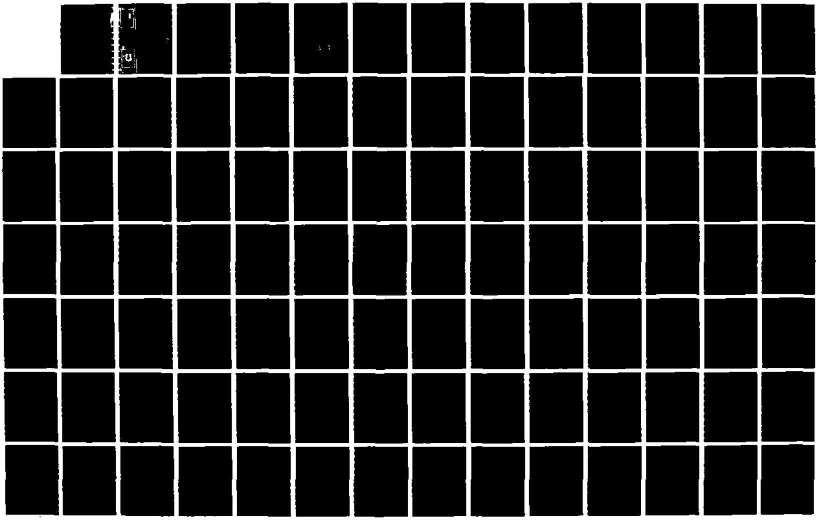
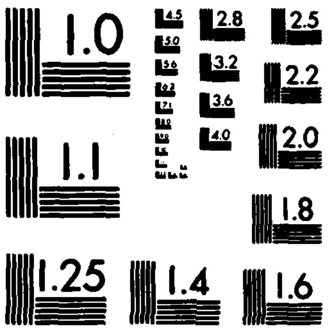
AD-A172 100 MOPAC MANUAL (VERSION 310) A GENERAL MOLECULAR ORBITAL 1/2
PACKAGE -- CRAY-XMP VERSION(U) FRANK J SEILER RESEARCH
LAB UNITED STATES AIR FORCE ACADEMY C.. J J STEWART
UNCLASSIFIED JUN 86 FJSRL-TR-86-0008 F/G 7/4 NL




1.0
4.5
5.0
5.6
6.3
7.1
8.0
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

①

FJSRL-TR-86-0008

FRANK J. SEILER RESEARCH LABORATORY

MOPAC MANUAL (VERSION 3.10) A GENERAL MOLECULAR ORBITAL PACKAGE -- CRAY-XMP VERSION

James J. P. Stewart

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

June 1986
AIR FORCE SYSTEMS COMMAND

UNITED STATES AIR FORCE

FJSRL-TR-86-0008

This document was prepared by the Energetic Materials Division, Directorate of Chemical Sciences, Frank J. Seiler Research Laboratory, United States Air Force Academy, CO. The research was conducted under Project Work Unit number 2303-F3-07. Jon T. Swanson was the project scientist.

Inquiries concerning the technical content of this document should be addressed to the Frank J. Seiler Research Laboratory (AFSC), FJSRL/NC, USAF Academy, Colorado Springs, CO 80840. Phone AC 303-472-2655.

This report has been reviewed by the Commander and is releasable to the National Technical Information Service (NTIS). At NTIS it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Jon T Swanson
JON T. SWANSON, Capt, USAF
Project Scientist

CHESTER J. DYMEK, JR., Lt Col, USAF
Director, Chemical Sciences

Kenneth E Siegenthaler
KENNETH E. SIEGENTHALER, Lt Col, USAF
Chief Scientist

Printed in the United States of America. Qualified requestors may obtain additional copies from the Defense Documentation Center. All others should apply to:

National Technical Information Service
6285 Port Royal Road
Springfield VA 22161

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

AD-A172100

REPORT DOCUMENTATION PAGE

ADA

1a. REPORT SECURITY CLASSIFICATION Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S) FJSRL-TR-86-0008

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION Frank J. Seiler Research Laboratory

6b. OFFICE SYMBOL (If applicable) FJSRL/NC

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS (City, State and ZIP Code)

USAF Academy
Colorado Springs CO 80840-6528

7b. ADDRESS (City, State and ZIP Code)

8a. NAME OF FUNDING/SPONSORING ORGANIZATION AF Office of Scientific Research

8b. OFFICE SYMBOL (If applicable) AFOSR

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State and ZIP Code)

Bldg. 410
Bolling AFB DC 20332

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2303	F3	07

11. TITLE (Include Security Classification) MOPAC Manual (Vers. 3.10) A General Molecular Orbital Package - Cray-XMP Version

12. PERSONAL AUTHOR(S) James J. P. Stewart

13a. TYPE OF REPORT Tech Report

13b. TIME COVERED FROM ________ TO Jun 86

14. DATE OF REPORT (Yr., Mo., Day) 86 Jun

15. PAGE COUNT 147

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.
0704	1901	2102

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) MOPAC, MNDO, AM1, Semi-empirical calculations

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

MOPAC is a general-purpose, semi-empirical molecular orbital program for the study of chemical reactions involving molecules, ions, and linear polymers. It implements the semi-empirical Hamiltonians MNDO, AM1 and MINDO/3 and combines the calculations of vibrational spectra, thermodynamic quantities, isotopic substitution effects, and force constants in a fully integrated program. Elements parametrized at the MNDO level include H, Li, Be, B, C, N, O, F, Al, Si, P, S, Cl, Cr, Ge, Br, Sn, Hg, Pb, and I. Within the electronic part of the calculation, molecular and localized orbitals excited states up to sextets, chemical bond indices, charges, etc. are computed. Both intrinsic and dynamic reaction coordinates can be calculated. A transition-state location routine and two transition-state optimizing routines are available for studying chemical reactions.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT

UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL Jon T. Swanson

22b. TELEPHONE NUMBER (Include Area Code) (303) 472-2655

22c. OFFICE SYMBOL FJSRL/NC

MOPAC MANUAL

(Version 3.10)

A GENERAL MOLECULAR ORBITAL PACKAGE

CRAY-XMP Version

Written by
James J.P. Stewart,
Frank J. Seiler Research Laboratory
United States Air Force Academy
Colorado Springs, CO 80840

CONTENTS

CHAPTER I PREFACE TO FIRST EDITION, VERSION 3.10

CHAPTER 1 INTRODUCTION

CHAPTER 2 KEY-WORDS

CHAPTER 3 GEOMETRY SPECIFICATION

CHAPTER 4 EXAMPLES

CHAPTER 5 TESTDATA

CHAPTER 6 BACKGROUND

Acknowledgements

The initial writing of MOPAC took about six months, with the current version taking several more months, spread over three years. During this time several co-workers provided invaluable assistance. Some contributed code, some ideas, and some identified bugs. Of those who helped, I would like to recognise the following people for their assistance during the writing of MOPAC.

For his patience and considerable technical expertise, I am indebted to Dr. Santiago Olivella. Dr. Olivella wrote ENPART, and read the entire program, identifying a large number of coding and theoretical errors.

Prof. Peter Pulay developed the theory used in DIAG, and was the driving force behind its development. He also developed the theory behind, and helped in the writing of, the sophisticated SCF converger PULAY.

Major Donn Storch, at the Air Force Academy, has been involved during the entire development of MOPAC, taking a professional interest in its design and structure. Many improvements, normally unseen in routine use, are due to his practical suggestions.

No large program is bug-free, and discovering bugs is very demanding. This arduous task was undertaken enthusiastically by Mr. James Friedheim, who identified what felt like hundreds of bugs.

For advice and encouragement I thank Dr. Eamon Healy. With his expert knowledge of quantum mechanics, Dr. Healy was able to check the soundness of each new option provided.

For describing the FRAME device for separating translations and rotations in normal coordinate analysis, and for providing code for the Camp-King converger, I thank Dr. John McKelvey, of Eastman-Kodak. He also made available a program to format the FORTRAN code to facilitate reading and debugging.

For providing professional assistance with the development of MOPAC on the CRAY computers, I thank Dr. Erich Wimmer of Cray Research Inc.

For her unflagging patience in checking the manual for clarity of expression, and for drawing to my attention innumerable spelling and grammatical errors, I thank my wife, Anna.

Finally, I wish to thank Prof. Michael J.S. Dewar for providing the facilities and funds during the initial development, the staff of the Seiler Research Laboratory for their support, and the National Research Council for the award of an Associateship to continue the development of MOPAC.

CHAPTER 7 PROGRAM

CHAPTER 8 ERROR MESSAGES PRODUCED BY MOPAC

CHAPTER 9 CRITERIA

CHAPTER 10 DEBUGGING

CHAPTER 11 INSTALLING MOPAC

APPENDIX A FORTRAN FILES

APPENDIX B SUBROUTINE CALLS IN MOPAC

APPENDIX C DESCRIPTION OF SUBROUTINES IN MOPAC

APPENDIX D HEATS OF FORMATION OF SOME MNDO AND AM1 COMPOUNDS

APPENDIX E REFERENCES

CHAPTER I

PREFACE TO FIRST EDITION, VERSION 3.10

Although this is the first CRAY version of MOPAC distributed by QCPE, it has been given the version number 3.1 in order to allow comparison with equivalent MOPAC programs on other machines. All options described in the manual should function as described.

As the program and manual were initially written for a VAX computer, there will be minor differences between the output given here and that obtained using a CRAY supercomputer. These differences are not significant, reflecting only different machine precisions.

The CRAY version of MOPAC has been tested on systems of 65 heavy atoms and 62 light atoms, and shown to work correctly for geometry optimization and reaction path work. Due to the long times required for such calculations - one SCF requires 58 seconds of CRAY time - such testing of the other options on large systems has not been done. However, testing of most other options has been done on smaller systems.

MOPAC is the core program of a series of programs for the theoretical study of chemical phenomena. This version is the third in an on-going development, and efforts are being made to continue its further evolution. In order to make the use of MOPAC easier, four other programs have also been written. Users of MOPAC are recommended to use all four programs. Efforts will be made to continue the development of these programs.

Recommended Programs

DRAW

DRAW, written by Maj. Donn Storch, is a powerful editing program specifically written to interface with MOPAC. Among the various facilities it offers are:

1. The on-line editing and analysis of a data file, starting from scratch or from an existing data file, an archive file, or from a results file.

2. The option of continuous graphical representation of the system being studied. Several types of terminal are supported, including DIGITAL, HEATHKIT, TEKTRONIX, and TERAK terminals.

3. The drawing of electron density contour maps generated by DENSITY on graphical devices.

4. The drawing of solid-state band structures generated by MOSOL.

5. The sketching of molecular vibrations, generated by a normal coordinate analysis.

DENSITY

DENSITY, written by Dr. James J. P. Stewart, is an electron-density plotting program. It accepts data-files directly from MOPAC, and is intended to be used for the graphical representation of electron density distribution, individual M.O.s and difference maps.

MOHELP

MOHELP is an on-line help facility, written by Maj. Donn Storch and Dr. James J. P. Stewart, to allow non-VAX users access to the VAX HELP libraries for MOPAC, DRAW, and DENSITY.

MOSOL

MOSOL is a full solid-state MNDO program written by Dr. James J. P. Stewart. In comparison with MOPAC, MOSOL is extremely slow. As a result, while geometry optimization, force constants, and other functions can be carried out by MOSOL, these slow calculations are best done using the solid-state facility within MOPAC. MOSOL should only be used for generating band-structures and densities of states, a task that MOPAC cannot perform.

IMPORTANT CHANGES FROM EARLIER VERSIONS OF MOPAC

Users must be aware of two important modifications. These are

(1) In FORCE calculations, atomic weights, not isotopic masses, are used. This change is justified by the fact that most experimental measurements are made using the natural abundance isotopes. The option of using specific isotopes is provided.

(2) In the FORCE calculation, if the geometry is not already optimized, then, by default, the Davidon-Fletcher-Powell optimization will be used to optimize it, instead of Bartell's gradient minimization. The option of using Bartell's method is provided.

CHAPTER 1

INTRODUCTION

MOPAC is a general-purpose semi-empirical molecular orbital package for the study of chemical reactions. The semi-empirical Hamiltonians MNDO, MINDO/3, 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, a transition-state location routine and two 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 MOPAC 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, MOPAC 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.

NOTES (1) This is the "third edition". MOPAC has undergone a considerable expansion since the second version, and users of the first or second editions are recommended to familiarize themselves with the changes which are described in this manual. If any errors are found, or if MOPAC does not perform as described, please contact Dr. James J. P. Stewart, Frank J. Seiler Research Laboratory, U.S Air Force Academy, Colorado Springs, CO 80840. (2) MOPAC runs sucessfully on normal CDC, Data General, Gould, and Digital computers, and also on the CDC 205 "supercomputer".

1.1 SUMMARY OF MOPAC CAPABILITIES

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

First of all, there are at most four possible types of data-files for MOPAC, but the simplest data-file is the most commonly used. Rather than define it, an example is shown below.

1.2.1 Example Of Data For Ethylene

```
Line  1 :    UHF PULAY MINDO3 VECTORS DENSITY LOCAL T=300
Line  2 :      EXAMPLE OF DATA FOR MOPAC
Line  3 :        MINDO/3 UHF CLOSED-SHELL D2D ETHYLENE
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 :
```

As can be seen, the first three lines are textual. The first line consists of key-words (here seven key-words 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.

These three lines are obligatory. If no name or comment is wanted, leave blank lines. If no key-words are specified, leave a blank line. A common error is to have a blank line before the key-word 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 decided, the three lines, blank or otherwise, are obligatory.

The next set of lines defines the geometry. In the example, the numbers are all neatly lined up; this is not necessary, but does make it easier when looking 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, for example symmetry data, would follow line 5. An explanation of the geometry definition is given in the chapter "GEOMETRY SPECIFICATION".

Summarizing, then, the structure for a MOPAC data-file is:

Line 1: Key-Words. (See chapter 2 on definitions of key-words)
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 to terminate the geometry definition.

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

(a) Symmetry data: This follows the geometric data, and is ended by a blank line.

(b) Reaction path: After all geometry and symmetry data (if any) are read in, points on the reaction coordinate are defined.

(c) Saddle data: A complete second geometry is input. The second geometry follows the first geometry and symmetry data (if any)

1.2.2 Example Of Data 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 4A:  H    1.113970 1   90.374545 1  126.799259 1  10 12 11
Line 4B:  H    1.114347 1   90.255788 1 -126.709810 1  10 12 11
Line 4C: Tv   12.299490 1    0.000000 0    0.000000 0   1 11 10
Line 5 :  0    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 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 4B use these dummy atoms, as does the translation vector on line 4C.

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

CHAPTER 2

KEY-WORDS

2.1 SPECIFICATION OF KEY-WORDS

All control data are entered in the form of key-words, which form the first line of a data-file. The order in which key-words appear is not important although they must be separated by a space. Some key-words can be abbreviated (for example 1ELECTRON can be entered as 1ELECT) but in general the full key-word is preferred, primarily in order to more clearly document the calculation, and secondarily to obviate the possibility that an abbreviated key-word might not be recognized. If there is insufficient space in the first line for all the key-words needed, then consider abbreviating the longer words. One type of key-word, those ending in an equal sign, such as, BAR=0.05, may not be abbreviated, and the full word needs to be supplied.

If two key-words which are incompatible, like UHF and C.I., are supplied, or a key-word which is incompatible with the species supplied, for instance TRIPLET and a methyl radical, then error trapping will normally occur, and an error message printed. This usually takes an insignificant time, so data are quickly checked for obvious errors.

2.2 FULL LIST OF KEY-WORDS USED IN MOPAC

1ELECTRON- FINAL ONE-ELECTRON MATRIX TO BE PRINTED
0SCF - READ IN DATA, THEN STOP
1SCF - DO ONE SCF AND THEN STOP
AM1 - THE AM1 HAMILTONIAN TO BE USED
BAR=n.n - REDUCE BAR LENGTH BY A MAXIMUM OF n.n
BIRADICAL- SYSTEM HAS TWO UNPAIRED ELECTRONS
BONDS - FINAL BOND-ORDER MATRIX TO BE PRINTED
C.I. - A MULTI-ELECTRON CONFIGURATION INTERACTION SPECIFIED
CHARGE=n - CHARGE ON SYSTEM = n (e.g. NH4 => CHARGE=1)
COMPFG - PRINT HEAT OF FORMATION CALCULATED IN COMPFG
CYCLES - PERFORM MAXIMUM NUMBER OF CYCLES IN NLLSQ
DCART - PRINT DETAILS OF WORKING IN DCART
DEBUG - DEBUG OPTION TURNED ON
DEBUGPULAY PRINT DETAILS OF WORKING IN PULAY
DENOUT - DENSITY MATRIX OUTPUT (CHANNEL 10)
DENSITY - FINAL DENSITY MATRIX TO BE PRINTED
DEP - GENERATE FORTRAN CODE FOR PARAMETERS FOR NEW ELEMENTS
DEPVAR=n - TRANSLATION VECTOR IS A MULTIPLE OF BOND-LENGTH
DERIV - PRINT PART OF WORKING IN DERIV
DFORCE - FORCE CALCULATION SPECIFIED, ALSO PRINT FORCE MATRIX.
DOUBLET - RHF DOUBLET STATE REQUIRED
DRC - DYNAMIC REACTION COORDINATE CALCULATION
ECHO - DATA ARE ECHOED BACK BEFORE CALCULATION STARTS
EIGS - ALL EIGENVALUES IN ITER TO BE PRINTED
ENPART - ENERGY TO BE PARTITIONED INTO COMPONENTS
ESR - RHF UNPAIRED SPIN DENSITY TO BE CALCULATED
EXCITED - FIRST EXCITED SINGLET STATE IS TO BE OPTIMIZED
EXTERNAL - MNDO OR AM1 PARAMETERS TO BE READ OFF DISK
FILL=n - IN RHF OPEN AND CLOSED SHELL, FORCE M.O. n TO BE FILLED
FLEPO - PRINT DETAILS OF GEOMETRY OPTIMIZATION
FMAT - PRINT DETAILS OF WORKING IN FMAT
FOCK - LAST FOCK MATRIX TO BE PRINTED
FORCE - FORCE CALCULATION SPECIFIED
FULSCF - FULL SCF CALCN'S TO BE DONE IN SEARCHES
GEO-OK - OVERRIDE INTERATOMIC DISTANCE CHECK
GNORM=n.n- FLEPO EXIT WHEN GRADIENT NORM BELOW n.n
GRADIENTS- ALL GRADIENTS TO BE PRINTED
GRAPH - GENERATE FILE FOR GRAPHICS
HCORE - PRINT DETAILS OF WORKING IN HCORE
H-PRIO - HEAT OF FORMATION TAKES PRIORITY IN DRC
IRC - INTRINSIC REACTION COORDINATE CALCULATION
ISOTOPE - FORCE MATRIX WRITTEN TO DISK (CHANNEL 9)
ITER - PRINT DETAILS OF WORKING IN ITER
ITRY=N - SET LIMIT OF NUMBER OF SCF ITERATIONS TO N.
KINETIC - EXCESS KINETIC ENERGY ADDED TO DRC CALCULATION
LARGE - EXPANDED OUTPUT TO BE PRINTED
LET - DO NOT REDUCE GRADIENTS IN FORCE
LINMIN - PRINT DETAILS OF WORKING IN LINMIN
LOCALIZE - LOCALIZED ORBITALS TO BE PRINTED

LOCMIN - PRINT DETAILS OF WORKING IN LOCMIN
MINDO/3 - THE MINDO/3 HAMILTONIAN TO BE USED
MECI - PRINT DETAILS OF MECI CALCULATION
MOLDAT - DETAILS OF WORKING IN MOLDAT TO BE PRINTED
MULLIK - PRINT THE MULLIKEN POPULATION ANALYSIS
NLLSQ - GRADIENTS TO BE MINIMIZED USING NLLSQ.
NOINTER - INTERATOMIC DISTANCES NOT TO BE PRINTED
NOXYZ - CARTESIAN COORDINATES NOT TO BE PRINTED
OLDENS - INITIAL DENSITY MATRIX READ OFF DISK
OPEN - OPEN-SHELL RHF CALCULATION REQUESTED
PI - RESOLVE DENSITY MATRIX INTO SIGMA AND PI BONDS
PL - MONITOR CONVERGANCE OF DENSITY MATRIX IN ITER
POWSQ - PRINT DETAILS OF WORKING IN POWSQ
PRECISE - CRITERIA TO BE INCREASED BY 100 TIMES
PULAY - PULAY'S CONVERGER TO BE USED TO OBTAIN A SCF
QUARTET - RHF QUARTET STATE REQUIRED
QUINTET - RHF QUINTET STATE REQUIRED
RESTART - CALCULATION RESTARTED
ROOT=n - ROOT n TO BE OPTIMIZED IN A C.I. CALCULATION
ROT=n - THE SYMMETRY NUMBER OF THE SYSTEM IS n.
SADDLE - TRANSITION STATE TO BE OPTIMIZED
SCFCRT=.n- DEFAULT SCF CRITERION REPLACED BY THE VALUE SUPPLIED
SEARCH - PRINT DETAILS OF WORKING IN SEARCH
SEXTET - RHF SEXTET STATE REQUIRED
SHIFT=n - A DAMPING FACTOR OF n DEFINED
SIGMA - GRADIENTS TO BE MINIMIZED USING SIGMA
SINGLET - RHF SINGLET STATE REQUIRED
SPIN - FINAL UHF SPIN MATRIX TO BE PRINTED
STEP1=n - STEP SIZE n FOR FIRST COORDINATE IN GRID CALCULATION
STEP2=n - STEP SIZE n FOR SECOND COORDINATE IN GRID CALCULATION
SYMMETRY - SYMMETRY CONDITIONS TO BE IMPOSED
T=n - A TIME OF n SECONDS REQUESTED
THERMO - A THERMODYNAMICS CALCULATION IS TO BE PERFORMED.
TIMES - TIMES OF VARIOUS STAGES TO BE PRINTED
T-PRIO - TIME TAKES PRIORITY IN DRC
TRANS - THE SYSTEM IS A TRANSITION STATE (USED IN THERMODYNAMICS CALCULATION)
TRIPLET - TRIPLET STATE REQUIRED
UHF - UNRESTRICTED HARTREE-FOCK CALCULATION
VECTORS - FINAL EIGENVECTORS TO BE PRINTED
X-PRIO - GEOMETRY CHAGES TAKE PRIORITY IN DRC
XYZ - ALL GEOMETRIC OPERATIONS TO BE DONE IN CARTESIAN COORDINATES.

2.3 DEFINITIONS OF KEY-WORDS

1ELECTRON

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

H(i,j) = 0.5 x (beta(i) +beta(j)) x overlap(i,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 (a) the charge, and (b) the electron density. It is only a function of the geometry. Abbreviation: 1ELEC.

0SCF

The data can be read in and output, but no actual calculation is performed when this key-word is used. This is useful for checking purposes.

1SCF

When a single geometry is to be studied, then 1SCF should be used. All the key-words relevant to output can be used. If the gradients are to be calculated, then GRADIENTS should be specified as they are not calculated by default.

If the key-word 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. MOPAC normally performs many SCF calculations, and in order to minimize output when following the working of the SCF calculation, 1SCF is very useful.

AM1

The new AM1 method is to be used. By default MNDO is run.

BAR=

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 reduces. 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. See also SADDLE.

BIRADICAL

NOTE: BIRADICAL is a redundant key-word, and represents a particular configuration interaction calculation. Experienced users of MECI can duplicate the effect of the key-word BIRADICAL by using the MECI key-words OPEN(2,2) and SINGLET.

For molecules which are believed to have biradicaloid character the option exists to optimize the lowest singlet energy state which results from the mixing of three states. These states are, in order, (1) the (micro)state arising from a one electron excitation from the HOMO to the LUMO, which is combined with the microstate resulting from the time-reversal operator acting on the parent microstate, the result being a full singlet state; (2) the state resulting from de-excitation from the formal LUMO to the HOMO; and (3) the state resulting from the single electron in the formal HOMO being excited into the LUMO.

```
                    Microstate 1          Microstate 2     Microstate 3

             Alpha Beta   Alpha Beta     Alpha  Beta      Alpha  Beta

    LUMO       *                *                           *     *
              ---  ---     ---  ---       ---   ---        ---   ---

                        +

    HOMO             *     *               *     *
              ---  ---     ---  ---       ---   ---        ---   ---
```

A configuration interaction calculation is involved here. A biradical calculation done without C.I. at the RHF level would be meaningless. Either rotational invariance would be lost, as in the D2d form of ethylene, or very artificial barriers to rotations would be found, such as in a methane molecule "orbiting" a D2d ethylene. In both cases the inclusion of limited configuration interaction corrects the error. 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. See also MECI, C.I., OPEN, SINGLET.

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-duodempotency of the density matrix. For an RHF wavefunction the square of the density matrix is equal to twice the density matrix.

C.I.

Normally configuration interaction is invoked if any of the key-words which imply a C.I. calculation are used, such as BIRADICAL, TRIPLET or QUARTET. Note that ROOT= does not imply a C.I. calculation: ROOT= is only used when a C.I. calculation is done. However, as these implied C.I.s involve the minimum number of configurations practical, the user may want to define a larger than minimum C.I., in which case the key-word C.I.=n can be used. 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 will do nothing for a closed-shell system, and lead 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 being used. Numbers of microstates implied by the use of the key-word C.I.=n on its own are as follows:

Keyword	Even-electron systems			Odd-electron systems		
	No. of electrons,		configs	No. of electrons,		configs
	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 key-words also used, see below)					

If a change of spin is defined, then larger numbers of M.O.s can be used up to a maximum of 10. The C.I. matrix is of size 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 key-words defining spin-states.

Note that for any system, use of C.I.=5 or higher normally implies the diagonalization of a 100 by 100 matrix. As a geometry optimization using a C.I. requires the derivatives to be calculated using full SCF calculations, geometry optimization with large C.I.s will require a considerable amount of time.

Associated key-words: MECI, ROOT=, SINGLET, DOUBLET, etc.

CHARGE=

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 -1 or -2 or -3, etc.

EXAMPLES

ION	KEYWORD	ION	KEYWORD
NH4(+)	CHARGE=1	CH3COO(-)	CHARGE=-1
C2H5(+)	CHARGE=1	(COO)(=)	CHARGE=-2
SO4(=)	CHARGE=-2	PO4(3-)	CHARGE=-3
HSO4(-)	CHARGE=-1	H2PO4(-)	CHARGE=-1

CYCLES=

In Bartel's method of gradient norm minimization, NLLSQ, the default number of cycles (100) is replaced by the number n specified by CYCLES=n.

DCART

The cartesian derivatives which are calculated in DCART for variationally optimized systems are printed if the key-word DCART is present. The derivatives are in units of kcals/Angstrom, and the coordinates are displacements in x, y, and z.

DEBUG

Certain key-words 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 a key-word is used, e.g. COMPFG.

DENOUT

The density matrix at the end of the calculation is to be output in a form suitable for input in another job. If an automatic dump due to the time being exceeded occurs during the current run then DENOUT is invoked automatically. (see RESTART)

DENSITY

At the end of a job, when the results are being printed, the density matrix is also printed. For RHF the normal density matrix is printed, for UHF the addition of the alpha and beta density matrices is printed.

If density is not requested, then the diagonal of the density matrix, that is, the electron density on the atomic orbitals, will be 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_6H_{12} 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.

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, then DOUBLET can be specified, and all other spin states, while calculated, are ignored in the choice of root to be used.

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

A Dynamic Reaction Coordinate calculation is to be run. By default, total energy is conserved, so that as the "reaction" proceeds in time energy is transferred 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

Data are echoed back if ECHO is specified. Only useful if data are suspected to be corrupt.

ENPART

This is a very useful tool for analyzing the energy terms within a system. The total energy, in eV, obtained by the addition of the electronic and nuclear terms, is partitioned into mono- and bi-centric contributions, and these contributions in turn are divided into nuclear and one- and two-electron terms.

ESR

The unpaired spin density arising from an odd-electron system can be calculated 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 key-words 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 in, for example, the phenoxy radical, several states should be mixed.

EXCITED

The state to be calculated is the first excited open-shell singlet state. If the ground state is a singlet, then the state calculated will be S(1); if the ground state is a triplet, then S(2). This state would normally be the state resulting from a one-electron excitation from the HOMO to the LUMO. Exceptions would be if the lowest singlet state were a biradical, in which case the EXCITED state could 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 key-word, and represents a particular configuration interaction calculation. Experienced users of MECI can duplicate the effect of the key-word EXCITED by using the MECI key-words OPEN(2,2), SINGLET, and ROOT=2.

EXTERNAL=

Normally, AM1 and MNDO parameters are taken from the BLOCK DATA files within MOPAC. When the supplied parameters are not suitable, as in an element recently parameterized, and the parameters not yet installed in the user's copy of MOPAC, then the new parameters can be inserted at run time by use of EXTERNAL=<filename>, where <filename> is the name of the file which 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 SI.

When new parameters for elements are published, they can be typed in as shown. This file is ended by a blank line, the word END or nothing, i.e., no end-of-file delimiter. An example of a parameter data file would be

```
Start of line£      (Put at least 2 spaces before and after parameter name)

      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
      Line  8:      GSS     Si        9.82
      Line  9:      GPP     Si        7.31
      Line 10:      GSP     Si        8.36
      Line 11:      GP2     Si        6.54
      Line 12:      HSP     Si        1.32
```

Derived parameters do no 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: EXTERNAL can only be used to input parameters for MNDO or AM1. It is unlikely, however, that any more MINDO/3 parameters will be published.

See also DEP to make a permanent change.

FILL=

The n'th M.O. in an RHF calculation is constrained to be filled. 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 n'th M.O. is stored, and, if occupied, no further action is taken at that time. If unoccupied, then the HOMO and the n'th M.O.'s are swapped around, so that the n'th 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. As FILL is known to give difficulty at times, consider also using C.I.=n and ROOT=m.

FORCE

A force-calculation is to be run. The Hessian, that is the matrix (in millidynes per Angstrom) of second derivatives of the energy with respect to displacements 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 about 30 reciprocal centimeters). A transition state is further characterized by one, and exactly one, negative force constant.

A FORCE calculation is a prerequisite for a THERMO calculation.

Before a FORCE calculation is started, a check is made to ensure that a stationary point is being used. This check involves calculating the gradient norm (GNORM) and if it is significant, the GNORM will be reduced using NLLSQ (Bartel's 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 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 D.F.P. routine has difficulty optimizing the geometry. GEO-OK will over-ride 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 system contains, or may give rise to, a Hydrogen molecule. GEO-OK will override other geometric safety checks such as the unstable gradient in a geometry optimization preventing reliable optimization.

See also the message "GRADIENTS OF OLD GEOMETRY, GNORM= nn.nnnn"

GNORM=

The D.F.P. geometry optimization termination criteria can be over-ridden by specifying a gradient norm requirement. For example, GNORM=20 would allow the D.F.P. 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. WARNING: If a very small value is chosen, the D.F.P. procedure 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 time. GRADIENTS allows the gradients to be calculated. All gradients are then calculated, whether marked for calculation or not, and printed. An exception is when the 1SCF was used in conjunction with the key-word RESTART, in which case only the coordinates being optimized would have their gradients printed. 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 data essential for the graphics package, q.v., are then output.

H-PRIORITY

In a DRC calculation, results will be printed whenever the calculated heat of formation changes by 0.1 Kcal/mole. 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/mole.

IRC

An Intrinsic Reaction Coordinate calculation is to be run. All kinetic energy is shed at every point in the calculation. See Background.

IRC=n

An Intrinsic Reaction Coordinate calculation to be run; an initial perturbation in the direction of normal coordinate n to be applied. If n is negative, then perturbation is reversed, i.e., initial motion is in the opposite direction to the normal coordinate.

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 re-calculate the frequencies, etc., starting at the end of the force matrix calculation specify RESTART.

The two key-words 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 re-define 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 a SCF have been exhausted, e.g. SHIFT and PULAY etc.

KINETIC=n.nnn

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

LARGE

Most of the time the output invoked by key-words is sufficient. LARGE will cause less-commonly wanted, but still useful, output to be printed. Currently LARGE only applies to the MECI.

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 significant. If the user does not want the refinement to be carried out, then LET is provided to let the calculation proceed.

LOCALIZE

The occupied eigenvectors are transformed into a localized set of M.O.'s by a series of 2 by 2 rotations which maximize <psi**4>. The value of 1/<psi**4> is a direct measure of the number of centers involved in the M.O., thus for H2 the value of 1/<psi**4> is 2.0, for a three-center bond is 3.0, and a lone pair would be 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 tne most stable conformer of some large compound. However, the localized bonds generated will normally be quite acceptable for routine applications. Abbreviation: LOCAL.

MECI

At the end of the calculation details of the Multi Electron Interaction Calculation are printed if MECI is specified. The state vectors can be printed by specifying VECTORS. The MECI calculation is either invoked automatically, or explicitly invoked by the use of the C.I.=n keyword.

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 read in.

Format for Microstates

After the geometry data plus any symmetry data are read in, data defining each microstate is read in, using format 20I1, one microstate per line.

For a system with n M.O.s in the C.I. (use OPEN=(n1,n) or C.I.=n to do this), 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 component of spin = 1, the state 111000110100, while having a 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 component of spin for every microstate must be the same, except for teaching purposes. Two microstates of different 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 components of a manifold of degenerate M.O.s should be either included or excluded. If only some, but not all, components are used, the required degeneracy of the states will be missing.

As an example, for the tetrahedral methane cation, if the user supplies the microstates corresponding to a component of spin = 3/2, neglecting Jahn-Teller distortion, the minimum number of states that can 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 20I1 so that spaces can be used for empty M.O.s.

MINDO/3

The default Hamiltonian within MOPAC is MNDO, with the alternatives of AM1 and MINDO/3. To use the MINDO/3 Hamiltonian the key-word MINDO/3 should be used. Acceptable alternatives to the key-word MINDO/3 are MINDO and MINDO3.

MULLIK

A full Mulliken Population analysis is to be 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.
(3) The overlap population is formed from P(i,j)*S(i,j).
(4) Half the off-diagonals are added onto the diagonals.

NLLSQ

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

(a) A minimum in the 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.

(b) 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.

(c) 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 MOPAC is to be used to start the current calculation. This can be used in attempts to obtain a SCF when a previous calculation *ended successfully but* a subsequent run failed to go SCF.

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. OPEN(1,1) 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.

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 on-diagonal terms give the hybridization state, so that an sp2 hybridized system would be represented as s-sigma 1.0, p-sigma 2.0, p-pi 1.0

POWSQ

Details of the working of POWSQ are printed out. This is only useful in debugging.

PRECISE

The criteria for terminating all optimizations, electronic and geometric, are to be increased by a factor, normally, 100. This can be used where more precise results are wanted. If the results are going to be used in a FORCE calculation, where the geometry needs to be known quite precisely, then PRECISE is recommended; for small systems the extra cost in CPU time is minimal.

PULAY

The default converger in the SCF calculation is to be replaced by Pulay's procedure as soon as the density matrix is sufficiently stable. A considerable improvement in speed can be achieved by the use of PULAY. If a large number of SCF calculations are envisaged, a sample calculation using 1SCF and PULAY should be compared with using 1SCF on its own, and if a saving in time results, then PULAY should be used in the full calculation. PULAY should be used with care in that its use will prevent the combined package of convergers (SHIFT, PULAY and the CAMP-KING convergers) from automatically being used in the event that the system fails to go SCF in (ITRY-10) iterations.

The combined set of convergers very seldom fails.

QUARTET

The desired spin-state is a quartet. That is, the state with 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, for component of spin = 3/2. 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 root to be used. 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, that is, the state with 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, for component of spin = 2. 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 root to be used. 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 has been stopped, for whatever reason, and intermediate results have been stored, then the calculation can be restarted at the point where it stopped by specifying RESTART. The most common cause of a job stopping before completion is its exceeding the time allocated. A saddle-point calculation has no restart, but the output file contains information which can easily be used to start the calculation from a point near to where it stopped.

It is not necessary to change the geometric data to reflect the new geometry, as a result the geometry printed at the start of a restarted job will be that of the original data, not that of the restarted 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 1: In the FORCE calculation two restarts are possible. These are (a) a restart in FLEPO if the geometry was not optimized fully before FORCE was called, and (b) the normal restart in the construction of the force matrix. If the restart is in FLEPO within FORCE then the key-word FORCE should be deleted, and the key-word RESTART used on its own. Forgetting this point is a frequent cause of failed jobs. NOTE 2: Two restarts also exist in the IRC calculation. If an IRC calculation stops while in the FORCE calculation, then a normal restart can be done. If the job stops while doing the IRC calculation itself then the keyword IRC=n should be changed to IRC, or it can be omitted if DRC is also specified. The absence of the string "IRC=" is used to indicate that the FORCE calculation was completed before the restart files were written.

ROOT=n

The n'th root of a C.I. calculation is to be used in the calculation. If a key-word specifying the spin-state is also present, e.g. SINGLET or TRIPLET, then the n'th 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 key-word would not be used. It is retained for educational and research purposes. Unusual care should be exercised when ROOT= is specified.

ROT=n

In the calculation of the 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 ----

C1 CI CS	1	D2 D2D D2H	4	C(INF)V	1
C2 C2V C2H	2	D3 D3D D3H	6	D(INF)H	2
C3 C3V C3H	3	D4 D4D D4H	8	T TD	12
C4 C4V C4H	4	D6 D6D D6H	12	OH	24
C6 C6V C6H	6	S6	3		

SADDLE

The transition state in a simple chemical reaction is to be optimized. Extra data are required. After the first geometry, specifying the reactants, and any symmetry functions have been defined, the second geometry, specifying the products, is defined, using the same format as that of the first geometry.

SADDLE often fails to work successfully. Frequently this is due to dihedral angles being different by about 360 degrees rather than zero degrees. As the choice of dihedral can be difficult, users should consider running this calculation with the key-word XYZ. There is normally no ambiguity in the definition of cartesian coordinates. See also BAR=.

Many of the bugs in SADDLE have been removed in this version. Use of the XYZ option is strongly recommended.

SCFCRT=

The default SCF criterion is to be replaced by that defined by SCFCRT=. The SCF criterion can be varied from about 0.0001 to 0.0000000001. To find a suitable value 1SCF and various values of SCFCRT=n.nnn should be used; a 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 a SCF, and consequent failure of the run.

SEXTET

The desired spin-state is a sextet: the state with component of spin = 5/2 and spin = 5/2.

The sextet states are the highest spin states normally calculable using MOPAC in its unmodified form. If SEXTET is used on its own, then 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=

In an attempt to obtain an SCF by damping oscillations which are slowing down the convergence or preventing a SCF being achieved, SHIFT can be used. The principle is that 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, the oscillations being attributed to autoregenerative charge fluctuations. 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 an SCF being generated. 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 the CAMP-KING convergers) from automatically being used in event that the system fails to go SCF in (ITRY-10) iterations.

The combined set of convergers almost never fails.

SIGMA

The McIver-Komornicki gradient norm minimization routines, POWSQ and SEARCH are to be used. These are very rapid routines, but do not work for all species. If the gradient norm is low, i.e., less than about 5 units, then SIGMA will probably work; in most cases, NLLSQ is recommended. SIGMA first calculates a quite accurate Hessian matrix, a slow step, then works out the direction of fastest decent, 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 found. Clearly, if the Hessian changes markedly as a result of the line-search, the update done will be inaccurate, and the new search direction will be faulty.

Of course, SIGMA should be avoided if at all possible when non-variationally optimized calculations are being done.

SINGLET

When a configuration interaction calculation is done, all spin states are calculated simultaneously, either for 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 root to be used.

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, is to be printed. If the system has a closed-shell ground state, e.g. methane run 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 the 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 the two parameters is given by n.nnn.

SYMMETRY

Symmetry data defining related bond lengths, angles and dihedrals can be included by supplying additional data after the geometry has been entered. If there are any other data, such as values for the reaction coordinates, or a second geometry, as required by SADDLE, then it would 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 because many derivatives do not need 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 run. (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 D6h symmetry. Any deformation from this symmetry must involve error, so by imposing symmetry some error is removed.

The layout of the symmetry data is:

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

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
 0    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 a multiple of 60 degrees.

Spaces, tabs or commas can be used to separate data. Note that only three parameters are marked to be optimized. To end symmetry data use 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=

This is a facility to allow the program to shut down in an orderly manner on computers with execution time C.P.U. limits.

The total C.P.U. time allowed for the current job is limited to *nn.nn* seconds; by default this is one hour, i.e., 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 then stop. The safety margin is 100 percent; that is, to do another cycle, enough time to do at least two full cycles must remain.

An alternative specification is T=nn.nnM, this allows time in minutes to be defined, thus T=60M will define 3600 seconds, or 60 minutes.

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, and checking of the lower bound of the temperature range is done to prevent temperatures of less than 100K being used.

Another limitation, for which no checking is done, is that there should be no internal rotations. If any exist, they will not be recognized as such, and the calculated quantities will be too low as a result.

If THERMO is specified on its own, then the default values of the temperature range are assumed. This starts at 200K and increases in steps of 10 degrees to 400K. Three options exist for overriding the default temperature range. These are:

THERMO(nnn)

The thermodynamic quantities for a 200 degree range of temperatures, starting at nnnK and with an interval of 10 degrees are to 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, the step size being calculated in order to give approximately 20 points, and a reasonable value for the step. The size of the step in Kelvin degrees will be 1, 2, or 5, or a power of 10 times these numbers.

THERMO(nnn,mmm,lll)

As for THERMO(nnn,mmm) only now the user can explicitly define the step size. The step size cannot be less than 1K.

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.

TRANS

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

3N-5 for a linear ground state system,

3N-6 for a non-linear ground state system, or

3N-6 for a linear transition-state complex,

3N-7 for a non-linear transition-state complex.

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

TRANS=n

The facility exists to allow the THERMO calculation to handle systems with internal rotations. TRANS=n will remove the n lowest vibrations. Note that TRANS=1 is equivalent to TRANS on its own. For xylene, for example, TRANS=2 would be suitable.

TRIPLET

The triplet state is defined. 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 numbers of alpha and beta electrons are set equal. This does not necessarily correspond to a singlet.

RHF interpretation.

A RHF MECI calculation is performed to calculate the triplet state. If no other C.I. key-words 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,..), that is, one electron is put in each of the two highest occupied M.O.s.

See key-words C.I.=n and OPEN(n1,n2).

UHF

The unrestricted Hartree-Fock Hamiltonian is to be used.

VECTORS

The eigenvectors are to be printed. In UHF calculations both alpha and beta eigenvectors are printed; in all cases the full set, occupied and virtual, are output. The eigenvectors are normalized to unity, that is 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. This is useful in a learning context, but would normally be very undesirable.

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 quite often fails due to faulty definition of the second geometry because the dihedrals give a lot of difficulty. To make this option easier to use, XYZ was developed. Using XYZ the calculation runs entirely in cartesian coordinates, this eliminating the problems associated with dihedrals. The connectivity of the two systems can be different, but the numbering must be the same. Dummy atoms can be used; these will be removed at the start of the run. A new numbering system will be generated by the program, when necessary.

CHAPTER 3

GEOMETRY SPECIFICATION

FORMAT: The geometry is read in using essentially "Free-Format" of FORTRAN-77. In fact, a character input is used in order to accommodate the chemical symbols, but the numeric data can be regarded as "free-format". This means that integers and real numbers can be interspersed, 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.

The geometry can be defined in terms of either internal or cartesian coordinates.

INTERNAL COORDINATE DEFINITION

For any one atom (i) this consists of an interatomic distance in Angstroms from an already-defined atom (j), an interatomic angle in degrees between atom i and j and an already defined k, (k and j must be different atoms), and finally a torsional angle in degrees between atom i, j, k, and an already 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 can be connected to atom 1 or 2, and must make an angle with atom 2 or 1 (thus - 3-2-1 or 3-1-2); no dihedral is possible for atom 3. By default, atom 3 is connected to atom 2.

3.1 CONSTRAINTS

1. Interatomic distances must be greater than zero. Zero Angstroms is not acceptable. The only exception is if the parameter is symmetry-related to another atom, and is the dependent function.

2. Angles must be positive. This constraint is for the benefit of the user only; negative angles are the result of errors in the construction of the geometry.

3. Dihedrals can normally only assume definable angles. 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. During the calculation this constraint is checked continuously, and if atoms j, k, and l lie within 0.1 Angstroms of a straight line the calculation will output an error message and then stop. Two exceptions to this constraint are:

   (a) if the angle is zero or 180 degrees, in which case the dihedral is not used.

   (b) if atoms j, k, and l lie in an exactly straight line (usually the result of a symmetry constraint), as in acetylene, acetonitrile, but-2-yne, etc.

If the exceptions are used, care must be taken to ensure that the program does not violate these constraints during any optimizations or during any calculations of derivatives - see also FORCE.

CARTESIAN COORDINATE DEFINITION

Cartesian coordinates consist of the chemical symbol or atomic number, then the cartesian coordinates and optimization flags but no connectivity.

MOPAC uses the lack of connectivity to indicate that cartesian coordinates are to be used. A unique case is the triatomics for which only internal coordinates are allowed. This is to avoid conflict of definitions: the user does not need to define the connectivity of atom 2, and can elect to use the default connectivity for atom 3. As a result, a triatomic may have no explicit connectivity defined, the user thus taking advantage of the default connectivity. Since internal coordinates are more commonly used than cartesian, the above choice was made.

If the key-word XYZ is absent then every coordinate must be marked for optimization. If any coordinates are not to be optimized, then the key-word XYZ must be present. The coordinates of all atoms, including atoms 1, 2 and 3 can be optimized. Dummy atoms should not be used, for obvious reasons.

3.2 DEFINITION OF ELEMENTS AND ISOTOPES

Elements are defined in terms of their atomic numbers or their chemical symbols. Acceptable symbols for MNDO are:

1	3	5	6	7	8	9	11	13	14	15	16	17	19	24	32	35
H	Li	B	C	N	O	F	Na	Al	Si	P	S	Cl	K	Cr	Ge	Br
	LI						NA	AL	SI			CL		CR	GE	BR

50	53	80	92	99	102	103	104	105	106	107
Sn	I	Hg	Pb	Xx	Cb	++	+	--	-	Tv
SN		HG	PB	XX	CB					TV

Old parameters for some elements are available. These are provided to allow compatibility with earlier copies of MOPAC. To use these older parameters, use a keyword composed of the chemical symbol followed by the year of publication of the parameters. Keywords currently available: Si1978 S1978.
and for AM1, acceptable symbols are

1	6	7	8	9	17	35	53	99	102	103	104	105	106	107
H	C	N	O	F	Cl	Br	I	Xx	Cb	++	+	--	-	Tv
					CL	BR		XX	CB					TV

Diatomics Parameterized 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 parameterized within MINDO/3.

Extra entities available to MNDO, MINDO/3 and AM1

\+ A 100% ionic alkali metal.

++ A 100% ionic alkaline earth metal.

\- A 100% ionic halogen-like atom

-- A 100% ionic group VI-like atom.

Cb A special type of monovalent atom

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

Element 102, symbol Cb, is designed to satisfy valency requirements of atoms for which some bonds are not completed. Thus in "solid" diamond the usual way to complete the normal valency in a cluster model is to use hydrogen atoms. This approach has the defect that the electronegativity of hydrogen is different to that of carbon. The "Capped bond" atom, Cb, is designed to satisfy these valency requirements without acquiring a net charge.

Cb behaves like a monovalent atom, with the exception that it can alter its electronegativity to achieve an exactly zero charge in whatever environment it finds itself. It is thus all things to all atoms. On bonding to hydrogen it behaves similar to a hydrogen atom. On bonding to fluorine it behaves like a very electronegative atom. If several capped bond atoms are used, each will behave independently. Thus if the two hydrogen atoms in formic acid were replaced by Cb's then each Cb would independently become electroneutral.

Capped bonds should not be optimized. They are still very new and not enough is known yet. A fixed bond-length of 1.7 A is recommended, if two Cb are on one atom, a contained angle of 109.471221 degrees is suggested, and if three Cb are on one atom, a contained dihedral of -120 degrees should be used.

Element 99, or XX is known as a dummy atom, and is used in the definition of the geometry; it is deleted automatically from any cartesian coordinate geometry files. Dummy atoms are pure mathematic points, and are useful in defining geometries; for example, in ammonia the definition of C3v symmetry is facilitated by using one dummy atom and symmetry relating the three hydrogens to it.

Output normally only gives chemical symbols.

Isotopes are used in conjunction with chemical symbols. If no isotope is specified, the average isotopic mass is used, thus chlorine is 35.453. This is different from all previous versions of MOPAC, 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 a user wishes to specify any specific isotope it should immediately follow the chemical symbol (no space), thus: H2, H2.0140, C13, C13.00335.

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

Each internal coordinate is followed by an integer, to indicate 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:

Only one reaction coordinate is allowed, but this can be made more versatile by the use of SYMMETRY. If a reaction coordinate is used, the values of the reaction coordinate should follow immediately after the geometry and any symmetry data. No terminator is required, and free-format-type input is acceptable.

If two "reaction coordinates" are used, then MOPAC assumes that the two-dimensional space in the region of the supplied geometry is to be mapped. The two dimensions to be mapped are in the plane defined by the "-1" labels. Step sizes in the two directions must be supplied using STEP1 and STEP2 on the key-word line.

Using internal coordinates, the first atom has three unoptimizable coordinates, the second atom two, (the bond-length can be optimized) and the third atom has one unoptimizable coordinate. None of these six unoptimizable coordinates at the start of the geometry should be marked for optimization. If any are so marked, a warning is given, and the calculation will continue.

In cartesian coordinates all parameters can be optimized.

3.3 EXAMPLES OF COORDINATE DEFINITIONS.

First, formaldehyde. By definition atom 1, here oxygen, is at the origin. Atom 2 is defined as being along the "x" axis, and is bonded to atom 1. Because of these definitions, the user does not need to specify the connectivity of atom 2, but can do so if desired. Further, by default, but not by definition, atom 3 is connected to atom 2, and makes an angle with atom 1. If the user wishes, either the default connectivity for atom 3, or the alternative connectivity, to atom 1, making an angle with atom 2, can be explicitly defined.

In this example, the minimum data necessary to define the geometry has been entered. Thus for atoms 2 and 3 the connectivity has been omitted: the default connectivity is used. For atom 4, however, the full connectivity must be explicitly defined; no defaults are allowed.

```
O                                   The first atom has no coordinates.
C         1.2 1                     The C-O bond length is to be optimized.
H2.0140   1.0 1 120   1             The third atom is a Deuterium.
1         0.0 0   0.0 1  180 0 2 1 3 Atomic number of Hydrogen is used, and
                                    SYMMETRY must have been specified, in
order to not have to give a bond-length or an angle. The dihedral is point-
group defined as 180 degrees.
```

MOPAC can generate data-files, both in the Archive files, and at the end of the normal output file, when a job ends prematurely due to time restrictions. Here the coordinate definition for formic acid is shown.

Note that all coordinates are generated, as is the full connectivity. Also, the data are all neatly lined up. This is, of course, characteristic of machine-generated data, but is useful when checking for errors.

Format of internal coordinates in ARCHIVE file

```
 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
 0    0.000000 0     0.000000 0     0.000000 0   0  0  0
```

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.

Example of polymer coordinates from ARCHIVE file

```
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    1.542344  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  7  8
 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. With MOPAC comes two sets of data for running calculations. One of these is called MNRSD1.DAT, and this will now be described.

4.1 MNRSD1 TEST DATA FILE FOR FORMALDEHYDE

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

```
Line  1:        SYMMETRY
Line  2:  Formaldehyde, for Demonstration Purposes
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
Line  2:  Formaldehyde, for Demonstration Purposes
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 files.

In all geometric specifications care must be taken in defining the internal coordinates to ensure that no three atoms being used to define a fourth atom's dihedral angle ever fall into 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 would become ill-defined.

4.2 MOPAC OUTPUT FOR TEST-DATA FILE MNRSD1

```
**********************************************************************
*** FRANK J SEILER RES. LAB., U.S. AIR FORCE ACADEMY, COLO. SPGS., CO. 80840 ***
**********************************************************************

                          MNDO CALCULATION RESULTS              Note 1


**********************************************************************
*                      VERSION  3.10                            Note 2
*  SYMMETRY - SYMMETRY CONDITIONS TO BE IMPOSED
**********************************************************************



     PARAMETER DEPENDENCE DATA

        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                                                    Note 3
  Formaldehyde, for Demonstration Purposes


         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                                                  Note 4
          2       C        1.20000 *                                   1
          3       H        1.10000 *    120.00000 *                    2   1
          4       H        1.10000      120.00000      180.00000       2   1   3


            CARTESIAN COORDINATES

    NO.       ATOM        X         Y          Z
```

```
    1          O          0.0000    0.0000    0.0000
    2          C          1.2000    0.0000    0.0000                  Note 5
    3          H          1.7500    0.9526    0.0000
    4          H          1.7500   -0.9526    0.0000
 H: (MNDO):  M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4907, (1977)
 C: (MNDO):  M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4907, (1977)
 O: (MNDO):  M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4907, (1977)


          RHF CALCULATION, NO. OF DOUBLY OCCUPIED LEVELS =  6


           INTERATOMIC DISTANCES

              O  1       C  2       H  3       H  4
------------------------------------------------------
    O  1   0.000000
    C  2   1.200000   0.000000                                     Note 6
    H  3   1.992486   1.100000   0.000000
    H  4   1.992486   1.100000   1.905256   0.000000
CYCLE:  1 TIME:   0.06 TIME LEFT:   3599.9 GRAD.:    36.370 HEAT: -32.86455
CYCLE:  2 TIME:   0.03 TIME LEFT:   3599.9 GRAD.:     8.250 HEAT: -32.88052
HEAT OF FORMATION TEST SATISFIED                                   Note 7
PETERS TEST SATISFIED                                              Note 8
  SYMMETRY                                                         Note 9
 Formaldehyde, for Demonstration Purposes                          Note 10


    PETERS TEST WAS SATISFIED IN FLETCHER-POWELL OPTIMISATION     Note 11
    SCF FIELD WAS ACHIEVED                                        Note 12


                            MNDO    CALCULATION                   Note 13
                                                   VERSION  3.10




          FINAL HEAT OF FORMATION =    -32.88189 KCAL             Note 14


          ELECTRONIC ENERGY       =   -870.74066 EV
          CORE-CORE REPULSION     =    392.62149 EV

          IONISATION POTENTIAL    =     11.04150
                                                      06/03/86
          NO. OF FILLED LEVELS    =     6
          MOLECULAR WEIGHT        =     30.026


          SCF CALCULATIONS  =           15
          COMPUTATION TIME  =            0.18 SECONDS            Note 15
```

```
    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                                                     Note 16
     2       C        1.21646 *                                     1
     3       H        1.10601 *     123.51050 *                     2   1
     4       H        1.10601       123.51050      180.00000        2   1   3

          INTERATOMIC DISTANCES

              O  1        C  2        H  3        H  4
------------------------------------------------------------
   O  1   0.000000
   C  2   1.216455   0.000000
   H  3   2.046609   1.106012   0.000000
   H  4   2.046609   1.106012   1.844352   0.000000

                EIGENVALUES

-42.99227 -25.11986 -16.95504 -16.29961 -14.17985 -11.04150   0.85990   3.67652
  3.84972   7.12622                                                   Note 17

              NET ATOMIC CHARGES AND DIPOLE CONTRIBUTIONS

          ATOM NO.   TYPE          CHARGE        ATOM  ELECTRON DENSITY
            1          O           -0.2902         6.2902
            2          C            0.2922         3.7078             Note 18
            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                   Note 19
SUM             2.164     0.000     0.000     2.164

         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               1.8271    0.9222    0.0000
    4         H               1.8271   -0.9222    0.0000

         ATOMIC ORBITAL ELECTRON POPULATIONS

  1.88260   1.21603   1.89109   1.30047   1.25526   0.86214   0.89092   0.69953
  1.00098   1.00098                                                   Note 20
```

NOTES ON RESULTS 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 MOPAC, and refers to the program, not to the Hamiltonians used. The version number should be cited in any correspondence regarding MOPAC.

All the key-words used, along with a brief explanation, should be printed at this time. If a key-word is not printed, it has not been recognized by the program. Key-words can be in upper or lower case letters, or any mixture. Originally only uppercase letters were allowed, as use of lower case letters involves an extension to the FORTRAN-77 standard. It was found, however, that all computers used support lower case letters, so the ability to use lower case letters has been included, for the sake of convenience.

NOTE 3: Symmetry information is output to allow the user to verify that the requested symmetry functions have in fact been recognized and used.

NOTE 4: The data for this example used a mixture of atomic numbers and chemical symbols, but the internal coordinate output is consistently in chemical symbols.

The atoms in the system are, in order:

Atom 1, an oxygen atom; this is defined as being at the origin.

Atom 2, the carbon atom. Defined as being 1.2 Angstroms from the oxygen atom, it is located in the +x direction. This distance is marked for optimization.

Atom 3, a hydrogen atom. It is defined as being 1.1 Angstroms from the carbon atom, and making an angle of 120 degrees with the oxygen atom. The asterisks indicate that the bond length and angle are both to be optimized.

Atom 4, a hydrogen atom. The bond length supplied has been overwritten with the symmetry-defined C-H bond length. Atom 4 is defined as being 1.1 Angstroms from atom 2, making a bond-angle of 120 degrees with atom 1, and a dihedral angle of 180 degrees with atom 3.

None of the coordinates of atom 4 are marked for optimization. The bond-length and angle are symmetry-defined by atom 3, and the dihedral is group-theory symmetry-defined as being 180 degrees. (The molecule is flat.)

NOTE 5: The cartesian coordinates are calculated as follows:

Stage 1: The coordinate of the first atom is defined as being at the origin of cartesian space, while the coordinate of the second atom is defined as being 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 x-y plane, from either atom 1 or 2 as defined in the data, or from atom 2 if no numbering is given. 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. As this particular system contains no dummy atoms, nothing is done.

NOTE 6: The interatomic distances are output for the user's advice, and a simple check made to insure that the smallest interatomic distance is greater than 0.8 Angstroms.

NOTE 7: The geometry is optimized in a series of cycles, each cycle consisting of a line-search and calculation of the gradients. The time given is the C.P.U. time for the cycle; time left is the total time requested (here 100 seconds) less the C.P.U. time since the start of the calculation (which is earlier than the start of the first cycle!). These times can vary considerably from cycle to cycle due to different options being used, for example whether or not the density matrix is being updated at every point in the line search. The gradient is the scalar length in kcal/mole/Angstrom of the gradient vector.

NOTE 8: 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.

NOTE 9 and 10: The key-words used, titles and comments are reproduced here to remind the user of the name of the calculation.

NOTE 11 and 12: Two messages are given here. The first is a reminder of how the geometry was obtained, whether from the D.F.P., Bartel's or the McIver-Komornicki methods. For any further results to be printed the second message must be as shown; when no SCF is obtained no results will be printed.

NOTE 13: Again, the results are headed with either MNDO or MINDO/3 banners, and the version number.

NOTE 14: 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 15: Advice on time required for the calculation. This is obviously useful in estimating the times required for other systems.

NOTE 16: 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 17: The roots are the eigenvalues or energy levels in electron volts of the molecular orbitals. There are six filled levels, therefore the HOMO has an energy of -11.041eV, and is mainly lone-pair on oxygen. The eigenvectors form an orthonormal set.

NOTE 18: 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, or atom electron density on the atom, here 1.0010, 3.7079, and 6.2902 respectively.

NOTE 19: 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 20: 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.

CHAPTER 5

TESTDATA

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

5.1 DATA FILE FOR A FORCE CALCULATION

```
Line 1 : ROT=2 THERMO(298,298,) PRECISE FORCE     ISOTOPE SYMMETRY
Line 2 :   DEMONSTRATION OF MOPAC - FORCE AND THERMODYNAMICS CALCULATION
Line 3 :   FORMALDEHYDE, MNDO ENERGY = -32.8819
Line 4a:   0     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:   0     0.000000  0     0.000000  0     0.000000  0    0  0  0
Line 5a:   3,    1,    4,
Line 5b:   3,    2,    4,
```

5.2 RESULTS FILE FOR THE FORCE CALCULATION

```
*******************************************************************************
*** FRANK J SEILER RES. LAB., U.S. AIR FORCE ACADEMY, COLO. SPGS., CO. 80840 ***
*******************************************************************************

                         MNDO CALCULATION RESULTS


*******************************************************************************
*                      VERSION  3.10
*  SYMMETRY - SYMMETRY CONDITIONS TO BE IMPOSED
*  FORCE    - FORCE CALCULATION SPECIFIED
*  PRECISE  - CRITERIA TO BE INCREASED BY 100 TIMES
*  ISOTOPE  - FORCE MATRIX WRITTEN TO DISK (CHAN. 9 )
*  THERMO   - THERMODYNAMIC QUANTITIES TO BE CALCULATED
*  ROT      - SYMMETRY NUMBER OF  2 SPECIFIED
```

```
*******************************************************************************


    PARAMETER DEPENDENCE DATA

        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
 ROT=2 THERMO(298,298,) PRECISE FORCE    ISOTOPE SYMMETRY          Note 1
DEMONSTRATION OF MOPAC - FORCE AND THERMODYNAMICS CALCULATION
FORMALDEHYDE, MNDO ENERGY = -32.8819

        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 *                                     1
         3       H        1.10619 *     123.51502 *                     2   1
         4       H        1.10619 *     123.51502 *   179.99569 *       2   1   3


        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        1.8273    0.9223    0.0000
     4         H        1.8273   -0.9223    0.0001
 H: (MNDO):  M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4907, (1977)
 C: (MNDO):  M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4907, (1977)
 O: (MNDO):  M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4907, (1977)

        RHF CALCULATION, NO. OF DOUBLY OCCUPIED LEVELS =  6


          INTERATOMIC DISTANCES

               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
```

```
          INTERNAL COORDINATE DERIVATIVES

  ATOM  AT. NO.  BOND        ANGLE       DIHEDRAL

    1     O
    2     C     0.020543                                              Note 2
    3     H     0.140798  0.010501
    4     H     0.000000  0.000000 -0.002392

          GRADIENT NORM =   0.14270

          TIME FOR SCF CALCULATION =    0.03

          TIME FOR DERIVATIVES     =    0.02                          Note 3

          SYMMETRY WAS SPECIFIED, BUT CANNOT BE USED HERE

          MOLECULAR WEIGHT =   30.03


           PRINCIPAL MOMENTS OF INERTIA IN CM(-1)

          A =    9.831716   B =    1.261967   C =    1.118411


           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

     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

          DEFAULT TIME OF 3600.00 SECONDS ALLOCATED FOR THIS STEP

          ESTIMATED TIME TO COMPLETE CALCULATION =     1.24 SECONDS
  STEP:   1 TIME =     0.06 SECS, INTEGRAL =      0.06 TIME LEFT:   3599.82
  STEP:   2 TIME =     0.06 SECS, INTEGRAL =      0.12 TIME LEFT:   3599.75
  STEP:   3 TIME =     0.06 SECS, INTEGRAL =      0.19 TIME LEFT:   3599.69
  STEP:   4 TIME =     0.06 SECS, INTEGRAL =      0.25 TIME LEFT:   3599.63
  STEP:   5 TIME =     0.06 SECS, INTEGRAL =      0.31 TIME LEFT:   3599.57
```

```
STEP:   6 TIME =    0.06 SECS, INTEGRAL =     0.37 TIME LEFT:    3599.51
STEP:   7 TIME =    0.06 SECS, INTEGRAL =     0.43 TIME LEFT:    3599.45
STEP:   8 TIME =    0.06 SECS, INTEGRAL =     0.48 TIME LEFT:    3599.39
STEP:   9 TIME =    0.07 SECS, INTEGRAL =     0.56 TIME LEFT:    3599.32
STEP:  10 TIME =    0.06 SECS, INTEGRAL =     0.62 TIME LEFT:    3599.26
STEP:  11 TIME =    0.06 SECS, INTEGRAL =     0.68 TIME LEFT:    3599.20
STEP:  12 TIME =    0.07 SECS, INTEGRAL =     0.75 TIME LEFT:    3599.13


          FORCE MATRIX IN MILLIDYNES/ANGSTROM

                O  1      C  2      H  3      H  4

-----------------------------------------------------
    O  1   9.557769
    C  2   8.683380  11.426325
    H  3   0.598735   2.552483   3.033980
    H  4   0.598769   2.552451   0.303925   3.033952


         HEAT OF FORMATION =  -32.881894 KCALS/MOLE


         TRIVIAL VIBRATIONS, SHOULD BE ZERO
                                                                Note 4
  0.0000=TX   0.0000=TY   0.0000=TZ   0.0022=RX   0.0003=RY   0.0001=RZ


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

  0.94406   1.35190   1.86920   4.07261   6.74514  18.72184


          ASSOCIATED EIGENVECTORS



                                                                Note 5

   ROOT NO.    1          2          3          4          5          6

            0.94406    1.35190    1.86920    4.07261    6.74514   18.72184

        1  -0.33016    0.00000   -0.00002    0.41812    0.00000   -0.68277
        2   0.00000    0.60834   -0.00002    0.00000   -0.04811    0.00000
        3  -0.00001   -0.00003   -0.28948    0.00000    0.00000    0.00000
        4  -0.34275    0.00000   -0.00002    0.32022    0.00000    0.72800
        5   0.00000   -0.34892    0.00004    0.00000    0.79262    0.00000
        6  -0.00001    0.00003    0.86603    0.00004   -0.00003    0.00000
        7   0.33646   -0.48710    0.00003   -0.36917   -0.21481   -0.02261
        8  -0.52302   -0.12971    0.00001   -0.47438   -0.37226   -0.03759
        9   0.00003   -0.00001   -0.28827    0.00000    0.00000    0.00000
       10   0.33645    0.48710    0.00001   -0.36917    0.21480   -0.02261
       11   0.52303   -0.12971   -0.00003    0.47438   -0.37226    0.03759
       12  -0.00001    0.00000   -0.28827   -0.00004    0.00003    0.00000
```

```
          ZERO POINT ENERGY       18.150 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



   I         1          2          3          4          5          6
FREQ(I) 1209.9907  1215.0493  1490.5916  2114.5759  3255.3657  3301.5822

MASS(I)    1.93736    1.50020    0.66394    6.79206    0.75075    0.66443

DIPX(I)   -0.00002    0.00033   -0.34444   -3.36646   -0.00009   -0.34788
DIPY(I)   -0.85473   -0.00087   -0.00006    0.00002    0.78304   -0.00016
DIPZ(I)   -0.00010    0.12739   -0.00002   -0.00008   -0.00003   -0.00001

DIPT(I)    0.85473    0.12739    0.34444    3.36646    0.78304    0.34788
                                                                           Note 6

   I         7          8          9         10         11         12
FREQ(I)    -0.0056     0.0005    -0.0019    87.0251    11.1816    10.6530

MASS(I)    0.00000    0.00000    0.00000    0.50395    0.00000    0.00000

DIPX(I)    0.00002    0.00001    0.00023    0.00000   -0.00028   -0.00001
DIPY(I)   -0.00004   -0.00005   -0.00005    0.00000    0.00005    0.46525
DIPZ(I)    0.00000    0.00000    0.00004    0.00001    0.56207   -0.00002

DIPT(I)    0.00005    0.00005    0.00024    0.00002    0.56207    0.46525


          NORMAL VECTORS



                                                                           Note 7
       ROOT NO.   1           2           3           4           5           6

           1209.99070  1215.04928  1490.59161  2114.57589  3255.36573  3301.58223

         1    0.00000     0.00000    -0.04332     0.52497     0.00000     0.00070
         2    0.07922     0.00007     0.00000     0.00000     0.00432     0.00000
         3    0.00004    -0.04399     0.00000     0.00000     0.00000     0.00000
         4    0.00000     0.00000    -0.03977    -0.66818    -0.00001    -0.06524
         5   -0.15855    -0.00015     0.00000     0.00000     0.08917    -0.00002
         6   -0.00016     0.17532     0.00000     0.00001     0.00000     0.00000
         7   -0.62003    -0.00060     0.58074    -0.18535    -0.41965     0.38325
         8    0.31599     0.00030    -0.40126    -0.32346    -0.56544     0.59260
         9    0.00067    -0.69546     0.00001    -0.00002     0.00000    -0.00001
        10    0.62003     0.00060     0.58074    -0.18533     0.41979     0.38310
        11    0.31600     0.00025     0.40126     0.32344    -0.56567    -0.59239
```

```
         12   0.00065    -0.69546    -0.00002    -0.00004    0.00005     0.00004



      ROOT NO.     7           8           9          10          11          12

              -0.00557     0.00049    -0.00194    87.02506    11.18157    10.65295

          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


 VIB.  1     ATOMS   C 2   AND   H 3   SHIFT  0.78  ANGSTROMS    6.8%  RADIALLY
 FREQ.     1209.99   C 2         H 4          0.78               6.8%

 VIB.  2     ATOMS   C 2   AND   H 3   SHIFT  0.87  ANGSTROMS    0.0%  RADIALLY
 FREQ.     1215.05   C 2         H 4          0.87               0.0%

 VIB.  3     ATOMS   C 2   AND   H 3   SHIFT  0.74  ANGSTROMS    1.1%  RADIALLY
 FREQ.     1490.59   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.     3255.37   C 2         H 4          0.78             100.0%

 VIB.  6     ATOMS   C 2   AND   H 3   SHIFT  0.74  ANGSTROMS   99.8%  RADIALLY
 FREQ.     3301.58   C 2         H 4          0.74              99.8%

 VIB. 10     ATOMS   C 2   AND   H 3   SHIFT  0.71  ANGSTROMS    0.0%  RADIALLY
 FREQ.       87.03   C 2         H 4          0.71               0.0%

 VIB. 11     ATOMS   O 1   AND   C 2   SHIFT  0.63  ANGSTROMS    0.0%  RADIALLY
 FREQ.       11.18   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.       10.65   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 MOPAC - 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
```

Note 8

```
                    CALCULATED THERMODYNAMIC PROPERTIES

   TEMP. (K)   PARTITION FUNCTION     ENTHALPY      HEAT CAPACITY    ENTROPY
                                       CAL/MOL       CAL/K/MOL      CAL/K/MOL


   298  VIB.          1.007          23.37536292     0.47807894     0.09142732
        ROT.      709.               888.305         2.981          16.026
        INT.      714.               911.681         3.459          16.117
        TRA.      0.159E+27         1480.509         4.968          36.113
        TOT.                        2392.1893        8.4271         52.2301
```

NOTE 1: 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 C2v point-group is 2.

NOTE 2: Internal coordinate derivatives are in Kcal/Angstrom or Kcal/radian. Values of less than about 0.2 are quite acceptable.

NOTE 3: In larger calculations, the time estimates are useful. In practice they are pessimistic, and only about 70% of the time estimated will be used, usually. 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 4: 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 about 0.0050, and Rx, Ry, and Rz less than about 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.

NOTE 5: These are the "genuine" force constants: there will normally be 3N-6 of these.

NOTE 6: The transition dipoles are useful in determining the intensity of I.R. absorption bands. The reduced mass indicates which atoms are involved in a vibration; thus, vibration 10 principally involves carbon and oxygen.

NOTE 7: These are the normal modes of vibration. The last three, five or six are the "trivial" vibrations, and should be small.

NOTE 8: In order, the thermodynamic quantities calculated are:

(1) The vibrational contribution,

(2) The rotational contribution,

(3) The sum of (1) and (2), this gives the internal contribution,

(4) The translational contribution.

For partition functions the various contributions are multiplied together.

5.3 EXAMPLE OF 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:    C
Line  5:    C    1.479146 1
Line  6:    H    1.109475 1  111.328433 1
Line  7:    H    1.109470 0  111.753160 0  120.000000 0   2  1  3
Line  8:    H    1.109843 0  110.103163 0  240.000000 0   2  1  3
Line  9:    H    1.082055 0  121.214083 0   60.000000 -1  1  2  3
Line 10:    H    1.081797 0  121.521232 0  180.000000 0   1  2  3
Line 11:    H    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.

CHAPTER 6

BACKGROUND

6.1 INTRODUCTION

While all the theory used in MOPAC is in the literature, so that in principle one to could read and understand the algorithm, many parts of the code involve programming concepts or constructions which, while not of sufficient importance to warrant publication, are described here in order to facilitate understanding.

6.2 CONVERGENCE TESTS IN SUBROUTINE ITER

Self-Consistency Test

The SCF iterations are stopped when two tests are satisfied. These are (1) when 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 SELCON, and (2) the difference in density matrix elements on two successive iterations falls below a preset limit, which is a multiple of SELCON.

SELCON is set initially to 0.00001 kcal/mole; this can be made 100 times smaller by specifying PRECISE or FORCE. It can be over-ridden by explicitly defining the SCF criterion via SCFCRT=0.0000001.

SELCON is further modified by the value of the gradient norm, if known. If GNORM is large, then a more lax SCF criterion is acceptable, and SCFCRT can be relaxed up to 50 times its default value. As the gradient norm drops, the SCF criterion returns to its default value.

The SCF test is performed using the energy calculated from the Fock matrix which arises from a density matrix, and not from the density matrix which arises from a Fock. In the limit, the two energies would be identical, but the first converges faster than the second, without loss of precision.

6.3 CONVERGENCE IN SCF CALCULATION

A brief description of the convergence techniques used in subroutine ITER follows.

ITER, the SCF calculation, employs six methods to achieve a self-consistent field. In order of usage, these are:

(a) Intrinsic convergence by virtue of the way the calculation is carried out. Thus a trial Fock gives rise to a trial density matrix, which in turn is used to generate a better Fock matrix.

This is normally convergent, but many exceptions are known. The main situations when the intrinsic convergence does not work are:

> (1) A bad starting density matrix. This normally occurs when the default starting density matrix is used. This is a very crude approximation, and is only used to get the calculation started. A large charge is generated on an atom in the first iteration the second iteration overcompensates, and an oscillation is generated.
>
> (2) The equations are only very slowly convergent. This can be due to a long-lived oscillation or to a slow transfer of charge.

(b) Oscillation damping. If, on any two consecutive iterations, a density matrix element changes by more than 0.05, then 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 the old density matrices are not filled until after iteration 2. For this reason they are not used in the damping before iteration 3.

(c) Three-point interpolation of the density matrix. Subroutine CNVG monitors the number of iterations, and if this is exactly divisible by three, and certain other conditions relating to the density matrices are satisfied, a three-point interpolation is performed. This is the default converger, and is very effective with normally convergent calculations. It fails in certain systems, usually those where significant charge buildup is present.

(d) Energy-level shift technique. The virtual M.O. energy levels are shifted to more positive energy. This has the effect of damping oscillations, and intrinsically divergent equations can often be changed to intrinsically convergent form. (Invoked by the keyword SHIFT=n.nnn)

(e) Pulay's method. If requested, when the largest change in density matrix elements on two consecutive iterations has dropped below 0.1, then routine CNVG is abandoned in favour of a multi-Fock matrix interpolation. This relies on the fact that the eigenvectors of the density and Fock matrices are identical at self-consistency, so [P.F]=0 at SCF. The extent to which this condition does not occur is a measure of the deviance from self-consistency. Pulay's method uses this relationship to calculate that linear combination of Fock matrices which minimize [P.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. (invoked by the keyword PULAY)

(f) The Camp-King converger. If all else fails, the Camp-King converger is just about guaranteed to work every time. However, it is very expensive in time, and therefore is invoked as a last resort.

It evaluates that linear combination of old and current eigenvectors which minimize the total energy. One of its strengths is that systems which otherwise oscillate due to charge surges, e.g. CHO-H, the C-H distance being very large, will converge using this very sophisticated converger.

6.4 CAUSES OF FAILURE TO ACHIEVE A SCF

In a system where a biradical can form, such as ethane decomposing into two CH3 units, the normal RHF procedure can fail to go self-consistent. If the system has marked biradicaloid character, then BIRADICAL or UHF and TRIPLET can often prove successful. These options rely on the assumption that two unpaired electrons can represent the open shell part of the wave-function.

Consider H-Cl, with the interatomic distance being steadily increased. At first the covalent bond will be strong, and a self-consistent field is readily obtained. Gradually the bond will become more ionic, and eventually the charge on chlorine will become very large. The hydrogen, meanwhile, will become very electropositive, and there will be an increased energy advantage to any one electron to transfer from chlorine to hydrogen. If this in fact occurred, the hydrogen would suddenly become very electron-rich and would, on the next iteration, lose its extra electron to the chlorine. A sustained oscillation would then be initiated. To prevent this, if BIRADICAL is specified, exactly one electron will end up on hydrogen. A similar result can be obtained by specifying TRIPLET in a UHF calculation.

6.5 TORSION ANGLE COHERENCY

MOPAC calculations do not distinguish between enantiomers, 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 MOPAC 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 353.

6.6 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, 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 is examined to see if there is a large relative motion between them. By bonded is meant within the Van der Waals' distance. 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.

6.7 REACTION COORDINATES

The Intrinsic Reaction Coordinate method pioneered and developed by Mark Gordon has been incorporated in a modified form into MOPAC. As this facility is quite complicated all the key-words associated with the IRC have been grouped together in this section.

Definitions of Terms

DRC

The Dynamic Reaction Coordinate is the path followed by all the atoms in a system assuming conservation of energy, i.e., as the potential energy changes the kinetic energy of the system changes in exactly the opposite way so that the total energy (kinetic plus potential) is a constant. If started at a ground state geometry, no significant motion should be seen. Similarly, starting at a transition state geometry should not produce any motion - after all it is a stationary point and during the lifetime of a calculation it is unlikely to accumulate enough momentum to travel far from the starting position.

In order to calculate the DRC path from a transition state, either an initial deflection is necessary or some initial momentum must be supplied.

Because of the time-dependent nature of the DRC the time elapsed since the start of the reaction is meaningful, and is printed.

IRC

The Intrinsic Reaction Coordinate is the path followed by all the atoms in a system assuming all kinetic energy is completely lost at every point, i.e., as the potential energy changes the kinetic energy generated is annihilated so that the total energy (kinetic plus potential) is always equal to the potential energy only.

The IRC is intended for use starting with the transition state geometry. A normal coordinate is chosen, usually the reaction coordinate, and the system is displaced in either the positive or negative direction along this coordinate. The internal modes are obtained by calculating the mass-weighted Hessian matrix in a force calculation and translating the resulting cartesian normal mode eigenvectors to conserve momentum. That is, the initial cartesian coordinates are displaced by a small amount proportional to the eigenvector coefficients plus a translational constant; the constant is required to ensure that the total translational momentum of the system is conserved as zero. At the present time there may be small residual rotational components which are not annihilated; these are considered unimportant.

General Description of the DRC and IRC.

As the IRC usually requires a normal coordinate a force constant calculation normally to be done first. If IRC is specified on its own a normal coordinate is not used and the IRC calculation is performed on the supplied geometry.

A recommended sequence of operations to start an IRC calculation is as follows:

1. Calculate the transition state geometry. If the T/S is not first optimized, then the IRC calculation may give very misleading results. For example, if NH3 inversion is defined as the planar system but without the N-H bond length being optimized the first normal coordinate might be for N-H stretch rather than inversion. In that case the IRC will relax the geometry to the optimized planar structure.

2. Do a normal FORCE calculation, specifying ISOTOPE in order to save the FORCE matrices. Do not attempt to run the IRC directly unless you have confidence that the FORCE calculation will work as expected. If the IRC calculation is run directly, specify ISOTOPE anyway: that will save the FORCE matrix and if the calculation has to be re-done then RESTART will work

correctly.

3. Using IRC=n and RESTART run the IRC calculation. If RESTART is specified with IRC=n then the restart is assumed to be from the FORCE calculation. If RESTART is specified without IRC=n, say with IRC on its own, then the restart is assumed to be from an earlier IRC calculation that was shut down before going to completion.

A DRC calculation is simpler in that a force calculation is not a prerequisite, however most calculations of interest normally involve use of an internal coordinate. For this reason IRC=n can be combined with DRC to give a calculation in which the initial motion (0.3Kcal worth of kinetic energy) is supplied by the IRC, and all subsequent motion obeys conservation of energy. The DRC motion can be modified in three ways.

1. It is possible to calculate the reaction path followed by a system in which the generated kinetic energy decays with a finite half-life. This can be defined by DRC=n.nnn, where n.nnn is the half-life in femtoseconds. If n.nn is 0.0 this corresponds to infinite damping simulating the IRC. A limitation of the program is that time only has meaning when DRC is specified without a half-life.

2. Excess kinetic energy can be added to the calculation by use of KINETIC=n.nn. After the kinetic energy has built up to 0.2Kcal/mole or if IRC=n is used then n.nn Kcal/mole of kinetic energy is added to the system. The excess kinetic energy appears as a velocity vector in the same direction as the initial motion.

3. The RESTART file <filename>.RES can be edited to allow the user to modify the velocity vector or starting geometry. This file is formatted.

Frequently DRC leads to a periodic, repeating orbit. One special type - the orbit in which the direction of motion is reversed so that the system retraces its own path - is sensed for and if detected the calculation is stopped after exactly one cycle. If the calculation is to be continued, the key-word GEO-OK will allow this check to be by-passed.

Due to the potentially very large output files that the DRC can generate extra key-words are provided to allow selected points to be printed. After the system has changed by a preset amount the following key-words can be used to invoke a print of the geometry.

Key-Word	Default	User Specification
X-PRIO	0.05 Angstroms	X-PRIORITY=n.nn
T-PRIO	0.10 Femtoseconds	T-PRIORITY=n.nn
H-PRIO	0.10 Kcal/mole	H-PRIORITY=n.nn

Option to allow only extrema to be output

In the geometry specification, if an internal coordinate is marked for optimization then when that internal coordinate passes throught an extremum a message will be printed and the geometry output.

Difficulties can arise from the way internal coordinates are processed. The internal coordinates are generated from the cartesian coordinates, so an internal coordinate supplied may have an entirely different meaning on output. In particular the connectivity may have changed. For obvious reasons dummy atoms should not be used in the supplied geometry specification. If there is any doubt about the internal coordinates or if the starting geometry contains dummy atoms then run a 1SCF calculation specifying XYZ. This will produce an ARC file with the "ideal" numbering - the internal numbering system used by MOPAC. Use this ARC file to construct a data file suitable for the DRC or IRC.

Notes

1. Any coordinates marked for optimization will result in only extrema being printed.
2. If extrema are being printed then kinetic energy extrema will also be printed.

Key-words for use with the IRC and DRC

1. Setting up the transition state: NLLSQ SIGMA.
2. Constructing the FORCE matrix: FORCE or IRC=n, ISOTOPE, LET.
3. Starting an IRC: RESTART and IRC=n, T-PRIO, X-PRIO, H-PRIO.
4. Starting a DRC: DRC or DRC=n.nn, KINETIC=n.nn.
5. Starting a DRC from a transition state: (DRC or DRC=n) and IRC=n, KINETIC=n.
6. Restarting an IRC: RESTART and IRC.
7. Restarting a DRC: RESTART and (DRC or DRC=n.nn).
8. Restarting a DRC starting from a transition state: RESTART and (DRC or DRC=n.nn).

Other key-words, such as T=nnn or GEO-OK can be used anytime.

Examples of DRC/IRC data

Use of the IRC/DRC facility is quite complicated. In the following examples various "reasonable" options are illustrated for a calculation on water.

It is assumed that an optimized transition-state geometry is available.

Example 1: A Dynamic Reaction Coordinate, starting at the transition state for water inverting, initial motion opposite to the transition normal mode, with 6kcal of excess kinetic energy added in. Every point calculated is to be printed (Note all coordinates are marked with a zero, and T-PRIO, H-PRIO and X-PRIO are all absent). The results of an earlier calculation using the same key-words is assumed to exist. The earlier calculation would have constructed the force matrix. While the total cpu time is specified, it is in fact redundant in that the calculation will run to completion in less than 600 seconds.

```
 KINETIC=6 RESTART  IRC=-1 DRC T=600
    WATER

   H    0.000000  0     0.000000  0     0.000000  0    0  0  0
   O    0.911574  0     0.000000  0     0.000000  0    1  0  0
   H    0.911574  0   180.000000  0     0.000000  0    2  1  0
   0    0.000000  0     0.000000  0     0.000000  0    0  0  0
```

Example 2: An Intrinsic Reaction Coordinate calculation. Here the restart is from a previous IRC calculation which was stopped before the minimum was reached. Recall that RESTART with IRC=n implies a restart from the FORCE calculation. Since this is a restart from within an IRC calculation the keyword IRC=n has been replaced by IRC. IRC on its own (without the "=n") implies an IRC calculation from the starting position - here the RESTART position - without initial displacement.

```
RESTART  IRC  T=600
WATER

H    0.000000 0    0.000000 0    0.000000 0   0  0  0
O    0.911574 0    0.000000 0    0.000000 0   1  0  0
H    0.911574 0  180.000000 0    0.000000 0   2  1  0
0    0.000000 0    0.000000 0    0.000000 0   0  0  0
```

6.8 SPARKLES

Four extra "elements" have been put into MOPAC. 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 MOPAC.

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 Angstroms. Any two sparkles of opposite sign will form an ion-pair with a interatomic separation of 1.4A.

3. They have a zero heat of atomization, no orbitals, and no ionization potential.

They can be regarded as unpolarizable ions of diameter 1.4A. 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, if the alkaline metal salt of formic acid was run, the formula would be:

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 H_2O+, 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, and a better description would be of an ion, of diameter 1.4A, and the charge delocalized over its surface. Computationally, a sparkle is an integer charge at the center of a repulsion sphere of form exp(-alpha*r). The hardness of the sphere is such that other atoms or sparkles can approach within about 2 Angstroms quite easily, but only with great difficulty come closer than 1.4A.

Uses of Sparkles

1. They can be used as counterions, e.g. for acid anions or for cations. Thus, if the ionic form of an acid is wanted, then the moieties H.X, H.-, and +.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 Angstroms. If ++ and -- are used, the separation would be D/9.606 Angstroms. If the inter-sparkle separation is less than 1.0 Angstroms (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 cubic Angstroms of a molecule in any particular orientation can then easily be calculated.

6.9 MECHANISM OF THE 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.

6.10 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 Ne electrons, there will be No = Ne /2 occupied molecular orbitals and Nvir = N-Nv 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 Fov can then be annihilated by a series of 2x2 rotations in the manner of Jacobi. Unlike Jacobi, however, the method need not be cycled to exactly diagonalize Fov; 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 2x2 rotations can be ignored. In an exact diagonalization the off-diagonal matrix elements formed by an elementary 2 by 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 Fov. 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 by 2 rotation will, of course, not eliminate fully even those elements which we do choose to operate on, but again the pseudoeigenvalue 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 by 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 a 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 should be 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 Fov matrix. In virtual memory computers this would normally not present any problem. A criterion is also needed to decide when to perform a 2 by 2 rotation and elimination. After a few trial calculations, it was found to be efficient to eliminate all off-diagonal elements whose modulus was 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 the new method rapidly converge. An interesting side-effect of the new 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.

6.11 DYNAMIC REACTION COORDINATE

Introduction

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, there being no damping of the kinetic energy allowed, and (b) liquid phase, in which kinetic energy is always set to zero, the motion of the atoms being infinitely damped.

All possible degrees of damping are allowed. In addition, the facility exists to dump energy into the system, appearing as kinetic energy. 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 0.2Kcal/mole.

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

Equations Used

Force acting on any atom

$$g(i) + g'(i)t + g''(i)t**2 = dE/dx(i) + d**2E/dx(i)**2 + d**3E/dx(i)**3$$

Acceleration due to force acting on each atom

$$a(i) = (g(i)+g'(i)t+g''(i)t**2)/M(i)$$

New velocity

$$V(o)+Dt*g(i)/M(i)+1/2*Dt**2*g'(i)/M(i)+/3*Dt**3*g''(i)/M(i)$$

or

$$V(i) = V(i) + V'(i)t + V''(i)t**2 + V'''(i)t**3$$

That is, the change in velocity is equal to the integral over the time interval of the acceleration.

New position of atoms

$$X(i) = X(o) + V(o)t + 1/2*V't**2 + 1/3*V''t**3 + 1/4*V'''t**4$$

That is, the change in position is equal to the integral over the time interval of the velocity.

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.

Even with all this, errors creep in and 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 0.2Kcal/mole.

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.

6.12 CONFIGURATION INTERACTION

MOPAC contains a very large Multi-Electron Configuration Interaction calculation, 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 an expert has available an exhaustive set of key-words from which a specific C.I. can be tailored.

A MECI calculation involves the interaction of microstates representing specific permutations of electrons in a set of M.O.s. Starting with a set electronic configuration, either closed shell or open shell, but unconditionally restricted Hartree-Fock, the first step in a MECI calculation is the removal from the M.O.s of the electrons 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, and the spin-states of the state functions calculated.

General Overview of Key-Words

Key-words associated with the operations of MECI are:

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

Each key-word may imply others; thus TRIPLET implies an open-shell system, therefore OPEN(2,2), and C.I.=2 are implied, if not user specified.

Starting Electronic Configuration

MECI is restricted to RHF calculations, but with that single restriction any starting configuration will be supported. Examples of starting configurations would be

System	Key-Words 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 starting 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 key-words define the same starting configuration. This can be verified by monitoring the convergance using PL, for which both key-words give the same SCF energy.

Removal of Electrons from Starting Configuration

For a starting configuration of alpha M.O. occupancies O(i), O(i) being in the range 0.0 to 1.0, the energies of the M.O.s involved in the MECI can be calculated from

$$E(i) = Sum(j)(2J(i,j)-K(i,j))O(j)$$

where J(i,j) and K(i,j) 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 MECI only. Most MECI calculations will involve between 1 and 5 M.O.s, so a system with about 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 arbitary zero of energy in a MECI calculation is the starting 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

```
GSE = Sum(i)[ E(i)O(i) + J(i,i) * O(i)*O(i)
    + Sum(j<i)( 2(2J(i,j) - K(i,j) ) * O(i)*O(j) ) ]
```

Formation of Microstate Configuration

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

Microstates for 5 electrons in 5 M.O.s

	Electron Configuration				Electron Configuration		
	Alpha 1 2 3 4 5	Beta 1 2 3 4 5	M(s)		Alpha 1 2 3 4 5	Beta 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 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 MECI, then all possible microstates which give rise to a 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

(1,1) = 1

(0,2) = 0

(1,2) = 10
01

(1,3) = 100
010
001

(2,3) = 110
101
011

(2,4) = 1100
1010
1001
0110
0101
0011

(1,4) = 1000
0100
0010
0001

(3,5) = 11100
11010
11001
10110
10101
10011
01110
01101
01011
00111

(2,5) = 11000
10100
10010
10001
01100
01010
01001
00110
00101
00011

Sets of Microstates for Various MECI 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 MECI 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 component of spin = 0, will thus be the positive combination.

$$(\bar{1})(2) \quad + \quad (1)(\bar{2})$$

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. E(p,p) is given simply by

```
E(p,p)=Sum(i)Oa(i,p)*[Eig(i) + Sum(1/2(<ii£jj>-<ij£ij>)*Oa(j,p)
                           +             <ii£jj>          *Ob(j,p))

      + Sum(i)Ob(i,p)*[Eig(i) + Sum(1/2(<ii£jj>-<ij£ij>)*Ob(j,p))
```

   Oa(i,p) = Occupancy of alpha M.O. i in Microstate p
   Ob(i,p) = 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.

```
E(p,q) = Sum(k) [ <ij£kk> - <ik£jk> ) * (Occa(k) - Occg(k))
                + <ij£kk>             * (Occb(k) - Occg(k)]
```

E(p,q) 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.

Where Occa(k) is the alpha molecular orbital occupancy in the configuration interaction.

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:

1. 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

   Q(p,q) = <ik£jl> - <il£jk>
   E(p,q) 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.

2. Both M.O.s have beta spin:

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

3. One M.O. has alpha spin, and one beta spin:

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

   Q(p,q) = <ik£jl>

   E(p,q) 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 MECI calculation invoked by use of the key-word C.I.=n normally gives rise to states of quantized spins. When C.I. is used without any other modifying key-words, the following states will be obtained.

No. of M.O.s	States Arising From Odd Electron Systems			States Arising From Even Electron Systems		
in MECI	Doublets			Singlets	Triplets	
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 of spin states 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 S2 is calculated.

S2 = S(S+1) = Sz**2 + 2*S(+)S(-)

= Ne -
Sum(i) [C(i,k)*C(i,k)*(1/4*(Na(i)-Nb(i))**2
+ Sum(l) Oa(l,i)*Ob(l,i))
+Sum(j) 2[C(i,k)*C(j,k)*(Kronekerdelta(C(i,k)(S(+)S(-))C(j,k)]]

Where Ne = No. of electrons in C.I.
C(i,k) = Coefficient of Microstate i in State k
Na(i) = Number of alpha electrons in Microstate i
Nb(i) = Number of beta electrons in Microstate i
Oa(l,k) = Occupancy of alpha M.O. l in Microstate k
Ob(l,k) = Occupancy of beta M.O. l in Microstate k
S(+) = Spin shift up or step up operator
S(-) = Spin shift down or step down operator

The Kronekerdelta is 1 if the two terms in brackets following it are identical.

The spin state is calculated from S = 1/2 (Sqrt(1+4*S2) - 1)

In practice, S is calculated to be exactly integer, or half integer. That is, there is insignificant error due to approximations used. This does not mean, however, that the method is accurate. The spin calculation is completely precise, in the group theoretic sense, but the accuracy of the calculation is limited by the Hamiltonian used, a space-dependent function.

Choice of State to be Optimized

MECI 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, when used on its own, select the n'th state, irrespective of its total spin. By default n=1. If ROOT=n is used in conjunction with a key-word 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 4th 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.

6.13 REDUCED MASSES IN A FORCE CALCULATION

Reduced masses for a diatomic are given by

$$\frac{(mass1) * (mass2)}{(mass1) + (mass2)}$$

For a Hydrogen molecule the reduced mass is thus 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.

6.14 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:   C    0.000000 0    0.000000 0    0.000000 0   0  0  0
Line  5:   C    1.479146 1    0.000000 0    0.000000 0   1  0  0
Line  6:   H    1.109475 1  111.328433 1    0.000000 0   2  1  0
Line  7:   H    1.109470 1  111.753160 1  120.288410 1   2  1  3
Line  8:   H    1.109843 1  110.103163 1  240.205278 1   2  1  3
Line  9:   H    1.082055 1  121.214083 1   38.110989 1   1  2  3
Line 10:   H    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:   C    0.000000 0    0.000000 0    0.000000 0   0  0  0
Line 13:   C    1.479146 1    0.000000 0    0.000000 0   1  0  0
Line 14:   H    1.109475 1  111.328433 1    0.000000 0   2  1  0
Line 15:   H    1.109470 1  111.753160 1  120.288410 1   2  1  3
Line 16:   H    2.109843 1   30.103163 1  240.205278 1   2  1  3
Line 17:   H    1.082055 1  121.214083 1   38.110989 1   1  2  3
Line 18:   H    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 length 3N-6 defining the difference between the two geometries is calculated.

3. The scalar P of the difference vector is reduced by some fraction, normally about 5 to 15 percent.

4. Identify the geometry of lower energy; call this G.

5. Optimize G, subject to the constraint that it maintains a constant distance P from the other geometry.

6. If the newly-optimized geometry is higher in energy then the other geometry, then go to 1. If it is higher, and the last two steps involved the same geometry moving, make the other geometry G without modifying P, and go to 5.

7. Otherwise go back to 2.

The mechanism of 5 involves the coordinates of the moving geometry being perturbed by an amount equal to the product of the discrepancy between the calculated and required P and the vector R.

As the specification of the geometries is quite difficult, in that the difference vector depends on angles (which are, of necessity ill-defined by 360 degrees) SADDLE can be made to run in cartesian coordinates using the key-word XYZ. If this option is chosen then the initial steps of the calculation are as follows:

1. Both geometries are converted into 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.

LIMITATIONS:

The two geometries must be related by a continuous deformation of the coordinates. By default, internal coordinates are used in specifying geometries, and while bond lengths and bond angles are unambiguously defined (being 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 being used, a completely incorrect vector was used, with disastrous results.

To correct this, ensure that one geometry can be obtained from the other by a continuous deformation, or use the XYZ option.

6.15 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, use is made of an alternating electric field. If the molecule has an intrinsic dipole, then the molecule will be stabilized in one direction. When the field is reversed, the molecule will be destabilized, but, on averaging the two effects, the result is a net stabilization due only to the induced dipole.

MOPAC 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 a*l, a being the cube root of 1/2. At the origin, the electric field gradient in volts per metre due to the charges is as follows.

Due to +Q charge: -QC/(l*l)

Due to -Q/2 charge: (Q/2)*C/(a*l*a*l)

Therefore total field: C*Q*(1-1/2**(1/3)) / (l*l).

(C has the value 1.6029*10**(-19) Coulombs)

The slope of the total gradient (that is, the addition of the two gradients) will be zero.
Differentiating -C*Q/(l*l) + (Q/2)*C/(a*l*a*l) with respect to x gives

+C*Q/(2*l*l*l) - (Q/2)*C/(2*a*l*a*l*a*l)

= +(Q/2)*C/(l*l*l) - (Q/2)*C/(l*l*l) = 0.

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 -a*l 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 therefore

E = C*Q*2(1-1/2**(-1/3))/(l*l*4*pi*E(o)),

the 4*pi*E(o) 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.

Heat = H.o.F + V*E(Charge) + dV/dx*E(Dipole)

+ d2V/dx2*E(polarizability)

That is, 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 = (2/23.061)*d**2H/dE**2

The second derivitive of H with respect to E is given by

d**2H/dE**2 = (H(E)+H(-E)-2*H(0))/(2*E),

H(E) being the heat of formation in the electric field.

The polarizability volume, Vol, is calculated from the polarizability by

Vol=P/(E*4*pi*E(o)) = 2/(E*23.061*4*pi*E(o)) * d**2H/dE**2

Substituting for E we have

Vol=2*1**4*pi*E(o)/(23.061*Q*Q*C*C*(1-1/2**(-1/3))**2) * 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(i,j), 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 2	+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, in Kcal/mole, the polarizability can be calculated in units of cubic angstroms via

```
Vol = (Heat(2)+Heat(4)-Heat(1)-Heat(3))*(1*1*1*1)*2*pi*Eo
          23.061 * (1-a) * (1-a) *  Q * Q * C * C
```

```
1eV is 1.60219 * 10-19 Joules

Eo is 8.854188 * 10-12 Joules**(-1).C**2.M**(-1)

or 8.854188 * 10-22 Joules**(-1).C**2.Angstroms**(-1)

Vol = (eV * 1**4 * J**(-1) * C**2 * M**(-1))
        C**2

Vol = 2 * 3.1415926 * 8.854188*10-22 / (23.061 * 1.60219*10-19)

=(Heat(2)+Heat(4)-Heat(1)-Heat(3))*0.0015056931*(1*1*1*1)
               (1-a)*(1-a)*Q*Q
```

Monopolar and dipolar terms are eliminated in this treatment.

Finally, monatomic additive terms are included when MNDO is used.

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

6.16 SOLID STATE CAPABILITY

Currently MOPAC can only handle up to one-dimensional extended systems, but work is under way to extend it to two and three dimensions. As the solid-state method used is unusual, details are given at this point.

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 visualised as being done entirely in reciprical 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" 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 call has an insignificant overlap with the atomic orbitals at the ends of the unit cell. In practice, a translation vector of more that about 7 or 8 Angstroms 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 Angstroms is needed.

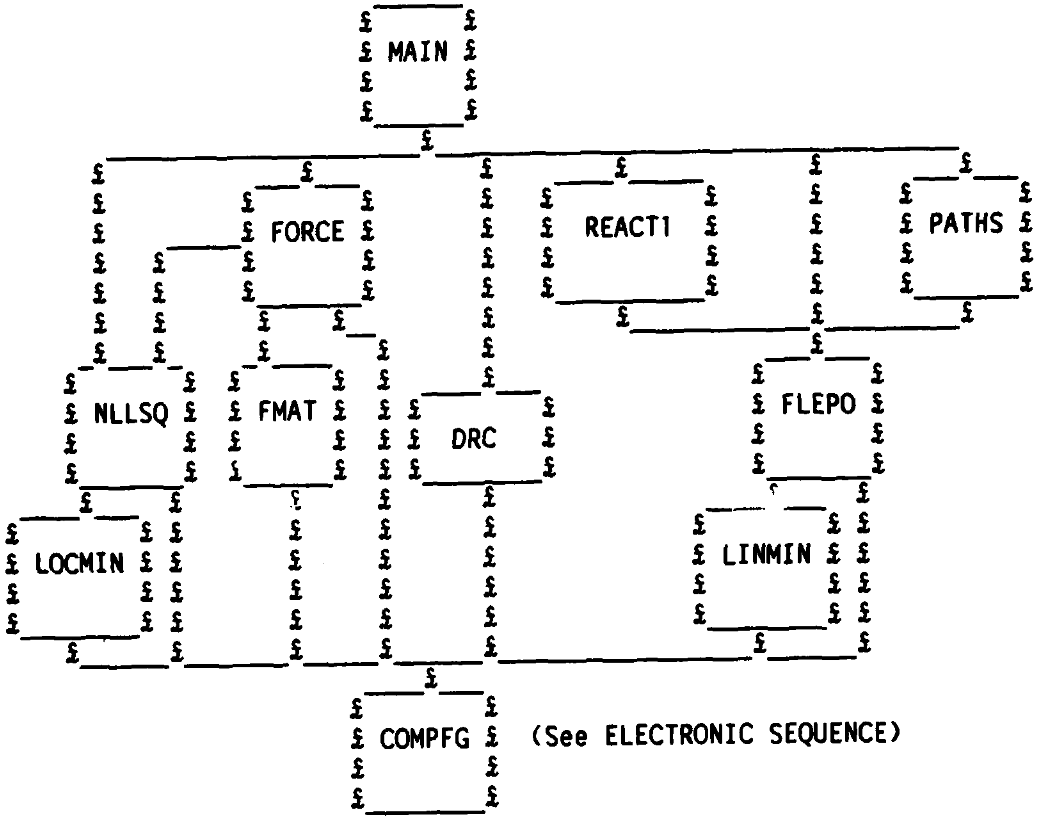
CHAPTER 7

PROGRAM

The logic within MOPAC is best understood by use of flow-diagrams.

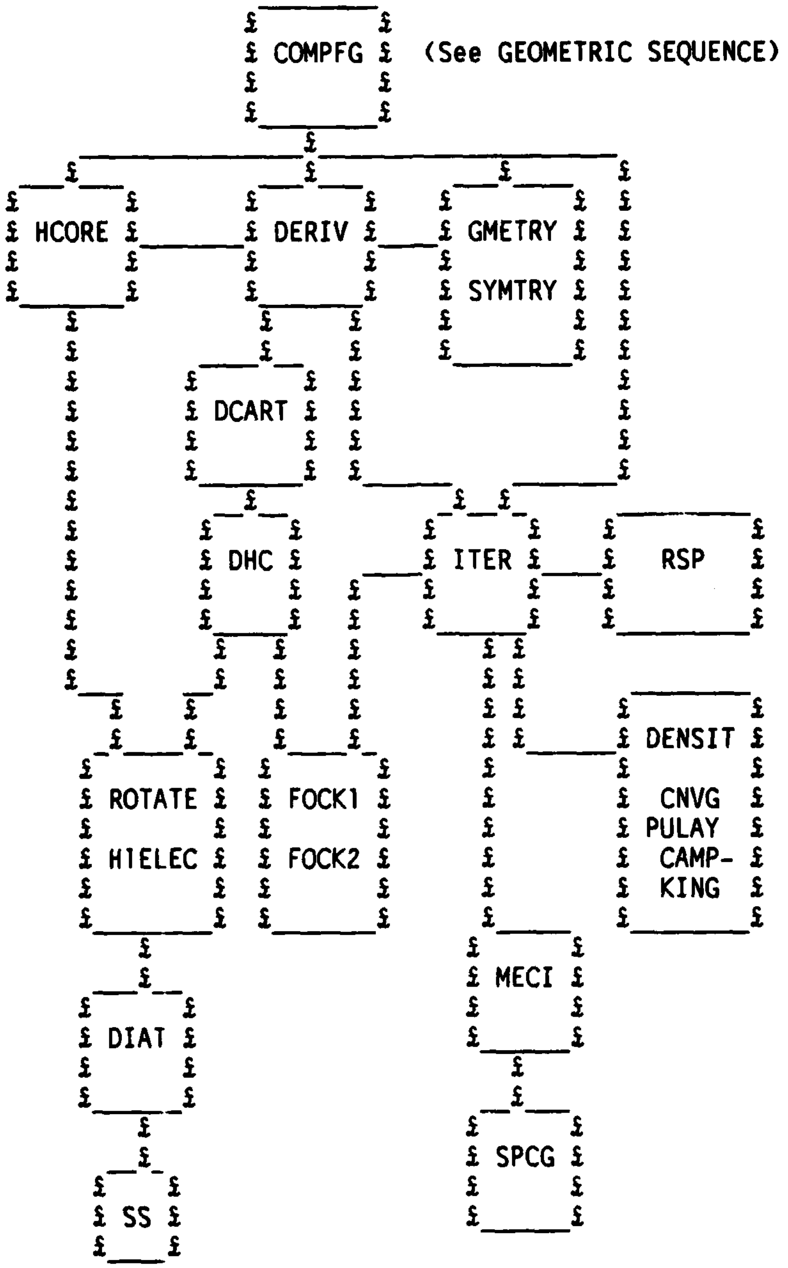
There are two main sequences, geometric and electronic. These join only at one common subroutine COMPFG. It is possible, therefore, to understand the geometric or electronic sections in isolation, without having studied the other section.

7.1 MAIN GEOMETRIC SEQUENCE

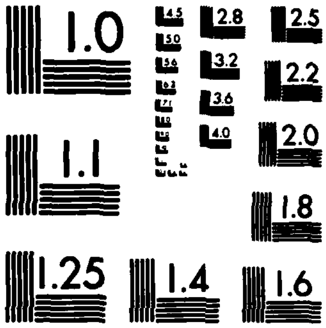


(See ELECTRONIC SEQUENCE)

7.2 MAIN ELECTRONIC FLOW


(See GEOMETRIC SEQUENCE)

AD-A172 100 MOPAC MANUAL (VERSION 310) A GENERAL MOLECULAR ORBITAL 2/2
PACKAGE -- CRAY-XMP VERSION(U) FRANK J SEILER RESEARCH
LAB UNITED STATES AIR FORCE ACADEMY C.. J J STEWART
UNCLASSIFIED JUN 86 FJSRL-TR-86-0008 F/G 7/4 NL


1.0
4.5
5.0
5.6
6.3
7.1
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

7.3 CONTROL WITHIN MOPAC

Almost all the control information is passed via the single datum "KEYWRD", a string of 80 characters, which is read in at the start of the job.

Each subroutine is made independent, as far as possible, even at the expense of extra code or calculation. Thus, for example, the SCF criterion is set in subroutine ITER, and nowhere else. Similarly subroutine DERIV has exclusive control of the step size in the finite-difference calculation of the energy derivatives. If the default values are to be reset, then the new value is supplied in KEYWRD, and extracted via INDEX and READA. The flow of control is decided by the presence of various key-words in KEYWRD.

When a subroutine is called, it assumes that all data required for its operation are available in either common blocks or arguments. Normally no check is made as to the validity of the data received. All data are "owned" by one, and only one, subroutine. By ownership is implied the permission and ability to change the data. Thus MOLDAT "owns" the number of atomic orbitals, in that it calculates this number, and stores it in the variable NORBS. Many subroutines use NORBS, but none of them is allowed to change it. For obvious reasons no exceptions should be made to this rule. To illustrate the usefulness of this convention, consider the eigenvectors, C and CBETA. These are owned by ITER. Before ITER is called, C and CBETA are not calculated, after ITER has been called C and CBETA are known, so any subroutine which needs to use the eigenvectors can do so in the certain knowledge that they exist.

Any variables which are only used within a subroutine are not passed outside the subroutine unless an overriding reason exists. This is found in PULAY and CNVG, among others where arrays used to hold spin-dependent data are used, and these cannot conveniently be defined within the subroutines. In these examples, the relevant arrays are "owned" by ITER.

A general subroutine, of which ITER is a good example, handles three kinds of data: First, data which the subroutine is going to work on, for example the one and two electron matrices; second, data necessary to manipulate the first set of data, such as the number of atomic orbitals; third, the calculated quantities, here the electronic energy, and the density and Fock matrices.

Reference data are entered into a subroutine by way of the common blocks. This is to emphasize their peripheral role. Thus the number of orbitals, while essential to ITER, is not central to the task it has to perform, and is passed through a common block.

Data the subroutine is going to work on are passed via the argument list. Thus the one and two electron matrices, which are the main reason for ITER's existence, are entered as two of the four arguments. As ITER does not own these matrices it can use them but may not change their contents. The other arguments are EE, the electronic energy, and FULSCF, a logical. EE is owned by ITER even though it first appears before ITER is called. FULSCF, on the other hand, is not owned by ITER, and is used, but not changed.

Sometimes common block data should more correctly appear in an argument list. This is usually not done in order to prevent obscuring the main role the subroutine has to perform. Thus ITER calculates the density and Fock matrices, but these are not represented in the argument list as the calling subroutine never needs to know them; instead, they are stored in common.

SUBROUTINE GMETRY: Description for programmers.

GMETRY has two arguments, GEO and COORD. On input GEO contains either (a) internal coordinates or (b) cartesian coordinates. On exit COORD contains the cartesian coordinates.

The normal mode of usage is to supply the internal coordinates, in which case the connectivity relations are found in common block GEOKST.

If the contents of NA(1) is zero, as required for any normal system, then the normal internal to cartesian conversion is carried out.

If the contents of NA(1) is 99, then the coordinates found in GEO are assumed to be cartesian, and no conversion is made. This is the situation in a FORCE calculation.

A further option exists within the internal to cartesian conversion. If STEP, stored in common block REACTN, is non-zero, then a reaction path is assumed, and the internal coordinates are adjusted radially in order that the "distance" in internal coordinate space from the geometry specified in GEO is STEPP away from the geometry stored in GEOA, stored in REACTN.

During the internal to cartesian conversion, the angle between the three atoms used in defining a fourth atom is checked to ensure that it is not near to 0 or 180 degrees, if it is near to these angles, then there is a high probability that a faulty geometry will be generated and to prevent this the calculation is stopped and an error message printed.

NOTE 1: If the angle is exactly 0 or 180 degrees, then the calculation is not terminated: This is the normal situation in a high-symmetry molecule such as propyne.

NOTE 2: The check is only made if the fourth atom has a bond angle which is not zero or 180 degrees.

CHAPTER 8

ERROR MESSAGES PRODUCED BY MOPAC

MOPAC 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. To correct this, identify geometries corresponding to points on each side of the T/S. (Two geometries on the output separated by the message "SWAPPING...") and make up a new data-file using these geometries. This corresponds to points on the reaction path near to the T/S. Run a new job using these two geometries, but with BAR set to a third or a quarter of its original value, e.g. BAR=0.05. This normally allows the T/S to be located.

C.I. NOT ALLOWED WITH UHF

There is no UHF configuration interaction calculation in MOPAC. 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, this message will not be triggered. The good news is that the criterion used to trigger this message was set too coarsely. The criterion has been tightened so that this message now does not often appear. Geometric integrity does not appear to be compromized.

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 MOPAC all sorts of weird and wonderful systems are calculated, systems the authors of the convergers never dreamed of. MOPAC 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 involking 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. Most likely line 4 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.

THE LINE MINIMISATION 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 badly wrong, and you should then suspect MOPAC is faulty in some way.

**** MAX. NUMBER OF ATOMS ALLOWED:....

At compile time the maximum sizes of the arrays in MOPAC 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 MOPAC should be re-compiled and re-linked.

**** MAX. NUMBER OF ORBITALS:....

At compile time the maximum sizes of the arrays in MOPAC 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 MOPAC should be re-compiled and re-linked.

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

At compile time the maximum sizes of the arrays in MOPAC 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 MOPAC should be re-compiled and re-linked.

NAME NOT FOUND

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

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

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.

Try using SIGMA instead of POWSQ.

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 MOPAC, 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 key-words line or 3,600 seconds, if no time was specified, is likely to be exceeded if another cycle of calculation were to be 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.

CHAPTER 9

CRITERIA

MOPAC uses various criteria which control the precision of its stages. These criteria are chosen as the best compromise between speed and acceptable errors in the results. The user can override the default settings by use of key-words; however, care should be exercised as increasing a criterion can introduce the potential for infinite loops, and decreasing a criterion can result in unacceptably imprecise results. These are usually characterized by 'noise' in a reaction path, or large values for the trivial vibrations in a force calculation.

9.1 SCF CRITERION

Name:	SCFCRT.	
Defined in	ITER.	
Default value	0.00001 kcal/mole	
Basic Test	Change in energy in kcal/mole on successive iterations is less than SCFCRT.	
Exceptions:	If PRECISE is specified,	SCFCRT=0.0000001
	If non-variational wavefunction	SCFCRT=0.0000001
	If a polarization calculation or gradient minimization	SCFCRT=0.0000001
	If SCFCRT=n.nnn is specified	SCFCRT=n.nnn

Secondary tests: (1) Change in density matrix elements on two successive iterations must be less than 0.001
(2) Change in energy in eV on three successive iterations must be less than 10 x SCFCRT.

9.2 GEOMETRIC OPTIMIZATION CRITERIA

Name:	TOLERX "Test on X Satisfied"
Defined in	FLEPO
Default value	0.0001
Basic Test	The projected change in geometry is less than TOLERX Angstroms.
Exceptions	If PRECISE is specified, TOLERX= 0.00001

Name:	EYEAD "Herbert's Test Satisfied"
Defined in	FLEPO
Default value	0.001
Basic Test	The projected decrease in energy is less than EYEAD Kcals/mole.
Exceptions	If PRECISE is specified, EYEAD=0.00001

Name:	TOLERG "Test on Gradient Satisfied"
Defined in	FLEPO
Default value	1.0
Basic Test	The gradient norm in Kcals/mole/Angstrom is less than TOLERG.
Exceptions	If PRECISE is specified, TOLERG=0.002 If GNORM=n.nnn is specified, TOLERG=n.nnn If a SADDLE calculation, TOLERG is made a function of the last gradient norm.

Name:	TOLERF "Heat of Formation Test Satisfied"
Defined in	FLEPO
Default value	0.002
Basic Test	The calculated heats of formation on two successive cycles differ by less than TOLERF.
Exceptions	If PRECISE is specified, TOLERF=0.00004
Secondary Tests	For the TOLERG, TOLERF, and TOLERX tests, a second test in which individual component of the gradient should be larger than TOLERG divided by the square root of the number of variables must be satisfied.
Other Tests	If, after the TOLERG, TOLERF, or TOLERX test has been satisfied three consecutive times, and the heat of formation has dropped by less than 0.3Kcal/mole, the optimization is stopped.

Name:	TOL2
Defined in	POWSQ
Default value	0.4
Basic Test	The absolute value of the largest component of the gradient is less than TOL2

Exceptions	If PRECISE is specified, TOL2=0.01

Name:	TOLS1
Defined in	NLLSQ
Default Value	0.000 000 000 001
Basic Test	The square of the ratio of the projected change in the geometry to the actual geometry is less than TOLS1.

Name:	<none>
Defined in	NLLSQ
Default Value	0.2
Basic Test	Every component of the gradient is less than 0.2.

CHAPTER 10

DEBUGGING

There are three potential sources of difficulty in using MOPAC, each of which requires special attention. There can be problems with data, due to errors in the data, or MOPAC may be called upon to do calculations for which it was not designed. There are intrinsic errors in MOPAC 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 MOPAC, either to make it compatible with the host computer, or to implement local features.

For whatever reason, the user may need to have access to more information than the normal key-words can provide, and a second set, specifically for debugging, is provided. These key-words give information about the working of individual subroutines, and do not affect the course of the calculation.

10.1 DEBUGGING KEY-WORDS

FULL LIST OF KEY-WORDS FOR DEBUGGING SUBROUTINES
INFORMATION PRINTED

1ELEC	the one-electron matrix.	Note 1
COMPFG	Heat of Formation.	
DCART	Cartesian derivatives.	
DEBUG		Note 2
DEBUGPULAY	Pulay matrix, vector, and error-function.	Note 3
DENSITY	Every density matrix.	Note 1
DERIV	All gradients, and other data in DERIV.	
DFORCE	Print Force Matrix.	
EIGS	All eigenvalues.	
FLEPO	Details of D.F.P. minimization.	
FMAT		
FOCK	Every Fock matrix	Note 1
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 4
SEARCH	Function values, step sizes at all points in SEARCH.	
TIMES	Times of stages within ITER.	
VECTORS	All eigenvectors on every iteration.	Note 1

Note the space before PL: this space is obligatory.

NOTES

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

2. DEBUG is not intended to increase the output, but does allow other key-words to have a special meaning.

3. PULAY is already a key-word, so DEBUGPULAY was an obvious alternative.

4. 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 when investigating options for increasing the rate of convergance of the SCF calculation.

SUGGESTED PROCEDURE FOR LOCATING BUGS

Users are supplied with the source code for MOPAC, and, while the original code is fairly bug-free, after it has been modified there is a possibility that bugs may have been introduced. In these circumstances the author of the changes is obviously responsible for removing the offending bug, and the following ideas might prove useful in this context.

First of all, and most important, before any modifications are done a back-up copy of the standard MOPAC should be made. This will prove invaluable in pinpointing deviations from the standard working. This point cannot be over-emphasized - MAKE A BACK-UP BEFORE MODIFYING MOPAC!!!!

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 (debugging is, of necessity, time consuming) and to minimize output.

The two sets of subroutines in MOPAC, those involved with the electronics and those involved in the geometrics, are kept strictly separate, so 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 geometrics; if faulty, in the electronics.

Bug in the Electronics Subroutines.

Use formaldehyde for this test. The supplied data-file MNRSD1.DAT could be used as a template for this operation. Use key-words 1SCF, DEBUG, and any others necessary.

The main steps are:

(1) Check the starting one-electron matrix and two-electron integral string, using the key-word HCORE. It is normally sufficient to verify that the two hydrogen atoms are equivalent, and that the pi system involves only pz 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.

If the starting matrices are correct, go on to step (2).

(2) Check the density or Fock matrix on every iteration, with the words FOCK or DENSITY. Check the equivalence of the two hydrogen atoms, and the pi system, as in (1).

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 (guessed) 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 RSP.

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, COMPFG 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.

Keep in mind the fact that MOPAC is a large calculation, and while intended to be versatile, many combinations of options have not been tested. If a bug is found in the original code, please communicate details to the Academy, to Dr. James J. P. Stewart, Frank J. Seiler Research Laboratory, U.S. Air Force Academy, Colorado Springs, CO 80840

CHAPTER 11

INSTALLING MOPAC

CRAY-XMP supercomputers are front-ended by a conventional computer, for example a VAX 11-780, to pre-process files to be used. Of necessity, then, the instructions that follow may have to be modified depending on the front-end computer. The following instructions apply to a CRAY computer with a Digital VAX computer (11-786, 11-782, 11-780, 11-750, 11-730) front end.

MOPAC is distributed on a magnetic tape as a set of FORTRAN-77 files, along with ancillary documents such as command, help, data and results files. The format of the tape is that of the VAX-11/780 computer.

1. Put the magnetic tape on the tape drive, write protected.
2. Allocate the tape drive with a command such as $ALLOCATE MTA0:
3. Go into an empty directory which is to hold MOPAC
4. Mount the magnetic tape with the command $MOUNT MTA0: MOPAC
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, called <filename>.CRAY, see Appendix A.
2. The command files CONCAT.COM, MOPAC.COM, PREUP.JOB and INSTALL.JOB.
3. A help file called MOPAC.HLP.

4. A text file MOPAC.MAN.

5. Some test-data files, and results files.

6. There are four parameters within AAAA.CRAY that the user can modify: MAXHEV, MAXLIT, MAXTIM and MAXDMP. MAXLIT is assigned a value equal to the largest number of hydrogen atoms that a MOPAC job is expected to run, MAXHEV is assigned the corresponding number of heavy (non-hydrogen) atoms. The ratio of light to heavy atoms should not be less than 1/2. Some molecular orbital eigenvector arrays are overlapped with Hessian arrays, and to prevent compilation time error messages, the number of allowed A.O.s must be greater than, or equal to three times the number of allowed real atoms. MAXTIM is the default maximum time in seconds a job is allowed to run before either completion or a restart file being written. MAXDMP is the default time in seconds for the automatic writing of the restart files. If your computer is very reliable, and disk space is at a premium, you might want to set MAXDMP as MAXDMP=999999. Alter MAXHEV and MAXLIT in AAAA.CRAY to suit local conditions. Remember that the CRAY is not a virtual machine you may want several copies of MOPAC available, each of different sizes, say 20 heavy atoms, 40, 50 and 60.

7. Modify MOPAC.COM, INSTALL.JOB and PREUP.JOB to suit your CRAY account number and password.

8. Run CONCAT.COM to form MOPAC.CFT. The command here is "@CONCAT", as usual read the COM file first to ensure that you know what will happen.

9. Use the CRAY command language to submit PREUP to the CRAY. This will add to the almost standard FORTRAN-77 program MOPAC.CFT CRAY specific instructions. The resulting file, MOPAC.UPS, will be returned to the VAX.

   This command should be of form "$CRAY SUBMIT PREUP"

10. Submit INSTALL to the CRAY. This will compile MOPAC and store the executable on the CRAY filestore.

    This command should be of form "$CRAY SUBMIT INSTALL"

    There will be two sets of error-messages. Two subroutines are names READ and WRITE and the CRAY compiler will warn that these represent multiple definitions. This is a minor bug - it will be fixed in future releases - and is nothing to worry about.

    MOPAC on the CRAY-2 will give more warnings - again these are harmless. All involve replacing exponentiations by repeated multiplications and all have been checked for logic.

11. There is no advantage to storing three copies of MOPAC on the VAX, so removal of two or three copies is recommended. If you plan to modify MOPAC, then delete MOPAC.CFT and MOPAC.UPS. If all you want to do is change the array sizes, keep MOPAC.UPS and delete all .CRAY files and MOPAC.CFT. Whatever you do, MOPAC.CFT is a fairly useless file and should be deleted.

12. Insert into your LOGIN.COM file on the VAX the following line

    "$MOPAC:==@<mopacdirectory>MOPAC" This defines the MOPAC command.

To familiarize yourself with the system, the following operations might be useful.

1. Run the supplied test molecules, MNRSD1.DAT and CHYMO.DAT, and verify that MOPAC is producing "acceptable" results.
2. Make some simple modifications to the datafiles supplied in order to test your understanding of the data format
3. When satisfied that MOPAC is working, and that data files can be made, begin production runs.

In case of difficulty, please contact Dr. James J.P. Stewart, Frank J. Seiler Res. Lab, Air Force Academy, Colorado Springs, CO 80840, (303)-472-2655.

Specific instructions for mounting MOPAC on other computers have been left out due to limitations of space in the Manual.

How to use MOPAC

The COM file to run the MOPAC can be accessed using the command "MOPAC" followed by none, one or two arguments. Possible options are:

```
MOPAC   MYDATAFILE 120
MOPAC   MYDATAFILE
```

In the latter case it is assumed that the default time (15 seconds) will be adequate.

```
MOPAC
```

In this case you will be prompted for the datafile. Restarts should be user transparent. If MOPAC does make any restart files, do not change them (It would be hard to do anyhow, as they're in machine code), as they will be used when you run a RESTART job. The main files that are produced are:

<filename>.OUT	Results
<filename>.ARC	Archive or summary
<filename>.RES	Restart
<filename>.DEN	Density matrix (in binary)

Size of MOPAC

As MOPAC is intended to fit into any computer of more than 500K addressable memory, it can assume almost any size desired. However, it is useful for programmers to have an idea of how large various portions are, and the following data might prove useful.

Sizes of Various Parts of MOPAC

Executable Code on its own (includes local constants)	200,000 bytes
"Fixed" arrays and variables without MECI	140,000 bytes
"Fixed" arrays and variables with MECI	410,000 bytes
Array space not defined in terms of MAXHEV and MAXLIT	77,000 bytes
Space used by common blocks defined by MAXHEV and MAXLIT	Variable
Size of MOPAC, 2 heavy atoms and 2 light atoms, no MECI INTERP, or PULAY	450,000 bytes
Size of MOPAC, 2 heavy atoms and 2 light atoms, including MECI INTERP, or PULAY	800,000 bytes
Size of MOPAC, 40 heavy atoms and 40 light atoms	9,000,000 bytes
Approximate size of MOPAC, for n heavy atoms and n light atoms	800,000 + bytes 5,000 * n * n
No. of lines in program = 15,000 code + 3,000 comment	18,000
Number of subroutines plus functions	108

APPENDIX A

FORTRAN FILES

TABLE 1

NAMES OF FORTRAN-77 FILES

AABABC	ANAVIB	AXIS	BLOCK	BONDS
CALPAR	CAPCOR	CHRGE	CNVG	COMPFG
DATIN	DCART	DENROT	DENSIT	DEPVAR
DERIV	DFPSAV	DIAG	DIAT	DIAT2
DIPOLE	DOT	DRC	ENPART	EXCHNG
FLEPO	FMAT	FOCK1	FOCK2	FOCK2D
FORCE	FORSAV	FRAME	FREQCY	GEOUT
GETGEO	GETSYM	GMETRY	GRID	H1ELEC
HADDON	HCORE	HELECT	IJKL	INTERP
ITER	LINMIN	LOCAL	LOCMIN	MAMULT
MATOUT	MECI	MNDO	MOLDAT	MULLIK
MULT	NLLSQ	NUCHAR	OSINV	PARSAV
PATHS	PERM	POLAR	POWSAV	POWSQ
PULAY	REACT1	READ	READA	REFS
REPP	ROTATE	RSP	SEARCH	SECOND
SPCG	SWAP	SYMTRY	THERMO	TIMOUT
UPDATE	VECPRT	WRITE	WRTKEY	XYZINT

APPENDIX B

SUBROUTINE CALLS IN MOPAC

A list of the program segments which call various subroutines. Note that function calls are not included here!

Note	Subroutine	Called by
	AABABC	MECI
	AABACD	MECI
	AABBCD	MECI
	AINTGS	DIAT2
	AM1	MAIN
	ANAVIB	FORCE
	AXIS	FORCE FRAME
	BABBBC	MECI
	BABBCD	MECI
	BANGLE	XYZINT
	BFN	DIAT
	BONDS	WRITE
a	CALPAR	AM1
	CAPCOR	ITER
	CHRGE	FMAT FOCK1 WRITE DRC
a	CNVG	ITER
	COMPFG	FLEPO FMAT FORCE LINMIN MNDO POWSQ REACT1 SEARCH DRC LOCMIN NLLSQ POLAR
a	COE	DIAT

a	DANG	XYZINT
a	DATE	WRITE
	DCART	DERIV
a	DENROT	WRITE
	DENSIT	ITER MULLIK
	DEPVAR	HADDON
	DERIV	COMPFG WRITE
a	DFPSAV	FLEPO PATHS
a	DHC	DERIV
	DIAG	ITER
a	DIAGI	MECI
	DIAT	H1ELEC
	DIAT2	DIAT
a	DIHED	XYZINT
	DIPOLE	WRITE FMAT
	DOT	FLEPO FORCE LOCMIN NLLSQ PULAY REACT1 SEARCH WRITE DRC POWSQ POWSAV
	DRC	MNDO FORCE
a	ENPART	WRITE
	EXCHNG	LINMIN LOCMIN
b	FLEPO	MNDO PATHS REACT1 FORCE GRID
	FMAT	FORCE
b	FOCK1	ITER
b	FOCK2	DCART ITER
b	FOCK2D	DCART
	FORCE	MNDO
a	FORSAV	FMAT
a	FRAME	FMAT FORCE FREQCY

	FREQCY	FORCE
	GEOUT	FLEPO GMETRY POWSQ REACT1 READ WRITE NLLSQ
	GETGEO	REACT1 READ
	GETSYM	READ
	GMETRY	COMPFG DERIV FMAT FORCE MOLDAT READ WRITE DENROT DRC MULLIK POLAR REACT1
	GRID	MAIN
b	H1ELEC	DCART DIAT HCORE
a	HADDON	DEPVAR SYMTRY
b	HCORE	COMPFG DERIV
	HELECT	DCART ITER
	IJKL	MECI
	INTERP	ITER
b	ITER	COMPFG DERIV
	LINMIN	FLEPO POWSQ
	LOCAL	WRITE
	LOCMIN	NLLSQ
a	MAMULT	PULAY
	MATOUT	FORCE WRITE FMAT ITER LOCAL MECI POLAR
	MECI	ITER WRITE
b	MOLDAT	MNDO AM1
	MULLIK	WRITE
	MULT	MULLIK
	NLLSQ	FORCE MNDO
	NUCHAR	GETSYM READ
	OSINV	PULAY
a	PARSAV	NLLSQ
	PATHS	MNDO

	PERM	MECI
	POLAR	MNDO
	POWSAV	POWSQ
	POWSQ	MNDO
	PULAY	ITER
	REACT1	MNDO
b	READ	MNDO
	READA	AM1 DEPVAR DERIV DRC FLEPO FMAT FORCE GETGEO GRID ITER MECI MOLDAT NLLSQ NUCHAR POWSQ REACT1 THERMO WRITE
	REFS	MOLDAT
a	REPP	ROTATE
b	ROTATE	DERIV HCORE
	RSP	FMAT FORCE INTERP ITER MULLIK POLAR FREQCY MECI POWSQ AXIS
	SECOND	DERIV DRC FLEPO FMAT FORCE ITER MNDO NLLSQ PATHS POWSQ REACT1 WRITE
	SCHMIT	INTERP
	SCHMIB	INTERP
	SEARCH	POWSQ
a	SET	DIAT2
	SOLROT	DCART HCORE
	SPCG	IJKL
	SPLINE	INTERP
	SS	DIAT
a	SWAP	ITER
	SYMTRY	COMPFG DERIV READ WRITE
	THERMO	FORCE
	VECPRT	BONDS FORCE HCORE ITER MECI MOLDAT MULLIK POLAR POWSQ WRITE

b	WRITE	FORCE ITER MNDO PATHS REACT1
a	WRTKEY	READ
	XYZINT	GEOUT DFPSAV WRITE DRC FORCE
	XYZGEO	XYZINT

NOTES:

a It is not important to know how these subroutines work in order to understand the working of the program.

b These subroutines form the backbone of the program, and should be among the first to be looked at when learning how it works.

A list of subroutines called by various segments (the inverse of the first list)

SUBROUTINE	CALLS
AABABC	
AABACD	
AABBCD	
AINTGS	
AM1	CALPAR MOLDAT CALPAR
ANAVIB	
AXIS	RSP
BONDS	VECPRT
COMPF	COMPFG
COMPFG	SYMTRY GMETRY HCORE ITER DERIV
DCART	DHC
DENROT	GMETRY COE
DEPVAR	
DERIV	SYMTRY GMETRY HCORE ITER DCART
DFPSAV	XYZINT
DHC	H1ELEC ROTATE FOCK2D SOLROT VECPRT
DIAT	COE BFN
DIAT2	SET AINTGS BINTGS
DRC	GMETRY COMPFG CHRGE XYZINT
EXCHNG	
FLEPO	DFPSAV COMPFG GEOUT LINMIN
FMAT	GMETRY FORSAV COMPFG CHRGE COMPFG VECPRT FRAME RSP
FOCK1	CHRGE
FORCE	GMETRY COMPFG NLLSQ WRITE FMAT RSP DRC MATOUT FREQCY THERMO FRAME XYZINT ANAVIB

FRAME	AXIS
FREQCY	RSP FRAME
GEOUT	XYZINT
GETGEO	XYZINT
GETSYM	NUCHAR
GMETRY	GEOUT
GRID	FLEPO
H1ELEC	DIAT
HADDON	DEPVAR
HCORE	H1ELEC ROTATE VECPRT SOLROT
INTERP	RSP SCHMIT SCHMIB SPLINE
ITER	VECPRT FOCK2 FOCK1 PULAY RSP DIAG DENSIT CNVG SWAP WRITE MATOUT INTERP
LINMIN	COMPFG EXCHNG
LOCAL	MATOUT
LOCMIN	COMPF EXCHNG
MECI	RSP IJKL PERM VECPRT MATOUT
MNDO	READ MOLDAT REACT1 WRITE PATHS NLLSQ COMPFG FORCE POWSQ FLEPO POLAR AM1 GRID DRC
MOLDAT	GMETRY VECPRT REFS
MULLIK	RSP GMETRY MULT DENSIT VECPRT
NLLSQ	PARSAV COMPFG GEOUT LOCMIN
PATHS	FLEPO WRITE DFPSAV
POLAR	GMETRY COMPFG VECPRT RSP MATOUT
POWSAV	
POWSQ	POWSAV COMPFG VECPRT RSP LINMIN SEARCH
PULAY	MAMULT OSINV
REACT1	GETGEO GEOUT FLEPO COMPFG WRITE
READ	GETGEO WRTKEY GETSYM NUCHAR SYMTRY GEOUT GMETRY

ROTATE	REPP
RSP	EPSETA TRED2 TQLRAT TQL2 TRBAK3
SEARCH	COMPFG
SECOND	TIMCLK
SOLROT	ROTATE
SYMTRY	HADDON
TIMOUT	
VALUE	COMPFG
VECPRT	
WRITE	DATE DERIV VECPRT MATOUT GMETRY GEOUT CHRGE BONDS LOCAL ENPART XYZINT MULLIK SYMTRY
XYZINT	XYZGEO BANGLE DIHED DANG

APPENDIX C

DESCRIPTION OF SUBROUTINES IN MOPAC

AABABC Utility: Calculates the configuration interaction matrix element between two configurations differing by exactly one alpha M.O. Called by MECI only.

AABACD Utility: Calculates the configuration interaction matrix element between two configurations differing by exactly two alpha M.O.s. Called by MECI only.

AABBCD Utility: 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.

ANAVIB Utility: 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

AXIS Utility: Works out 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 cm**(-1) and 10**(-40) gram-cm-cm.

BABBBC Utility: Calculates the configuration interaction matrix element between two configurations differing by exactly one beta M.O. Called by MECI only.

BABBCD Utility: Calculates the configuration interaction matrix element between two configurations differing by exactly two beta M.O.s. Called by MECI only.

BONDS Utility: 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.

CALPAR Utility: When external parameters are read in via EXTERNAL=, the derived parameters are worked out using CALPAR. Note that all derived parameters are calculated for all parameterized elements at the same time.

CAPCOR Utility: Capping atoms, of type Cb, should not contribute to the energy of a system. CAPCOR calculates the energy contribution due to the Cb and subtracts it from the electronic energy.

CHRGE Utility: 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.

CNVG Utility: 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.

COMPFG Main Sequence: Evaluates the total heat of formation of the supplied geometry, and the derivatives, if requested. This is the nodal point connecting the electronic and geometric parts of the program. Main arguments: on input: geometry, on output: heat of formation, gradients.

DENSIT Utility: 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.

DENROT Utility: 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 handling "d" functions, if present.

DEPVAR Utility: A symmetry-defined "bond length" is related to another bond length by a multiple. This special symmetry funtion is intended for use in Cluster calculations. Called by HADDON.

DERIV Main Sequence: 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.

DFPSAV Utility: 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.

DIAG Utility: 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 or ion. 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, but with the two-electron matrix in double precision. Called by DHC.

FORCE Main sequence: 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 key-words. Isotopic substitution is allowed. Thermochemical quantities are calculated. Called by MAIN.

FORSAV Utility: Saves and restores data used in FMAT in FORCE calculation. Called by FMAT.

FRAME Utility: Applies a very rigid constraint on the translations and rotations of the system. Used to separate the trivial vibrations in a FORCE calculation.

FREQCY Main sequence: Final stage of a FORCE calculation. Evaluates and prints the vibrational frequencies and modes.

GEOUT Utility: Prints out the current geometry. Can be called at any time. Does not change any data.

GETGEO Utility: Reads in geometry in character mode from specified channel, and stores parameters in arrays. Some error-checking is done. Called by READ and REACT1.

GETSYM Utility: Reads in symmetry data. Used by READ.

GMETRY Utility: 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.

H1ELEC Utility: Given any two atoms in cartesian space, H1ELEC calculates the one-electron energies of the off-diagonal elements of the atomic orbital matrix.

H(i,j) = -S(i,j)*(beta(i)+beta(j))/2.

Called by HCORE and DERIV.

HADDON Utility: The symmetry operation subroutine, HADDON relates two geometric variables by making one a dependent function of the other. Called by SYMTRY only.

HCORE Main sequence: 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.

HELECT Utility: 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.

HQRII Utility: Rapid diagonalization routine. Accepts a secular determinant, and produces a set of eigenvectors and eigenvalues. The secular determinant is destroyed.

IJKL Utility: Fills the large two-electron array over a M.O. basis set. Calls SPCG, and is called by MECI.

INTERP Utility: Runs the Camp-King converger. q.v.

ITER Main sequence: Given the one and two electron matrices, ITER calculates the Fock and density matrices, and the electronic energy. Called by COMPFG.

LINMIN Main sequence: Dedicated to FLEPO, LINMIN 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 by FLEPO.

LOCAL Utility: Given a set of occupied eigenvectors, produces a canonical set of localized bonding orbitals, by a series of 2 x 2 rotations which maximize <psi**4>. Called by WRITE.

LOCMIN Main sequence: 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.

MAMULT Utility: 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.

MATOUT Utility: 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.

MECI Main sequence: 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.

MNDO Main sequence: MAIN program. MNDO 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.

MOLDAT Main Sequence: 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 MNDO only.

MULLIK Utility: Constructs and prints the Mulliken Population Analysis. Available only for RHF calculations. Called by WRITE.

MULT Utility: Used by MULLIK only, MULT multiplies two square matrices together.

NLLSQ Main sequence: Used in the gradient norm minimization.

OSINV Utility: Inverts a square matrix. Called by PULAY only.

PARSAV Utility: Stores and restores data used in the gradient-norm minimization calculation.

PATHS Main sequence: 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 MNDO only.

PERM Utility: Permutes n1 electrons of alpha or beta spin among n2 M.O.s.

POWSAV Utility: Calculation store and restart for SIGMA calculation. Called by POWSQ.

POWSQ Main sequence: The McIver - Komornicki gradient minimization routine. Constructs a full Hessian matrix and proceeds by line-searches Called from MAIN when SIGMA is specified.

POLAR Utility: 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 Utility: A new converger. 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 Main sequence: 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 MNDO only.

READ Main sequence: Almost all the data are read in through READ. There is a lot of data-checking in READ, but very little calculation. Called by MNDO.

READA Utility: 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.

REFS Utility: Prints the original references for atomic data. If an atom does not have a reference, i.e. it has not been parametrized, then a warning message will be printed and the calculation stopped.

REPP Utility: 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 Utility: 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.

RSP Utility: Rapid diagonalization routine. Accepts a secular determinant, and produces a set of eigenvectors and eigenvalues. The secular determinant is destroyed.

SEARCH Utility: 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 Utility: 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 Utility: 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.

SPCG Utility: Calculates two-electron integral between any four M.O.'s at the MNDO or MINDO/3 level. Called by MECI and WRITE.

SWAP Utility: Used with FILL=, SWAP ensures that a specified M.O. is filled. Called by ITER only.

SYMTRY Utility: Calculates values for geometric parameters from known geometric parameters and symmetry data. Called whenever GMETRY is called.

THERMO Main sequence: 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.

VECPRT Utility: 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 VECPRT.

WRITE Main sequence: 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.

XYZINT Utility: Converts from cartesian coordinates into internal. XYZINT sets up its own numbering system, so no connectivity is needed. Used only in conjunction with key-word XYZ.

APPENDIX D

HEATS OF FORMATION OF SOME MNDO AND AM1 COMPOUNDS

In order to verify that MOPAC is working correctly, a large number of tests need to be done. These take about 45 minutes on a VAX 11-780, and even then the many potential bugs remain undetected. It is obviously impractical to ask users to test MOPAC. However, users must be able to verify the basic working of MOPAC, and to do this the following tests for the elements have been provided.

Each element can be tested by making up a data-file using estimated geometries and running that file using MOPAC. The optimized geometries should give rise to heats of formation as shown. Any difference greater than 0.1 Kcal/mole indicates a serious error in the program.

Caveats

1. Geometry definitions must be correct.
2. Heats of formation may be too high for certain compounds. This is due to a poor starting geometry trapping the system in an excited state. (Affects ICl at times)

Element	Test Compound	Heat of Formation MNDO	AM1
Hydrogen	CH4	-11.9	-8.8
Lithium	LiH	+23.2	
Beryllium	BeO	+38.6	
Boron	BF3	-261.0	
Carbon	CH4	-11.9	-8.8
Nitrogen	NH3	-6.0	-7.3
Oxygen	CO2	-75.0	-79.8
Fluorine	CF4	-214.2	-225.7
Aluminium	AlF	-83.6	
Silicon	SiH	+90.2	
Phosphorus	PF3	-229.3	
Sulfur	H2S	+1.7	
Chlorine	HCl	-15.3	-24.6
Germanium	GeF	-16.4	
Bromine	HBr	+3.6	-10.5
Tin	SnF	-20.4	
Iodine	ICl	-6.7	-4.6
Mercury	HgO	+101.6	
Lead	PbF	-22.6	

APPENDIX E

REFERENCES

On 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).

Original References for Elements Parametrized in MNDO

H M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

Li Parameters taken from the MNDOC program, written by Walter Thiel, Quant. Chem. Prog. Exch. No. 438; 2, 63, (1982)

Be M.J.S. Dewar, H.S. Rzepa, J. Am. Chem. Soc, 100, 777, (1978)

B M.J.S. Dewar, M.L. McKee, J. Am. Chem. Soc., 99, 5231, (1977).

C M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

N M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

O M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

F M.J.S. Dewar, H.S. Rzepa, J. Am. Chem. Soc., 100, 58, (1978).

Al L.P. Davis, R.M. Guidry, J.R. Williams, M.J.S. Dewar, H.S. Rzepa J. Comp. Chem., 2 433, (1981).

Si (a) M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc., 100, 3607 (1978). *
(c) M.J.S. Dewar, J. Friedheim, G. Grady, E.F. Healy, J.J.P. Stewart, Organometallics, 5, 375 (1986).

P M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc., 100, 3607 (1978).

S (a) M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc., 100, 3607 (1978). *
(b) M.J.S. Dewar, C. H. Reynolds, J. Comp. Chem., 7, 140 (1986).

Cl (a) M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc., 100, 3607 (1978). *
(b) M.J.S. Dewar, H.S. Rzepa, J. Comp. Chem., 4, 158, (1983)

Ge (no reference available yet - currently (2-Jun-86) in gally-proof form.

Br M.J.S. Dewar, E.F. Healy, J. Comp. Chem., 4, 542, (1983)

I M.J.S. Dewar, E.F. Healy, J.J.P. Stewart, J. Comp. Chem., 5, 358, (1984)

Sn M.J.S. Dewar, G.L. Grady, J.J.P. Stewart, J. Am. Chem. Soc., 106, 6771 (1984).

Hg M.J.S. Dewar, G.L. Grady, K. Merz, J.J.P. Stewart, Organometallics, 4, 1964, (1985).

Pb M.J.S. Dewar, M. Holloway, G.L. Grady, J.J.P. Stewart, Organometallics, 4, 1973, (1985).

* - Parameters defined here are obsolete.

On MINDO/3

Part XXVI, Bingham, R.C., Dewar, M.J.S., Lo, D.H, J. Am. Chem. Soc., 97, (1975).

On 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).

Original References for Elements Parametrized in AM1

H M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc., 107, 3902-3909 (1985).
C M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc., 107, 3902-3909 (1985).
N M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc., 107, 3902-3909 (1985).
O M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc., 107, 3902-3909 (1985).
F No reference available at this time.
Cl No reference available at this time.
Br No reference available at this time.
I No reference available at this time.

On Shift

"Unconditional Convergence in SCF Theory: A General Level Shift Technique", CARBO, R., HERNANDEZ, J.A., SANZ, F., Chem. Phys. Lett., 47, 581, (1977).

On Half-Electron

"Ground States of Conjugated Molecules. IX. Hydrocarbon Radicals and Radical Ions", M.J.S. Dewar,

J.A. Hashmall, C.G. Venier, J.A.C.S. 90, 1953 (1968).
"Triplet States of Aromatic Hydrocarbons", M.J.S. Dewar, N. Trinajstic, Chem. Comm., 646, (1970).
"Semiempirical SCF-MO Treatment of Excited States of Aromatic Compounds" M.J.S. Dewar, N. Trinajstic, J. Chem. Soc., (A), 1220, (1971).

On Pulay's Converger
"Convergance Acceleration of Iterative Sequences. The Case of SCF Iteration", Pulay, P., Chem. Phys. Lett., 73, 393, (1980).

On Pseudodiagonalization
"Fast Semiempirical Calculations", Stewart. J.J.P., Csaszar, P., Pulay, P., J. Comp. Chem., 3, 227, (1982).

On Localization
"A New Rapid Method for Orbital Localization." P.G. Perkins and J.J.P. Stewart, J.C.S. Faraday (II) 77, 000, (1981).

On Diagonalization
Beppu, Y., Computers and Chemistry, Vol.6 Page 00. (1982).

On MECI
"Molecular Orbital Theory for the Excited States of Transition Metal Complexes", D.R. Armstrong, R. Fortune, P.G. Perkins, and J.J.P. Stewart, J. Chem. Soc., Faraday 2, 68 1839-1846 (1972)

On Davidon-Fletcher-Powell Method
Fletcher, R., Powell, M. J. D., Computer Journal, VOL. 6, p. 163, (1963), Davidon. W. C., ibid, 10, 406 (1968).

On 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).

On Thermodynamics
"Ground States of Molecules. 44 MINDO/3 Calculations of Absolute Heat Capacities and Entropies of Molecules without Internal Rotations." Dewar, M.J.S., Ford, G.P., J. Am. Chem. Soc., 99, 7822 (1977).

On SIGMA Method
Komornicki, A., McIver, J. W., Chem. Phys. Lett., 10, 303, (1971).
Komornicki, A., McIver, J. W., J. Am. Chem. Soc., H

94, 2625 (1971)

On Bonds

"Bond Indices and Valency", Armstrong, D.R., Perkins, P.G., Stewart, J.J.P., J. Chem. Soc., Dalton, 838 (1973).

On 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)

On Dipole Moments for Ions

"Molecular Quadrupole Moments", A.D. Buckingham, Quarterly Reviews, 182 (1958 or 1959)

INDEX

END

10-86

DTIC