

# **AMPAC**

**Version 2.1**

**A GENERAL MOLECULAR ORBITAL PACKAGE**

*A Product of  
The Dewar Research Group  
University of Texas at Austin*



# **AMPAC MANUAL**

**Version 2.1**

## **A GENERAL MOLECULAR ORBITAL PACKAGE**

**Daniel A. Liotard, Eamonn F. Healy, James M. Ruiz, Michael J. S. Dewar**

*A Product of  
The Dewar Research Group  
University of Texas at Austin*

**Manual Written and Edited by:  
Roy D. Dennington, II & Eamonn F. Healy**

**Copyright (c) 1989 University of Texas at Austin**

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



## TABLE OF CONTENTS

<b>TABLE OF CONTENTS.....</b>	<b>i</b>
<b>PREFACE.....</b>	<b>iii</b>
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>1</b>
1.1 SUMMARY OF AMPAC CAPABILITIES.....	1
1.2 LAYOUT OF DATA.....	1
Example Data File for D2d Ethylene.....	2
Example Data File for Polytetrahydrofuran.....	3
<b>CHAPTER 2: KEYWORDS.....</b>	<b>5</b>
2.1 SPECIFICATION OF KEYWORDS.....	5
2.2 FULL LIST OF KEYWORDS.....	5
2.3 DEFINITIONS OF KEYWORDS.....	7
<b>CHAPTER 3: GEOMETRY SPECIFICATION.....</b>	<b>27</b>
3.1 INTERNAL COORDINATE DEFINITION.....	27
3.2 CARTESIAN COORDINATE DEFINITION.....	28
3.3 DEFINITION OF ELEMENTS AND ISOTOPES.....	29
3.4 EXAMPLES OF COORDINATE DEFINITIONS.....	30
<b>CHAPTER 4: EXAMPLES.....</b>	<b>33</b>
4.1 MNRS1 TEST DATA FILE FOR FORMALDEHYDE.....	33
4.2 OUTPUT FILE FOR TEST DATA FILE MNRS1.....	34
NOTES ON OUTPUT FILE.....	36
4.3 DATA FILE FOR A FORCE CALCULATION.....	38
4.4 OUTPUT FILE FROM A FORCE CALCULATION.....	38
NOTES ON FORCE OUTPUT FILE.....	43
4.5 DATA FILE FOR A REACTION PATH WITH SYMMETRY.....	45
4.6 DATA FILE FOR A SADDLE CALCULATION.....	46
4.7 DATA FILE FOR A CHAIN CALCULATION.....	46
4.8 OUTPUT FILE FROM A CHAIN CALCULATION.....	47
NOTES ON OUTPUT FILE.....	50
<b>CHAPTER 5: BACKGROUND.....</b>	<b>53</b>
5.1 THE SELF CONSISTENT FIELD PROCEDURE.....	53
Criteria and Accuracy in SCF Convergence.....	53
Convergence Techniques.....	53
Failure to Achieve Self-Consistence.....	54
5.2 GEOMETRY OPTIMIZATION.....	54
5.3 TORSION ANGLE COHERENCY.....	55
5.4 PSEUDODIAGONALIZATION-- SUBROUTINE DIAG.....	56
Theory.....	56
Computational Aspects.....	56
5.5 CONFIGURATION INTERACTION.....	57
Reference Electronic Configuration.....	57
Removal of Electrons.....	58
Formation of Microstate Configurations.....	58
Multi Electron Configuration Interaction.....	60
States Arising from Various Calculations.....	61
Calculation of Spin States.....	62
Choice of State to be Optimized.....	62
Calculation of Unpaired Spin Density.....	62
5.6 REDUCED MASSES IN A FORCE CALCULATION.....	63

5.7	FRAME IN THE FORCE CALCULATION .....	63
5.8	VIBRATIONAL ANALYSIS .....	63
5.9	DYNAMIC REACTION COORDINATE .....	64
	Equations Used .....	64
5.10	POLARIZABILITY CALCULATION.....	65
5.11	SOLID STATE CAPABILITY .....	67
5.12	SPARKLES.....	68
	Behavior of Sparkles in AMPAC .....	68
	Uses of Sparkles .....	68
5.13	INTEGRATION OF A REACTION PATH IN PATH.....	69
	Differential Equations of a Path.....	69
	The Choice of a Numerical Integrator.....	69
5.14	USE OF SADDLE CALCULATION.....	70
5.15	TRANSITION STATE BY THE CHAIN METHOD.....	71
5.16	ANALYTICAL GRADIENT COMPUTATION.....	72
<b>CHAPTER 6:</b>	<b>INSTALLATION, DEBUGGING AND PROGRAMMING.....</b>	<b>73</b>
6.1	INSTALLING AMPAC ON A VAX/VMS SYSTEM.....	73
6.2	DESCRIPTION OF THE COMMAND FILES.....	74
	COMPILE.COM .....	74
	AMPAC.COM.....	75
	RAMPAC.COM.....	75
6.3	RUNNING AMPAC.....	75
6.4	SHUTDOWN COMMAND .....	76
6.5	DEBUGGING.....	76
	List of KEYWORDS for Debugging Subroutines.....	76
	Suggested Procedure for Locating Bugs.....	77
6.6	DIFFERENCES FROM AMPAC VERSION 1.0.....	78
	Criteria and Accuracy .....	78
	Optimization Package .....	79
	New KEYWORDS .....	80
6.7	AMPAC PROGRAM FLOW DIAGRAM.....	81
6.8	FORTRAN-77 FILES.....	82
6.9	SUBROUTINE CALLS .....	82
6.10	DESCRIPTION OF MAJOR SUBROUTINES.....	85
<b>APPENDIX A:</b>	<b>ERROR MESSAGES .....</b>	<b>93</b>
<b>APPENDIX B:</b>	<b>REFERENCES.....</b>	<b>103</b>
<b>INDEX.....</b>		<b>107</b>

## PREFACE

The AMPAC program has proved, and is proving, of general value to chemists. It provides a simple and very effective means for carrying out calculations using the semiempirical procedures that have been developed by our group, in particular AM1, and it requires less computing time than its only rival (MOPAC).

Here we present a new version (AMPAC 2.1) of AMPAC which represents a major enhancement over the original one.

- (a) The basic routines have been rewritten and now run significantly faster, particularly on vector computers.
- (b) Inclusion of analytical derivatives of the energy for CI wave functions now allows open shell RHF calculations (BIRADICAL option) to be carried out at reasonable cost, the computing time required being only 2-3 times greater than for UHF ones.
- (c) A new and very effective procedure (CHAIN option) has been developed for locating transition states.
- (d) Several other utilities have been added for Hessian evaluation and characterization of stationary points.
- (e) AM1 parameters are now available for a number of additional elements. Those for sulfur and phosphorus reproduce properties of molecules containing their higher valence states. AM1 parameters are now available for all the "MNDO elements" except beryllium and lead.

Of the four authors of AMPAC listed, the major contribution has been by Professor Daniel Liotard. He developed an easily vectorized version of the basic code, which also runs faster on nonvector computers, and he was also responsible for the inclusion of analytical derivatives of CI energies and the CHAIN method for locating transition states. Dr. Eamonn Healy and Dr. James Ruiz have been responsible for developing and debugging a final version of AMPAC 2.1. My own contribution has been to provide encouragement and the necessary support. The manual was written by Roy D. Dennington and Eamonn Healy. We hope the fruits of our effort will prove useful to the chemical community.

One final change: we have decided to copyright AMPAC 2.1 solely to be able to protect it from mutilation. One of my major concerns has always been to try to prevent proliferation of versions of our procedures that represent only minor improvements. One of the major advantages of MINDO/3, MNDO and AM1 is that there has been only one version of each, so results obtained in different laboratories are strictly comparable. It is clearly important that the same should be true for computer programs.

The development of AMPAC 2.1, like all our work on semiempirical procedures, has been supported by grants from the Air Force Office of Scientific Research, the Robert A. Welch Foundation and the National Science Foundation.

*Michael J. S. Dewar*



## CHAPTER 1: INTRODUCTION

AMPAC is a general purpose semiempirical molecular orbital package for the study of chemical reactions. The semiempirical Hamiltonians MINDO/3, MNDO, and AM1 are implemented, and calculations of vibrational spectra, thermodynamic quantities, isotopic substitution effects and force constants for molecules, radicals, ions, and polymers are combined in a fully integrated package. Within the electronic part of the calculation, eigenvectors and localized orbitals, chemical bond indices, charges, molecular orbitals etc. are calculable. For studying chemical reactions two transition state location routines and transition state optimizing routines are available. For potential users to get the most out of the program they must understand how the program works, how to enter data, how to interpret the results, and what to do when things go wrong.

While AMPAC calls upon many concepts in quantum theory and thermodynamics and uses some fairly advanced mathematics, the user who is not familiar with these specialized topics should not feel excluded from using it. On the contrary, AMPAC is written with the non-theoretician in mind. To this end the data are kept as simple as possible; this means that users can give their attention to the chemistry involved, and not concern themselves with quantum and thermodynamic exotica.

### 1.1 SUMMARY OF AMPAC CAPABILITIES

1. AM1, MNDO, and MINDO/3 Hamiltonians
2. RHF and UHF methods
3. Extensive Configuration Interaction (C.I.)
  - a. 100 configurations
  - b. Singlets, Doublets, Triplets, Quartets, Quintets, and Sextets
  - c. Excited states
  - d. Geometry optimization, etc., on specified states
4. Single SCF calculation
5. Geometry optimization
6. Gradient minimization
7. Transition state location
8. Reaction path calculation
9. Force constant calculation
10. Normal mode analysis
11. Full vibrational frequency analysis
12. Transition dipole calculation
13. Thermodynamic properties calculation
14. Localized orbitals
15. Covalent bond orders
16. Bond analysis into sigma and pi
17. One dimensional polymer calculation
18. Dynamic Reaction Coordinate calculation

### 1.2 LAYOUT OF DATA

This section is aimed at the complete novice -- someone who knows nothing at all about the structure of a AMPAC data file.

While there are several types of data files for AMPAC, the simplest data file is the most commonly used. Rather than define it, an example is given below.

### Example Data File for D<sub>2d</sub> Ethylene

```

Line 1 :   UHF PULAY MINDO/3 VECTORS DENSITY LOCAL T=300
Line 2 :           MINDO/3 UHF CLOSED SHELL D2d ETHYLENE
Line 3 :           EXAMPLE DATA FILE FOR AMPAC
Line 4a: C
Line 4b: C      1.400118  1
Line 4c: H      1.098326  1 123.572063  1
Line 4d: H      1.098326  1 123.572063  1 180.000000  0  2  1  3
Line 4e: H      1.098326  1 123.572063  1  90.000000  0  1  2  3
Line 4f: H      1.098326  1 123.572063  1 270.000000  0  1  2  3
Line 5 :
```

The first line consists of keywords (here seven keywords are shown). These control the calculation. The next two lines are comments or titles. The user might want to put the name of the molecule and why it is being run on these two lines.

If no keywords are specified, leave a blank line. If comments are unwanted, include blank lines. A common error is to have a blank line before the keyword line. This error is quite tricky to find, so be careful not to have four lines before the start of the geometric data. Whatever is used, the three lines are necessary.

The next set of lines define the geometry. In this example, the numbers are all neatly lined up. While this is not necessary, it is easier to look for errors in the data. The geometry is defined in lines 4a to 4f; line 5 terminates both the geometry and the data file. Any additional data, such as symmetry data, would follow line 5.

Summarizing, then, the structure for a AMPAC data file is:

- Line 1: Keywords (See chapter on *Definitions of Keywords*);
- Line 2: Title of the calculation, e.g. the name of the molecule or ion;
- Line 3: Other information describing the calculation;
- Line 4: Internal or Cartesian coordinates (See chapter on *Specification of Geometry*);
- Line 5: Blank line or a line of 0's to terminate geometry definition.

Other layouts for data files involve additions to this simple layout. These additions occur at the end of the data file, after line 5. The four most common additions are:

1. Symmetry data: Following the geometric data, symmetry definitions are included and ended by a blank line.
2. Reaction path: After all geometry and symmetry data (if any) are specified, points on the reaction coordinate are defined.
3. Saddle data: A complete second geometry is specified. The second geometry follows the first geometry and symmetry data (if any).

- Chain data: First, an approximate transition state is specified, including any symmetry definitions. Next, the geometry of the reactant and the geometry of the product are given using the same format and linkage as the first geometry.

A more exotic example of a data file utilizes AMPAC's solid state capabilities.

### Example Data File for Polytetrahydrofuran

```

Line 1 :      T=40000
Line 2 :      POLY-TETRAHYDROFURAN (C4 H8 O)2
Line 3 :
Line 4a:  C      0.000000  0      0.000000  0      0.000000  0      0  0  0
Line 4b:  C      1.551261  1      0.000000  0      0.000000  0      1  0  0
Line 4c:  O      1.401861  1    108.919034  1      0.000000  0      2  1  0
Line 4d:  C      1.401958  1    119.302489  1   -179.392581  1      3  2  1
Line 4e:  C      1.551074  1    108.956238  1    179.014664  1      4  3  2
Line 4f:  C      1.541928  1    113.074843  1    179.724877  1      5  4  3
Line 4g:  C      1.551502  1    113.039652  1    179.525806  1      6  5  4
Line 4h:  O      1.402677  1    108.663575  1    179.855864  1      7  6  5
Line 4i:  C      1.402671  1    119.250433  1   -179.637345  1      8  7  6
Line 4j:  C      1.552020  1    108.665746  1   -179.161900  1      9  8  7
Line 4k:  XX     1.552507  1    112.659354  1   -178.914985  1     10  9  8
Line 4l:  XX     1.547723  1    113.375266  1   -179.924995  1     11 10  9
Line 4m:  H      1.114250  1      89.824605  1    126.911018  1      1  3  2
Line 4n:  H      1.114708  1      89.909148  1   -126.650667  1      1  3  2
Line 4o:  H      1.123297  1      93.602831  1    127.182594  1      2  4  3
Line 4p:  H      1.123640  1      93.853406  1   -126.320187  1      2  4  3
Line 4q:  H      1.123549  1      90.682924  1    126.763659  1      4  6  5
Line 4r:  H      1.123417  1      90.679889  1   -127.033695  1      4  6  5
Line 4s:  H      1.114352  1      90.239157  1    126.447043  1      5  7  6
Line 4t:  H      1.114462  1      89.842852  1   -127.140168  1      5  7  6
Line 4u:  H      1.114340  1      89.831790  1    126.653999  1      6  8  7
Line 4v:  H      1.114433  1      89.753913  1   -126.926618  1      6  8  7
Line 4w:  H      1.123126  1      93.644744  1    127.030541  1      7  9  8
Line 4x:  H      1.123225  1      93.880969  1   -126.380511  1      7  9  8
Line 4y:  H      1.123328  1      90.261019  1    127.815464  1      9 11 10
Line 4z:  H      1.123227  1      91.051403  1   -125.914234  1      9 11 10
Line 4aa: H      1.113970  1      90.374545  1    126.799259  1     10 12 11
Line 4bb: H      1.114347  1      90.255788  1   -126.709810  1     10 12 11
Line 4cc: Tv     12.299490  1      0.000000  0      0.000000  0      1 11 10
Line 5 :  O      0.000000  0      0.000000  0      0.000000  0      0  0  0

```

Polytetrahydrofuran has a repeat unit of (C4 H8 O)2; i.e., twice the monomer unit. This is necessary in order to allow the lattice to repeat after a translation through 12.3 Angstroms. (See the section on *Solid State Capability* in the *Background* chapter for further details.)

Note the two dummy atoms on lines 4k and 4l. These are useful, but not essential, for defining the geometry. The atoms on lines 4y to 4bb use these dummy atoms, as does the translation vector on line 4cc.

The translation vector has only the length marked for optimization. The reason for this is also explained in the *Background* chapter.



## CHAPTER 2: KEYWORDS

### 2.1 SPECIFICATION OF KEYWORDS

All control data are entered in the form of keywords, which form the first line of a data file. The order in which keywords appear is not important although they must be separated by a space. Some keywords can be abbreviated (for example 1ELECTRON can be entered as 1ELECT), but in general the full keyword is preferred. If there is insufficient space in the first line for all the keywords needed, then consider abbreviating the longer words. Keywords containing an equal sign, such as BAR=0.05, may not be abbreviated.

If two keywords that are incompatible (like UHF and C.I.) are given, or a keyword which is incompatible with the species given (for instance TRIPLET and a methyl radical) then error trapping will normally occur and an error message printed. As this usually takes an insignificant amount of time, data files are easily checked for obvious errors.

Note that if a keyword is misspelled, AMPAC will ignore it and no error message will be generated.

### 2.2 FULL LIST OF KEYWORDS USED IN AMPAC

0SCF	- READ IN DATA, THEN STOP
1ELECTRON	- FINAL ONE-ELECTRON MATRIX WILL BE PRINTED
1SCF	- DO 1 SCF CALCULATION AND THEN STOP
AM1	- THE AM1 HAMILTONIAN WILL BE USED
BAR=nn	- REDUCE BAR LENGTH BY A MAX. OF nn
BIRADICAL	- SYSTEM HAS TWO UNPAIRED ELECTRONS
BONDS	- FINAL BOND ORDER MATRIX WILL BE PRINTED
C.I.=n	- A MULTI-ELECTRON CONFIGURATION INTERACTION SPECIFIED
CHAIN	- TRANSITION STATE WILL BE LOCATED
CHARGE=n	- CHARGE ON SYSTEM = n (e.g. NH4 => CHARGE=1)
COMPFG	- PRINT HEAT OF FORMATION CALCULATED IN COMPFG
CYCLES=nn	- MAXIMUM NUMBER OF CYCLES is nn
DCART	- PRINT DETAILS OF WORKING IN DCART
DEBUG	- DEBUG OPTION TURNED ON
DEBUGPULAY	- PRINT DETAILS OF WORKING IN PULAY
DENMAT	- DENSITY MATRIX WRITTEN TO DISK IN ASCII FORMAT
DENOUT	- DENSITY MATRIX WRITTEN TO DISK IN BINARY FORMAT
DENSITY	- FINAL DENSITY MATRIX WILL BE PRINTED
DEP	- GENERATE FORTRAN CODE FOR PARAMETER UPDATES
DEPVAR=n.nn	- TRANSLATION VECTOR IS A MULTIPLE OF BOND LENGTH
DERI1	- PRINT WORKINGS OF SUBROUTINE DERI1
DERI2	- PRINT WORKINGS OF SUBROUTINE DERI2
DERINU	- FORCE DERIVATIVES WILL BE COMPUTED NUMERICALLY
DERIV	- PRINT WORKINGS OF SUBROUTINE DERIV
DFORCE	- FORCE CALCULATION SPECIFIED, FORCE MATRIX PRINTED
DOUBLET	- RHF DOUBLET STATE REQUIRED
DRC=n.nnn	- DYNAMIC REACTION COORDINATE CALCULATION
ECHO	- DATA ARE ECHOED BACK BEFORE CALCULATION STARTS
EIGS	- ALL EIGENVALUES IN ITER WILL BE PRINTED
ENPART	- ENERGY WILL BE PARTITIONED INTO COMPONENTS
ESR	- UNPAIRED SPIN DENSITY WILL BE CALCULATED

EXCITED	- FIRST EXCITED SINGLET STATE WILL BE OPTIMIZED
EXTERNAL=	- MNDO OR AM1 PARAMETERS WILL BE READ FROM A FILE
FILL=n	- IN RHF OPEN AND CLOSED SHELL, M.O. n WILL BE FILLED
FLEPO	- PRINT DETAILS OF GEOMETRY OPTIMIZATION
FMAT	- PRINT DETAILS OF WORKING IN FMAT
FOCK	- LAST FOCK MATRIX WILL BE PRINTED
FORCE	- FORCE CALCULATION SPECIFIED
FULSCF	- IN SEARCHES, FULL SCF CALCULATIONS WILL BE DONE
GEO-OK	- OVERRIDE INTERATOMIC DISTANCE CHECK
GNORM=n.n	- FLEPO EXIT WHEN GRADIENT NORM BELOW n.n
GRADIENTS	- ALL GRADIENTS WILL BE PRINTED
GRAPH	- GENERATE FILE FOR GRAPHICS
H-PRIO	- HEAT OF FORMATION TAKES PRIORITY IN DRC
HCORE	- PRINT DETAILS OF WORKING IN HCORE
ISOTOPE	- FORCE MATRIX WRITTEN TO DISK (CHANNEL 9)
ITER	- PRINT DETAILS OF WORKING IN ITER
ITRY=nn	- SET LIMIT OF NUMBER OF SCF ITERATIONS TO nn
KINETIC=n.nnn	- n.nnn EXCESS KINETIC ENERGY ADDED TO DRC CALCULATION
LET	- DO NOT REDUCE GRADIENTS IN FORCE
LINMIN	- PRINT DETAILS OF WORKING IN LINMIN
LOCALIZE	- LOCALIZED ORBITALS WILL BE PRINTED
LOCMIN	- PRINT DETAILS OF WORKING IN LOCMIN
LTRD	- MINIMIZE GRADIENT USING FULL NEWTON ALGORITHM
MECI	- PRINT THE FINAL C.I. MATRIX
MICROS=n	- NUMBER OF MICROSTATES SPECIFIED IN THE C.I. IS n
MINDO/3	- THE MINDO/3 HAMILTONIAN WILL BE USED
MOLDAT	- DETAILS OF WORKING IN MOLDAT WILL BE PRINTED
MULLIK	- PRINT "MULLIKEN-TYPE" POPULATION ANALYSIS
NEWTON	- MINIMIZE ENERGY USING FULL NEWTON ALGORITHM
NLLSQ	- GRADIENTS WILL BE MINIMIZED USING NLLSQ
NOINTER	- INTERATOMIC DISTANCES WILL NOT BE PRINTED
NOXYZ	- CARTESIAN COORDINATES WILL NOT BE PRINTED
OLDENS	- INITIAL DENSITY MATRIX READ FROM BINARY FILE
OLDMAT	- INITIAL DENSITY MATRIX READ FROM ASCII FILE
OPEN(n1,n2)	- OPEN SHELL RHF CALCULATION REQUESTED
PATH	- FOLLOW THE DESCENDING REACTION PATH
PI	- RESOLVE DENSITY MATRIX INTO SIGMA AND PI BONDS
PL	- MONITOR CONVERGENCE OF DENSITY MATRIX IN ITER
POWELL	- MINIMIZE GRADIENT USING POWELL METHOD
POWSQ	- PRINT DETAILS OF WORKING IN POWSQ
PRECISE	- TERMINATION CRITERIA INCREASED BY 100 TIMES
PRINT=n	- PRINTOUT LEVEL IN OPTIMIZATION = n
PULAY	- PULAY'S METHOD WILL BE USED IN SCF
QUARTET	- RHF QUARTET STATE REQUIRED
QUINTET	- RHF QUINTET STATE REQUIRED
RESTART	- CALCULATION RESTARTED
ROOT=n	- IN A C.I. CALCULATION, ROOT n WILL BE OPTIMIZED
ROT=n	- THE SYMMETRY NUMBER OF THE SYSTEM IS n
SADDLE	- TRANSITION STATE WILL BE LOCATED
SCFCRT=.n	- DEFAULT SCF CRITERION REPLACED BY THE VALUE SUPPLIED
SEARCH	- PRINT DETAILS OF WORKING IN SEARCH
SEXTET	- RHF SEXTET STATE REQUIRED
SHIFT=n.n	- A DAMPING FACTOR OF n.n DEFINED
SIGMA	- GRADIENTS WILL BE MINIMIZED USING SIGMA
SINGLET	- RHF SINGLET STATE REQUIRED
SPIN	- FINAL UHF SPIN MATRIX WILL BE PRINTED
STEP1=n.nnn	- STEP SIZE FOR FIRST COORDINATE IN GRID CALCULATION
STEP2=n.nnn	- STEP SIZE FOR SECOND COORDINATE IN GRID CALCULATION

SYMMETRY	- SYMMETRY CONDITIONS WILL BE IMPOSED
T-PRIO	- TIME TAKES PRIORITY IN DRC
T.V.	- TRANSITION VECTOR TO BE PROVIDED FOR PATH
T=nnn	- A TIME OF nnn SECONDS REQUESTED
THERMO( )	- A THERMODYNAMICS CALCULATION WILL BE PERFORMED
TIMES	- TIMES OF VARIOUS STAGES TO BE PRINTED
TRANS	- THE SYSTEM IS A TRANSITION STATE (USED IN THERMODYNAMICS CALCULATION)
TRIPLET	- TRIPLET STATE REQUIRED
UHF	- UNRESTRICTED HARTREE-FOCK CALCULATION
VECTORS	- FINAL EIGENVECTORS WILL BE PRINTED
WEIGHT	- WEIGHT WILL BE PROVIDED FOR PATH
X-PRIO	- GEOMETRY CHANGES TAKE PRIORITY IN DRC
XYZ	- GEOMETRY GIVEN IN CARTESIAN SPACE

## 2.3 DEFINITIONS OF KEYWORDS

### 0SCF

The data is read and printed in the output file, but no actual calculation is performed by using this keyword. This is useful for checking data files.

### 1ELECTRON

The final one-electron matrix is printed out. This matrix is composed of atomic orbitals; the element between orbitals *i* and *j* in MNDO and AM1 on different atoms is given by:

$$H_{ij} = \frac{1}{2} S_{ij} (\beta_i + \beta_j)$$

The matrix elements between orbitals *i* and *j* on the same atom are calculated from the electron-nuclear attraction energy, and also from the *U(i)* value, if *i=j*.

The one-electron matrix is *unaffected* by the charge, and the electron density. It is only a function of the geometry. *Abbreviation:* 1ELEC.

### 1SCF

When studying a single specific geometry, 1SCF should be used. All the keywords relevant to output can be used. If the gradients are to be calculated, then GRADIENTS should be specified. They are not calculated by default.

If the keyword RESTART is also present, then the geometric parameters which were being optimized will be used in the gradient calculation.

1SCF is helpful in a learning situation. AMPAC normally performs many SCF calculations, and in order to minimize output when following the working of the SCF calculation, 1SCF is very useful. Additionally, 1SCF can be used to calculate specific properties from previously optimized geometries.

**AM1**

The new AM1 method is used. MNDO is taken as the default.

**BAR=nn**

In the SADDLE calculation the distance between the two geometries is steadily reduced until the transition state is located. Sometimes, however, the user may want to alter the maximum rate at which the distance between the two geometries are reduced. BAR is a ratio, normally 0.15, or 15 percent. This represents a maximum rate of reduction of the bar of 15 percent per step. Alternative values that might be considered are BAR=0.05 or BAR=0.10, although other values may be used.

**BIRADICAL**

This keyword or UHF should be specified when attempting to calculate properties of molecules which have biradicaloid character. First, a Half Electron RHF SCF (HE-RHF) calculation is done to generate the reference wavefunction. Then, the following microstates of the SCF reference wavefunction are mixed in a post SCF C.I. calculation:

	Microstate 1				Microstate 2		Microstate 3	
	Alpha	Beta	Alpha	Beta	Alpha	Beta	Alpha	Beta
LUMO	+	—	—	+	—	—	+	+
					+			
HOMO	—	+	+	—	+	+	—	—

The first microstate is a linear combination of the two states formed by placing one electron in the HOMO and one in the LUMO, the result being a full singlet state. The second microstate results from placing both electrons in the HOMO. The third microstate results from placing both electrons in the LUMO.

NOTE: Experienced users of MECI can duplicate the effect of the keyword BIRADICAL by using the MECI keywords OPEN(2,2) and SINGLET.

BIRADICAL should not be used if either the HOMO or LUMO is degenerate; in this case, the full manifold of HOMO x LUMO should be included in the C.I., using MECI options. The user should be aware of this situation. When the biradical calculation is performed correctly, the result is normally a net stabilization. However, if the first singlet excited state is much higher in energy than the closed shell ground state, BIRADICAL can lead to a destabilization. *Abbreviation:* BIRAD.

**BONDS**

The rotationally invariant bond order between all pairs of atoms is printed. In this context a bond is defined as the sum of the squares of the density matrix elements connecting any two atoms. For ethane, ethylene, and acetylene the carbon-carbon bond orders are roughly 1.00, 2.00, and 3.00 respectively. The diagonal terms are the valencies calculated from the atomic terms only and are defined as the sum of the bonds the atom makes with other atoms. In UHF and non-variationally optimized wavefunctions the calculated valency will be incorrect, the degree of error being proportional to the non-idempotency of the density matrix. For an RHF wavefunction the square of the density matrix is equal to twice the density matrix.

**C.I.=n**

Normally configuration interaction is invoked if any of the keywords which imply a C.I. calculation are used, such as BIRADICAL, TRIPLET, QUARTET, etc. Note that ROOT=n does not imply a C.I. calculation; ROOT=n is only used when a C.I. calculation is done. Since “implied” C.I.s involve the minimum number of configurations, users may want to define a larger than minimum C.I., in which case the keyword C.I.=n must be included. When C.I.=n is specified, the n M.O.s which “bracket” the occupied-virtual energy levels will be used. Thus C.I.=2 will include both the HOMO and the LUMO, while C.I.=1 (implied for odd electron systems) will only include the HOMO. This has no effect for a closed shell system and leads to Dewar’s half-electron correction for odd electron systems. Users should be aware of the rapid increase in the size of the C.I. with increasing numbers of M.O.s. The number of microstates implied by use of the keyword C.I.=n are as follows:

Keyword	Even Electron Systems			Odd Electron Systems		
	No. Electrons		No. Configurations	No. Electrons		No. Configurations
	Alpha	Beta		Alpha	Beta	
C.I.=1	1	1	1	1	0	1
C.I.=2	1	1	4	1	0	2
C.I.=3	2	2	9	2	1	9
C.I.=4	2	2	36	2	1	24
C.I.=5	3	3	100	3	2	100
C.I.=6	3	3	400	3	2	300
C.I.=7	4	4	1225	4	3	1225
C.I.=8	(Do not use unless other keywords are also used, see below.)					

If the spin state is specified, then larger numbers of M.O.s can be used up to a maximum of 10. The C.I. matrix is 100 x 100. For calculations involving up to 100 configurations, the spin states are exact eigenstates of the spin operators. For systems with more than 100 configurations, the 100 configurations of lowest energy are used. (See also MICROS and the keywords defining spin states.)

Note that for any system, use of C.I.=5 or higher normally requires the diagonalization of a 100 x 100 matrix costing considerable CPU time.

**CHAIN**

A transition state lying between two given minima will be located by the chain method. This algorithm is very successful at finding the highest transition state in a multi-step mechanism. Three separate geometries must be given. The first geometry is an approximate transition state, the second and third geometries are left and right minima, respectively. It is important to use the same format and connectivity for all three geometries. (See the chapters on *Examples* and *Background* for more information.)

**CHARGE=n**

When the system being studied is an ion, the charge (n) on the ion can be supplied by CHARGE=n. For cations n can be 1 or 2 or 3, etc; for anions, n is -1 or -2 or -3, etc.

## EXAMPLES

<u>ION</u>	<u>KEYWORD</u>	<u>ION</u>	<u>KEYWORD</u>
(NH <sub>4</sub> ) <sup>+</sup>	CHARGE=1	(CH <sub>3</sub> COO) <sup>-</sup>	CHARGE=-1
(C <sub>2</sub> H <sub>5</sub> ) <sup>+</sup>	CHARGE=1	(COO) <sup>=</sup>	CHARGE=-2
(SO <sub>4</sub> ) <sup>=</sup>	CHARGE=-2	(PO <sub>4</sub> ) <sup>3-</sup>	CHARGE=-3
(HSO <sub>4</sub> ) <sup>-</sup>	CHARGE=-1	(H <sub>2</sub> PO <sub>4</sub> ) <sup>-</sup>	CHARGE=-1

**CYCLES=n**

The number of optimization cycles can be increased by explicitly specifying CYCLES=n, where n is the maximum number of cycles. For most calculations, the DEFAULT values are sufficient.

**DCART**

The Cartesian derivatives which are calculated in DCART for variationally optimized systems are printed if the keyword DCART is present. The derivatives are in units of kcal/Angstrom, and the components are given as displacements in x, y, and z.

**DEBUG**

Certain keywords have specific output control meanings, such as FOCK, VECTORS and DENSITY. If they are used, only the final arrays of the relevant type are printed. If DEBUG is supplied, then all arrays are printed. This is useful in debugging ITER. DEBUG can also increase the amount of output produced when other keywords unrelated to output control are used, e.g. COMPG.

**DENMAT**

The density matrix at the end of the calculation is printed as a formatted ASCII text file. This allows the user to manually alter a density matrix and start a new calculation with this altered matrix as the default. (See OLDMAT.)

**DENOUT**

The density matrix at the end of the calculation is printed in binary form suitable for input with another job. If a calculation stops because the time limit is exceeded, DENOUT is invoked automatically. (See RESTART.)

**DENSITY**

At the end of a job, the final density matrix is printed in the output file. For RHF the normal density matrix is printed, for UHF both alpha and beta density matrices are printed.

If density is not requested, then the diagonal of the density matrix (i.e., the electron density on atomic orbitals) is printed.

## DEP

For use only with EXTERNAL=. When new parameters are published, they can be entered at run time by using EXTERNAL=, but as this is somewhat clumsy, a permanent change can be made by use of DEP.

If DEP is invoked, a complete block of FORTRAN code will be generated, and this can be inserted directly into the BLOCK DATA file.

Note that this is designed only for use with MNDO or AM1 parameters. Only code for AM1 will be generated. To convert the FORTRAN code to define MNDO parameters, insert the letter M before every left parenthesis; thus, convert “(” to read “M(”.

## DEPVAR=n.nn

In polymers the translation vector is frequently a multiple of some internal distance. For example, in polythene it is the C1-C3 distance. If a cluster unit cell of C<sub>6</sub>H<sub>12</sub> is used, then symmetry can be used to tie together all the carbon atom coordinates and the translation vector distance. In this example DEPVAR=3.0 would be suitable.

## DERINU

The effect of this keyword is to force all energy derivatives to be performed by numerical methods. This keyword is useful when the analytical C.I. derivatives are suspected of being in error.

## DOUBLET

When a configuration interaction calculation is done, all spin states are calculated simultaneously, either for component of spin = 0 or 1/2. When only doublet states are of interest DOUBLET can be specified. While all other spin states are calculated, they are ignored in the choice of root.

Note that while almost every odd electron system will have a doublet ground state, DOUBLET should still be specified if the desired state must be a doublet.

DOUBLET has no meaning in a UHF calculation.

## DRC

This keyword is used to indicate a Dynamic Reaction Coordinate calculation. By default, total energy is conserved. As the “reaction” proceeds in time, energy transfers between kinetic and potential forms.

## DRC=n.nnn

In a DRC calculation, the “half-life” for loss of kinetic energy is defined as n.nnn x 10 femtoseconds. If n.nnn is set to zero, infinite damping simulating a very condensed phase is obtained.

## ECHO

Geometry data will be echoed back if ECHO is specified. This is only useful for checking the data file for errors.

**ENPART**

The total energy (in eV) obtained by the addition of the electronic and nuclear terms is partitioned into mono- and bicentric contributions, and these contributions in turn are divided into nuclear, one- and two-electron terms.

**ESR**

The unpaired spin density arising from an odd electron system can be calculated for both RHF and UHF. In a UHF calculation the alpha and beta M.O.s have different spatial forms, so unpaired spin density can naturally be present on in-plane hydrogen atoms such as in the phenoxy radical.

In the RHF formalism a MECI calculation is performed. If the keywords OPEN and C.I.= are both absent then only a single state is calculated. The unpaired spin density is then calculated from the state function. In order to have unpaired spin density on the hydrogens (e.g., in the phenoxy radical) several states should be mixed.

**EXCITED**

The state that will be calculated is the first excited open shell singlet state. If the ground state is a singlet, then the state calculated will be S(2); if the ground state is a triplet, then S(1). This state would normally be the one resulting from a one electron excitation from the HOMO to the LUMO. If the lowest singlet state is a biradical, the EXCITED state will be a closed shell.

The EXCITED state will be calculated from a BIRADICAL calculation in which the second root of the C.I. matrix is selected. Note that the eigenvector of the C.I. matrix is not used in the current formalism.  
*Abbreviation:* EXCI.

NOTE: EXCITED is a redundant keyword, and represents a particular configuration interaction calculation. Experienced users of MECI can duplicate the effect of the keyword EXCITED by using the MECI keywords OPEN(2,2), SINGLET, and ROOT=2.

**EXTERNAL=**

Normally, AM1 and MNDO parameters are taken from the BLOCK DATA files within AMPAC. As new parameters become available, they can be inserted at run time by use of EXTERNAL=<filename>, where <filename> is the name of the file that contains the new parameters.

<filename> consists of a series of parameter definitions in the format:

<Parameter>   <Element>   <Value of parameter>

where the possible parameters are USS, UPP, UDD, ZS, ZP, ZD, BETAS, BETAP, BETAD, GSS, GSP, GPP, GP2, HSP, ALP, FNnm (n=1,2, or 3, and m=1 to 10), and the elements are defined by their chemical symbols, such as Si or S1.

When new parameters for elements are published, they can be typed in as shown. This file is ended by a blank line, or the word END. An example of a parameter data file would be:

Line	1:	USS	Si	-34.08201495
Line	2:	UPP	Si	-28.03211675
Line	3:	BETAS	Si	-5.01104521
Line	4:	BETAP	Si	-2.23153969
Line	5:	ZS	Si	1.28184511
Line	6:	ZP	Si	1.84073175
Line	7:	ALP	Si	2.18688712

Derived parameters do not need to be entered. They will be calculated from the optimized parameters. All "constants" such as the experimental heat of atomization are already inserted for all elements.

NOTE: Only MNDO or AM1 parameters can be specified with EXTERNAL. It is unlikely that any more MINDO/3 parameters will be published.

(See also DEP to make a permanent change.)

### **FILL=n**

The nth M.O. in an RHF calculation is constrained to be occupied. It has no effect on a UHF calculation. After the first iteration (NOTE: not after the first SCF calculation, but after the first iteration within the first SCF calculation) the nth M.O. is stored, and, if occupied, no further action is taken at that time. If unoccupied, then the HOMO and the nth M.O.'s are swapped, so that the nth M.O. is now filled. On all subsequent iterations the M.O. nearest in character to the stored M.O. is forced to be filled, and the stored M.O. replaced by that M.O. This is necessitated by the fact that in a reaction a particular M.O. may change its character very considerably. A useful procedure is to run 1SCF and DENOUT first, in order to identify the M.O.'s; the complete job is then run with OLDENS and FILL=nn, so that the eigenvectors at the first iteration are fully known.

### **FORCE**

The Hessian, that is the matrix (in millidynes per Angstrom) of second derivatives of the energy with respect to displacement of all pairs of atoms in x, y, and z, is calculated. On diagonalization this gives the force constants for the molecule. The force matrix, weighted for isotopic masses, is then used for calculating the vibrational frequencies. The system can be characterized as a ground state or a transition state by the presence of five (for a linear system) or six eigenvalues which are very small (less than 30 reciprocal centimeters). A transition state is further characterized by one, and exactly one, negative force constant.

Proper characterization of all stationary points requires a FORCE or LTRD calculation. Also, a FORCE calculation is a prerequisite for a THERMO calculation.

Before a FORCE calculation is started, the gradient norm (GNORM) is checked to ensure that the geometry is at a stationary point. If it is larger than 5, the GNORM will be reduced using the BFGS method. All internal coordinates are optimized, and any symmetry constraints are ignored at this point. An implication of this is that if the specification of the geometry relies on any angles being exactly 180 or zero degrees, the calculation may fail.

The geometric definition supplied to FORCE should not rely on angles or dihedrals assuming exact values. The test of exact linearity is sufficiently slack so that most molecules that are linear, such as acetylene and but-2-yne, should not be stopped. (See also THERMO, LET, TRANS, ISOTOPE.)

## FULSCF

In line searches the option exists to require all energy evaluations to be done using full SCF calculations. Normally full SCF calculations are not carried out during a line search as the density matrix is normally not changing very fast. The only important exception is in non-variationally optimized wavefunctions, such as occur in half-electron or C.I. calculations.

## GEO-OK

Normally the program will stop with a warning message if two atoms are within 0.8 Angstroms of each other, or (more rarely) the BFGS routine has difficulty optimizing the geometry. GEO-OK will override the job termination sequence, and allow the calculation to proceed. In practice most jobs that terminate due to these checks contain errors in data, so caution should be exercised if GEO-OK is used. An important exception to this warning is if the calculation contains or produces a hydrogen molecule.

## GNORM=n.n

Geometry optimization termination criteria can be overridden by specifying a gradient norm requirement. For example, GNORM=20 would allow the optimization routine to exit as soon as the gradient norm dropped below 20.0, the default being 1.0. A GNORM=0.01 could be used to refine a geometry beyond the normal limits. However, if the calculation passes other tests it may stop before reaching the gradient norm specified. WARNING: If a very small value is chosen, the calculation may not terminate in a reasonable time. A reasonable lower bound for GNORM is 0.1.

## GRADIENTS

In a 1SCF calculation, gradients are not calculated by default. In non-variationally optimized systems this would take an excessive amount of time. When the keyword GRADIENTS is specified with 1SCF *all* gradients are calculated (whether marked for calculation or not) and printed. An exception is when the 1SCF is used in conjunction with the keyword RESTART, in which case only the coordinates flagged for optimization would have their gradients printed.

In a geometry optimization calculation, the final gradient components are printed in the output file and the Gradient Norm is printed in the archive file. Only the coordinates being optimized will have their gradients printed. Since some of the optimization packages do not print gradient information by default, the GRADIENTS keyword is recommended in most cases. *Abbreviation:* GRAD.

## GRAPH

Information needed to generate electron density contour maps can be written to a file by calling GRAPH. GRAPH first enters MULLIK in order to generate the inverse square root of the overlap matrix, which is required for the re-normalization of the eigenvectors. All necessary data is written to disk.

## H-PRIORITY

In a DRC calculation, results will be printed whenever the calculated heat of formation changes by 0.1 kcal/mol. *Abbreviation:* H-PRIO.

### **H-PRIORITY=n.nn**

In a DRC calculation, results will be printed whenever the calculated heat of formation changes by n.nn kcal/mol.

### **ISOTOPE**

The FORCE matrix is very time consuming to generate, and in isotopic substitution studies several vibrational calculations may be needed. To allow the frequencies to be calculated from the (constant) force matrix, ISOTOPE is used. When a FORCE calculation is completed, ISOTOPE will cause the force matrix to be stored, regardless of whether or not any intervening restarts have been made. To recalculate the frequencies, etc., start at the end of the force matrix calculation and specify RESTART.

The two keywords RESTART and ISOTOPE can be used together. For example, if a normal FORCE calculation runs for a long time, the user may want to divide it up into stages and save the final force matrix. Once ISOTOPE has been used, it does not need to be used on subsequent RESTART runs.

### **ITRY=nn**

The default maximum number of SCF iterations is 200. When this limit presents difficulty, ITRY=nn can be used to redefine it. For example, if ITRY=400 is used, the maximum number of iterations will be set to 400. ITRY should normally not be changed until all other means of obtaining an SCF have been exhausted, e.g. SHIFT and PULAY.

### **KINETIC=n.nnn**

In a DRC calculation n.nnn kcal/mol of excess kinetic energy is added to the system as soon as the kinetic energy builds up to 0.2 kcal/mol. The excess energy is added to the velocity vector, without change of direction.

### **LET**

Before the Hessian matrix is calculated in a FORCE calculation the geometry will be refined by a gradient minimization routine if the gradient norm is larger than 5. If the user does not want the refinement to be carried out, LET is provided to let the calculation proceed. Extreme caution should be exercised with this keyword. FORCE results are meaningless unless the gradient norm is sufficiently low.

### **LOCALIZE**

The occupied eigenvectors are transformed into a localized set of M.O.'s by a series of 2 x 2 rotations which maximize  $\langle \Psi^4 \rangle$ . The value of  $1/\langle \Psi^4 \rangle$  is a direct measure of the number of centers involved in the M.O., thus for H<sub>2</sub> the value of  $1/\langle \Psi^4 \rangle$  is 2.0, for a three-center bond it is 3.0, and for a lone pair it is 1.0. Higher degeneracies than allowed by point group theory are readily obtained. For example, benzene would give rise to a 6-fold degenerate C-H bond, a 6-fold degenerate C-C sigma bond and a three-fold degenerate C-C pi bond. In principle, there is no single step method to unambiguously obtain the most localized set of M.O.'s in systems where several canonical structures are possible, just as no simple method exists for finding the most stable conformer of some large compound. However, the localized bonds generated from this procedure will usually be acceptable for routine applications. *Abbreviation:* LOCAL.

**LTRD**

The gradient norm will be minimized by the full Newton algorithm. Inflection points are avoided by switching the quadratic search direction to that of the eigenvector associated with the vanishing eigenvalue. This method is extremely accurate and expensive; a full Hessian matrix is constructed at each iteration. LTRD can also be used with CYCLES=O to build and diagonalize a Hessian matrix in a given subset of internal coordinates. Note that when testing stationary points, the gradient norm must be less than 5 or the results will not be valid.

LTRD provides a standard deviation for the eigenvalues, and uses either an analytical-numerical method or a twice numerical method according to the nature of the gradient computation. Simple forward difference or central finite difference is used depending on whether the keyword PRECISE is specified.

**MECI**

The C.I. coefficients are printed. This keyword is recommended with every C.I. calculation.

**MICROS=n**

The microstates used by MECI are normally generated by use of a permutation operator. When individually defined microstates are desired, then MICROS=n can be used, where n defines the number of microstates to be specified.

*Format for Microstates:* After the geometry and any symmetry data are given, data defining each microstate is specified, using format 20I1, one microstate per line.

For a system with n M.O.s in the C.I. (OPEN=(n1,n) or C.I.=n), the populations of the n alpha M.O.s are defined, followed by the n beta M.O.s. Allowed occupancies are zero and one. For n=6 the closed shell ground state would be defined as 111000111000, meaning one electron in each of the first three alpha M.O.s, and one electron in each of the first three beta M.O.s.

Users are warned that they are responsible for completing any spin manifolds. Thus while the state 111100110000 is a triplet state with z-component of spin = 1, the state 111000110100, while having a z-component of spin = 0 is neither a singlet nor a triplet. In order to complete the spin manifold the microstate 110100111000 must also be included.

If a manifold of spin states is not complete, then the eigenstates of the spin operator will not be quantized. When 100 or fewer microstates are supplied, this is the only cause of loss of spin quantization.

There are two other limitations on possible microstates. First, the number of electrons in every microstate should be the same. If they differ, a warning message will be printed, and the calculation continued (but the results will almost certainly be nonsense). Second, the z-component of spin for every microstate must be the same, except for teaching purposes. Two microstates of different z-components of spin will have a zero matrix element connecting them. No warning will be given as this is a reasonable operation in a teaching situation. For example, if all states arising from two electrons in two levels are to be calculated, say for teaching Russel-Saunders coupling, then the following microstates would be used:

Microstate	No. of alpha,	beta electrons	Ms	State
1100	2	0	1	Triplet
1010	1	1	0	Singlet
1001	1	1	0	Mixed
0110	1	1	0	Mixed
0101	1	1	0	Singlet
0011	0	2	-1	Triplet

Constraints on the space manifold are just as rigorous, but much easier to satisfy. If the energy levels are degenerate, then all z-components of a manifold of degenerate M.O.s should be either included or excluded. If only some, but not all, z-components are used, the required degeneracy of the states will be missing.

Taking tetrahedral methane cation with a z-component of spin = 3/2 (neglecting Jahn-Teller distortion) as an example, the minimum number of microstates that should be supplied is  $90 = (6!/(1!.5!))*(6!/(4!.2!))$ .

While the total number of electrons should be the same for all microstates, this number does not need to be the same as the number of electrons supplied to the C.I.; thus in the example above, a cationic state could be 110000111000.

The format is defined as 2011 so that spaces can be used for empty M.O.s.

### MINDO/3

The default Hamiltonian within AMPAC is MNDO. To use the MINDO/3 Hamiltonian the keyword MINDO/3 must be specified. Acceptable alternatives are MINDO and MINDO3.

### MULLIK

A full Mulliken-type Population analysis is done on the final RHF wavefunction. This involves the following steps:

- (1) The eigenvector matrix is divided by the square root of the overlap matrix, S;
- (2) The Coulson type density matrix, P, is formed from the modified eigenvector matrix;
- (3) The overlap population is formed from  $P_{ij} * S_{ij}$ ;
- (4) Half the off-diagonals are added onto the diagonals.

### NEWTON

Minimize the energy using the full Newton algorithm. The search direction is chosen as either steepest descent, the eigenvector associated with the lowest eigenvalue of the Hessian, or the full Newton direction (For further discussion see LTRD).

### NLLSQ

The gradient norm will be minimized by Bartel's method (a Non Linear Least Squares gradient minimization routine). Gradient minimization will locate one of three possible points:

1. A *minimum* on the potential energy surface. The gradient norm will go to zero, and the lowest five or six eigenvalues resulting from a FORCE calculation will be approximately zero.

2. A *transition state*. The gradient norm will vanish, as in (a), but in this case the system is characterized by one, and only one, negative force constant.
3. A *local minimum* in the gradient norm space. In this (normally unwanted) case the gradient norm is minimized, but does not go to zero. A FORCE calculation will not give the five or six zero eigenvalues characteristic of a stationary point. While normally undesirable, this is sometimes the only way to obtain a geometry. For instance, if a system is formed which cannot be characterized as an intermediate, and at the same time is not a transition state, but nonetheless has some chemical significance, then that state can be refined using NLLSQ.

## OLDENS

A density matrix produced by an earlier run of AMPAC is used to start the current calculation. When a previous calculation ended successfully but a subsequent run fails to achieve an SCF, an old density matrix can sometimes be used to successfully obtain an SCF.

## OLDMAT

An ASCII text file containing a formatted density matrix produced by an earlier run of AMPAC using the keyword DENMAT is used to start the current calculation.

## OPEN(n1,n2)

The M.O. occupancy during the SCF calculation can be defined in terms of doubly occupied, empty, and fractionally occupied M.O.s. The fractionally occupied M.O.s are defined by OPEN(n1,n2), where n1 = number of electrons in the open shell manifold, and n2 = number of open shell M.O.s; n1 must be in the range 0 to 2 x n2. OPEN(1,) will be assumed for odd electron systems unless an OPEN keyword is used. Errors introduced by use of fractional occupancy are automatically corrected in a MECI calculation when OPEN(n1,n2) is used.

## PATH

Starting from a transition state in the direction of the transition vector (if provided), the method follows a reaction path tangential to the weighted force. This is a smooth and efficient method to discover a weak stable intermediate along a specified path. Extra data are required if the keywords T.V. or WEIGHT are specified to provide the transition vector or a weighting, respectively.

## PI

The normal density matrix is composed of atomic orbitals, that is s, px, py and pz. PI allows the user to see how each atom-atom interaction is split into sigma and pi bonds. The resulting "density matrix" is composed of the following basis functions: s-sigma, p-sigma, p-pi, d-sigma, d-pi, d-dell. The diagonal terms give the hybridization state, so that an sp<sup>2</sup> hybridized system would be represented as s-sigma 1.0, p-sigma 2.0, p-pi 1.0.

## POWELL

Minimize the gradient norm by the Powell algorithm. The search direction is chosen as the steepest descent, the full Newton quadratic direction, or a judicious hybrid of both. POWELL provides the index of

the Hessian at the starting point, an estimate of the index at the final point, and a statistical estimate of the accuracy of the final geometry.

This method is probably the best choice for refining the geometry of a transition state that has been approximately located by either the CHAIN method, SADDLE or a reaction coordinate procedure.

### **POWSQ**

Details of the subroutine POWSQ are printed out. This information is useful only for debugging.

### **PRECISE**

The criteria for terminating all optimizations, electronic and geometric, will be increased by a factor of 100. This keyword should be used where more exact results are desired. If the results are going to be used in a FORCE calculation, where the geometry needs to be known precisely, then PRECISE is recommended. For small systems the extra cost in CPU time is minimal.

### **PRINT=n**

This keyword controls the level of detail printed by the optimization methods at each iteration. The level of printout ranges from 0 to 5, zero being minimal printout. Error messages and final results are always printed.

### **PULAY**

The default converger in the SCF calculation will be replaced by Pulay's procedure as soon as the density matrix is sufficiently stable. A considerable improvement in speed can be achieved by use of PULAY. If a large number of SCF calculations is expected, a sample calculation using 1SCF and PULAY should be compared with one using just 1SCF. If the PULAY calculation time is faster, it should be used in the full calculation.

The keyword PULAY should be used with care since it prevents the automatic use of the combined package of convergers (SHIFT, PULAY and the CAMP-KING convergers) when the system fails to achieve SCF after 200 iterations. The combined set of SCF convergers very seldom fails.

### **QUARTET**

The desired spin state is a quartet, with z-component of spin = 3/2 and spin = 3/2. When a configuration interaction calculation is done, all spin states of spin equal to, or greater than 3/2 are calculated simultaneously. From these states the quartet states are selected when QUARTET is specified, and all other spin states, while calculated, are ignored in the choice of the root. If QUARTET is used on its own, then a single state corresponding to an alpha electron in each of three M.O.s is calculated.

QUARTET has no meaning in a UHF calculation.

### **QUINTET**

The desired spin state is a quintet, with z-component of spin = 2 and spin = 2. When a configuration interaction calculation is done, all spin states of spin equal to, or greater than 2 are calculated

simultaneously. From these states the quintet states are selected when QUINTET is specified, and the septet states, while calculated, will be ignored in the choice of the root. If QUINTET is used on its own, then a single state, corresponding to an alpha electron in each of four M.O.s is calculated.

QUINTET has no meaning in a UHF calculation.

## RESTART

When a job is stopped by the program *and* the program has created a restart file, the calculation can be continued by specifying RESTART. The most common cause of a job stopping prematurely is the time limit is exceeded. A SADDLE calculation has no restart, but the output file contains information which can easily be used to restart the calculation.

It is not necessary to change the geometric data to reflect the new geometry. The geometry printed at the start of a restarted job will be that of the original data, not that of the restart file.

A convenient way to monitor a long run is to specify 1SCF and RESTART; this will give a normal output file at very little cost.

NOTE: In a FORCE calculation two restarts are possible: a restart in BFGS (subroutine FLEPO) if the geometry was not optimized fully before entering FORCE, and the normal restart in the construction of the force matrix. If the restart is in FLEPO the keyword FORCE should be deleted, and the keyword RESTART should be specified. Forgetting this point is a frequent cause of failed jobs.

## ROOT=n

The nth root of a C.I. calculation will be used in the calculation. If a keyword specifying the spin state is also present, e.g. SINGLET or TRIPLET, then the nth root of that state will be selected. Thus ROOT=3 and SINGLET will select the third singlet root. If ROOT=3 is used on its own, then the third root will be used, which may be a triplet, the third singlet, or the second singlet (the second root might be a triplet). In normal use, this keyword would not be used. It is retained for educational and research purposes. Unusual care should be exercised when ROOT=n is specified.

## ROT=n

In calculation of rotational contributions to the thermodynamic quantities, the symmetry number of the molecule must be supplied. The symmetry number of a point group is the number of equivalent positions attainable by pure rotations. No reflections or improper rotations are allowed. This number cannot be assumed by default, and may be affected by subtle modifications to the molecule, such as isotopic substitution. A list of the most important symmetry numbers follows:

---- TABLE OF SYMMETRY NUMBERS ----

C <sub>1</sub>	C <sub>i</sub>	C <sub>s</sub>	1	D <sub>2</sub>	D <sub>2d</sub>	D <sub>2h</sub>	4	C <sub>∞v</sub>	1
C <sub>2</sub>	C <sub>2v</sub>	C <sub>2h</sub>	2	D <sub>3</sub>	D <sub>3d</sub>	D <sub>3h</sub>	6	D <sub>∞h</sub>	2
C <sub>3</sub>	C <sub>3v</sub>	C <sub>3h</sub>	3	D <sub>4</sub>	D <sub>4d</sub>	D <sub>4h</sub>	8	T <sub>d</sub>	12
C <sub>4</sub>	C <sub>4v</sub>	C <sub>4h</sub>	4	D <sub>6</sub>	D <sub>6d</sub>	D <sub>6h</sub>	12	O <sub>h</sub>	24
C <sub>6</sub>	C <sub>6v</sub>	C <sub>6h</sub>	6	S <sub>6</sub>			3		

## SADDLE

The transition state in a simple chemical reaction  $A \rightarrow B$  will be located. The geometries of both A and B must be supplied. After the first geometry and any symmetry data have been defined, the second geometry is given using the same format and connectivity as that of the first geometry.

SADDLE often fails to work successfully. The most common cause of failure is improper definition of the dihedral angles. As the choice of dihedral can be difficult, users should consider running this calculation with the keyword XYZ. There is normally no ambiguity in the definition of Cartesian coordinates. For a more complete explanation of SADDLE, see the *Examples* and *Background* chapters.

## SCFCRT=n.nnn

The default SCF criterion will be replaced by that defined by SCFCRT=n.nnn. The SCF criterion can be varied from 0.0001 to 0.0000000001. To find a suitable value, 1SCF and various values of SCFCRT=n.nnn should be used; an SCFCRT which allows evaluation of the heat of formation to an acceptable precision can thus be found rapidly. An overly tight criterion can lead to failure to achieve SCF, and consequent failure of the run. Extreme care should be used with this option. The Default value should be acceptable in all but a few special cases.

## SEXTET

The desired spin state is a sextet, with z-component of spin = 5/2 and spin = 5/2. The sextet states are the highest spin states normally calculable using AMPAC in its unmodified form. If SEXTET is used on its own, then a single state, corresponding to one alpha electron in each of five M.O.s, is calculated. If several sextets are to be calculated, say the second or third, then OPEN(n1,n2) should be used.

SEXTET has no meaning in a UHF calculation.

## SHIFT=n.n

SHIFT can be used to obtain an SCF in situations where oscillations (attributed to autoregenerative charge fluctuations) are hindering or preventing convergence. If the virtual M.O.'s are raised in energy relative to the occupied set, then the polarizability of the occupied M.O.'s will decrease. A SHIFT of 20 will raise the virtual M.O.'s by 20 eV above their correct value. The disadvantage of SHIFT is that a large value can lead to excessive damping, and thus prevent convergence. As some virtual M.O.'s are used in non-variationally optimized calculations, SHIFT is automatically annulled at the end of the SCF in these circumstances. All effects of SHIFT are removed before the results are printed. SHIFT should be used with care in that the use of any value for SHIFT will prevent the combined package of convergers (SHIFT, PULAY and CAMP-KING) from automatically being used in the event that the system fails to achieve SCF in 200 iterations. The combined set of convergers almost never fails.

## SIGMA

The McIver-Komornicki gradient norm minimization routines, POWSQ and SEARCH will be used. These routines do not converge to a stationary point for all species; if the gradient norm is low (less than 5 units), SIGMA will probably work. In most cases, NLLSQ or POWELL are recommended. SIGMA first calculates an accurate Hessian matrix (a very slow step), then calculates the direction of fastest descent, and searches along that direction until the gradient norm is minimized. The Hessian is then partially updated in light of the new gradients, and a fresh search direction calculated. Clearly, if the Hessian changes markedly as a result of the line search, the update will be inaccurate, and the new search direction will be faulty.

If at all possible, SIGMA should be avoided when doing non-variationally optimized calculations.

### **SINGLET**

When used with a configuration interaction calculation, all spin states are calculated simultaneously, either for z-component of spin = 0 or 1/2. When only singlet states are of interest, then SINGLET can be specified, and all other spin states, while calculated, are ignored in the choice of the root.

Note that while almost every even electron system will have a singlet ground state, SINGLET should still be specified if the desired state must be a singlet.

SINGLET has no meaning in a UHF calculation, but see also TRIPLET.

### **SPIN**

The spin matrix, defined as the difference between the alpha and beta density matrices, will be printed. If the system has a closed shell ground state, e.g. methane with UHF, the spin matrix will be null.

If SPIN is not requested in a UHF calculation, then the diagonal of the spin matrix, that is the spin density on the atomic orbitals, will be printed.

### **STEP1=n.nnn**

In a grid calculation the step size (in degrees or Angstroms) for the first of two parameters is given by n.nnn. 11 steps in each direction are calculated, giving a total of 121 steps. The origin is in the center at position (6,6).

### **STEP2=n.nnn**

In a grid calculation the step size (in degrees or Angstroms) for the second of two parameters is given by n.nnn.

### **SYMMETRY**

Symmetry data defining related bond lengths, angles and dihedrals must be entered. The additional data is supplied after the geometry. If there are any other data, such as values for the reaction coordinates, or a second geometry (as required by SADDLE or CHAIN), they follow the symmetry data. Symmetry data are terminated by one blank line. For non-variationally optimized systems, symmetry constraints can save a lot of time by reducing the number of derivatives that have to be calculated. At the same time, there is a risk that the geometry may be wrongly specified, e.g. if methane radical cation is defined as being tetrahedral, no indication that this is faulty will be given until a FORCE calculation is done. (This system undergoes spontaneous Jahn-Teller distortion.)

Usually a lower heat of formation can be obtained when SYMMETRY is specified. To see why, consider the geometry of benzene. If no assumptions are made regarding the geometry, then all the C-C bond lengths will be very slightly different, and the angles will be almost, but not quite 120 degrees. Fixing all angles at 120 degrees, dihedrals at 180 or 0 degrees, and only optimizing one C-C and one C-H bond length will result in a 2-D optimization, and exact  $D_{6h}$  symmetry. Any deformation from this symmetry must involve error, so by imposing symmetry the error is removed.

The layout of the symmetry data is:

<defining atom>, <symmetry relation>, <defined atom>, <defined atom>, ...

Commas or spaces are acceptable delimiters. For example, ethane, with three independent variables, can be defined as:

```

SYMMETRY
ETHANE, D3D

C      0.000000 0      0.000000 0      0.000000 0      0 0 0
C      1.528853 1      0.000000 0      0.000000 0      1 0 0
H      1.105161 1 110.240079 1      0.000000 0      2 1 0
H      1.105161 0 110.240079 0 120.000000 0      2 1 3
H      1.105161 0 110.240079 0 240.000000 0      2 1 3
H      1.105161 0 110.240079 0 60.000000 0      1 2 3
H      1.105161 0 110.240079 0 180.000000 0      1 2 3
H      1.105161 0 110.240079 0 300.000000 0      1 2 3
O      0.000000 0      0.000000 0      0.000000 0      0 0 0
3,      1,      4,      5,      6,      7,      8,
3,      2,      4,      5,      6,      7,      8,

```

Here atom 3 (a hydrogen) is used to define the bond length (function 1) of atoms 4, 5, 6, 7 and 8; similarly, its angle (function 2) is used to define the bond angle of atoms 4, 5, 6, 7 and 8. The other angles are point group symmetry defined as multiples of 60 degrees. Note that only three parameters are marked for optimization.

Spaces, tabs or commas can be used to separate data. Symmetry data is ended with a blank line.

The full list of available symmetry relations is as follows:

#### SYMMETRY FUNCTIONS

- |    |   |  |
|----|---|--|
| 1  | BOND LENGTH   | IS SET EQUAL TO THE REFERENCE BOND LENGTH        |
| 2  | BOND ANGLE  | IS SET EQUAL TO THE REFERENCE BOND ANGLE         |
| 3  | DIHEDRAL ANGLE  | IS SET EQUAL TO THE REFERENCE DIHEDRAL ANGLE     |
| 4  | DIHEDRAL ANGLE  | VARIES AS 90 DEGREES - REFERENCE DIHEDRAL        |
| 5  | DIHEDRAL ANGLE  | VARIES AS 90 DEGREES + REFERENCE DIHEDRAL        |
| 6  | DIHEDRAL ANGLE  | VARIES AS 120 DEGREES - REFERENCE DIHEDRAL       |
| 7  | DIHEDRAL ANGLE  | VARIES AS 120 DEGREES + REFERENCE DIHEDRAL       |
| 8  | DIHEDRAL ANGLE  | VARIES AS 180 DEGREES - REFERENCE DIHEDRAL       |
| 9  | DIHEDRAL ANGLE  | VARIES AS 180 DEGREES + REFERENCE DIHEDRAL       |
| 10 | DIHEDRAL ANGLE  | VARIES AS 240 DEGREES - REFERENCE DIHEDRAL       |
| 11 | DIHEDRAL ANGLE  | VARIES AS 240 DEGREES + REFERENCE DIHEDRAL       |
| 12 | DIHEDRAL ANGLE  | VARIES AS 270 DEGREES - REFERENCE DIHEDRAL       |
| 13 | DIHEDRAL ANGLE  | VARIES AS 270 DEGREES + REFERENCE DIHEDRAL       |
| 14 | DIHEDRAL ANGLE  | VARIES AS THE NEGATIVE OF THE REFERENCE DIHEDRAL |
| 15 | BOND LENGTH VARIES AS HALF THE REFERENCE BOND LENGTH    |  |
| 16 | BOND ANGLE VARIES AS HALF THE REFERENCE BOND ANGLE      |  |
| 17 | BOND ANGLE VARIES AS 180 DEGREES - REFERENCE BOND ANGLE |  |
| 18 | BOND LENGTH IS A MULTIPLE OF REFERENCE BOND LENGTH      |  |

Function 18 is intended for use in polymers, in which the translation vector may be a multiple of some bond length. 1, 2, 3 and 14 are most commonly used. *Abbreviation:* SYM.

SYMMETRY is not available for use with Cartesian coordinates.

**T=nnn**

This option allows the program to shut down in an orderly manner on computers with execution time CPU limits.

The total CPU time allowed for the current job is limited to nnn seconds. By default, this is 3600 seconds. If the next cycle of the calculation cannot be completed without running a risk of exceeding the assigned time, the calculation will write a restart file and stop. The safety margin is 100 percent; therefore, enough time to do at least two full cycles must remain.

An alternative specification is T=nn.nnM. This allows time to be defined in minutes, thus T=60M is equivalent to 3600 seconds.

**T-PRIORITY**

In a DRC calculation, results will be printed whenever the calculated time changes by 0.1 femtoseconds. *Abbreviation:* T-PRIO.

**T-PRIORITY=n.nn**

In a DRC calculation, results will be printed whenever the calculated time changes by n.nn femtoseconds.

**T.V.**

The oriented transition vector at the starting point of a PATH calculation will be given in the data file. The components of the transition vector are entered in the same order as the optimized coordinates, in free format, after the weights (if any, see the keyword WEIGHT). If no transition vector is provided, the PATH method will start in the direction of the vanishing driving force encountered.

**THERMO**

The thermodynamic quantities internal energy, heat capacity, partition function, and entropy can be calculated for translation, rotation and vibrational degrees of freedom for a single temperature, or a range of temperatures. Special situations such as linear systems and transition states are accommodated. The approximations used in the THERMO calculation are invalid below 100K.

The program does not automatically recognize internal rotations. If any exist, the calculated quantities will be too low. The TRANS=n keyword can be used to correct this problem.

The TRANS keyword must be used in conjunction with THERMO for transition states.

If THERMO is specified, the default values of the temperature range are assumed (200K to 400K in steps of 10 degrees). Three options exist for overriding the default temperature range, and they are given below.

**THERMO(nnn)**

The thermodynamic quantities for a 200 degree range of temperatures, starting at nnn Kelvin in steps of 10 degrees, will be calculated.

### **THERMO(nnn,mmm)**

The thermodynamic quantities for the temperature range limited by a lower bound of nnn Kelvin and an upper bound of mmm Kelvin will be calculated. The step size is calculated to give approximately 20 points, and a reasonable value for the step. The size of the step will be 1, 2, or 5, times a power of 10.

### **THERMO(nnn,mmm,ooo)**

The thermodynamic quantities for the temperature range limited by a lower bound of nnn Kelvin and an upper bound of mmm Kelvin will be calculated. The step size is given by ooo Kelvin and cannot be less than 1K.

### **TRANS**

The imaginary frequency due to the reaction vector in a transition state calculation must not be included in a thermochemical calculation. The number of genuine vibrations considered are:

3N-5 for a linear ground state system;  
3N-6 for a non-linear ground state system;  
3N-6 for a linear transition state complex;  
3N-7 for a non-linear transition state complex.

This keyword must be used in conjunction with THERMO if a transition state is being calculated.

### **TRANS=n**

TRANS=n can be used with THERMO to handle systems with internal rotations. TRANS=n will remove the n lowest vibrations. Note that TRANS=1 is equivalent to TRANS. For example, TRANS=2 would be suitable for Xylene.

### **TRIPLET**

The triplet state is calculated. If the system has an odd number of electrons, an error message will be printed.

*UHF interpretation:* The number of alpha electrons exceeds that of the beta electrons by 2. If TRIPLET is not specified, then the number of alpha and beta electrons are equal. This does not necessarily correspond to a singlet.

*RHF interpretation:* An RHF MECI calculation is performed to calculate the triplet state. If no other C.I. keywords are used, then only one state is calculated by default. The occupancy of the M.O.s in the SCF calculation is defined as (..2, 1, 1, 0,..). One electron is placed in each of the two highest occupied M.O.s.

See keywords C.I.= and OPEN(n1,n2).

### **UHF**

The unrestricted Hartree-Fock Hamiltonian will be used.

## VECTORS

The full set of eigenvectors (both occupied and virtual) will be printed. In UHF calculations both alpha and beta eigenvectors will be printed. The eigenvectors are normalized to unity, such that the sum of the squares of the coefficients is exactly one. If DEBUG is specified, then ALL eigenvectors on every iteration of every SCF calculation will be printed resulting in incredibly large output files.

## WEIGHT

The weights for a PATH calculation will be given in the data file. Weight data should be entered after the geometry, symmetry, and the microstates data (if any) in the same order as the optimized coordinates, in free format (data separated by a space, a comma or a semi-colon, up to 80 characters per line). If no weights are provided then all the weighting factors are set to 1. This is analogous to following the path of steepest descent in the coordinate system used.

## X-PRIORITY

In a DRC calculation, results will be printed whenever the calculated geometry changes by 0.05 Angstroms. The geometry change is defined as the linear sum of the translation vectors of motion for all atoms in the system. *Abbreviation:* X-PRIO.

## X-PRIORITY=n.nn

In a DRC calculation, results will be printed whenever the calculated geometry changes by n.nn Angstroms.

## XYZ

The SADDLE calculation often fails because of faulty dihedral definitions in the second geometry. To make SADDLE easier to use, XYZ was developed. The keyword XYZ causes the calculation to run entirely in Cartesian coordinates, eliminating the problems associated with properly defining dihedrals. The connectivity of the two systems can be different, but the atom numbering must be the same. Dummy atoms can be used, but will be removed at the start of the run. A new numbering system will be generated by the program, if necessary.

Note that this keyword is unnecessary for molecules that are larger than three atoms. The program automatically checks for the connectivity list required for internal coordinates. If no list is given, the program assumes the geometry is given in Cartesian.

## CHAPTER 3: GEOMETRY SPECIFICATION

All input data in a data file are entered in “free format”. This means that integers and real numbers can be interspersed, and numbers can be separated by one or more spaces, a tab and/or by one comma. If a number is not specified, its value is set to zero. Chemical symbols can be given as numeric data or character data (e.g. “6” or “C” for carbon).

The geometry can be defined in terms of either internal or Cartesian coordinates, but not both simultaneously in the same data file.

### 3.1 INTERNAL COORDINATE DEFINITION

The internal coordinates of a new atom **i** are defined by an interatomic distance (in Angstroms) from a previously defined atom **j**, an interatomic angle (in degrees) between atoms **i**, **j** and a previously defined atom **k** (**k** and **j** must be different atoms), and finally a torsional angle in degrees between atoms **i**, **j**, **k**, and a previously defined atom **l** (**l** cannot be the same as **k** or **j**)

Exceptions:

1. Atom 1 has no coordinates at all: this is the origin.
2. Atom 2 must be connected to atom 1 by an interatomic distance only.
3. Atom 3 should be connected to atom 2, and make an angle with atom 1; no dihedral is possible for atom 3.

Constraints:

1. Interatomic distances must be greater than zero.
2. Angles must be positive.
3. Dihedrals can normally assume definable angles only. If atom **i** makes a dihedral with atoms **j**, **k**, and **l**, and the three atoms **j**, **k**, and **l** are in a straight line, then the dihedral has no definable angle. Thus, atom **i** should not be defined in this manner and a different dihedral must be chosen. During the calculation this constraint is checked continuously, and if atoms **j**, **k**, and **l** lie within 0.1 Angstroms of a straight line then calculation will terminate with an error message. This problem can be corrected by redefining the atom connectivity for atom **i**. NOTE: This does not mean that atoms cannot be in a straight line. Acetylene, for example, must be defined with at least one dummy atom. (See Section 5.3 for more information on dihedrals)

The internal coordinates for each atom are given on a single line of not more than 80 characters. The order in which the internal coordinate data must be entered is given below:

1. Chemical symbol (atom or atomic number).
2. Interatomic distance in Angstroms.
3. Optimization flag for the interatomic distance.

4. Interatomic angle in degrees.
5. Optimization flag for the interatomic angle.
6. Dihedral angle in degrees. The sign is determined by I.U.P.A.C. convention.
7. Optimization flag for the dihedral angle.
8. NA: Atom number defining the interatomic distance ("current atom"-NA).
9. NB: Atom number defining the interatomic angle ("current atom"-NA-NB).
10. NC: Atom number defining the dihedral angle ("current atom"-NA-NB-NC).

An optimization flag is associated with each internal coordinate. The first atom has three unoptimizable coordinates, the second atom has two (the bond length can be optimized), and the third atom has one unoptimizable coordinate. These six unoptimizable coordinates should not be marked for optimization.

Optimization flags are integers that appear after each internal coordinate, indicating the action to be taken.

Integer:	Action:
1	Optimize the internal coordinate.
0	Do not optimize the internal coordinate.
-1	Reaction coordinate, or grid index.

Remarks:

1. Only one reaction coordinate is allowed, but this can be made more versatile by use of SYMMETRY. If a reaction coordinate is used, the values of the reaction coordinate (also in free format) should follow immediately after the geometry and any symmetry data. No terminator is required.
2. If two reaction coordinates are specified, then the two-dimensional space in the region of the supplied geometry will be mapped. The two dimensions to be mapped are defined by the "-1" labels. Step sizes in the two directions must be supplied as STEP1= and STEP2= on the keyword line.

### 3.2 CARTESIAN COORDINATE DEFINITION

The Cartesian coordinates for each atom are given on a single line of less than 80 characters in length. The order in which the Cartesian coordinate data must be entered is given below:

1. Chemical symbol or Atomic Number.
2. X-coordinate in Angstroms.
3. Optimization flag for the X-coordinate.
4. Y-coordinate in Angstroms.
5. Optimization flag for the Y-coordinate.

6. Z-coordinate in Angstroms.
7. Optimization flag for the Z-coordinate.

AMPAC uses the lack of connectivity data as an indication that Cartesian coordinates are being used. A unique case is the triatomics for which only internal coordinates are allowed. This is necessary since connectivity is implied for the first three atoms.

If the keyword XYZ is absent, then every coordinate must be marked for optimization. If any coordinates are not being optimized, then the keyword XYZ must be present. The coordinates of all atoms, including atoms 1, 2 and 3 can be optimized. Dummy atoms should not be used, since they are unnecessary.

### 3.3 DEFINITION OF ELEMENTS AND ISOTOPES

Elements are defined in terms of their atomic numbers or their chemical symbols. Elements parametrized in this version for MNDO and AM1 are marked with a star (\*):

Element	H	Be	B	C	N	O	F	Al
Atomic #	1	4	5	6	7	8	9	13
MNDO	*	*	*	*	*	*	*	*
AM1	*		*	*	*	*	*	*

Element	Si	P	S	Cl	Zn	Ge	Br	Sn
Atomic #	14	15	16	17	30	32	35	50
MNDO	*	*	*	*	*	*	*	*
AM1	*	*	*	*	*	*	*	*

Element	I	Hg	Pb
Atomic #	53	80	82
MNDO	*	*	*
AM1	*	*	

Diatomics Parametrized within the MINDO/3 Formalism:

	H	B	C	N	O	F	Si	P	S	Cl
H	*	*	*	*	*	*	*	*	*	*
B	*	*	*	*	*	*				
C	*	*	*	*	*	*	*	*	*	*
N	*	*	*	*	*	*			*	*
O	*	*	*	*	*	*			*	*
F	*	*	*	*	*	*			*	
Si	*		*				*			
P	*		*					*		*
S	*		*	*	*	*			*	*
Cl	*		*	*	*			*	*	*

A star (\*) indicates that the atom-pair is parametrized within MINDO/3.

Extra entities available to MNDO and AM1:

Symbol	Li	Na	K	XX	++	+	--	-	Tv
#	3	11	19	99	103	104	105	106	107

- + A 100% ionic alkali metal.
- ++ A 100% ionic alkaline earth metal.
- A 100% ionic halogen-like atom
- A 100% ionic group VI-like atom.

Elements 103, 104, 105, and 106 are the sparkles; elements 3, 11 and 19 are sparkles tailored to look like the alkaline metal ions; Tv is the translation vector for polymer calculations.

Element 99, or XX is known as a dummy atom, and is used in the definition of geometries; it is deleted automatically from any Cartesian coordinate geometry files. Dummy atoms are mathematical points that have no chemical meaning. For example, in ammonia the definition of  $C_{3v}$  symmetry is facilitated by using one dummy atom and symmetry relating the three hydrogens to it.

Isotopes are used in conjunction with chemical symbols. If no isotope is specified, the average isotopic mass is used (e.g., chlorine is 35.453). This is a change to previous versions of AMPAC, in which the most abundant isotope was used by default. This change is justified by the removal of any ambiguity in the choice of isotope. Also, the experimental vibrational spectra involve a mixture of isotopes. If specific isotopes are required, they should immediately follow the chemical symbol (no space), thus: H2, H2.0140, C13, C13.00335.

The sparkles ++, +, --, and - have no mass; if they are used in a force calculation, then appropriate masses should be specified.

### 3.4 EXAMPLES OF COORDINATE DEFINITIONS

Use of internal coordinates can best be illustrated by example. The geometry of mono-deuterated formaldehyde can be defined as follows:

Atom 1: Oxygen is at the origin.

Atom 2: Carbon is defined along the x-axis and bonded to atom 1. Connectivity of atom 2 is implied, but can be given if desired.

Atom 3: Deuterium is connected to atom 2 and makes an angle with atom 1 (angle 3-2-1). Again, the connectivity is implied.

Atom 4: Hydrogen is connected to atom 2, makes an angle with atom 1 (angle 4-2-1), and a dihedral with atom 3 (angle 4-2-1-3). Here, the full connectivity must be explicitly defined.

## Chapter 3: Geometry Specification

O				The first atom has no coordinates.						
C	1.2	1		The C-O bond length will be optimized.						
H2.0140	1.0	1	120	1	The third atom is a Deuterium.					
1	1.0	0	120.0	0	180	0	2	1	3	The fourth uses the atomic number of Hydrogen, and SYMMETRY is specified for bond length and angle. The dihedral is point group defined as 180 degrees.
0	0	0								The geometry is terminated with a row of zeros. Symmetry data follows.
	4,	1,	3							
	4,	2,	3							

AMPAC also generates data files, both in the archive files and at the end of a normal output file when a job ends prematurely due to time restrictions. Here, the coordinate definition for formic acid is shown:

O	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.209615	1	0.000000	0	0.000000	0	1	0	0
O	1.313679	1	116.886168	1	0.000000	0	2	1	0
H	0.964468	1	115.553316	1	0.000000	0	3	2	1
H	1.108040	1	128.726078	1	180.000000	0	2	1	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

Note that all coordinates are generated, as is the full connectivity. Also, the data are all neatly lined up. This, of course, is characteristic of machine generated data and is not absolutely necessary for data files.

Polymers are defined by the presence of a translation vector. In the following example, polyethylene, the translation vector spans three monomeric units, and is 7.7 Angstroms long. Note in this example the presence of two dummy atoms. These not only make the geometry definition easier but also allow the translation vector to be specified in terms of distance only, rather than both distance and angles.

T=20000									
POLYETHYLENE, CLUSTER UNIT : C6H12									
C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.540714	1	0.000000	0	0.000000	0	1	0	0
C	1.542585	1	113.532306	1	0.000000	0	2	1	0
C	1.542988	1	113.373490	1	179.823613	1	3	2	1
C	1.545151	1	113.447508	1	179.811764	1	4	3	2
C	1.541777	1	113.859804	1	-179.862648	1	5	4	3
XX	1.527344	1	108.897076	1	-179.732346	1	6	5	4
XX	1.540749	1	108.360151	1	-178.950271	1	7	6	5
H	1.114786	1	90.070026	1	126.747447	1	1	3	2
H	1.114512	1	90.053136	1	-127.134856	1	1	3	2
H	1.114687	1	90.032722	1	126.717889	1	2	4	3
H	1.114748	1	89.975504	1	-127.034513	1	2	4	3
H	1.114474	1	90.063308	1	126.681098	1	3	5	4
H	1.114433	1	89.915262	1	-126.931090	1	3	5	4
H	1.114308	1	90.028131	1	127.007845	1	4	6	5
H	1.114434	1	90.189506	1	-126.759550	1	4	6	5
H	1.114534	1	88.522263	1	127.041363	1	5	7	6
H	1.114557	1	88.707407	1	-126.716355	1	5	7	6
H	1.114734	1	90.638631	1	127.793055	1	6	8	7
H	1.115150	1	91.747016	1	-126.187496	1	6	8	7
Tv	7.746928	1	0.000000	0	0.000000	0	1	3	4
0	0.000000	0	0.000000	0	0.000000	0	0	0	0



## CHAPTER 4: EXAMPLES

In this chapter various examples of data files are described. AMPAC comes with two sets of data for running calculations. One of these is called MNRSD1.DAT, and this will be described now.

### 4.1 MNRSD1 TEST DATA FILE FOR FORMALDEHYDE

The following file is suitable for generating the results described in the next section, and is suitable for debugging data:

```
Line 1: SYMMETRY GRADIENTS GNORM=1 T=590
Line 2: Examples: Formaldehyde
Line 3:
Line 4: O
Line 5: C 1.2 1
Line 6: H 1.1 1 120 1
Line 7: H 1.1 0 120 0 180 0 2 1 3
Line 8:
Line 9: 3 1 4
Line 10: 3 2 4
Line 11:
```

This data could be more neatly written as:

```
Line 1: SYMMETRY GRADIENTS GNORM=1 T=590
Line 2: Examples: Formaldehyde
Line 3:
Line 4: O 0.000000 0 0.000000 0 0.000000 0 0 0 0
Line 5: C 1.200000 1 0.000000 0 0.000000 0 1 0 0
Line 6: H 1.100000 1 120.000000 1 0.000000 0 2 1 0
Line 7: H 1.100000 0 120.000000 0 180.000000 0 2 1 3
Line 8: 0 0.000000 0 0.000000 0 0.000000 0 0 0 0
Line 9: 3, 1, 4,
Line 10: 3, 2, 4,
Line 11:
```

These two data files will produce identical results.

In all geometry specifications care must be taken in defining the internal coordinates to ensure that no three atoms used to define a fourth atom's dihedral angle can ever form a straight line. This can happen in the course of a geometry optimization, in a SADDLE calculation or in following a reaction coordinate. If such a condition should develop, then the position of the dependent atom will become ill defined.

## 4.2 OUTPUT FILE FOR TEST DATA FILE MNRSD1

\*\*\*\*\*

## MNDO CALCULATION RESULTS

Note 1

\*\*\*\*\*

\* VERSION 2.10 Note 2

\* - MINIMIZE ENERGY USING B-F-G-S METHOD

\* GNORM= - OPTIMIZATION EXIT WHEN GRADIENT NORM BELOW 1.000

\* SYMMETRY - SYMMETRY CONDITIONS TO BE IMPOSED

\* GRADIENTS- ALL GRADIENTS TO BE PRINTED

\* T= - A TIME OF 590.0 SECONDS REQUESTED

\*\*\*\*\*

## PARAMETER DEPENDENCE DATA

Note 3

REFERENCE ATOM	FUNCTION NO.	DEPENDENT ATOM(S)
3	1	4
3	2	4

## DESCRIPTIONS OF THE FUNCTIONS USED

1 BOND LENGTH IS SET EQUAL TO THE REFERENCE BOND LENGTH

2 BOND ANGLE IS SET EQUAL TO THE REFERENCE BOND ANGLE

SYMMETRY GRADIENTS GNORM=1 T=590

Examples: Formaldehyde

ATOM NUMBER (I)	CHEMICAL SYMBOL	BOND LENGTH (ANGSTROMS) NA:I	BOND ANGLE (DEGREES) NB:NA:I	TWIST ANGLE (DEGREES) NC:NB:NA:I	NA	NB	NC
1	O						
2	C	1.20000 *			1		
3	H	1.10000 *	120.00000 *		2	1	
4	H	1.10000	120.00000	180.00000	2	1	3

Note 4

## CARTESIAN COORDINATES

Note 5

NO.	ATOM	X	Y	Z
1	8	0.0000	0.0000	0.0000
2	6	1.2000	0.0000	0.0000
3	1	1.7500	0.9526	0.0000
4	1	1.7500	-0.9526	0.0000

RHF CALCULATION, NO. OF DOUBLY OCCUPIED LEVELS = 6

## INTERATOMIC DISTANCES

Note 6

0

	O 1	C 2	H 3	H 4
O 1	0.000000			
C 2	1.200000	0.000000		
H 3	1.992486	1.100000	0.000000	
H 4	1.992486	1.100000	1.905256	0.000000

STANDARD DEVIATION ON ENERGY (KCAL) 0.00029121 Note 7

STEP LENGTH FOR FIRST DERIVATIVES 0.00010000 0.00010000 0.00010000

STANDARD DEVIATION OF GRADIENT (KCAL/A,RD) 0.05433654 0.05433654 0.05433654

USE 1-POINT FORWARD FINITE DIFFERENCE IN INTEGRAL DERIVATIVES AND JACOBIAN

SYMMETRY GRADIENTS GNORM=1 T=590 Note 8

Examples: Formaldehyde

PETERS TEST WAS SATISFIED IN FLETCHER-POWELL OPTIMISATION

SCF FIELD WAS ACHIEVED

Note 9

# Chapter 4: Examples

MNDO CALCULATION

VERSION 2.10

FINAL HEAT OF FORMATION = -32.881899 KCAL

**Note 10**

ELECTRONIC ENERGY = -870.727873 EV  
CORE-CORE REPULSION = 392.608700 EV

GRADIENT NORM = 0.044038  
IONISATION POTENTIAL = 11.041363

2-SEP-88

NO. OF FILLED LEVELS = 6  
MOLECULAR WEIGHT = 30.026

SCF CALCULATIONS = 13  
COMPUTATION TIME = 8.73 SECONDS

**Note 11**

FINAL POINT AND DERIVATIVES

**Note 12**

PARAMETER	ATOM	TYPE	VALUE	GRADIENT	
1	2 C	BOND	1.216482	-0.022588	KCAL/ANGSTROM
2	3 H	BOND	1.106104	-0.037802	KCAL/ANGSTROM
3	3 H	ANGLE	123.510547	-0.000366	KCAL/RADIAN

ATOM NUMBER (I)	CHEMICAL SYMBOL	BOND LENGTH (ANGSTROMS) NA:I	BOND ANGLE (DEGREES) NB:NA:I	TWIST ANGLE (DEGREES) NC:NB:NA:I	NA	NB	NC
-----------------------	--------------------	------------------------------------	------------------------------------	--	----	----	----

1	O						
2	C	1.21648 *			1		
3	H	1.10610 *	123.51055 *		2	1	
4	H	1.10610	123.51055	180.00000	2	1	3

**Note 13**

INTERATOMIC DISTANCES

		O 1	C 2	H 3	H 4
O 1		0.000000			
C 2		1.216482	0.000000		
H 3		2.046712	1.106104	0.000000	
H 4		2.046712	1.106104	1.844503	0.000000

EIGENVALUES

**Note 14**

-42.99130	-25.11902	-16.95440	-16.29939	-14.17950	-11.04136	0.85977	3.67568
3.84927	7.12606						

NET ATOMIC CHARGES AND DIPOLE CONTRIBUTIONS

**Note 15**

	ATOM NO.	TYPE	CHARGE	ATOM ELECTRON DENSITY
	1	O	-0.2902	6.2902
	2	C	0.2921	3.7079
	3	H	-0.0010	1.0010
	4	H	-0.0010	1.0010
DIPOLE	X	Y	Z	TOTAL
POINT-CHG.	1.690	0.000	0.000	1.690
HYBRID	0.475	0.000	0.000	0.475
SUM	2.165	0.000	0.000	2.165

**Note 16**

CARTESIAN COORDINATES				
NO.	ATOM	X	Y	Z
1	O	0.0000	0.0000	0.0000
2	C	1.2165	0.0000	0.0000
3	H	1.8272	0.9223	0.0000
4	H	1.8272	-0.9223	0.0000

ATOMIC ORBITAL ELECTRON POPULATIONS							<b>Note 17</b>
1.88260	1.21602	1.89108	1.30049	1.25530	0.86214	0.89091	0.69951
1.00097	1.00097						

## NOTES ON OUTPUT FILE

**NOTE 1:** The banner indicates whether the calculation uses a MNDO, MINDO/3 or AM1 Hamiltonian. Here, the default MNDO Hamiltonian is used.

**NOTE 2:** The version number is a constant for any release of AMPAC, and refers to the program, not to the Hamiltonians used. The version number should be cited in any correspondence regarding AMPAC.

All the keywords used, along with a brief explanation, are also printed at this time. If a keyword is not printed, it has not been recognized by the program. Keywords can be in upper or lower case letters, or any mixture.

**NOTE 3:** Symmetry information is printed to verify that the requested symmetry functions have been recognized and are being used.

**NOTE 4:** The initial internal coordinates are printed along with chemical symbols for each atom. Only the geometric parameters with asterisks beside them will be optimized during geometry optimization.

**NOTE 5:** The Cartesian coordinates are calculated as follows:

Stage 1: The coordinate of the first atom is defined to be the origin of Cartesian space, while the coordinate of the second atom is displaced by its defined bond length along the positive x-axis. The coordinate of the third atom is displaced by its bond length in the x-y plane, from atom 2. The angle it makes with atoms 1 and 2 is that given by its bond angle. The dihedral, which first appears in the fourth atom, is defined according to the I.U.P.A.C. convention.

**NOTE:** This is different from previous versions of MNDO and MINDO/3, where the dihedral had the opposite chirality to that defined by the I.U.P.A.C. convention.

Stage 2: Any dummy atoms are removed from the Cartesian output. As this particular system contains no dummy atoms, nothing is done.

**NOTE 6:** The interatomic distances are printed for the user's convenience, and a simple check made to insure that the smallest interatomic distance is greater than 0.8 Å. If such distances are encountered, the program terminates execution. If the user wishes to continue the calculation with bond lengths less than 0.8 Å, then the keyword GEO-OK should be specified.

**NOTE 7:** The geometry is optimized in a series of cycles, each cycle consisting of a line search and calculation of the gradients. Various criteria for the SCF cycle, derivative calculation and geometry optimization are given. By specifying a higher print option (PRINT=n) more information can be obtained. Note that in this example PRECISE was not specified and thus only one step forward finite difference was used in the gradient calculation.

*NOTE 8:* The keywords used, titles and comments are reproduced here to remind the user of the name of the calculation.

*NOTE 9:* Two messages are given here. The first is a reminder of how the geometry was obtained. At the end of the Davidon-Fletcher-Powell geometry optimization a message is given which indicates how the optimization ended. All “normal” termination messages contain the word “satisfied”; other terminations may give acceptable results, but more care should be taken, particularly regarding the gradient vector.

For any further results to be printed the message: SCF FIELD WAS ACHIEVED must be as shown. When SCF is not obtained no results will be printed.

*NOTE 10:* The total energy of the system is the addition of the electronic and nuclear terms. The heat of formation is relative to the elements in their standard state. The I.P. is the negative of the energy level of the highest occupied, or highest partially occupied molecular orbital (in accordance with Koopmans' theorem).

*NOTE 11:* C.P.U. time used for the calculation. This number is of course highly machine dependent.

*NOTE 12:* Gradient components for each geometric parameter flagged for optimization are printed when the GRADIENTS option is specified. These numbers will be small if the geometry is at a stationary point on the potential surface.

*NOTE 13:* The fully optimized geometry is printed here. If a parameter is not marked for optimization, it will not be changed unless it is a symmetry related parameter.

*NOTE 14:* The eigenvalues or energy levels in electron volts of the molecular orbitals are printed here. There are six filled levels, therefore the HOMO has an energy of -11.041 eV, and is mainly a lone pair on oxygen. The eigenvectors form an orthonormal set, and can be printed with the keyword VECTOR.

*NOTE 15:* The charge on an atom is the sum of the positive core charge (for hydrogen, carbon, and oxygen these numbers are 1.0, 4.0, and 6.0, respectively) and the negative of the number of valence electrons (ATOM ELECTRON DENSITY, here 1.0010, 3.7079, and 6.2902 respectively).

*NOTE 16:* The dipole is the scalar of the dipole vector in Cartesian coordinates. The components of the vector coefficients are the point charge dipole and the hybridization dipole. In formaldehyde there is no z-dipole since the molecule is flat.

*NOTE 17:* MNDO AM1, and MINDO/3 all use the Coulson density matrix. Only the diagonal elements of the matrix representing the valence orbital electron populations will be printed, unless the keyword DENSITY is specified.

### 4.3 DATA FILE FOR A FORCE CALCULATION

This example is taken from the first data file in TESTDATA.DAT, and illustrates the working of a FORCE calculation.

```

Line 1 : ROT=2 THERMO(298,298,) PRECISE FORCE      ISOTOPE SYMMETRY
Line 2 : DEMONSTRATION OF AMPAC - FORCE AND THERMODYNAMICS CALCULATION
Line 3 :   FORMALDEHYDE, MNDO ENERGY = -32.8819
Line 4a:  O   0.000000  0   0.000000  0   0.000000  0   0  0  0
Line 4b:  C   1.216486  1   0.000000  0   0.000000  0   1  0  0
Line 4c:  H   1.106188  1  123.515021  1   0.000000  0   2  1  0
Line 4d:  H   1.106188  1  123.515021  1  179.995690  1   2  1  3
Line 4e:  O   0.000000  0   0.000000  0   0.000000  0   0  0  0
Line 5a:  3,   1,   4,
Line 5b:  3,   2,   4,

```

### 4.4 OUTPUT FILE FROM A FORCE CALCULATION

```

*****
                                MNDO CALCULATION RESULTS                                Note 1
*****

*                               VERSION 2.10                                         Note 2
*  FORCE      - FORCE CALCULATION SPECIFIED
*  THERMO    - THERMODYNAMIC QUANTITIES TO BE CALCULATED
*  ROT       - SYMMETRY NUMBER OF 2 SPECIFIED
*  PRECISE   - OPTIMIZATION CRITERIA TO BE INCREASED BY 10 TIMES,
*             - S.C.F. CRITERIA BY 100 TIMES,
*             - AND USE ACCURATE FINITE DIFFERENCE FORMULA IN HESSIAN
*  SYMMETRY  - SYMMETRY CONDITIONS TO BE IMPOSED
*  ISOTOPE   - FORCE MATRIX WRITTEN TO DISK (CHAN. 9 )
*  T=        - A TIME OF 3590.0 SECONDS REQUESTED
*****

PARAMETER DEPENDENCE DATA                                                         Note 3

REFERENCE ATOM      FUNCTION NO.      DEPENDENT ATOM(S)
    3                1                4
    3                2                4

DESCRIPTORS OF THE FUNCTIONS USED

1      BOND LENGTH      IS SET EQUAL TO THE REFERENCE BOND LENGTH
2      BOND ANGLE       IS SET EQUAL TO THE REFERENCE BOND ANGLE
ROT=2 THERMO(298,298,) PRECISE FORCE ISOTOPE SYMMETRY T=3590
DEMONSTRATION OF AMPAC - FORCE AND THERMODYNAMICS CALCULATION
FORMALDEHYDE, MNDO ENERGY = -32.8819
Note 4

ATOM      CHEMICAL      BOND LENGTH      BOND ANGLE      TWIST ANGLE
NUMBER    SYMBOL        (ANGSTROMS)    (DEGREES)      (DEGREES)
(I)              NA:I          NB:NA:I        NC:NB:NA:I    NA  NB  NC

    1      O
    2      C      1.21649 *
    3      H      1.10619 *      123.51502 *
    4      H      1.10619 *      123.51502 *      179.99569 *
                                1      2      1      3
                                Note 5

```

## Chapter 4: Examples

CARTESIAN COORDINATES					Note 6
NO.	ATOM	X	Y	Z	
1	8	0.0000	0.0000	0.0000	
2	6	1.2165	0.0000	0.0000	
3	1	1.8273	0.9223	0.0000	
4	1	1.8273	-0.9223	0.0001	
RHF CALCULATION, NO. OF DOUBLY OCCUPIED LEVELS = 6					
INTERATOMIC DISTANCES					Note 7
0	O 1	C 2	H 3	H 4	
	O 1	0.000000			
	C 2	1.216486	0.000000		
	H 3	2.046832	1.106188	0.000000	
	H 4	2.046832	1.106188	1.844549	0.000000
HEAT OF FORMATION = -32.881894 KCALS/MOLE					Note 8
STANDARD DEVIATION ON ENERGY (KCAL)					0.00000077
STEP LENGTH FOR FIRST DERIVATIVES					0.00010000 0.00010000 0.00010000
STANDARD DEVIATION OF GRADIENT (KCAL/A,RD)					0.00023513 0.00023513 0.00023513
USE 2-POINTS CENTRAL FINITE DIFFERENCES IN INTEGRAL DERIVATIVES AND JACOBIAN					
INTERNAL COORDINATE DERIVATIVES					Note 9
ATOM	AT. NO.	BOND	ANGLE	DIHEDRAL	
1	8				
2	6	0.020562			
3	1	0.140775	0.010503		
4	1	0.000000	0.000000	-0.002859	
GRADIENT NORM = 0.14268					Note 10
TIME FOR SCF CALCULATION = 1.10					Note 11
TIME FOR DERIVATIVES = 0.96					
SYMMETRY WAS SPECIFIED, BUT CANNOT BE USED HERE					
MOLECULAR WEIGHT = 30.03					
PRINCIPAL MOMENTS OF INERTIA IN CM(-1)					
A = 9.831942 B = 1.261996 C = 1.118437					
PRINCIPAL MOMENTS OF INERTIA IN UNITS OF 10**(-40)*GRAM-CM**2					
A = 2.847177 B = 22.181748 C = 25.028925					
ORIENTATION OF MOLECULE IN FORCE CALCULATION					Note 12
NO.	ATOM	X	Y	Z	
1	8	-0.6093	0.0000	0.0000	
2	6	0.6072	0.0000	0.0000	
3	1	1.2180	0.9223	0.0000	
4	1	1.2180	-0.9223	0.0001	

FIRST DERIVATIVES WILL BE USED IN THE CALCULATION OF SECOND DERIVATIVES

**Note 13**

TIME DEFINED FOR THIS STEP = 3590.00 SECONDS

ESTIMATED TIME TO COMPLETE CALCULATION = 49.44 SECONDS

STEP:	1	TIME =	2.31 SECS,	INTEGRAL =	2.31	TIME LEFT:	3582.90
STEP:	2	TIME =	2.59 SECS,	INTEGRAL =	4.90	TIME LEFT:	3580.31
STEP:	3	TIME =	2.61 SECS,	INTEGRAL =	7.51	TIME LEFT:	3577.70
STEP:	4	TIME =	2.33 SECS,	INTEGRAL =	9.84	TIME LEFT:	3575.37
STEP:	5	TIME =	2.51 SECS,	INTEGRAL =	12.35	TIME LEFT:	3572.86
STEP:	6	TIME =	2.35 SECS,	INTEGRAL =	14.70	TIME LEFT:	3570.51
STEP:	7	TIME =	2.28 SECS,	INTEGRAL =	16.98	TIME LEFT:	3568.23
STEP:	8	TIME =	2.25 SECS,	INTEGRAL =	19.23	TIME LEFT:	3565.98
STEP:	9	TIME =	2.63 SECS,	INTEGRAL =	21.86	TIME LEFT:	3563.35
STEP:	10	TIME =	2.46 SECS,	INTEGRAL =	24.32	TIME LEFT:	3560.89
STEP:	11	TIME =	2.23 SECS,	INTEGRAL =	26.55	TIME LEFT:	3558.66
STEP:	12	TIME =	2.69 SECS,	INTEGRAL =	29.24	TIME LEFT:	3555.97

FORCE MATRIX WRITTEN TO DISK

FORCE MATRIX IN MILLIDYNES/ANGSTROM

**Note 14**

0

	O	1	C	2	H	3	H	4
O 1	9.557770							
C 2	8.683373	11.426273						
H 3	0.598735	2.552464	3.033930					
H 4	0.598754	2.552449	0.304236	3.033911				

HEAT OF FORMATION = -32.881894 KCALS/MOLE

TRIVIAL VIBRATIONS, SHOULD BE ZERO

**Note 15**

0.0000=TX 0.0000=TY 0.0000=TZ 0.0010=RX 0.0002=RY 0.0001=RZ

FORCE CONSTANTS IN MILLIDYNES/ANGSTROM (= 10\*\*5 DYNES/CM)

**Note 16**

0.94407 1.35190 1.86908 4.07259 6.74507 18.72183

ASSOCIATED EIGENVECTORS

**Note 17**

ROOT NO.	1	2	3	4	5	6
	0.94407	1.35190	1.86908	4.07259	6.74507	18.72183
1	-0.33016	0.00000	-0.00002	0.41812	0.00000	0.68277
2	0.00000	0.60834	-0.00007	0.00000	-0.04810	0.00000
3	-0.00001	-0.00005	-0.28948	0.00000	0.00000	0.00000
4	-0.34275	0.00000	-0.00002	0.32022	0.00000	-0.72800
5	0.00000	-0.34893	0.00006	0.00000	0.79262	0.00000
6	0.00000	0.00009	0.86602	0.00004	-0.00003	0.00000
7	0.33646	-0.48710	0.00006	-0.36917	-0.21481	0.02261
8	-0.52302	-0.12970	0.00002	-0.47438	-0.37226	0.03759
9	0.00003	-0.00003	-0.28827	0.00000	0.00000	0.00000
10	0.33645	0.48710	-0.00003	-0.36917	0.21481	0.02261
11	0.52303	-0.12970	-0.00002	0.47437	-0.37226	-0.03758
12	-0.00001	-0.00002	-0.28827	-0.00004	0.00003	0.00000

ZERO POINT ENERGY 18.105 KILOCALORIES PER MOLE

THE LAST 6 VIBRATIONS ARE THE TRANSLATION AND ROTATION MODES  
 THE FIRST THREE OF THESE BEING TRANSLATIONS IN X, Y, AND Z, RESPECTIVELY

## FREQUENCIES, REDUCED MASSES AND VIBRATIONAL DIPOLES

Note 18

I	1	2	3	4	5	6
FREQ(I)	1209.9655	1214.9634	1490.6029	2114.5763	3255.3630	3301.5739
MASS(I)	0.00000	0.55288	0.00000	0.01407	0.16617	0.21482
DIPX(I)	-0.00002	-0.00002	0.34437	3.36649	-0.00004	-0.34785
DIPY(I)	0.85477	-0.00379	0.00005	0.00001	0.78301	-0.00005
DIPZ(I)	-0.00063	-0.12739	0.00001	0.00008	-0.00003	-0.00001
DIPT(I)	0.85477	0.12745	0.34437	3.36649	0.78301	0.34785
I	7	8	9	10	11	12
FREQ(I)	-0.0052	0.0002	-0.0028	57.3134	11.5938	9.0138
MASS(I)	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
DIPX(I)	0.00006	-0.00002	-0.00001	-0.00013	-0.00002	0.00001
DIPY(I)	-0.00003	0.00004	0.00012	0.00001	0.00018	0.46524
DIPZ(I)	0.00000	0.00001	-0.00004	0.00001	0.56208	-0.00002
DIPT(I)	0.00007	0.00004	0.00013	0.00013	0.56208	0.46524

## NORMAL VECTORS

Note 19

ROOT NO.	1	2	3	4	5	6
	1209.96549	1214.96342	1490.60293	2114.57633	3255.36297	3301.57392
1	0.00000	0.00000	0.04332	-0.52496	0.00000	0.00070
2	-0.07921	0.00037	0.00000	0.00000	0.00431	0.00000
3	0.00021	0.04399	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	0.03976	0.66818	0.00000	-0.06524
5	0.15855	-0.00074	0.00000	0.00000	0.08917	-0.00001
6	-0.00082	-0.17532	0.00000	0.00000	0.00000	0.00000
7	0.62001	-0.00290	-0.58075	0.18532	-0.41969	0.38321
8	-0.31598	0.00148	0.40127	0.32347	-0.56551	0.59254
9	0.00324	0.69545	-0.00002	-0.00001	0.00000	-0.00001
10	-0.62004	0.00286	-0.58073	0.18530	0.41975	0.38314
11	-0.31600	0.00151	-0.40125	-0.32346	-0.56560	-0.59245
12	0.00326	0.69545	0.00001	0.00001	0.00004	0.00004
ROOT NO.	7	8	9	10	11	12
	-0.00518	0.00022	-0.00283	57.31344	11.59385	9.01384
1	0.50000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	-0.50000	0.00000	0.00000	-0.00001	-0.26200
3	0.00000	0.00000	0.50000	0.00000	-0.31646	0.00001
4	0.50000	0.00000	0.00000	0.00000	0.00000	0.00000
5	0.00000	-0.50000	0.00000	0.00000	0.00001	0.26110
6	0.00000	0.00000	0.50000	0.00000	0.31537	-0.00001
7	0.50000	0.00000	0.00000	0.00000	-0.00002	-0.39659
8	0.00000	-0.50000	0.00000	-0.00001	0.00002	0.52375
9	0.00000	0.00000	0.50000	-0.70711	0.63261	-0.00002
10	0.50000	0.00000	0.00000	0.00000	-0.00002	0.39659
11	0.00000	-0.50000	0.00000	0.00005	0.00002	0.52375
12	0.00000	0.00000	0.50000	0.70711	0.63261	-0.00002

## DESCRIPTION OF VIBRATIONS

Note 20

VIB.	1	ATOMS	C 2	AND	H 3	SHIFT	0.78	ANGSTROMS	6.8%	RADIALLY
FREQ.		1209.97	C 2		H 4		0.78		6.8%	
VIB.	2	ATOMS	C 2	AND	H 3	SHIFT	0.87	ANGSTROMS	0.0%	RADIALLY
FREQ.		1214.96	C 2		H 4		0.87		0.0%	

VIB.	3	ATOMS	C 2	AND	H 3	SHIFT	0.74	ANGSTROMS	1.1%	RADIALLY
FREQ.		1490.60	C 2		H 4		0.74		1.1%	
VIB.	4	ATOMS	O 1	AND	C 2	SHIFT	1.19	ANGSTROMS	100.0%	RADIALLY
FREQ.		2114.58	C 2		H 3		0.58		0.5%	
			C 2		H 4		0.58		0.5%	
VIB.	5	ATOMS	C 2	AND	H 3	SHIFT	0.78	ANGSTROMS	100.0%	RADIALLY
FREQ.		3255.36	C 2		H 4		0.78		100.0%	
VIB.	6	ATOMS	C 2	AND	H 3	SHIFT	0.74	ANGSTROMS	99.8%	RADIALLY
FREQ.		3301.57	C 2		H 4		0.74		99.8%	
VIB.	10	ATOMS	C 2	AND	H 3	SHIFT	0.71	ANGSTROMS	0.0%	RADIALLY
FREQ.		57.31	C 2		H 4		0.71		0.0%	
VIB.	11	ATOMS	O 1	AND	C 2	SHIFT	0.63	ANGSTROMS	0.0%	RADIALLY
FREQ.		11.59	C 2		H 3		0.32		0.0%	
			C 2		H 4		0.32		0.0%	
VIB.	12	ATOMS	O 1	AND	C 2	SHIFT	0.52	ANGSTROMS	0.0%	RADIALLY
FREQ.		9.01	C 2		H 3		0.48		0.0%	
			C 2		H 4		0.48		0.0%	

SYSTEM IS A GROUND STATE

FORMALDEHYDE, MNDO ENERGY = -32.8819  
 DEMONSTRATION OF AMPAC - FORCE AND THERMODYNAMICS CALCULATION

MOLECULE IS NOT LINEAR

THERE ARE 6 GENUINE VIBRATIONS IN THIS SYSTEM  
 THIS THERMODYNAMICS CALCULATION IS LIMITED TO  
 MOLECULES WHICH HAVE NO INTERNAL ROTATIONS

#### CALCULATED THERMODYNAMIC PROPERTIES

**Note 21**

TEMP. (K)	PARTITION FUNCTION	ENTHALPY CAL/MOL	HEAT CAPACITY CAL/K/MOL	ENTROPY CAL/K/MOL
298	VIB. 1.007	23.37963966	0.47814559	0.09144464
	ROT. 709.	888.305	2.981	16.026
	INT. 714.	911.685	3.459	16.117
	TRA. 0.159E+27	1480.509	4.968	36.113
	TOT.	2392.1936	8.4272	52.2300
COMPUTATION TIME	33.84 SECONDS			

#### NOTES ON FORCE OUTPUT FILE

**NOTE 1:** The banner indicates whether the calculation uses a MNDO, MINDO/3 or AM1 Hamiltonian. Here, the default MNDO Hamiltonian is used.

**NOTE 2:** The version number is a constant for any release of AMPAC, and refers to the program, not to the Hamiltonians used. The version number should be cited in any correspondence regarding AMPAC.

All three words, ROT, FORCE, and THERMO are necessary in order to obtain thermodynamic properties. In order to obtain results for only one temperature, THERMO has the first and second arguments identical. The symmetry number for the C<sub>2v</sub> point group is 2. If the system was a transition state then the keyword TRANS must be specified to ensure that the imaginary frequency is not included in the thermochemical computation.

**NOTE 3:** Symmetry information is printed. All symmetry definitions are ignored in a force calculation.

*NOTE 4:* The keywords used, titles and comments are reproduced here to remind the user of the name of the calculation.

*NOTE 5:* The initial internal coordinates are printed along with chemical symbols for each atom.

*NOTE 6:* The Cartesian coordinates are calculated as follows:

Stage 1: The coordinate of the first atom is defined to be at the origin of Cartesian space, while the coordinate of the second atom is defined to be displaced by its defined bond length along the positive x-axis. The coordinate of the third atom is defined as being displaced by its bond length in the xy plane, from 2. The angle it makes with atoms 1 and 2 is given by its bond angle. The dihedral, which first appears in the fourth atom, is defined according to the I.U.P.A.C. convention.

*NOTE:* This is different from previous versions of MNDO and MINDO/3, where the dihedral had the opposite chirality to that defined by the I.U.P.A.C. convention.

Stage 2: Any dummy atoms are removed. As this particular system contains no dummy atoms, nothing is done.

*NOTE 7:* The interatomic distances are printed.

*NOTE 8:* Heat of Formation and the calculation criteria are printed.

*NOTE 9:* Internal coordinate derivatives are in kcal/Å or kcal/radian. Values of less than 0.2 are quite acceptable.

*NOTE 10:* If the gradient norm is larger than 5, the program will attempt a geometry optimization before doing the force calculation. In this case, the gradient norm is sufficiently small so the force calculation continues. To inhibit automatic re-optimization the keyword LET should be specified.

*NOTE 11:* In larger calculations, the time estimates are useful. In practice they are pessimistic, and only about 70% of the time estimated will be used.

The principal moments of inertia can be directly related to the microwave spectrum of the molecule. They are simple functions of the geometry of the system, and are usually predicted with very high accuracy.

*NOTE 12:* Proper orientation of the molecule is necessary to examine the force eigenvectors. Always use these Cartesian coordinates when doing a Normal Mode analysis.

*NOTE 13:* Second derivatives are calculated for the construction of the Hessian, along with required times.

*NOTE 14:* The force matrix is printed and written to disk for the isotope calculation.

*NOTE 15:* The trivial vibrations are (in order) translation in x, y, and z, and rotation in x, y, and z. Errors in the force calculation appear in Tx, Ty, and Tz, and in the geometry in Rx, Ry, and Rz. Ideally, all six should be zero, but values of Tx, Ty, and Tz less than 0.0050, and Rx, Ry, and Rz less than 0.0080 are acceptable. For linear systems there will only be 5 trivial vibrations, and for polymers only 4. For polymers, only the three translations are considered as trivial; the single trivial rotation is considered in the genuine force constants, and should be very small.

*NOTES 16 and 17:* Force constants and their associated eigenvectors are printed. Since all are positive, the stationary point is a minimum. These are the “genuine” force constants; there will normally be 3N-6 of these.

**NOTE 18:** The transition dipoles are useful in determining the intensity of I.R. absorption bands. The intensity is proportional to the square of the transition dipole times the vibrational frequency.

**NOTE 19 and 20:** These are the normal modes of vibration followed by a description of the vibrations. This description is incomplete and is meant only to be a guide to the interpretation of the vibrational modes. The last three, five or six are the “trivial” vibrations, and should be small.

**NOTE 21:** In order, the thermodynamic quantities calculated are:

- (1) The vibrational contribution.
- (2) The rotational contribution.
- (3) The sum of (1) and (2) gives the internal contribution.
- (4) The translational contribution.

For partition functions the various contributions are multiplied together.

## 4.5 DATA FILE FOR A REACTION PATH WITH SYMMETRY

In this example, one methyl group in ethane is rotated relative to the other and the geometry is optimized at each point. As the reaction coordinate involves three hydrogen atoms moving, symmetry is imposed to ensure equivalence of all hydrogens.

```

Line 1:          SYMMETRY   T=600
Line 2:    ROTATION OF METHYL GROUP IN ETHANE
Line 3:    EXAMPLE OF A REACTION PATH CALCULATION
Line 4:      6
Line 5:      6   1.479146 1
Line 6:      1   1.109475 1 111.328433 1
Line 7:      1   1.109470 0 111.753160 0 120.000000 0 2 1 3
Line 8:      1   1.109843 0 110.103163 0 240.000000 0 2 1 3
Line 9:      1   1.082055 0 121.214083 0 60.000000 -1 1 2 3
Line 10:     1   1.081797 0 121.521232 0 180.000000 0 1 2 3
Line 11:     1   1.081797 0 121.521232 0 -60.000000 0 1 2 3
Line 12:     0   0.000000 0 0.000000 0 0.000000 0 0 0 0
Line 13:     3 1 4 5 6 7 8
Line 14:     3 2 4 5 6 7 8
Line 15:     6 7 7
Line 16:     6 11 8
Line 17:
Line 18:      70 80 90 100 110 120 130 140 150

```

Points to note:

1. The dihedrals of the second and third hydrogens are not marked for optimization. The dihedrals follow from point group symmetry.
2. All six C-H bond lengths and H-C-C angles are related by symmetry (see lines 13 and 14).
3. The dihedral on line 9 is the reaction coordinate, while the dihedrals on lines 10 and 11 are related to it by symmetry functions on lines 15 and 16 (see SYMMETRY for definitions of functions 1, 2, 7, and 11).
4. Symmetry data are ended by a blank line.
5. The reaction coordinate data are ended by the end of file; several lines of data are allowed.

6. The files generated by the calculation will contain results for each reaction coordinate specified.

## 4.6 DATA FILE FOR A SADDLE CALCULATION

A SADDLE calculation uses two complete geometries, as shown in the following data file for the ethyl radical hydrogen migration from one methyl group to the other.

```

Line 1:          UHF SADDLE
Line 2:          ETHYL RADICAL HYDROGEN MIGRATION
Line 3:
Line 4:    6      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 5:    6      1.479146 1      0.000000 0      0.000000 0      1 0 0
Line 6:    1      1.109475 1     111.328433 1      0.000000 0      2 1 0
Line 7:    1      1.109470 1     111.753160 1     120.288410 1      2 1 3
Line 8:    1      1.109843 1     110.103163 1     240.205278 1      2 1 3
Line 9:    1      1.082055 1     121.214083 1      38.110989 1      1 2 3
Line 10:   1      1.081797 1     121.521232 1     217.450268 1      1 2 3
Line 11:   0      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 12:   6      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 13:   6      1.479146 1      0.000000 0      0.000000 0      1 0 0
Line 14:   1      1.109475 1     111.328433 1      0.000000 0      2 1 0
Line 15:   1      1.109470 1     111.753160 1     120.288410 1      2 1 3
Line 16:   1      2.109843 1      30.103163 1     240.205278 1      2 1 3
Line 17:   1      1.082055 1     121.214083 1      38.110989 1      1 2 3
Line 18:   1      1.081797 1     121.521232 1     217.450268 1      1 2 3
Line 19:   0      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 20:

```

Points to note:

The two geometries must be related by a continuous deformation of the coordinates. By default, internal coordinates are used in specifying geometries. While bond lengths and bond angles are unambiguously defined (both positive), the dihedral angles can be positive or negative. For example, 300 degrees can also be specified as -60 degrees. A wrong choice of dihedral angle would yield a reaction vector that does not follow a smooth path from reactant to product, and produces disastrous results.

This can be corrected by defining all dihedrals to be positive for the first geometry and considering whether the dihedral increases or decreases for each coordinate in the second geometry file. Problems arise only for dihedrals which cross 0 degrees. Two cases must be considered:

1. *Dihedral decreases and crosses 0 degrees.* Define the dihedral for the second geometry to be negative, i.e. subtract the angle change from the starting dihedral and leave as a negative number.
2. *Dihedral increases and crosses 360 degrees.* Add the change in dihedral to the first dihedral to obtain the dihedral for the second geometry. The resulting number should be larger than 360.

As the specification of the geometries is quite difficult, SADDLE can be made to run in Cartesian coordinates using the keyword XYZ.

## 4.7 DATA FILE FOR A CHAIN CALCULATION

Behavior of the CHAIN method is illustrated by the inversion of ammonia through a planar transition state, without symmetry constraint. As is the case for the SADDLE method, care must be taken in choosing the

coordinates and linkage, in order to avoid meaningless intermediate geometries or a poor definition of some dihedral along the path.

The coordinates must be pictured as continuously varying from the left minimum to the right one via the transition state. It is important to avoid having more than one atom in the same location along the path.

```

Line 1:      CHAIN PRINT=1
Line 2:      SAMPLE NH3 INVERSION WITHOUT SYMMETRY CONSTRAINT
Line 3:
Line 4:      N
Line 5:      H      1.1
Line 6:      XX      1.0000  0      90.0  0
Line 7:      H      1.1000  1      110.0  1      0.0  0      1  2  3
Line 8:      H      1.1000  1      110.0  1      180.0  1      1  2  3
Line 9:      O      0.0000  0      0.0  0      0.0  0      0  0  0
Line 10:     N
Line 11:     H      1.0075  1
Line 12:     XX      1.0000  0      90.0  0
Line 13:     H      1.0075  1      105.2  1      0.0  0      1  2  3
Line 14:     H      1.0075  1      105.2  1      110.8  1      1  2  3
Line 15:     O
Line 16:     N
Line 17:     H      1.0075  1
Line 18:     XX      1.0000  0      90.0  0
Line 19:     H      1.0075  1      105.2  1      0.0  0      1  2  3
Line 20:     H      1.0075  1      105.2  1      249.2  1      1  2  3
Line 21:     O

```

Points to note:

1. Lines 4-9: Geometry of the hypothetical transition state is given.
2. Lines 10-15: Geometry of the reactant (the 'left' minimum along the reaction path) follows the geometry of the transition state. Note that the atom numbering and connectivity must match the first geometry.
3. Lines 16-21: Geometry of the product (the 'right' minimum along the path) is given last. Again, all three geometries must have the same numbering and connectivity.
4. The dihedral angles must be related by a continuous deformation (left--TS--right). If the dihedral angle on line 20 was specified as -110.8 then the reactant and product would not be related by a continuous deformation, and the calculation would fail.

## 4.8 OUTPUT FILE FROM A CHAIN CALCULATION

```

*****
                                MNDO CALCULATION RESULTS
*****
*                                VERSION  2.10
*  CHAIN      - TRANSITION STATE TO BE OPTIMISED
*  PRINT      - PRINTOUT LEVEL IN OPTIMISATION =    1
*****
CHAIN PRINT=1
SAMPLE NH3 INVERSION WITHOUT SYMMETRY CONSTRAINT

```

## Chapter 4: Examples

ATOM NUMBER (I)	CHEMICAL SYMBOL	BOND LENGTH (ANGSTROMS) NA:I	BOND ANGLE (DEGREES) NB:NA:I	TWIST ANGLE (DEGREES) NC:NB:NA:I	NA	NB	NC
1	N						
2	H	1.10000			1		
3	XX	1.00000	90.00000		2	1	
4	H	1.10000 *	110.00000 *	0.00000	1	2	3
5	H	1.10000 *	110.00000 *	180.00000 *	1	2	3

CARTESIAN COORDINATES

NO.	ATOM	X	Y	Z
1	7	0.0000	0.0000	0.0000
2	1	1.1000	0.0000	0.0000
3	1	-0.3762	1.0337	0.0000
4	1	-0.3762	-1.0337	0.0000

RHF CALCULATION, NO. OF DOUBLY OCCUPIED LEVELS = 4

INTERATOMIC DISTANCES

	N 1	H 2	H 3	H 4
N 1	0.000000			
H 2	1.100000	0.000000		
H 3	1.100000	1.802134	0.000000	
H 4	1.100000	1.802134	2.067324	0.000000

TOTAL TIME ALLOWED FOR THIS RUN : 3600.00 SECONDS

1SADDLE POINT RESEARCH... CHAIN METHOD

DISCRIMINATION ANGLE: 30.0 DEGREES

DIMENSION: 5 MAX. ITERATION: 45

RMS GRAD CV THRESHOLD: 5.0E+00

MAXIMUM LINK LENGTH: 0.300

PRINTOUT LEVEL = 1

COORDINATES OF MINIMA AND IMPOSED POINT:

	LEFT MINI	INTERMEDIATE	RIGHT MINI
N	1.007500	1.100000	1.007500
1	1.836086	1.919862	1.836086
2	1.007500	1.100000	1.007500
3	1.836086	1.919862	1.836086
4	1.933825	3.141593	4.349360

ENERGY -2.1576E+00 3.1660E+01 -2.1576E+00

INSERTION ON THE LEFT ENERGY= 1.20154E+00

INSERTION ON THE LEFT ENERGY= 8.95881E+00

INSERTION ON THE LEFT ENERGY= 1.83997E+01

INSERTION ON THE LEFT ENERGY= 2.66921E+01

INSERTION ON THE RIGHT ENERGY= 1.20148E+00

INSERTION ON THE RIGHT ENERGY= 8.95853E+00

INSERTION ON THE RIGHT ENERGY= 1.83999E+01

INSERTION ON THE RIGHT ENERGY= 2.66921E+01

ITERATION 1 HIGH POINT...ENERGY: 3.16600140E+01 RMS GRAD.:7.02E+01

<G!G>=2.462952E+04 INDEX: -5 CHAIN LENGTH... LEFT: 5 RIGHT: 5

DESCENT BY PROJECTED GRADIENT SLOPE: -1.569E+02

ITERATION 2 HIGH POINT...ENERGY: 2.66971643E+01 RMS GRAD.:6.18E+01

<G!G>=1.912087E+04 INDEX: -4 CHAIN LENGTH... LEFT: 6 RIGHT: 4

DESCENT BY PROJECTED GRADIENT SLOPE: -1.371E+02

ITERATION 3 HIGH POINT...ENERGY: 2.66971643E+01 RMS GRAD.:6.18E+01

<G!G>=1.912430E+04 INDEX: -4 CHAIN LENGTH... LEFT: 4 RIGHT: 6

DESCENT BY PROJECTED GRADIENT SLOPE: -1.371E+02

ITERATION 4 HIGH POINT...ENERGY: 1.84130664E+01 RMS GRAD.:5.06E+01

<G!G>=1.278515E+04 INDEX: -4 CHAIN LENGTH... LEFT: 7 RIGHT: 3

DESCENT BY PROJECTED GRADIENT SLOPE: -1.131E+02

ITERATION 5 HIGH POINT...ENERGY: 1.84130667E+01 RMS GRAD.:5.06E+01

<G!G>=1.279032E+04 INDEX: -3 CHAIN LENGTH... LEFT: 3 RIGHT: 7

MAXIMUM DISTANCE REDUCED TO : 0.2571

DESCENT BY PROJECTED GRADIENT SLOPE: -1.131E+02

ITERATION 6 HIGH POINT...ENERGY: 1.53017894E+01 RMS GRAD.:1.56E+01

<G!G>=1.218598E+03 INDEX: -1 CHAIN LENGTH... LEFT: 5 RIGHT: 5

DESCENT BY PROJECTED GRADIENT SLOPE: -3.491E+01

Note 1

Note 2

Note 3

Note 4

INSERTION ON THE LEFT ENERGY= 1.28567E+01 **Note 5**  
 INSERTION ON THE RIGHT ENERGY= 1.28557E+01  
 ITERATION 7 HIGH POINT...ENERGY: 1.41113243E+01 RMS GRAD.:1.66E+01  
 <G!G>=1.380304E+03 INDEX: 2 CHAIN LENGTH... LEFT: 4 RIGHT: 8  
 DESCENT BY PROJECTED GRADIENT SLOPE: -3.710E+01  
 INSERTION ON THE LEFT ENERGY= 9.67729E+00  
 ITERATION 8 HIGH POINT...ENERGY: 1.41110895E+01 RMS GRAD.:1.66E+01  
 <G!G>=1.380296E+03 INDEX: 2 CHAIN LENGTH... LEFT: 9 RIGHT: 4 **Note 6**  
 DESCENT BY PROJECTED GRADIENT SLOPE: -3.712E+01  
 INSERTION ON THE RIGHT ENERGY= 9.79050E+00  
 ITERATION 9 HIGH POINT...ENERGY: 1.28568333E+01 RMS GRAD.:1.00E+01  
 <G!G>=5.020268E+02 INDEX: 1 CHAIN LENGTH... LEFT: 6 RIGHT: 8  
 DESCENT BY PROJECTED GRADIENT SLOPE: -2.207E+01  
 ITERATION 10 HIGH POINT...ENERGY: 1.28567650E+01 RMS GRAD.:1.00E+01  
 <G!G>=5.017896E+02 INDEX: 1 CHAIN LENGTH... LEFT: 8 RIGHT: 6  
 MAXIMUM DISTANCE REDUCED TO : 0.2250  
 QUADRATIC DESCENT SLOPE: -9.262E+00  
 ELIMINATION ON THE LEFT FROM 8 TO 8  
 INSERTION ON THE RIGHT ENERGY= 1.16270E+01

CONVERGENCE ACHIEVED AFTER 49 CALLS OF ENERGY 17 OF WHICH WITH GRADIENT

ITERATION 11 HIGH POINT...ENERGY: 1.21735310E+01 RMS GRAD.:1.93E+00  
 <G!G>=1.872096E+01 INDEX: 1 CHAIN LENGTH... LEFT: 7 RIGHT: 7  
 COORDINATES : 0.97986 2.10547 0.97716 2.05532 3.17296  
 GRADIENT : 0.92670 1.88132 -3.22735 -1.72321 -0.96828

ESTIMATED EIGENVALUES AND EIGENVECTORS OF THE HESSIAN

1 -4.41811E+01 ... -0.0018 0.0604 0.0037 0.0155 -0.9980  
 2 1.67039E+02 ... -0.0120 -0.9672 -0.0883 -0.2295 -0.0624

LEFT CHAIN LENGTH: 7 ENERGY AND COORDINATES(A,RAD) ...

-2.1576290E+00 ... 1.0075E+00 1.8361E+00 1.0075E+00 1.8361E+00 1.9338E+00 **Note 8**  
 1.2015404E+00 ... 1.0260E+00 1.8528E+00 1.0260E+00 1.8528E+00 2.1754E+00  
 8.9588117E+00 ... 1.0445E+00 1.8696E+00 1.0445E+00 1.8696E+00 2.4169E+00  
 1.0333323E+01 ... 9.6953E-01 1.9142E+00 9.6950E-01 1.9142E+00 2.6215E+00  
 9.6772910E+00 ... 9.8177E-01 1.9924E+00 9.8172E-01 1.9923E+00 2.7017E+00  
 1.0474534E+01 ... 9.9401E-01 2.0705E+00 9.9394E-01 2.0703E+00 2.7818E+00  
 1.2160291E+01 ... 9.6656E-01 2.0480E+00 9.6653E-01 2.0478E+00 2.9930E+00

RIGHT CHAIN LENGTH: 7 ENERGY AND COORDINATES(A,RAD) ...

1.1627017E+01 ... 9.8683E-01 2.0882E+00 9.8548E-01 2.0630E+00 3.3383E+00  
 1.0450457E+01 ... 9.9381E-01 2.0708E+00 9.9380E-01 2.0707E+00 3.5036E+00  
 9.7904989E+00 ... 9.8900E-01 1.9903E+00 9.8900E-01 1.9903E+00 3.5797E+00  
 1.0392567E+01 ... 9.8419E-01 1.9099E+00 9.8419E-01 1.9098E+00 3.6559E+00  
 8.9585299E+00 ... 1.0445E+00 1.8696E+00 1.0445E+00 1.8696E+00 3.8663E+00  
 1.2014831E+00 ... 1.0260E+00 1.8528E+00 1.0260E+00 1.8528E+00 4.1078E+00  
 -2.1576290E+00 ... 1.0075E+00 1.8361E+00 1.0075E+00 1.8361E+00 4.3494E+00

CHAIN PRINT=1

SAMPLE NH3 INVERSION WITHOUT SYMMTRY CONSTRAINT

GEOMETRY OPTIMISED : GRADIENT NORM MINIMISED  
 SCF FIELD WAS ACHIEVED

MNDO CALCULATION

VERSION 2.10

FINAL HEAT OF FORMATION = 12.173523 KCAL

ELECTRONIC ENERGY = -434.394382 EV  
 CORE-CORE REPULSION = 184.944252 EV

GRADIENT NORM = 4.326773  
 IONISATION POTENTIAL = 8.965246

NO. OF FILLED LEVELS = 4

SCF CALCULATIONS = 52  
 COMPUTATION TIME = 0.34 SECONDS

ATOM NUMBER (I)	CHEMICAL SYMBOL	BOND LENGTH (ANGSTROMS) NA:I	BOND ANGLE (DEGREES) NB:NA:I	TWIST ANGLE (DEGREES) NC:NB:NA:I	NA	NB	NC
1	N						
2	H	1.10000			1		
3	XX	1.00000	90.00000		2	1	
4	H	0.97986 *	120.63449 *	0.00000	1	2	3
5	H	0.97716 *	117.76089 *	-178.20301 *	1	2	3

## Chapter 4: Examples

```

      INTERATOMIC DISTANCES
      N 1      H 2      H 3      H 4
-----
N 1      0.000000
H 2      1.100000      0.000000
H 3      0.979863      1.807924      0.000000
H 4      0.977164      1.779373      1.708161      0.000000

      EIGENVALUES
-31.00325 -16.85674 -15.37828 -8.96525  3.93849  6.70444  8.13881

      NET ATOMIC CHARGES AND DIPOLE CONTRIBUTIONS
      ATOM NO.  TYPE      CHARGE      ATOM ELECTRON DENSITY
1              N      -0.5164      5.5164
2              H       0.1784      0.8216
3              H       0.1725      0.8275
4              H       0.1655      0.8345
DIPOLE          X          Y          Z      TOTAL
POINT-CHG.     0.167      0.011      0.022      0.169
HYBRID         -0.051     -0.010      0.046      0.070
SUM            0.116      0.002      0.068      0.134

      CARTESIAN COORDINATES
      NO.      ATOM      X          Y          Z
1              N          0.0000      0.0000      0.0000
2              H          1.1000      0.0000      0.0000
3              H         -0.4993      0.8431      0.0000
4              H         -0.4551     -0.8643      0.0271

      ATOMIC ORBITAL ELECTRON POPULATIONS
1.45446  1.03037  1.03229  1.99931  0.82158  0.82750  0.83449

```

### NOTES ON OUTPUT FILE

**NOTE 1:** Bond lengths are in Angstroms; angles are in radians.

**NOTE 2:** Intermediate points along the path are generated.

**NOTE 3:** A negative index gives the number of missing gradients to build a Hessian estimate.

**NOTE 4:** Done periodically to insure convergence.

**NOTE 5:** Add some points in order to maintain a short distance between 2 successive points.

**NOTE 6:** An estimate of the Hessian becomes available.

**NOTE 7:** A saddle point has been located

**NOTE 8:** Geometries of points along the chain, but this is not a 'true' reaction path.

**NOTE 9:** End of specific results provided by the CHAIN algorithm. The .ARC file will contain the geometry of the approximate transition state.



## CHAPTER 5: BACKGROUND

This chapter provides a brief summary of some of the more important algorithms found in AMPAC. A full reference listing is provided in Appendix B.

### 5.1 THE SELF CONSISTENT FIELD PROCEDURE

#### Criteria and Accuracy in SCF Convergence

Two criteria are used to test for SCF convergence, and both must be met before a self-consistent field is achieved.

**SELCON:** If the difference in electronic energy, in eV, between any two consecutive iterations drops below SELCON, *and* the difference between any three consecutive iterations drops below ten times the value of SELCON, then the first test is satisfied. SELCON has a default value of 0.0001 eV, but this can be modified in one of two ways:

1. If any of the keywords PRECISE, FORCE, LTRD or NEWTON are specified, this criterion is made two orders of magnitude smaller, thus increasing the precision of the calculation.
2. The user can choose their own convergence criterion by using the keyword SCFCRT=. A lower bound of 1.D-11 is placed on this value.

**PLTEST:** This is the convergence threshold for the diagonal elements of the density matrix. This criterion is not a simple value, but is coupled with the energy threshold via an empirical equation. By thus relating these two values one obtains a smooth and more confident behavior of the SCF response versus the required SCFCRT criterion. This is required in the frequent cases where the energy appears stationary but the density matrix is still oscillating (e.g. molecules with heteroatoms involved in a donor-acceptor bond). With PLTEST, one can obtain a comparable level of self-consistence for both the energy and density matrix, and thus more accurate results as the value of SCFCRT is lowered.

An estimate of the standard deviation of the SCF electronic energy is also provided. This value is based on an empirical formula involving both SELCON and PLTEST.

#### Convergence Techniques in the SCF Calculation

A number of methods are available to subroutine ITER to promote convergence. If a calculation does not converge in 100 cycles, these routines are called automatically.

1. *Oscillation damping:* If on any two consecutive iterations a density matrix element changes by more than 0.05, the density matrix element is set equal to the old element shifted by 0.05 in the direction of the calculated element. Thus, if on iterations 3 and 4 a certain density matrix element was 0.55 and 0.78, respectively, then the element would be set to 0.60 ( $=0.55+0.05$ ) on iteration 4. The density matrix from iteration 4 would then be used in the construction of the next Fock matrix. The arrays which hold old density matrices are not filled until after iteration 2 and can not be used in the damping before iteration 3.
2. *Three-point interpolation of the density matrix:* This is the default converger, and is very effective with normally convergent calculations. If the number of iterations is divisible by three and certain

other conditions relating to the density matrices are satisfied, subroutine CNVG performs a three-point interpolation. It fails in certain systems, usually where significant charge buildup is present.

3. *Energy level shift technique:* The virtual M.O. energy levels are shifted to more positive energy. This has the effect of damping oscillations in occupancy between nearly degenerate virtual and filled M.O.s. Intrinsically divergent equations can often be changed to intrinsically convergent forms. This method can also be invoked manually by using the keyword SHIFT=n.
4. *Pulay's method:* When the largest change in density matrix elements on two consecutive iterations has dropped below 0.1, then routine CNVG is abandoned in favor of a multi-Fock matrix interpolation. PULAY relies on the fact that the eigenvectors of the density and Fock matrices are identical at self-consistency, so  $[\mathbf{P}, \mathbf{F}] = 0$  at SCF. The extent to which this condition is not met is a measure of the deviance from self-consistency. Pulay's method uses this relationship to calculate a linear combination of Fock matrices which minimize  $[\mathbf{P}, \mathbf{F}]$ . This new Fock matrix is then used in the SCF calculation.

Under certain circumstances, Pulay's method can cause very slow convergence, but sometimes it is the only way to achieve a self-consistent field. At other times the procedure gives a ten-fold increase in speed, so care must be exercised in its use. This method can also be invoked manually by using the keyword PULAY.

5. *The Camp-King converger:* If all else fails, the Camp-King converger will work almost every time. It evaluates a linear combination of old and current eigenvectors that minimize the total energy. One of its strengths is that systems which otherwise oscillate due to charge surges, e.g. CHO-H with a very large C-H distance, will converge using this very sophisticated converger. Because it is very expensive in time, this method is invoked only as a last resort.

### Failure to Achieve Self-Consistence

In a system where a biradical can form, such as ethane decomposing into two  $\text{CH}_3$  units, the normal RHF procedure can fail to go self-consistent. If the system has significant biradicaloid character, BIRADICAL or UHF/TRIPLET can often prove successful.

This problem can be demonstrated with H-Cl by steadily increasing the interatomic distance. At bonding distance the covalent bond will be strong, and a self-consistent field is readily obtained. As the bond is stretched it will become more ionic. Electrons will be polarized toward the chlorine and the hydrogen will become electropositive. As the energy increases electrons are attracted by the positive charge on the hydrogen. At some point, an electron will suddenly "jump" from the chlorine to the hydrogen causing a reorganization of the density matrix. There is a good chance the electron will oscillate between the hydrogen and the chlorine, and SCF will not be achieved. This can be prevented by using BIRADICAL, and exactly one electron will end up on hydrogen. A similar result can be obtained by specifying TRIPLET in a UHF calculation.

## 5.2 GEOMETRY OPTIMIZATION

The default procedure for geometry optimization is based on the BFGS algorithm (for reference see Appendix B). This is implemented in subroutine FLEPO, which minimizes a real valued function (in this case the heat of formation) evaluated by routine COMPFG. The following convergence tests, listed according to the messages generated, are used within FLEPO.

1. *HERBERTS TEST SATISFIED - GEOMETRY OPTIMIZED*

The estimated distance from the current geometry to the minimum is less than EYEAD. The default value of EYEAD is 0.001, with PRECISE this drops to 0.0001.

2. *TEST ON GRADIENT SATISFIED*

The gradient norm is less than TOLERG times the square root of the number of geometric variables. The default value is 1.0 kcal mol<sup>-1</sup> Å<sup>-1</sup> (or kcal mol<sup>-1</sup> rad<sup>-1</sup>); with PRECISE this is refined by a factor of 0.2. If the keyword GNORM is specified, then TOLERG is replaced by the value supplied.

3. *TEST ON X SATISFIED*

The relative change in the geometry, as measured by its norm, over any two successive iterations has dropped below TOLERX. The default value is 0.0001 Å. When precise is specified, this value is lowered to 0.00001 Å.

4. *HEAT OF FORMATION TEST SATISFIED*

The calculated values of the heat of formation on two successive cycles differ by less than TOLERF. Default value is 0.002 kcal mol<sup>-1</sup>, becoming 0.0002 kcal mol<sup>-1</sup> when precise is specified.

5. *PETERS TEST SATISFIED*

In addition to the TOLERG, TOLERF and TOLERX tests, a second test in which no individual component of the gradient norm is larger than TOLERG must be satisfied. If on three successive cycles PETERS test fails, while all other criteria are met and the heat of formation has dropped by less than 0.3 kcal/mol, then the calculation is stopped after issuing the message:

*THERE HAVE BEEN 3 ATTEMPTS TO REDUCE THE GRADIENT  
DURING THESE ATTEMPTS THE ENERGY DROPPED BY LESS THAN 0.3 KCAL/MOL  
FURTHER CALCULATION IS NOT JUSTIFIED AT THIS TIME*

### 5.3 TORSION ANGLE COHERENCY

AMPAC calculations do not distinguish between diastereoisomers; consequently, the sign of the dihedrals can be multiplied by -1 and the calculations will be unaffected. However, if chirality is important, a user should be aware of the sign convention used.

The dihedral angle convention used in AMPAC is that defined by Klyne and Prelog in *Experientia* **16**, 521 (1960). In this convention, four atoms, AXYB, with a dihedral angle of 90 degrees, will have atom B rotated by 90 degrees clockwise relative to A when X and Y are lined up in the direction of sight, X being nearer to the eye. In their words, "To distinguish between enantiomeric types, the angle  $\tau$  is considered as positive when it is measured clockwise from the front substituent A to the rear substituent B, and negative when it is measured anticlockwise." The alternative convention was used in all earlier programs, including QCPE 355.

## 5.4 PSEUDODIAGONALIZATION-- SUBROUTINE DIAG

The basis of subroutine DIAG is the observation that accurate matrix diagonalization of the secular determinant is not a prerequisite in the SCF procedure for obtaining a self-consistent density matrix in a variationally optimized calculation. To have a self-consistent density matrix, it is sufficient to have annihilated all energy matrix elements connecting the occupied and virtual molecular orbitals.

### Theory

Given a basis set of  $N$  atomic orbitals and  $N_e$  electrons, there will be  $N_o = N_e / 2$  occupied molecular orbitals and  $N_{vir} = N - N_o$  virtual orbitals. If the approximate form of the molecular orbitals is known, perhaps from an accurate, standard diagonalization of the first trial secular determinant, then the interaction matrix can be constructed. The off-diagonal matrix elements in  $F_{ov}$  can then be annihilated by a series of  $2 \times 2$  rotations in the manner of Jacobi. Unlike Jacobi, however, the method need not be cycled to exactly diagonalize  $F_{ov}$ ; it is sufficient to have only one sweep. This is due to the fact that the Fock equations form a pseudo-eigenvalue problem, and it is necessary to iterate to obtain a self-consistent field. For the same reason the second-order effects of the  $2 \times 2$  rotations can be ignored. In an exact diagonalization the off-diagonal matrix elements formed by an elementary  $2 \times 2$  rotation would have to be eliminated. These are normally less than one tenth of the matrix element being annihilated, and as the SCF procedure does not converge at one magnitude per cycle the second-order errors introduced can be absorbed into the Fock matrix of the following cycle. Also, since second-order effects in the "diagonalization" are being ignored it is equally valid to eliminate only those matrix elements which are comparable with the largest off-diagonal elements in  $F_{ov}$ . A further advantage of the pseudo-eigenvalue nature of the SCF equations appears when we come to evaluate the diagonal terms of the secular determinant. For this, we can equate these elements with the eigenvalues resulting from the exact diagonalization, and hold them exactly constant throughout the entire calculation, right up to self-consistency. At first sight this would appear to introduce errors in the final SCF density matrix, as obviously the sum of the eigenvalues cannot be constant in an exact calculation, and thus the final sum of occupied energy levels must be in error. However, to obtain a SCF density matrix we not only do not need to know the exact eigenvalues, we have no need to know the sum of the occupied energy levels. Using the initial set of eigenvalues, a  $2 \times 2$  rotation will, of course, not eliminate fully even those elements which we do choose to operate on, but again the pseudo-eigenvalue nature of the problem comes to our rescue. As the iterations proceed those errors introduced are rapidly eliminated, so that at self-consistency an exact density matrix is generated, but we have no knowledge of the values of the eigenvalues, eigenvectors or two-electron energy. This completes the definition of the secular determinant, and the  $2 \times 2$  rotations needed to pseudo-diagonalize it. The unitary matrix that results from the set of rotations would normally need to be multiplied by the original set of eigenvectors to obtain the correct molecular orbital matrix. There is no reason to start the pseudo-diagonalization with a unit matrix, and this step can be eliminated by starting the pseudo-diagonalization using the old set of molecular orbitals.

### Computational Aspects

Clearly, an initial exact diagonalization is needed in order to obtain a good starting set of eigenvectors. This is, however, not a sufficient condition for initiating the use of the new method. During the first one or two cycles of an SCF procedure the order of occupancy of the molecular orbitals may change. That is, the occupancy of the M.O.'s in the first two iterations may correspond to an excited singlet state. For this reason it is recommended that the new method be used only after the initial large fluctuations in the density matrix have died down. Our arbitrary criterion was that the largest change in the diagonal elements of the density matrix between any two iterations are less than 0.05.

The method proposed uses three matrices: the C matrix of eigenvector coefficients, the F matrix, and a working matrix to hold the new  $F_{OV}$  matrix. In virtual memory computers this would normally not present any problem. A criterion is also needed to decide when to perform a 2x2 rotation and elimination. After a few trial calculations, it was found to be efficient to eliminate all off-diagonal elements whose moduli were larger than 0.01 times the modulus of the largest off-diagonal element. No important change in the number of iterations was observed. The electronic energy calculated after each iteration does indeed differ following the introduction of the new method. This is a result of all the second-order effects being introduced, but the differences between electronic energies calculated by the exact diagonalization and by this method rapidly converge. An interesting side effect of the non-rigorous method is that some damping is introduced; in all molecules examined the number of iterations required to achieve self-consistency either stayed constant or dropped by one or two.

## 5.5 CONFIGURATION INTERACTION

AMPAC contains a Multi-Electron Configuration Interaction calculation, subroutine MECI, which allows almost any configuration interaction calculation to be performed. Because of its complexity, two distinct levels of input are supported; the default values will be of use to the novice while the expert has available an exhaustive set of keywords from which a specific C.I. can be tailored.

A C.I. calculation involves the interaction of microstates representing specific permutations of electrons in a set of M.O.s. Starting with a reference set of M.O.s from an RHF or RHF/half-electron calculation, the first step in a C.I. calculation is the removal of the electrons from the M.O.s to be used in the C.I.

Each microstate is then constructed from these empty M.O.s by adding in electrons according to a prescription. The energy of the configuration is evaluated, as is the energy of interaction with all previously defined configurations. Diagonalization then results in state functions. From the eigenvectors the expectation value of  $S^2$  is calculated.

Keywords associated with the operations of subroutine MECI are:

SINGLET	DOUBLET	EXCITED
TRIPLET	QUARTET	BIRADICAL
QUINTET	SEXTET	ESR
OPEN(n1,n2)	C.I.=n	MECI
SIZE=n	ROOT=n	

Some of the keywords are implicit, thus BIRADICAL is equivalent to OPEN(2,2), SINGLET. Each keyword is discussed in the *Keywords* Chapter.

### Reference Electronic Configuration

C.I. is restricted to RHF/HE wavefunctions; however, any reference configuration is supported. Examples of reference configurations are:

System	Keywords used	Starting Configuration
Methane	<none>	2.00 2.00 2.00 2.00 2.00
Methyl Radical	<none>	2.00 2.00 2.00 2.00 1.00
Twisted Ethylene	TRIPLET	2.00 2.00 2.00 1.00 1.00
Twisted Ethylene	OPEN(2,2)	2.00 2.00 2.00 1.00 1.00
Twisted Ethylene Cation	OPEN(1,2)	2.00 2.00 2.00 0.50 0.50
Methane Cation	CHARGE=1 OPEN(5,3)	2.00 2.00 1.67 1.67 1.67

Choice of reference configuration is important. For example, if twisted ethylene (a ground state triplet) is not defined using TRIPLET or OPEN(2,2), then the closed shell ground state structure will be calculated. Obviously this configuration is a legitimate microstate, but from the symmetry of the system a better choice would be to define one electron in each of the two formally degenerate pi-type M.O.s. The initial SCF calculation does not distinguish between OPEN(2,2) and TRIPLET since both keywords define the same starting configuration. This can be verified by monitoring the convergence using PL, for which both keywords give the same SCF energy.

### Removal of Electrons from Reference Configuration

For a reference configuration of alpha M.O. occupancies  $n_j$  ( $0.0 \leq n_j \leq 1.0$ ) the energies of the M.O.s involved in the C.I. can be calculated from:

$$\epsilon_i = h_{ii} + \sum_j n_j (2 J_{ij} - K_{ij})$$

where  $h_{ii}$  is the one-electron energy of M.O.  $i$ , and  $J_{ij}$  and  $K_{ij}$  are the coulomb and exchange integrals between M.O.s  $i$  and  $j$ . The M.O. index  $j$  runs over those M.O.s involved in the C.I. only. Most C.I. calculations will involve between 1 and 5 M.O.s, so a system with 30 filled or partly filled M.O.s could have M.O.s 25-30 involved. The resulting eigenvalues correspond to those of the cationic system resulting from removal of  $N$  electrons, where  $N$  is twice the sum of the orbital occupancies of those M.O.s involved in the C.I.

The arbitrary zero of energy in a C.I. calculation is the reference ground state, without any correction for errors introduced by the use of fractional occupancies. In order to calculate the energy of the various configurations, the energy of the vacuum state (i.e., the state resulting from removal of the electrons used in the C.I.) needs to be evaluated. This energy is defined by:

$$E_o = \sum_i \left[ \epsilon_i n_i - J_{ii} n_i^2 - 2 \sum_{j < i} (2 J_{ij} - K_{ij}) n_i n_j \right]$$

### Formation of Microstate Configurations

Microstates are particular electron configurations. Thus if there are 5 electrons in 5 levels, then examples of various microstates could be as follows:

Electron Configuration						Electron Configuration					
Alpha			Beta			Alpha			Beta		
1	2	3	4	5	M(s)	1	2	3	4	5	M(s)
1	1,1,1,0,0	1,1,0,0,0	1/2			4	1,1,1,1,1	0,0,0,0,0	5/2		
2	1,1,0,0,0	1,1,1,0,0	-1/2			5	1,1,0,1,0	1,1,0,0,0	1/2		
3	1,1,1,0,0	0,0,0,1,1	1/2			6	1,1,0,1,0	1,0,1,0,0	1/2		

For 5 electrons in 5 M.O.s there are 252 microstates ( $10!/(5!5!)$ ), but as states of different spin do not mix, we can use a smaller number. If doublet states are needed then 100 states ( $5!/(2!3!)(5!/(3!2!))$ ) are needed. If only quartet states are of interest then 25 states ( $5!/(1!4!)(5!/(4!1!))$ ) are needed and if the sextet state is required, then only one state is calculated.

In the microstates listed, state 1 is the ground state configuration. This can be written as (2,2,1,0,0), meaning that M.O.s 1 and 2 are doubly occupied, M.O. 3 is singly occupied by an alpha electron, and M.O.s 4 and 5 are empty. Microstate 1 has a component of spin of 1/2, and is a pure doublet. By Kramer's degeneracy (sometimes called time-inversion symmetry) microstate 2 is also a doublet, and has a spin of 1/2 and a z-component of spin of -1/2.

Microstate 3, while it has a component of spin of 1/2, is not a doublet, but is in fact a component of a doublet, a quartet and a sextet. The coefficients of these states can be calculated from the Clebsch-Gordon 3-J symbol. For example, the coefficient in the sextet is 1/Sqrt(5).

Microstate 4 is a pure sextet. If all 100 microstates of component of spin = 1/2 were used in a C.I., one of the resulting states would have the same energy as the state resulting from microstate 4.

Microstate 5 is an excited doublet, and microstate 6 is an excited state of the system, but not a pure spin state.

By default, if n M.O.s are included in the C.I., then all possible microstates which give rise to a z-component of spin = 0 for even electron systems, or 1/2 for odd electron systems, will be used.

#### Permutations of Electrons among Molecular Orbitals

(0,1) = 0	(2,4) = 1100	(3,5) = 11100	(2,5) = 11000
	1010	11010	10100
(1,1) = 1	1001	11001	10010
	0110	10110	10001
(0,2) = 0	0101	10101	01100
	0011	10011	01010
(1,2) = 10		01110	01001
01	(1,4) = 1000	01101	00110
	0100	01011	00101
(1,3) = 100	0010	00111	00011
010	0001		
001			
(2,3) = 110			
101			
011			

#### Sets of Microstates for Various C.I. Calculations

Odd Electron Systems				Even Electron Systems			
	Alpha	Beta	No. of Configs.		Alpha	Beta	No. of Configs.
C.I. = 1	(1,1)	* (0,1)	= 1	(1,1)	* (1,1)	= 1	
2	(1,2)	* (0,2)	= 2	(1,2)	* (1,2)	= 4	
3	(2,3)	* (1,3)	= 9	(2,3)	* (2,3)	= 9	
4	(2,4)	* (1,4)	= 24	(2,4)	* (2,4)	= 36	
5	(3,5)	* (2,5)	= 100	(3,5)	* (3,5)	= 100	

#### Multi Electron Configuration Interaction

The numbering of the M.O.s used in the C.I. is standard, and follows the Aufbau principle. The order of filling is in order of energy, and alpha before beta. This point is critically important in deciding the sign of matrix elements. For a 5 M.O. system, then, the order of filling is:

$$(1)(\bar{1})(2)(\bar{2})(3)(\bar{3})(4)(\bar{4})(5)(\bar{5})$$

A triplet state arising from two microstates, each with a z-component of spin = 0, will thus be the positive combination:

$$|\bar{\Phi}_1\Phi_2\rangle + |\Phi_1\bar{\Phi}_2\rangle$$

This is in variance with the sign convention used in earlier programs for running MNDO. This standard sign convention was chosen in order to allow the signs of the microstate coefficients to conform to those resulting from the spin step-down operator.

Matrix elements between all pairs of microstates are calculated in order to form the secular determinant. Many elements will be identically zero, due to the interacting determinants differing by more than two M.O.s. For the remaining interactions the following types can be identified:

1. The two determinants are identical.

No permutations are necessary in order to calculate the sign of the matrix element.  $H_{pp}$  is given simply by:

$$H_{pp} = \sum_i \left( \begin{array}{l} n_{ip}^\alpha \left[ \epsilon_i^\alpha - \sum_j \left[ \frac{1}{2} n_{jp}^\alpha (J_{ij} - K_{ij}) + n_{jp}^\beta J_{ij} \right] \right] \\ + n_{ip}^\beta \left[ \epsilon_i^\beta - \sum_j \left[ \frac{1}{2} n_{jp}^\beta (J_{ij} - K_{ij}) + n_{jp}^\alpha J_{ij} \right] \right] \end{array} \right)$$

$n_{ip}^\alpha$  = Occupancy of alpha M.O.  $i$  in Microstate  $p$

$n_{ip}^\beta$  = Occupancy of beta M.O.  $i$  in Microstate  $p$

2. Determinants differing by exactly one M.O.

The differing M.O. can be of type alpha or beta. It is sufficient to evaluate the case in which both M.O.s are of alpha type, the beta form is obtained in like manner:

$$H_{pq} = h_{ij} + \sum_k \left[ n_k^\alpha \left( \langle ik | jk \rangle - \langle ij | kk \rangle \right) + n_k^\beta \langle ik | jk \rangle \right]$$

$H_{pq}$  may need to be multiplied by -1, if the number of two electron permutations required to bring M.O.s  $i$  and  $j$  into coincidence is odd.

3. Determinants differing by exactly two M.O.s.

The two M.O.s can have the same or opposite spins. Three cases can be identified:

- a. Both M.O.s have alpha spin.

For the first microstate having M.O.s  $i$  and  $j$ , and the second microstate having M.O.s  $k$  and  $l$ , the matrix element connecting the two microstates is given by:

$$H_{pq} = \langle ik | jl \rangle - \langle il | jk \rangle$$

$H_{pq}$  may need to be multiplied by -1, if the number of two electron permutations required to bring M.O. **i** into coincidence with M.O. **k** and M.O. **j** into coincidence with M.O. **l** is odd.

- b. Both M.O.s have beta spin.

The matrix element is calculated in the same manner as in the previous case.

- c. One M.O. has alpha spin, and one beta spin.

For the first microstate having M.O.s  $i^\alpha$  and  $j^\beta$ , and the second microstate having M.O.s  $k^\alpha$  and  $l^\beta$ , the matrix element connecting the two microstates is given by:

$$H_{pq} = \langle ik | jl \rangle$$

$H_{pq}$  may need to be multiplied by -1, if the number of two electron permutations required to bring M.O. **i** into coincidence with M.O. **k** and M.O. **j** into coincidence with M.O. **l** is odd.

### States Arising from Various Calculations

Each C.I. calculation invoked by use of the keyword C.I.=n normally gives rise to states of quantized spins. When C.I. is used without any other modifying keywords, the following states will be obtained:

No. of M.O.s in C.I.	States Arising From Odd Electron Systems			States Arising From Even Electron Systems		
	Doublets	Quartets	Sextets	Singlets	Triplets	Quintets
1	1			1		
2	2			3	1	
3	8	1		6	3	
4	20	4		20	15	1
5	75	24	1	50	45	5

These numbers will be obtained irrespective of the chemical nature of the system.

### Calculation of Spin States

In order to calculate the spin state, the expectation value of  $S^2$  is calculated.

$$\begin{aligned}
 S^2 &= S(S+1) = S_z^2 + 2S^-S^+ \\
 &= N_{\text{tot}} - \sum_i C_{ik}^2 \left[ \frac{1}{4} (N_i^\alpha - N_i^\beta)^2 + \sum_l n_{li}^\alpha n_{li}^\beta + 2 \sum_j C_{ik} C_{jk} (\delta_{C_{ik} S^+ S^- C_{jk}}) \right]
 \end{aligned}$$

Where	$N_{\text{tot}}$	=	No. of electrons in C.I.
	$C_{ik}$	=	Coefficient of Microstate <b>i</b> in State <b>k</b>
	$N_i^\alpha$	=	Number of alpha electrons in Microstate <b>i</b>
	$N_i^\beta$	=	Number of beta electrons in Microstate <b>i</b>
	$n_{li}^\alpha$	=	Occupancy of alpha M.O. <b>l</b> in Microstate <b>i</b>
	$n_{li}^\beta$	=	Occupancy of beta M.O. <b>l</b> in Microstate <b>i</b>
	$S^+$	=	Spin shift up or step up operator
	$S^-$	=	Spin shift down or step down operator
	$\delta_{C_{ik} S^+ S^- C_{jk}}$	=	The Kroneker delta is 1 if the two terms are identical, zero otherwise.

The spin state is calculated from  $S = \frac{1}{2} \left( \sqrt{1 + 4 S^2} - 1 \right)$ .

In practice,  $S$  is calculated to be exactly integer or half integer, thus there is insignificant error due to approximations used. This does not imply that the method is accurate. The spin calculation is completely precise in the “group theory” sense, but the accuracy of the calculation is limited by the space-dependent Hamiltonian.

### Choice of State to be Optimized

C.I. can calculate a large number of states of various total spin. Two schemes are provided to allow a given state to be selected. First, ROOT= $n$  will select the  $n$ th state, irrespective of its total spin. By default  $n=1$ . If ROOT= $n$  is used in conjunction with a keyword from the set SINGLET, DOUBLET, TRIPLET, QUARTET, QUINTET, or SEXTET, then the  $n$ th root of that spin state will be used. For example, ROOT=4 and SINGLET will select the fourth singlet state. If there are two triplet states below the fourth singlet state then this will mean that the sixth state will be selected.

### Calculation of Unpaired Spin Density

Starting with the state functions as linear combinations of configurations, the unpaired spin density, corresponding to the alpha spin density minus the beta spin density, will be calculated for the first few states. This calculation is straightforward for diagonal terms and only those terms are used.

## 5.6 REDUCED MASSES IN A FORCE CALCULATION

Reduced masses for a diatomic are given by:

$$\frac{m_1 m_2}{m_1 + m_2}$$

For a Hydrogen molecule the reduced mass is 0.5; for heavily hydrogenated systems, e.g. methane, the reduced mass can be very low. A vibration involving only heavy atoms, e.g. a C-N in cyanide, should give a large reduced mass.

For the “trivial” vibrations the reduced mass is ill-defined, and where this happens the reduced mass is set to zero.

## 5.7 FRAME IN THE FORCE CALCULATION

The FORCE calculation uses Cartesian coordinates and all 3N modes are calculated, where N is the number of atoms in the system. Clearly, there will be 5 or 6 “trivial” vibrations, which represent the three translations and two or three rotations. If the molecule is exactly at a stationary point, then these “vibrations” will have a force constant and frequency of precisely zero. If the force calculation was done correctly, and the molecule was not exactly at a stationary point, then the three translations should be exactly zero, but the rotations would be non-zero. The extent to which the rotations are non-zero is a measure of the error in the geometry.

If the distortions are non-zero, the trivial vibrations can interact with the low lying genuine vibrations or rotations, and with the transition vibration if present.

To prevent this the analytic form of the rotations and vibrations is calculated, and arbitrary eigenvalues assigned; these are 500, 600, 700, 800, 900, and 1000 millidynes/angstrom for Tx, Ty, Tz, Rx, Ry and Rz (if present), respectively. The rotations are about the principal axes of inertia for the system, taking into account isotopic masses. The “force matrix” for these trivial vibrations is determined, and added on to the calculated force matrix. After diagonalization the arbitrary eigenvalues are subtracted off the trivial vibrations, and the resulting numbers are the “true” values. Interference with genuine vibrations is thus avoided.

## 5.8 VIBRATIONAL ANALYSIS

Analyzing normal coordinates is very tedious. Users are normally familiar with the internal coordinates of the system they are studying, but not familiar with the Cartesian coordinates. To help characterize the normal coordinates in FORCE, a very simple analysis is done automatically, and users are strongly encouraged to use this analysis first, and then to look at the normal coordinate eigenvectors.

In the analysis, each pair of bonded atoms within the Van der Waals' distance is examined to see if there is a large relative motion between them. If there is such a motion, the indices of the atoms, the relative distance in Angstroms, and the percentage radial motion are printed. Radial plus tangential motion adds to 100%, but as there are two orthogonal tangential motions and only one radial, the radial component is printed.

Alternatively, LTRD lists the mass-weighted eigenvectors (analogous to Cartesian normal modes) in internal coordinates making analysis somewhat more straightforward.

## 5.9 DYNAMIC REACTION COORDINATE

The course of a molecular vibration can be followed by calculating the potential and kinetic energy at various times. Two extreme conditions can be identified: (a) *gas phase*, in which the total energy is a constant through time and no damping of the kinetic energy is allowed, and (b) *liquid phase*, in which kinetic energy is always set to zero and the motion of the atoms are infinitely damped.

All possible degrees of damping are allowed. In addition, the facility exists to dump kinetic energy into the system. As kinetic energy is a function of velocity, a vector quantity, the energy appears as energy of

motion in the direction in which the molecule would naturally move. If the system is a transition state, then the excess kinetic energy is added after the intrinsic kinetic energy has built up to at least 0.2 kcal/mol.

For ground-state systems, the excess energy sometimes may not be added; if the intrinsic kinetic energy never rises above 0.2 kcal/mol then the excess energy will not be added.

### Equations Used

Force acting on any atom:

$$g_i + g_i' t + g_i'' t^2 = \frac{dE}{dx_i} + \frac{d^2E}{dx_i^2} + \frac{d^3E}{dx_i^3}$$

Acceleration due to force acting on each atom:

$$a_i = \frac{g_i + g_i' t + g_i'' t^2}{m_i}$$

The change in velocity is equal to the integral over the time interval of the acceleration. New velocity:

$$v_o + \frac{dt g_i}{m_i} + \frac{dt^2 g_i'}{2 m_i} + \frac{dt^3 g_i''}{3 m_i}$$

$$v_i = v_i + v_i' t + v_i'' t^2 + v_i''' t^3$$

The change in position is equal to the integral over the time interval of the velocity. New position of atoms:

$$x_i = x_o + v_o t + \frac{1}{2} v' t^2 + \frac{1}{3} v'' t^3 + \frac{1}{4} v''' t^4$$

The velocity vector is accurate to the extent that it takes into account the previous velocity, the current acceleration, the predicted acceleration, and the change in predicted acceleration over the time interval. Very little error is introduced due to higher order contributions to the velocity; those that do occur are absorbed in a re-normalization of the magnitude of the velocity vector after each time interval.

The magnitude of dt, the time interval, is determined mainly by the factor needed to re-normalize the velocity vector. If it is significantly different from unity, dt will be reduced; if it is very close to unity, dt will be increased.

Despite this correction, errors still occur. A system started at the transition state is unlikely to return precisely to the transition state unless an excess kinetic energy is supplied, for example 0.2 kcal/mol.

The calculation is carried out in Cartesian coordinates, and converted into internal coordinates for display. All Cartesian coordinates must be allowed to vary, in order to conserve angular and translational momentum.

## 5.10 POLARIZABILITY CALCULATION

If the electrons in a molecule are easily moved as the result of a stimulus, then the molecule is easily polarizable. Thus, if an applied electric field can easily induce a dipole, then the polarizability is large. Any induced dipole will lower the energy of the system, but this stabilization might be masked by the presence of a permanent dipole. To avoid this, an alternating electric field is used. If the molecule has an intrinsic dipole, then it will be stabilized in one direction. When the field is reversed, the molecule will be destabilized. On averaging the two effects, the result will be a net stabilization due only to the induced dipole.

AMPAC calculates the polarizability of molecules, radicals, and ions by use of a shaped electric field.

The applied electric field is produced by adding individual fields, each generated by four point charges. Consider first a single electric field in the x-direction. There are two point charges of value  $+Q$  and  $-Q/2$  electrons, at locations  $l$  and  $al$ ,  $a$  being the cube root of  $1/2$ . At the origin, the electric field gradient (in V/m) is:

$$\begin{aligned} \text{Due to } +Q \text{ charge:} & \quad \frac{-QC}{l^2} \\ \text{Due to } \frac{-Q}{2} \text{ charge:} & \quad \frac{QC}{2a^2 l^2} \\ \text{Therefore, the total field:} & \quad \frac{-QC}{l^2} \left( 1 - \sqrt[3]{\frac{1}{2}} \right) \\ \text{Differentiating } \frac{-QC}{l^2} + \frac{QC}{2a^2 l^2} \text{ with respect to } x: & \\ \frac{QC}{2l^3} - \frac{QC}{4a^3 l^3} = 0 & \end{aligned}$$

In like manner, it can be shown that in the plane perpendicular to the line joining the point charges, the derivative of the voltage gradient is also exactly zero.

A molecule positioned near to the origin will thus experience a uniform electric field gradient due to these two point charges. The other two point charges are of value  $+Q/2$  and  $-Q$  at points  $-al$  and  $-l$ , respectively. They produce an electric field gradient at the origin identical to the first two point charges. The total field,  $E$ , midway between the two sets of charges is:

$$E = \frac{2CQ \left( 1 - \sqrt[3]{\frac{1}{2}} \right)}{l^2 4\pi \epsilon_0}$$

The  $4\pi\epsilon_0$  allows for vacuum permittivity.

The heat of formation of the molecule in this field is then calculated. This quantity can be expressed as a series sum:

$$H_f(E) = H_f^0 + V E (\text{Charge}) + \frac{dV}{dx} E (\text{Dipole}) + \frac{d^2V}{dx^2} E (\text{Polarizability})$$

The heat of formation in the field is the sum of the basic heat of formation, plus the electric potential times any charge, any dipole times the electric field gradient, and any polarizability times the square of the electric field gradient.

We are interested in the polarizability, **P**:

$$P = \frac{2}{23.061} \frac{d^2H}{dE^2}$$

The second derivative of **H** with respect to **E** is given by:

$$\frac{d^2H}{dE^2} = \frac{H_f(E) + H_f(-E) - 2 H_f^0}{2 E}$$

**H<sub>f</sub>(E)** being the heat of formation in the electric field.

The polarizability volume, **V**, is calculated from the polarizability by:

$$V = \frac{P}{E \ 4 \pi \ \epsilon_0} = \frac{2}{E \ (23.061) \ 4 \pi \ \epsilon_0} \frac{d^2H}{dE^2}$$

Substituting for **E** we have:

$$V = \frac{2 \ l^4 \ \pi \ \epsilon_0}{(23.061) \ Q^2 \ C^2 \left( 1 - \sqrt[3]{\frac{1}{2}} \right)^2} \frac{d^2H}{dE^2}$$

It is a simple matter to evaluate the value of this second rank tensor by calculating the heats of formation of the molecule subject to four different electric field gradients. For the tensor component  $V_{ij}$ ,  $i=x$  or  $y$  or  $z$ ,  $j=x$  or  $y$  or  $z$ , the directions of the four different fields are defined by:

Field 1	+i, +j	Field 3	+i, -j
Field 2	-i, -j	Field 4	-i, +j

Thus if  $i=x$  and  $j=x$  the four fields are:

Field 1	+x
Field 2	0
Field 3	-x
Field 4	0

Using these four heats of formation (kcal/mol) the polarizability can be calculated in units  $\text{\AA}^3$  via:

$$V = 2 \pi \ \epsilon_0 \ (23.061) \ Q^2 \ C^2 \ ( H_f(2) + H_f(4) - H_f(1) - H_f(3) ) \ l^4 \ (1 - a)^2$$

Monopolar and dipolar terms are eliminated in this treatment, but monatomic additive terms are included when MNDO or AM1 is used.

A polarization matrix of size 3x3 is constructed and diagonalized, and the resulting eigenvalues are the calculated independent polarization volumes in cubic Angstroms; the vectors are the independent polarization vectors.

## 5.11 SOLID STATE CAPABILITY

Currently AMPAC can only handle up to one dimensional extended systems, but work is under way to extend it to two and three dimensions.

If a polymer unit cell is large enough, then a single point in k-space, the Gamma point, is sufficient to specify the entire Brillouin zone. The secular determinant for this point can be constructed by adding together the Fock matrix for the central unit cell plus those for the adjacent unit cells. The Born-von Karman cyclic boundary conditions are satisfied, and diagonalization yields the correct density matrix for the Gamma point.

At this point in the calculation, conventionally, the density matrix for each unit cell is constructed. Instead, the Gamma point density and one electron density matrices are combined with a "Gamma point like" Coulomb and exchange integral strings to produce a new Fock matrix. The calculation can be visualized as being done entirely in reciprocal space, at the Gamma point.

Most solid state calculations take a very long time. These calculations, called "Cluster" calculations after the original publication, require between 1.3 and 2 times the equivalent molecular calculation.

A minor "fudge factor" is necessary to make this method work. The contribution to the Fock matrix element arising from the exchange integral between an atomic orbital and its equivalent in the adjacent unit cells is ignored. This is necessitated by the fact that the density matrix element involved is invariably large.

The unit cell must be large enough that an atomic orbital in the center of the unit cell has an insignificant overlap with the atomic orbitals at the ends of the unit cell. In practice, a translation vector of more than 7 or 8 Å is sufficient. For one rare group of compounds a larger translation vector is needed. Polymers with delocalized pi-systems and polymers with very small band gaps will require a larger translation vector, in order to accurately sample k-space. For these systems, a translation vector in the order of 15-20 Å is needed.

## 5.12 SPARKLES

Four extra "elements" are included in AMPAC. These represent pure ionic charges, roughly equivalent to the following chemical entities:

Chemical Symbol	Equivalent to:
+	Tetramethyl ammonium radical, Potassium atom or Cesium atom.
++	Barium atom.
-	Borohydride radical, Halogen, or Nitrate radical.
--	Sulfate, oxalate.

For the purposes of discussion these entities are called "sparkles"; the name arises from consideration of their behavior.

### Behavior of Sparkles in AMPAC

Sparkles have the following properties:

1. Their nuclear charge is integer, and is +1, +2, -1, or -2; there are an equivalent number of electrons to maintain electroneutrality, 1, 2, -1, and -2 respectively.
2. They all have an ionic radius of 0.7 Å. Any two sparkles of opposite sign will form an ion pair with an interatomic separation of 1.4 Å.
3. They have a zero heat of atomization, no orbitals, and no ionization potential.

They can be regarded as unpolarizable ions of diameter 1.4 Å. They do not contribute to the orbital count, and cannot accept or donate electrons.

Since they appear as uncharged species which immediately ionize, attention should be given to the charge on the whole system. For example the alkaline metal salt of formic acid would have the formula  $\text{HCOO}^+$ , where + is the unipositive sparkle. The charge on the system would then be zero.

A water molecule polarized by a positive sparkle would have the formula  $\text{H}_2\text{O}^+$ , and the charge on the system would be +1.

At first sight, a sparkle would appear to be too ionic to be a point charge and would combine with the first charge of opposite sign it encountered. This representation is faulty; a better description would be of an ion of diameter 1.4 Å and the charge delocalized over its surface. Computationally, a sparkle is an integer charge at the center of a repulsion sphere of form  $e^{-\alpha r}$ . The hardness of the sphere is such that other atoms or sparkles can approach within about 2 Å quite easily, but only with great difficulty come closer than 1.4 Å.

### Uses of Sparkles

1. Sparkles can be used as counterions, e.g. for acid anions or for cations. Thus if the ionic form of an acid is wanted, the moieties  $\text{H}^+\text{X}$ ,  $\text{H}^-$ , and  $^+\text{X}$  could be examined.
2. Two sparkles of equal and opposite sign can form a dipole for mimicking solvation effects. Thus water could be surrounded by six dipoles to simulate the solvent cage. A dipole of value D can be made by using the two sparkles + and -, or using ++ and --. If + and - are used, the inter-sparkle separation would be  $D/4.803$  Å. If ++ and -- are used, the separation would be  $D/9.606$  Å. If the inter-sparkle separation is less than 1.0 Å (a situation that cannot occur naturally) then the energy due to the dipole on its own is subtracted from the total energy.
3. They can operate as polarization functions. A controlled, shaped electric field can easily be made from two or more sparkles. The polarizability in Å<sup>3</sup> of a molecule in any particular orientation can then easily be calculated.

## 5.13 INTEGRATION OF A REACTION PATH IN PATH

### Differential Equations of a Path

A reaction path is a special case of a steepest descent path, i.e. a trajectory with the force vector as a tangent. With  $s$  (a parameter of integration) and  $E$  (the energy as a function of a set of coordinates,  $\mathbf{x}$ ), such a trajectory is defined by the first order system of differential equations:

$$\frac{dx_i}{ds} = -w_i \frac{dE}{dx_i} \quad (i = 1, n)$$

where  $w_i$  is defined as a positive weight, e.g. the inverse of the masses of the nuclei.

With  $s=0$  as a starting point, a unique trajectory is defined and always ends at a critical point (most of the time, a minimum) when  $s$  grows to infinity. The length  $l$  of the trajectory from  $s=0$  to  $s=t$  is given by the following integral along the path:

$$l(t) = \int_0^t \sqrt{\sum_i \left( \frac{dx_i}{ds} \right)^2} ds$$

Strictly speaking, a reaction path in chemistry is defined as such a trajectory only in the weighted Cartesian coordinates. However, for purposes of connectivity (a topological concept invariant under a change of the coordinates), it can be extended to any internal coordinates system.

In chemical terms the answer to the question, “What is the destiny of a molecule starting from a given transition state in the (given) direction of the transition vector?”, does not depend on the coordinates system, but on the details of the trajectory followed (which are coordinate dependent).

### The Choice of a Numerical Integrator

Various kinds of numerical integrators are available in the literature and have been extensively tested in chemical kinetics using semi-classical trajectories. None of the standard ones (e.g. Runge-Kutta or Predictor-Corrector, both in various orders) turns out to be efficient for integration of a reaction path as soon as the bottom of a valley is followed.

Elsewhere, special integrators have been derived for systems with a large time constant of the kind:

$$\frac{dx_i}{ds} = -r x(i,s) + f(i,s)$$

with  $r$  a “large” positive number and  $f(i,s)$  a slowly varying function of  $s$ . Such systems are encountered in physics (e.g. in electronics) when rapid exponential transient responses are superimposed on a slowly varying permanent solution.

Curiously, this is also the case (with  $f(i,s)=0$ ) for a reaction path expressed in the normal modes of a quadratic potential. Anharmonic effects can be taken into account in  $f(i,s)$ .

Owing to these properties, one can derive a mixed strategy using both a standard Predictor-Corrector when  $r$  is small or negative and an exponential integrator when  $r$  is large enough. The Hessian matrix needed for rotation onto the local normal mode basis should not be explicitly computed and is updated using only the gradient evaluations along the path. Such a mixed integrator works 3 to 20 times faster than a standard one when applied to a reaction path expressed in internal coordinates.

## 5.14 USE OF SADDLE CALCULATION

A SADDLE calculation uses two complete geometries, as shown on the following data file for the ethyl radical hydrogen migration from one methyl group to the other.

```

Line 1:          UHF  SADDLE
Line 2:          ETHYL RADICAL HYDROGEN MIGRATION
Line 3:
Line 4:      6      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 5:      6      1.479146 1      0.000000 0      0.000000 0      1 0 0
Line 6:      1      1.109475 1     111.328433 1      0.000000 0      2 1 0
Line 7:      1      1.109470 1     111.753160 1     120.288410 1      2 1 3
Line 8:      1      1.109843 1     110.103163 1     240.205278 1      2 1 3
Line 9:      1      1.082055 1     121.214083 1      38.110989 1      1 2 3
Line 10:     1      1.081797 1     121.521232 1     217.450268 1      1 2 3
Line 11:     0      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 12:     6      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 13:     6      1.479146 1      0.000000 0      0.000000 0      1 0 0
Line 14:     1      1.109475 1     111.328433 1      0.000000 0      2 1 0
Line 15:     1      1.109470 1     111.753160 1     120.288410 1      2 1 3
Line 16:     1      2.109843 1     30.103163 1     240.205278 1      2 1 3
Line 17:     1      1.082055 1     121.214083 1      38.110989 1      1 2 3
Line 18:     1      1.081797 1     121.521232 1     217.450268 1      1 2 3
Line 19:     0      0.000000 0      0.000000 0      0.000000 0      0 0 0
Line 20:

```

Details of the mathematics of SADDLE appeared in print in 1984, so only a superficial description will be given here.

The main steps in the saddle calculation are as follows:

1. The heats of formation of both systems are calculated.
2. A vector **R** of dimension 3N-6 defining the difference between the two geometries is calculated.
3. The scalar length **P** of the difference vector is reduced by some fraction, normally about 5 to 15 percent.
4. The geometry of lowest energy is called **G**.
5. **G** is optimized subject to the constraint that it maintains a constant distance **P** from the other geometry.
6. Repeat.

In the mechanism of Step 5, the coordinates of the moving geometry are perturbed by an amount equal to the product of the discrepancy between the calculated and required **P** and the vector **R**.

The two geometries must be related by a continuous deformation of the coordinates. By default, internal coordinates are used in specifying geometries. While bond lengths and bond angles are unambiguously defined (both positive), the dihedral angles can be positive or negative. Clearly 300 degrees could equally well be specified as -60 degrees. A wrong choice of dihedral would mean that instead of the desired reaction vector, a completely incorrect vector would be specified, with disastrous results.

As the specification of the geometries is quite difficult, SADDLE can also be made to run in Cartesian coordinates using the keyword XYZ. If this option is chosen then the initial steps of the calculation are as follows:

1. Both geometries are converted to Cartesian coordinates.
2. Both geometries are centered about the origin of Cartesian space.

3. One geometry is rotated until the difference vector is a minimum; this minimum is within 1 degree of the absolute bottom.
4. The SADDLE calculation then proceeds as described above.

### 5.15 TRANSITION STATE BY THE CHAIN METHOD

Optimization of a transition state can have several pitfalls which are not encountered during a standard energy minimization:

1. The “natural” reaction coordinate method not only generates expensive intermediate optimizations, but also can fail to locate a saddle point when a non-continuous path is generated.
2. The SADDLE method is a sophisticated variant of the preceding method but does not avoid its deficiencies.
3. The gradient norm minimization techniques are not selective and can find any nearby critical point including local minima.

The main problem is to ensure stability toward a true transition state lying between a given set of minima (respectively the reactants and the products of the chemical reaction studied). A mathematical answer involves stepping along the gradient vector field of an arbitrary continuous path between the reactants and the products. This leads to a limiting path where the highest energetic point is a saddle point, or in the case of a multi-step mechanism the highest energy transition state will be located.

The practical algorithm consists of replacing a chain of points  $C(n)=(R,...,p(i),...,P)$  running from reactants R to products P by a new one  $C(n+1)$  at iteration  $n+1$ . In order to maintain the connectivity of the path, each distance between two successive points is required not to exceed a given length (0.3 Å in AMPAC). An iteration roughly consists of skipping the current highest point of the chain along either a descending path (thus insuring the energetic relaxation of the whole path) or an ascending path (interpolation of a point along the path). New points are inserted as soon as a link length becomes too long. The successive evaluations of the gradient are used to update a quadratic local estimate of the potential, providing quadratic termination properties.

### 5.16 ANALYTICAL GRADIENT COMPUTATION

Analytical gradients are calculated for open shell and C.I. systems. The energy of a C.I. wave function is the sum of two distinct terms, the SCF energy and the C.I. correction. Each term depends on the coordinates of the nuclei (via the one and two electron integrals in a given basis set of atomic orbitals) and on the electronic cloud spanned by the molecular orbitals. Finally the M.O.'s are themselves dependant on the nuclear coordinates via the SCF equations, which dictate stationary conditions to be fulfilled by the electronic cloud at each nuclear geometry.

Thus the problem raised by the analytical derivation of the energy,  $E$ , reduces to the derivatives of a function  $E(\mathbf{q},\mathbf{e})$  with respect to the coordinates  $\{\mathbf{q}\}$ , the electronic cloud  $\{\mathbf{e}\}$  being an implicit function of  $\{\mathbf{q}\}$  through the SCF equations  $\mathbf{f}(\mathbf{q},\mathbf{e}) = 0$ . The answer is provided by differentiation with respect to a coordinate  $\mathbf{x}$  using the chain rule:

$$\frac{dE}{d\mathbf{x}} = E\left(\frac{d\mathbf{q}}{d\mathbf{x}}, \mathbf{e}\right) + E(\mathbf{q}, \frac{d\mathbf{e}}{d\mathbf{x}})$$

Differentiation of  $f(\mathbf{q},\mathbf{e}) = 0$  yields:

$$\frac{df}{dx} + \frac{df}{de} \frac{de}{dx} = 0 \quad \text{or} \quad \frac{de}{dx} = - \frac{de}{df} \frac{df}{dx}$$

In this notation, summation over all the variables is implicit, thus  $(df/de)$  is the electronic Hessian of the SCF equations. The second relation above gives the relaxation of the electronic cloud in response to a perturbation  $\mathbf{x}$  of the molecular geometry. Computationally, the electronic Hessian must be inverted. This is the only time consuming step for semi-empirical methods. The M.O. derivatives  $d\mathbf{C}$  are obtained from a rotation matrix  $\mathbf{B}$  expressed in the M.O. basis  $\mathbf{C}$ :

$$d\mathbf{C} = - \mathbf{C} \mathbf{B}$$

The derivative of the C.I. correction term follows from the derivatives of the C.I. matrix elements. This is a sum of the derivatives of the eigenvalues of the Fock matrix and the derivatives of the two electron integrals in the M.O. basis:

$$d \langle ij | kl \rangle = \langle ij | d(i) | kl \rangle + \langle (di)j | kl \rangle + \langle i(dj) | kl \rangle + \langle ij | (dk)l \rangle + \langle ij | k(dl) \rangle$$

The first term is the non-relaxed component and requires a full four index transformation. This transformation is not time consuming within the NDDO formalism, due to the relatively small number of non-vanishing two center integrals of type  $\langle pq | d(1/r) | rs \rangle$  over the atomic orbital basis set.

The four remaining terms represent the relaxation component of the derivative. These issue from a fast one index transformation involving the M.O. derivatives (the  $\mathbf{B}$  matrix) and the  $\langle nj | kl \rangle$  integrals, where  $\mathbf{n}$  is over all M.O.'s and  $\mathbf{j}, \mathbf{k}$  and  $\mathbf{l}$  run over all C.I. active M.O.'s:

$$\langle (di)j | kl \rangle = \sum_n B_{in} \langle nj | kl \rangle$$

## CHAPTER 6:        INSTALLATION, DEBUGGING AND PROGRAMMING

Installation of AMPAC requires a FORTRAN-77 compiler.

### 6.1        INSTALLING AMPAC ON A VAX/VMS SYSTEM

AMPAC is distributed on a magnetic tape as a set of FORTRAN files, along with command, help, data and results files. The standard tape format for a VAX-11/780 is used. The following instructions are VAX specific.

1. Put the magnetic tape on the tape drive, write protected.
2. Allocate the tape drive with a command such as `$ALLOCATE MTA0:.`
3. Create an empty directory to hold AMPAC.
4. Mount the magnetic tape with the command `$MOUNT MTA0: AMPAC.`
5. Copy all the files from the tape with the command:

```
$ COPY MTA0:*. * *
```

A useful operation after this would be to make a hard copy of the directory. You should now have the following sets of files in the directory:

1. A set of FORTRAN-77 files, see Section 6.9 for a listing.
2. The command files `COMPILE`, `AMPACCOM`, `AMPAC`, and `RAMPAC`.
3. A file, `AMPAC.OPT`, which lists all the object modules used by AMPAC.
4. A help file called `AMPAC.HLP`.
5. Several test-data files, and results files.

A recommended sequence of operations to get AMPAC up and running would be:

1. Edit the file `DIMSIZES.DAT` and modify the array sizes if necessary. The default sizes are 20 heavy atoms and 20 light atoms.
2. Edit the file `AMPAC.COM` and replace the `QUEUE` names with ones that are specific to your system.
3. Edit the file `RAMPAC.COM` if the default filenames are not acceptable.
4. Assign `AMPACDIRECTORY` to the disk and directory which contains AMPAC as a system-wide definition.
5. Edit the `SYLOGIN.COM` files to insert the following line:

```
$ AMPAC := @<AMPAC-directory>AMPAC.COM
```

6. Execute the modified SYLOGIN command so that the new commands are effective.
7. Run COMPILE.COM.
8. Test AMPAC with the test molecules that are supplied on the distribution tape.

A more complete discussion of this procedure is given in the next section.

## 6.2 DESCRIPTION OF THE COMMAND FILES

### COMPILE.COM

The parameter file DIMSIZES.DAT should be examined and modified, if necessary, *before* COMPILE is executed, since most of the FORTRAN files use the array sizes given in this file. If DIMSIZES.DAT needs to be changed later, COMPILE must be re-executed, as modules compiled with a different DIMSIZES.DAT will be incompatible.

The five adjustable parameters in DIMSIZES.DAT are:

MAXHEV: Maximum number of non-hydrogen atoms allowed in a calculation.

MAXLIT: Maximum number of hydrogen atoms allowed in a calculation.

NCHAIN: Maximum number of points allowed on the path of a chain calculation.

NMECI: Maximum number of active molecular orbitals allowed in a C.I. calculation.

NRELAX: Maximum array size allowed in analytical C.I. gradient calculation.

COMPILE assigns DIMSIZES.DAT the logical name "SIZES". This is a temporary assignment, but could be made permanent by suitably modifying LOGIN file(s).

All the FORTRAN files are then compiled, using the array sizes given in DIMSIZES.DAT. This operation takes several minutes, and should be run "interactively", as operator assistance is required.

If compilation is successful, the object files will be linked into an executable image called AMPAC.EXE.

To allow user access to AMPAC, the following system-wide assignment should be made in the SYLOGIN.COM command file:

```
$ AMPAC := @<AMPAC-directory>AMPAC
```

where <AMPAC-directory> is the name of the disk and directory which holds all the AMPAC files. For example if DUA0:[AMPAC] is the appropriate directory, then:

```
$ AMPAC := @DUA0:[AMPAC]AMPAC
```

AMPAC.COM and RAMPAC.COM should be modified once to accommodate local definitions of the directory which is to hold AMPAC.

## **AMPAC.COM**

This command file submits an AMPAC job to a batch queue and must be modified to suit local conditions. The user's VAX is assumed to run three queues, called QUEUE1, QUEUE2, and QUEUE3. The user should substitute the actual names of the VAX queues for these symbolic names.

## **RAMPAC.COM**

RAMPAC is the command file submitted for batch processing by AMPAC.COM and contains all the necessary channel assignments. Modification of channel assignments and file extensions, though not encouraged, require appropriate changes in RAMPAC.COM.

### **6.3        RUNNING   AMPAC**

The AMPAC.COM file that executes the new program can be accessed using the command "AMPAC" followed by appropriate arguments. Possible options are:

```
AMPAC <data file> <queue> <priority>
AMPAC <data file> <queue>
AMPAC <data file>
AMPAC
```

In the latter case the user will be prompted for further information. For example:

Enter the command:

```
$AMPAC
```

You will receive the message:

```
"What file? :"
```

to which the reply should be the actual data file name, e.g. "MNRSD1". The file is assumed to end in ".DAT". This file is supplied on the distribution tape.

You will then be prompted for the queue:

```
"What queue? :"
```

Any queue defined in AMPAC.COM will suffice:

```
SYS$BATCH
```

Finally, the priority will be requested:

```
"What priority? [5]:"
```

to which any value between 1 and 5 will suffice. Note that the maximum priority is set by the system (manager).

AMPAC produces the following files:

<filename>.OUT	Results
<filename>.ARC	Archive or summary
<filename>.RES	Restart
<filename>.DEN	Density matrix (in binary)
<filename>.GPT	Data for program DENSITY (in binary)

Note: Only a “.OUT” file is guaranteed. The remaining files depend on the condition of the calculation.

## 6.4 SHUTDOWN COMMAND

If a job needs to be stopped prematurely, the command \$SHUT <jobnumber> can be used to execute a small command-language file that modifies the file called <filename>.END that is only present during execution. The next time AMPAC calls the function SECOND, the apparent elapsed CPU time is increased by 1,000,000 seconds, exceeding the time limit for the calculation. Since there is not enough time left to do another cycle, the restart files are generated and the job is stopped.

SHUTDOWN is completely machine independent, and works successfully on the GOULD and Data General computers, in addition to the VAX 11-780.

## 6.5 DEBUGGING

There are three potential sources of difficulty in using AMPAC, each of which requires special attention. There can be problems with data, due to errors in the data, or AMPAC may be called upon to do calculations for which it was not designed. There are intrinsic errors in AMPAC which extensive testing has not yet revealed, but which a user's novel calculation uncovers. Finally there can be bugs introduced by the user modifying AMPAC, either to make it compatible with the host computer, or to implement local features.

For whatever reason, the user may need more information than the normal keywords provide. Additional keywords specific to debugging give information about the working of individual subroutines, but do not affect the course of the calculation.

### List of KEYWORDS for Debugging Subroutines

1ELEC	the one-electron matrix.	<i>Note 1</i>
COMPFG	Heat of Formation.	
DCART	Cartesian derivatives.	
DEBUG		<i>Note 2</i>
DENSITY	Every density matrix.	<i>Note 1</i>
DERIV	All gradients, and other data in DERIV.	
DFORCE	Print Force Matrix.	
EIGS	All eigenvalues.	
FLEPO	Details of D.F.P. minimization.	
FMAT	Force constant matrix.	
FOCK	Every Fock matrix.	<i>Note 1</i>
HCORE	The one electron matrix and two electron integrals.	
ITER	Values of variables and constants in ITER.	
LINMIN	Function values, step sizes at all points in LINMIN.	
LOCMIN	Function values, step sizes at all points in LOCMIN.	
MOLDAT	Molecular data, number of orbitals, “U” values, etc.	
MECI	C.I. matrices, M.O. indices, etc.	

PL	Differences between density matrix elements in ITER. Note the space before PL; this space is obligatory.	<i>Note 3</i>
SEARCH	Function values, step sizes at all points in SEARCH.	
TIMES	Times of stages within ITER.	
VECTORS	All eigenvectors on every iteration.	<i>Note 1</i>

*NOTE 1:* These key words are activated by the keyword DEBUG. Thus if DEBUG and FOCK are both specified, every Fock matrix on every iteration will be printed.

*NOTE 2:* DEBUG is not intended to increase the normal output, but does allow other keywords to have a special meaning.

*NOTE 3:* PL initiates the output of the value of the largest difference between any two density matrix elements on two consecutive iterations. This is very useful in cases of failure to achieve self consistence.

### **Suggested Procedure for Locating Bugs**

While AMPAC has been tested extensively, any modification of the source code can introduce new errors. Before wasting time on a “bug chase”, make sure that the version you are starting with is the QCPE released unmodified version. In circumstances where modifications must be made, the author of the changes is obviously responsible for removing offending bugs. The following ideas might prove useful in this context.

Before making any modifications, make a back-up copy of the standard AMPAC. This will prove invaluable in pinpointing deviations from standard operation. This point cannot be over-emphasized - **MAKE A BACK-UP BEFORE MODIFYING AMPAC!!!!**

Clearly, a bug can occur almost anywhere, and a logical search sequence is necessary in order to minimize the time taken to locate it. If possible, perform the debugging with a small molecule, in order to save time and minimize output.

There are two main mutually exclusive sequences in AMPAC, geometric and electronic. Subroutine COMPGF joins the geometric part to the electronic part. Since each part may be studied separately, the first question to be answered is which set contains the bug. If the heats of formation, derivatives, I.P.s, and charges, etc., are correct, the bug lies in the geometry section; if they are faulty, the bug may be found in the electronic section.

Use formaldehyde for testing. The supplied data file MNRSD1.DAT could be used as a template for this operation. Use keywords such as 1SCF, DEBUG, etc. The main steps are:

Check the starting one-electron matrix and two-electron integral string, using the keyword HCORE. It is normally sufficient to verify that the two hydrogen atoms are equivalent, and that the pi system involves only  $p_z$  on oxygen and carbon. Note that numerical values are not checked, but only relative values.

If an error is found, use MOLDAT to verify the orbital character, etc.

If faulty, the error lies in READ, GETGEO or MOLDAT.

Otherwise, the error lies in HCORE, H1ELEC or ROTATE.

Check the density or Fock matrix on every iteration, with the keywords FOCK or DENSITY. Check the equivalence of the two hydrogen atoms, and the pi system, as discussed above.

If an error is found, check the first Fock matrix. If faulty, the bug lies in ITER, probably in the Fock subroutines FOCK1 or FOCK2. or in the initial guess density matrix (MOLDAT). An exception is in the UHF closed-shell calculation, where a small asymmetry is introduced to initiate the separation of the alpha and beta UHF wavefunctions.

If no error is found, check the second Fock matrix. If faulty, the error lies in the density matrix DENSIT, or the diagonalization HQRIL.

If the Fock matrix is acceptable, check all the Fock matrices. If the error starts in iterations 2 to 4, the error probably lies in CNVG, if after that, in PULAY, if used.

If SCF is achieved, and the heat of formation is faulty, check HELECT. If C.I. was used check MECI.

If the derivatives are faulty, use DCART to verify the Cartesian derivatives. If these are faulty, check DCART and DHC. If they are correct, or not calculated, check the DERIV finite difference calculation.

If the geometric calculation is faulty, use FLEPO and LINMIN to monitor the optimization, DERIV may also be useful here.

For the FORCE calculation, DCART or DERIV are useful for variationally optimized functions, COMPGF for non-variationally optimized functions.

For reaction paths, verify that FLEPO is working correctly; if so, then PATHS is faulty.

For saddle-point calculations, verify that FLEPO is working correctly; if so, then REACT1 is faulty.

## 6.6 DIFFERENCES FROM AMPAC VERSION 1.0

### Criteria and Accuracy

SCF convergence criteria:

The “secondary” criterion (see *Background*) on the density matrix is now coupled with the “first” one on the energy in order to obtain a smooth and more confident behavior of the SCF response versus the required criterion SCFCRT. This is required in the frequent cases where the energy seems stationary but the density matrix is still oscillating. (e.g. molecules with heteroatoms engaged in a donor-acceptor bond.) One obtains a comparable level of self-consistency for both the energy and the density matrix and more accurate results as the SCFCRT is lowered.

Gradient computation:

For closed shell systems, the gradient accuracy depends on the density matrix only, but in other cases, it depends on the energy via a finite difference calculation of the change in energy with respect to small changes in the coordinates. The damped behavior of the SCF response vs SCFCRT now produces a smooth increase in the gradient accuracy for a regular decrease in SCFCRT.

For open shell cases the step size of the selected finite difference formula becomes a function of SCFCRT, thus minimizing the error on the gradient components. A statistical estimate of the error is provided in both closed and open shell cases as a function of SCFCRT. Analytical gradients are now computed for half-electron and CI systems.

Optimization criteria:

Most of the time, the standard criterion on the gradient norm has been replaced by a criterion on the root mean square of the gradient:

$$|G|_{\text{rms}} = \sqrt{\sum_i \frac{g_i^2}{n}}$$

thus taking into account size effects for both small and large molecules. Elsewhere, the gradient's error is taken as a lower bound for the convergence threshold of the various optimization algorithms.

Keyword "PRECISE":

The impact of this keyword has been completely modified in order to actually obtain the accuracy needed. It acts on the SCF criterion SCFCRT (factor 100), on the optimization criterion (factor 10) and on the finite difference formula used in gradient and Hessian computation. Be aware of a factor two in the gradient computation for open-shell systems.

Optimization of the code:

Systematic replacement of multiplications by additions has been done in the computation of the indices of the matrix elements when stored linearly. The code is inelegant but significantly more efficient delivering a 20% increase in speed in the diagonalization routine HQRIL.

### Optimization Package

Five new algorithms have appeared in the arsenal available in AMPAC. They are accessed by keyword and provide various alternatives to existing routines. Restart procedures have been implemented in each case.

Type of Calculation:	New Option:	Old Option (still available):
energy minimization	NEWTON	DFP
gradient minimization	POWELL , LTRD	NLLSQ,SIGMA
Hessian computation	LTRD , NEWTON	FORCE
transition state search	CHAIN	SADDLE
reaction path integration	PATH	DRC(damped)

Some keywords can affect control in several routines. For example, CYCLES=n can be used with either POWELL or NLLSQ. Keywords that provide control in more than one routine are given below:

CYCLES=n	- maximum number of iterations	n = 0,...
GNORM=n.n	- convergence threshold on the rms gradient	kcal/mol
PRINT=n	- printout level in each iteration	n = 0,...,5
T=n	- maximum allowed cpu time	n = seconds

For most applications in chemistry, a computation of the matrix of the second derivatives (Hessian) is required for only one of the two following purposes:

- full thermodynamic calculations and/or infra-red spectrum,
- characterization of a stationary point in terms of the number of negative eigenvalues.

In the first case, an accurate Hessian is required either in the full 3N Cartesian space (FORCE) or alternatively in the full 3N-6 configuration space (NEWT or LTRD with CYCLES=0 and PRECISE). In

the second case, a less expensive estimation can be performed using a less accurate, but faster, finite difference formula (NEWT or LTRD without PRECISE).

The only limitation on the lack of accuracy is that it must not affect the signs of the eigenvalues. This is achieved in NEWT or LTRD with variable step sizes in the finite difference formula and statistical estimates of the errors of the eigenvalues, via the previous estimated errors on both the energy and the gradients.

Typical discrepancy between FORCE and LTRD + PRECISE < 0.5 cm<sup>-1</sup>

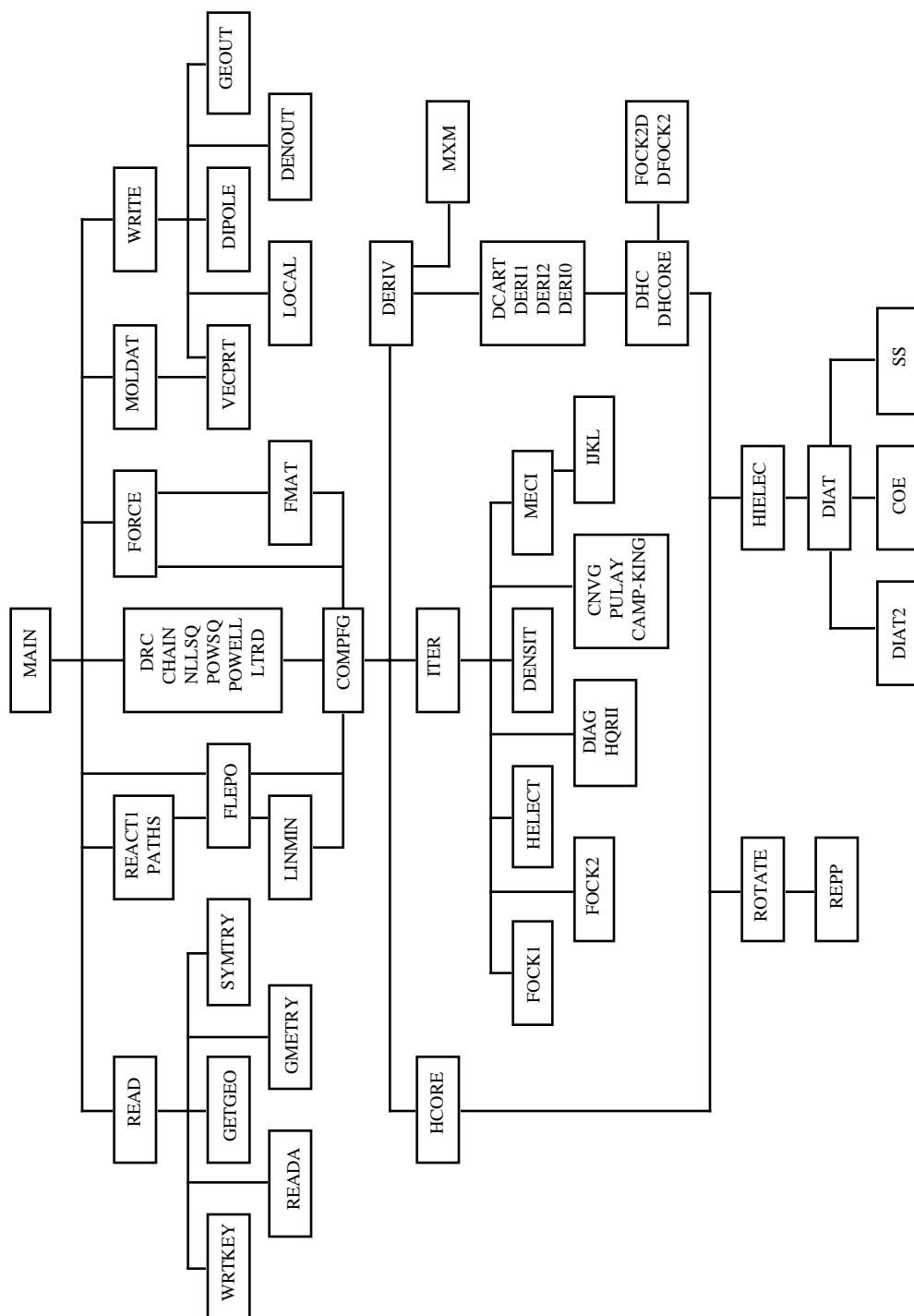
Typical discrepancy between FORCE and LTRD < 5.0 cm<sup>-1</sup>

### New KEYWORDS Implemented in AMPAC

Several new keywords have been added and are discussed in detail in the *Keywords* chapter. They are listed below.

PRINT	- PRINTOUT LEVEL IN OPTIMIZATION = n
NEWTON	- MINIMIZE ENERGY USING FULL-NEWTON
LTRD	- MINIMIZE GRADIENT USING FULL-NEWTON
POWELL	- MINIMIZE GRADIENT USING POWELL METHOD
PATH	- FOLLOW THE DESCENDING REACTION PATH
CHAIN	- TRANSITION STATE TO BE OPTIMIZED
T.V.	- TRANSITION VECTOR TO BE PROVIDED FOR PATH
WEIGHT	- WEIGHT TO BE PROVIDED FOR PATH

## 6.7 AMPAC PROGRAM FLOW DIAGRAM



## 6.8 FORTRAN-77 FILES

AABABC	ANAVIB	AXIS	BLOCK	BONDS
CALPAR	CHAIN	CHRG,	CNVG	COMPFG
DATIN	DCART	DENROT	DENSIT	DEPVAR
DERIV	DERI0	DERI1	DERI2	DFPSAV
DIAG	DIAT	DIAT2	DIPOLE	DRC
ENPART	EXCHNG	FLEPO	FMAT	FOCK1
FOCK2	FOCK2D	FORCE	FORSAV	FRAME
FREQCY	GEOUT	GETGEO	GETSYM	GMETRY
GRID	H1ELEC	HADDON	HCORE	HELECT
HQRII	IJKL	INTERP	ITER	LINMIN
LOCAL	LOCMIN	LTRD	MAIN	MATOUT
MATPAK	MECI	MOLDAT	MULLIK	MULT
NLLSQ	NUCHAR	OSINV	PATH	PATHS
PERM	POLAR	POWELL	POWSQ	PULAY
REACT1	READ	READA	REPP	ROTATE
SAVOPT	SEARCH	SECOND	SOLROT	SWAP
SYMTRY	THERMO	UPDATE	VECPRT	VECRE
VECWRT	WRITE	WRTKEY	XYZINT	ZPE

## 6.9 SUBROUTINE CALLS IN AMPAC

The program segments which call various subroutines are listed below. Note that function calls are not included here!

<u>SUBROUTINE</u>	<u>CALLS</u>
AABABC (F)	
AABACD (F)	
AABBBCD (F)	
AINTGS	
AM1	MOLDAT CALPAR
ANAVIB	
AXIS	HQRII
BABBBC (F)	
BABBBCD (F)	
BANGLE	
BFN	
BINTGS	
BONDS VECPR	
CALPAR	
CARDAN	
CHAIN	SAVOPT GETGEO COMPFG COL
CHRG	
CNVG	
COE	
COL	VOISIN RABOT PTHAUT QUADR ECRIT QUADRI
COMPFG	GMETRY HCORE ITER DERIV
CVMGT (F)	
DANG	
DCART	
DENROT	GMETRY COE

DENSIT	
DEPVAR	
DERI0	
DERI1	DHCORE SCOPY DFOCK2 SUPDOT MTXM
	MXM DIJKL1 MECID MECIH
DERI2	MXMT DERI21 SCOPY DERI22 MXM OSINV MTXM
	DERI23 DIJKL2 MECID MECIH SUPDOT
DERI21	MTXMC DIAGIV MXM
DERI22	MXM MXMT FOCK2 FOCK1 SUPDOT MTXM
DERI23	SCOPY
DERIV	GMETRY HCORE ITER DERI0 DERI1 DERI2 SCOPY
	DCART JCARIN MXM
DFOCK2	
DFPSAV	XYZINT
DHC	H1ELEC ROTATE SOLROT FOCK2D
DHCORE	H1ELEC ROTATE WCANON
DIAG	
DIAGI (F)	
DIAGIV	SAXPY
DIAT	COE DIAT2
DIAT2	SET
DIHED	DANG
DIJKL1	DPQTKL MTXMC
DIJKL2	
DIPOLE (F)	
DOT (F)	
DPQTKL	MXM SAXPY
DRC	GMETRY COMPGF CHRGE XYZINT
ECRIT	
ENPART	
EULER	
EXCHNG	
FINDS	SUPDOT INVRT1 CARDAN SCOPY
FLEPO	DFPSAV COMPGF SCOPY GEOUT SUPDOT LINMIN
FMAT	FORSAV COMPGF CHRGE FRAME HQRII MATOUT
FOCK1	CHRGE
FOCK2	
FOCK2D	
FORCE	GMETRY COMPGF NLLSQ FLEPO WRITE AXIS FMAT
	VECPRT FRAME HQRII MATOUT FREQCY ANAVIB
	THERMO
FORSAV	
FRAME	AXIS
FREQCY	FRAME HQRII
GEOUT	XYZINT
GETGEO	XYZINT EXIT
GETSYM	NUCHAR
GMETRY	INTCAR
GRID	FLEPO
H1ELEC	DIAT
HADDON	DEPVAR
HCORE	H1ELEC VECPR
HELECT (F)	
HQRII	
IJKL	PQTKL MTXMC MXM
INTCAR	GEOUT
INTERP	HQRII SCHMIT SCHMIB SYSTEM SPLINE

ITER	VECREG SCOPY VECPRG FOCK2 FOCK1 INTERP PULAY HQRH DIAG MATOUT DENSIT CNVG WRITE
JCARIN	GMETRY
JNCAR	RLOCAL
LDATA	
LINMIN	COMPFG EXCHNG
LOCAL	MATOUT
LOCMIN	SAXPY COMPFG EXCHNG SCOPY
LTRD	SAVOPT SCOPY POINT1 COMPFG POINT2 POINT3 POINT4 POINT5 FINDS STAT ZPE
MAIN	READ AM1 MOLDAT DRC PATH REACT1 CHAIN PATHS GRID NLLSQ POWSQ POWELL FORCE LTRD FLEPO WRITE POLAR
MAMULT	
MATOUT	
MECI (F)	IJKL PERM MECID MECIH VECPRG HQRH MATOUT
MECID	
MECIH	
MECIP	MXM
MOLDAT	GMETRY VECPRG
MOPOUT	EULER
MTXM	
MTXMC	MXM
MULLIK	HQRH GMETRY MULT DENSIT VECPRG
MULT	
MXM	
MXMT	
NLLSQ	SAVOPT SCOPY COMPFG GEOUT LOCMIN SAXPY
NUCHAR	
OSINV	
PATH	SAVOPT SCOPY COMPFG READVT PATH1
PATH1	PATH2 MXM DIAGIV SUPDOT
PATH2	
PATHS	DFPSAV FLEPO WRITE
PERM	
POINT1	DIAGIV
POINT3	DIAGIV
POLAR	GMETRY COMPFG VECPRG HQRH MATOUT
POWEL1	INVERT
POWELL	SAVOPT SCOPY COMPFG POWEL1 POWEL2 POWEL3
POWSQ	SAVOPT SCOPY COMPFG INVERT VECPRG HQRH SUPDOT SEARCH SAXPY
PQTKL	MXM SAXPY
PULAY	SCOPY MAMULT MXM OSINV
QUADRI	INVERT DIAGIV INVRT1 SUPDOT
RABIOT	
REACT1	GETGEO GEOUT GMETRY FLEPO COMPFG WRITE
READ	GETGEO WRTKEY NUCAR GEOUT GMETRY
READA (F)	
READVT	
REPP	
RLOCAL	
ROTATE	REPP
SAVOPT	
SAXPY	
SCHMIB	

SCHMIT	
SCOPY	
SDOT (F)	
SEARCH	COMPFG SCOPY
SECOND (F)	TIMCLK
SELECT	
SET	AINTGS BINTGS
SOLROT	
SPLINE	
SS (F)	BFN
STAT	STAT1 COMPFG STAT2 STAT3
STAT1	SCOPY STATUS CARDAN SELECT
STATUS	
SUPDOT	
SWAP	
SYMTRY	HADDON
THERMO	
TIMBGN	
UPDATE	
VECPRT	
VECREG	
VECWRT	
WCANON	
WNONCA	
WRITE	DATE GEOUT DERIV IJKL GMETRY VECPR MATOUT MECIP CHRGE DENROT BONDS LOCAL ENPART VECWRT MULLIK
WRTKEY	
XYZGEO	BANGLE DIHED
XYZINT	XYZGEO
ZPE	INTCAR JINCAR SUPDOT VECPR DIAGIV

## 6.10 DESCRIPTION OF MAJOR SUBROUTINES IN AMPAC

<i>AABABC Utility:</i>	Calculates the configuration interaction matrix element between two configurations differing by exactly one alpha M.O. Called by MECI only.
<i>AABACD Utility:</i>	Calculates the configuration interaction matrix element between two configurations differing by exactly two alpha M.O.s. Called by MECI only.
<i>AABBBCD Utility:</i>	Calculates the configuration interaction matrix element between two configurations differing by exactly two M.O.s; one configuration has alpha M.O. "A" and beta M.O. "C" while the other configuration has alpha M.O. "B" and beta M.O. "D". Called by MECI only.
<i>ANAVIB Utility:</i>	Gives a brief interpretation of the modes of vibration of the molecule. The principal pairs of atoms involved in each vibration are identified, and the mode of motion (tangential or radial) is output.
<i>AXIS Utility:</i>	Calculates the three principal moments of inertia of a molecule. If the system is linear, one moment of inertia is zero. Prints moments in units of $\text{cm}^{-1}$ and $10^{-40}$ gram- $\text{cm}^2$ .

<i>BABBBC Utility:</i>	Calculates the configuration interaction matrix element between two configurations differing by exactly one beta M.O. Called by MECI only.
<i>BABBCD Utility:</i>	Calculates the configuration interaction matrix element between two configurations differing by exactly two beta M.O.s. Called by MECI only.
<i>BONDS Utility:</i>	Evaluates and prints the valencies of atoms and bond-orders between atoms. Main argument: density matrix. No results are passed to the calculation, and no data are changed. Called by WRITE only.
<i>CALPAR Utility:</i>	When external parameters are entered via EXTERNAL=, the derived parameters are worked out using CALPAR. Note that all derived parameters are calculated for all parametrized elements at the same time.
<i>CHARGE Utility:</i>	Calculates the total number of valence electrons on each atom. Main arguments: density matrix, array of atom charges (empty on input). Called by ITER only.
<i>CNVG Utility:</i>	Used in SCF cycle. CNVG does a three-point interpolation of the last three density matrices. Arguments: Last three density matrices, Number of iterations, measure of self-consistency (empty on input). Called by ITER only.
<i>COMPFG Main Sequence:</i>	Evaluates the total heat of formation of the supplied geometry, and the derivatives, if requested. This routine connects the electronic and geometric parts of the program. Main arguments on input: geometry, on output: heat of formation, gradients.
<i>DENSIT Utility:</i>	Constructs the Coulson electron density matrix from the eigenvectors. Main arguments: Eigenvectors, No. of singly and doubly occupied levels, density matrix (empty on input) Called by ITER.
<i>DENROT Utility:</i>	Converts the ordinary density matrix into a condensed density matrix over basis functions s (sigma), p (sigma) and p (pi), i.e., three basis functions. Useful in hybridization studies. Has capability to handle "d" functions, if present.
<i>DEPVAR Utility:</i>	A symmetry-defined "bond length" is related to another bond length by a multiple. This special symmetry function is intended for use in Cluster calculations. Called by HADDON.
<i>DERIV Main Sequence:</i>	Calculates the derivatives of the energy with respect to the geometric variables. This is done either by using initially Cartesian derivatives (normal mode) or by full SCF calculations (half-electron and C.I. mode). Arguments on input: geometry, on output: derivatives. Called by COMPFG.
<i>DFPSAV Utility:</i>	Saves and restores data used by the Davidon-Fletcher-Powell geometry optimization. Main arguments: parameters being optimized, gradients of parameters, last heat of formation, integer and real control data. Called by FLEPO.
<i>DIAG Utility:</i>	Rapid pseudo-diagonalization. Given a set of vectors which almost block-diagonalize a secular determinant, DIAG modifies the vectors so that the

	block-diagonalization is more exact. Main arguments: Old vectors, Secular Determinant, New vectors (on output). Called by ITER.
DIAGI Utility:	Calculates the electronic energy arising from a given configuration. Called by MECI.
DIAT Utility:	Calculates overlap integrals between two atoms in general Cartesian space. Principal quantum numbers up to 6, and angular quantum numbers up to 2 are allowed. Main arguments: Atomic numbers and Cartesian coordinates in Angstroms of the two atoms, Diatomic overlaps (on exit). Called by H1ELEC.
DIAT2 Utility:	Calculates reduced overlap integrals between atoms of principal quantum numbers 1, 2, and 3, for s and p orbitals. Faster than the SS in DIAT. This is a dedicated subroutine, and is unable to stand alone without considerable backup. Called by DIAT.
DIPOLE Utility:	Evaluates and, if requested, prints dipole components and dipole for the molecule. Arguments: Density matrix, Charges on every atom, coordinates, dipoles (on exit). Called by WRITE and FMAT.
DOT Utility:	Given two vectors, X and Y, of length N, function DOT returns with the dot product X.Y; i.e., if X=Y, then DOT = the square of X. Called by FLEPO.
DRC	The dynamic reaction coordinate follows the mass-weighted path of a reaction in Cartesian coordinates.
ENPART Utility:	Partitions the energy of a molecule into its monatomic and diatomic components. Called by WRITE when the keyword ENPART is specified. No data are changed by this call.
EXCHNG Utility:	Dedicated procedure for storing 3 parameters and one array in a store. Used by LINMIN.
FLEPO Main Sequence:	Optimizes a geometry by minimizing the energy. Makes use of the first and estimated second derivatives to achieve this end. Arguments: Parameters to be optimized, (overwritten on exit with the optimized parameters), Number of parameters, final optimized heat of formation. Called by MAIN, REACT1, and FORCE.
FMAT Main sequence:	Calculates the exact Hessian matrix for a system This is done by either using differences of first derivatives (normal mode) or by four full SCF calculations (half electron or C.I. mode). Called by FORCE.
FOCK1 Utility:	Adds on to Fock matrix the one-center two electron terms. Called by ITER only.
FOCK2 Utility:	Adds on to Fock matrix the two-center two electron terms. Called by ITER and DERIV. In ITER the entire Fock matrix is filled; in DERIV, only diatomic Fock matrices are constructed.
FOCK2D Utility:	Virtually identical to FOCK2, except that it is used in derivative calculations. Called by DHC.

<i>FORCE Main sequence:</i>	Performs a force-constant and vibrational frequency calculation on a given system. If the starting gradients are large, the geometry is optimized to reduce the gradient norm, unless LET is specified in the keywords. Isotopic substitution is allowed. Thermochemical quantities are calculated. Called by MAIN.
<i>FORSAV Utility:</i>	Saves and restores data used in FMAT in FORCE calculation. Called by FMAT.
<i>FRAME Utility:</i>	Applies a very rigid constraint on the translations and rotations of the system. Used to separate the trivial vibrations in a FORCE calculation.
<i>FREQCY Main sequence:</i>	Final stage of a FORCE calculation. Evaluates and prints the vibrational frequencies and modes.
<i>GEOOUT Utility:</i>	Prints out the current geometry. Can be called at any time. Does not change any data.
<i>GETGEO Utility:</i>	Reads in geometry in character mode from specified channel, and stores parameters in arrays. Some error-checking is done. Called by READ and REACT1.
<i>GETSYM Utility:</i>	Reads in symmetry data. Used by READ.
<i>GMETRY Utility:</i>	Fills the Cartesian coordinates array. Data are supplied from the array GEO, GEO can be (a) in internal coordinates, or (b) in Cartesian coordinates. If STEP is non-zero, then the coordinates are modified in light of the other geometry and STEP. Called by HCORE, DERIV, READ, WRITE, MOLDAT, etc.
<i>H1ELEC Utility:</i>	Given any two atoms in Cartesian space, H1ELEC calculates the one-electron energies of the off-diagonal elements of the atomic orbital matrix. Called by HCORE and DERIV.
<i>HADDON Utility:</i>	The symmetry operation subroutine, HADDON relates two geometric variables by making one a dependent function of the other. Called by SYMTRY only.
<i>HCORE Main sequence:</i>	Sets up the energy terms used in calculating the SCF heat of formation. Calculates the one and two electron matrices, and the nuclear energy. Called by COMPFG.
<i>HELECT Utility:</i>	Given the density matrix, and the one electron and Fock matrices, calculates the electronic energy. No data are changed by a call of HELECT. Called by ITER and DERIV.
<i>HQRH Utility:</i>	Rapid diagonalization routine. Accepts a secular determinant, and produces a set of eigenvectors and eigenvalues. The secular determinant is destroyed.
<i>IJKL Utility:</i>	Fills the large two-electron array over a M.O./ basis set. Calls SPCG, and is called by MECI.
<i>INTERP Utility:</i>	Runs the Camp-King converger.

<i>ITER Main sequence:</i>	Given the one and two electron matrices, ITER calculates the Fock and density matrices, and the electronic energy. Called by COMPG.
<i>LINMIN Main sequence:</i>	Does a line-search to minimize the heat of formation of a system. Arguments: starting geometry, search direction, initial step size, initial heat of formation; on exit: optimized geometry, optimized heat of formation, flags. Called exclusively by FLEPO.
<i>LOCAL Utility:</i>	Given a set of occupied eigenvectors, produces a canonical set of localized bonding orbitals, by a series of 2 x 2 rotations. Called by WRITE.
<i>LOCMIN Main sequence:</i>	In a gradient minimization, LOCMIN does a line-search to find the gradient norm minimum. Main arguments: current geometry, search direction, step, current gradient norm; on exit: optimized geometry, gradient norm.
<i>MAMULT Utility:</i>	Matrix multiplication. Two matrices, stored as lower half triangular packed arrays, are multiplied together, and the result stored in a third array as the lower half triangular array. Called from PULAY.
<i>MATOUT Utility:</i>	Matrix printer. Prints a square matrix, and a row-vector, usually eigenvectors and eigenvalues. The indices printed depend on the size of the matrix: they can be either over orbitals, atoms, or simply numbers, thus M.O.'s are over orbitals, vibrational modes are over numbers. Called by WRITE, FORCE.
<i>MECI Main sequence:</i>	Main function for Configuration Interaction, MECI constructs the appropriate C.I. matrix, and evaluates the roots, which correspond to the electronic energy of the states of the system. The appropriate root is then returned. Called by ITER only.
<i>MAIN Main sequence:</i>	MAIN program. MAIN first reads in data using READ, then calls either FLEPO to do geometry optimization, FORCE to do a FORCE calculation, PATHS for a reaction with a supplied coordinate, NLLSQ for a gradient minimization or REACT1 for locating the transition state. Starts the timer.
<i>MOLDAT Main Sequence:</i>	Sets up all the invariant parameters used during the calculation, e.g. number of electrons, initial atomic orbital populations, number of open shells, etc. Called once by MAIN only.
<i>MULLIK Utility:</i>	Constructs and prints the Mulliken Population Analysis. Available only for RHF calculations. Called by WRITE.
<i>MULT Utility:</i>	Used by MULLIK only, MULT multiplies two square matrices together.
<i>NLLSQ Main sequence:</i>	Used in the gradient norm minimization.
<i>OSINV Utility:</i>	Inverts a square matrix. Called by PULAY only.
<i>PARSAV Utility:</i>	Stores and restores data used in the gradient-norm minimization calculation.
<i>PATHS Main sequence:</i>	Given a reaction coordinate as a row-vector, PATHS performs a FLEPO geometry optimization for each point, the later geometries being initially guessed from a knowledge of the already optimized geometries, and the current step. Called by MAIN only.

PERM <i>Utility:</i>	Permutes n1 electrons of alpha or beta spin among n2 M.O.s.
POWSAV <i>Utility:</i>	Calculation store and restart for SIGMA calculation. Called by POWSQ.
POWSQ <i>Main sequence:</i>	The McIver - Komornicki gradient minimization routine. Constructs a full Hessian matrix and proceeds by line-searches Called from MAIN when SIGMA is specified.
POLAR <i>Utility:</i>	Calculates the polarizability volumes for a molecule or ion. Uses 19 SCF calculations, so appears after WRITE has finished. Cannot be used with FORCE, but can be used anywhere else. Called by WRITE
PULAY <i>Utility:</i>	Uses a powerful mathematical non-iterative method for obtaining the SCF Fock matrix. Principle is that at SCF the eigenvectors of the Fock and density matrices are identical, so [F.P] is a measure of the non-self consistency. While very powerful, PULAY is not universally applicable. Used by ITER.
REACT1 <i>Main sequence:</i>	Uses reactants and products to find the transition state. A hypersphere of N dimensions is centered on each moiety, and the radius steadily reduced. The entity of lower energy is moved, and when the radius vanishes, the transition state is reached. Called by MAIN only.
READ <i>Main sequence:</i>	Almost all the data are read in through READ. There is a lot of data-checking in READ, but very little calculation. Called by MAIN.
READA <i>Utility:</i>	General purpose character number reader. Used to enter numerical data in the control line as “ <variable>=n.nnn ” where <variable> is a mnemonic such as SCFCRT or CHARGE. Called by READ, FLEPO, ITER, FORCE, and many other subroutines.
REPP <i>Utility:</i>	Calculates the 22 two-electron reduced repulsion integrals, and the 8 electron-nuclear attraction integrals. These are in a local coordinate system. Arguments: atomic numbers of the two atoms, interatomic distance, and arrays to hold the calculated integrals. Called by ROTATE only.
ROTATE <i>Utility:</i>	All the two-electron repulsion integrals, the electron-nuclear attraction integrals, and the nuclear-nuclear repulsion term between two atoms are calculated here. Typically 100 two-electron integrals are evaluated.
SEARCH <i>Utility:</i>	Part of the SIGMA gradient minimization. The line-search subroutine, SEARCH locates the gradient minimum and calculates the second derivative of the energy in the search direction. Called by POWSQ.
SECOND <i>Utility:</i>	Contains VAX specific code. Function SECOND returns the number of CPU seconds elapsed since an arbitrary starting time. If the SHUTDOWN command has been issued, the CPU time is in error by exactly 1,000,000 seconds, and the job usually terminates with the message “time exceeded”.
SOLROT <i>Utility:</i>	For Cluster systems, adds all the two-electron integrals of the same type, between different unit cells, and stores them in a single array. Has no effect on molecules.

<i>SWAP Utility:</i>	Used with FILL=, SWAP ensures that a specified M.O. is filled. Called by ITER only.
<i>SYMTRY Utility:</i>	Calculates values for geometric parameters from known geometric parameters and symmetry data. Called whenever GMETRY is called.
<i>THERMO Main sequence:</i>	After the vibrational frequencies have been calculated, THERMO calculates thermodynamic quantities such as internal energy, heat capacity, entropy, etc, for translational, vibrational, and rotational, degrees of freedom.
<i>VECPRT Utility:</i>	Prints out a packed, lower-half triangular matrix. The labeling of the sides of the matrix depend on the matrix's size: if it is equal to the number of orbitals, atoms, or other. Arguments: The matrix to be printed, size of matrix. No data are changed by a call of VECPRY.
<i>WRITE Main sequence:</i>	Most of the results are printed here. All relevant arrays are assumed to be filled. A call of WRITE only changes the number of SCF calls made, this is reset to zero. No other data are changed. Called by MAIN, FLEPO, FORCE.
<i>XYZINT Utility:</i>	Converts from Cartesian coordinates into internal. XYZINT sets up its own numbering system, so no connectivity is needed. Used only in conjunction with keyword XYZ.



## **APPENDIX A: ERROR MESSAGES PRODUCED BY AMPAC**

AMPAC produces several hundred messages, all of which are intended to be self-explanatory. However, when an error occurs it is useful to have more information than is given in the standard messages.

The following alphabetical list gives more complete definitions of the messages printed.

### **AN UNOPTIMIZABLE GEOMETRIC PARAMETER...**

When internal coordinates are supplied, six coordinates cannot be optimized. These are the three coordinates of atom 1, the angle and dihedral on atom 2 and the dihedral on atom 3. An attempt has been made to optimize one of these. This is usually indicative of a typographic error, but might simply be an oversight. Either way, the error will be corrected and the calculation will not be stopped here.

### **ATOM NUMBER nn IS ILLDEFINED**

The rules for definition of atom connectivity are:

1. Atom 2 must be connected to atom 1 (default - no override)
2. Atom 3 must be connected to atom 1 or 2, and make an angle with 2 or 1.
3. All other atoms must be defined in terms of already-defined atoms: these atoms must all be different. Thus atom 9 might be connected to atom 5, make an angle with atom 6, and have a dihedral with atom 7. If the dihedral was with atom 5, then the geometry definition would be faulty.

If any of these rules is broken, a fatal error message is printed, and the calculation stopped.

### **ATOMIC NUMBER nn IS NOT AVAILABLE...**

An element has been used for which parameters are not available. Only if a typographic error has been made can this be rectified. This check is not exhaustive, in that even if the elements are acceptable there are some combinations of elements within MINDO/3 that are not allowed. This is a fatal error message.

### **ATOMIC NUMBER OF nn ?**

An atom has been specified with a negative or zero atomic number. This is normally caused by forgetting to specify an atomic number or symbol. This is a fatal error message.

### **ATOMS nn AND nn ARE SEPARATED BY nn.nnnn ANGSTROMS**

Two genuine atoms (not dummies) are separated by a very small distance. This can occur when a complicated geometry is being optimized, in which case the user may wish to continue. This can be done by using the key-word GEO-OK. More often, however, this message indicates a mistake, and the calculation is, by default, stopped.

**ATTEMPT TO GO DOWNHILL IS UNSUCCESSFUL...**

A quite rare message, produced by Bartel's gradient norm minimization. Bartel's method attempts to minimize the gradient norm by searching the gradient space for a minimum. Apparently a minimum has been found, but not recognized as such. The program has searched in all (3N-6) directions, and found no way down, but the criteria for a minimum have not been satisfied. No advice is available for getting round this error.

**BOTH SYSTEMS ARE ON THE SAME SIDE...**

A non-fatal message, but still cause for concern. During a SADDLE calculation the two geometries involved are on opposite sides of the transition state. This situation is verified at every point by calculating the cosine of the angle between the two gradient vectors. For as long as it is negative, then the two geometries are on opposite sides of the T/S. If, however, the cosine becomes positive, then the assumption is made that one moiety has fallen over the T/S and is now below the other geometry. That is, it is now further from the T/S than the other, temporarily fixed, geometry.

**C.I. NOT ALLOWED WITH UHF**

There is no UHF configuration interaction calculation in AMPAC. Either remove the key-word that implies C.I. or the word UHF.

**CALCULATION ABANDONED AT THIS POINT**

A particularly annoying message! In order to define an atom's position, the three atoms used in the connectivity table must not accidentally fall into a straight line. This can happen during a geometry optimization or gradient minimization. If they do, and if the angle made by the atom being defined is not zero or 180 degrees, then its position becomes ill-defined. This is not desirable, and the calculation will stop in order to allow corrective action to be taken. Note that if the three atoms are in an exactly straight line, then this message will not be triggered.

**CARTESIAN COORDINATES READ IN, AND CALCULATION...**

If Cartesian coordinates are read in, but the calculation is to be carried out using internal coordinates, then either all possible geometric variables must be optimized, or none can be optimized. If only some are marked for optimization then ambiguity exists. For example, if the "X" coordinate of atom 6 is marked for optimization, but the "Y" is not, then when the conversion to internal coordinates takes place, the first coordinate becomes a bond-length, and the second an angle. These bear no relationship to the "X" or "Y" coordinates. This is a fatal error.

**CARTESIAN COORDINATES READ IN, AND SYMMETRY...**

If Cartesian coordinates are read in, but the calculation is to be carried out using internal coordinates, then any symmetry relationships between the Cartesian coordinates will not be reflected in the internal coordinates. For example, if the "Y" coordinates of atoms 5 and 6 are equal, it does not follow that the internal coordinate angles these atoms make are equal. This is a fatal error.

### **ELEMENT NOT FOUND**

When an external file is used to redefine MNDO or AM1 parameters, the chemical symbols used must correspond to known elements. Any that do not will trigger this fatal message.

### **ERROR DURING READ AT ATOM NUMBER...**

Something is wrong with the geometry data. In order to help find the error, the geometry already read in is printed. The error lies either on the last line of the geometry printed, or on the next (unprinted) line. This is a fatal error.

### **FAILED IN SEARCH, SEARCH CONTINUING**

Not a fatal error. The McIver-Komornicki gradient minimization involves use of a line-search to find the lowest gradient. This message is merely advice. However, if SIGMA takes a long time, consider doing something else, such as using NLLSQ, or refining the geometry a bit before resubmitting it to SIGMA.

**<<<<----\*\*\*\* FAILED TO ACHIEVE SCF. \*\*\*\*---->>>>**

The SCF calculation failed to go to completion; an unwanted and depressing message that unfortunately appears every so often.

To date three unconditional convergers have appeared in the literature: the SHIFT technique, Pulay's method, and the Camp-King converger. It would not be fair to the authors to condemn their methods. In AMPAC all sorts of weird and wonderful systems are calculated, systems the authors of the convergers never dreamed of. AMPAC uses a combination of all three convergers at times. Normally only a quadratic damper is used.

If this message appears, suspect first that the calculation might be faulty, then, if you feel confident, try altering the SHIFT, or invoking PULAY or CAMP-KING on their own.

If nothing works, then consider slackening the SCF criterion. This will allow heats of formation to be calculated with reasonable precision, but the gradients are likely to be imprecise.

### **GEOMETRY TOO UNSTABLE FOR EXTRAPOLATION...**

In a reaction path calculation the initial geometry for a point is calculated by quadratic extrapolation using the previous three points.

If a quadratic fit is likely to lead to an inferior geometry, then the geometry of the last point calculated will be used. The total effect is to slow down the calculation, but no user action is recommended.

### **\*\* GRADIENT IS TOO LARGE TO ALLOW...**

Before a FORCE calculation can be performed the gradient norm must be so small that the third and higher order components of energy in the force field are negligible. If, in the system under examination, the gradient norm is too large, the gradient norm will first be reduced using FLEPO, unless LET has been specified. In some cases the FORCE calculation may be run only to decide if a state is a ground state or a

transition state, in which case the results have only two interpretations. Under these circumstances, LET may be warranted.

### **GRADIENT IS VERY LARGE...**

In a calculation of the thermodynamic properties of the system, if the rotation and translation vibrations are non-zero, as would be the case if the gradient norm was significant, then these “vibrations” would interfere with the low-lying genuine vibrations. The criteria for THERMO are much more stringent than for a vibrational frequency calculation, as it is the lowest few genuine vibrations that determine the internal vibrational energy, entropy, etc.

### **GRADIENTS OF OLD GEOMETRY, GNORM= nn.nnnn**

This is an error message. For the D.F.P. geometry optimization to work efficiently the gradient norm must not rise very much on making a small displacement to the geometry in the direction calculated to reduce the gradient norm. If this happens, then there is a finite chance that the D.F.P. will fail to give good results. Normally the cause of the error is a badly specified geometry, and information is printed to allow the user to correct the fault. If the user is satisfied that there is no fault, then the key-word “GEO-OK” will allow the calculation to proceed, but great care must then be taken in insuring that the results are satisfactory. In particular, verify that the gradients have been reduced to about zero by the D.F.P. at the end of the calculation.

### **ILLEGAL ATOMIC NUMBER**

An element has been specified by an atomic number which is not in the range 1 to 107. Check the data: the first datum on one of the lines is faulty.

### **IMPOSSIBLE NUMBER OF OPEN SHELL ELECTRONS**

The keyword OPEN(n1,n2) has been used, but for an even-electron system n1 was specified as odd or for an odd-electron system n1 was specified as even. Either way, there is a conflict which the user must resolve.

### **IMPOSSIBLE OPTION REQUESTED**

A general catch-all. This message will be printed if two incompatible options are used, such as both MINDO/3 and AM1 being specified. Check the key-words, and resolve the conflict.

### **INTERNAL COORDINATES READ IN, AND CALCULATION...**

If internal coordinates are read in, but the calculation is to be carried out using Cartesian coordinates, then either all possible geometric variables must be optimized, or none can be optimized. If only some are marked for optimization, then ambiguity exists. For example, if the bond-length of atom 6 is marked for optimization, but the angle is not, then when the conversion to Cartesian coordinates takes place, the first coordinate becomes the “X” coordinate and the second the “Y” coordinate. These bear no relationship to the bond length or angle. This is a fatal error.

### **INTERNAL COORDINATES READ IN, AND SYMMETRY...**

If internal coordinates are read in, but the calculation is to be carried out using Cartesian coordinates, then any symmetry relationships between the internal coordinates will not be reflected in the Cartesian coordinates. For example, if the bond-lengths of atoms 5 and 6 are equal, it does not follow that these atoms have equal values for their "X" coordinates. This is a fatal error.

### **JOB STOPPED BY OPERATOR**

Any AMPAC calculation, for which the SHUTDOWN command works, can be stopped by a user who issues the command \$SHUT <filename>, from the directory which contains <filename>.DAT

AMPAC will then stop the calculation at the first convenient point, usually after the current cycle has finished. A restart file will be written and the job ended. The message will be printed as soon as it is detected, which would be the next time the timer routine is accessed.

### **THE LINE MINIMIZATION FAILED TWICE IN A ROW**

This is usually found in more exotic calculations. For some reason, the heat of formation at the start of a line minimization within FLEPO was lower than any point on the line minimization, a situation that cannot occur naturally.

Look at the gradient, if printed. If it is acceptably small, ignore the message. The gradient will not be printed if it was very small: less than 2.0. The time to worry is when the gradient norm is larger than, say, 20: if larger than 50 something is seriously wrong, and you should then suspect AMPAC is faulty in some way.

### **\*\*\*\* MAX. NUMBER OF ATOMS ALLOWED...**

At compile time the maximum sizes of the arrays in AMPAC are fixed. The system being run exceeds the maximum number of atoms allowed. To rectify this, modify the file DIMSIZES.DAT to increase the number of heavy and light atoms allowed. If DIMSIZES.DAT is altered, then the whole of AMPAC should be re-compiled and re-linked.

### **\*\*\*\* MAX. NUMBER OF ORBITALS...**

At compile time the maximum sizes of the arrays in AMPAC are fixed. The system being run exceeds the maximum number of orbitals allowed. To rectify this, modify the file DIMSIZES.DAT to change the number of heavy and light atoms allowed. If DIMSIZES.DAT is altered, then the whole of AMPAC should be re-compiled and re-linked.

### **\*\*\*\* MAX. NUMBER OF TWO ELECTRON INTEGRALS...**

At compile time the maximum sizes of the arrays in AMPAC are fixed. The system being run exceeds the maximum number of two-electron integrals allowed. To rectify this, modify the file DIMSIZES.DAT to modify the number of heavy and light atoms allowed. If DIMSIZES.DAT is altered, then the whole of AMPAC should be re-compiled and re-linked.

**NAME NOT FOUND**

Various atomic parameters can be modified in AMPAC by use of EXTERNAL=. These comprise:

Uss	Betas	Gp2	GSD
Upp	Betap	Hsp	GPD
Udd	Betad	AM1	GDD
Zs	Gss	Expc	FN1
Zp	Gsp	Gaus	FN2
Zd	Gpp	Alp	FN3

Thus to change the Uss of hydrogen to -13.6 the line:

```
USS      H      -13.6
```

could be used. If an attempt is made to modify any other parameters, then an error message is printed, and the calculation terminated.

**NO POINT LOWER IN ENERGY...**

This is a non-fatal error message, caused by the Powell line-search routine in the D.F.P. failing to lower the energy. Two explanations are possible: first, the geometry might be almost fully optimized, in which case the smallest step possible in the line search is still too big; and second, the calculated heat of formation of a previous step was too low. This would be an algorithmic error.

This error message occurs rarely in routine use, but can easily be induced if program parameters such as the SCF criterion are altered.

**NUMBER OF PARTICLES, nn GREATER THAN...**

When user-defined microstates are not used, the MECI will calculate all possible microstates that satisfy the space and spin constraints imposed. This is done in PERM, which permutes N electrons in M levels. If N is greater than M, then no possible permutation is valid. This is not a fatal error - the program will continue to run, but no C.I. will be done.

**NUMBER OF PERMUTATIONS TOO GREAT, LIMIT 60**

The number of permutations of alpha or beta microstates is limited to 60. Thus if 3 alpha electrons are permuted among 5 M.O.s, that will generate  $10 = 5!/(3!*2!)$  alpha microstates, which is an allowed number. However if 4 alpha electrons are permuted among 8 M.O.s, then 70 alpha microstates result and the arrays defined will be insufficient. Note that 60 alpha and 60 beta microstates will permit 3600 microstates in all, which should be more than sufficient for most purposes. (An exception would be for excited radical icosohedral systems.)

**SINCE COS HAS JUST BEEN RESET...**

A non-fatal error message. After two attempts to lower the heat of formation have failed, the D.F.P. optimization is terminated. If the gradients are acceptable, no further action is necessary; otherwise examine the geometry to see if a more reasonable numbering system can be made.

A rarer cause of this message is when the geometry supplied is already almost fully optimized, and it becomes optimized within the first cycle.

#### **SYMMETRY SPECIFIED, BUT CANNOT BE USED IN DRC**

This is self-explanatory. The DRC requires all geometric constraints to be lifted. Any symmetry constraints will first be applied, to symmetrize the geometry, and then removed to allow the calculation to proceed.

#### **SYSTEM DOES NOT APPEAR TO BE OPTIMIZABLE**

This is a gradient norm minimization message. These routines will only work if the nearest minimum to the supplied geometry in gradient-norm space is a transition state or a ground state. Gradient norm space can be visualized as the space of the scalar of the derivative of the energy space with respect to geometry. To a first approximation, there are twice as many minima in gradient norm space as there are in energy space.

It is unlikely that there exists any simple way to refine a geometry that results in this message. While it is appreciated that a large amount of effort has probably already been expended in getting to this point, users should steel themselves to writing off the whole geometry. It is not recommended that a minor change be made to the geometry and the job re-submitted.

#### **TEMPERATURE RANGE STARTS TOO LOW...**

The thermodynamics calculation assumes that the statistical summations can be replaced by integrals. This assumption is only valid above 100K, so the lower temperature bound is set to 100, and the calculation continued.

#### **THERE HAVE BEEN 3 ATTEMPTS...**

In FLEPO, the energy minimization routine, three attempts have been made to lower the heat of formation, and thus the gradient, but while the heat of formation has reached a fairly constant value, the gradient norm is still large. This is a non-fatal error, and almost certainly the results will be useful. Try using "XYZ", or re-arranging the connectivity.

#### **THERE IS A RISK OF INFINITE LOOPING...**

The SCF criterion has been reset by the user, and the new value is so small that the SCF test may never be satisfied. This is a case of user beware!

#### **THIS MESSAGE SHOULD NEVER APPEAR, CONSULT A PROGRAMMER!**

This message should never appear; a fault has been introduced into AMPAC, most probably as a result of a programming error.

#### **THREE ATOMS BEING USED TO DEFINE...**

If the Cartesian coordinates of an atom depend on the dihedral angle it makes with three other atoms, and those three atoms fall in an almost straight line, then a small change in the Cartesian coordinates of one of

those three atoms can cause a large change in its position. This is a potential source of trouble, and the data should be changed to make the geometric specification of the atom in question less ambiguous.

This message can appear at any time, particularly in reaction path and saddle-point calculations.

An exception to this rule is if the three atoms fall into an exactly straight line. For example, if, in propyne, the hydrogens are defined in terms of the three carbon atoms, then no error will be flagged. In such a system the three atoms in the straight line must not have the angle between them optimized, as the finite step in the derivative calculation would displace one atom off the straight line and the error-trap would take effect.

Correction involves re-defining the connectivity. LET and GEO-OK will not allow the calculation to proceed.

### **- - - - - TIME UP - - - - -**

The time defined on the keyword line (3,600 seconds, if no time is specified) will be exceeded if another cycle is performed. A controlled termination of the run would follow this message. The job may terminate earlier than expected: this is ordinarily due to one of the recently completed cycles taking unusually long, and the safety margin has been increased to allow for the possibility that the next cycle might also run for much longer than expected.

### **TRIPLET SPECIFIED WITH ODD NUMBER OF ELECTRONS**

If TRIPLET has been specified the number of electrons must be even. Check the charge on the system, the empirical formula, and whether TRIPLET was intended.

### **""""""""""UNABLE TO ACHIEVE SELF-CONSISTENCY**

See the error-message:

<<<<---\*\*\* FAILED TO ACHIEVE SCF. \*\*\*--->>>>

### **UNDEFINED SYMMETRY FUNCTION USED**

Symmetry operations are restricted to those defined, i.e. in the range 1-18. Any other symmetry operations will trip this fatal message.

### **UNRECOGNISED ELEMENT NAME**

In the geometric specification a chemical symbol which does not correspond to any known element has been used. The error lies in the first datum on a line of geometric data.

### **\*\*\* WARNING \*\*\***

Don't pay too much attention to this message. Thermodynamics calculations require a higher precision than vibrational frequency calculations. In particular, the gradient norm should be very small. However, it is frequently not practical to reduce the gradient norm further, and to date no one has determined just how slack the gradient criterion can be before unacceptable errors appear in the thermodynamic quantities. The 0.4 gradient norm is only a suggestion.

**WARNING: INTERNAL COORDINATES...**

Triatomics are, by definition, defined in terms of internal coordinates. This warning is only a reminder. For diatomics, Cartesian and internal coordinates are the same. For tetra-atomics and higher, the presence or absence of a connectivity table distinguishes internal and Cartesian coordinates, but for triatomics there is an ambiguity. To resolve this, Cartesian coordinates are not allowed for the data input for triatomics.



## APPENDIX B: REFERENCES

### MNDO:

“Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters.”, M.J.S. Dewar, W. Thiel, *J. Am. Chem. Soc.* **99**, 4899 (1977).

### Elements Parametrized in MNDO:

Al	L.P. Davis, R.M. Guidry, J.R. Williams, M.J.S. Dewar, H.S. Rzepa, <i>J. Comp. Chem.</i> <b>2</b> , 433, (1981)
B	M.J.S. Dewar, M.L. McKee, <i>J. Am. Chem. Soc.</i> <b>99</b> , 5231 (1977).
Be	M.J.S. Dewar, H.S. Rzepa, <i>J. Am. Chem. Soc.</i> <b>100</b> , 777 (1978).
Br	M.J.S. Dewar, E.F. Healy, <i>J. Comp. Chem.</i> <b>4</b> , 542 (1983).
C	M.J.S. Dewar, W. Thiel, <i>J. Am. Chem. Soc.</i> <b>99</b> , 4907 (1977).
Cl	(a) M.J.S. Dewar, M.L. McKee, H.S. Rzepa, <i>J. Am. Chem. Soc.</i> <b>100</b> , 3607 (1978). (b) M.J.S. Dewar, H.S. Rzepa, <i>J. Comp. Chem.</i> <b>4</b> , 158 (1983).
F	M.J.S. Dewar, H.S. Rzepa, <i>J. Am. Chem. Soc.</i> <b>100</b> , 58 (1978).
Ge	M.J.S. Dewar, E.F. Healy, G.L. Grady, <i>Organometallics</i> <b>6</b> , 186 (1987).
H	M.J.S. Dewar, W. Thiel, <i>J. Am. Chem. Soc.</i> <b>99</b> , 4907 (1977).
Hg	M.J.S. Dewar, G.L. Grady, K.M. Merz, J.J.P. Stewart, <i>Organometallics</i> <b>4</b> , 1964 (1985).
I	M.J.S. Dewar, E.F. Healy, J.J.P. Stewart, <i>J. Comp. Chem.</i> <b>5</b> , 358 (1984).
N	M.J.S. Dewar, W. Thiel, <i>J. Am. Chem. Soc.</i> <b>99</b> , 4907 (1977).
O	M.J.S. Dewar, W. Thiel, <i>J. Am. Chem. Soc.</i> <b>99</b> , 4907 (1977).
P	M.J.S. Dewar, M.L. McKee, H.S. Rzepa, <i>J. Am. Chem. Soc.</i> <b>100</b> , 3607 (1978).
Pb	M.J.S. Dewar, M.K. Holloway, G.L. Grady, J.J.P. Stewart, <i>Organometallics</i> <b>4</b> , 1973 (1985).
S	M.J.S. Dewar, C. H. Reynolds, <i>J. Comp. Chem.</i> <b>7</b> , 140 (1986).
Si	M.J.S. Dewar, E.F. Healy, J.J.P. Stewart, J.E. Friedheim, G.L. Grady, <i>Organometallics</i> <b>5</b> , 375 (1986).
Sn	M.J.S. Dewar, G.L. Grady, J.J.P. Stewart, <i>J. Am. Chem. Soc.</i> <b>106</b> , 6771 (1984).
Zn	M.J.S. Dewar, K.M. Merz, <i>Organometallics</i> <b>5</b> , 1494 (1986).

### MINDO/3:

“Ground States of Molecules. 25. MINDO/3. An Improved Version of the MINDO Semiempirical SCF MO Method.”, M.J.S. Dewar, R.C. Bingham, D.H. Lo, *J. Am. Chem. Soc.* **97**, 1285 (1975).

### AM1:

“AM1: A New General Purpose Quantum Mechanical Molecular Model”, M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, *J. Am. Chem. Soc.* **107**, 3902-3909 (1985).

### Elements Parametrized in AM1:

Al	M.J.S. Dewar, A.J. Holder, to be published.
B	M.J.S. Dewar, C. Jie, E.G. Zoebisch, <i>Organometallics</i> <b>7</b> , 513 (1988).
Br	M.J.S. Dewar, E.G. Zoebisch, <i>J. Mol. Struct. (Theochem)</i> <b>180</b> , 1 (1988).
C	M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, <i>J. Am. Chem. Soc.</i> <b>107</b> , 3902 (1985).
Cl	M.J.S. Dewar, E.G. Zoebisch, <i>J. Mol. Struct. (Theochem)</i> <b>180</b> , 1 (1988).
F	M.J.S. Dewar, E.G. Zoebisch, <i>J. Mol. Struct. (Theochem)</i> <b>180</b> , 1 (1988).
Ge	M.J.S. Dewar, C. Jie, <i>Organometallics</i> <b>8</b> , 1544 (1989).
Ge	M.J.S. Dewar, C. Jie, <i>Organometallics</i> <b>8</b> , 1547 (1989).
H	M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, <i>J. Am. Chem. Soc.</i> <b>107</b> , 3902 (1985).
I	M.J.S. Dewar, E.G. Zoebisch, <i>J. Mol. Struct. (Theochem)</i> <b>180</b> , 1 (1988).
N	M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, <i>J. Am. Chem. Soc.</i> <b>107</b> , 3902 (1985).
O	M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, <i>J. Am. Chem. Soc.</i> <b>107</b> , 3902 (1985).
P	M.J.S. Dewar, C. Jie, <i>J. Mol. Struct. (Theochem)</i> <b>187</b> , 1 (1989).

S	M.J.S. Dewar, Y-C Yuan, <i>J. Inorg. Chem.</i> , submitted for publication.
Si	M.J.S. Dewar, C. Jie, <i>Organometallics</i> <b>6</b> , 1486 (1987).
Sn	M.J.S. Dewar, D.R. Kuhn, E.F. Healy, to be published.
Zn	M.J.S. Dewar, K.M. Merz, <i>Organometallics</i> <b>7</b> , 522 (1988).

**Bonds:**

“Bond Indices and Valency”, D.R. Armstrong, P.G. Perkins, J.J.P. Stewart, *J. Chem. Soc., Dalton*, 838 (1973).

**C.I. Gradients:**

D. Liotard, M.J.S. Dewar, *J. Mol. Struct. (Theochem)*, accepted for publication June, 1989.

**CHAIN Method:**

D. Liotard, Thesis, University of Pau, France (1979).

**Davidon-Fletcher-Powell Method:**

R. Fletcher, M.J.D. Powell, *Computer Journal* **6**, 163 (1963); W.C. Davidon, *ibid.* **10**, 406 (1968).

**Diagonalization:**

Y. Beppu, *Computers and Chemistry* **6**, 87 (1982).

**Half-Electron:**

“Ground States of Conjugated Molecules. IX. Hydrocarbon Radicals and Radical Ions”, M.J.S. Dewar, J.A. Hashmall, C.G. Venier, *J. Am. Chem. Soc.* **90**, 1953 (1968).

“Triplet States of Aromatic Hydrocarbons”, M.J.S. Dewar, N. Trinajstić, *Chem. Comm.*, 646 (1970).

“Semiempirical SCF-MO Treatment of Excited States of Aromatic Compounds” M.J.S. Dewar, N. Trinajstić, *J. Chem. Soc. (A)*, 1220 (1971).

**Localization:**

“A New Rapid Method for Orbital Localization.” P.G. Perkins, J.J.P. Stewart, *J. Chem. Soc., Faraday Trans. 2* **78**, 285 (1982).

**Locating Transition States:**

“Location of Transition States in Reaction Mechanisms”, M.J.S. Dewar, E.F. Healy, J.J.P. Stewart, *J. Chem. Soc., Faraday Trans. 2* **3**, 227 (1984).

**LTRD and NEWTON:**

D. Liotard, Thesis, University of Pau, France (1979).

D. Poppinger, *Chem. Phys. Letters* **35**, 550 (1975).

**MECI:**

“Molecular Orbital Theory for the Excited States of Transition Metal Complexes”, D.R. Armstrong, R. Fortune, P.G. Perkins, J.J.P. Stewart, *J. Chem. Soc., Faraday Trans. 2* **68**, 1839-1846 (1972).

**Polarizability:**

“A New Procedure for Calculating Molecular Polarizabilities: Applications Using MNDO.” M.J.S. Dewar, J.J.P. Stewart, *Chem. Phys. Lett.* **111**, 416 (1984).

**Powell algorithm:**

M. J. D. POWELL, 'Numerical Methods for Nonlinear Algebraic Equations', Ed. by Rabinowitz, Gordon & Breach: New York, Ch. 6 & 7 (1970).

**Pseudodiagonalization:**

“Fast Semiempirical Calculations”, J.J.P. Stewart, P. Csaszar, P. Pulay, *J. Comp. Chem.* **3**, 227 (1982).

**Pulay's Converger:**

“Convergence Acceleration of Iterative Sequences. The Case of SCF Iteration”, P. Pulay, *Chem. Phys. Lett.* **73**, 393 (1980).

**Reaction path definition and integration:**

K. Fukui, *J. Phys. Chem.* **74**, 4161 (1970).

A. Tachibana, K. Fukui, *Theoret. Chim. Acta.* **49**, 321 (1978).

K. Ishida, K. Morokuma, A. Komornicki, *J. Chem. Phys.* **66**, 2153 (1977).

A. Ralston, A. Certaine, 'Mathematical Methods for Digital Computer', ed. by A. Ralston, J. Wiley & Sons: New York (1960).

D. Liotard, to be published.

**Shift:**

“Unconditional Convergence in SCF Theory: A General Level Shift Technique”, R. Carbo, J.A. Hernandez, F. Sanz, *Chem. Phys. Lett.* **47**, 581 (1977).

**SIGMA Method:**

A. Komornicki, J.W. McIver, *Chem. Phys. Lett.* **10**, 303 (1971).

A. Komornicki, J.W. McIver, *J. Am. Chem. Soc.* **94**, 2625 (1972).

**Thermodynamics:**

“Ground States of Molecules. 44. MINDO/3 Calculations of Absolute Heat Capacities and Entropies of Molecules without Internal Rotations.” M.J.S. Dewar, G.P. Ford, *J. Am. Chem. Soc.* **99**, 7822 (1977).



## INDEX

- 0SCF 7
- 1ELEC 76
- 1ELECTRON 7
- 1SCF 7
- alkali metal 30
- AM1 8
- AMPAC.COM 75
- AMPAC.EXE 74
- ANALYTICAL GRADIENTS 72
- BAR=nn 8
- Bartel's method 17
- BFGS algorithm 54
- BIRADICAL 8
- BLOCK DATA file 11
- bond order 8
- BONDS 8
- C.I.
  - Formation of Microstates 58
  - Keywords 57
  - Matrix elements 60
  - reference configuration 57
  - Removal of Electrons 58
  - Spin States 62
- C.I., increasing size 9
- C.I.=n 9
- Camp-King converger 54
- CAPABILITIES of AMPAC 1
- Cartesian coordinates 28
- Cartesian derivatives 10
- CHAIN 9, 71
- characterizing a transition state 13
- CHARGE=n 9
- COMPFG 76
- COMPILE.COM 74
- compiling AMPAC 74
- configuration interaction 9, 57
- contour maps 14
- Convergence Techniques 53
- counterions 68
- CYCLES=n 10
- D2d Ethylene, example data file 2
- data file, structure of 2
- DCART 10, 76
- DEBUG 10, 76
- DEBUG keywords 76
- DENMAT 10
- DENOUT 10
- DENSITY 10, 76
- density matrix 10
- DEP 11
- DEPVAR=n.nn 11
- DERINU 11
- DERIV 76
- Derived parameters 13
- DFORCE 76
- DIAG 56
- diastereoisomers 55
- Diatomics Parametrized in MINDO/3 29
- dihedral angle sign convention 55
- DIMSIZES.DAT 74
- DOUBLET 11
- DRC 11
- DRC=n.nn 11
- dummy atoms 30
- Dynamic Reaction Coordinate 11, 64
- ECHO 12
- eigenvectors 26
- EIGS 76
- Elements parametrized 29
- Energy level shift technique 54
- ENPART 12
- entropy 24
- ERROR MESSAGES 93
- ESR 12
- EXCITED 12
- EXTERNAL= 12
- EYEAD 55
- FILL=n 13
- FLEPO 76
- FLOW DIAGRAM 81
- FOCK 76
- FORCE 13
- force calculation
  - reduced mass 63
  - trivial vibrations 63
  - vibrational analysis 63
- force constants 13
- force matrix 13
- FORTTRAN-77 FILES 82
- free format 26, 27
- full SCF calculations 14
- FULSCF 14
- GEO-OK 14
- GNORM= 14
- gradient components 14
- Gradient computation 78
- Gradient minimization
  - by NLLSQ 17
  - by POWELL 19
  - by SIGMA 22
- gradient norm 14, 79
- GRADIENTS 14
- GRAPH 14
- grid calculation 22
- H-PRIORITY 15
- H-PRIORITY=n.nn 15
- HCORE 76
- heat capacity 24

HEAT OF FORMATION TEST SATISFIED 55  
 HERBERTS TEST SATISFIED 55  
 Hessian  
     accuracy 80  
 Hessian matrix 13, 16  
 hybridization 18  
 Installation of AMPAC 73  
 internal coordinates 27  
 internal energy 24  
 ISOTOPE 15  
 isotopes 30  
 isotopic substitution 15  
 ITER 76  
 ITRY=nn 15  
 KINETIC=n.nnn 15  
 LET 15  
 LINMIN 76  
 LIST OF KEYWORDS 5  
 LOCALIZE 15  
 Locating Bugs 77  
 LOCMIN 76  
 LTRD 16  
 MAXHEV 74  
 MAXLIT 74  
 McIver-Komornicki minimization 22  
 MECI 16, 57, 77  
 MICRO=n 16  
 microstate format 16  
 MINDO/3 17  
 minimization of the gradient norm 19  
 MNRSD1.DAT 33  
 MOLDAT 76  
 MULLIK 17  
 Mulliken-type Population analysis 17  
 NCHAIN 74  
 new algorithms 79  
 New KEYWORDS 80  
 NEWTON 17  
 NLLSQ 17  
 NMECI 74  
 Normal Mode analysis 44  
 NRELAX 74  
 numerical energy derivatives 11  
 OLDENS 18  
 OLDMAT 18  
 one-electron matrix 7  
 OPEN(n1,n2) 18  
 Optimization criteria 79  
 Optimization of the code 79  
 Oscillation damping 53  
 parameter data file 13  
 partition function 24  
 PATH 18  
 PETERS TEST SATISFIED 55  
 PI 18  
 PL 77  
 PLTEST 53  
 polarizability 65  
 Polytetrahydrofuran, example data file 3  
 POWELL 19  
 POWSQ 19  
 PRECISE 19, 79  
 PRINT=n 19  
 PSEUDODIAGONALIZATION 56  
 PULAY 19  
 Pulay's method 54  
 QUARTET 19  
 QUINTET 20  
 RAMPAC.COM 75  
 reaction coordinates 45  
 reaction path 69  
 RESTART 20  
 ROOT=n 20  
 ROT=n 20  
 Russel-Saunders coupling 17  
 SADDLE 21, 70  
     Use of BAR= 8  
 SCF convergence criteria 53, 78  
 SCFCRT 53  
 SCFCRT=n.nnn 21  
 SEARCH 77  
 SELCON 53  
 SEXTET 21  
 SHIFT=n.n 21  
 SHUTDOWN 76  
 SIGMA 22  
 SINGLET 22  
 SIZES 74  
 Solid State method 67  
 solvation effects 68  
 sparkles 30, 68  
 SPIN 22  
 spin density 22  
 spin manifolds 16  
 spin matrix 22  
 spin states  
     DOUBLET 11  
     EXCITED 12  
     QUARTET 19  
     QUINTET 20  
     SEXTET 21  
     SINGLET 22  
     TRIPLET 25  
 STEP1=n.nnn 22  
 STEP2=n.nnn 22  
 Subroutine  
     AABABC 85  
     AABACD 85  
     AABBCD 85  
     ANAVIB 85  
     AXIS 85  
     BABBBC 86  
     BABBCD 86  
     BONDS 86  
     CALPAR 86  
     CHARGE 86

CNVG 86  
 COMPG 86  
 DENROT 86  
 DENSIT 86  
 DEPVAR 86  
 DERIV 86  
 DFPSAV 86  
 DIAG 87  
 DIAGI 87  
 DIAT 87  
 DIAT2 87  
 DIPOLE 87  
 DOT 87  
 DRC 87  
 ENPART 87  
 EXCHNG 87  
 FLEPO 87  
 FMAT 87  
 FOCK1 87  
 FOCK2 87  
 FOCK2D 88  
 FORCE 88  
 FORSAV 88  
 FRAME 88  
 FREQCY 88  
 GEOUT 88  
 GETGEO 88  
 GETSYM 88  
 GMETRY 88  
 H1ELEC 88  
 HADDON 88  
 HCORE 88  
 HELECT 88  
 HQRII 88  
 IJKL 89  
 INTERP 89  
 ITER 89  
 LINMIN 89  
 LOCAL 89  
 LOCMIN 89  
 MAIN 89  
 MAMULT 89  
 MATOUT 89  
 MECI 89  
 MOLDAT 89  
 MULLIK 89  
 MULT 89  
 NLLSQ 89  
 OSINV 89  
 PARSAV 90  
 PATHS 90  
 PERM 90  
 POLAR 90  
 POWSAV 90  
 POWSQ 90  
 PULAY 90  
 REACT1 90  
 READ 90  
 READA 90  
 REPP 90  
 ROTATE 90  
 SEARCH 90  
 SECOND 90  
 SOLROT 91  
 SWAP 91  
 SYMTRY 91  
 THERMO 91  
 VECPR 91  
 WRITE 91  
 XYZINT 91  
 SUBROUTINE CALLS 82  
 SUBROUTINES 85  
 SYMMETRY 22  
 T-PRIORITY 24  
 T-PRIORITY=n.nn 24  
 T.V. 24  
 T=nnn 24  
 TEST ON GRADIENT SATISFIED 55  
 TEST ON X SATISFIED 55  
 THERMO 24  
 THERMO(nnn) 25  
 THERMO(nnn,mmm) 25  
 THERMO(nnn,mmm,ooo) 25  
 thermodynamic quantities 24  
 Three-point interpolation of the density matrix 53  
 TIMES 77  
 TOLERF 55  
 TOLERG 55  
 TOLERX 55  
 TRANS 25  
 TRANS=n 25  
 transition state locating  
     by the CHAIN method 9  
     by the SADDLE method 21  
 transition vector 18, 24  
 translation vector  
     DEPVAR=n.nn 11  
 triatomics 29  
 TRIPLET 25  
 UHF 26  
 unpaired spin density 12  
 Updating Parameters 12  
     Using DEP 11  
 VECTORS 26, 77  
 velocity vector 15  
 vibrational frequencies 13  
 WEIGHT 26  
 X-PRIORITY 26  
 X-PRIORITY=n.nn 26  
 XX 30  
 XYZ 26, 29