

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 5. INTEGRAL TRANSFORMATION & DIRECT-CI

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

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

<b>1 Introduction</b>	<b>1</b>
<b>2 Integral Transformation</b>	<b>1</b>
<b>3 Directives Controlling Integral Transformation</b>	<b>2</b>
3.1 TRACC . . . . .	2
3.2 PASS . . . . .	2
3.3 ACTIVE . . . . .	3
3.4 CORE . . . . .	4
<b>4 RUNTYPE and Restarting the Transformation</b>	<b>5</b>

<b>5</b>	<b>Direct-CI Calculations</b>	<b>6</b>
5.1	Direct-CI - Default CISD Calculations	7
5.1.1	Closed-shell Systems	7
5.1.2	Open-shell Systems	9
<b>6</b>	<b>Direct-CI Data Input</b>	<b>10</b>
6.1	DIRECT	10
6.2	THRESH	10
6.3	MAXCYC	11
6.4	SHIFT	11
6.5	ALTERNAT	11
6.6	DIAGMODE	12
6.7	PRINTVAR	12
6.8	TRIAL	13
6.9	JACDAV	14
6.10	VPRINT	15
6.11	SPIN	15
6.12	NATORB	16
6.13	EXCIT	17
6.14	CONF	20
6.15	CONF-CARDS	30
6.16	REFGEN	31
6.17	CASGEN	32
6.18	SCREEN	33
6.19	RESTRICT	33
6.20	PRCONF	34
6.21	CEPA	34
6.22	MP	35
6.23	Spin Functions	36
6.24	Using GVB Orbitals	37
6.25	FP Geometry Optimisations	39
6.26	Calculating the $^1A_1$ states of $H_2O$	41

## 1 Introduction

In this chapter we turn our attention to the post-Hartree Fock modules within GAMESS-UK, considering initially the integral transformation routines and associated data input, and then the Direct-CI module. Note that the transformation module acts not only as a precursor to Direct-CI, but finds more widespread usage in, for example, both OVGf and TDA Green's function calculations, and in the semi-direct Table-CI module.

## 2 Integral Transformation

Before detailing the directives associated with the transformation of 1 and 2-electron integrals over atomic orbitals to the corresponding set over molecular orbitals, we describe briefly the data sets used. Note that the algorithm employed is basically that due to Yoshimine [2]. In addition to the Mainfile, Dumpfile and Scratchfile, the following files will be used, with the associated space requirements considered below:

- **Sortfile:** A dataset assigned to SORT will be used as a scratchfile in sorting operations. The space requirements are slightly more than twice the length of the Mainfile if a single pass sorting is adopted.
- **Secondary Mainfile:** Partially transformed 2-electron integrals are output to a dataset referred to as the Secondary Mainfile, and the user may direct this dataset to any of the files, ED0-ED19 and MT0-MT19. The LFN ED4 is used in default: the SFILE directive described in Part 3 may be used to re-assign Secondary Mainfile output. Care should be taken if the user assigns the Secondary Mainfile to the same file as the Mainfile. The Secondary Mainfile should not be allowed to overwrite the Mainfile except where the integrals sort has been completed in one pass (NPASS=1). The Mainfile should not be over-written if it is wished to perform a 2-index transformation of a Fock operator. Generally it is advisable to route the Secondary Mainfile to a different file, rather than to use the Mainfile.
- **Transformed Integral file:** Fully transformed 2-electron integrals over the molecular orbitals are output to datasets referred to as the Transformed Integral File, written in default to ED6. The user may re-direct this dataset to any file, ED0-ED19 and MT0-MT19 using the FFILE directive of Part 3. The Transformed Integral file may overwrite the Mainfile, but should not be allowed to overwrite the Secondary Mainfile except where the second integrals sort has been completed in one pass (NPASS2=1). Again overwriting of the Mainfile is not allowed if a Fock operator is to be transformed. It is advisable to route the Transformed Integral file to a different file than is allocated to either the Mainfile and the Secondary Mainfile.

The following approximation may prove useful in the considering the space requirements of the Secondary Mainfile, Transformed Integral File and Sortfile. Given NBASIS basis functions and NACT active orbitals, and defining,

$$M = \text{NBASIS} * (\text{NBASIS}+1) / 2.$$

$$N = \text{NACT} * (\text{NACT}+1) / 2.$$

$$L = \text{Number of blocks in the Mainfile (1 block = 512 words)}.$$

Then approximate space requirements are given by:

$$\begin{aligned} \text{Secondary Mainfile} &= S = (2*L*N) / M \\ \text{Transformed Integral file} &= (L*N*N) / (M*M) \\ \text{Sortfile} &= (2*L) / \text{NPASS1 or } S / \text{NPASS2} \\ &\text{(whichever is the greater)} \end{aligned}$$

### 3 Directives Controlling Integral Transformation

#### 3.1 TRACC

The TRACC directive consists of a single line read to variables TEXT, K using the format (A,I).

- TEXT should be set to the character string TRACC.
- The integer K is used to compute the threshold factor  $\text{ACC} = 10^{-K}$ , and if the absolute value of a transformed 2-electron integral is less than ACC, that integral will not be output to the Transformed Integral file. A factor  $\text{ACC1} = \text{ACC} * 10^{-2}$  is also computed, and if the absolute value of a partially trans-formed 2-electron integral is less than ACC1, that integral will not be output to the Secondary Mainfile.

The TRACC directive may be omitted, when the default value  $K = 10$  is assumed. The smaller the value of K, the shorter will be the size of the Secondary and Transformed Integral file, and the shorter the computation time in the second phase of the 4-index transformation.

#### 3.2 PASS

The PASS directive consists of a single dataline read to variables TEXT, NPASS1, NPASS2 using format (A,2I).

- TEXT should be set to the character string PASS.
- NPASS1 is an integer specifying the minimum number of passes of the Mainfile in the first phase of the 4-index transformation.
- NPASS2 is an integer specifying the minimum number of passes of the Secondary Mainfile in the second phase of the 4-index transformation.

In the absence of a PASS directive the module will calculate the minimum number of passes required, which will depend on the basis set size and the amount of memory available to the module. The user should note the following:

- The size of the Sortfile is usually inversely proportional to either NPASS1 or NPASS2. To reduce the size of the Sortfile multi-passing of the Mainfile and Secondary Mainfile must be employed.
- The program forms a dump enabling a restart of the 4-index transformation process at the end of each pass. The more passes, the shorter time interval between dumps.

**Example:**

```
PASS 3 2
```

Specifies a 3 and 2 pass sort of the Mainfile and Secondary Mainfile respectively.

### 3.3 ACTIVE

This directive specifies those members of the molecular orbital set which are deemed 'active' in the integral transformation, so that integrals of the form  $\langle ij/kl \rangle$  will be computed and output to the Transformed Integral File if all four orbitals are specified using the ACTIVE directive. The first data field consists of the character string ACTIVE in the first data field. Subsequent data lines are read to an array (IACTIV(I),I=1,NACT) using free I-format. The last data field presented should be the character string END.

**Example 1**

```
ACTIVE
10 11 12 13 14 15 16 17 18 20
END
```

Molecular Orbitals 10 to 18 and 20 are made active, and will be re-indexed 1 to 10 respectively, with the Direct-CI module, for example, referring to the orbitals in this re-indexed convention.

**Example 2**

```
ACTIVE
10 TO 18 20
END
```

This example shows the use of the string TO to abbreviate consecutive sequences of integers, and is equivalent to example 1.

### 3.4 CORE

The CORE directive allows the user to:

- route a 2-index transformed Fock operator to a specific section of the Dumpfile.
- factor frozen doubly occupied orbitals (such orbitals retain their double occupation in all configurations generated by the Direct-CI module, for example) into the Fock Operator (F) where:

$$F = H + 2 J[R] - K[R]$$

H denotes the usual one-electron operator (sum of kinetic and nuclear attraction), R denotes the CORE shells density matrix, and J and K are coulomb and exchange matrices constructed therefrom.

When employing the CORE directive in conjunction with factoring out molecular orbitals, it must be remembered that the molecular orbitals must be doubly occupied and no partial occupied orbitals can be factored out in this manner.

The first line is read to TEXT,NSECT in format (A,I).

- TEXT should be set to the character string CORE, although ONELEC is also acceptable.
- NSECT specifies the section number on the Dumpfile where the transformed 1-electron integrals are to be placed. If omitted the integrals are routed to section 466. If specified, NSECT must lie between 1 and 350.

Subsequent lines specify the frozen doubly occupied orbitals as a sequence of integers, corresponding to the MO ordering that came from the SCF module. The sequence can be abbreviated using the string TO, and is terminated by a line containing the string END in the first datafield.

#### Example 1

```
CORE 200
1 2
3 4 5 6 7 8 9 10
END
```

This example routes the transformed 1-electron integrals to section 200 of the Dumpfile. Molecular Orbitals 1 to 10 have been declared to be doubly occupied and frozen.

#### Example 2

```
CORE 200
1 TO 10
END
```

This example shows the use of the string TO to shorten the data input, and is equivalent to example 1.

### Example 3

```
CORE
1 TO 20
END
```

This example assumes default routing of the transformed 1-electron integrals, to section 466. MOs 1 to 20 will be frozen.

### Example 4

The CORE directive can be omitted, when no MOs will be frozen, and the transformed 1-electron integrals will be routed to section 466 of the Dumpfile. Omission is equivalent to:

```
CORE
END
```

## 4 RUNTYPE and Restarting the Transformation

In most applications, the transformation module will be run as part of either Direct-CI or Green's function calculations, under control of RUNTYPE CI, RUNTYP GF or RUNTYPE TDA specification. In some circumstances it may be necessary to generate the transformed integrals only, and a specific RUNTYPE, code-named TRANSFORM, has been provided for this purpose. Note that TRANSFORM processing includes both the SCF step and subsequent integral transformation, with restarts possible in both steps. For such restarting, the Mainfile, Secondary Mainfile and Transformed Integral file should have been permanent files in the startup job, exactly the same files being presented to the restart job, and, if used, identical MFILE, SFILE and FFILE directives should be used in the startup and restart jobs. The Sortfile need not be preserved between jobs.

### Example

In this example we are transforming the integrals from a TZVP calculation on H<sub>2</sub>CO. The first data file represents the startup job, the second the restart data, assuming the processing in the startup job did not complete in the allocated time.

### The Startup Job

```
TITLE
H2CO - TZVP BASIS - TRANSFORMATION
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 TZVP
RUNTYPE TRANSFORM
PASS 2 2
TRACC 9
ENTER
```

### The Restart Job

```
RESTART TRANSFORM
TITLE
H2CO - TZVP BASIS - TRANSFORMATION
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 TZVP
RUNTYPE TRANSFORM
PASS 2 2
TRACC 9
ENTER
```

## 5 Direct-CI Calculations

The Direct-CI module performs general multi-reference singles and doubles configuration interaction (CI) calculations. The method (direct or integral driven) used in the package is described in [1]. For optimal running the main memory allocation should be at least 3 times longer than the number of configuration state functions (CSFs) in the CI expansion. The module will execute in less memory, but at the cost of increased disc input/output, and higher overall job cost. The following data sets will be used by the program.

- Transformed Integral File: Integrals over the molecular orbitals (MOs) will be read from this file. The user may direct that this dataset be read from any file except ED5 or ED8 by means of the FFILE directive (see above). The default is to use ED6.
- Direct-CI File: The file ED5 is used to hold control information, accommodate the partial Hamiltonian matrix elements and to store the update vectors created by the Davidson diagonalization procedure. Twice the CI expansion length will be added at each iterative cycle. The DIAGMODE directive (see below) may be used to reduce the maximum size of the Davidson sub-space, thus limiting the ultimate size of ED5, possibly at the expense of an inferior rate of convergence. If it is thought possible that the diagonalization procedure may not converge in one job, ED5 must be retained to allow for restarts.



- The P-Sortfile: A dataset normally assigned using the local file name (LFN) PSORT will be used as a scratchfile in a pre-sort of the transformed molecular integrals. The space requirements of the P-Sortfile are about 1.5 times that of the Transformed integral File produced by the transformation module.
- the Sortfile: A dataset normally assigned using the LFN SORT will be used as a scratchfile in a post-sort of the transformed molecular integrals. The maximum space requirements of the SORT FILE are about twice that of the Transformed integral File produced by the transformation module, although this will be much reduced in high symmetry, by an inverse factor approaching the order of the point group involved.
- Overflow File: During the construction of the partial matrix elements a scratchfile allocated as file ED8 may be required. This is particularly likely to occur if a large number of reference CSFs are specified, while its likelihood is decreased if a large main memory allocation is used. ED8 is usually of the order 1000 to 4000 blocks long if it is required.

Data input characterising the CI calculation commences with the DIRECT data line, and is typically followed by a sequence of directives, terminated by presenting a valid *Class 2* directive, such as VECTORS or ENTER. The directives may be presented in any order, although in some cases directives are inter-related and care should be taken when presenting them, since the order in which the directives are presented is in such cases often significant; this is particularly the case for the EXCIT, CONF, REFGEN and CASGEN directives.

Before describing each of the directives in detail, we cater for those users who wish to "fast forward" through the directive descriptions by outlining how to perform "default" single reference CISD calculations. This provides a set of default attributes that bypasses the requirement for explicit data specification; while of somewhat limited applicability, it does provide a starting point for users, and a route to subsequent, more extensive calculations.

## 5.1 Direct-CI - Default CISD Calculations

In order to simplify the process of configuration specification and data preparation, the Direct-CI module now provides a set of default options that require little or no data input. To illustrate this default working of the module, we consider below a number of example calculations based on those that will be described in more detail in the subsequent sections.

### 5.1.1 Closed-shell Systems

A Direct-CI calculation is to be performed on the formaldehyde molecule. Given the following data sequence:

```
TITLE
H2CO - 3-21G DEFAULT DIRECT-CI CISD OPTION
ZMAT ANGSTROM
C
0 1 1.203
```

```

H 1 1.099 2 121.8
H 1 1.099 2 121.8 3 180.0
END
RUNTYPE CI
ENTER

```

then the calculation undertaken will be based on the following;

1. The format of the 2e-integral file will be automatically set to the required "SUPER OFF NOSYM", triggered by the presence of the CI runtype.
2. Integral transformation will use the set of orbitals from section 1, the default section for output of the closed-shell SCF eigenvectors. All orbitals will be deemed ACTIVE in the transformation.
3. The Direct-CI module is the default module loaded under RUNTYPE CI control, so that the DIRECT directive is not required.
4. The division of the molecular orbital space into an internal and external space, typically specified by the DIRECT directive, is now handled automatically, with the internal space comprising all doubly occupied SCF MOs orbitals, the external space all SCF virtual MOs. All electrons will be deemed active in the CI.
5. The SYMMETRY and SPIN of the CI wavefunction are taken to be those of the SCF wavefunction.
6. A single reference configuration will be employed, just the SCF configuration; the final configuration space will include all single and double excitations from this SCF reference configuration.
7. The spinfree natural orbitals will be written to section 11 of the Dumpfile.

The full data specification corresponding to the defaults generated from the above data file is shown below; the role of each of the directives will be described in later sections.

```

TITLE
H2CO - 3-21G CISD DIRECT-CI 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
DIRECT 16 8 14
CONF
2 2 2 2 2 2 2
NATORB 11 0 PRINT
ENTER

```

### 5.1.2 Open-shell Systems

Let us now consider a Direct-CI calculation on the  $^2B_2$  state of  $H_2CO^+$ , again using default options available within the module. A valid data sequence for performing such a calculation is shown below:

```
TITLE
H2CO+ 2B2 3-21G - DEFAULT CISD DIRECT-CI OPTION
MULT 2
CHARGE 1
ZMAT 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
ENTER
```

As with the closed-shell run above, no explicit data is required to define the nature of the CI calculation. In practice the defaults adopted correspond to the following:

1. The CI will be based on the high-spin open-shell RHF calculation.
2. The set of vectors used in the transformation will be the energy-ordered SCF orbitals from section 5 of the Dumpfile, the default section in the absence of section specification on the ENTER directive.
3. The symmetry and spin of the CI wavefunction will be deduced from the preceding SCF calculation i.e. a doublet CI wavefunction of  $B_2$  symmetry (corresponding to SPIN 2).
4. The number of active electrons in the CI will be set to be those involved in the SCF calculation (i.e. 15).
5. The reference configuration to be employed will be just the open-shell SCF configuration. The internal space comprises the doubly plus singly occupied SCF orbitals, with the external space comprising the SCF virtual orbitals. All electrons will be deemed active in the CI.
6. The spinfree natural orbitals will be written to section 11 and the spin natural orbitals to section 12 of the Dumpfile.

The full data specification corresponding to the defaults generated from the above data file is shown below; the role of each of the directives will again be described in later sections.

```
TITLE
H2CO+ - 2B2 - 3-21G CISD DIRECT-CI CALCULATION
SUPER OFF NOSYM
CHARGE 1
MULT 2
```

```
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
OPEN 1 1
DIRECT 15 8 14
SPIN DOUBLET
CONF
2 2 2 2 2 2 1
NATORB 11 12 PRINT
ENTER
```

## 6 Direct-CI Data Input

### 6.1 DIRECT

DIRECT consists of one line read to variables TEXT, NELEC, NINT, NEXT using format (A,3I).

- TEXT should be set to the character string DIRECT.
- NELEC specifies the number of electrons in the CI calculation. Notice that any inner shell electrons frozen out using the CORE directive of the transformation module should not be included.
- NINT specifies the number of internal MOs. These will be used to construct reference CSFs. If a MO is not occupied in any reference CSF it should not ordinarily be classified as internal, unless high levels of internal excitation are contemplated (see EXCIT directive below). The internal MOs normally correspond to that set capable of producing a qualitatively correct wavefunction. Notice that  $NINT*2$  must be greater than or equal to NELEC.
- NEXT specifies the number of external MOs. Such MOs will be unoccupied in all reference CSF. Single and double excitations from the internal to the external MOs cause the latter to contribute to the CI expansion. Notice that  $NINT+NEXT$  must be less than or equal to the number of active MOs as specified under control of the ACTIVE directive. If less than, then some active MOs will not take part in the CI, and in the absence of a REORDER directive (see below), these will be the highest indexed active MOs.

### 6.2 THRESH

This directive consists of a single line read to variables TEXT, C, K using format (A,F,I).

- TEXT should be set to the character string THRESH.

- C,K: The diagonalization is converged to a threshold (T) such that  $T=C/(10^{**}K)$ . If K is not set, it will be given the value 0. The lowest value to which T may be set is 1E-8, and this minimum will be selected if the user attempts to set a smaller T value.

The THRESH directive may be omitted, when T will be set to 3E-4.

### Example

```
THRESH 2 5
```

```
THRESH 2E-5
```

```
THRESH 0.00002
```

are equivalent, causing T to be set to 2E-5.

## 6.3 MAXCYC

This directive consists of one line read to variables TEXT, MAXC using format (A,I).

- TEXT should be set to the character string MAXCYC.
- MAXC specifies the maximum number of iterative cycles to be carried out by the Davidson diagonalizer.

The directive may be omitted, when MAXC will take the default value 50.

## 6.4 SHIFT

This directive consists of one line read to variables TEXT, SHIF using format (A,F).

- TEXT should be set to the character string SHIFT.
- SHIF should be set to the desired value of the level shifter to be used in the CI diagonalization phase.

If the SHIFT directive is omitted, the default SHIF=0.0 will be taken. For ground states small values (between 0.0 and 0.2) provide an optimal rate of convergence, and usually there is little point in using the SHIFT directive. For excited states, the rate of convergence may sometimes be markedly improved by using a SHIF value of between 0.3 and 0.5, particularly if the ALTERNAT directive (see below) is used.

## 6.5 ALTERNAT

This directive consists of one line which should contain the character string ALTERNAT in the first data field. If presented, the directive causes the sign of the SHIF parameter (see the SHIFT

directive above) to be altered at each iterative cycle of the CI diagonalization, and this may improve the convergence rate for excited states. We do not recommend use of ALTERNAT except in cases where severe convergence problems are encountered.

## 6.6 DIAGMODE

This directive consists of a single data line read to variables TEXT, ATEXT, NDAVID using format (2A,I).

- TEXT should be set to the character string DIAGMODE.
- ATEXT should be set to one of the character strings EMIN, VMIN or LOCK. EMIN causes the program to unconditionally minimize the total energy, and is normally the best option for ground states. VMIN causes the program to minimize the variance (sum of squares of residuals in the secular problem), and is usually the best option if convergence to an excited state is required. LOCK causes the program to seek a solution to the CI problem looking most like the trial wavefunction (see TRIAL directive below), and is therefore another way of trying to converge onto an excited state.
- NDAVID specifies the maximum size of the sub-space to be used in the Davidson diagonalization procedure. If omitted, the default NDAVID=50 will be taken. It may be necessary to use a smaller sub-space in order to limit the size of ED5, as explained above. The largest possible sub-space is 50; attempts to set NDAVID larger than 50 will cause the program to use this maximum.

The DIAGMODE directive may be omitted, when the defaults ATEXT=EMIN and NDAVID=50 will be taken. It is possible to omit the TEXT parameter of this directive.

### Example

```
DIAGMODE VMIN 50  
  
VMIN 50  
  
VMIN
```

all have an equivalent effect, causing the variance minimization option to be selected, with a maximum sub-space of 50.

## 6.7 PRINTVAR

This directive consists of a single line, whose first field should contain one of the character strings PRINTVAR or VARPRINT. The EMIN and LOCK modes of the diagonalizer (see DIAGMODE directive above) do not ordinarily compute the variance, and therefore do not output this quantity. The PRINTVAR directive may be used to turn such printing on, and is redundant if the VMIN option of the DIAGMODE directive is used, since the variance is always printed in this case.

## 6.8 TRIAL

This directive may be used to define the trial CI wavefunction as a linear combination of CSFs. There are two ways to accomplish this:

1. Specify the coefficients of the CSFs in the input file. In this case the first data line should contain the character string TRIAL in the first data field. There may be up to 20 lines following this directive initiator, each being read to variables ICSF, CCSF using format (I,F).
  - ICSF should be set to the index of a CSF in the CI wavefunction. Normally, such an index will only be known after running the program once, so that the configuration generator output can be studied. The user should understand the order in which the program generates spin states belonging to the same space occupancy pattern (see below) before using this directive, particularly where CSFs involving large numbers of singly occupied MOs are involved.
  - CCSF should be set to the coefficient of the CSF in the trial wavefunction. The trial wavefunction will be a linear combination of the indicated CSFs with coefficients given by the CCSF parameters.

### Example

```
TRIAL
1 0.5
2 -0.5
```

The trial wavefunction will consist of a linear combination of the first and second CSF in the CI expansion, with coefficients 0.5 and -0.5 respectively. It will be subsequently normalized by the program.

2. Select a subspace of CSFs and compute the eigenvector of the matrix corresponding to the subspace. In this case the syntax is:

```
TEXT TEXTA [TEXTB [NCSF] or TEXTC] [ TEXTD ISTATE ] [ TEXTE ]
```

where

- TEXT should be the literal string TRIAL.
- TEXTA should be the literal string DIAG.
- TEXTB may be set to SELECT or to FIRST. If the string SELECT is supplied a number of CSFs with the lowest energies will be selected. If the string FIRST is supplied a number of CSFs that are the first in order will be selected. Only a maximum of 100 CSFs can be selected. SELECT is default.
- TEXTC may be set to 'ref' or 'vac' selecting the whole reference space or vacuum space respectively.

- TEXTD may be set to 'state' in which case ISTATE is an integer specifying which eigenvector from the subspace matrix should be used. The default value is 1 i.e., the eigenvector with the lowest eigenvalue will be selected.
- TEXTE may be set to the literal string PRINT. If supplied the trial vector will be printed. By default the trial vector will NOT be printed. if PRINT is given twice the selected H-matrix is also printed

### Example

```
TRIAL DIAG SELECT 25
```

This will cause the program to select the 25 CSFs with the lowest eigenvalues. The matrix in this basis is solved for the eigenvector to obtain the trial vector.

### Default:

The TRIAL directive may be omitted, when the trial wavefunction will be selected from the reference space.

## 6.9 JACDAV

The JACDAV directive sets the controls of the Jacobi-Davidson preconditioning method. The syntax of this directive is

```
JACDAV SUBDIR [ SUBDIR [ .. ] ]
```

where JACDAV is a literal string, acting as the directive initiator, and SUBDIR is a valid subdirective. Each subdirective consists of a literal string eventually followed by an integer, real, or string argument. The supported subdirectives, OFF, ON, SHIFT, THRESH, MAXCYC and PRINT are detailed below:

1. **OFF**: This is a literal string, it switches the Jacobi-Davidson preconditioner off. By default the preconditioner is switched on.
2. **ON**: This is a literal string, that switches the Jacobi-Davidson preconditioner on. This is the default.
3. **SHIFT DYNAMIC / RSHIFT**: This subdirective sets the level shifter for the preconditioner. Here SHIFT is a literal string that is followed either by the literal string DYNAMIC or a real value for RSHIFT. If a real value RSHIFT is supplied then that value will be used for the level shifter. If the string DYNAMIC is supplied the level shifter will be automatically adjusted to force convergence. The latter is the default.
4. **THRESH RTHRS**: This subdirective sets the convergence threshold for the preconditioner. Here THRESH is a literal string and RTHRS is a real value. By default the threshold is set to half the CI threshold at time of calling JACDAV.



5. **MAXCYC IMAXC**: This subdirective sets the maximum number of cycles for the preconditioner. Here MAXCYC is a literal string and IMAXC is an integer value. By default IMAXC is set to 100.
6. **PRINT**: This subdirective controls the output level of the preconditioner. Here PRINT is a literal string. Every time this string is supplied the output level is increased, causing the preconditioner to generate a more detailed output. By default the printing is off.

NOTE: Because the default threshold is determined from the CI threshold at time of calling JACDAV interchanging the order of the THRESH directive and the JACDAV directive may lead to different convergence behaviour.

### 6.10 VPRINT

This directive consists of a single data line read to variables TEXT, NPR, TPR, ATEXT, BTEXT, CTEXT using format (A,I,F,3A), and is used to control printing of the CI wavefunction.

- TEXT should be set to the character string VPRINT.
- NPR specifies the maximum number of CI coefficients to be printed.
- CI coefficients less than TPR in absolute magnitude will not be printed.
- ATEXT, BTEXT, CTEXT may each be set to one of the character strings VAC, N-1, or N-2. If VAC is specified, CI coefficients of all the vacuum states (those with no electrons in the external MO space) will be printed, irrespective of the values of NPR and TPR. Similarly the parameters N-1 and N-2 control printing of those CSFs with one or two electrons in the external space respectively.

This directive may be omitted, when the defaults NPR=100 and TPR=1E-7 will be taken.

#### Example

```
VPRINT 50 0.02 VAC
```

will cause all vacuum state coefficients to be printed. All coefficients greater in absolute magnitude than 0.02 will be printed, unless there are more than 50 of these, in which case only the largest 50 will be printed.

### 6.11 SPIN

This directive consists of one line read to variables TEXT, NSPIN using format (A,I).

- TEXT should be set to the character string SPIN.

- NSPIN is used to specify the spin degeneracy of the CI wavefunction, using the values 1,2,3 etc. for singlet, doublet, triplet states etc. respectively. It is also possible to use one of the character strings SINGLET, DOUBLET, TRIPLET, QUARTET, QUINTET, SEXTET, SEPTET, OCTET and NONET to specify NSPIN.

The SPIN directive may be omitted, when the program will set NSPIN to 1 or 2 if NELEC is even or odd respectively.

### Example

```
SPIN 4
SPIN QUARTET
```

are equivalent; the wavefunction will be four-fold spin degenerate.

## 6.12 NATORB

The NATORB directive consists of a single dataline read to variables TEXT, KSPACE, KSPIN, ATEXT using format (A,2I,A).

- TEXT should be set to the character string NATORB.
- KSPACE is an integer (between 0 and 350 inclusive) specifying the section number of the Dumpfile where the spin-free NOs are to be placed. If KSPACE=0, spin-free NOs will not be routed to the Dumpfile.
- KSPIN is an integer (between 0 and 350 inclusive) specifying the section number of the Dumpfile where the spin NOs are to be placed. If KSPIN=0, spin NOs will not be routed to the Dumpfile. Notice that spin NOs will not be produced for singlet wavefunctions (see SPIN directive above) because they have an occupation number of zero in this case.
- ATEXT may be set to the character string PRINT, when the NOs will be printed. If ATEXT is omitted, the NOs will not be sent to the output.

### Example

```
NATORB 12 14 PRINT
```

The spin-free and spin NOs are output to sections 12 and 14 respectively of the Dumpfile, and routed to the output.

**Note:** In the absence of the NATORB directive (Version 6.3 onwards) both spin-free and, if appropriate, spin natural orbitals will be generated in default, and routed to sections 10 and 11 of the Dumpfile. This default thus corresponds to presenting the data line:

```
NATORB 10 11 PRINT
```

so that the NATORB directive now need only be presented to override these defaults.

### 6.13 EXCIT

This directive is used to define the excitation pattern allowed for a set of reference CSFs defined using the CONF directive (see below). The EXCIT directive may be used more than once in the data input to allow the user to specify different excitation patterns for different reference CSFs, and normally consists of a single dataline, in which the first data field should contain the character string EXCIT. The second data field may also be read in A format, and if so used should be set to one of the character strings OCTAL or BINARY. If this second A format field is omitted, the program takes OCTAL as default. Subsequent data fields are read in I-format, and should contain either octal numbers (valid range 0 to 7) or binary numbers (valid range 0 to 1), according to the character string contained in the second data field. The octal or binary integers may be continued onto subsequent lines if necessary. We now explain the significance of these integers if OCTAL input mode is selected. The first octal integer specifies the external excitation pattern for the reference CSFs, (where no internal excitations have been applied). This octal number should be translated into a binary format, such that:

OCTAL NUMBER		BINARY NUMBER
=====		=====
0	=	0 0 0
1	=	0 0 1
2	=	0 1 0
3	=	0 1 1
4	=	1 0 0
5	=	1 0 1
6	=	1 1 0
7	=	1 1 1

The left-most binary integer defines a double external excitation, where two electrons are promoted from the internal to the external space. If it is 0 or 1 the process is forbidden or allowed respectively. The middle binary integer defines the single external excitation process. If this is set to 1, the process is allowed, if it is 0, single external excitations are forbidden. The right-most binary integer corresponds to a no external excitation process, the reference CSFs being left as they are. If set to 0 or 1 the reference CSFs will be eliminated from or retained in the final list of CSFs for the CI. Note that CSFs may also be eliminated from the CI list because they are of the wrong spin/space symmetry.

#### Example 1

```
EXCIT OCTAL 7
```

This excitation mask will cause the reference CSFs and the single and double external excitations generated from them to be included in the final CI list of CSFs.

#### Example 2

```
EXCIT OCTAL 5
```

Will cause all the reference CSFs and all double external excitations generated therefrom to be added to the list of CSFs for the CI.

### Example 3

```
EXCIT OCTAL 4
```

Will cause the double external excitations of the reference states to appear in the final list of CSFs. Note the reference and single external excited CSFs are excluded from the CI list. If a second octal integer is defined, the reference CSFs have undergone a single internal excitation process. That is, a transfer of an electron from one internal MO to another. This second integer defines the external excitation mask on these newly constructed internal CSFs. A third octal integer defines the external excitation pattern after a double internal excitation. Additional octal integers may be presented, up to a maximum to 21 integers. Thus it is possible to define a state which is up to 20 fold internally excited and up to doubly externally excited.

### Example 4

```
EXCIT OCTAL 7 3 1
```

corresponds to a CI list containing the reference CSFs plus all single and double excitations (internally and externally). If the BINARY option is chosen, then the full binary patterns of each equivalent octal integer must be given.

### Example 5

```
EXCIT OCTAL 7 3 1
```

```
EXCIT BINARY 1 1 1 0 1 1 0 0 1
```

are equivalent. It is possible to give different excitation patterns to different reference CSFs.

### Example 6

```
EXCIT OCTAL 7 3 1
```

```
CONF
```

```
.
```

```
.
```

```
.
```

```
.
```

```
EXCIT OCTAL 5 0 1
```

```
CONF
```

```
.
```

```
.
```

```
.
```

The first set of reference CSFs are associated with the excitation pattern 7 3 1, while the second set have an excitation pattern 5 0 1. If the EXCIT directive is not invoked the excitation pattern will default to the setting 7 3 1, corresponding to all single and double excitations. An EXCIT directive presented without parameters will cause restoration of the 7 3 1 default.

Table 1: Resolution of the  $C_{\infty v}$  Species into the  $C_{2v}$  Species

Orbital		IRrep
$C_{\infty v}$	$C_{2v}$	Sequence No.
$\sigma$	$a_1$	1
$\delta_{x^2-y^2}$		
$\pi_x$	$b_1$	2
$\pi_y$	$b_2$	3
$\delta_{xy}$	$a_2$	4

Table 2: Resolution of the  $D_{\infty h}$  Species into the  $D_{2h}$  Species

Orbital		IRrep
$D_{\infty h}$	$D_{2h}$	Sequence No.
$\sigma_g$	$a_g$	1
$\delta_{g,x^2-y^2}$		
$\pi_{u,x}$	$b_{3u}$	2
$\pi_{u,y}$	$b_{2u}$	3
$\delta_{g,xy}$	$b_{1g}$	4
$\sigma_u$	$b_{1u}$	5
$\delta_{u,x^2-y^2}$		
$\pi_{g,x}$	$b_{2g}$	6
$\pi_{g,y}$	$b_{3g}$	7
$\delta_{u,xy}$	$a_u$	8

## 6.14 CONF

The CONF directive is used to specify the reference CSFs for the CI expansion. CONF may be presented more than once in the data input, usually in conjunction with a different excitation pattern (see the EXCIT directive above) acting on the reference CSFs. The first line of the CONF directive is set to the character string CONF. Each subsequent line specifies a reference CSF by giving the number of electrons (0,1 or 2) in each internal MO. Thus each reference CSF is defined by NINT numbers, the ordering of which should conform to the order of the internal MOs as specified under control of the ACTIVE directive. If necessary, CSF defining data may be carried over to further lines. The CONF directive may also be used to determine a reference set based on CARDS data dumped by a preceding CASSCF, e.g. CONF CARDS, see below, We now illustrate CONF usage through a series of examples that will be subsequently used in Part 6 when describing configuration input for the Table-CI module.

### Example 1

Consider performing a valence-CI calculation on the PH<sub>3</sub> molecule using a 6-31G(\*) basis. While the molecular symmetry is C<sub>3v</sub>, the symmetry adaptation and subsequent direct-CI will be conducted in the C<sub>s</sub> point group. An examination of the SCF output reveals the following orbital analysis.

```

=====
IRREP  NO. OF SYMMETRY ADAPTED
        BASIS FUNCTIONS
=====
      1         18
      2          7
=====

```

and the following orbital assignments characterising the closed-shell SCF configuration:

$$1a_1^2 2a_1^2 1e^4 3a_1^2 4a_1^2 2e^4 5a_1^2 \quad (1)$$

or, in the C<sub>s</sub> symmetry representation:

$$1a'^2 2a'^2 3a''^2 1a''^2 4a'^2 5a'^2 6a'^2 2a''^2 7a'^2 \quad (2)$$

```

=====
M.O.  IRREP  ORBITAL ENERGY  ORBITAL OCCUPANCY
=====
      1     1   -79.93661395      2.0000000
      2     1    -7.48916431     2.0000000
      3     1    -5.38319410     2.0000000
      4     2    -5.38319405     2.0000000
      5     1    -5.38149104     2.0000000
      6     1    -0.85610769     2.0000000
      7     1    -0.52191424     2.0000000
=====

```

8	2	-0.52191424	2.0000000
9	1	-0.38579686	2.0000000
10	1	0.16819544	0.0000000
11	2	0.16819544	0.0000000
12	1	0.26587776	0.0000000
13	1	0.46072690	0.0000000
14	2	0.46072690	0.0000000
15	1	0.47871033	0.0000000
16	1	0.56106989	0.0000000
17	1	0.89229884	0.0000000
18	2	0.89229885	0.0000000
19	2	0.91131383	0.0000000
20	1	0.91131383	0.0000000
21	1	0.93118300	0.0000000
22	1	1.17900613	0.0000000
23	2	1.45058658	0.0000000
24	1	1.45058658	0.0000000
25	1	3.78674557	0.0000000

=====

Assume that we wish to freeze the five inner shell orbitals:

$$1a'^2 2a'^2 3a'^2 1a''^2 4a'^2 \quad (3)$$

requiring the following data lines for the transformation

```
CORE
1 TO 5 END
ACTIVE
6 TO 25 END
```

To perform an 8-electron valence-CI calculation, involving the SCF configuration and two degenerate (1e)' to (2e)' doubly-excited configurations

$$5a'^2 8a'^2 2a''^2 7a'^2 \quad (4)$$

and

$$5a'^2 6a'^2 3a''^2 7a'^2 \quad (5)$$

would require the following CONF data:

```
CONF
2 2 2 2 0 0
2 0 2 2 2 0
2 2 0 2 0 2
```

where there are six orbitals in the internal space, the four doubly occupied valence SCF MOs, and the two components of the (2e)' virtual orbital. Note that no re-ordering of the MOs is required since the (2e)' orbitals are the two lowest unoccupied VMOs. The complete data file for performing the SCF and subsequent CI would then be as follows:

```

TITLE
PH3 * 6-31G*  VALENCE-CI 3M/1R
ZMAT
P
H 1 RPH
H 1 RPH 2 THETA
H 1 RPH 2 THETA 3 THETA  1
VARIABLES
RPH 2.685
THETA 93.83
END
BASIS 6-31G*
RUNTYPE CI
CORE
1 TO 5 END
ACTIVE
6 TO 25 END
DIRECT 8 6 14
CONF
2 2 2 2  0 0
2 0 2 2  2 0
2 2 0 2  0 2
ENTER

```

Note that the "SUPER OFF NOSYM" constraint on the two-electron integral file generated during the initial SCF will be automatically imposed by virtue of the nominated RUNTYPE. The SUPER directive need only be presented in an initial SCF calculation that is driven under RUNTYP SCF control, when the user plans to subsequently access this file in a separate CI step under control of the BYPASS directive.

### Example 2

In this example we wish to perform a valence-CI calculation on the CuCl molecule using a 3-21G basis. While the molecular symmetry is  $C_{\infty v}$ , the symmetry adaptation and subsequent CI will be conducted in the  $C_{2v}$  point group. The resolution of the  $C_{\infty v}$  into the  $C_{2v}$  orbital species is given in Table 1. An examination of the SCF output reveals the following orbital analysis.

```

=====
IRREP  NO. OF SYMMETRY ADAPTED
        BASIS FUNCTIONS
=====
   1           22
   2            9
   3            9
   4            2
=====

```

and the following orbital assignments from the converged closed shell SCF:

```

=====
M.O. IRREP  ORBITAL ENERGY  ORBITAL OCCUPANCY
=====

```



1	1	-326.84723972	2.0000000
2	1	-104.02836336	2.0000000
3	1	-40.71695637	2.0000000
4	1	-35.46377378	2.0000000
5	3	-35.45608069	2.0000000
6	2	-35.45608068	2.0000000
7	1	-10.42193940	2.0000000
8	1	-7.88512031	2.0000000
9	2	-7.88222844	2.0000000
10	3	-7.88222844	2.0000000
11	1	-5.07729175	2.0000000
12	1	-3.38247056	2.0000000
13	3	-3.35978308	2.0000000
14	2	-3.35978307	2.0000000
15	1	-1.01099628	2.0000000
16	3	-0.53702948	2.0000000
17	2	-0.53702947	2.0000000
18	4	-0.49640067	2.0000000
19	1	-0.49640067	2.0000000
20	1	-0.44715317	2.0000000
21	3	-0.39988537	2.0000000
22	2	-0.39988537	2.0000000
23	1	-0.35127248	2.0000000
24	1	0.00023285	0.0000000
25	3	0.06300102	0.0000000
26	2	0.06300102	0.0000000
27	1	0.12855448	0.0000000
28	1	0.19287013	0.0000000
29	3	0.25729975	0.0000000
30	2	0.25729975	0.0000000
31	1	0.39720201	0.0000000
32	1	0.86197727	0.0000000
33	2	0.88942618	0.0000000
34	3	0.88942618	0.0000000
35	1	1.01877167	0.0000000
36	1	2.16694989	0.0000000
37	3	3.96181512	0.0000000
38	2	3.96181512	0.0000000
39	4	3.98212497	0.0000000
40	1	3.98212497	0.0000000
41	1	4.08851360	0.0000000
42	1	24.51368240	0.0000000

=====

Assume that we wish to freeze the first 14 inner shell orbitals:

$$1\sigma^2 2\sigma^2 3\sigma^2 4\sigma^2 1\pi^4 5\sigma^2 6\sigma^2 2\pi^4 7\sigma^2 8\sigma^2 3\pi^4 \quad (6)$$

requiring the following data lines for the transformation

```

CORE
1 TO 14 END
ACTIVE
15 TO 42 END

```

To perform an 18-electron valence-CI calculation, based on the SCF configuration

$$9\sigma^2 4\pi^4 1\delta^4 10\sigma^2 5\pi^4 11\sigma^2 \quad (7)$$

would require the following CONF data:

```
CONF
2 2 2 2 2 2 2 2 2
```

The complete data file for performing the SCF and subsequent CI would then be as follows:

```
TITLE\CUCL .. 3-21G
ZMAT ANGSTROM\CU\CL 1 CUCL\
VARIABLES\CUCL 2.093 \END
BASIS 3-21G
RUNTYPE CI
CORE
1 TO 14 END
ACTIVE
15 TO 42 END
DIRECT 18 9 19
CONF
2 2 2 2 2 2 2 2 2
ENTER
```

The inclusion of a second reference configuration corresponding to the doubly excited configuration

$$9\sigma^2 4\pi^4 1\delta^4 10\sigma^2 5\pi^4 12\sigma^2 \quad (8)$$

would require incorporating the  $12\sigma$  orbital into the internal space, leading to 10 internal and 18 external MOs. The CONF data would then appear as follows:

```
CONF
2 2 2 2 2 2 2 2 2 0
2 2 2 2 2 2 2 2 0 2
```

and the overall data file,

```
RESTART NEW
TITLE\CUCL .. 3-21G
BYPASS SCF
ZMAT ANGSTROM\CU\CL 1 CUCL\
VARIABLES\CUCL 2.093 \END
BASIS 3-21G
RUNTYPE CI
CORE
1 TO 14 END
ACTIVE
15 TO 42 END
DIRECT 18 10 18
CONF
2 2 2 2 2 2 2 2 2 0
2 2 2 2 2 2 2 2 0 2
ENTER
```

where we have, assuming the Mainfile and Dumpfile to have been retained, by-passed the SCF and modified the DIRECT, CONF and VECTORS data.

### Example 3

Consider performing a valence-CI calculation on the  $\text{SiH}_4$  molecule using a 6-31G(\*) basis. While the molecular symmetry is  $T_d$ , the symmetry adaptation and subsequent CI will be conducted in the  $C_{2v}$  point group. An examination of the SCF output reveals the following orbital analysis.

```

=====
IRREP  NO. OF SYMMETRY ADAPTED
        BASIS FUNCTIONS
=====
  1          9
  2          6
  3          6
  4          6
=====

```

and the following orbital assignments from the converged closed shell SCF:

```

=====
M.O.  IRREP  ORBITAL ENERGY  ORBITAL OCCUPANCY
=====
  1    1    -68.77130710    2.0000000
  2    1     -6.12943325    2.0000000
  3    2     -4.23503117    2.0000000
  4    3     -4.23503117    2.0000000
  5    4     -4.23503117    2.0000000
  6    1     -0.73046864    2.0000000
  7    4     -0.48480821    2.0000000
  8    3     -0.48480821    2.0000000
  9    2     -0.48480821    2.0000000
 10    2      0.16291387    0.0000000
 11    3      0.16291387    0.0000000
 12    4      0.16291387    0.0000000
 13    1      0.25681257    0.0000000
 14    1      0.33606346    0.0000000
 15    3      0.37087856    0.0000000
 16    2      0.37087856    0.0000000
 17    4      0.37087856    0.0000000
 18    1      0.79946861    0.0000000
 19    1      0.79946861    0.0000000
 20    4      0.86232544    0.0000000
 21    3      0.86232544    0.0000000
 22    2      0.86232544    0.0000000
 23    1      1.23833149    0.0000000
 24    4      1.44033091    0.0000000
 25    3      1.44033091    0.0000000
 26    2      1.44033091    0.0000000
 27    1      3.13181655    0.0000000
=====

```

Assume that we wish to freeze the first 5 silicon inner shell orbitals requiring the following CORE and ACTIVE directives:

```
CORE
1 TO 5 END
ACTIVE
6 TO 27 END
```

To perform a 8-electron valence-CI calculation, based on the SCF configuration would require the following CONF data:

```
CONF
2 2 2 2
```

The complete data file for performing the SCF and subsequent CI would then be as follows:

```
TITLE
SIH4 * 6-31G* DIRECT VALENCE-CI 1M/1R
ZMAT
SI
H 1 SIH
H 1 SIH 2 109.471
H 1 SIH 2 109.471 3 120.0
H 1 SIH 2 109.471 4 120.0
VARIABLES
SIH 2.80
END
BASIS 6-31G*
RUNTYPE CI
CORE
1 TO 5 END
ACTIVE
6 TO 27 END
DIRECT 8 4 18
CONF
2 2 2 2
ENTER
```

#### Example 4

In this example we wish to perform a valence-CI calculation on the  $N_2$  molecule using a 4-31G(\*) basis. While the molecular symmetry is  $D_{\infty h}$ , the symmetry adaptation and subsequent CI will be conducted in the  $D_{2h}$  point group. The resolution of the  $D_{\infty h}$  into the  $D_{2h}$  orbital species is given in Table 2. An examination of the SCF output reveals the following orbital analysis.

```
=====
IRREP NO. OF SYMMETRY ADAPTED
      BASIS FUNCTIONS
=====
1           8
```

```

2      3
3      3
4      1
5      8
6      3
7      3
8      1
=====

```

and the following orbital assignments from the converged closed shell SCF:

```

=====
M.O. IRREP  ORBITAL ENERGY  ORBITAL OCCUPANCY
=====
  1    1   -15.65951533      2.0000000
  2    5   -15.65474750      2.0000000
  3    1    -1.50615941      2.0000000
  4    5    -0.75782277      2.0000000
  5    1    -0.63244925      2.0000000
  6    3    -0.63135826      2.0000000
  7    2    -0.63135826      2.0000000
  8    6     0.20154861      0.0000000
  9    7     0.20154861      0.0000000
 10    5     0.63883097      0.0000000
 11    1     0.82491489      0.0000000
 12    3     0.89634343      0.0000000
 13    2     0.89634343      0.0000000
 14    1     0.91812387      0.0000000
 15    7     1.10036132      0.0000000
 16    6     1.10036132      0.0000000
 17    5     1.17625689      0.0000000
 18    5     1.66995008      0.0000000
 19    4     1.70518236      0.0000000
 20    1     1.70518236      0.0000000
 21    3     1.91001614      0.0000000
 22    2     1.91001614      0.0000000
 23    8     2.29436539      0.0000000
 24    5     2.29436539      0.0000000
 25    1     2.84356916      0.0000000
 26    7     3.00847817      0.0000000
 27    6     3.00847817      0.0000000
 28    5     3.37447679      0.0000000
 29    1     3.71753400      0.0000000
 30    5     4.09917273      0.0000000
=====

```

Assume that we wish to freeze the two N1s inner shell orbitals, thus

```

CORE
1 2 END
ACTIVE
3 TO 30 END

```

To perform a 10-electron valence-CI calculation, based on the SCF configuration

$$2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_u^4 \quad (9)$$

and associated  $\pi$  to  $\pi^*$  excitations

$$2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_{u,x}^2 1\pi_{g,y}^2 \quad (10)$$

$$2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_{u,x}^2 1\pi_{g,x}^2 \quad (11)$$

$$2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_{u,y}^2 1\pi_{g,x}^2 \quad (12)$$

$$2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_{u,y}^2 1\pi_{g,y}^2 \quad (13)$$

$$2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 (1\pi_{u,x} 1\pi_{g,x})(1\pi_{u,y} 1\pi_{g,y}) \quad (14)$$

would require the following CONF data:

```
CONF
2 2 2 2 2 0 0
2 2 2 0 2 0 2
2 2 2 0 2 2 0
2 2 2 2 0 2 0
2 2 2 2 0 0 2
2 2 2 1 1 1 1
```

with an internal space of 7 orbitals, an external space of 21; note again that the ordering of the virtual MOs is such that no reordering is required within the ACTIVE directive. The complete data file for performing the SCF and subsequent CI would then be as follows:

```
TITLE\N2 .. 4-31G*
ZMAT ANG5\N\N 1 NN
VARIABLES\NN 1.05 \END
BASIS 4-31G*
RUNTYPE CI
CORE
1 2 END
ACTIVE
3 TO 30 END
DIRECT 10 7 21
CONF
2 2 2 2 2 0 0
2 2 2 0 2 0 2
2 2 2 0 2 2 0
2 2 2 2 0 2 0
2 2 2 2 0 0 2
2 2 2 1 1 1 1
NATORB 10 0 PRINT
ENTER
```

### Example 5

In this example we wish to perform a valence-CI calculation on the  $\text{CaH}_2$  molecule using a 3-21G basis. While the molecular symmetry is  $D_{\infty h}$ , the symmetry adaptation and subsequent CI will be conducted in the  $D_{2h}$  point group. An examination of the SCF output reveals the following orbital analysis.

```

=====
IRREP  NO. OF SYMMETRY ADAPTED
        BASIS FUNCTIONS
=====
  1         7
  2         4
  3         4
  5         6
=====

```

and the following orbital assignments from the converged closed shell SCF:

```

=====
M.O.  IRREP  ORBITAL ENERGY  ORBITAL OCCUPANCY
=====
  1     1    -148.37173884      2.0000000
  2     1    -16.76521275      2.0000000
  3     3    -13.55586861      2.0000000
  4     2    -13.55586861      2.0000000
  5     5    -13.55460610      2.0000000
  6     1     -2.26357685      2.0000000
  7     3     -1.36160958      2.0000000
  8     2     -1.36160958      2.0000000
  9     5     -1.35089927      2.0000000
 10     1     -0.34923025      2.0000000
 11     5     -0.31649941      2.0000000
 12     2      0.02334207      0.0000000
 13     3      0.02334207      0.0000000
 14     1      0.04980631      0.0000000
 15     5      0.09478404      0.0000000
 16     1      0.12395484      0.0000000
 17     3      0.13549605      0.0000000
 18     2      0.13549605      0.0000000
 19     5      0.28345574      0.0000000
 20     1      1.32404002      0.0000000
 21     5      1.45900204      0.0000000
=====

```

Assume that we wish to freeze the nine Ca inner shell orbitals, thus

```

CORE
1 TO 9 END
ACTIVE
10 TO 21 END

```

To perform a 4-electron valence-CI calculation, based on the SCF configuration would simply require the following CONF data:

```
CONF
2 2
```

The complete data file for performing the SCF and subsequent CI would then be as follows:

```
TIME 60
TITLE\CAH2 .. 3-21G
ZMAT ANGS\CA\X 1 1.0\ H 1 CAH 2 90.0\H 1 CAH 2 90.0 3 THETA
VARIABLES\CAH 2.148 \THETA 180.0 \END
BASIS 3-21G
RUNTYPE CI
CORE
1 TO 9 END
ACTIVE
10 TO 21 END
DIRECT 4 2 10
CONF
2 2
ENTER
```

## 6.15 CONF-CARDS

The directive consists of a single line, where the first 2 character strings are either CONF CARDS or CONF FILE or CONF ASCII . On the same line sub-directives may be specified,

- FILE string

A 44 character string may be supplied, specifying the file to read the configuration information from. The file must be as generated by the CARDS CASSCF directive i.e. fixed format like

```
1 222220000 .9644755897
```

- COEF value

Specify a minimum absolute value for a coefficient of a configuration to be included. Different spin-paths are combined.

- WEIGHT value

Specify the minimum weight of a configurations (including all its spin possibilities) to be included.

- DOC ndoc

Specify the number of doubly occupied orbitals not in the configurations in the list to be prepended to each configuration.

- NFRZ nfrz

Specify the number of orbitals to be frozen and kept doubly occupied. NFRZ and NDOC can cancel each other.

The reference configurations are printed, so the working of this directive is easily checked.



## 6.16 REFGEN

The first line should consist of the character string REFGEN in the first data field. Subsequent lines are read using I-format, to paired integers IA and IC. As many such lines as required may be presented. IA and IC define annihilation and creation operators respectively, which will be allowed to operate on the set of reference CSFs in existence at the time when REFGEN is called. Thus at least one CONF directive must have been presented before using REFGEN, and the IA ,IC integers refer to internal MOs. The result of the annihilation/creation process will be further reference CSFs, whose excitation mask will be that of the most recently issued EXCIT directive.

### Example 1

```
EXCIT OCTAL 7 3 1
CONF
2 2 0 0
EXCIT OCTAL 5
REFGEN
1 3
2 4
```

will result in a further two reference states, to give three in all, of the form:

CSFS	EXCITATION MASK	ORIGIN
=====	=====	=====
2 2 0 0	7 3 1	CONF directive
1 2 1 0	5	REFGEN directive
2 1 0 1	5	REFGEN directive

### Example 2

```
EXCIT OCTAL 7 3 1
CONF
2 2 0 0
EXCIT OCTAL 5
REFGEN
1 3 2 4
```

is equivalent to example 1; more than one IA/IC pair may be given on a single line.

### Example 3

```
EXCIT OCTAL 7 3 1
CONF
2 2 0 0
EXCIT OCTAL 5
REFGEN
1 3
EXCIT OCTAL 3
REFGEN
2 4
```

will produce 4 reference CSFs, of the form:

CSFS	EXCITATION MASK	ORIGIN
=====	=====	=====
2 2 0 0	7 3 1	CONF directive
1 2 1 0	5	First REFGEN directive
2 1 0 1	3	Second REFGEN directive
1 1 1 1	3	Second REFGEN directive

#### Example 4

```

EXCIT 7 3 1
CONF
2 2 0 0
REFGEN
1 3 1 4 2 3 2 4
REFGEN
1 3 1 4 2 3 2 4
REFGEN
1 3 1 4 2 3 2 4
REFGEN
1 3 1 4 2 3 2 4

```

will produce a reference space consisting of all possible CSFs that can be generated by distributing 4 electrons in 4 MOs.

### 6.17 CASGEN

The CASGEN directive is on 1 card and is meant to produce a selective Complete Active Space reference function. More than one CASGEN directives may be given. Each one working with respect to the previous one. The first word on the card is the textstring CASGEN (A).

The following may be specified on the same card :

- DOC NDOC (A,I) : Specifies number of orbitals to remain doubly occupied
- SDOC ND1 ND2 .. (A,I,I,I): As DOC but per symmetry; The number of integers is equal to the number of representations. DOC and SDOC are mutually exclusive
- NORB NORBC (A,I): Number of orbitals taken into account
- MAXEX MAX (A,I): Maximum excitation level with respect to the current reference set
- EXCIT (A): Use current excitation allowance for the newly generated configurations only, Often sensible if restrictions (like DOC) are employed. The default is the excitation mask 7b (no internal excitations), which is appropriate for a real CAS.
- NOSYM (A): Do not select the configurations on symmetry (default)
- SYM (A): Select configurations on symmetry
- SPIN (A): Select configurations on spin-symmetry

- PRINT (A): Print the generated reference configurations

To generate a 4-orbital casscf wavefunction for water, having the 3 inner orbitals doubly occupied in the reference function, and doing a Single-Double CI from this reference function one specifies

```
EXCIT OCTAL 7 3 1
CASGEN DOC 3 NORB 7 EXCIT
```

## 6.18 SCREEN

This directive consists of one line whose first data field should contain the character string SCREEN. If this directive is presented all reference configurations will be checked to see if they are of the same spin/space symmetry as that of the required CI wavefunction. If they are not, they will be eliminated from the reference space, and will take no part in the excitation process. Normally, this directive may only be required if a REFGEN directive is presented, since reference states of undesired properties will not usually be presented under control of the CONF directive.

## 6.19 RESTRICT

This directive provides a means of selectively eliminating CSFs from the CI expansion by specifying a minimum and maximum number of electrons which user specified sets of internal MOs may carry. The first line should contain the character string RESTRICT in the first data field. Subsequent lines are read to variables TEXT, MIN, MAX,(IORBS(i),i=1,m), using format (A,2I,mI). TEXT should be set to one of the character strings REF, VAC, N-1 or N-2. If REF is chosen the restrictions will apply to the reference CSFs; should such a state fail to comply with the restriction applied it will take no part in the excitation process used to generate the CI list, and will most probably be used when REFGEN directives are used. The character strings VAC, N-1 and N-2 refer to CSFs with 0,1 or 2 electrons in the external space respectively. MIN,MAX specify the minimum and maximum number of electrons to be allowed to populate the internal MOs defined by the IORBS parameters. IORBS A sequence of internal MO indices terminated by the integer 0. This data may be carried over to subsequent lines if necessary, and the character string TO may be used to shorten the data if desired. These parameters refer to the reordered MO list if a REORDER directive (see above) has been presented.

### Example

```
REFGEN
VAC 7 8 1 TO 4 0
N-1 7 8 1 TO 4 0
N-2 7 8 1 TO 4 0
```

Internal MOs 1 to 4 inclusive are allowed to carry either 7 or 8 electrons in vacuum, N-1 and N-2 CSFs.

## 6.20 PRCONF

This directive consists of one line, read to variables TEXT, IPR using format (A,I).

- TEXT should be set to the character string PRCONF.
- IPR specifies that every IPR'th occupation pattern generated by the configuration generator is to be printed. If IPR=1, all occupation patterns will be printed.

The directive may be omitted, when no occupation patterns generated by the configuration generator will be printed. The main use of the PRCONF directive is to generate a detailed occupation pattern listing.

### Example

```
PRCONF 1
```

## 6.21 CEPA

The CEPA directive allows the user to calculate the unlinked cluster correction to the final CI energy. This correction factor is more accurate than the Davidson correction factor, as it is done in an iterative way and the correction (shift) may be different for different n-2 states. Both the classical CEPA variants (0,1,2) for closed shell single determinant reference states and Multi-Reference variants are provided.

This directive consists of a single data line read to the variables TEXT, CEPA-variant (A,A) followed by optional options. TEXT should specify the character string CEPA and CEPA-variant is the text-string specifying which CEPA (or approximated Coupled Cluster) approach is requested. Options are:

- 0 : CEPA(0) (closed shell; shift is correlation energy) Cf. [3].
- 1 : CEPA(1) (closed shell; recommended;default) Cf. [3].
- 2 : CEPA(2) (closed shell; not invariant for mixing of occupied orbitals) Cf. [3].
- MR0 : Straight multi-reference variant of CEPA(0)
- ACPF : Averaged Coupled Pair Functional (Shift is modified correlation energy) [10] The correlation energy is (default) the difference with the (variationally determined) reference function.
- AQCC : Averaged Quadratic Coupled Cluster (Shift is modified correlation energy) [9] The correlation energy is (default) the difference with the (variationally determined) reference function.
- MRD : Multi Reference CEPA (taking Variationally Included (VI) terms into account) [7]

- MR1 : Multi Reference CEPA (taking VI and EPV terms into account) [11, 12] The EPV terms may be determined for the "inactive" space (DOC) or for all occupied orbitals (ALL; default); The choice may be made by specifying the additional string DOC or ALL.

Additional options, applicable to some or all of the variants mentioned above may be specified on the same card :

- For the approaches that use the (modified) correlation energy one may specify how this is calculated. Choices are PROJECT, where the expectation value of the "CI"-function projected onto the reference space is used as reference, VARIA, where the energy of the optimised reference space is employed or PAIRS, computing the correlation energy as sum of pair energies (this is ( ) project). The default is VARIA, which is Ahlrichs' choice.
- Using the keywords SIN or NOSIN, one may specify if single excitations should be shifted (SIN) or not. The default is that singles are shifted, except for single reference CEPA2. For the multi-reference variant, the option is dubious.
- Using the textstring IT one may specify after what iteration the CEPA mode may start, using IT ITCEPA (A,I). Default is 3.
- Using the textstring CRIT, one may specify at which particular threshold determined by the TESTER in the diagonalisation phase of the CI calculation the CEPA is switched on. The format is : CRIT CRITC (A,F). Default of CRITC is 0.01.
- PRINT requests intermediate printing within the CEPA mode.
- PAUL refers only to the CEPA 2 mode, and invokes an unpublished EPV correction formula due to P.J.A. Ruttinck.
- MICRO allows one to control the CEPA micro iterations, when the CI vector and correlation energy and shifts are updated without a matrix-vector product in between. The format is MICRO MCYC CRIT (A,I,F). MCYC is the maximum number of micro iterations and CRIT the relative convergence criterion. Defaults are 3 and 0.01.

The directive can be invoked without parameters, which will result in the following and crash for a multi-reference case:

```
CEPA 1 SIN IT 3 CRIT 0.01 MICRO 3 0.01
```

## 6.22 MP

The MP directive allows one to perform multi-reference Møller-Plesset calculations [4, 5, 6]. To generate the reference wavefunction one should run an MCSCF calculation on the required state first to obtain the correct orbitals and use the TRIAL DIAG directive to rebuild the MCSCF wavefunction. Details on this process will be given in an example in the Direct-CI - Multi-reference MP section of Part 2.

Once the reference wavefunction has been constructed the MP directive should be used to control the perturbation theory applied to it. The syntax is

```
TEXT [ TEXTA ] [ TEXTB TEXTC or IMODEL ] [ TEXTD M E ]
```

where

- TEXT should specify the character string MP
- TEXTA specifies the order of perturbation theory required. One can choose from '2' and '3'.
- TEXTB should specify the character string MODEL to choose the form of the zeroth order Hamiltonian. This Hamiltonian can be specified by name using
  - 'RUTTINK' to select  $H_{S,D}^{(0)-minimal}$  in [4] in honour of his introduction of the excitation classes in MRCEPA [7]
  - 'PULAY' to select  $H_{S,D}^{(0)}$  in [4] in honour of his work on MRMP methods [6]
  - 'ANDERSSON' to select the Pulay zeroth order Hamiltonian but removing the excitations within the reference space. The name was chosen in honour of her work on CASPT2 [8]

Alternatively the model may be specified by a number where

- the units stand for; 1 using a projector operator on the combined single and double excitation space, 2 same as 1 but eliminating all single excitations, 3 using projector operators onto the space of single and doubly excitations separately
- the decades stand for; 0 no modification, 1 removing all excitations within the reference space
- the sign stands for; + no modification, - removing all parts of the zeroth order Hamiltonian that connect different excitation classes.

i.e. 'RUTTINK' can also be entered as -1, 'PULAY' as 1, and 'ANDERSSON' as 11.

- TEXTD can be used to specify algorithm to orthogonalise the set of single and double excitations. One can choose from 'SCHMIDT' for modified Gramm-Schmidt orthogonalisation, 'LOWDIN' for orthogonalising by diagonalisation of the overlap matrix, or 'HOUSEHOLDER' for applying the House-Holder method. The floating point number M and the integer E specify the orthogonalisation accuracy  $M * 10^{-E}$ .

The default settings can be specified by either of the 2 following

```
MP
MP 2 RUTTINK HOUSEHOLDER 1. 8
```

## 6.23 Spin Functions

It may be necessary for the user to understand the nature and order of spin functions associated with a given occupation pattern. The program makes use of Yamanouchi-Kotani genealogical spin functions, the coupling order being such that higher indexed MOs are coupled before

lower indexed MOs. The ordering of the MOs is as defined using the ACTIVE directive of the transformation module (see above). Use the digits 0 and 1 to denote down and up spin coupling respectively. Proceeding from the highest to the lowest indexed singly occupied MO, write down the digitized representation of the possible spin functions, the digits being written from left to right. The resultant binary number defines the lexical ordering of the members of the spin canonical set, the higher the number, the higher the lexical index.

### Example

Consider 5 doubly occupied MOs coupled to a doublet. The possible spin functions in digitized representation are, in order of increasing lexical index:

- 10101 - Spin function 1
- 10110 - Spin function 2
- 11001 - Spin function 3
- 11010 - Spin function 4
- 11100 - Spin function 5

## 6.24 Using GVB Orbitals

In this section we briefly outline specification of a multi-reference direct-CI based on a GVB-1/PP wavefunction. Consider again the example of section 4.4.2, where a 4-pair GVB/PP calculation on H<sub>2</sub>CO is described, in which the two C-H bonds and two C-O orbitals are treated within the perfect pairing approximation. The sequence of calculations included:

- performing the closed shell SCF calculation;
- localising the set of valence SCF orbitals;
- performing the GVB calculation using the set of LMO input under control of the VECTORS option NOGEN.

An examination of the GVB output reveals the following orbital assignments, with orbitals 5, 7, 9 and 11 corresponding to the strongly occupied orbitals, and orbitals 6, 8, 10 and 12 to the weakly occupied orbitals of the GVB pairs.

M.O.	IRREP	ORBITAL ENERGY	ORBITAL OCCUPANCY
1	1	-20.48204464	2.0000000
2	1	-11.25090140	2.0000000
3	1	-1.12106354	2.0000000
4	1	-0.55607328	2.0000000

5	1	-1.42285552	1.9843054
6	1	-0.02615928	0.0156946
7	1	-1.42285519	1.9843052
8	1	-0.02615932	0.0156948
9	1	-1.93008469	1.9904691
10	1	-0.02325080	0.0095309
11	1	-1.07958389	1.9032736
12	1	-0.09164423	0.0967264
13	1	0.62048723	0.0000000
14	1	0.87018377	0.0000000
15	1	0.87310154	0.0000000
16	1	0.92519148	0.0000000
17	1	1.03015198	0.0000000
18	1	1.38072957	0.0000000
19	1	1.79683989	0.0000000
20	1	1.81440429	0.0000000
21	1	1.97615975	0.0000000
22	1	3.26852884	0.0000000

=====

Assume that we wish to freeze the O1s and C1s orbitals, thus

```
CORE\1 2 \END
ACTIVE\3 TO 22\END
```

To perform a 12-electron valence-CI calculation based on the leading term in the GVB expansion, together with those doubly excited configurations corresponding to each GVB pair, would require the following CONF data:

```
CONF
2 2 2 0 2 0 2 0 2 0
2 2 0 2 2 0 2 0 2 0
2 2 2 0 0 2 2 0 2 0
2 2 2 0 2 0 0 2 2 0
2 2 2 0 2 0 2 0 0 2
```

with 10 orbitals in the internal space and 10 in the external space, given that the inner shell orbitals have been frozen. The complete data file for performing the GVB-CI would then appear as follows, where the canonicalised GVB orbitals are restored from the default section, section 5, of the Dumpfile.

```
RESTART NEW
TITLE
H2CO - 3-21G 4PAIR GVB
SUPER OFF NOSYM
BYPASS SCF
ADAPT OFF
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
SCFTYPE GVB 4
CORE\1 2 \END
ACTIVE\3 TO 22\END
DIRECT 12 10 10
CONF
2 2 2 0 2 0 2 0 2 0
2 2 0 2 2 0 2 0 2 0
2 2 2 0 0 2 2 0 2 0
2 2 2 0 2 0 0 2 2 0
2 2 2 0 2 0 2 0 0 2
ENTER

```

We show below the output from the resulting CI calculation.

```

TOTAL ENERGY -113.4467990858

0.95233628      1  VACUUM    0  0      1  -- 2220202020
-0.08236057     1  VACUUM    0  0      2  -- 2202202020
-0.08236123     1  VACUUM    0  0      3  -- 2220022020
-0.05407166     1  VACUUM    0  0      4  -- 2220200220
-0.16888657     1  VACUUM    0  0      5  -- 2220202002
-0.03606575     1  VACUUM    0  0     163 -- 2121202011
-0.03606465     1  VACUUM    0  0     224 -- 2120212011
-0.07796058     1  VACUUM    0  0     279 -- 2220201111

```

```

SUM OF SQUARES OF VACUUM CSF CI COEFFICIENTS= 0.973137161834E+00
SUM OF SQUARES OF      5-MAIN CI COEFFICIENTS= 0.951957435390E+00

```

## 6.25 FP Geometry Optimisations

Energy-only optimisation for direct-CI wavefunctions may be performed using a variant of the RUNTYPE OPTIMIZE directive. The data line

```
RUNTYPE OPTIMIZE CI
```

requests use of the Fletcher Powell (FP) optimiser, with subsequent data used to characterize the direct-CI wavefunction to be employed in the energy calculation. We illustrate such usage below for the case of a direct-CI calculation on the H<sub>2</sub>CO cation, performing the calculation in several steps. The first two steps carry out an RHF open shell geometry optimisation, and the third the corresponding CI optimisation.

### Runs I and II: The SCF Optimisation

```

TITLE
H2CO - TZVP - CLOSED SHELL STARTUP
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 TZVP
ENTER

```

The first step is merely used to generate a suitable set of MOS for initiating the SCF geometry optimisation on the ion below.

```

RESTART NEW
TITLE
H2CO+ - TZVP - GEOMETRY OPTIMISATION SCF
MULT 2
CHARGE 1
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 TZVP
RUNTYPE OPTIMIZE
OPEN 1 1
ENTER

```

### Run III: The Direct-CI Calculation

```

RESTART NEW
TITLE
H2CO+ - TZVP - CI/1M AT RHF GEOMETRY
MULT 2
CHARGE 1
ZMATRIX ANGSTROM
C
O 1 CO
H 1 CH 2 HCO
H 1 CH 2 HCO 3 180.0
VARIABLES
CO      1.2063534 HESS      .792378
CH      1.0876643 HESS      .681619
HCO     117.8503752 HESS     .709462
END
BASIS TZVP
RUNTYPE OPTIMIZE CI
OPEN 1 1

```

```
DIRECT 15 8 34
CONF
2 2 2 2 2 2 1
ENTER
```

Note again that some care must be taken when reducing the orbital space in FP CI optimisations. In open shell calculations, the CI step will derive the orbital set at each point from the second section specified on the ENTER directive i.e., the energy ordered MOs. If this ordering varies from point to point in the FP optimisation, and symmetry is used in minimising the configuration space, it is quite likely that this space will vary during successive points, with disastrous consequences on the optimisation pathway. As a general rule, the user should only consider freezing or discarding orbitals that are well separated from those MOs included in the CI space i.e. inner shell or inner-shell complement MOs.

## 6.26 Calculating the $^1A_1$ states of $H_2O$

To conclude our discussion of the Direct-CI module, we work through a typical example of using the Direct-CI method in calculating the energetics and properties of the three low lying  $^1A_1$  states of the  $H_2O$  molecule. The basis set employed is the TZVP triple-zeta plus polarisation set; this is augmented with a diffuse s- and p-orbital on the oxygen to provide a reasonable description of the known Rydberg character of the states of interest. The computation is split into a number of separate jobs, in which we

1. perform the initial SCF;
2. carry out an initial CI for each state, where the reference set employed acts to provide at least a qualitative description of the states of interest;
3. based on the output from the initial CIs, we augment the reference set to provide a quantitative description of the first three states.

We now consider various aspects of each job in turn.

### Job 1: The SCF

```
TITLE
**** H2O TZVP + DIFFUSE S,P SCF *
SUPER OFF NOSYM
ZMAT ANGSTROM
O
H 1 0.951
H 1 0.951 2 104.5
END
BASIS
TZVP O
TZVP H
S 0
1.0 0.02
P 0
```

```

1.0 0.02
END
ENTER

```

The only point to note here is the use of the SUPER directive in suppressing skeletonisation, given the user wishes to access this file in a subsequent CI step under BYPASS control.

### Jobs 2-4: The Initial CIs

An examination of the SCF output reveals the following orbital analysis.

```

=====
IRREP  NO. OF SYMMETRY ADAPTED
        BASIS FUNCTIONS
=====
   1         18
   2          6
   3         10
   4          2
=====

```

and the following orbital assignments characterising the closed-shell SCF configuration:

$$1a_1^2 2a_1^2 1b_2^2 3a_1^2 1b_1^2 \quad (15)$$

```

=====
M.O.  IRREP  ORBITAL ENERGY  ORBITAL OCCUPANCY
=====
   1     1    -20.56084959      2.0000000
   2     1     -1.35696939      2.0000000
   3     3     -0.72200122      2.0000000
   4     1     -0.58247942      2.0000000
   5     2     -0.50858566      2.0000000
   6     1      0.02724259      0.0000000
   7     3      0.04894440      0.0000000
   8     2      0.05589681      0.0000000
   9     1      0.06133571      0.0000000
  10     1      0.20403420      0.0000000
  11     3      0.22824210      0.0000000
  12     3      0.53700802      0.0000000
  13     1      0.56235022      0.0000000
  14     2      0.58645643      0.0000000
  15     1      0.66887228      0.0000000
  16     3      0.74805617      0.0000000
  17     1      1.07690608      0.0000000
  18     1      1.88545053      0.0000000
  19     4      1.92243836      0.0000000
  20     2      2.12944874      0.0000000
  21     3      2.20541910      0.0000000
  22     1      2.34202871      0.0000000
  23     3      2.39946430      0.0000000
  24     3      2.69788310      0.0000000

```

25	1	2.72651832	0.0000000
26	2	2.73832720	0.0000000
27	1	3.07664215	0.0000000
28	3	3.26840142	0.0000000
29	2	3.54616570	0.0000000
30	1	3.58631019	0.0000000
31	4	3.59701772	0.0000000
32	1	3.84174131	0.0000000
33	1	4.84610143	0.0000000
34	3	5.14220270	0.0000000
35	1	7.73115986	0.0000000
36	1	47.56758932	0.0000000

=====

Assuming that we wish to freeze the O1s inner shell orbitals and discard the inner shell complement orbital, the following data lines should be presented in the transformation:

```
CORE\1\END
ACTIVE\2 TO 35\END
```

Note that the virtual SCF MOs dominated by the diffuse oxygen basis functions are the  $4a_1$ , the  $2b_2$ , the  $2b_1$  and the  $5a_1$ , with SCF sequence numbers 6,7,8 and 9 respectively. The re-ordered sequence numbers, allowing for the effective removal of the two  $a_1$  orbitals, are 5,6,7 and 8 respectively. To perform a balanced valence-CI treatment of the three states of interest will require a three-root 8-electron reference set, based on the SCF configurations of the ground and excited Rydberg states, involving the single excitations ( $1b_1$  to  $2b_1$ ) and ( $3a_1$  to  $4a_1$ ). This specification will require the following CONF data:

```
CONF
2 2 2 2 0 0 0 0
2 2 2 1 0 0 1 0
2 2 1 2 1 0 0 0
```

In contrast to the Table-CI module (see Part 6), where the three-state CI may be performed in a single job, the Direct-CI treatment will require three separate jobs, each job looking to describe a specific state. This is achieved in the case of the excited states through use of the TRIAL directive, which identifies the particular state under investigation. The data files for these three jobs are given below:

### Job 2: Direct-CI Treatment of the $X^1A_1$

```
RESTART NEW
TITLE
**** H2O TZVP + DIFFUSE S,P DIRECT-CI 3M X1A1*
SUPER OFF NOSYM
BYPASS SCF
ZMAT ANGSTROM
0
H 1 0.951
```

```

H 1 0.951 2 104.5
END
BASIS
TZVP 0
TZVP H
S 0
1.0 0.02
P 0
1.0 0.02
END
RUNTYPE CI
CORE\1\END
ACTIVE\2 TO 35\END
DIRECT 8 8 26
CONF
2 2 2 0 0 0 0
2 2 2 1 0 0 1 0
2 2 1 2 1 0 0 0
ENTER

```

The following points should be noted:

- the SCF computation is BYPASS'ed;
- the CORE and ACTIVE directives act to freeze and discard the two  $a_1$  MOs;
- the parameters on the DIRECT data line reflect the number of active electrons (8), number of internal orbitals (8) and the number of external orbitals (26);
- in this job we are describing the  $X^1A_1$ , so that the default diagonalisation controls will prove satisfactory.

### Job 3: Direct-CI Treatment of the $1^1A_1$ state

```

RESTART CI
TITLE
**** H2O TZVP + DIFFUSE S,P DIRECT-CI 3M 1A1*
SUPER OFF NOSYM
BYPASS TRANSFORM
ZMAT ANGSTROM
0
H 1 0.951
H 1 0.951 2 104.5
END
BASIS
TZVP 0
TZVP H
S 0
1.0 0.02
P 0
1.0 0.02
END
RUNTYPE CI
CORE\1\END

```

```

ACTIVE\2 TO 35\END
DIRECT 8 8 26
CONF
2 2 2 2 0 0 0 0
2 2 2 1 0 0 1 0
2 2 1 2 1 0 0 0
SHIFT 0.5\ALTERNATE
TRIAL
2 1.0
VMIN
ENTER

```

The following points should be noted:

- the SCF and Transformation are BYPASS'ed: this assumes that the Transformed Integral file from the initial CI has been saved;
- in this job we are describing an excited state, the  $1^1A_1$ , so that the default diagonalisation controls will no longer prove satisfactory. The specification of SHIFT, ALTERNATE and VMIN is typical in such calculations. The TRIAL directive is now specifying a starting vector with the second configuration in the CONF list as the dominant term.

#### Job 4: Direct-CI Treatment of the $2^1A_1$ state

```

RESTART CI
TITLE
**** H2O TZVP + DIFFUSE S,P DIRECT-CI 3M 2A1*
SUPER OFF NOSYM
BYPASS TRANSFORM
ZMAT ANGSTROM
O
H 1 0.951
H 1 0.951 2 104.5
END
BASIS
TZVP O
TZVP H
S 0
1.0 0.02
P 0
1.0 0.02
END
RUNTYPE CI
CORE\1\END
ACTIVE\2 TO 35\END
DIRECT 8 8 26
CONF
2 2 2 2 0 0 0 0
2 2 2 1 0 0 1 0
2 2 1 2 1 0 0 0
TRIAL
3 1.0
SHIFT 0.5\ALTERNATE

```

```

VMIN
ENTER

```

The following points should be noted:

- the SCF and Transformation are again BYPASS'ed;
- we are again describing an excited state, the  $2^1A_1$ , so that SHIFT, ALTERNATE and VMIN are again specified. The TRIAL directive is now specifying a starting vector with the third configuration in the CONF list as the dominant term.

### The Final 16-Reference CI Jobs

An examination of the output from the initial CI calculations reveals that the dominant configurations have, as expected, been included. We show below the final CI vectors for each of the states: not surprisingly the ground state is more accurate, by virtue of its SCF MOs having been employed. Augmenting the reference set to improve the description of the two excited states follows straightforwardly from the statistics below:

#### Description of the $X^1A_1$ state

```

TOTAL ENERGY -76.2725934815

*****
COEFFICIENT INTERNAL EXTERNAL EXTERNAL INTERNAL
                SPIN    SPIN    MOS    CONF  -- OCC.
*****

SPIN-COUPLING REFERS TO REORDERED ORBITALS
*****

0.97419222      1  VACUUM    0  0      1  -- 22220000
-0.03737671     1  SINGLET   13 13     201 -- 22200000
-0.03086972     1  SINGLET   10 15     202 -- 20220000
-0.03143494     1  SINGLET   15 15     202
0.03223099      1  SINGLET   14 15     218 -- 21120000

SUM OF SQUARES OF VACUUM CSF CI COEFFICIENTS= 0.949366016991E+00
SUM OF SQUARES OF      3-MAIN CI COEFFICIENTS= 0.949241228068E+00

```

#### Description of the $1^1A_1$ state

```

TOTAL ENERGY -75.9018772961

*****
COEFFICIENT INTERNAL EXTERNAL EXTERNAL INTERNAL
                SPIN    SPIN    MOS    CONF  -- OCC.
*****

SPIN-COUPLING REFERS TO REORDERED ORBITALS
*****

```



-0.91631377	1	VACUUM	0	0	2	--	22210010
-0.25186711	1	VACUUM	0	0	3	--	22121000
-0.05309578	1	VACUUM	0	0	12	--	22200020
-0.08694410	1	DOUBLET	0	9	87	--	22120000
0.03713591	1	DOUBLET	0	9	88	--	12210010
0.04438742	1	DOUBLET	0	12	89	--	22110010
-0.04467943	1	DOUBLET	0	14	89		
0.03994355	2	DOUBLET	0	9	89		
-0.04976149	2	DOUBLET	0	12	89		
0.05925473	2	DOUBLET	0	14	89		
0.05097226	1	DOUBLET	0	13	124	--	22210000
0.10693810	1	DOUBLET	0	13	125	--	22200010
0.08035163	1	DOUBLET	0	10	153	--	21210010
0.07987526	1	DOUBLET	0	15	153		

SUM OF SQUARES OF VACUUM CSF CI COEFFICIENTS= 0.909083899548E+00  
 SUM OF SQUARES OF 3-MAIN CI COEFFICIENTS= 0.903281844886E+00

### Description of the $2^1A_1$ state

TOTAL ENERGY -75.8837522007

\*\*\*\*\*  
 COEFFICIENT INTERNAL EXTERNAL EXTERNAL INTERNAL  
 SPIN SPIN MOS CONF -- OCC.  
 \*\*\*\*\*

SPIN-COUPLING REFERS TO REORDERED ORBITALS

\*\*\*\*\*  

0.26477982	1	VACUUM	0	0	2	--	22210010
-0.88348347	1	VACUUM	0	0	3	--	22121000
-0.08640544	1	VACUUM	0	0	6	--	22120001
-0.04120749	1	VACUUM	0	0	9	--	22111010
0.03104762	1	VACUUM	0	0	15	--	22022000
-0.23849192	1	DOUBLET	0	9	87	--	22120000
0.04891462	1	DOUBLET	0	9	91	--	22021000
-0.05841301	1	DOUBLET	0	12	91		
0.07064985	1	DOUBLET	0	14	91		
-0.03000718	1	DOUBLET	0	13	125	--	22200010
0.09148369	1	DOUBLET	0	13	126	--	22111000
0.07828915	1	DOUBLET	0	10	154	--	21121000
0.07660116	1	DOUBLET	0	15	154		

SUM OF SQUARES OF VACUUM CSF CI COEFFICIENTS= 0.863430795611E+00  
 SUM OF SQUARES OF 3-MAIN CI COEFFICIENTS= 0.850807372991E+00

Taking as the criterion for inclusion a coefficient of 0.05, the final reference set to be employed is constructed based on both

- the appropriate external MOs in the above coefficient lists, which now must be assigned internal orbital status.
- the internal configurations specified in the ci-vector output above.

This information is now provided directly in printing the final CI-vector; it may also be derived from the print of the occupation patterns, assuming the data line PRCONF 1 had been presented in the 3-reference job.

```

*** VACUUM STATES (# CONF      85 # STATES      127) ***
    1  (#      1  )  2 2 2 2 0 0 0 0
    2  (#      1  )  2 2 2 1 0 0 1 0
    3  (#      1  )  2 2 1 2 1 0 0 0
    6  (#      1  )  2 2 1 2 0 0 0 1
   12  (#      1  )  2 2 2 0 0 0 2 0

*** DOUBLET STATES (# CONF     112 # STATES     1738) ***

SYM  1  ONE      # CONF      38  EXTERNAL DIMENSION      12
    87  (#      1  )  2 2 1 2 0 0 0 0
    89  (#      2  )  2 2 1 1 0 0 1 0
    90  (#      2  )  1 2 1 2 1 0 0 0
    91  (#      1  )  2 2 0 2 1 0 0 0

SYM  2  TWO      # CONF      28  EXTERNAL DIMENSION      4
   124  (#      1  )  2 2 2 1 0 0 0 0
   125  (#      1  )  2 2 2 0 0 0 1 0
   126  (#      2  )  2 2 1 1 1 0 0 0

SYM  3  THREE    # CONF      28  EXTERNAL DIMENSION      8
   153  (#      2  )  2 1 2 1 0 0 1 0
   154  (#      2  )  2 1 1 2 1 0 0 0

```

The final 16-reference-CI jobs are shown below. Since the ordering of the orbitals has now changed, we repeat the transformation with a revised ACTIVE list designed to incorporate the external orbitals referenced in the coefficient lists above into the internal space. Note that the external MO indexing within the CI module does not take into account the inner-shell frozen in the transformation: adding one to the indices referenced in the CI will provide the integers for specification in the revised ACTIVE data i.e., 9,10,11,13,14,15 and 16, thus

```

ACTIVE\2 TO 11 13 TO 16
12 17 TO 35\END

```

### Final-CI Treatment of the $X^1A_1$

```

RESTART NEW
TITLE
**** H2O TZVP + DIFFUSE S,P DIRECT-CI 16M X1A1*
SUPER OFF NOSYM
BYPASS SCF
ZMAT ANGSTROM
0
H 1 0.951
H 1 0.951 2 104.5
END

```

```

BASIS
TZVP 0
TZVP H
S 0
1.0 0.02
P 0
1.0 0.02
END
RUNTYPE CI
CORE\1\END
ACTIVE\2 TO 11 13 TO 16
12 17 TO 35\END
DIRECT 8 14 20
CONF
2 2 2 2 0 0 0 0 0 0 0 0 0 0 0
2 2 2 1 0 0 1 0 0 0 0 0 0 0 0
2 2 1 2 1 0 0 0 0 0 0 0 0 0 0
2 2 1 2 0 0 0 1 0 0 0 0 0 0 0
2 2 2 0 0 0 2 0 0 0 0 0 0 0 0
2 2 1 2 0 0 0 0 1 0 0 0 0 0 0
2 2 1 1 0 0 1 0 0 0 0 0 0 1 0
2 2 0 2 1 0 0 0 0 0 1 0 0 0 0
2 2 0 2 1 0 0 0 0 0 0 0 0 1 0
2 2 2 1 0 0 0 0 0 0 0 0 1 0 0
2 2 2 0 0 0 1 0 0 0 0 0 1 0 0
2 2 1 1 1 0 0 0 0 0 0 0 1 0 0
2 1 2 1 0 0 1 0 0 1 0 0 0 0 0
2 1 2 1 0 0 1 0 0 0 0 0 0 0 1
2 1 1 2 1 0 0 0 0 1 0 0 0 0 0
2 1 1 2 1 0 0 0 0 0 0 0 0 0 1
ENTER

```

The following points should be noted:

- There are now 14 orbitals in the internal space and 20 in the external space.
- We again assume that the Transformed Integral File is retained from the above job, enabling the integral transformation to be bypassed in the jobs below.
- We show below the data file for the final CI on the  $1^1A_1$  state: that on the third state follows in straightforward fashion.

### Final-CI Treatment of the $1^1A_1$ state

```

RESTART CI
TITLE
**** H2O TZVP + DIFFUSE S,P DIRECT-CI 16M 1A1*
SUPER OFF NOSYM
BYPASS TRANSFORM
ZMAT ANGSTROM
0
H 1 0.951
H 1 0.951 2 104.5
END

```

```
BASIS
TZVP 0
TZVP H
S 0
1.0 0.02
P 0
1.0 0.02
END
RUNTYPE CI
CORE\1\END
ACTIVE\2 TO 11 13 TO 16
12 17 TO 35\END
DIRECT 8 14 20
CONF
2 2 2 2 0 0 0 0 0 0 0 0 0 0 0
2 2 2 1 0 0 1 0 0 0 0 0 0 0 0
2 2 1 2 1 0 0 0 0 0 0 0 0 0 0
2 2 1 2 0 0 0 1 0 0 0 0 0 0 0
2 2 2 0 0 0 2 0 0 0 0 0 0 0 0
2 2 1 2 0 0 0 0 1 0 0 0 0 0 0
2 2 1 1 0 0 1 0 0 0 0 0 0 1 0
2 2 0 2 1 0 0 0 0 0 1 0 0 0 0
2 2 0 2 1 0 0 0 0 0 0 0 0 1 0
2 2 2 1 0 0 0 0 0 0 0 0 1 0 0
2 2 2 0 0 0 1 0 0 0 0 0 1 0 0
2 2 1 1 1 0 0 0 0 0 0 0 1 0 0
2 1 2 1 0 0 1 0 0 1 0 0 0 0 0
2 1 2 1 0 0 1 0 0 0 0 0 0 0 1
2 1 1 2 1 0 0 0 0 1 0 0 0 0 0
2 1 1 2 1 0 0 0 0 0 0 0 0 0 1
SHIFT 0.5\ALTERNATE
TRIAL
2 1.0
VMIN
ENTER
```

## References

- [1] P.E.M. Siegbahn, *J. Chem. Phys.* **72** (1980) 1647, doi:[10.1063/1.439365](https://doi.org/10.1063/1.439365); V.R. Saunders and J.H. van Lenthe, *Mol. Phys.* **48** (1983) 923, doi:[10.1080/00268978300100661](https://doi.org/10.1080/00268978300100661).
- [2] M. Yoshimine, *J. Comp. Phys.* **11** (1973) 449; P.S. Bagus, B. Liu, A.D. McLean and M. Yoshimine, *Energy, Structure and Reactivity*, edited by D.W. Smith and W.B. McRae, (Wiley) (1973) 130; G.H.F. Diercksen, *Theor. Chim. Acta* **33** (1974) 1, doi:[10.1007/BF00527620](https://doi.org/10.1007/BF00527620); S.T. Elbert, *Numerical Algorithms in Chemistry : Algebraic Methods-LBL 8158*, edited by C. Moler and I. Shavitt, (Lawrence Berkeley Laboratory, University of California, Berkeley) (1978) 1298
- [3] C. Zirz and R. Ahlrichs, in 'Electron Correlation: Proceedings of the Daresbury Study Weekend', eds. M.F. Guest and S. Wilson (Daresbury Laboratory Report DL/SCI/R14) (1980) 83.
- [4] H.J.J. van Dam, J.H. van Lenthe, and P.J.A. Ruttink, 'Exact size consistency of multi-reference Møller-Plesset perturbation theory', *Int. J. Quant. Chem.* **72** (1999) 549–558.
- [5] H.J.J. van Dam, J.H. van Lenthe, and P. Pulay, 'The size consistency of multi-reference Møller-Plesset perturbation theory', *Mol. Phys.* **93** (1998) 431–439, doi:[10.1080/002689798169122](https://doi.org/10.1080/002689798169122).
- [6] K. Wolinski, H.L. Sellers, and P. Pulay, *Chem. Phys. Lett.* **140** (1987) 225, doi:[10.1016/0009-2614\(87\)80448-7](https://doi.org/10.1016/0009-2614(87)80448-7).
- [7] P.J.A. Ruttink, J.H. van Lenthe, R. Zwaans, and G.C. Groenenboom, *J. Chem. Phys.* **94** (1991) 7212–7220, doi:[10.1063/1.460204](https://doi.org/10.1063/1.460204).
- [8] K. Andersson, P.-Å. Malmqvist, and B.O. Roos, *J. Chem. Phys.* **96** (1992) 1218, doi:[10.1063/1.462209](https://doi.org/10.1063/1.462209).
- [9] P.G. Szalay, R.J. Bartlett, *Chem. Phys. Lett.* **214** (1993) 481, doi:[10.1016/0009-2614\(93\)85670-J](https://doi.org/10.1016/0009-2614(93)85670-J).
- [10] R. Gdanitz, R. Ahlrichs, *Chem. Phys. Lett.* **143** (1988) 413, doi:[10.1016/0009-2614\(88\)87388-3](https://doi.org/10.1016/0009-2614(88)87388-3).
- [11] L. Fusti-Molnar, P.G. Szalay, *J. Phys. Chem.* **100** (1996) 6288, doi:[10.1021/jp952840j](https://doi.org/10.1021/jp952840j).
- [12] P.J.A. Ruttink, J.H. van Lenthe, P. Todorov, *Mol. Phys.* **103** (2005) 2497, doi:[10.1080/00268970500180725](https://doi.org/10.1080/00268970500180725).