1271 (1981)

54. R.L. DeKock, E.J. Baerends, P.M. Boerrigter and J.G. Snijders, *On the nature of the first excited states of the uranyl ion*, Chemical Physics Letters **105**, 308 (1984)

55. R.L. DeKock, E.J. Baerends, P.M. Boerrigter and R. Hengelmolen, *Electronic structure and bonding of $Hg(CH_3)_2$, $Hg(CN)_2$, $Hg(CH_3)(CN)$, $Hg(CCCH_3)_2$, and $Au(PMe)_3)(CH_3)$*, Journal of the American Chemical Society **106**, 3387 (1984)

56. P.M. Boerrigter, *Spectroscopy and bonding of heavy element compounds*, 1987, Vrije Universiteit.

57. P.M. Boerrigter, M.A. Buijse and J.G. Snijders, *Spin-Orbit interaction in the excited states of the dihalogen ions F_2^+, Cl_2^+ and Br_2^+*, Chemical Physics **111**, 47 (1987)

58. P.M. Boerrigter, E.J. Baerends and J.G. Snijders, *A relativistic LCAO Hartree-Fock-Slater investigation of the electronic structure of the actinocenes $M(COT)_2$, M=Th, Pa, U, Np and Pu*, Chemical Physics **122**, 357 (1988)

59. T. Ziegler, V. Tschinke, E.J. Baerends, J.G. Snijders and W. Ravenek, *Calculation of bond energies in compounds of heavy elements by a quasi-relativistic approach*, Journal of Physical Chemistry **93**, 3050 (1989)

60. J. Li, G. Schreckenbach and T. Ziegler, *A Reassessment of the First Metal-Carbonyl Dissociation Energy in $M(CO)_4$ (M = Ni, Pd, Pt), $M(CO)_5$ (M = Fe, Ru, Os), and $M(CO)_6$ (M = Cr, Mo, W) by a Quasirelativistic Density Functional Method*, Journal of the American Chemical Society **117**, 486 (1995)

61. E. van Lenthe, A.E. Ehlers and E.J. Baerends, *Geometry optimization in the Zero Order Regular Approximation for relativistic effects*, Journal of Chemical Physics **110**, 8943 (1999)

62. E. van Lenthe, E.J. Baerends and J.G. Snijders, *Relativistic regular two-component Hamiltonians*, Journal of Chemical Physics **99**, 4597 (1993)

63. E. van Lenthe, E.J. Baerends and J.G. Snijders, *Relativistic total energy using regular approximations*, Journal of Chemical Physics **101**, 9783 (1994)

64. E. van Lenthe, J.G. Snijders and E.J. Baerends, *The zero-order regular approximation for relativistic effects: The effect of spin-orbit coupling in closed shell molecules*, Journal of Chemical Physics **105**, 6505 (1996)

65. E. van Lenthe, R. van Leeuwen, E.J. Baerends and J.G. Snijders, *Relativistic regular two-component Hamiltonians*, International Journal of Quantum Chemistry **57**, 281 (1996)

66. C.C. Pye and T. Ziegler, *An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package*, Theoretical Chemistry Accounts **101**, 396 (1999)

67. A. Klamt and G. Schüürmann, *COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient*, Journal of the Chemical Society: Perkin Transactions **2**, 799 (1993)

68. A. Klamt, *Conductor-like Screening Model for real solvents: A new approach to the quantitative calculation of solvation phenomena*, Journal of Physical Chemistry **99**, 2224 (1995)

69. A. Klamt and V. Jones, *Treatment of the outlying charge in continuum solvation models*, Journal of Chemical Physics **105**, 9972 (1996)

70. J.L. Pascual-ahuir, E. Silla and I. Tuñon, *GEPOL: An improved description of molecular surfaces. III. A new algorithm for the computation of a solvent-excluding surface*, Journal of Computational Chemistry **15**, 1127 (1994)

71. S.J.A. van Gisbergen, J.G. Snijders and E.J. Baerends, *Implementation of time-dependent density functional response equations*, Computer Physics Communications **118**, 119 (1999)

72. S.J.A. van Gisbergen, *Molecular Response Property Calculations using Time Dependent Density Functional Theory*, in *Chemistry*. 1998, Vrije Universiteit: Amsterdam. p. 190.

73. E.K.U. Gross, J.F. Dobson and Petersilka, in *Density Functional Theory*, R.F. Nalewajski, Editor. 1996, Springer: Heidelberg.

74. S.J.A. van Gisbergen, V.P. Osinga, O.V. Gritsenko, R. van Leeuwen, J.G. Snijders and E.J. Baerends, *Improved density functional theory results for frequency-dependent polarizabilities, by the use of an exchange-correlation potential with correct asymptotic behavior*, Journal of Chemical Physics **105**, 3142 (1996)

75. S.J.A. van Gisbergen, J.G. Snijders, and E.J. Baerends, *Time-dependent Density Functional Results for the Dynamic Hyperpolarizability of C_{60}*, Physical Review Letters **78**, 3097 (1997)

76. S.J.A. van Gisbergen, J.G. Snijders and E.J. Baerends, *Calculating frequency-dependent hyperpolarizabilities using time-dependent density functional theory*, Journal of Chemical Physics **109**, 10644 (1998)

77. S.J.A. van Gisbergen, J.G. Snijders and E.J. Baerends, *Accurate density functional calculations on frequency-dependent hyperpolarizabilities of small molecules*, Journal of Chemical Physics **109**, 10657 (1998)

78. M.E. Casida, C. Jamorski, K.C. Casida and D.R. Salahub, *Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold*, Journal of Chemical Physics **108**, 4439 (1998)

79. V.P. Osinga, S.J.A. van Gisbergen, J.G. Snijders and E.J. Baerends, *Density functional results for isotropic and anisotropic multipole polarizabilities and C_6, C_7, and C_8 Van der Waals dispersion coefficients for molecules*, Journal of Chemical Physics **106**, 5091 (1997)

80. J. Autschbach and T. Ziegler, *Calculating molecular electric and magnetic properties from time-dependent density functional response theory*, Journal of Chemical Physics **116**, 891 (2002)

81. J. Autschbach, T. Ziegler, S.J.A. van Gisbergen and E.J. Baerends, *Chiroptical properties from time-dependent density functional theory. I. Circular dichroism spectra of organic molecules*, Journal of Chemical Physics **116**, 6930 (2002)

82. S.J.A. van Gisbergen, F. Kootstra, P.R.T. Schipper, O.V. Gritsenko, J.G. Snijders and E.J. Baerends, *Density-functional-theory response-property calculations with accurate exchange-correlation potentials*, Physical Review A **57**, 2556 (1998)

83. S.J.A. van Gisbergen, A. Rosa, G. Ricciardi and E.J. Baerends, *Time-dependent density functional calculations on the electronic absorption spectrum of free base porphin*, Journal of Chemical Physics **111**, 2499 (1999)

84. A. Rosa, G. Ricciardi, E.J. Baerends and S.J.A. van Gisbergen, *The Optical Spectra of NiP, Nipz, NiTBP, and NiPc. Electronic effects of meso-tetraaza substitution and tetrabenzoannulation*, Journal of Physical Chemistry A **105**, 3311 (2001)

85. G. Ricciardi, A. Rosa and E.J. Baerends, *Ground and Excited States of Zinc Phthalocyanine studied by Density Functional Methods*, Journal of Physical Chemistry A **105**, 5242 (2001)

86. S.J.A. van Gisbergen, J.A. Groeneveld, A. Rosa, J.G. Snijders and E.J.Baerends, *Excitation energies for transition metal compounds from time-dependent density functional theory. Applications to $MnO4^-$, $Ni(CO)_4$ and $Mn^2(CO)_{10}$*, Journal of Physical Chemistry A **103**, 6835 (1999)

87. A. Rosa, E.J. Baerends, S.J.A. van Gisbergen, E. van Lenthe, J.A. Groeneveld and J. G. Snijders, *Article Electronic Spectra of $M(CO)_6$ (M = Cr, Mo, W) Revisited by a Relativistic TDDFT Approach*, Journal of the American Chemical Society **121**, 10356 (1999)

88. S.J.A. van Gisbergen, C. Fonseca Guerra and E.J. Baerends, *Towards excitation energies and (hyper)polarizability calculations of large molecules. Application of parallelization and linear scaling*

techniques to time-dependent density functional response theory, Journal of Computational Chemistry **21**, 1511 (2000)

89. S.J.A. van Gisbergen, J.G. Snijders and E.J. Baerends, *A Density Functional Theory study of frequency-dependent polarizabilities and van der Waals dispersion coefficients for polyatomic molecules*, Journal of Chemical Physics **103**, 9347 (1995)

90. B. Champagne, E.A. Perpète, S.J.A. van Gisbergen, E.J. Baerends, J.G. Snijders, C. Soubra-Ghaoui, K.A. Robins and B.Kirtman, *Assessment of conventional density functional schemes for computing the polarizabilities and hyperpolarizabilities of conjugated oligomers: An ab initio investigation of polyacetylene chains*, Journal of Chemical Physics **109**, 10489 (1998)
Erratum: Journal of Chemical Physics **111**, 6652 (1999)

91. D.M. Bishop, *Aspects of Non-Linear-Optical Calculations*, Advances in Quantum Chemistry **25**, 3 (1994)

92. A. Willets, J.E. Rice, D.M. Burland and D.P. Shelton, *Problems in comparison of experimental and theoretical hyperpolarizabilities*, Journal of Chemical Physics **97**, 7590 (1992)

93. D.P. Shelton and J.E. Rice, *Measurements and calculations of the hyperpolarizabilities of atoms and small molecules in the gas phase*, Chemical Reviews **94**, 3 (1994)

94. J. Autschbach, S. Patchkovskii, T. Ziegler, S.J.A. van Gisbergen and E.J. Baerends, *Chiroptical properties from time-dependent density functional theory. II. Optical rotations of small to medium sized organic molecules*, Journal of Chemical Physics **117**, 581 (2002)

95. E. van Lenthe, A. van der Avoird and P.E.S. Wormer, *Density functional calculations of molecular g-tensors in the zero order regular approximation for relativistic effects*, Journal of Chemical Physics **107**, 2488 (1997)

96. E. van Lenthe, A. van der Avoird and P.E.S. Wormer, *Density functional calculations of molecular hyperfine interactions in the zero order regular approximation for relativistic effects*, Journal of Chemical Physics **108**, 4783 (1998)

97. E. van Lenthe and E.J. Baerends, *Density functional calculations of nuclear quadrupole coupling constants in the zero-order regular approximation for relativistic effects*, Journal of Chemical Physics **112**, 8279 (2000)

98. R.E. Bulo, A.W. Ehlers, S. Grimme and K. Lammertsma, *Vinylphosphirane.Phospholene Rearrangements: Pericyclic [1,3]-Sigmatropic Shifts or Not?* Journal of the American Chemical Society **124**, 13903 (2002)

99. R. Pauncz, *Spin Eigenfunctions*, ISBN13: 9780306401411, 1979, New York: Plenum Press

100. A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry*, ISBN13: 9780070627390, 1st ed. revised ed. 1989: McGraw-Hill

101. H. Eschrig and V.D.P. Servedio, *Relativistic density functional approach to open shells*, Journal of Computational Chemistry **20**, 23 (1999)

102. C. van Wüllen, *Spin densities in two-component relativistic density functional calculations: Noncollinear versus collinear approach*, Journal of Computational Chemistry **23**, 779 (2002)

103. S.G. Wang and W.H.E. Schwarz, *Simulation of nondynamical correlation in density functional calculations by the optimized fractional orbital occupation approach: Application to the potential energy surfaces of O_3 and SO_2*, Journal of Chemical Physics **105**, 4641 (1996)

104. P.M. Boerrigter, G. te Velde and E.J. Baerends, *Three-dimensional Numerical Integration for Electronic Structure Calculations*, International Journal of Quantum Chemistry **33**, 87 (1988)

105. G. te Velde and E.J. Baerends, *Numerical Integration for Polyatomic Systems*, Journal of Computational Physics **99**, 84 (1992)

106. A. Bérces and T. Ziegler, *The harmonic force field of benzene calculated by local density functional theory*, Chemical Physics Letters **203**, 592 (1993)

107. A. Bérces and T. Ziegler, *The harmonic force field of benzene. A local density functional study*, Journal of Chemical Physics **98**, 4793 (1993)

108. F. Neese and E. I. Solomon, *MCD C-Term Signs, Saturation Behavior, and Determination of Band Polarizations in Randomly Oriented Systems with Spin S ≥ 1/2. Applications to S = 1/2 and S = 5/2*, Inorganic Chemistry **38**, 1847 (1999)

109. G. te Velde, *Numerical integration and other methodological aspects of bandstructure calculations*, in *Chemistry*. 1990, Vrije Universiteit: Amsterdam.

110. T. Ziegler and A. Rauk, *A theoretical study of the ethylene-metal bond in complexes between copper(1+), silver(1+), gold(1+), platinum(0) or platinum(2+) and ethylene, based on the Hartree-Fock-Slater transition-state method*, Inorganic Chemistry **18**, 1558 (1979)

111. L. Noodleman, and E.J. Baerends, *Electronic Structure, Magnetic Properties, ESR, and Optical Spectra for 2-Fe Ferredoxin Models by LCAO-Xa Valence Bond Theory*, Journal of the American Chemical Society **106**, 2316 (1984)

112. F.M. Bickelhaupt, N.M. Nibbering, E.M. van Wezenbeek and E.J. Baerends, *The Central Bond in the Three CN* Dimers NC_CN, CN-CN, and CN-NC: Electron Pair Bonding and Pauli Repulsion Effects*, Journal of Physical Chemistry **96**, 4864 (1992)

113. G. Schreckenbach and T. Ziegler, *The calculation of NMR shielding tensors using GIAO's and modern density functional theory*, Journal of Physical Chemistry **99**, 606 (1995)

114. G. Schreckenbach and T. Ziegler, *The calculation of NMR shielding tensors based on density functional theory and the frozen-core approximation*, International Journal of Quantum Chemistry **60**, 753 (1996)

115. G. Schreckenbach and T. Ziegler, *Calculation of NMR shielding tensors based on density functional theory and a scalar relativistic Pauli-type Hamiltonian. The application to transition metal complexes*, International Journal of Quantum Chemistry **61**, 899 (1997)

116. S.K. Wolff and T. Ziegler, *Calculation of DFT-GIAO NMR shifts with inclusion of spin-orbit coupling*, Journal of Chemical Physics **109**, 895 (1998)

117. S.K. Wolff, T. Ziegler, E. van Lenthe and E.J. Baerends, *Density functional calculations of nuclear magnetic shieldings using the zeroth-order regular approximation (ZORA) for relativistic effects: ZORA nuclear magnetic resonance*, Journal of Chemical Physics **110**, 7689 (1999)

118. J. Autschbach and T. Ziegler, *Nuclear spin-spin coupling constants from regular approximate density functional calculations. I. Formalism and scalar relativistic results for heavy metal compounds*, Journal of Chemical Physics **113**, 936 (2000)

119. J. Autschbach, and T. Ziegler, *Nuclear spin-spin coupling constants from regular approximate relativistic density functional calculations. II. Spin-orbit coupling effects and anisotropies*, Journal of Chemical Physics **113**, 9410 (2000)

120. G. Schreckenbach and T. Ziegler, *Calculation of the G-tensor of electron paramagnetic resonance spectroscopy using Gauge-Including Atomic Orbitals and Density Functional Theory*, Journal of Physical Chemistry A **101**, 3388 (1997)

121. S. Patchkovskii and T. Ziegler, *Calculation of the EPR g-Tensors of High-Spin Radicals with Density Functional Theory*, Journal of Physical Chemistry A **105**, 5490 (2001)

122. C. Edmiston and K. Rudenberg, *Localized Atomic and Molecular Orbitals*, Reviews of Modern Physics **35**, 457 (1963)

123. J.M Foster and S.F. Boys, *Canonical Configurational Interaction Procedure*, Reviews of Modern Physics **32**, 300 (1960)

124. W. von Niessen, *Density Localization of Atomic and Molecular Orbitals. I*, Journal of Chemical Physics **56**, 4290 (1972)

125. F.L. Hirshfeld, *Bonded-atom fragments for describing molecular charge densities*, Theoretica Chimica Acta **44**, 129 (1977)

126. K.B. Wiberg and P.R. Rablen, *Comparison of atomic charges derived via different procedures*, Journal of Computational Chemistry **14**, 1504 (1993)

127. F.M. Bickelhaupt, N.J.R. van Eikema Hommes, C. Fonseca Guerra and E.J. Baerends, *The Carbon-Lithium Electron Pair Bond in $(CH3Li)_n$ (n = 1, 2, 4)*, Organometallics **15**, 2923 (1996)

128. C. Fonseca Guerra, J.-W. Handgraaf, E. J. Baerends and F. M. Bickelhaupt, *Voronoi Deformation Density (VDD) charges. Assessment of the Mulliken, Bader, Hirshfeld, Weinhold and VDD methods for Charge Analysis*, Journal of Computational Chemistry **25**, 189 (2004)

129. C. Fonseca Guerra, F.M. Bickelhaupt, J.G. Snijders and E.J. Baerends, *The Nature of the Hydrogen Bond in DNA Base Pairs: The Role of Charge Transfer and Resonance Assistance*, Chemistry - A European Journal **5**, 3581 (1999)

130. K. Kitaura and K. Morokuma, *A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation*, International Journal of Quantum Chemistry **10**, 325 (1976)

131. T. Ziegler and A. Rauk, *Carbon monoxide, carbon monosulfide, molecular nitrogen, phosphorus trifluoride, and methyl isocyanide as sigma donors and pi acceptors. A theoretical study by the Hartree-Fock-Slater transition-state method*, Inorganic Chemistry **18**, 1755 (1979)

132. H. Fujimoto, J. Osamura and T. Minato, *Orbital interaction and chemical bonds. Exchange repulsion and rehybridization in chemical reactions*, Journal of the American Chemical Society **100**, 2954 (1978)

133. S. Wolfe, D.J. Mitchell and M.-H. Whangbo, *On the role of steric effects in the perturbational molecular orbital method of conformational analysis*, Journal of the American Chemical Society **100**, 1936 (1978)

134. A.J. Stone and R.W. Erskine, *Intermolecular self-consistent-field perturbation theory for organic reactions. I. Theory and implementation; nucleophilic attack on carbonyl compounds*, Journal of the American Chemical Society **102**, 7185 (1980)

135. F. Bernardi, A. Bottoni, A. Mangini and G. Tonachini, *Quantitative orbital analysis of ab initio SCF=MO computations : Part II. Conformational preferences in H2N---OH and H2N---SH*, Journal of Molecular Structure: THEOCHEM **86**, 163 (1981)

136. P.J. van den Hoek and E.J. Baerends, *Chemical bonding at metal-semiconductor interfaces*, Applied Surface Science **41/42**, 236 (1989)

137. J. Autschbach, *On the calculation of relativistic effects and how to understand their trends in atoms and molecules*, in *Chemistry*. 1999, University of Siegen: Siegen.

138. M.A. Watson, N.C. Handy and A.J. Cohen, *Density functional calculations, using Slater basis sets, with exact exchange*, Journal of Chemical Physics **119**, 6475 (2003)

139. Ö. Farkas and H.B. Schlegel, *Methods for optimizing large molecules. Part III. An improved algorithm for geometry optimization using direct inversion in the iterative subspace (GDIIS)*, Physical Chemistry Chemical Physics **4**, 11 (2002)

140. I. Mayer, *Charge, bond order and valence in the ab inition SCF theory*, Chemical Physics Letters **97**, 270 (1983)

141. L. Jensen, P.T. van Duijnen and J.G. Snijders, *A discrete solvent reaction field model within density functional theory*, Journal of Chemical Physics **118**, 514 (2003)

142. L. Jensen, P.T. van Duijnen and J.G. Snijders, *A discrete solvent reaction field model for calculating molecular linear response properties in solution*, Journal of Chemical Physics **119**, 3800 (2003)

143. L. Jensen, P.T. van Duijnen and J.G. Snijders, *A discrete solvent reaction field model for calculating frequency-dependent hyperpolarizabilities of molecules in solution*, Journal of Chemical Physics **119**, 12998 (2003)

144. L. Jensen, M. Swart and P.T. van Duijnen, *Microscopic and macroscopic polarization within a combined quantum mechanics and molecular mechanics model*, Journal of Chemical Physics **122**, 34103 (2005)

145. L. Jensen, *Modelling of optical response properties: Application to nanostructures*, PhD thesis, Rijksuniversiteit Groningen, 2004.

146. P.T. van Duijnen and M. Swart, *Molecular and Atomic Polarizabilities: Thole's Model Revisited*, Journal of Physical Chemistry A **102**, 2399 (1998)

147. L. Jensen, P.-O. Astrand, A. Osted, J. Kongsted and K.V. Mikkelsen, *Polarizability of molecular clusters as calculated by a dipole interaction model*, Journal of Chemical Physics **116**, 4001 (2002)

148. A. Michalak, R.L. De Kock and T. Ziegler, *Bond Multiplicity in Transition-Metal Complexes: Applications of Two-Electron Valence Indices*, Journal of Physical Chemistry A **112**, 7256 (2008)

149. R.F. Nalewajski and J. Mrozek, *Modified valence indices from the two-particle density matrix*, International Journal of Quantum Chemistry **51**, 187 (1994)

150. R.F. Nalewajski, J. Mrozek and A. Michalak, *Two-electron valence indices from the Kohn-Sham orbitals*, International Journal of Quantum Chemistry **61**, 589 (1997)

151. R.F. Nalewajski, J. Mrozek and A. Michalak, *Exploring Bonding Patterns of Molecular Systems Using Quantum Mechanical Bond Multiplicities*, Polish Journal of Chemistry **72**, 1779 (1998)

152. R.F. Nalewajski, J. Mrozek and G. Mazur, *Quantum chemical valence indices from the one-determinantal difference approach*, Canadian Journal of Chemistry **74**, 1121 (1996)

153. M.S. Gopinathan and K. Jug, *Valency. I. A quantum chemical definition and properties*, Theoretica Chimica Acta 1983 **63**, 497 (1983)

154. F. Wang and T. Ziegler, *Excitation energies of some d1 systems calculated using time-dependent density functional theory: an implementation of open-shell TDDFT theory for doublet-doublet excitations*, Molecular Physics **102**, 2585 (2004)

155. Z. Rinkevicius, I. Tunell, P. Salek, O. Vahtras and H. Agren, *Restricted density functional theory of linear time-dependent properties in open-shell molecules*, Journal of Chemical Physics **119**, 34 (2003)

156. F. Wang and T. Ziegler, *Time-dependent density functional theory based on a noncollinear formulation of the exchange-correlation potential*, Journal of Chemical Physics **121**, 12191 (2004)

157. F. Wang and T. Ziegler, *The performance of time-dependent density functional theory based on a noncollinear exchange-correlation potential in the calculations of excitation energies*, Journal of Chemical Physics **122**, 74109 (2005)

158. S. Hirata and M. Head-Gordon, *Time-dependent density functional theory within the Tamm-Dancoff approximation*, Chemical Physics Letters **314**, 291 (1999)

159. G. Henkelman, B.P. Uberuaga and H. Jonssón, *A climbing image nudged elastic band method for finding saddle points and minimum energy paths*, Journal of Chemical Physics **113**, 9901 (2000)

160. G. Vignale and W. Kohn, *Current-Dependent Exchange-Correlation Potential for Dynamical Linear Response Theory*, Physical Review Letters **77**, 2037 (1996)

161. G. Vignale and W. Kohn, in *Electronic Density Functional Theory: Recent Progress and New Direction*, ISBN13: 9780306458347, J. F. Dobson, G. Vignale, and M. P. Das, Editors. 1998, Plenum: New York.

162. M. van Faassen, P. L. de Boeij, R. van Leeuwen, J. A. Berger and J. G. Snijders, *Ultranonlocality in Time-Dependent Current-Density-Functional Theory: Application to Conjugated Polymers*, Physical Review Letters **88**, 186401 (2002)

163. M. van Faassen, P. L. de Boeij, R. van Leeuwen, J. A. Berger and J. G. Snijders, *Application of time-dependent current-density-functional theory to nonlocal exchange-correlation effects in polymers*, Journal of Chemical Physics **118**, 1044 (2003)

164. M. van Faassen and P. L. de Boeij, *Excitation energies for a benchmark set of molecules obtained within time-dependent current-density functional theory using the Vignale-Kohn functional*, Journal of Chemical Physics **120**, 8353 (2004)

165. M. van Faassen and P. L. de Boeij, *Excitation energies of Π-conjugated oligomers within time-dependent current-density-functional theory*, Journal of Chemical Physics **121**, 10707 (2004)

166. M. van Faassen, *Time-Dependent Current-Density-Functional Theory for Molecules*, PhD thesis, Rijksuniversiteit Groningen, 2004.

167. R. Nifosi, S. Conti and M. P. Tosi, *Dynamic exchange-correlation potentials for the electron gas in dimensionality D=3 and D=2*, Physical Review B **58**: p. 12758 (1998)

168. Z. X. Qian and G. Vignale, *Dynamical exchange-correlation potentials for the electron liquid in the spin channel*, Physical Review B **68**, 195113 (2003)

169. M. Stener, G. Fronzoni and M. de Simone, *Time dependent density functional theory of core electrons excitations*, Chemical Physics Letters **373**, 115 (2003)

170. M. Swart, P.Th. van Duijnen and J.G. Snijders, *A charge analysis derived from an atomic multipole expansion*, Journal of Computational Chemistry **22**, 79 (2001)

171. T.W. Keal and D.J. Tozer, *The exchange-correlation potential in Kohn.Sham nuclear magnetic resonance shielding calculations*, Journal of Chemical Physics **119**, 3015 (2003)

172. X. Xu and W.A. Goddard III, *The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties*, Proceedings of the National Academy of Sciences **101**, 2673 (2004)

173. J. Baker, A. Kessi and B. Delley, *The generation and use of delocalized internal coordinates in geometry optimization*, Journal of Chemical Physics 1996 **105**, 192 (1996)

174. C. Adamo and V. Barone, *Physically motivated density functionals with improved performances: The modified Perdew.Burke.Ernzerhof model*, Journal of Chemical Physics 1996 **116**, 5933 (1996)

175. M. Swart, A.W. Ehlers and K. Lammertsma, *Performance of the OPBE exchange-correlation functional*, Molecular Physics 2004 **102**, 2467 (2004)

176. P.J. Stephens, F.J. Devlin, C.F. Chabalowski and M.J. Frisch, *Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields*, Journal of Physical Chemistry **98**, 11623 (1994)

177. M. Reiher, O. Salomon and B.A. Hess, *Reparameterization of hybrid functionals based on energy differences of states of different multiplicity*, Theoretical Chemistry Accounts **107**, 48 (2001)

178. C. Adamo and V. Barone, *Toward reliable adiabatic connection models free from adjustable parameters*, Chemical Physics Letters **274**, 242 (1997)

179. J.K. Kang and C.B. Musgrave, *Prediction of transition state barriers and enthalpies of reaction by a new hybrid density-functional approximation*, Journal of Chemical Physics **115**, 11040 (2001)

180. A.J. Cohen and N.C. Handy, *Dynamic correlation*, Molecular Physics **99**, 607 (2001)

181. B.J. Lynch, P.L. Fast, M. Harris and D.G. Truhlar, *Adiabatic Connection for Kinetics*, Journal of Physical Chemistry A **104**, 4811 (2000)

182. F. Wang, T. Ziegler, E. van Lenthe, S.J.A. van Gisbergen and E.J. Baerends, *The calculation of excitation energies based on the relativistic two-component zeroth-order regular approximation and time-dependent density-functional with full use of symmetry*, Journal of Chemical Physics **122**, 204103 (2005)

183. F. Wang and T. Ziegler, *Theoretical study of the electronic spectra of square-planar platinum (II) complexes based on the two-component relativistic time-dependent density-functional theory*, Journal of Chemical Physics **123**, 194102 (2005)

184. T.A. Wesolowski and A. Warshel, *Frozen Density Functional Approach for ab-initio Calculations of Solvated Molecules*, Journal of Physical Chemistry **97**, 8050 (1993)

185. J. Neugebauer, C.R. Jacob, T.A. Wesolowski and E.J. Baerends, *An Explicit Quantum Chemical Method for Modeling Large Solvation Shells Applied to Aminocoumarin C151*, Journal of Physical Chemistry A **109**, 7805 (2005)

186. M.E. Casida and T.A. Wesolowski, *Generalization of the Kohn-Sham equations with constrained electron density formalism and its time-dependent response theory formulation*, International Journal of Quantum Chemistry **96**, 577 (2004)

187. T.A. Wesolowski, *Hydrogen-Bonding-Induced Shifts of the Excitation Energies in Nucleic Acid Bases: An Interplay between Electrostatic and Electron Density Overlap Effects*, Journal of the American Chemical Society **126**, 11444 (2004)

188. T.A. Wesolowski, *Density functional theory with approximate kinetic energy functionals applied to hydrogen bonds*, Journal of Chemical Physics **106**, 8516 (1997)

189. C.R. Jacob, T.A. Wesolowski and L. Visscher, *Orbital-free embedding applied to the calculation of induced dipole moments in $CO_2^{\cdots}X$ (X=He, Ne, Ar, Kr, Xe, Hg) van der Waals complexes*, Journal of Chemical Physics **123**, 174104 (2005)

190. C.R. Jacob, J. Neugebauer, L. Jensen and L. Visscher, *Comparison of frozen-density embedding and discrete reaction field solvent models for molecular properties*, Physical Chemistry Chemical Physics **8**, 2349 (2006)

191. J. Neugebauer, M.J. Louwerse, E.J. Baerends and T.A. Wesolowski, *The merits of the frozen-density embedding scheme to model solvatochromic shifts*, Journal of Chemical Physics **122**, 94115 (2005)

192. J. Neugebauer, M.J. Louwerse, P. Belanzoni, T.A. Wesolowski and E.J. Baerends, *Modeling solvent effects on electron-spin-resonance hyperfine couplings by frozen-density embedding*, Journal of Chemical Physics **123**, 114101 (2005)

193. J. Neugebauer and E.J. Baerends, *Exploring the Ability of Frozen-Density Embedding to Model Induced Circular Dichroism*, Journal of Physical Chemistry A **110**, 8786 (2006)

194. L.H. Thomas, *The calculation of atomic fields*, Mathematical Proceedings of the Cambridge Philosophical Society **23**, 542 (1927)

195. E. Fermi, *Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente*, Zeitschrift für Physik **48**, 73 (1928)

196. C.F. von Weizsäcker, *Zur Theorie der Kernmassen*, Zeitschrift für Physik **96**, 431 (1935)

197. A. Lembarki and H. Chermette, *Obtaining a gradient-corrected kinetic-energy functional from the Perdew-Wang exchange functional*, Physical Review A **50**, 5328 (1994)

198. H. Lee, C. Lee and R.G. Parr, *Conjoint gradient correction to the Hartree-Fock kinetic- and exchange-energy density functionals*, Physical Review A **44**, 768 (1991)

199. J.P. Perdew and Wang Yue, *Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation*, Physical Review B **33**, 8800 (1986), ErratumPhysical Review B **40**, 3399 (1989)

200. H. Ou-Yang and M. Levy, *Approximate noninteracting kinetic energy functionals from a nonuniform scaling requirement*, International Journal of Quantum Chemistry **40**, 379 (1991)

201. A.J. Thakkar, *Comparison of kinetic-energy density functionals*, Physical Review A **46**, 6920 (1992)

202. J. Neugebauer, E.J. Baerends, E. Efremov, F. Ariese and C. Gooijer, *Combined Theoretical and Experimental Deep-UV Resonance Raman Studies of Substituted Pyrenes*, Journal of Physical Chemistry A **109**, 2100 (2005)

203. J. Neugebauer, E.J. Baerends and M. Nooijen, *Vibronic coupling and double excitations in linear response time-dependent density functional calculations: Dipole-allowed states of N_2*, Journal of Chemical Physics **121**, 6155 (2004)

204. J. Neugebauer, Vibronic Coupling Calculations using ADF, documentation on the VIBRON module available on request.

205. T.A. Wesolowski, in: Computational Chemistry: Reviews of Current Trends - Vol. 10, World Scientific, 2006.

206. M. Zbiri, M. Atanasov, C. Daul, J.-M. Garcia Lastra and T.A. Wesolowski, *Application of the density functional theory derived orbital-free embedding potential to calculate the splitting energies of lanthanide cations in chloroelpasolite crystals*, Chemical Physics Letters **397**, 441 (2004)

207. M. Zbiri, C.A. Daul and T.A. Wesolowski, *Effect of the f-Orbital Delocalization on the Ligand-Field Splitting Energies in Lanthanide-Containing Elpasolites*, Journal of Chemical Theory and Computation **2**, 1106 (2006)

208. A. Bérces, R. M. Dickson, L. Fan, H. Jacobsen, D. Swerhone and T. Ziegler, *An implementation of the coupled perturbed Kohn-Sham equations: perturbation due to nuclear displacements*, Computer Physics Communications **100**, 247 (1997)

209. H. Jacobsen, A. Bérces, D. Swerhone and T. Ziegler, *Analytic second derivatives of molecular energies: a density functional implementation*, Computer Physics Communications **100**, 263 (1997)

210. S. K. Wolff, *Analytical second derivatives in the Amsterdam density functional package*, International Journal of Quantum Chemistry **104**, 645 (2005)

211. S. Grimme, *Accurate description of van der Waals complexes by density functional theory including empirical corrections*, Journal of Computational Chemistry **25**, 1463 (2004)

212. M. Ernzerhof and G. Scuseria, *Assessment of the Perdew.Burke.Ernzerhof exchange-correlation functional*, Journal of Chemical Physics **110**, 5029 (1999)

213. C. Adamo and V. Barone, *Toward reliable density functional methods without adjustable parameters: The PBE0 model*, Journal of Chemical Physics **110**, 6158 (1999)

214. A.D. Buckingham, P.W. Fowler and J.M. Hutson, *Theoretical studies of van der Waals molecules and intermolecular forces*, Chemical Reviews **88**, 963 (1988)

215. J.M. Ducéré and L. Cavallo, *Parametrization of an Empirical Correction Term to Density Functional Theory for an Accurate Description of .-Stacking Interactions in Nucleic Acids*, Journal of Physical Chemistry B **111**, 13124 (2007)

216. V.P. Nicu J. Neugebauer S.K. Wolff and E.J. Baerends, *A vibrational circular dichroism implementation within a Slater-type-orbital based density functional framework and its application to hexa- and hepta-helicenes*, Theoretical Chemical Accounts **119**, 245 (2008)

217. C.R. Jacob, J. Neugebauer and L. Visscher, *A flexible implementation of frozen-density embedding for use in multilevel simulations*, Journal of Computational Chemistry **29**, 1011 (2008)

218. C.R. Jacob and L. Visscher, *Calculation of nuclear magnetic resonance shieldings using frozen-density embedding*, Journal of Chemical Physics **125**, 194104 (2006)

219. V. Bakken and T. Helgaker, *The efficient optimization of molecular geometries using redundant internal coordinates*, Journal of Chemical Physics **117**, 9160 (2002)

220. C.R. Jacob, S.M., Beyhan and L. Visscher, *Exact functional derivative of the nonadditive kinetic-energy bifunctional in the long-distance limit*, Journal of Chemical Physics **126**, 234116 (2007)

221. Y. Zhao, N.E. Schultz and D.G. Truhlar, *Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions*, Journal of Chemical Physics **123**, 161103 (2005)

222. 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*, Journal of Chemical Theory and Computation **2**, 364 (2006)

223. Y. Zhao and D.G. Truhlar, *A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions*, Journal of Chemical Physics **125**, 194101 (2006)

224. 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*, Theoretical Chemical Accounts **120**, 215 (2008)

225. M. Swart and F.M. Bickelhaupt, *Optimization of strong and weak coordinates*, International Journal of Quantum Chemistry **106**, 2536 (2006)

226. S. Grimme, *Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction*, Journal of Computational Chemistry **27**, 1787 (2006)

227. S. Grimme, , J. Antony, T. Schwabe and C. Mück-Lichtenfeld, *Density Functional Theory with Dispersion Corrections for Supramolecular Structures, Aggregates, and Complexes of (Bio)Organic Molecules*, Organic & Biomolecular Chemistry **5**, 741 (2007)

228. J.I. Rodríguez, A.M. Köster, P.W. Ayers, A. Santos-Valle, A. Vela and G. Merino, *An efficient grid-based scheme to compute QTAIM atomic properties without explicit calculation of zero-flux surfaces*, Journal of Computational Chemistry **30**, 1082 (2009)

229. J.I. Rodríguez, R.F.W. Bader, P.W. Ayers, C. Michel, A.W. Götz and C. Bo, *A high performance grid-based algorithm for computing QTAIM properties*, Chemical Physics Letters **472**, 149 (2009)

230. M. Krykunov and J. Autschbach, *Calculation of static and dynamic linear magnetic response in approximate time-dependent density functional theory*, Journal of Chemical Physics **126**, 24101 (2007)

231. M. Krykunov, M.D. Kundrat and J. Autschbach, *Calculation of CD spectra from optical rotatory dispersion, and vice versa, as complementary tools for theoretical studies of optical activity using time-dependent density functional theory*, Journal of Chemical Physics **125**, 194110 (2006)

232. M. Krykunov and J. Autschbach, *Calculation of origin independent optical rotation tensor components for chiral oriented systems in approximate time-dependent density functional theory*, Journal of Chemical Physics **125**, 34102 (2006)

233. J. Autschbach, L. Jensen, G.C. Schatz, Y.C.E. Tse and M. Krykunov, *Time-dependent density functional calculations of optical rotatory dispersion including resonance wavelengths as a potentially useful tool for determining absolute configurations of chiral molecules*, Journal of Physical Chemistry A **110**, 2461 (2006)

234. M. Krykunov and J. Autschbach, *Calculation of optical rotation with time-periodic magnetic field-dependent basis functions in approximate time-dependent density functional theory*, Journal of Chemical Physics **123**, 114103 (2005)

235. A. Baev, M. Samoc, P.N. Prasad, M. Krykunov and J. Autschbach, *A Quantum Chemical Approach to the Design of Chiral Negative Index Materials*, Optics Express **15**, 5730 (2007)

236. M. Krykunov, A. Banerjee, T. Ziegler and J. Autschbach, *Calculation of Verdet constants with time-dependent density functional theory: Implementation and results for small molecules*, Journal of Chemical Physics **122**, 74105 (2005)

237. P. Cortona, *Self-consistently determined properties of solids without band-structure calculations*, Physical Review B **44**, 8454 (1991)

238. T.A. Wesolowski and J. Weber, *Kohn-Sham equations with constrained electron density: The effect of various kinetic energy functional parametrizations on the ground-state molecular properties*, International Journal of Quantum Chemistry **61**, 303 (1997)

239. T.A. Wesolowski, H. Chermette and J. Weber, *Accuracy of Approximate Kinetic Energy Functionals in the Model of Kohn-Sham Equations with Constrained Electron Density: the FH···NCH complex as a Test Case*, Journal of Chemical Physics **105**, 9182 (1996)

240. T.A. Wesolowski and J. Weber, *Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules*, Chemical Physics Letters **248**, 71 (1996)

241. Y.A. Bernard, M. Dulak, J.W. Kaminski and T.A. Wesolowski, *The energy-differences based exact criterion for testing approximations to the functional for the kinetic energy of non-interacting electrons*, Journal of Physics A **41**, 55302 (2008)

242. D.A. Kirzhnits, Soviet Physics JETP-USSR **5**, 64 (1957)

243. P. Fuentealba and O. Reyes, *Further evidence of the conjoint correction to the local kinetic and exchange energy density functionals*, Chemical Physics Letters **232**, 31 (1995)

244. O.V. Gritsenko, P.R.T. Schipper and E.J. Baerends, *Approximation of the exchange-correlation Kohn-Sham potential with a statistical average of different orbital model potentials*, Chemical Physics Letters **302**, 199 (1999)

245. B. Delley, *The conductor-like screening model for polymers and surfaces*, Molecular Simulation **32**, 117 (2006)

246. J. Tao, J.P. Perdew, V.N. Staroverov and G.E. Scuseria, *Climbing the Density Functional Ladder: Nonempirical MetaGeneralized Gradient Approximation Designed for Molecules and Solids* Physical Review Letters **91**, 146401 (2003)

247. V.N. Staroverov, G.E. Scuseria, J. Tao and J.P. Perdew, *Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes* Journal of Chemical Physics **119**, 12129 (2003)

248. S. Ivanov, S. Hirata, R. J. Bartlett, *Exact Exchange Treatment for Molecules in Finite-Basis-Set Kohn-Sham Theory*, Physical Review Letters **83**, 5455 (1999)

249. A. F. Izmaylov, V. N. Staroverov, G. E. Scuseria, E. R. Davidson, G. Stoltz, E. Cancès, *The effective local potential method: Implementation for molecules and relation to approximate optimized effective potential techniques*, Journal of Chemical Physics **126**, 084107 (2007)

250. M. Krykunov and T. Ziegler, *On the use of the exact exchange optimized effective potential method for static response properties*, International Journal of Quantum Chemistry **109**, 3246 (2009)

251. M. L. Connolly, *Solvent-accessible surfaces of proteins and nucleic acids*, Science, **221**, 709 (1983)

252. L. Onsager, *Electric moments of molecules in liquids*, Journal of the American Chemical Society **58**, 1486 (1936)

253. S. Miertus, E. Scrocco and J. Tomasi, *Electrostatic interaction of a solute with a continuum: a direct utilization of ab initio molecular potentials for the prevision of solvent effects*, Chemical Physics **55**, 117 (1981)

254. J. Tomasi, R. Bonaccorsi, R. Cammi and F.J. Olivares del Valle, *Theoretical chemistry in solution. Some results and perspectives of the continuum methods and in particular of the polarizable continuum model*, Journal of Molecular Structure: THEOCHEM **234**, 401 (1991)

255. J.L. Chen, L. Noodleman, D.A. Case and D. Bashford, *Incorporating solvation effects into density functional electronic structure calculations*, Journal of Physical Chemistry **98**, 11059 (1994)

256. J.-M. Mouesca, J.L. Chen, L. Noodleman, D. Bashford and D.A. Case, *Density functional/Poisson-Boltzmann calculations of redox potentials for iron-sulfur clusters*, Journal of the American Chemical Society **116**, 11898 (1994)

257. A. Fortunelli and J. Tomasi, *The implementation of density functional theory within the polarizable continuum model for solvation*, Chemical Physics Letters **231**, 34 (1994)

258. C.M. Breneman and K.B. Wiberg, *Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis*, Journal of Computational Chemistry **11**, 361 (1990)

259. F.M. Richards, *Areas, volumes, packing and protein structures*, Annual Review of Biophysics and Bioengineering **6**, 151 (1977)

260. T. You and D. Bashford, *An analytical algorithm for the rapid determination of the solvent accessibility of points in a three-dimensional lattice around a solute molecule*, Journal of Computational Chemistry **16**, 743 (1995)

261. M. Mitoraj, A. Michalak and T. Ziegler, *A Combined Charge and Energy Decomposition Scheme for Bond Analysis*, Journal of Chemical Theory and Computation **5**, 962 (2009)

262. M. Mitoraj, A. Michalak and T. Ziegler, *On the Nature of the Agostic Bond between Metal Centers and Beta-Hydrogen Atoms in Alkyl Complexes. An Analysis Based on the Extended Transition State Method and the Natural Orbitals for Chemical Valence Scheme (ETS-NOCV)*, Organometallics **28**, 3727 (2009)

263. A.W. Götz, S.M. Beyhan and L. Visscher, *Performance of Kinetic Energy Functionals for Interaction Energies in a Subsystem Formulation of Density Functional Theory*, Journal of Chemical Theory and Computation **5**, 3161 (2009)

264. M. Ernzerhof, *The role of the kinetic energy density in approximations to the exchange energy*, Journal of Molecular Structure: THEOCHEM **501-502**, 59 (2000)

265. J.P. Perdew, *Generalized gradient approximation for the fermion kinetic energy as a functional of the density*, Physics Letters A **165**, 79 (1992)

266. L. Jensen, L. Zhao, J. Autschbach and G.C. Schatz, *Theory and method for calculating resonance Raman scattering from resonance polarizability derivatives*, Journal of Chemical Physics **123**, 174110 (2005)

267. L. Jensen, L. Zhao, J. Autschbach and G.C. Schatz, *Resonance Raman Scattering of Rhodamine 6G as calculated using Time-Dependent Density Functional Theory*, Journal of Physical Chemistry A **110**, 5973 (2006)

268. L.L. Zhao, L. Jensen and G.C. Schatz, *Pyridine - Ag_{20} Cluster: A Model System for Studying Surface-Enhanced Raman Scattering*, Journal of the American Chemical Society **128**, 2911 (2006)

269. L. Jensen, L.L. Zhao and G.C. Schatz, *Size-Dependence of the Enhanced Raman Scattering of Pyridine Adsorbed on Ag_n (n=2-8,20) Clusters*, Journal of Physical Chemistry C **111**, 4756 (2007)

270. J. Autschbach, *Magnitude of Finite-Nucleus-Size Effects in Relativistic Density Functional Computations of Indirect NMR Nuclear Spin-Spin Coupling Constants*, ChemPhysChem **10**, 2274 (2009)

271. S. Høst, J. Olsen, B. Jansík, L. Thøgersen, P. Jørgensen and T. Helgaker, *The augmented Roothaan-Hall method for optimizing Hartree-Fock and Kohn-Sham density matrices*, Journal of Chemical Physics **129**, 124106 (2008)

272. M. Krykunov, M. Seth, T. Ziegler and J. Autschbach, *Calculation of the magnetic circular dichroism B term from the imaginary part of the Verdet constant using damped time-dependent density functional theory*, Journal of Chemical Physics **127**, 244102 (2007)

273. S.B. Piepho and P. N. Schatz, *Group Theory in Spectroscopy With Application to Magnetic Circular Dichroism*, (Wiley, New York, 1983).

274. W.R. Mason, *A Practical Guide to Magnetic Circular Dichroism Spectroscopy*, (Wiley, New Jersey, 2007).

275. M. Seth and T. Ziegler, *Formulation of magnetically perturbed time-dependent density functional theory*, Journal of Chemical Physics **127**, 134108 (2007)

276. M. Seth, M. Krykunov, T. Ziegler, J. Autschbach and A. Banerjee, *Application of magnetically perturbed time-dependent density functional theory to magnetic circular dichroism: Calculation of B terms*, Journal of Chemical Physics **128**, 144105 (2008)

277. M. Seth, M. Krykunov, T. Ziegler and J. Autschbach, *Application of magnetically perturbed time-dependent density functional theory to magnetic circular dichroism. II. Calculation of A terms*, Journal of Chemical Physics **128**, 234102 (2008)

278. M. Seth, T. Ziegler and J. Autschbach, *Application of magnetically perturbed time-dependent density functional theory to magnetic circular dichroism. III. Temperature-dependent magnetic circular dichroism induced by spin-orbit coupling*, Journal of Chemical Physics **129**, 104105 (2008)

279. J.M. Garcia Lastra, J.W. Kaminski and T.A. Wesolowski, *Orbital-free effective embedding potential at nuclear cusps*, Journal of Chemical Physics **129**, 074107 (2008)

280. F. Wang and T. Ziegler, *A simplified relativistic time-dependent density-functional theory formalism for the calculations of excitation energies including spin-orbit coupling effect*, Journal of Chemical Physics **123**, 154102 (2005)

281. W.-G. Han, T. Liu, T. Lovell and L. Noodleman, *DFT calculations of ^{57}Fe Mössbauer isomer shifts and quadrupole splittings for iron complexes in polar dielectric media: Applications to methane monooxygenase and ribonucleotide reductase*, Journal of Computational Chemistry **27**, 1292 (2006)

282. A. Ghysels, D. Van Neck, V. Van Speybroeck, T. Verstraelen and M. Waroquier, *Vibrational Modes in partially optimized molecular systems*, Journal of Chemical Physics**126**, 224102 (2007)

283. A. Ghysels, D. Van Neck and M. Waroquier, *Cartesian formulation of the Mobile Block Hessian Approach to vibrational analysis in partially optimized systems*, Journal of Chemical Physics **127**, 164108 (2007)

284. J. J. Mortensen, K. Kaasbjerg, S. L. Frederiksen, J. K. Nørskov, J. P. Sethna, and K. W. Jacobsen, *Bayesian Error Estimation in Density-Functional Theory*, Physical Review Letters **95**, 216401 (2005)

285. J.P. Perdew, A. Ruzsinszky, G.I. Csonka, O.A. Vydrov, G.E. Scuseria, *Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces*, Physical Review Letters **100**, 136406 (2008)

286. M. Swart, M. Solà and F.M. Bickelhaupt, *A new all-round DFT functional based on spin states and S_N2 barriers*, Journal of Chemical Physics **131**, 094103 (2009)

287. M. Swart, M. Solà and F.M. Bickelhaupt, *Switching between OPTX and PBE exchange functionals*, Journal of Computational Methods in Science and Engineering **9**, 69 (2009)

288. J.P. Perdew and Y. Wang, *Accurate and simple analytic representation of the electron-gas correlation energy*, Physical Review B **45**, 13244 (1992)

289. K.N. Kudin, G.E. Scuseria and E. Cances, *A black-box self-consistent field convergence algorithm: One step closer*, Journal of Chemical Physics **116**, 8255 (2002)

290. N.L. Allinger, X. Zhou, J. Bergsma, *Molecular mechanics parameters*, Journal of Molecular Structure: THEOCHEM **312**, 69 (1994)

291. X. Hu and W. Yang, *Accelerating self-consistent field convergence with the augmented Roothaan-Hall energy function*, Journal of Chemical Physics **132**, 054109 (2010)

292. S. Grimme, J. Anthony, S. Ehrlich, and H. Krieg, *A consistent and accurate* ab initio *parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu*, Journal of Chemical Physics **132**, 154104 (2010).

293. M.D. Newton, *Quantum chemical probes of electron-transfer kinetics: the nature of donor-acceptor interactions*, Chemical Reviews **91**, 767 (1991).

294. K. Senthilkumar, F.C. Grozema, F.M. Bickelhaupt, and L.D.A. Siebbeles, *Charge transport in columnar stacked triphenylenes: Effects of conformational fluctuations on charge transfer integrals and site energies*, Journal of Chemical Physics **119**, 9809 (2003).

295. K. Senthilkumar, F.C. Grozema, C. Fonseca Guerra, F.M. Bickelhaupt, F.D. Lewis, Y.A. Berlin, M.A. Ratner, and L.D.A. Siebbeles, *Absolute Rates of Hole Transfer in DNA*, Journal of the American Chemical Society **127**, 14894 (2005)

296. J. Neugebauer, *Couplings between electronic transitions in a subsystem formulation of time-dependent density functional theory*, Journal of Chemical Physics **126**, 134116 (2007).

297. J. Neugebauer, *Photophysical Properties of Natural Light-Harvesting Complexes Studied by Subsystem Density Functional Theory*, Journal of Physical Chemistry B **112**, 2207 (2008)

298. J. Neugebauer, *On the calculation of general response properties in subsystem density functional theory*, Journal of Chemical Physics **131**, 084104 (2009).

299. T.N. Truong and E.V. Stefanovich, *A new method for incorporating solvent effect into the classical, ab initio molecular orbital and density functional theory frameworks for arbitrary shape cavity*, Chemical Physics Letters **240**, 253 (1995)

300. S. Gusarov, T. Ziegler, and A. Kovalenko, *Self-Consistent Combination of the Three-Dimensional RISM Theory of Molecular Solvation with Analytical Gradients and the Amsterdam Density Functional Package*, Journal of Physical Chemistry A **110**, 6083 (2006)

301. D. Casanova, S. Gusarov, A. Kovalenko, and T. Ziegler, *Evaluation of the SCF Combination of KS-DFT and 3D-RISM-KH; Solvation Effect on Conformational Equilibria, Tautomerization Energies, and Activation Barriers*, Journal of Chemical Theory and Computation **3**, 458 (2007)

302. A. Kovalenko and F. Hirata, *Self-consistent description of a metal-water interface by the Kohn-Sham density functional theory and the three-dimensional reference interaction site model*, Journal of Chemical Physics **110**, 10095 (1999)

303. A. Kovalenko and F. Hirata, *Potentials of mean force of simple ions in ambient aqueous solution. I. Three-dimensional reference interaction site model approach*, Journal of Chemical Physics **112**, 10391 (2000)

304. A. Kovalenko, *Three-dimensional RISM theory for molecular liquids and solid-liquid interfaces*, In Molecular Theory of Solvation; Hirata, Fumio, Ed.; Understanding Chemical Reactivity (series); Mezey, Paul G., Series Ed.; Kluwer Acadamic Publishers: Dordrecht, The Netherlands, 2003; Vol. 24, pp 169-275.

305. J.W. Kaminski, S. Gusarov, A. Kovalenko, T.A. Wesolowski, *Modeling solvatochromic shifts using the orbital-free embedding potential at statistically mechanically averaged solvent density*, Journal of Physical Chemistry A **114**, 6082 (2010)

306. L. Jensen, J. Autschbach, M. Krykunov, and G.C. Schatz, *Resonance vibrational Raman optical activity: A time-dependent density functional theory approach*, Journal of Chemical Physics **127**, 134101 (2007)

307. M. Krykunov, A. Banerjee, T. Ziegler and J. Autschbach, *Calculation of Verdet constants with time-dependent density functional theory. Implementation and results for small molecules*, Journal of Chemical Physics **122**, 074105 (2005)

308. E.J. Baerends, D.E. Ellis and P. Ros, *Self-consistent molecular Hartree-Fock-Slater calculations I. The computational procedure*, Chemical Physics **2**, 41 (1973)

309. E.J. Baerends and P. Ros, *Evaluation of the LCAO Hartree-Fock-Slater method: Applications to transition-metal complexes*, International Journal of Quantum Chemistry **14, S12**, 169 (1978)

310. G. te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler, *Chemistry with ADF*, Journal of Computational Chemistry **22**, 931 (2001)

311. A. Devarajan, A. Gaenko, and J. Autschbach, *Two-component relativistic density functional method for computing nonsingular complex linear response of molecules based on the zeroth order regular approximation*, Journal of Chemical Physics **130**, 194102 (2009)

312. M. R. Pederson, S.N. Khanna, *Magnetic anisotropy barrier for spin tunneling in $Mn_{12}O_{12}$ molecules*, Physical Review B **60**, 9566 (1999)

313. F. Neese, *Calculation of the zero-field splitting tensor on the basis of hybrid density functional and Hartree-Fock theory*, Journal of Chemical Physics **127**, 164112 (2007)

314. C. van Wüllen, *Magnetic anisotropy from density functional calculations. Comparison of different approaches: $Mn_{12}O_{12}$ acetate as a test case*, Journal of Chemical Physics **130**, 194109 (2009)

315. S. Schmitt, P. Jost, C. van Wüllen, *Zero-field splittings from density functional calculations: Analysis and improvement of known methods*, Journal of Chemical Physics **134**, 194113 (2011)

316. S. N. Steinmann, and C. Corminboeuf, *Comprehensive Benchmarking of a Density-Dependent Dispersion Correction*, Journal of Chemical Theory and Computation **7**, 3567 (2011).

317. R.C. Raffenetti, *Eventempered atomic orbitals. II. Atomic SCF wavefunctions in terms of eventempered exponential bases*, Journal of Chemical Physics **59**, 5936 (1973)

318. D.P. Chong, *Completeness profiles of one-electron basis sets*, Canadian Journal of Chemistry **73**, 79 (1995)

319. G.D. Zeiss, W.R. Scott, N. Suzuki, D.P. Chong, S.R. Langhoff, *Finite-field calculations of molecular polarizabilities using field-induced polarization functions: second- and third-order perturbation correlation corrections to the coupled Hartree-Fock polarizability of H_2O*, Molecular Physics **37**, 1543 (1979)

320. E. van Lenthe and E.J. Baerends, *Optimized Slater-type basis sets for the elements 1-118*, Journal of Computational Chemistry **24**, 1142 (2003)

321. D.P. Chong, E. van Lenthe, S.J.A. van Gisbergen and E.J. Baerends, *Even-tempered Slater-Type orbitals revisited: From Hydrogen to Krypton*, Journal of Computational Chemistry **25**, 1030 (2004)

322. D.P. Chong, *Augmenting basis set for time-dependent density functional theory calculation of excitation energies: Slater-type orbitals for hydrogen to krypton*, Molecular Physics **103**, 749 (2005)

323. T. Ziegler, A. Rauk and E.J. Baerends, *On the calculation of Multiplet Energies by the Hartree Fock Slater method*, Theoretica Chimica Acta **43**, 261 (1977)

324. C. Daul, *DFT applied to excited states*, International Journal of Quantum Chemistry **52**, 867 (1994)

325. E.J. Baerends, V. Branchadell and M. Sodupe, *Atomic reference energies for density functional calculations*, Chemical Physics Letters **265**, 481 (1997)

326. P.J. van den Hoek, E.J. Baerends, and R.A. van Santen, *Ethylene epoxidation on silver(110): the role of subsurface oxygen*, Journal of Physical Chemistry **93**, 6469 (1989)

327. J. Autschbach, S. Zheng, and R.W. Schurko, *Analysis of Electric Field Gradient Tensors at Quadrupolar Nuclei in Common Structural Motifs*, Concepts in Magnetic Resonance Part A **36A**, 84 (2010)

328. A.J. Rossini, R.W. Mills, G.A. Briscoe, E.L. Norton, S.J. Geier, I. Hung, S. Zheng, J. Autschbach, and R.W. Schurko, *Solid-State Chlorine NMR of Group IV Transition Metal Organometallic Complexes*, Journal of the American Chemical Society **131**, 3317 (2009)

329. J. Autschbach, *Analyzing NMR shielding tensors calculated with two-component relativistic methods using spin-free localized molecular orbitals*, Journal of Chemical Physics **128**, 164112 (2008)

330. J. Autschbach and S. Zheng, *Analyzing Pt chemical shifts calculated from relativistic density functional theory using localized orbitals: The role of Pt 5d lone pairs*, Magnetic Resonance in Chemistry **46**, S45 (2008)

331. J. Autschbach and S. Zheng, *Relativistic computations of NMR parameters from first principles: Theory and applications*, Annual Reports on NMR Spectroscopy **67**, 1 (2009)

332. J. Autschbach, *Analyzing molecular properties calculated with two-component relativistic methods using spin-free Natural Bond Orbitals: NMR spin-spin coupling constants* Journal of Chemical Physics **127**, 124106 (2007)

333. J. Autschbach and B. Le Guennic, *Analyzing and interpreting NMR spin-spin coupling constants from molecular orbital calculations*, Journal of Chemical Education **84**, 156 (2007)

334. A.M.A. Boshaala, S.J. Simpson, J. Autschbach and S. Zheng, *Synthesis and Characterization of the Trihalophosphine Compounds of Ruthenium [$RuX_2(\eta^6$-cymene)(PY_3)] (X = Cl, Br, Y = F, Cl, Br) and the Related $PF_2(NMe_2)$ and $P(NMe_2)_3$ Compounds; Multinuclear NMR Spectroscopy and the X-ray Single Crystal Structures of [$RuBr_2(\eta^6$-cymene$)(PF_3)$], [$RuBr_2(\eta^6$-cymene$)(PF_2\{NMe_2\})$], and [$RuI_2(\eta^6$-cymene$)(P\{NMe_2\}_3)$]*, Inorganic Chemistry **47**, 9279 (2008)

335. A. Michalak, M. Mitoraj, and T. Ziegler, *Bond Orbitals from Chemical Valence Theory*, Journal of Physical Chemistry A **112**, 1933 (2008)

336. E. Clementi, Carla Roetti, *Roothaan-Hartree-Fock atomic wavefunctions: Basis functions and their coefficients for ground and certain excited states of neutral and ionized atoms, Z≤54*, Atomic Data and Nuclear Data Tables **14**, 177 (1974)

337. A.D. McLean, R.S. McLean, *Roothaan-Hartree-Fock atomic wave functions Slater basis-set expansions for Z = 55-92*, Atomic Data and Nuclear Data Tables **26**, 197 (1981)

338. J.G. Snijders, P. Vernooijs, E.J. Baerends, *Roothaan-Hartree-Fock-Slater atomic wave functions: Single-zeta, double-zeta, and extended Slater-type basis sets for $_{87}Fr$-$_{103}Lr$*, Atomic Data and Nuclear Data Tables **26**, 483 (1981)

339. J. Autschbach and B. Pritchard, *Calculation of molecular g-tensors using the zeroth-order regular approximation and density functional theory: expectation value versus linear response approaches* , Theoretical Chemistry Accounts **129**, 453 (2011)

340. J. Autschbach, S. Patchkovskii, and B. Pritchard, *Calculation of Hyperfine Tensors and Paramagnetic NMR Shifts Using the Relativistic Zeroth-Order Regular Approximation and Density Functional Theory*, Journal of Chemical Theory and Computation **7**, 2175 (2011)

341. S. Moon, and S. Patchkovskii, *First-principles calculations of paramagnetic NMR shifts*, in Calculation of NMR and EPR parameters, ISBN13: 9783527307791, M. Kaupp, M. Bühl, V.G. Malkin, Editors, (Wiley, Weinheim, 2004).

342. P. Hrobárik, Ro. Reviakine, A.V. Arbuznikov, O.L. Malkina, V.G. Malkin, F.H. Köhler, and M. Kaupp, *Density functional calculations of NMR shielding tensors for paramagnetic systems with arbitrary spin multiplicity: Validation on 3d metallocenes* , Journal of Chemical Physics **126**, 024107 (2007)

343. J. Li, M.R. Nelson, C.Y. Peng, D. Bashford, and L. Noodleman, *Incorporating Protein Environments in Density Functional Theory: A Self-Consistent Reaction Field Calculation of Redox Potentials of [2Fe2S] Clusters in Ferredoxin and Phthalate Dioxygenase Reductase*, Journal of Physical Chemistry A **102**, 6311 (1998)

344. T. Liu, W.-G Han, F. Himo, G.M. Ullmann, D. Bashford, A. Toutchkine, K.M. Hahn, and L. Noodleman, *Density Functional Vertical Self-Consistent Reaction Field Theory for Solvatochromism Studies of Solvent-Sensitive Dyes*, Journal of Physical Chemistry A **108**, 3545 (2004)

345. W.-G. Han, T. Liu, F. Himo, A. Toutchkine, D. Bashford, K.M. Hahn, L. Noodleman, *A Theoretical Study of the UV/Visible Absorption and Emission Solvatochromic Properties of Solvent-Sensitive Dyes*, ChemPhysChem **4**, 1084 (2003)

346. A. Kovyrshin, J. Neugebauer, *State-selective optimization of local excited electronic states in extended systems*, Journal of Chemical Physics **133**, 174114 (2010)

347. M. Swart, E. Rösler, and F. M. Bickelhaupt, *Proton affinities of maingroup-element hydrides and noble gases: Trends across the periodic table, structural effects, and DFT validation*, Journal of Computational Chemistry **27**, 1486 (2006)

348. M. Swart, and F. M. Bickelhaupt, *Proton Affinities of Anionic Bases: Trends Across the Periodic Table, Structural Effects, and DFT Validation*, Journal of Chemical Theory and Computation **2**, 281 (2006).

349. A. Kovalenko and F. Hirata, *Potentials of mean force of simple ions in ambient aqueous solution. II. Solvation structure from the three-dimensional reference interaction site model approach, and comparison with simulations*a, Journal of Chemical Physics **112**, 10403 (2000)

350. M. Seth, G. Mazur, and T. Ziegler, *Time-dependent density functional theory gradients in the Amsterdam density functional package: geometry optimizations of spin-flip excitations*, Theoretical Chemistry Accounts **129**, 331 (2011)

351.S.Y. Quek, L. Venkataraman, H.J. Choi, S.G. Louie, M.S. Hybertsen and J.B. Neaton, *mine-Gold Linked Single-Molecule Circuits: Experiment and Theory*, Nano Letters **7**, 3477 (2007)

352. M. Pavanello and J. Neugebauer, *Modelling charge transfer reactions with the frozen density embedding formalism*, Journal of Chemical Physics **135**, 234103 (2011)

353. M. Pavanello, T. Van Voorhis, L. Visscher, and J. Neugebauer, *An accurate and linear-scaling method for calculating charge-transfer excitation energies and diabatic couplings*, Journal of Chemical Physics **138**, 054101 (2013)

354. U. Ekström, L. Visscher, R. Bast, A.J. Thorvaldsen, and K. Ruud, *Arbitrary-Order Density Functional Response Theory from Automatic Differentiation*, Journal of Chemical Theory and Computation **6**, 1971 (2010)

355. M. Seth and T. Ziegler, *Range-Separated Exchange Functionals with Slater-Type Functions*, Journal of Chemical Theory and Computation **8**, 901 (2012)

356. E.R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García A.J. Cohen, and W. Yang, *Revealing Non-Covalent Interactions*, Journal of the American Chemical Society **132**, 6498 (2010)

357. J. Contreras-García E.R. Johnson, S. Keinan, R. Chaudret, J-P. Piquemal, D.N. Beratan, and W. Yang, *NCIPLOT: A Program for Plotting Noncovalent Interaction Regions*, Journal of Chemical Theory and Computation **7**, 625 (2011)

358. P. de Silva, J. Korchowiec, T.A. Wesolowski, *Revealing the Bonding Pattern from the Molecular Electron Density Using Single Exponential Decay Detector: An Orbital-Free Alternative to the Electron Localization Function*, ChemPhysChem **13**, 3462 (2012)

359. R. De Francesco, M. Stener, and G. Fronzoni, *Theoretical Study of Near-Edge X-ray Absorption Fine Structure Spectra of Metal Phthalocyanines at C and N K-Edges*, Journal of Physical Chemistry A, **116** 2285 (2012)

360. B. Delley, *An all-electron numerical method for solving the local density functional for polyatomic molecules*, Journal of Chemical Physics **92**, 508 (1992)

361. A.D. Becke, *A multicenter numerical integration scheme for polyatomic molecules*, Journal of Chemical Physics **88**, 2547 (1988)

362. J.L. Payton, S.M. Morton, Justin E. Moore, and Lasse Jensen, *A discrete interaction model/quantum mechanical method for simulating surface-enhanced Raman spectroscopy*, Journal of Chemical Physics **136**, 214103 (2012)

363. J.I. Rodríguez, *An Efficient Method for Computing the QTAIM Topology of a Scalar Field: The Electron Density Case*, Journal of Computational Chemistry **34**, 681 (2013)

364. C. König, N. Schlüter. J. Neugebauer, *Direct Determination of Exciton Couplings from Subsystem TDDFT within the Tamm-Dancoff Approximation*, Journal of Chemical Physics **138**, 034104 (2013)

365. C.J.O. Verzijl, J.S. Seldenthuis, and J.M. Thijssen, *Applicability of the wide-band limit in DFT-based molecular transport calculations*, Journal of Chemical Physics **138**, 094102 (2013)

366. H. Kim, J.-M. Choi, W.A. Goddard, *Universal Correction of Density Functional Theory to Include London Dispersion (up to Lr, Element 103)*, Journal of Physical Chemistry Letters **3**, 360 (2012)

367. M.Swart, *A new family of hybrid density functionals*, Chemical Physics Letters **580**, 166 (2013)

368. J. Autschbach, C.D. Igna, T. Ziegler, *A theoretical investigation of the apparently irregular behavior of Pt-Pt spin-spin coupling constants* Journal of the American Chemical Society **125**, 1028 (2003)

369. B.L. Guennic, K. Matsumoto, J. Autschbach, *On the NMR properties of platinum thallium bonded complexes: Analysis of relativistic density functional theory results*, Magnetic Resonance in Chemistry **42**, S99 (2004)

370. J. Khandogin, T. Ziegler, *A density functional study of nuclear magnetic resonance spin-spin coupling constants in transition metal systems*, Spectrochimica Acta Part A **55**, 607 (1999)

371. N.F. Ramsey, *Electron Coupled Interactions between Nuclear Spins in Molecules*, Physical Review **91**, 303 (1953)

372. R.M. Dickson, T. Ziegler, *NMR Spin-Spin Coupling Constants from Density Functional Theory with Slater-Type Basis Functions*, Journal of Physical Chemistry **100**, 5286 (1996)

373. D. L. Bryce, R. Wasylishen, *Indirect Nuclear Spin-Spin Coupling Tensors in Diatomic Molecules: A Comparison of Results Obtained by Experiment and First Principles Calculations*, Journal of the American Chemical Society **122**, 3197 (2000)

374. G. Schreckenbach, S. K. Wolff, T. Ziegler, *Covering the Entire Periodic Table: Relativistic Density Functional Calculations of NMR Chemical Shifts in Diamagnetic Actinide Compounds*, in Modeling NMR chemical shifts, ACS Symposium Series, Vol 732, J.C. Facelli, A.C. de Dios, Editors (American Chemical Society, Washington DC, 1999), Chapter 7

375. M. Franchini, P.H.T. Philipsen, L. Visscher, *The Becke Fuzzy Cells Integration Scheme in the Amsterdam Density Functional Program Suite*, Journal of Computational Chemistry **34**, 1818 (2013).

376. A. Tkatchenko, R.A. DiStasio Jr., R. Car, M. Scheffler *Accurate and Efficient Method for Many-Body van der Waals Interactions*, Physical Review Letters **108**, 236402 (2012)

377. A. Ambrosetti, A.M. Reilly, Robert A. DiStasio Jr., A. Tkatchenko, *Long-range correlation energy calculated from coupled atomic response functions*, Journal of Chemical Physics **140**, 18A508 (2014)

378. A.V. Marenich, S.V. Jerome, C.J. Cramer, D.G. Truhlar, *Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases*, Journal of Chemical Theory and Computation **8**, 527 (2012)

379. M. Franchini, P.H.T. Philipsen, E. van Lenthe, L. Visscher, *Accurate Coulomb Potentials for Periodic and Molecular Systems through Density Fitting*, Journal of Chemical Theory and Computation **10**, 1994 (2014)

380. A.D. Becke, R.M. Dickson, *Numerical solution of Poisson's equation in polyatomic molecules*, Journal of Chemical Physics **89**, 2993 (1988)

381. M.-C. Kim, E. Sim, and K. Burke, *Understanding and Reducing Errors in Density Functional Calculations*, Physical Review Letters **111**, 2073003 (2013)

403. J. Autschbach and T. Ziegler, *Solvent Effects on Heavy Atom Nuclear Spin.Spin Coupling Constants: A Theoretical Study of Hg.C and Pt.P Couplings*, Journal of the American Chemical Society **123**, 3341 (2001)

404. J. Autschbach and T. Ziegler, *A Theoretical Investigation of the Remarkable Nuclear Spin.Spin Coupling Pattern in [(NC)$_5$Pt-Tl(CN)]*, Journal of the American Chemical Society **123**, 5320 (2001)

414. A. van der Avoird, P.E.S. Wormer, F. Mulder, R.M. Berns, Topics in Current Chemistry **93**, 1 (1980)

415. C.J. Pickard and F. Mauri, *First-Principles Theory of the EPR g Tensor in Solids: Defects in Quartz*, Physical Review Letters **88**, 86403 (2002)

416. S. Patchkovskii and G. Schreckenbach in Calculation of NMR and EPR parameters, ISBN13: 9783527307791, M. Kaupp, M. Bühl, V.G. Malkin, Editors, (Wiley, Weinheim, 2004).

417. S. Patchkovskii, R.S. Strong, C.J. Pickard and S. Un, *Gauge invariance of the spin-other-orbit contribution to the g-tensors of electron paramagnetic resonance*, Journal of Chemical Physics **122**, 214101 (2005)

418. J.S. Seldenthuis, H.S.J. van der Zant, M.A. Ratner and J.M. Thijssen, *Vibrational Excitations in Weakly Coupled Single-Molecule Junctions: A Computational Analysis*, ACS Nano 2, 1445 (2008)

419. G.M. Sando and K.G. Spears, *Ab Initio Computation of the Duschinsky Mixing of Vibrations and Nonlinear Effects*, Journal of Physical Chemistry A **105**, 5326 (2001)

420. P.T. Ruhoff and M.A. Ratner, *Algorithms for computing Franck-Condon overlap integrals*, International Journal of Quantum Chemistry **77**, 383 (2000)

421. J. Poater, E. van Lenthe and E.J. Baerends, *Nuclear magnetic resonance chemical shifts with the statistical average of orbital-dependent model potentials in Kohn.Sham density functional theory*, Journal of Chemical Physics **118**, 8584 (2003)

422. M. Krykunov, T. Ziegler and E. van Lenthe, *Hybrid density functional calculations of nuclear magnetic shieldings using Slater-type orbitals and the zeroth-order regular approximation*, International Journal of Quantum Chemistry **109**, 1676 (2009)

423. M. Krykunov, T. Ziegler and E. van Lenthe, *Implementation of a hybrid DFT method for calculating NMR shieldings using Slater-type orbitals with spin-orbital coupling included. Applications to ^{187}Os, ^{195}Pt and ^{13}C in heavy metal complexes*, Journal of Physical Chemistry A **113**, 11495 (2009)

425. J. Autschbach, *Two-component relativistic hybrid density functional computations of nuclear spin-spin coupling tensors using Slater-type basis sets and density-fitting techniques*, Journal of Chemical Physics **129**, 094105 (2008), Erratum: Journal of Chemical Physics **130**, 209901 (2009)

426. D.L. Bryce and J. Autschbach, *Relativistic hybrid density functional calculations of indirect nuclear spin-spin coupling tensors . Comparison with experiment for diatomic alkali metal halides*, Canadian Journal of Chemistry **87**, 927 (2009)

429. J. Autschbach, *Analyzing NMR shielding tensors calculated with two-component relativistic methods using spin-free localized molecular orbitals*, Journal of Chemical Physics **128**, 164112 (2008)

432. Y.A. Wang, C.Y. Yam, Y.K. Chen, G. Chen, *Communication: Linear-expansion shooting techniques for accelerating self-consistent field convergence*, Journal of Chemical Physics **134**, 241103 (2011)

433. W.L. Jorgensen, J.D. Madura, C.J. Swenson, *Optimized intermolecular potential functions for liquid hydrocarbons,* Journal of the American Chemical Society **106**, 6638 (1984)

434. A.E. Kobryn, A. Kovalenko, *Molecular theory of hydrodynamic boundary conditions in nanofluidics,* Journal of Chemical Physics **129**, 134701 (2008)

435. O. Acevedo, W.L. Jorgensen, *Influence of Inter- and Intramolecular Hydrogen Bonding on Kemp Decarboxylations from QM/MM Simulations,* Journal of the American Chemical Society **127**, 8829 (2005)

436. S. Grimme, S. Ehrlich, and L. Goerigk, *Effect of the Damping Function in Dispersion Corrected Density Functional Theory*, Journal of Computational Chemistry **32**, 1457 (2011).

437. Ph. Haas, F. Tran, P. Blaha, and K.H. Schwartz, *Construction of an optimal GGA functional for molecules and solids*, Physical Review B**83**, 205117 (2011).

438. J.P. Perdew, A. Ruzsinszky, G.I. Csonka, L.A. Constantin, and J. Sun, *Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry*, Physical Review Letters **103**, 026403 (2009).

Keywords

Index