

Computing for Science (CFS) Ltd.,
CCLRC Daresbury Laboratory.

Generalised Atomic and Molecular Electronic Structure System

G A M E S S - U K

USER'S GUIDE and REFERENCE MANUAL

Version 8.0 June 2008

PART 8. WAVEFUNCTION ANALYSIS and MODELS for SOLVATION

M.F. Guest, J. Kendrick, J.H. van Lenthe, P. Sherwood and A.H. de Vries

Copyright (c) 1993-2008 Computing for Science Ltd.

This document may be freely reproduced provided that it is reproduced
unaltered and in its entirety.

Contents

1 Introduction	1
2 Analysis Modules - Introduction	1
3 One-electron Properties	2
3.1 PROPERTY	2
3.2 CENTRES	2
3.3 NUCLIDIC	2
4 Simplified Property Specification	3
4.1 SCF Calculations	3
4.2 UHF Calculations	4

4.3	MP2 Calculations	5
4.4	CI Calculations	6
5	Localised Molecular Orbitals	7
5.1	LOCAL	8
6	Graphical Analysis	10
6.1	Introduction	10
6.2	Grid Definition Directives	11
6.3	Grid Definition - GDEF	11
6.4	Grid Definition - TITLE	11
6.5	Grid Definition - TYPE	11
6.5.1	TYPE 2D or TYPE 3D	12
6.5.2	TYPE SPHERE	12
6.5.3	TYPE CARDS ncards	12
6.5.4	TYPE CONTOUR value	12
6.5.5	TYPE WRAP value	12
6.6	Grid Definition - ORIG	12
6.7	Grid Definition - X and Y	13
6.8	Grid Definition - SIZE	13
6.9	Grid Definition - POINTS	13
6.10	Grid Definition - SECTION	13
6.11	Directives Requesting Data Calculation	13
6.12	Data Calculation - CALC	13
6.13	Data Calculation - TITLE	14
6.14	Data Calculation - TYPE	14
6.14.1	MO and GRAD MO	16
6.14.2	COMB	16
6.14.3	Van der Waals functions	16
6.15	Data Calculation - OCCDEF	17
6.16	Data Calculation - CONFIG	18
6.17	Data Calculation - SFAC	20
6.18	Data Calculation - RADII	20
6.19	Data Calculation - SECTION	20

6.20	Data Calculation - RESTORE	20
6.21	Data Calculation - SURF	20
6.21.1	Example Potential on an isodensity grid	21
6.22	Plot Requests	22
6.23	Plot Requests - PLOT	22
6.24	Plot Requests - TITLE	23
6.25	Plot Requests - TYPE	23
6.26	Plot Requests - CONT	23
6.27	Plot Requests - VIEW	23
6.28	Plot Requests - SCALE	24
6.29	Termination of GRAPHICS Input	25
7	Potential Derived Charges	25
8	Mulliken Analysis	26
8.1	MULLIK	26
9	Distributed Multipole Analysis	28
9.1	ADD	28
9.2	DELETE	28
9.3	RADIUS	28
9.4	LIMIT	29
9.5	SHIFT	29
9.6	LINEAR	29
9.7	GAUGE	29
9.8	NONUCLEAR	29
10	The DRF Model for Solvation	30
11	Total energy in the DRF model	31
12	Features of the DRF Model	32
12.1	Grouping of External Points	32
12.2	Damping Functions	33
12.3	CHARMM Model Repulsion	33
12.4	Estimate of the Dispersion Interaction	33

12.5 Expansion of Fields	34
12.6 Dielectric Response	35
13 DRF Directives	35
13.1 FIELD	36
13.2 DRFOUT	36
13.3 UNITS	36
13.4 GAMDRF	37
13.5 INCLPOL	37
13.6 EXPANDCM	37
13.7 ASSIGN	38
13.8 GROUPING	38
13.9 DSTGRP	39
13.10DSTMAX	39
13.11AGRPE	40
13.12AGRPM	40
13.13AGRPC	40
13.14DAMPING	41
13.15CLASDISP	41
13.16QMRADI	42
13.17CLASRADI	42
13.18EXTERNAL	43
13.19DIELECTRIC	47
13.19.1 SURFACE	48
13.19.2 SRADIUS	48
13.19.3 BEMLEV	48
13.19.4 JUFFER	49
13.19.5 CONNOLLY	49
13.19.6 SOLVENT	50
13.19.7 DIELTYP	50
13.19.8 DIELOUT	51
13.19.9 EPSSTAT	51
13.19.10EPSOPT	51
13.19.11KAPPAS	52

13.19.1 KAPPAO	52
13.20 CLASCLAS	52
13.21 HBOND	52
13.22 DRFTWOEL	53
14 DRF Output—Analysis of (D)RF Energies	53

1 Introduction

In the first part of this chapter we provide a description of the analysis options and associated data input available within GAMESS-UK, including (i) the calculation of 1-electron properties and localised molecular orbitals, (ii) the graphical analysis of wavefunctions by the calculation of charge densities, molecular orbitals, atom difference densities and electrostatic potentials on a grid of points, (iii) the calculation of potential derived charges with electrostatic potential data calculated using the graphics module to generate least-squares fitted point charges at the nuclei, and (iv) performing both Mulliken and Distributed Multipole Analyses.

We then describe the capabilities and data input associated with the (Direct) Reaction Field (DRF) model for solvation. This model, developed at the University of Groningen [1, 2], is an embedding technique enabling the computation of the interaction between a quantum-mechanically described molecule and its classically described surroundings.

2 Analysis Modules - Introduction

The analysis modules of GAMESS-UK are requested under control of the RUNTYPE ANALYSE directive. Note that at present only one mode of analysis may be carried out in a given step, when it is assumed that the vectors to be analysed are resident in the section nominated on the VECTORS directive. The following points on eigenvector specification should be noted:

- While the SCF modules now support default eigenvector section usage (removing the need for explicit section specification on both the VECTORS and ENTER directives in such jobs), the user is strongly advised to specify the section containing the eigenvectors to be analysed under RUNTYPE ANALYSE.
- With the exception of Localised Orbital generation, the analysis routines do not output a separate set of eigenvectors, so that no section specification is required on the ENTER directive. In such cases the following sequence is typical of that required:

```
VECTORS 1  
ENTER
```

with the eigenvectors to be analysed resident in section 1. In the case of localised orbital generation, the sequence

```
VECTORS 1  
ENTER 10
```

will act to route the localised orbitals to section 10 of the Dumpfile.

Each mode of analysis has an associated set of sub-directives, which are described below.

3 One-electron Properties

This section deals with the data input used to drive the 1-electron properties module. Note that the module can at present only run with basis sets comprising s,p and d basis functions (see however the simplified property specification described below).

3.1 PROPERTY

This directive is used to specify which molecular 1-electron properties are to be computed. The first data line consists of the character string PROPERTY in the first data field. Subsequent data lines, the 'property definition' lines, are read to variables NPROP,TAGA,ISECT using format (I,A,I). NPROP is a code number of the property to be computed (see Table 4 of Part 2) TAGA should be set to the TAG of one of the centres defined in in the ZMATRIX or GEOMETRY directive, or in a CENTRES directive. ISECT is the section number of the Dumpfile where the computed property integrals are to be stored. If ISECT is omitted, the integrals are not stored on the Dumpfile. The final data line consists of the character string END in the first data field. Only 100 property lines may be presented in any one job.

3.2 CENTRES

This directive permits the specification of additional (non-nuclear) centres at which the properties are to be evaluated. The first data line consists of the character string CENTRES in the first data field. Subsequent data lines are read to variables X,Y,Z,TAG using format (3F,A), where X,Y and Z are the Cartesian co-ordinates of an additional centre (in atomic units). TAG is a name (up to 8 non-blank characters) by which the centre will subsequently be known. TAG may be omitted, when the system will supply an ordinal default. The final line consists of the character string END in the first data field.

Example

```
CENTRES
-1.0 1.0 1.0 ADDC
END
```

3.3 NUCLIDIC

This directive is used to re-define the nuclidic mass by the program, which by default corresponds to the most abundant isotope. The first data line consists of the character string NUCLIDIC in the first data field. Subsequent lines are read to variables TAGB,CENMAS using format (A,F). TAGB should be set to the TAG of a previously defined atomic centre, while CENMAS should be set to the value of the nuclidic mass to be used for this centre. The final line consists of the character string END in the first data field.

Example

```
NUCLIDE
OXYGEN 17
END
```

4 Simplified Property Specification

In the section above we have assumed that property evaluation is to be conducted under control of RUNTYPE ANALYSE, with explicit specification of the required one-electron properties. A simplified mechanism for property evaluation can be requested through presenting the data line

```
PROPERTY ATOMS
```

after RUNTYPE and SCFTYPE specification. This will result in the default wavefunction analysis conducted after RUNTYPE processing being augmented with the computation of certain one-electron properties. The following points should be noted:

- the properties evaluated include the electrostatic potential, electric field, electric field gradient, and electron density at each of the atomic centres, plus the dipole, second moment, quadrupole moment, third and octupole moments, at the computed centre of mass of the system under study. In addition the spin densities will also be computed in the case of open shell systems.
- this analysis, if requested, is available on completion of SCF, OPTIMIZE, OPTXYZ, SADDLE, and CI processing.
- in contrast to the detailed property evaluation performed under RUNTYPE ANALYSE control, this default evaluation may be conducted with basis sets comprising s,p,d and f basis functions.

We illustrate below a number of considerations that arise as a function of RUNTYPE and SCFTYPE processing when invoking this default property specification.

4.1 SCF Calculations

The following data sequence would be required to generate the default list of properties on completion of an SCF calculation of the formaldehyde molecule.

```
TITLE
H2CO - 3-21G BASIS - SCF + DEFAULT 1-E PROPERTIES
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
RUNTYPE SCF
PROPERTY ATOMS
ENTER
```


In this example the set of MOs to be used in the property evaluation will be retrieved from the default section of the Dumpfile written to by the closed-shell SCF module i.e. section 1.

4.2 UHF Calculations

A somewhat different approach may be required when computing the one-electron properties derived from a wavefunction with more than one set of MOs (e.g., a UHF wavefunction), or in cases where only the total density matrix, and not an associated set of MOs, is available (e.g., in an MP2 calculation). In both cases, the user may need to ensure that the associated set of spinfree natural orbitals and, where relevant SPIN natural orbitals, are generated by specification of the NATORB directive(s), used to route the NOs to a nominated section on the Dumpfile.

We illustrate this effect by first considering the data requirements when performing a UHF wavefunction. The following data sequence would be required when evaluating the properties based on a direct-UHF calculation, with the alpha- and beta-UHF MOs routed to the default sections 2 and 3 of the Dumpfile in the absence of explicit section specification on the ENTER directive.

```
TITLE
H2CO - 3A2 UHF PROPERTIES - 3-21G BASIS
MULT 3
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
SCFTYPE DIRECT UHF
PROPERTY ATOMS
ENTER
```

The same calculation may be performed based on the spinfree and spin natural orbitals of the UHF wavefunction; in this case the NATORB data lines will be used to route the spinfree and spin natural orbitals to sections 10 and 11 of the Dumpfile respectively, and these orbitals will be used in computing the 1-electron properties, thus:

```
TITLE
H2CO - 3A2 UHF NO-BASED PROPERTIES - 3-21G BASIS
MULT 3
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
SCFTYPE DIRECT UHF
PROPERTY ATOMS
NATORB 10
NATORB SPIN 11
ENTER
```

The following data sequence would be required if the user wished to compute the properties of the annihilated UHF wavefunction:

```
TITLE
H2CO - 3A2 annihilated UHF properties 3-21G BASIS
MULT 3
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
SCFTYPE DIRECT UHF
PROPERTY ATOMS
NATORB 10 ANNIHILATE
NATORB SPIN 11 ANNIHILATE
ENTER
```

Note again that the NOs of the UHF and AUHF wave function are in fact identical, the only difference lying in the occupation numbers.

4.3 MP2 Calculations

Now let us consider the data requirements when computing properties at the optimum geometry derived from an MP2 calculation.

```
TITLE
H2CO - X1A1 - MP2 DZ BASIS - PROPERTIES
ZMATRIX ANGSTROM
C
O 1 CO
H 1 CH 2 HCO
H 1 CH 2 HCO 3 180.0
VARIABLES
CO 1.203
CH 1.099
HCO 121.8
END
BASIS DZ
RUNTYPE OPTIMISE
PROPERTY ATOMS
SCFTYPE MP2
NATORB 20
ENTER
```

Having generated the MP2 optimised geometry, the spinfree natural orbitals will be routed to section 20 on the Dumpfile, and used in the subsequent properties calculation.

4.4 CI Calculations

Computing the default set of one-electron properties at completion of CI processing may be readily accomplished through the addition of the PROPERTY ATOMS data line. Note that any such calculation requires the NATORB directive to specify the routing of the spinfree, and where relevant, the spin NOS to specified sections on the Dumpfile. Note also that property evaluation under PROPERTY ATOMS control is only available for Direct-CI calculations (and not Full-CI or coupled cluster calculations).

```
TITLE
H2CO - 3-21G CISD DCI + PROPERTIES CALCULATION
SUPER OFF NOSYM
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
RUNTYPE CI
PROPERTY ATOMS
DIRECT 16 8 14
CONF
2 2 2 2 2 2 2
NATORB 10 0 PRINT
ENTER
```

To further illustrate property evaluation for CI wavefunctions, let us consider a CI calculation on the $^3A^2$ state of H_2CO . First, the data for the open-shell SCF calculation;

```
TITLE
H2CO - DZ BASIS - 3A2 GRHF TOTAL ENERGY = -113.73954029 AU
MULT 3
SUPER OFF NOSYM
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
BASIS DZ
OPEN 1 1 1 1
ENTER
```

Having generated the SCF wavefunction, the following data sequence would be used for a single reference CI calculation; routing the spinfree and spin natural orbitals to sections 10 and 11 of the Dumpfile will permit subsequent property generation, requested by presenting the data line PROPERTY ATOMS.

```
RESTART NEW
TITLE
```

```

H2CO - DZ BASIS - 3A2 CISD DIRECT-CI -113.934177537 AU
MULT 3
SUPER OFF NOSYM
BYPASS SCF
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
BASIS DZ
RUNTYPE CI
PROPERTY ATOMS
OPEN 1 1 1 1
DIRECT 16 9 15
SPIN 3
CONF
2 2 2 2 2 2 1 1
NATORB 10 11 PRINT
ENTER

```

Note that properties could also have been calculated after the CI job by specifying the appropriate natural orbitals under RUNTYPE ANALYSE. The data below would compute the isotropic ESR coupling constants (property index 19) at carbon, oxygen and hydrogen, where the spin NOS are nominated on the VECTORS line.

```

RESTART NEW
TITLE
H2CO - DZ - 3A2 UHF SPIN DENSITIES
MULT 3
ZMATRIX ANGSTROM
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
BASIS DZ
RUNTYPE ANALYSIS
PROPERTY
19 C
19 O
19 H
END
VECTORS 11
ENTER

```

5 Localised Molecular Orbitals

The purpose of this module is the localisation of molecular orbitals according to either:

- the criterion of Foster and Boys [3];

- the overlap-based criterion due to Pipek and Mezey [4].

The particular technique to be employed, together with the specification of the orbitals involved, is requested by presenting the LOCAL directive:

5.1 LOCAL

This directive, which may comprise one or more data lines, is used to define those molecular orbitals which take part in the localisation process, and the localisation technique to be employed. The first data line comprises one or more data fields;

- The first field consists of the character string LOCAL;
- If specified, the second data field may be used to nominate the localisation technique to be employed, and comprises the character string BOYS (for the Foster-Boys technique) or OVERLAP (for the Pipek-Mezey overlap-based method). In the absence of this data field, the default Foster-Boys method will be used.
- If specified, the third data field may be used to define a default set of orbitals to be incorporated in the localisation process. Presenting the data field DEFAULT instructs the program to consider just the set of valence molecular orbitals, omitting all core orbitals from the process. The presence of the DEFAULT data character string signals the termination of LOCAL input.

An alternative to utilising the DEFAULT option above is to include additional data lines that explicitly nominate the molecular orbitals to be incorporated. Such data lines are read to an array (LMO(I),I=1,NACT) using free I-format. When specifying such an orbital set, the last data field presented should be the character string END.

Example 1

The single data line

```
LOCAL OVERLAP
 1 2 3 4 5 6 7 8 9 END
```

requests use of the overlap-based localisation technique due to Pipek and Mezey, with molecular orbitals 1-9 to be included.

Example 2

The single data line

```
LOCAL OVERLAP DEFAULT
```

requests use of the overlap-based localisation technique due to Pipek and Mezey, with only the set of valence molecular orbitals to be included.

Example 3

```
LOCAL  
2 3 4 5 6 7 9 END
```

Declares molecular orbitals 2 to 7 inclusive, plus molecular orbital 9 to be active in the Foster-Boys localisation process.

Example 4

```
LOCAL  
2 3 4 5 6 7  
9  
END
```

This sequence has an equivalent effect to that of Example 3.

Example 5

```
LOCAL  
2 TO 7 9 END
```

The above sequence shows an abbreviated form of specifying the list of molecular orbitals, invoking the character string TO, to link together a sequence of consecutive numbered active molecular orbitals. This sequence is equivalent to examples 3 and 4.

Example 6

```
LOCAL  
2 TO 7 9 TO 14  
END
```

Note:-

During the localisation process, under control of RUNTYPE ANALYSE, symmetry adaptation is automatically switched off to enable orbitals of different irreducible representation to mix (although the total wavefunction remains, of course, a unitary transformation of the SCF wavefunction). Assuming these orbitals are to be used in a subsequent SCF or GVB calculation, via the VECTORS directive, then the ADAPT OFF data line *must* be presented in any such job that utilises the LMOs. Failure to provide such a line will probably lead to an error condition when restoring the vectors.

6 Graphical Analysis

6.1 Introduction

GAMESS–UK supports the graphical analysis of wavefunctions by the calculation of charge densities, molecular orbitals, atom difference densities and electrostatic potentials on a grid of points. Versions of GAMESS–UK which are linked to the GHOST graphical library allow graphical display of those datasets which are calculated on regular two-dimensional grids as contour or surface plots. In general, the data values will be written to the formatted punchfile (See Part 11), and analysed using other visualisation software.

The module is invoked under control of RUNTYPE ANALYSE by presence of the GRAPHICS directive. Unlike the graphical analysis code present in previous versions of GAMESS–UK the GRAPHICS keyword may only appear once for each specification of RUNTYPE ANALYSE.

Note that the original restrictions in running this module - to basis sets comprising only s,p and d basis functions - have been lifted in the calculation of both molecular densities and potentials, where the full range of s,p,d,f and g-functions may now be used.

The subsequent directives are presented in groups with the following functions

1. Those that define the grid of points (a group of directives initiated by GDEF)
2. Those that request data calculation on a grid (a group of directives initiated by CALC)
3. The SURF directive for generation of molecular surfaces (a combination of both grid definition and calculation)
4. The RESTORE directive to retrieve grid definitions or data from the dumpfile.
5. Those that generate graphical output (when available, initiated by PLOT).

Presentation of each of the directives which appear as groups are initiated by the specific directives GDEF, CALC and PLOT as given above. Directives within a group are terminated by the first directive of another group, by SURF or RESTORE, or by any other valid GAMESS–UK *Class 2* directive.

The ordering of groups is significant (but the ordering of directives within a group is not), as follows:

1. A CALC directive must be preceded by a directives to generate (or restore) a grid definition, which will be used for the calculation. An exception is a calculation of a combination grid (TYPE COMB) which must follow the definition of a calculation.
2. A PLOT directive must be preceded directives to generate or restore a (suitable) array of data values to be plotted.
3. A SURF directive must follow the generation or restoration of an array of scalar data points on a regular 3D grid.

Table 1: Keywords of the Grid TYPE Directive

keyword	Grid Type
2D	Planar Rectangular 2D grid
3D	Orthogonal 3D grid
SPHERE	Spherical grid
CARDS	User specified points
CONTOUR	Generate points on an isovalue surface
WRAP	Generate points on an iso-electron-density surface
ATOM	Place grid points at the nuclear positions

All directions and positions specified must be given in the molecular coordinate system, (after reorientation by the GAMESS–UK symmetry routines) in units of a.u.

At present, the number of grids, the number of calculations, the number of plot requests and the number of dumpfile restore operations are each limited to 10 (this total including requests made implicitly by SURF directives, see below). Users who need to generate more data may use multiple sets of directives starting with RUNTYPE ANALYSE.

6.2 Grid Definition Directives

6.3 Grid Definition - GDEF

```
GDEF
```

Grid definition mode is initiated by a GDEF directive. This may (optionally) be followed by a character string (8 characters or less) which will be used in the output to reference the grid. Grid definition mode is terminated by a GDEF, CALC, PLOT, REST, SURF or valid *Class 2* directive.

6.4 Grid Definition - TITLE

Provide a title for the grid.

```
TITLE
.. title string ...
```

6.5 Grid Definition - TYPE

The type of grid is specified by the TYPE directive. TYPE is followed by a keyword specifying the type of grid to be generated, and other data as required by the grid type. Valid keywords are given in Table 1.

The individual grid generation modes are described below.

6.5.1 TYPE 2D or TYPE 3D

Generate regular 2 or 3D grids. The origin (centre of the grid) is set using the ORIG directive, and the X and Y directives set the orientation. The number of points is set with the POINTS directive, and the edge lengths using SIZE (see below).

6.5.2 TYPE SPHERE

Generate a grid of points on the surface of a sphere of given radius. The centre (default 0.0 0.0 0.0) can be set using the ORIG directive. The number of points is set with the POINTS directive. The radius is set using the SIZE directive.

The SPHERE directive may optionally be followed by one of the keywords RAND or SYMM. RAND (which is the default) requests that random starting latitude values be used for each ring of points at a given longitude. SYMM requests that the generated grid preserve axial rotation symmetry, and must be followed by an integer specifying the order of rotation symmetry required.

6.5.3 TYPE CARDS ncards

Read a set of points from the input file. The number points (ncards) must be specified. The directive is followed by ncards data records containing the x,y, and z coordinates of a data point, in a.u., in the coordinate system of the molecule after reorientation by the GAMESS-UK symmetry analysis routines.

6.5.4 TYPE CONTOUR value

Generate points on an isovalue surface. The section number or calc-id refers to an array of data values calculated on a regular 3D grid, which are to be contoured.

6.5.5 TYPE WRAP value

Points are generated on an isodensity surface (without storage of a 3D array of density points. If this option is used, it must follow the specification of a regular 3D grid, which defines the volume and mesh density for the contouring.

6.6 Grid Definition - ORIG

The ORIG directive specifies the origin (ie centre) of the grid, it applies only to the grid types 2D, 3D and SPHERE.

```
ORIG x y z
```

6.7 Grid Definition - X and Y

Set the direction of the plot axes for 2D and 3D grid types. By default, the first plot axis (referred to as X) lies along the +x direction of the molecular coordinate system, and the second along +y.

```
X xx xy xz
Y yx yy yz
```

6.8 Grid Definition - SIZE

Set the grid size (edge length) for 2D and 3D grids, or the radius for spherical grids. If only one value is specified it will be used for all dimensions, but extra values may be provided to set the y and (for 3D) z axis lengths independently.

```
SIZE sx <sy <sz>>
```

6.9 Grid Definition - POINTS

Set the mesh density. For 2 and 3D regular grids this sets the number of points along the axes. For spherical grids the number of longitudinal and latitudinal divisions should be given as a single integer argument.

```
POINTS nx <ny <nz>>
```

6.10 Grid Definition - SECTION

Specify a section number (isect) on the dumpfile where the grid is to be stored.

```
SECTION isect
```

Storage of the grid is required in the following circumstances:

- the grid is to be restored later in the job, or in a later GAMESS–UK run
- the grid is irregular (ie a type other than 2D or 3D) and is required to appear on the punchfile along with data calculated on the grid.

6.11 Directives Requesting Data Calculation

6.12 Data Calculation - CALC

The computation of data on a predefined grid is requested by the CALC directive. All subsequent directives up to the next GDEF, CALC, PLOT, SURF, REST or other *Class 2* directive serve to specify details of the calculation.

```
CALC
```

Table 2: Keywords of the Data TYPE Directive

keyword	Data Type
MO	Molecular Orbital
DENS	electron density
ATOM	atom difference electron density
POTE	electrostatic potential
GRAD DENS	gradient of the charge density
GRAD MO	gradient of MO
COMB	linear combination with the previous calculation
VDW	Van der Waals function
LVDW	log Van der Waals function
GRAD VDW	gradient of Van der Waals function
GRAD LVDW	gradient of Van der Waals function (sic)

6.13 Data Calculation - TITLE

Specify a title for the data array

```
TITLE
... title ...
```

6.14 Data Calculation - TYPE

```
TYPE keyword
```

Specify the property to be calculated. Valid keywords are given in Table 2.

A more detailed description of the functions available is given below, together with information regarding any other input data required.

Electron Density Functions

In depicting the spatial characteristics of the density associated with one or more molecular orbitals, the program computes densities according to the formula:

$$\rho(r) = \sum_i OCC_i \Theta_i(r)^2 \quad (1)$$

where Θ_i denotes the i 'th molecular orbital and OCC_i its occupation number. Plots of the amplitude of a single orbital, i :

$$a_i(r) = OCC_i \Theta_i(r) \quad (2)$$

may also be generated.

Density Difference Function

The atomic density difference function is defined as:

$$\Delta\rho(r) = \rho_{mol}(r) - \sum_j^{Atoms} \rho_j(r) \quad (3)$$

The first term is the total electron density associated with a molecule. The second term represents the sum of the electron densities of the atoms which constitute the molecule, these being placed at the same positions they occupy in the molecule, but which are assumed to have undergone no interactions with each other, and have remained undistorted, as in the free state. The atomic density difference function provides an indication of the overall rearrangement of density which occurs when the atoms come together upon molecular formation. The program incorporates an atomic SCF module, so that calculations on the ground states of the component atoms are performed in line, with the basis set of each atom the same as that used in the parent molecule. These plots are known as 'atomic-difference' plots; the program is capable of generating such plots of molecular systems with component atoms up to, and including, zinc.

A more general form of the density difference function, the molecular density difference:

$$\Delta\rho(r) = \sum_i \rho_i(r) - \sum_j \rho_j(r) \quad (4)$$

is used in the construction of molecular-difference plots, which allows the user to display the density resulting from the addition or subtraction of up to 15 component density functions. Two examples of the value of such plots are:

- Illustrating the effect on a molecular charge distribution, resulting from an extension of the basis set, so that a typical plot would be constructed using the function:

$$\Delta\rho(r) = \rho_{mol}^{extended}(r) - \rho_{mol}^{minimal}(r) \quad (5)$$

- In depicting the rearrangement of electron density which occurs when the component ligands and metal atom of a transition metal complex come together to form the molecular system, so that for a complex MX, the following difference function:

$$\Delta\rho(r) = \rho_{complex}(r) - \rho_M(r) - \sum_i^n \rho_{X_i}(r) \quad (6)$$

would be constructed, this being the molecular analogue of the atomic density difference function.

Electrostatic Potential Function

The value of the electrostatic potential created by the electronic distribution and nuclear charge of a molecule, in the different regions of space surrounding it, provides information about possible sites involved in protonation or in reactions with electrophilic agents. The interaction energy between a molecular distribution and an external unitary positive charge at a given point i , is given by:

$$V(r_i) = \sum_{\alpha}^{nucl} \frac{Z_{\alpha}}{r_{i\alpha}} - \int d\tau_1 \frac{\rho(1)}{r_{1i}} \quad (7)$$

where Z_{α} is the nuclear charge of nucleus alpha and $\rho(1)$ is the first order density function. The program permits the construction of plots of electrostatic interaction energies based on the density distribution arising from wavefunctions constructed in Gaussian orbital basis sets.

6.14.1 MO and GRAD MO

The MO and GRAD MO keywords must be followed by an integer specifying the MO of interest.

6.14.2 COMB

The COMB keyword is followed by specification of the data to be combined with the current grid and a scale factor as follows;

```
TYPE COMB [lfn iblk] isec scale
```

If lfn and iblk are given, they determine the foreign dumpfile and start block on which the data is to be found. If they are omitted the current dumpfile is used. isec is the section number of the required data (as specified using the SECTION directive when the data was generated)

The grid is multiplied by scale before being added to the current grid. If more than two grids are to be combined. Only one TYPE COMB directive may be present in a given CALC group of directives.

6.14.3 Van der Waals functions

These functions are defined in such a way that contouring them leads to VdW surfaces. The gradient functions generate unit normals to the surface (ie not strict gradients, but normals to the surface which are adequate for lighting calculations). The internal van der Waals radii are taken from Nuffield Advanced Science "Book of Data", R. D Harrison (ed), 1988. Additional (or replacement) radii may be provided using the RADII directive (see below).

The two functions available are VDW, defined by

$$\max (r^{vdw}_i - r_i)$$

and LVDW, defined by

$$\max(-\ln(r_i / r_i^{vdw}))$$

where r_i is the distance from the test point to nucleus i , r_i^{vdw} the van der Waals radius of atom i , and the max function runs over all nuclei in the molecule.

Both functions have the property that the isovalue surface for value 0.0 is the van der Waals surface. When contoured at a positive value (v), the VDW function gives rise to a surface inside the van der Waals surface, corresponding to a surface with all radii reduced by v . Similarly negative contour heights may be used to generate surfaces with a constant added to all radii. Although not identical, these surfaces are similar to Connolly surfaces with a probe sphere of radius v . Contouring the LVDW function at non-zero values gives rise to surfaces that correspond to scaled van der Waals radii. Positive values v thus lead to surfaces inside the van de Waals radius.

NB. It is important to note that at the moment the use of the LVDW function to generate the surfaces with scaled radii is not fully tested - ie there are sets of radii, and/or choices of contour levels which will result in a surface that does not correspond to that generated from the scaled radii.

6.15 Data Calculation - OCCDEF

The purpose of this directive is to allow the user to define the occupation numbers for the molecular orbitals to be analysed. In the absence of the OCCDEF directive, the occupation numbers will be taken from the section of the Dumpfile specified on the VECTORS directive. The first data line contains the character string, OCCDEF, in the first data field. Following the directive initiator are the occupation definition lines. The first data field of such lines is read in F-format, and should contain a specified occupation number. Subsequent data fields are read in I-format. Let the value of an integer specified in such a field be j , the j 'th molecular orbital will be assigned the occupation number specified in the first data field of the line. The following:

```
2.0 1 2 3 4 5 7
```

comprises a valid occupation definition line. Such lines may be shortened, if a sequence of consecutive integers appear, by means of the character string, TO. Thus, the abbreviated form of the above line is:

```
2.0 1 TO 5 7
```

The occupation definition lines are specified until all the orbitals to be assigned a finite occupancy have been declared. A data line containing the text, END, in the first data field, must be specified to terminate the OCCDEF directive.

The following points must be noted:

- Any orbital omitted from the list specified on the occupation definition lines will be assigned zero occupancy, and thus will make no contribution to the grid of function values to be constructed.
- It is envisaged that the OCCDEF directive will not be required when generating grids of total electron density, atomic density difference, and interaction potentials. In these three cases, the occupation numbers should reflect the overall orbital occupancy in the molecule, and should be just the values calculated during the construction of the molecular orbitals and output to the Dumpfile.
- The OCCDEF directive should be used when analysis of the electron density, associated with a certain subset of orbitals, is required.

Example

```
OCCDEF
2.0 1 TO 5 7
END
```

The grid of values will be generated assuming the first five molecular orbitals, together with orbital 7, are doubly-occupied. All other orbitals will be assigned zero occupancy.

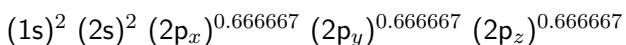
6.16 Data Calculation - CONFIG

The CONFIG directive is only applicable when generating a grid of atomic density difference, and may be used to specify the configuration to be used in computing the atomic density distribution, corresponding to the ground state of the atoms. In the absence of this directive, spherically symmetric atoms are chosen, with equal occupation of the degenerate open-shell orbitals.

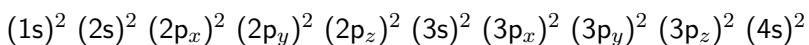
The user should note that applying a CONFIG specification will not change the configuration of the atom used in the atomic SCF calculation, but will result in modification of the occupations of the atomic orbitals just prior to the computation of the density. The computed density therefore does not correspond to a self-consistent atomic calculation.

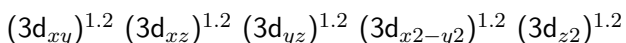
Example

The atomic configuration chosen, by default, for the carbon atom would be:



and for the iron atom (d^6s^2 high spin):





The CONFIG directive consists of three types of data line. The first line, the directive initiator, consists of the character string, CONFIG, in the first data field; the last line, the directive terminator, consists of the character string, END, in the first data field. Lines specified between the directive initiator and terminator, are the 'configuration definition' lines. If there are NAT atoms, whose configuration are to be specified, NAT 'configuration' lines are required. Each line consists of (NORB+1) data fields, where NORB is the total number of doubly-occupied or partially-occupied orbitals in the atom. The first data field is read to the variable ALAB, using format A, whilst the remainder of the data line should contain real numbers, read in F-format, to a vector (OCC(I), I=1,NORB). ALAB should be set to the label parameter of the nucleus, as specified by the corresponding 'nucleus definition' lines in the GEOMETRY or ZMATRIX directives. OCC(I) should be set to the occupation number of the i'th atomic orbital. The latter must be input in order of symmetry – s, p, d, etc., – with the partially-occupied orbitals preceded by the doubly-occupied orbitals, within each symmetry class.

Example 1

Suppose the user wishes the $(1s)^2 (2s)^2 (2p_x)^1 (2p_y)^1$ configuration for a carbon atom, which has been labelled as nucleus C1. Then the following configuration definition line should be specified:

```
C1 2.0 2.0 1.0 1.0 0.0
```

Example 2

To specify the configuration $(1s)^2 (2s)^2 (2p)^6 (3s)^2 (3p_z)^1$ for an aluminium atom, labelled AL by the ZMATRIX directive, the user must specify the following configuration definition line:

```
AL 2.0 2.0 2.0 2.0 2.0 2.0 0.0 0.0 1.0
```

where the first three 2.0 give the occupation of the s orbitals, and the remainder details the occupation of the p orbitals.

Example 3

To specify the configuration $(1s)^2 (2s)^2 (2p)^6 (3s)^2 (3p)^6 (4s)^2 (3d_{xy})^2 (3d_{xz})^2 (3d_{yz})^2$ for an iron atom, which has been labelled FE, the user should specify the following line:

```
FE 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 0.0 0.0
   < s orbitals > < p orbitals > < d orbitals >
```

Note that the occupation numbers of the three p orbitals should be input in the order (x,y,z), and those of the five d orbitals in the order: d_{xy} , d_{xz} , d_{yz} , $d_{x^2-y^2}$, d_{z^2}

6.17 Data Calculation - SFAC

The SFAC directive specifies a constant that will be used to multiply all calculated grid data values. The default scale factor is 1.0, except for electrostatic potential plots, in which case it is 627.707, (to convert the values to kcal/mol).

6.18 Data Calculation - RADII

The RADII directive allows the user to input a set of radius values for use in the van der Waals function calculation (see above). It is followed by a series of records, each specifying an atom label and the radius for all atoms with that label. The input is terminated with a record containing the string END. By default radii are expected in atomic units, angstroms may be used if the string ANGS is added on the RADII directive.

6.19 Data Calculation - SECTION

Specify a section number (isect) on the dumpfile where the calculated data is to be stored.

```
SECTION isect
```

Storage of the data is required in the following circumstances:

- the grid is to be written to the punchfile.
- the data is on a 3D regular grid and is to be used in a contour grid generation step.

6.20 Data Calculation - RESTORE

The RESTORE Directive is used to bring a grid or data from a dumpfile into memory. The data must have been written by a SECTION directive, (see above) and the first keyword on the directive, DATA or GRID, is used to specify the whether a grid definition or data is required. It is an error to specify DATA if the dumpfile section actually contains a grid definition, and *vice-versa*. The remaining data items are the lfn (eg ED3 for the current dumpfile), iblk is the starting block of the dumpfile (usually 1), and isec the dumpfile section as specified on the SECTION directive.

```
RESTORE [ GRID | DATA] lfn iblk isec
```

6.21 Data Calculation - SURF

The SURF directive is used to generate one or more isovalue surfaces from a 3D dataset, and calculate data at points on that surface. It is presented after directives requesting the calculation of the 3D dataset.

The syntax of the surf directive is as follows:

```
SURF type isec [imo] level1 [level2 ...]
```

where

- type is the property to be calculated. It may be one of the valid keywords of the calculation TYPE subdirective given in Table 2, and may be followed by an MO index where appropriate.
- isec is the a section number on the current dumpfile for output of the first grids generated. Subsequent data and grids will be saved to consecutive section numbers.
- level1 level2 .. are a series of values at which the isovalue surface is to be generated.

The result is the same as if the user had made a series of grid definition requests (of type CONTOUR) and data calculation requests. For each level requested, GAMESS–UK will generate a surface grid and write it to the dumpfile. If the 3D data being contoured is of type MO or DENS the gradient of the field at every point will be calculated and stored. The requested property will then be calculated for every point. When more than one level is required it is necessary to restore the 3D dataset prior to each contouring operation. This is performed automatically by the program, but a consequence of this, that the 3D data must be stored on the dumpfile, is the responsibility of the user. The user should note that there is a limit, noted above, to the number of of grid definitions, data calculations etc which may be requested for each invocation of RUNTYPE ANALYSE, and the totals include requests generated implicitly by SURFACE.

All grid definitions and data arrays are written to the dumpfile, with section numbers counting from the isec value given, in the order in which they are generated. The user is responsible for ensuring that any data required by other software is written to the punchfile. The grid definitions will be appear in the punchfile as part of the property (but not gradient) datasets and are not requested explicitly. The form of the punch directive required should be clear from the examples below.

6.21.1 Example Potential on an isodensity grid

The following GAMESS–UK input assumes that a job to calculate the closed-shell SCF wavefunction has already been completed

```
RESTART NEW
PUNCH GRID 161 162 164 165
TITLE
FORMALDEHYDE SURFACE
ZMAT ANGS
C
O 1 1.203
H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
BASIS DZ
```

```
RUNTYPE ANALYSE
GRAPHICS
GDEF
TITLE
3D GRID
TYPE 3D
POINTS 60
SIZE 8.0
CALC
TYPE DENS
SECTION 150
SURF POTE 160 0.02 0.04
VECTORS 1
ENTER
```

The result would be following job steps

1. A 3D grid, with edge length 8 a.u., centred on the origin is defined.
2. The density is calculated at all points on the grid and written to dumpfile section 150
3. A set of points on the surface of electron density 0.02 are generated and written to dumpfile section 160
4. The gradient of the density is calculated at each point, and written to dumpfile section 161
5. The potential is calculated at every point, and written to section 162
6. The 3D density grid is restored from the dumpfile
7. Steps 3,4, and 5 are repeated for a density value of 0.04, resulting in the new grid definition being written to section 163, that density gradient to section 164 and the potential to section 165.
8. The data from sections 161 162 164 and 165 are written to the punchfile. The grid definitions (from sections 160 and 163) also appear on the dumpfile, although not explicitly requested.

6.22 Plot Requests

Note that Users of versions of GAMESS-UK which support GHOST graphics can generate contour plots (TYPE CONT) and relief plots (TYPE SURF). Otherwise only lineprinter plots are available.

6.23 Plot Requests - PLOT

Plots are requested by a series of directives initiated by the PLOT directive.

Table 3: Keywords of the Plot TYPE Directive

keyword	Data Type
LINE	lineprinter plot
CONT	contour plot
SURF	surface plot

6.24 Plot Requests - TITLE

Provide a title for the plot

```
TITLE
... title...
```

6.25 Plot Requests - TYPE

Specify the type of graphical output required, Valid directives are given in Table 3.

6.26 Plot Requests - CONT

Set the contour heights for lineprinter and contour options. The directive is followed by one or more records containing the values required, terminated by end.

```
CONT
cont(1) cont(2) ...
...
END
```

If the CONT directive is omitted, the default set of contour values, given in Table 4 will be used.

Contour values of POTE, AMPLITUDE, ATOM, and DIFF setting also have corresponding negative values.

6.27 Plot Requests - VIEW

The VIEW directive is only relevant when generating perspective plots, and may be used to specify the angle of view and viewing distance. The grid of electron densities or potentials define the z-values of the surface on a two-dimensional (x,y) grid of points covering the specified area of the molecular plane under investigation.

The directive consists of a single data line, read to variables TEXT, THETA V, THETA H, and DIST, using format (A,3F).

Table 4: Default Contour Values in Graphical Analysis

GTYPE setting	POTE	DENS	AMPLITUDE	ATOM or DIFF
	210.0	64.7837	1.0	0.8691
	180.0	16.1959	0.5	0.43455
	150.0	4.0490	0.25	0.21727
	120.0	1.0122	0.125	0.10864
	90.0	0.5061	0.0625	0.05432
	75.0	0.2531	0.03125	0.02716
	60.0	0.1265	0.01562	0.01358
	40.0	0.0633	0.00781	0.00697
	20.0	0.0316	0.00391	0.00339
	10.0	0.0158	0.00195	0.00170
	5.0	0.0079	0.00098	0.00085
	2.0	0.0040	0.0049	0.00042
	0.0	0.0020	0.0	0.0
		0.0010		

- TEXT should be set to the character string, VIEW.
- THETA V specifies the elevation of the view axis above the horizontal base-plane in degrees
- THETA H specifies the rotation of the vertical axis through the centre of the grid in a clockwise direction, in degrees . For $0.0 < \text{THETA H} < 180.0$, the surface appears to rotate in a clockwise direction as THETA H increases.
- DIST specifies the viewing distance in Bohr.

If the VIEW directive is omitted, THETA V (thetav) and THETA H (thetah) will be given the value 30, and DIST will be set to the value of SIZE, specified by the PLANE directive.

Example

To view the surface edge-on along the y-axis, from a distance of 10 Bohr, the following data line should be specified:

```
VIEW 0.0 0.0 10.0
```

6.28 Plot Requests - SCALE

This directive is only relevant when generating perspective plots, and may be used to normalize the stereo-graphic projection to certain values of electron density or potential function. The directive consists of a single data line, read to variables TEXT, SCAMAX, SCAMIN, and FACTOR, using format (A,4F).

- TEXT should be set to the character string, SCALE.
- SCAMAX,FACTOR the grid of values to be plotted is scanned to detect all local maxima, which will appear as peaks on the projected plot, and thus, to determine VMAX, the value of the greatest maximum $< \text{SCAMAX}$. In the event that no such maximum is detected, VMAX will be set to SCAMAX. All maxima with a greater value than $(\text{VMAX} * \text{FACTOR})$ will appear as beheaded peaks on the final plot.
- SCAMIN,FACTOR the grid of values is scanned to detect all local minima, which will appear as troughs on the projected plot, and thus, to determine VMIN, the value of the lowest minimum $> \text{SCAMIN}$. In the event that no such minimum is detected, VMIN will be set to SCAMIN. All minima with a value less than $(\text{VMIN} * \text{FACTOR})$ will appear as beheaded troughs on the final plot.

The projected plot will be normalized to $(\text{VMAX} * \text{FACTOR} - \text{VMIN} * \text{FACTOR})$. If the SCALE directive is omitted, SCAMAX is set to 0.7, SCAMIN to -0.7, and FACTOR to 1.2.

6.29 Termination of GRAPHICS Input

Data input for Graphical Analysis is terminated by presenting a valid *Class 2* directive. This might typically be the VECTORS directive, instructing the analysis module as to the source of eigenvectors to be analysed.

7 Potential Derived Charges

The potential derived charges module uses electrostatic potential data calculated using the graphics module to generate least-squares fitted point charges at the nuclei.

The module is invoked by the POTFIT directive, under control of RUNTYPE ANALYSE.

An example of the input data is given below

```
TITLE
PDC CALCULATION
ZMAT ANGS
0
H 1 1.0
H 1 1.1 2 109.0
END
RUNTYPE SCF
ENTER
RUNTYPE ANALY
GRAPHICS
GDEF
TYPE 3D
POINTS 50
SIZE 6
SECTION 150
CALC
```

```
TYPE DENS
SECTION 151
TITLE
DENSITY ON 3D GRID
SURFACE POTE 170 0.02 0.04
VECTORS 1
ENTER
RUNTYPE ANALY
POTF 172 175 CHAR 0.0
VECTORS 1
ENTER
```

This job consists of three phases (each terminated by an ENTER directive)

- perform an SCF calculation, writing vectors to Dumpfile section 1
- generate the electron density on a 3D grid, and calculate the electrostatic potential at the 0.02 and 0.04 isodensity surfaces. This data is written to the dumpfile on section 172 and 175
- generate potential derived charges by fitting to the potential points from 2. The total charge is constrained to 0.0

In addition to the dumpfile sections for the potential data, the following keywords may appear on the POTFIT directive

- CHARGE *charge* constrain charge
- SYMMETRY constrain symmetry equivalent atoms to have the same charge
- DIPOLE *dx dy dz* constrain dipole (atomic units)
- CUTOFF *scale* exclude points closer than *scale* * the covalent radius from a nucleus.

8 Mulliken Analysis

The purpose of this module is to provide for an increased level of analysis of a given set of molecular orbitals. The default SCF options will typically provide a population analysis of the total SCF wavefunction. In some instances it is useful to probe the individual molecular orbitals, extracting quantities such as atom-atom overlap populations and sub-dividing the orbital electrons into s,p and d character (say) on the component atoms. Such an analysis may be requested under control of the MULLIK directive:

8.1 MULLIK

This directive consists of a single data line, and is used to define those molecular orbitals for which a detailed analysis is required, and to define the type of analysis to be performed.

The first data field consists of the character string MULLIK. One or two data fields may then be read in A-format to define the type of Mulliken analysis required. This analysis is performed by initially assigning basis functions to 'groups', and then performing the Mulliken analysis over these groups, rather than over individual basis functions. Two such groups are recognised by the program, and may be activated by appropriate keyword setting on the MULLIK directive. The groups, with appropriate keyword settings, are as follows:

- Specifying the keyword ATOM will assign all basis functions sited on a given atomic centre to the same group, with the group labelled by the centre name (based on that specified by the ZMATRIX or GEOMETRY directive). This will culminate in an atom-atom type of Mulliken analysis.
- Specifying the keyword ORBITAL will provide a more detailed analysis than above, with the atomic groups further classified into sub-groups of specific orbital (s, p, d or f) character.

Subsequent data fields are used to specify those molecular orbitals whose analysis is to be printed. If NMO is the total number of orbitals whose analysis is required, subsequent data fields should contain NMO integer numbers read in free-I format. An abbreviated form of this data specification allows the user to introduce the character string TO, read in free A-format, when specifying a sequence of consecutive orbitals. The last data field presented should be the character string END. Note that the integer and END data fields may be omitted, in which case only the Mulliken analysis of the total wavefunction will be printed.

The following notes may be helpful:

- Each molecular orbital is individually analysed as if it contained one electron. The occupation numbers (as retrieved from the Dumpfile section nominated on the VECTORS directive) are only used in the evaluation of the Mulliken populations of the total wavefunction.
- While all molecular orbitals regardless of occupation number, are analysed, the detailed analysis of a given orbital will only be printed if that orbital is referenced by the MULLIK directive.

Example 1

```
MULLIK ATOM ORBITAL 2 3 4 5 6 7 9 END
```

Requests an atom- and orbital-group based analysis for molecular orbitals 2 to 7 inclusive, plus molecular orbital 9

Example 2

The data line below has an equivalent effect to that of Example 1.

```
MULLIK ATOM ORBITAL 2 TO 7 9 END
```


has an equivalent effect to that of Example 1.

Example 3

The following data line will request an atom-grouped analysis for the total wavefunction alone, with no printed analysis of the individual orbitals.

```
MULLIK ATOM
```

9 Distributed Multipole Analysis

The DMA analysis [5] is instigated by the directive DMA.

By default, the DMA module will generate an expansion with multipoles at the atomic sites, with a maximum l value (rank) of 10 for the poles at each site. Each site is assigned a relative radius (see below) which is used in the partitioning of the overlap density between the sites. By default, all relative radii are set to 1.0.

A number of subdirectives may be presented after the DMA line, these serve to modify the selection of DMA sites, and the distribution algorithms for partitioning between the sites.

9.1 ADD

```
ADD name x y z <lmax <radius>>
```

Add a new site at (x,y,z) with the name specified. the multipole rank is limited to l_{max} if a value is specified, and a relative radius can be specified also.

9.2 DELETE

```
DELETE name
```

Delete all sites of the given name. DELETE ALL deletes all sites. DELETE CHARGE deletes nuclear charges on the atoms.

9.3 RADIUS

```
RADIUS name radius
```

Specify a relative radius for all sites with the name given. The actual distances from an overlap centre to the sites are scaled by dividing by the relative radii of the sites, and the contributions are moved to the site which is closest, in terms of scaled distances, to the overlap centre. The default is that all sites have relative radius 1.0.

9.4 LIMIT

```
LIMIT name lmax
```

Limit the rank of multipoles on sites with the name given to lmax at most. Contributions with higher ranks are moved to other sites. If no name is given the limit applies all sites. default (and maximum) is 20 for the linear version, 10 otherwise.

9.5 SHIFT

```
SHIFT tshift
```

Distribute multipoles from an overlap contribution around several dma sites, using a Gaussian weighting function. tshift is a cutoff parameter; maximum 1, minimum 1e-6. A value of 1.0 (the default) means distribution is to the nearest site only.

9.6 LINEAR

This directive invokes a faster version of the DMA program, which is applicable when all the atoms lie in a line parallel to the z axis and only the z components of the multipoles are required. In this case the maximum rank is 20. The option is revoked if the molecule is found not to be linear. If the molecule is subject to an external field, or is not in a singlet sigma state, there may be other non-vanishing multipole moments which will not be calculated; however the ql0 will still be correct.

9.7 GAUGE

```
GAUGE ox oy oz
```

The GAUGE directive resets the coordinate origin, and is followed by the coordinates of the new origin.

The distributed multipole analysis is not affected, but the total multipoles are referred to the new origin. The sites (as printed out and in the punchfile) are specified with respect to the *new* origin.

9.8 NONUCLEAR

The nuclear contribution to the multipoles is not evaluated.

10 The DRF Model for Solvation

The (Direct) Reaction Field model for solvation, developed at the University of Groningen [1, 2], is an embedding technique enabling the computation of the interaction between a quantum-mechanically described molecule and its classically described surroundings. The classical surroundings may be modelled in the following ways:

1. by point charges to model the electrostatic field due to the surroundings
2. by polarizabilities to model the (electronic) response of the surroundings
3. by an enveloping dielectric to model bulk response (both static and electronic) of the surroundings
4. by an enveloping ionic solution, characterized by its Debye screening length

The four representations may be combined freely to model all aspects of the surroundings. The best results with this model for solvation studies have been obtained by immersing the QM solute by 2-3 layers of explicitly described (point charges and polarizabilities) solvent molecules, enveloped by a surface defining the boundary between the microscopic system and a dielectric with bulk-solvent properties (dielectric constant). [2] The model has also been applied to active sites in proteins. [6] Special care has to be taken to avoid spurious electrostatic and reaction-field interactions with nearby atoms when bonds between the QM and classical systems are cut and the QM system is capped by H-atoms to satisfy valence. These aspects of embedding are the subject of ongoing research.

Having decided on the QM system and the representation of the surroundings, the embedding may be treated at the following levels:

1. electrostatic potential as a perturbation
The QM density is calculated as if the QM system were in vacuum. The interaction with the point charges is then calculated with the vacuum density.
2. electrostatic potential and reaction field as a perturbation
The QM density is calculated as if the QM system were in vacuum. The interaction with the point charges, polarizabilities, and dielectric is then calculated with the vacuum density.
3. electrostatic potential self-consistently
The QM density is calculated in the presence of the potential generated by the point charges by including this field in the one-electron hamiltonian.
4. electrostatic potential self-consistently and reaction field as a perturbation
The QM density is calculated in the presence of the potential generated by the point charges by including this field in the one-electron hamiltonian. The interaction with the polarizabilities and dielectric is then calculated with this density.

5. electrostatic potential and reaction field self-consistently

The QM density is calculated in the presence of the potential generated by the point charges, and the reaction field due to induced dipoles at the polarizabilities and surface polarization at the dielectric boundary, by including these fields in the one- and two-electron parts of hamiltonian, and Fock-matrix, respectively.

For many systems, the difference in total energy between the fourth and fifth levels is small; the self-consistent treatment of the electrostatic field is often found to change results substantially from a fully perturbative treatment.

The response of the surroundings is calculated through solving the set of linear equations that result from coupling the moments induced at the polarizabilities and at the boundary elements, the latter resulting from the linearized Poisson-Boltzmann equation. This formally requires the inversion of what is called the relay-matrix, which contains all the couplings. The procedure implemented here is the LU-decomposition of the relay-matrix. The LU-decomposition is the most expensive step in the embedding procedure. Care should therefore be taken to limit the number of polarizabilities and boundary elements. The relay-matrix is square matrix of a maximum of $(3*N_{pol} + N_{be}) \times (3*N_{pol} + N_{be})$ elements, where N_{pol} is the number of polarizabilities, and N_{be} is the number of boundary elements. In the case of specifying a non-zero ionic strength of the dielectric, another N_{be} is added to the dimensionality of the relay-matrix.

11 Total energy in the DRF model

The total energy in the DRF model is may be partitioned in various contributions:

1. energy of the quantum system

This is the energy of the quantum-mechanical system, as calculated with the supported wave functions. Any change in energy upon interaction with the classical system is to be measured against a vacuum calculation on the quantum-mechanical system.

2. energy of the classical system

This is the energy of the classical system, calculated as if there were no QM system present. The zero of energy depends on the specification of the classical system; for example, atoms may be defined to be part of molecules, excluding their interactions, making the infinitely separate molecules defining the zero of energy, rather than infinitely separate atoms.

In this energy interactions between classical subsystems (e.g. molecules) are included. They may be electrostatic, dispersion, repulsion, and induction interactions.

3. interaction energy

This is the sum of all separable interaction energies:

- (a) electrostatic interactions: interactions between point charges in the classical system and the QM charge distribution (nuclei and electrons).

- (b) Induction Interactions, also called screening. This is the interaction of one subsystem with the reaction field induced by another subsystem. The interaction of the subsystem with its own reaction field is also part of the interaction, and is used to calculate the polarization energy, which is half the energy gain from induction at equilibrium.
- (c) dispersion interactions: an estimate of the dispersion energy between QM and classical subsystems, based upon the Second-order Perturbation (SOP) expression for the dispersion interaction, may be calculated (see also section 12.4).
- (d) model repulsion energy: a molecular mechanics force-field expression from CHARMM [7] is used to model (Pauli) repulsion between subsystems.

NOTE: Electrons do not 'feel' the repulsion and care has to be taken to avoid close contacts, which may lead to electrons 'wandering off' the QM system. Devices to lessen this effect are available (see section 12.2)

Care should be taken in comparing total energies that the proper reference systems have been defined.

In a single calculation, one does not have access to the polarization energy of the QM system: this is the change of the expectation value of the vacuum hamiltonian of the QM distribution (and therefore internal QM energy) upon interaction with the surroundings. A separate calculation in the absence of the surroundings (or with the effect of the surroundings treated as a perturbation only) is necessary to obtain this (interaction) energy.

The definition of the zero of energy for the classical system has been discussed above. The total energy of QM + classical system is given on output under: configuration total energy. Contributions have been described in the Output section 14.

12 Features of the DRF Model

12.1 Grouping of External Points

Points that carry point charges and polarizabilities may be grouped together. Interactions between members of the same group (e.g. a molecule) may be excluded. Also, the dimensionality of the coupling matrix may be reduced by using group polarizabilities, rather than atomic polarizabilities. For instance, for a classical surroundings comprising of any number water molecules, described by 3 atoms each, a reduction of a factor 3 in the dimensionality of the coupling matrix may be achieved by using one molecular polarizability for each water molecule. For small molecules, the results from using group polarizabilities do not differ considerably from using distributed atomic polarizabilities if the molecules are distant enough from the inducing electrostatic field.

Criteria for grouping may be supplied by the user, and groupings may be suggested, which are then tested against the criteria. Grouping may also be enforced.

Thole has shown that for small organic molecules group polarizabilities from atom polarizabilities yield very good molecular polarizabilities, compared to experiment.[8] He derived a number of

parameter sets for the atomic polarizabilities, based on an assumption about the charge-density distribution around an atom. The shape of the charge density modifies the field at atoms within the molecule due to dipoles induced at other atoms in the molecule. The density–shape-functions implemented here are the homogeneous–conical and the exponentially decreasing spherical shapes.

12.2 Damping Functions

Problems with numerical stability and/or unphysical behaviour may arise when using the full Coulomb potentials and fields for the calculation of interactions between subsystems that are nearby. Use of full Coulomb potentials does not account for any overlap of the charge-clouds that would damp the potential if both subsystems were treated quantum-chemically.

Thole developed damping functions to account for overlap effects, and they may be used to damp the potentials and fields between QM system and classical systems and between classical subsystems.[8] Two types of damping function are available:

1. damping based on the assumption that the charge is distributed homogeneously in a cone of certain width around the nucleus/expansion centre
2. damping based on the assumption that the charge is exponentially decreasing radially from the nucleus/expansion centre

Both models have been used to construct molecular polarizabilities from atomic polarizabilities.[8]

12.3 CHARMM Model Repulsion

The repulsion between atoms in the QM system and classical system, and between the atoms in the classical subsystems is treated by using the CHARMM force-field expression:

$$E_{rep,CHARMM} = \sum_{i < j} \frac{3}{4} \frac{\alpha_i \alpha_j (r_i + r_j)^6}{\sqrt{\alpha_i/n_i + \sqrt{\alpha_j/n_j}}} r_{ij}^{-12}$$

where α_i is the polarizability of atom i , r_i is the radius of atom i , and n_i is the number of valence electrons of atom i , and r_{ij} is the distance between atoms i and j .

The parameters used here do not come from the CHARMM force-field, but are instead the DRF parameters. Several options for setting the radii and polarizabilities are available.

A special treatment of H-bonds is possible, as in the CHARMM force-field. The radius of the H-atom participating in a H-bond may be set differently from other H-atom radii.

12.4 Estimate of the Dispersion Interaction

The interaction of the QM system with the polarizable environment is calculated by taking expectation values of the reaction-field operator over the wave function. The expectation values are of two kinds:[9, 10]

1. average: the interaction of the charge density as a whole with the (dipole) moments induced by the charge distribution as a whole at the polarizability; in formula:

$$\langle 0|F \dagger |0\rangle\alpha\langle 0|F|0\rangle$$

where $|0\rangle$ is the wave function, F the electric field operator, and α the polarizability.

2. direct: the screening of the self-interaction through the interaction with the polarizable surrounding; in formula:

$$\langle 0|F \dagger \alpha F|0\rangle$$

The availability of these two expectation values enables an estimate of the dispersion interaction between QM systems and surroundings, through a fluctuation expression:

$$E_{disp,DRF} = \frac{1}{2}(\langle 0|F \dagger \alpha F|0\rangle - \langle 0|F \dagger |0\rangle\alpha\langle 0|F|0\rangle)$$

The factor $\frac{1}{2}$ comes in when accounting for the polarization costs.

This expression may be rewritten in the same form as the second-order perturbation (SOP) expression for the dispersion between two fragments by inserting the resolution of the identity in the first term. The $k = 0$ terms in the first and second terms cancel, and:

$$E_{disp,DRF} = \frac{1}{2}\sum_{k\neq 0}\langle 0|F \dagger |k\rangle\alpha\langle k|F|0\rangle$$

This expression is the same as the SOP expression in the Unsöld approximation, but for a factor

$$\frac{U_S}{U_A+U_S}$$

where U_S and U_A are mean excitation energies for the polarizable and QM systems, respectively. This factor must be supplied by the user. This factor may be estimated from e.g. ionization energies (experimental or calculated). If S and A are identical, the factor is $\frac{1}{2}$. The user may specify this factor to switch on this interaction. By default, this factor is set to zero, and therefore the QM system is embedded in the Average Reaction Field only.

The expressions above have been derived for a QM system interacting with one polarizability. By substituting the effective response of coupled polarizabilities and dielectric response for α , the estimate for the dispersion interaction with all the surroundings may be made.

The dispersion interaction between classically described molecules is given by the Slater-Kirkwood expression:

$$E_{disp,clas} = \frac{1}{4} \frac{Tr(\alpha_i T_{ij}^2 \alpha_j)}{\sqrt{\alpha_i/n_i + \sqrt{\alpha_j/n_j}}}$$

where α_i is the polarizability of atom i , n_i is the number of valence electrons of atom i , and T_{ij} is the dipole-dipole interaction tensor between atoms i and j . A minimum number of parameters is introduced if one uses the same polarizabilities defined for the reaction-field (induction) interaction in this expression and the CHARMM repulsion as well.

12.5 Expansion of Fields

In this DRF implementation the electrostatic and reaction field operators of the QM charge distribution are expanded to second order, which greatly reduces the computational effort because

the coupling equations need to be solved only for a limited number of source multipoles at each expansion centre, rather than for every single overlap distribution. Each overlap distribution is assigned to an expansion centre for which the static and reaction fields are calculated in advance for a unit charge, dipole, and quadrupole. These formal interactions are then simply multiplied by the overlap, dipole, and second moment integrals of the charge distribution(s) in question to yield one- and two-electron matrix elements.[1]

By default, the nuclei of the QM system are used as expansion centres, but the user may add expansion centres by hand, or by automated procedures.

Exact interaction with point charges may be obtained through using the BQ centres (see §3.8.4) specified in the ZMATRIX directive. As far as the DRF model is concerned, these BQ centres are part of the QM system.

12.6 Dielectric Response

The dielectric response due to the continuum surrounding the molecular system may be calculated separately for the static and optic components in a single calculation. To this end, the two dielectric constants may be given on input. If the Direct Reaction Field is active, only the optic component of the dielectric response is coupled directly to the charge distribution for calculation of the estimate of the dispersion interaction, since this part reflects the electronic part of the dielectric response; the static component is coupled through the Average Reaction Field.

The properties of the dielectric continuum enveloping the molecular system that may be specified are both the dielectric constant and the ionic strength. This is due to the implementation of the Poisson-Boltzmann equation-solver on the boundary surface, rather than just the Poisson equation-solver. This option may be useful for studying solvation effects in ionic solutions.

NOTE: At present, the analytical gradients (both on the QM system and on the classical points (charge and/or polarizabilities)) are not available, limiting geometry optimization capabilities of this module.

13 DRF Directives

The (Direct) Reaction Field model is invoked through specifying a block of data marked by the single-line directives REACT at the start and END at the end. Inside this block, the positions and magnitude of the point charges and polarizabilities, the definition of an enveloping surface to mark the beginning of a dielectric, and all DRF-options may be specified by their respective directives given below. The definition of the QM system remains to be specified through the ZMATRIX directive. Any capping atoms to be introduced when excising a QM system from a covalently bonded superstructure must be specified with the QM system, since there is no automatic generation for these atoms. Also, the user is responsible for avoiding close contacts between QM (capping) atoms and classical atoms.

The REACT subdirectives are given below.

13.1 FIELD

The FIELD directive specifies the level of coupling between QM system and surroundings. The directive consists of a single data line, starting with FIELD, which may contain a number of keywords followed by their values. The FIELD keywords are:

1. STAT specifies presence of static embedding potential, followed by the desired level
2. REAC specifies presence of response embedding potential, followed by the desired level. The level of embedding may be:
 - (a) NONE: this embedding is not taken into account
 - (b) PERT: this embedding is added as a perturbation after convergence of the wave function
 - (c) SCF: this embedding is treated self-consistently with the wave function

The default is FIELD STAT NONE REAC NONE

13.2 DRFOUT

The DRFOUT directive serves to specify the level of output for the DRF extension. The directive consists of a single data line, read to the variables TEXT, OUTOPT, using format (2A)

1. TEXT should be set to the character string DRFOUT
2. OUTOPT is the level of output for the DRF module. Valid options are:
 - (a) STANDARD:
 - (b) SOME: Information on classical surroundings
 - (c) MORE: More information on classical surroundings, printing of DRF settings, and flagging DRF-additions to hamiltonian
 - (d) MATRICES: MORE plus DRF matrices produced
 - (e) ONEEL: MORE plus one-electron information
 - (f) TWOEL: MORE plus two-electron information

The default is DRFOUT STANDARD.

NOTE: This directive must proceed the EXTERNAL directive to be effective there.

13.3 UNITS

The UNITS directive specifies the unit of length used throughout the REACT input directives. The directive consists of a single data line read to the variables TEXT, UNIT, using format (2A),

1. TEXT should be set to the character string UNITS

2. UNIT can be set to the character strings BOHR or ANGS

The default is UNITS BOHR.

NOTE: All REACT subdirectives will be assumed to be in the units given here, e.g. DSTGRP 15. will assign 15 Bohr to DSTGRP if UNITS BOHR, but 15 Angstrom if UNITS ANGS is specified.

NOTE: This directive must proceed the EXTERNAL directive to be effective there.

13.4 GAMDRF

The GAMDRF directive serves to specify the scaling factor to be applied to the estimate of the dispersion energy (see also section 12.4). A non-zero value also triggers the use of the Direct, rather than Average Reaction-Field coupling scheme. The directive is read to the variables TEXT, VALUE, using format (A,F)

1. TEXT should be set to the character string GAMDRF
2. VALUE is the scaling factor to be applied

The default is GAMDRF 0.0.

13.5 INCLPOL

The INCLPOL directive serves to specify the strength of the reaction field coupled to the QM system. The reaction field may either be coupled back in full, or at half-strength only, in which case the polarization energy is taken into account beforehand. The directive is read to the variables TEXT, OPTION, using format (2A)

1. TEXT should be set to the character string INCLPOL
2. OPTION may be either ON, or OFF. When INCLPOL ON, the polarization is taken into account beforehand

The default is INCLPOL ON.

13.6 EXPANDCM

The EXPANDCM directive specifies the use of the centre of nuclear charge of the QM partition as an extra expansion centre for the (D)RF integrals. The directive consists of a single data line read to the variables TEXT, EXPOPT, using format (2A),

1. TEXT should be set to the character string EXPANDCM
2. EXPOPT can be set to the character strings ON or OFF

The default is EXPANDCM OFF.

13.7 ASSIGN

The ASSIGN directive specifies the way the overlap distributions are assigned to the expansion centres. The centre of a charge distribution is defined by $\langle i|\hat{r}|j\rangle/\langle i|j\rangle$. The directive consists of a single data line read to the variables TEXT, ASGNOPT, using format (2A),

1. TEXT should be set to the character string ASSIGN
2. ASGNOPT can be set to the following character strings:
 - (a) DISTANCE: assign overlap distributions on the basis of distance of the centre of the overlap distribution to the expansion centres. The overlap distribution is assigned to the nearest expansion centre. If two or more centres are equally close, the assignment is arbitrarily to the last centre encountered in the search
 - (b) OLAPTOCM: assign all two-centre overlap distributions to the expansion centre at centre of nuclear charge (only with EXPANDCM ON)
 - (c) EQUATOCM: assign overlap distributions that are equally close to one or more expansion centres to the expansion centre at centre of nuclear charge (only with EXPANDCM ON)
 - (d) MIDPTS: add midpoints between atoms as expansion centres and assign as with DISTANCE
 - (e) ALLDISTR: expand each overlap distribution around its own centre. If any expansion centre is close to one already defined, the expansion centre will be shared

The default is ASSIGN DISTANCE.

NOTE: Extra (other) expansion centres may be specified by using BQ centres in the ZMATRIX directive. These centres will then be used as expansion centres. Note that if the BQ centres carry charge, interaction with the QM atoms will be included in the QM energy, interaction with classical atoms in the QM/classical interaction.

13.8 GROUPING

The GROUPING directive specifies the attempt to construct group polarizabilities from atomic polarizabilities on the basis of the atoms' group name (see the EXTERNAL directive). It also serves to define the number of characters in the group name of atoms used to exclude interactions between atoms. The directive consists of a single data line read to the variables TEXT, GROUPOPT, NGRNAM1, NGRNAM2, using format (2A,2I)

1. TEXT should be set to the character string GROUPING
2. GROUPOPT can be set to the character strings ON or OFF
3. NGRNAM1 is the number of characters at the start of the group name considered for exclusion of interactions

4. NGRNAM2 is the number of characters, after the name considered for exclusion of interaction, considered for attempting to group atom polarizabilities to a group polarizability

The default is `GROUPING OFF 2 2`.

NOTE: The `GROUPING` directive and the following grouping criteria are ignored if the user *forces* grouping of certain atoms. This may be done within the `EXTERNAL` data block by entering a data-line reading `GROUP` following a number of atoms. See also the `EXTERNAL` directive.

NOTE: `NGRNAM1` specifies the number of characters considered from the classical atoms' group name to exclude electrostatic, dispersion, and repulsion interactions between classical atoms. The `NGRNAM2` characters of the group name from `NGRNAM1` onward are considered when deciding whether to construct a group polarizability from the preceding atom polarizabilities. Any remaining characters may be used for additional labelling of the classical atoms. The energy does not depend on these remaining characters.

NOTE: The names of the classical atoms are stored in a 16-character string, the group name starting at character 7. There are thus a maximum of 10 characters available for the group name, and `NGRNAM1` may not exceed 8 to leave 2 characters for grouping polarizabilities.

NOTE: This directive must proceed the `EXTERNAL` directive to be effective.

13.9 DSTGRP

The `DSTGRP` directive specifies the minimum distance criterion for constructing group polarizabilities from atomic polarizabilities under the control of the `GROUPING` directive. The directive consists of a single data line read to the variables `TEXT`, `DSTGRP`, using format (A,F)

1. `TEXT` should be set to the character string `DSTGRP`
2. `DSTGRP` is the minimum distance any classical atom in the prospective group must be separated from any QM atom for the grouping to be considered. Grouping is not effected if any classical atom in the prospective group is closer to any QM atom than `DSTGRP`

The default is `DSTGRP 15.0`.

NOTE: This directive must proceed the `EXTERNAL` directive to be effective.

13.10 DSTMAX

The `DSTMAX` directive specifies the maximum distance criterion for constructing group polarizabilities from atomic polarizabilities under the control of the `GROUPING` directive. The directive consists of a single data line read to the variables `TEXT`, `DSTMAX`, using format (A,F)

1. `TEXT` should be set to the character string `DSTMAX`
2. `DSTMAX` is the maximum distance the nearest classical atom in the prospective group should be from any QM atom to effect grouping. If that atom is further removed, the whole group is discarded

The default is `DSTMAX 1000.0`.

NOTE: This directive must proceed the `EXTERNAL` directive to be effective.

13.11 AGRPE

The `AGRPE` directive specifies the polarization energy difference criterion for constructing group polarizabilities from atomic polarizabilities under the control of the `GROUPING` directive. The directive consists of a single data line read to the variables `TEXT`, `AGRPE`, using format (A,F)

1. `TEXT` should be set to the character string `AGRPE`
2. `AGRPE` is the maximum allowed energy difference between the reaction field interaction of a unit charge at the nearest QM atom as calculated with and without the grouping. If the energy difference is larger, grouping is not effected

The default is `AGRPE 0.001` (in Hartree).

NOTE: This directive must proceed the `EXTERNAL` directive to be effective.

13.12 AGRPM

The `AGRPM` directive specifies the induced dipole difference criterion for constructing group polarizabilities from atomic polarizabilities under the control of the `GROUPING` directive. The directive consists of a single data line read to the variables `TEXT`, `AGRPM`, using format (A,F)

1. `TEXT` should be set to the character string `AGRPM`
2. `AGRPM` is the maximum allowed induced dipole difference between the induced dipole due to a unit charge at the nearest QM atom as calculated with and without the grouping. If the induced dipole difference is larger, grouping is not effected

The default is `AGRPM 0.01` (in au).

NOTE: This directive must proceed the `EXTERNAL` directive to be effective.

13.13 AGRPC

The `AGRPC` directive specifies the cosine criterion for constructing group polarizabilities from atomic polarizabilities under the control of the `GROUPING` directive. The directive consists of a single data line read to the variables `TEXT`, `AGRPC`, using format (A,F)

1. `TEXT` should be set to the character string `AGRPC`
2. `AGRPC` is the minimum allowed cosine of the angle between the induced dipole moments due a unit charge at the nearest QM atom as calculated with and without the grouping. If the cosine is smaller, grouping is not effected

The default is AGRPC 0.99.

NOTE: This directive must proceed the EXTERNAL directive to be effective.

13.14 DAMPING

The DAMPING directive serves to specify the type of damping function and corresponding width-parameter to be used to damp electrostatic fields by to avoid the so-called polarization catastrophe (see section 12.2). It also controls the selection of default atom polarizabilities. The directive consists of a single data line, starting with DAMPING, which may contain a number of keywords (AFCT followed by its value). The DAMPING keywords are:

1. OFF to specify use of the full Coulomb operator
2. EXPO to specify the damping associated with an exponentially decreasing spherical shape function
3. CONE to specify the damping associated with a uniform conical shape function
4. AFCT to specify the width parameter, the value following the keyword. With EXPO, the width parameter is the exponential prefactor; with CONE, the width parameter is the radius of the cone. Beyond this radius, the charge distribution behaves as a point charge

The type of damping function used (CONE or EXPO) also determines the *default atom polarizabilities* assigned to classical atoms (see also 12.1), and possibly radii assigned to both QM and classical atoms (see subdirectives QMRADI and CLASRADI). If DAMPING OFF, the polarizabilities belonging to the damping function associated with the conical charge distribution will be used.

The default is DAMPING OFF AFCT 1.662.

NOTE: This directive must proceed the EXTERNAL directive to be effective.

13.15 CLASDISP

The CLASDISP directive serves to specify the treatment of the dispersion interaction between the classical atoms and groups. The directive consists of a single data line read to the variables TEXT, n(INTOPT), using format (nA)

1. TEXT should be set to the character string CLASDISP
2. INTOPT may be either blank or one of following keywords:
 - (a) GROUPOPOL or ATOMPOL to switch between the use of group polarizabilities and atom polarizabilities. If there are polarizable atoms that are not used in the construction of any group polarizability, the dispersion between those atom polarizabilities and any group polarizabilities is calculated if the GROUPOPOL option is specified.
 - (b) ORGPOL or EFFPOL to switch between the use of the input atom polarizabilities and the effective atom polarizabilities after group polarizability formation. This option affects atom polarizabilities only

- (c) ISODIS or NONISO to switch between the use of isotropic and anisotropic polarizabilities. Atom polarizabilities that have not been used in the construction of any group polarizability are always isotropic, and this option does not affect the dispersion interaction between such atoms

The default is CLASDISP ATOMPOL ORGPOL ISODIS.

NOTE: This directive must proceed the EXTERNAL directive to be effective in setting the correct polarizabilities.

13.16 QMRADI

The QMRADI directive specifies the method for calculating the radii of atoms in the QM partition for use in both construction of a boundary surface marking the dielectric and the CHARMM QM/classical repulsion, if appropriate. The directive consists of a single data line read to the variables TEXT, RADOPT, using format (2A),

1. TEXT should be set to the character string QMRADI
2. RADOPT is the option for selecting the radii. Valid options are:
 - (a) TABLE: look-up from internal table (Bondi's van der Waals radii)
 - (b) CONEPOL: calculated from Thole's polarizability optimized for the damping according to a conical charge-density
 - (c) EXPOPOL: calculated from Thole's polarizability optimized for the damping according to an exponential charge-density

The default is QMRADI TABLE

NOTE: The radius is calculated as $r = f_{shape}\alpha^{1/3}$, with f_{shape} equal AFCT, specified in subdirective DAMPING.

13.17 CLASRADI

The CLASRADI directive specifies the method for calculating the radii of atoms in the classical partition for use in both construction of a boundary surface marking the dielectric and the CHARMM QM/classical repulsion, if appropriate. The directive consists of a single data line read to the variables TEXT, RADOPT, using format (2A),

1. TEXT should be set to the character string CLASRADI
2. RADOPT is the option for selecting the radii. Valid options are:
 - (a) TABLE: look-up from internal table (Bondi's van der Waals radii)
 - (b) CONEPOL: calculated from Thole's polarizability optimized for the damping according to a conical charge-density

- (c) EXPOPOL: calculated from Thole's polarizability optimized for the damping according to an exponential charge-density
- (d) USERPOL: calculated from assigned polarizability, whether given on input (see EXTERNAL subdirective) or through look-up

The default is CLASRADI TABLE

NOTE: The radius is calculated as $r = f_{shape}\alpha^{1/3}$, with f_{shape} equal AFCT, specified in subdirective DAMPING.

NOTE: This directive must precede the EXTERNAL directive to be effective.

13.18 EXTERNAL

The EXTERNAL directive marks the beginning of the block input for x-, y-, z-coordinates, charge, polarizability, and radius of the atoms in the classically treated system. The block is closed with an END directive.

The attributes of each atom in the classically treated system are input on a single data line, in the following order, but with free format; first 2 character strings, then 6 numerics:

name (2 parts) – charge – x, y, z co-ordinate – polarizability – radius

1. name:

The name of a classical atom consists of two parts:

(a) chemical symbol:

The chemical symbol is a maximum of 2 characters. Apart from the elements, classical points may also be named XX, QQ, and E, which give them special attributes.

- i. QQ: point-charge only, no polarizability, no radius (polarizability and radius given on the data line are ignored, nominal unit polarizability for use in damping function only)
- ii. XX: point-charge with default unit polarizability and unit radius (polarizability and radius given on the data line are effective)
- iii. E : point-charge and polarizability, no radius (radius given on the data line is ignored)

NOTE: The extra atoms without radius (QQ and E) are ignored in the calculation of the model repulsion. XX-atoms do partake in the repulsion; they are assigned 1 valence electron. For use in the Slater–Kirkwood formula for the dispersion interaction, E-atoms are assigned 1 valence electron.

(b) group name:

The group name serves three purposes:

- i. to exclude interactions between classical atoms. The exclusions are controlled by the first part of the group name; the first NGRNAM1 characters are considered when deciding the exclusion of interactions

- ii. to define for which polarizabilities grouping to form a group polarizability should be attempted. The NGRNAM2 characters of the group name following the NGRNAM1 characters are considered when deciding whether to attempt grouping the preceding atom polarizabilities
- iii. label the classical atoms

NOTE: NGRNAM1 and NGRNAM2 may be specified on the GROUPING directive (see above).

NOTE: The names of the classical atoms are stored in a 16-character string, the group name starting at character 7. There are thus a maximum of 10 characters available for the group name, and NGRNAM1 may not exceed 8, to leave at least 2 characters for the grouping name.

NOTE: A group is formed on the basis of the group name if the relevant part of the group name changes from one atom to the next, and formation of group polarizabilities is requested (see the GROUPING directive).

NOTE: The specification of a group name is mandatory. If no use is to be made for grouping purposes, the best way to ensure independent treatment of the atoms is by using the group name to label the atoms uniquely.

NOTE: When grouping under the GROUPING directive, a change in group name at the NGRNAM2 characters governing grouping will signify the attempt to group the preceding atoms. If further down the list, the same group name for grouping is used again, the first set of atoms will not be considered again.

NOTE: The group name of the resulting group centre inherits the group name of the first atom in the newly formed group; the group centre gets this atom's index, and this atom is then put down as the last member of the group.

2. x-, y-, z-coordinates:

Cartesian co-ordinate input for classical atoms is supported only.

3. polarizability:

Atomic polarizabilities may be specified optionally. If not specified, default values are used, according to the chemical symbol, and the setting of the shape-function used for damping potentials (see section 12.2).

NOTE: if the potentials are undamped, the polarizabilities associated with the conical shape function will be used.

4. radius:

Atomic radii may be specified optionally. If not specified, default values are used, according to the chemical symbol, under control of the CLASRADI directive (see above).

NOTE: the radii are used in the CHARMM repulsion expression and for definition of a Juffer or Connolly surface if such a surface is requested (see the DIELECTRIC directive).

Within the EXTERNAL block one other directive may be specified to force grouping of the preceding polarizabilities:

GROUP ANAL n:

Forces construction of a group polarizability from the preceding atoms, following the previous group. This directive supersedes the formation of groups on the basis of group names (see above). Thus, a GROUP directive placed within a series of atoms with the same second part of the group name will result in two, rather than one group polarizabilities from those atoms.

The ANAL n, where n is an integer (maximum 10) is an optional addition to the forced grouping, causing analysis of the RF contributions to the total energy to be split into groups. n provides an index into the analysis groups. In this way, contributions from e.g. first and second solvent shells, or from different residues in a protein may be separated.

EXAMPLE: 2 classical water molecules

Here are a number of examples of 2 classical water molecules:

1. EXTERNAL part of input:

```
external
o w1 -0.796 3.459378 4.492394 0.0
h w1 0.398 3.459378 5.595494 +1.4325
h w1 0.398 3.459378 5.595494 -1.4325
group anal 1
o w2 -0.796 3.459378 -4.492394 0.0 8.4 3.1
h w2 0.398 3.459378 -5.595494 +1.4325 3.2 2.5
h w2 0.398 3.459378 -5.595494 -1.4325 3.2 2.5
group anal 2
end
```

With all other directives at their default values, the following attributes are given to the classical system (copied from output):

```
induced dipole criterion for grouping not met
WARNING: grouping forced by user, but grouping criteria not met
grouping information may be obtained by specifying the REACT subdirective DRFOOT MORE
```

```
6 points found on input of which 6 polarisable
constructed
8 (charged) points
0 point polarisabilities and
2 group polarisabilities
```

```
classical system specification
name x y z charge radius alfa (b**3) effalf (b**3) analysis
group w1gr 3.459378 5.092418 0.000000 0.000000 3.542 9.678 0.000 1
h w1 3.459378 5.595494 1.432500 0.398000 2.267 0.000 3.469 1
h w1 3.459378 5.595494 -1.432500 0.398000 2.267 0.000 3.469 1
o w1 3.459378 4.492394 0.000000 -0.796000 2.872 0.000 5.817 1
group w2gr 3.459378 -4.969410 0.000000 0.000000 3.688 10.925 0.000 2
h w2 3.459378 -5.595494 1.432500 0.398000 2.500 0.000 3.200 2
h w2 3.459378 -5.595494 -1.432500 0.398000 2.500 0.000 3.200 2
o w2 3.459378 -4.492394 0.000000 -0.796000 3.100 0.000 8.400 2
```

Points to note:

1. a warning is given that the grouping may not be of enough accuracy to be justified. The criteria are given in the directives AGRE, AGRPM, and AGRPC.
 2. the group centres head the group members; the first member of the group on input is now the last atom in the group. The first part of the centre's group name w1 is inherited from the first atom; the second part gr is added internally in the absence of further group-name input.
 3. the radius of the group centre is calculated from the group polarizability, and depends on the settings of the directive DAMPING.
 4. the radii and polarizabilities of the first water molecule are taken from internal data bases, and depend on the settings of the directives CLASRADI, and DAMPING, respectively.
 5. the radii and polarizabilities of the second water molecule are taken from input, overriding any internal settings. The group polarizability is, however, constructed under the settings of the DAMPING directive, as is the radius of the group centre.
 6. the values of *alfa* are the mean polarizabilities used in calculation of the relay matrix to yield induction and polarization terms. For group polarizabilities, the polarizability tensor need not be (and in general isn't) isotropic.
 7. the values of *effalf* are the (mean) polarizabilities that are used in calculation of the dispersion interaction between the classical subsystems. The effective polarizabilities used need not be isotropic. Their value and use depend on the settings of the CLASDISP directive. In this example, the isotropic atomic polarizabilities as given on input are used. (The default is CLASDISP ATOMPOL ORGPOL ISODIS.)
 8. the water molecules are put into two analysis groups (1 and 2). This will lead to final output in which the interaction between the first and second molecule is separated out, as is true for their respective interactions with the QM system.
2. The following EXTERNAL input is processed under the GROUPING ON 2 3, and DSTGRP 5.0 directives, with other directives at their default values:

```

external
o  w1gr1  -0.796      3.459378      4.492394      0.0
h  w1gr1   0.398      3.459378      5.595494     +1.4325
h  w1gr1   0.398      3.459378      5.595494     -1.4325
xx w2gr2  -0.796      3.459378     -4.492394      0.0      8.4  3.1
xx w2gr2   0.398      3.459378     -5.595494     +1.4325   3.2  2.5
xx w2gr2   0.398      3.459378     -5.595494     -1.4325   3.2  2.5
end

```

The relevant part of the output is:

```

induced dipole criterion for grouping not met
- - - criteria for grouping not met, grouping atoms      5 to      7
not effected

6 points found on input of which      6 polarisable

```

```

constructed
  7 (charged) points
  3 point polarisabilities and
  1 group polarisabilities

```

```

classical system specification
name      x      y      z      charge  radius  alfa (b**3)  effalf (b**3)  analysis
group w1gr1  3.459378  5.092418  0.000000  0.000000  3.542    9.678    0.000    1
h      w1gr1  3.459378  5.595494  1.432500  0.398000  2.267    0.000    3.469    1
h      w1gr1  3.459378  5.595494 -1.432500  0.398000  2.267    0.000    3.469    1
o      w1gr1  3.459378  4.492394  0.000000 -0.796000  2.872    0.000    5.817    1
xx     w2gr2  3.459378 -4.492394  0.000000 -0.796000  3.100    8.400    8.400    1
xx     w2gr2  3.459378 -5.595494  1.432500  0.398000  2.500    3.200    3.200    1
xx     w2gr2  3.459378 -5.595494 -1.432500  0.398000  2.500    3.200    3.200    1

```

Points to note:

1. grouping the first molecule has succeeded; all criteria are met, and the resulting attributes are the same as for the forced grouping.
2. grouping of the second molecule was not effected because the induced dipole criterion was not met, and the atoms are left as on input.
3. no separate analysis of the interactions of the two molecules will be made (it is not possible to specify an analysis group without forced grouping through the GROUP directive.
4. the energy of the classical system and the interaction of the classical system with the QM system will be different for cases 1. and 2. for the following type of interaction:
 - (a) dispersion between classical subsystems: this is because an XX centre is assigned 1 valence electron. In case 1. one of the XX centres is an O with 6, rather than 1 valence electron.
 - (b) repulsion between classical subsystems: again this is due to the assignment of just 1 valence electron to XX.
 - (c) induction and polarization energies: this is because in case 2., the polarizability of the second molecule is represented by a distributed, rather than group polarizability. This will affect the response of the system.
 - (d) dispersion energy between QM and classical system: this is again because of the difference in response of the classical system.

13.19 DIELECTRIC

The DIELECTRIC directive specifies the options for the dielectric continuum. It consists of a block data structure terminated by an END directive.

Below follow the descriptions of the DIELECTRIC subdirectives.

13.19.1 SURFACE

The SURFACE directive serves to specify the type of surface enveloping QM and classical atoms to be constructed. The directive consists of a single data line read to the variables TEXT, SURFOPT, using format (2A)

1. TEXT should be set to the character string SURFACE
2. SURFOPT is the type of surface to be constructed. The following options are valid:
 - (a) SPHERE: The surface is a triangulated sphere around (0,0,0). The radius and the number of surface points may be further specified by the DIELECTRIC subdirectives SRADIUS and BEMLEV, respectively
 - (b) JUFFER: The surface is generated by Juffer's method, which distorts a triangulated sphere to envelop the atoms according to the molecular shape. The surface is constructed by searching for the atom furthest away from the centre of the initial sphere along a number of spokes. Each spoke defines a cylinder within which an atom is to be found. If no atom is found within the cylinder, the program terminates. The number of surface points may be further specified by the DIELECTRIC subdirective BEMLEV, and the location of the surface w.r.t. the atom-positions by the subdirectives JUFFER, and CONNOLLY
 - (c) CONNOLLY: The surface is generated by using M.L. Connolly's surface program MSCON,[\[11\]](#) which has been made part of GAMESS-UK. Further specification of the Connolly surface may be given through the CONNOLLY directive

The default is SURFACE SPHERE.

13.19.2 SRADIUS

The SRADIUS directive serves to specify the radius of a triangulated sphere around (0,0,0) defining the dielectric boundary (with SURFACE SPHERE). The directive is read to the variables TEXT, SPHERAD, using format (A,F)

1. TEXT should be set to the character string SRADIUS
2. SRADIUS is the radius of the triangulated sphere

The default is SRADIUS 1000.0.

13.19.3 BEMLEV

The BEMLEV directive serves to specify the number of elements of the triangulated surface defining the dielectric boundary. This directive is active when the SURFACE keyword is either SPHERE or JUFFER. The directive is read to the variables TEXT, BEMLEV, using format (A,I)

1. TEXT should be set to the character string BEMLEV
2. BEMLEV is the level of detail of the triangulated sphere. Valid options are:
 - (a) 0 : results in 60 boundary elements
 - (b) 1 : results in 240 boundary elements
 - (c) 2 : results in 960 boundary elements

The default is BEMLEV 0.

13.19.4 JUFFER

The JUFFER directive serves to specify the generation of a Juffer triangulated surface enveloping QM (and classical, if present) atoms. The directive consists of a single data line, starting with JUFFER, which may contain a number of keywords followed by their values. The JUFFER keywords are:

1. CYLWDTH: width of a cylinder around a spoke. The atom within this cylinder which is furthest away from the centre of the molecular system defines the position of the surface element associated with the spoke. If CYLWDTH is not set, or < 0.0 , it defaults to $2 * \text{maxrad} / (\text{BEMLEV} / 2 + 1) + \text{RPROBE}$, where maxrad is the largest atomic radius of the QM and classical systems
2. SURFDIST: distance of the surface-defining point to the atom generating the surface. If SURFDIST is not set, or < 0.0 , it defaults to $\text{maxrad} + \text{RPROBE}$, where maxrad is the largest atomic radius of the QM and classical systems

The default is JUFFER CYLWDTH -1.0 SURFDIST -1.0.

NOTE: RPROBE may be defined in subdirective CONNOLLY.

13.19.5 CONNOLLY

The CONNOLLY directive serves to specify the generation of a Connolly surface enveloping QM (and classical, if present) atoms. The directive consists of a single data line, starting with CONNOLLY, which may contain a number of keywords followed by their values. The CONNOLLY keywords are:

1. MXSURPTS: maximum number of surface elements to be generated by the Connolly algorithm. The algorithm is satisfied if the number of surface points is between MNSURPTS and MXSURPTS
2. MNSURPTS: minimum number of surface elements to be generated by the Connolly algorithm. If unspecified, it is set equal $0.8 * \text{MXSURPTS}$

3. SPDENS: Connolly surface-point density. Because the surface-point density is optimized when setting MXSURPTS, specification of both MXSURPTS and SPDENS conflicts and results in an error message
4. RPROBE: probe radius for defining solvent-accessible surface. RPROBE 0.0 yields the Van der Waals surface

The default is `CONNOLLY MXSURPTS 200 RPROBE 1.0`.

NOTE: Care is required when specifying SPDENS; the number of surface points generated is unchecked, and a large relay-matrix to be LU-decomposed may result.

NOTE: RPROBE may also be relevant with `SURFACE JUFFER`.

13.19.6 SOLVENT

The `SOLVENT` directive serves to specify the solvent name associated with the dielectric continuum. If the solvent is known, solvent properties will be taken from the internal database. Values in the internal database supersede any other input. The directive is read to the variables `TEXT`, `SOLNAM`, using format (2A)

1. `TEXT` should be set to the character string `SOLVENT`
2. `SOLNAM` is the name of the solvent associated with the dielectric continuum (more than 8 characters are permissible). An internal database storing solvent static and optic dielectric constants, molecular masses, and densities is accessed through this directive. Specification of an unknown solvent will result in discontinuation of the calculation, except when `SOLNAM X` is specified

The default is `SOLVENT X`.

13.19.7 DIELTYP

The `DIELTYP` directive serves to specify the types of dielectric response taken into account. Static and/or optic dielectric response may be taken into account. If the estimate of the dispersion energy is requested (see `GAMDRF` directive), the appropriate direct reaction field coupling of the electrons is only to the optic part of the dielectric response, if the optic response is requested. The directive is read to the variables `TEXT`, `DIELOPT1` `DIELOPT2`, using format (3A)

1. `TEXT` should be set to the character string `DIELTYP`
2. `DIELOPT1` and `DIELOPT2` are the types of response requested. Either may be set to `STAT` or `OPT`, or left empty, requesting static and/or optic response, and absence of response, respectively

The default is `DIELTYP STAT`.

13.19.8 DIELOUT

The DIELOUT directive serves to specify the level of output associated with the surface defining the dielectric boundary. The directive is read to the variables TEXT, OUTDIEL, using format (2A)

1. TEXT should be set to the character string DIELOUT
2. OUTDIEL is the level of output requested. The actual output depends on the type of surface requested. Valid options are:
 - (a) STANDARD: Standard output
 - (b) SOME:
 - (c) MODERATE:
 - (d) EXTENDED:
 - (e) ALL: Maximum output

The default is DIELOUT STANDARD.

13.19.9 EPSSTAT

The EPSSTAT directive serves to specify the static dielectric constant associated with the dielectric continuum (DIELTYP STAT). The directive is read to the variables TEXT, EPS1, using format (A,F)

1. TEXT should be set to the character string EPSSTAT
2. EPS1 is the static dielectric constant. This should be a number larger than 1.0 to take effect. Values smaller than 1.0 are ignored; a value of 1.0 effectively means the dielectric is a vacuum and is therefore not active. (A finite ionic strength (see subdirective KAPPAS) may still be specified, though.)

The default is EPSSTAT 1.0.

13.19.10 EPSOPT

The EPSOPT directive serves to specify the optic dielectric constant associated with the dielectric continuum (DIELTYP OPT). The directive is read to the variables TEXT, EPS2, using format (A,F)

1. TEXT should be set to the character string EPSOPT
2. EPS2 is the optic dielectric constant. This should be a number larger than 1.0 to take effect. Values smaller than 1.0 are ignored; a value of 1.0 effectively means the dielectric is a vacuum and is therefore not active. (A finite ionic strength (see subdirective KAPPAO) may still be specified, though.)

The default is EPSOPT 1.0.

13.19.11 KAPPAS

The KAPPAS directive serves to specify the ionic strength associated with the static dielectric constant associated with the dielectric continuum (DIELTYP STAT). The directive is read to the variables TEXT, KAPPAS, using format (A,F)

1. TEXT should be set to the character string KAPPAS
2. KAPPAS is the ionic strength associated with the static dielectric constant. This should be a number larger than 0.0 to take effect. Values smaller than 0.0 are ignored.

The default is KAPPAS 0.0.

13.19.12 KAPPAO

The KAPPAO directive serves to specify the ionic strength associated with the optic dielectric constant associated with the dielectric continuum (DIELTYP OPT). The directive is read to the variables TEXT, KAPPAO, using format (A,F)

1. TEXT should be set to the character string KAPPAO
2. KAPPAO is the ionic strength associated with the optic dielectric constant. This should be a number larger than 0.0 to take effect. Values smaller than 0.0 are ignored.

The default is KAPPAO 0.0.

13.20 CLASCLAS

The CLASCLAS directive serves to specify the types of interaction between classical subsystems to be calculated. The directive consists of a single data line read to the variables TEXT, n(INTOPT), using format (nA)

1. TEXT should be set to the character string CLASCLAS
2. INTOPT may be either blank or one or more of the keywords ELST, NOELST, REPN, NOREPN, DISP, or NODISP, to switch calculation of the electrostatic, repulsion, and dispersion interactions between the classical fragments on or off, respectively.

The default is CLASCLAS ELST DISP REPN.

13.21 HBOND

The HBOND directive serves to specify the use of the special CHARMM repulsion parameters for H-bonds. The directive consists of a single data line, starting with HBOND which may contain a number of keywords followed by their values. The HBOND keywords are:

1. ON or OFF to switch special H-bond repulsion on or off, respectively
2. RADI to set the H-atom radius used in CHARMM-expression when within H-bond distance
3. DIST to set the distance within which H-bond radius is used

The default is HBOND OFF Specifying HBOND ON is equivalent to HBOND ON RADI 1.0 DIST 4.5.

13.22 DRFTWOEL

The DRFTWOEL directive serves to specify the treatment of the 2-electron reaction-field integrals. The directive consists of a single data line read to the variables TEXT, INTOPT, using format (2A)

1. TEXT should be set to the character string DRFTWOEL
2. INTOPT may be either blank or one of the keywords:
 - (a) DIRECT: 2-electron RF integrals are calculated on the fly in each SCF cycle
 - (b) DISK: 2-electron RF integrals are added to the vacuum 2-electron integrals stored on disk

The default is DRFTWOEL DIRECT

NOTE: This keyword only takes effect when the QM density is treated self-consistently with the reaction field.

NOTE: DRFTWOEL DISK is currently valid only in the following combination: SUPER OFF NOSYM and GAMDRF 1.0.

14 DRF Output—Analysis of (D)RF Energies

The (D)RF module produces additional output at the end of the normal output. This is a full analysis of the (D)RF contributions to the energy. The output is structured as follows:

1. energies associated with the quantum system (T: electron kinetic energy operator; V_0 : vacuum one-electron energy operator; g_0 : vacuum two-electron energy operator):
 - (a) vacuum nuclear repulsion
 - (b) one-electron energy: expectation value of $T + V_0$, using the final density
 - (c) electron kinetic energy: expectation value of T, using the final density
 - (d) nuclear attraction energy: expectation value of V_0 , using the final density
 - (e) two-electron energy: expectation value of g_0 , using the final density

- (f) scf energy: converged total energy of the QM system. This may include contributions from static and reaction fields if the QM density is treated self-consistently with these fields
 - (g) screening of the nuclear repulsion energy: interaction of all nuclei with the RF induced by all nuclei
 - (h) screening of the nuclear attraction energy: interaction of all nuclei with the RF induced by all electrons plus interaction of all electrons with the RF induced by all nuclei
 - (i) screening of the two-electron energy: interaction of all electrons with the RF induced by all electrons
 - (j) screening of the electronic self energy: sum of interaction of each electron with the RF induced by itself (only with the DIRECT RF option, see also the GAMDRF directive)
 - (k) molecular reaction field stabilisation: interaction of the QM system with the RF induced by the QM system
 - (l) equilibrium molecular polarisation energy: energy cost to induce dipole moments and surface polarisation by the QM system
2. energies associated with the classical system, collected per analysis group:
- (a) vacuum classical electrostatic interaction: interaction between point charges in the classical (sub)system(s)
 - (b) vacuum dispersion energy estimate: polarizability-based dispersion interaction between atoms in the classical (sub)system(s) (see also CLASDISP directive)
 - (c) vacuum repulsion energy estimate: CHARMM repulsion interaction between atoms in the classical (sub)system(s) (see also the HBOND directive)
 - (d) vacuum classical interaction energy: sum of the first three contributions
 - (e) screening of classical electrostatic energy: interaction of all charges in analysis group A with the RF induced by all charges in analysis group B, and *vice versa*
 - (f) equilibrium classical polarisation energy: energy cost to induce dipole moments and surface polarisation by the charges in the analysis group (only when A=B)
3. energies associated with the interaction between QM and classical system, collected per analysis group:
- (a) nuclei/ external charge interaction: electrostatic interaction between all QM nuclei (includes any BQ centres with a charge) and all charges in the analysis group
 - (b) electrons/ external charge interaction: electrostatic interaction between all electrons and all charges in the analysis group
 - (c) screening of the ext. charge/nuclear int.n: interaction of all nuclei (including BQ charges) with the RF induced by all charges in the analysis group
 - (d) screening of the ext. charge/electron int.n: interaction of all electrons with the RF induced by all charges in the analysis group

- (e) screening of the nuclei/ ext. charge int.n: interaction of all charges in analysis group with the RF induced by all QM nuclei (including BQ charges)
 - (f) screening of the electron/ext. charge int.n: interaction of all charges in analysis group with the RF induced by all electrons
 - (g) electrostatic qm-classical interaction: sum of first 2 terms
 - (h) screening electrostatic qm-classical int.n: sum of last 4 terms
 - (i) model repulsion energy: CHARMM model repulsion energy between QM system and all atoms in analysis group
4. total energies associated with the interaction between QM and classical system
- (a) as for individual analysis groups
 - (b) estimate of the dispersion energy: the DRF estimate for QM-classical dispersion, only in the Direct RF option (see section 12.4 and the GAMDRF directive)
5. summary of the contributions:
- (a) total energy of quantum system: expectation value of the vacuum hamiltonian (including nuclear repulsion) with the final density
 - (b) total energy of classical system
 - (c) quant. mech. /classical interaction energy: total interaction energy of QM system with surroundings
 - (d) equilibrium polarisation energy: total cost for inducing dipole moments and surface polarization
 - (e) configuration total energy: energy of the QM + classical system as a whole

References

- [1] A.H. de Vries, P.Th. van Duijnen, A.H. Juffer, J.A.C. Rullmann, J.P. Dijkman, H. Merenga, and B.T. Thole, *J. Comp. Chem.* **16** (1995) 37–55, doi:[10.1002/jcc.540160105](https://doi.org/10.1002/jcc.540160105) erratum 1445–1446, doi:[10.1002/jcc.540161113](https://doi.org/10.1002/jcc.540161113).
- [2] P.Th. van Duijnen and A.H. de Vries, *Int. J. Quant. Chem.*, **60** (1996) 1111, doi:[10.1002/\(SICI\)1097-461X\(1996\)60:6;1111::AID-QUA2;3.0.CO;2-2](https://doi.org/10.1002/(SICI)1097-461X(1996)60:6;1111::AID-QUA2;3.0.CO;2-2);
- [3] J.M. Foster and S.F. Boys, *Rev. Mod. Phys.* **32** (1960) 300, doi:[10.1103/RevModPhys.32.300](https://doi.org/10.1103/RevModPhys.32.300).
- [4] J. Pipek and P.G. Mezey, *J. Chem. Phys.* **90** (1989) 4916, doi:[10.1063/1.456588](https://doi.org/10.1063/1.456588).
- [5] A.J. Stone, *Chem. Phys. Lett.* **83** (1983) 233, doi:[10.1016/0009-2614\(81\)85452-8](https://doi.org/10.1016/0009-2614(81)85452-8).
- [6] J.P. Dijkman and P.Th. van Duijnen, *Int. J. Quant. Chem., Quant. Biol. Symp.* **18** (1991) 49, doi:[10.1002/qua.560400710](https://doi.org/10.1002/qua.560400710).
- [7] B.R. Brooks, R.E. Brucoleri, B.D. Olafson, D.J. States, S.J. Swaminathan, and M. Karplus, *J. Comp. Chem.* **4** (1983) 187, doi:[10.1002/jcc.540040211](https://doi.org/10.1002/jcc.540040211).
- [8] B.T. Thole, *Chem. Phys.* **59** (1981) 341, doi:[10.1016/0301-0104\(81\)85176-2](https://doi.org/10.1016/0301-0104(81)85176-2).
- [9] B.T. Thole and P.Th. van Duijnen, *Theor. Chim. Acta*, **55** (1980) 307, doi:[10.1007/BF00549429](https://doi.org/10.1007/BF00549429).
- [10] B.T. Thole and P.Th. van Duijnen, *Chem. Phys.*, **71** (1982) 211, doi:[10.1016/0301-0104\(82\)87020-1](https://doi.org/10.1016/0301-0104(82)87020-1).
- [11] M.L. Connolly, *Science*, **221** (1983) 709, doi:[10.1126/science.6879170](https://doi.org/10.1126/science.6879170).