GAUSSIAN 76 - An ab initio molecular orbital program.

J. S. Binkley, R. Whiteside, P. C. Hariharan, R. Seeger, W. J. Hehre,
W. A. Lathan, M. D. Newton, R. Ditchfield and J. A. Pople

Abstract

GAUSSIAN 76 is a general-purpose computer program for ab initio Hartree-Fock molecular orbital calculations. It can handle basis sets involving s, p and d-type gaussian functions. Certain standard sets (STO-3G, 4-31G, 6-31G* etc.) are stored internally for easy use. Closed shell (RHF) or unrestricted open shell (UHF) wavefunctions can be obtained. Facilities are provided for geometry optimization to potential minima and for limited potential surface scans.

MASTER

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

LIST OF CONTENTS

1. Introduction

2. General Features of GAUSSIAN 76

3. Input Decks for Simple Jobs

    A. Control Input

    B. Data Input

4. Input Decks for Compound Jobs

    A. External Initial Guess

    B. Restarts

5. List of Options

1. Introduction

GAUSSIAN 76 is a connected system of programs capable of performing *ab initio* molecular orbital (MO) calculations within the linear combination of atomic orbitals (LCAO) framework. It represents a further development of the GAUSSIAN 70 system already published.[1] The version of GAUSSIAN 76 distributed through the Quantum Chemistry Program Exchange (QCPE) has been made operational on the Brookhaven National Laboratory CDC 7600 computer.

The system contains programs for the calculation of the required one- and two-electron integrals using basis sets of s, p or d cartesian gaussian functions. These are followed by programs for determination of self-consistent field (SCF) single-determinant wave functions and associated total energies. For closed shell singlet states, the restricted Hartree-Fock (RHF) procedure of Roothaan[2] is used. For open shell states, wave functions and energies are obtained with the unrestricted Hartree-Fock (UHF) procedure.[3] This is followed by a program for Mulliken population analysis[4] and computation of the electric dipole moment.

Execution of GAUSSIAN 76 requires specification of nuclear geometry (by coordinates or by internuclear distances and angles), net charge and spin multiplicity, basis set details (either from internally stored standard sets or from cards) and an initial guess at the density matrix (which may be generated internally or may be obtained from the output of a previous execution).

In these notes, the general features of GAUSSIAN 76 are described and detailed instructions provided for execution. In Section 2, the arrangement of the program is outlined and certain key terminology is introduced. Various limitations (maximum number of basis functions,

maximum number of atoms, etc.) are given at the end of this section. Section 3 gives detailed instructions concerning the preparation of input decks for simple (self-contained) jobs. This includes a description of some of the options that exist within the program. Section 4 describes compound jobs, those which require use of output from a previous job. Finally, Section 5 contains a complete list of the options available in GAUSSIAN 76.

2. General Features of GAUSSIAN 76

GAUSSIAN 76 is a collection of thirteen interconnected links, each consisting of a main program and associated routines that occupy high-speed memory together. A group of links that are related by common objectives comprises an overlay. The overlays are numbered in logical sequence and have the following general functions:

Overlay 2: Nuclear geometry, net charge and spin multiplicity

Overlay 3: Basis set and integrals

Overlay 4: Initial guess at density matrix

Overlay 5: Self-consistent field calculations

Overlay 6: Analysis of results

With each overlay is associated a set of options, which control the way the overlay is executed.

Within an overlay, there are various links which are specified by a two-digit number following the overlay number. For example, the first link of overlay 3 is numbered 301 and given the label L301. Each link also has a descriptive name. When use of an overlay is required, instructions must be provided for the sequence of links and the options. Note that the set of options apply to the whole overlay and are the same for each link in the overlay.

The complete list of links in GAUSSIAN 76 is given in Table 1. It should be noted that the numbering is not sequential. This is because the published version contains only a subset of the links of a larger system.

The first three links (LINK0, LINK1 and LINK2) do not form part of any overlay. LINK0 is the only part of GAUSSIAN 76 which is permanently resident in high-speed memory and serves to initialize blank common (described below) and to call LINK1. The purpose of LINK1 is (a) to determine from input the required execution sequence of links and (b) to read in the options which are specified for each overlay. LINK2 is concerned only with repeated execution of a given route.

Data produced or required by the various links are internally communicated by means of shared high-speed memory (blank common), shared random-access mass storage files (disc or low-speed memory) and shared sequential-access mass storage files (tape or disc) for two-electron integrals.

To complete a calculation with GAUSSIAN 76, the following four types of data are required:

1. Nuclear geometry. This must always be specified externally, usually in terms of internuclear lengths and angles, supplied to L202. There is provision for geometry optimization to minimize the energy and for partial scan of a potential surface.

2. Net charge and spin multiplicity. This is always required by L202 and specifies the number of electrons of either α or β spin. For spin multiplicities greater than one (doublet, triplet ..... states etc.), the program determines the wave function component with the maximum possible number of α-electrons.

TABLE 1. List of Links in GAUSSIAN 76

Number	Name	Function
L000	LINK0	Initialization
L001	LINK1	Control of route through the links
L002	LINK2	Control of repeated execution.
L202	VGEOM2	Computation and modification of nuclear geometry
L301	GINPUT	Specification of basis set
L302	STVINT	Overlap, kinetic and potential one-electron integrals
L303	DIPOLE	Dipole one-electron integrals
L311	SHELL	Two-electron integrals involving s and p functions
L312	UNCON	Two-electron integrals involving s, p and d-functions
L401	GUESS	Initial guess at the density matrix
L501	CLOSED	Closed-shell (RHF) self-consistent-field
L502	OPEN	Open-shell (UHF) self-consistent field
L601	MULPOP	Mulliken population analysis and dipole computation

3. Basis set. The gaussian basis set may be specified externally for L301. However, certain types of basis set are stored internally and may be used automatically by appropriate choice of option. The internal sets are (i) the minimal STO-NG sets[5] (N=2,3,4,5 or 6), the split-valence N-31G sets[6] (N =4,5 or 6) and the polarized sets[7,8] 6-31G*, 6-31G** and STO-NG*.

4. Initial guess. The initial guess at the density matrix is usually generated internally by L401. There are options to interchange some occupied and unoccupied orbitals at this stage to give some control over which electronic state of the molecule is obtained. External specification of the initial density matrix is possible only if it is available on a file from a previous run.

If all the necessary information is provided, overlay 5 will attempt an iterative solution of the Roothaan equations[2]

$$\underline{FC} = \underline{SCE}$$

or their UHF equivalents.[3] Here $\underline{F}$ is the Fock matrix, $\underline{C}$ the matrix of LCAO coefficients, $\underline{S}$ the overlap matrix and $\underline{E}$ the diagonal matrix of MO eigenvalues. The initial guess is used to compute $\underline{F}$, equations (1) are solved for $\underline{C}$ and $\underline{E}$, $\underline{C}$ is used to calculate a new density matrix and this is in turn used to compute $\underline{F}$ for the next iteration. This process is repeated until convergence is achieved.

The two most common sources of difficulty are failure to achieve SCF convergence and convergence to a wave function for the wrong state. These can sometimes be overcome by experimentation (orbital switching) with the initial guess. A more reliable (but more expensive) procedure is available in the GAUSSIAN 78 system.[9]

There are upper limits to the size of computation that can be carried out with GAUSSIAN 76. These are:

Maximum number of atoms. This is 35 in the published version.

Maximum number of basis functions. This is 90 in the QCPE version. By appropriate redimensioning, this can be increased to 127 if sufficient high-speed memory is available.

Maximum number of shells. A shell is a set of basis functions on one center with shared gaussian exponents (e.g. the 2s, $2p_x$, $2p_y$, $2p_z$ functions on carbon in STO-3G constitute a shell with four members). The number of shells cannot exceed 80 in the published version.

Maximum degree of contraction. Each function φ in a shell is a linear combination of primitive gaussian functions

$$\varphi = \sum_{i=1}^{n} d_i \, P(x,y,z) \exp(-\alpha_i r^2)$$

where P is a polynomial. The number of terms n in such a sum is the degree of contraction for that particular shell. This cannot exceed six in GAUSSIAN 76.

Maximum number of independent gaussian exponents. This is the sum of the degrees of contraction of all the shells. The total number here cannot exceed 240 in the published version.

3. Input Decks for Simple Jobs

An input deck is a set of cards (or card images) which specifies the job to be executed. It consists of two parts--control input and data input. The control input, which comes first, specifies the route or sequence of links to be executed together with various options telling each link what to do. The system may be instructed, for example, to perform an RHF/STO-3G single point calculation or a UHF/6-31G* geometry optimization, depending upon the control input. Each link may be directed by the control input to read cards from the following data input. This may consist, for example, of the molecular geometry, charge, spin multiplicity and the basis set. Thus, the type of data input required is determined by the control input.

If only a single route is specified (which may be executed several times), the job will be termed simple. The complete control input must then precede the data input. If additional routes are specified which follow the execution of the first route and which make use of output from execution of the first route, the job will be described as compound. Simple jobs will be described first, compound jobs being left for Section 4.

3A. Control Input for Simple Jobs

The first control card called for by GAUSSIAN 76 is the system control card. Certain standard routes may be generated automatically by this card, in which case no further control input is necessary. Alternatively, the route may be specified by a following sequence of overlay control cards.

The following information is always required on the system control card:

Column 1: This must contain a '$'.

Column 2: This must be blank or contain an 'N'.

These features identify the card as a system control card. If 'N' appears in column 2, economical or compressed printing is called for throughout. This option is to save paper and printing charges and is not recommended until the user is familiar with the output.

Column 4: The parameter NTIMES, the number of times the prescribed route is to be executed. If greater than 9, this is read in I2 format from columns 3 and 4.

If the route is to be specified by overlay control cards (see section 3A.2), the remainder of the system control card is left blank. However, if the route is to be generated automatically, further entries are needed.

3A.1 Automatic Route Generation.

The additional parameters read from the system control card are:

Column 6: ITYPE ... Type of calculation to be performed.

0 (or blank) ... Read route from cards (see 3A.2)

1 ... Single point calculation

3 ... Geometry optimization (see section 3A.2)

4 ... Potential surface scan (see section 3A.2)

5 ... Restart of geometry optimizations. (See Section 4B).

Column 10: IBASIS ... Type of s,p basis set used

0 (or blank) ... One of the minimal STO-NG bases, using standard exponents. N is specified by the next parameter NGAUSS.

1 ... One of the split-valence N-31G bases, using standard scaling factors. N is specified by the next parameter, NGAUSS.

2 ... General basis set input from cards (see data input for L301).

Column 12: NGAUSS ... Number of gaussians in a contraction (see description under IBASIS above).

0 (or blank) ... Default value. For IBASIS = 0, this will give STO-3G; for IBASIS = 1, the default is 4-31G.

N ... If IBASIS = 0, values 2,3,4,5 or 6 are permitted. If IBASIS = 1, values 4,5 or 6 are permitted.

Column 14: IPOL ... Polarization functions in the basis set.

0 (or blank) ... No polarization functions

1 ... Include a set of standard polarization functions on each non-hydrogen atom.

2 ... In addition to the above, include a set of p-functions on each hydrogen.

Column 16: IGUESS ... Type of initial guess

0 (or blank) ... Obtained by projection from a Huckel minimal basis calculation.

1 ... Obtained by diagonalization of core Hamiltonian

2 ... Huckel-type guess for extended basis. However, this must not be used if IBASIS = 2.

Column 18: ISCF ... Type of SCF calculation performed

0 (or blank) ... Restricted closed-shell Hartree-Fock (RHF)

1 ... Unrestricted open-shell Hartree-Fock (UHF)

At this point, we give some examples to illustrate automatic route generation by the system control card.

<u>Example 3A-1</u>. A single closed shell RHF/STO-3G calculation

Card 1 : $^^1^1

Card 2 : First card of data input

Here ITYPE = 1 ; single point calculation

IBASIS = 0 ; minimal STO- NG basis

NGAUSS = 0 ; N = 3

IPOL = 0 ; No polarization functions

IGUESS = 0 ; Huckel-type minimal guess

ISCF = 0 ; closed shell SCF

<u>Example 3A.2</u> A single open shell UHF/6-31G* calculation

Card 1 : $ 1 1 1 6 1 1

Card 2 : First card of data input

ITYPE	= 1 ;	single point calculation
IBASIS	= 1 ;	split valence N-31G basis
NGAUSS	= 6 ;	N = 6
IPOL	= 1 ;	polarization functions on heavy atoms
IGUESS	= 0 ;	projected Hückel-type guess
ISCF	= 1 ;	open shell SCF

Example 3A.3 RHF/STO-3G geometry optimization with minimal printing

Card 1 : $N 1 3

Card 2 : First card of data input

ITYPE = 3 ; geometry optimization

IBASIS = 0 ; minimal STO-NG basis

NGAUSS = 0 ; N = 3

IPOL = 0 ; no polarization functions

IGUESS = 0 ; projected Hückel-type guess

ISCF = 0 ; closed shell RHF

Example 3A.4 UHF/4-31G potential surface grid scan

Card 1 : $ 1 4 1 1

Card 2 : First card of data input

ITYPE = 4 ; potential surface scan

IBASIS = 1 ; split valence N-31G basis

NGAUSS = 0 ; N = 4

IPOL = 0 ; no polarization function

IGUESS = 0 ; projected Hückel-type guess

ISCF = 1 ; unrestricted open shell Hartree-Fock (UHF)

3.A2. Route Generation by Overlay Control Cards

Although the type of control input described above is sufficient for many jobs, it does not reflect the full flexibility of the program. In this section, we describe the more detailed route generation by a sequence of overlay control cards, called for by ITYPE = 0 (or blank) in column 6 of the system control card. Each overlay control card prescribes a sequence of links to be executed and the set of options used. This sequence of overlay control cards (a maximum of 20) is terminated by a blank card.

The parameters read from an overlay control card are:

IOV,(IOP(I),I=5,34), (ISEG(J),J=1,18), JUMP with Format (I4,48OI,I4).

The functions of these parameters are as follows:

Column 4: IOV ... The number of the overlay called. If IOV=3, for example, this implies that one or more links from overlay 3 will be executed.

Columns 5-34: IOP(I) ... These are thirty options which are available in each overlay (not all are used). A detailed listing of the options is given in Section 5. In this description, we shall only mention some of those most commonly used in examples. The options are octal quantities (between 0 and 7), and the option number coincides with the column number. Thus the value of IOP(7) should appear in column 7.

<u>Columns 35-52</u>: ISEG(J) ... The entries in these columns must be either '1' or '0' and determine which links of the overlay are executed. Column 35 corresponds to the first link of the overlay, column 36 to the second and so forth. If a '1' appears, the link is executed, if a '0' or a blank, it is not. The links are called once in turn.

<u>Columns 53-55</u>: JUMP ... If this parameter is '0' (or blank) control is transferred to the next overlay control card. Non-zero values are used to generate loops in the route. For example JUMP = -2 will transfer control to the overlay control card two places before the current one. Use of this feature will become clear in some of the examples given below.

As an example of the use of these fields, consider the following card image:

```
Col. 1                          35
     ↓   316                    ↓
     ^^^                        1 1
                                  ^
```

This overlay control card requests the execution of the first and third links (cols. 35 and 37) of overlay 3 (Col. 4). These are links L301 and L303. In addition, options 5 and 6 are set at non-zero values, IOP(5)=1 and IOP(6)=6.

Example 3A.5

As an example of route specification, consider a single closed shell RHF/STO-3G calculation. The steps required are:

Step	Link Required
Input geometry	202
Specify basis set	301
Overlap kinetic and potential integrals	302
Dipole integrals	303
Electron repulsion integrals	311
Generate initial guess	401
Solve SCF equations	501
Mulliken population analysis	601

Reference to Section 5 shows that the route can be executed without setting any options to non-zero values. Hence the calculation can be executed with the following cards:

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:         $  1
CARD  2:            2                                   1
CARD  3:            3                                 111         1
CARD  4:            4                                 1
CARD  5:            5                                 1
CARD  6:            6                                 1
CARD  7:
CARD  8:         (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Example 3A.5 (cont'd)

Card 1 is the system control card with ITYPE=0 (col 6), indicating that overlay control cards follow.

Card 2 indicates that link 202 is executed with all options set to zero. This performs the first step of the calculation by reading from cards the molecular geometry in the form of a Z-matrix (see Section 3B).

Card 3 calls for links 301, 302, 303 and 311 of overlay 3. Since all options are zero, the STO-3G basis is selected and appropriate integrals are evaluated.

Card 4 calls for the initial guess link 401. This forms the density matrix based on the minimal Hückel guess.

Card 5 calls for link 501 to solve the closed shell restricted Hartree-Fock equations.

Card 6 calls for link 601 to carry out a Mulliken population analysis and evaluate the dipole moment.

Card 7 is blank indicating the end of the route specification

Card 8 is the beginning of the data input which will be called for during execution of the specified links. After this execution is complete, the program will seek another system control card to indicate the beginning of another job.

The route generated in example 3A.5 is identical with that produced automatically by the system control card in example 3A.1. However, the longer version has advantages of greater flexibility since various options may be changed.

Example 3A.6

Suppose that in example 5, a print of the initial guess MO coefficients is required. This is called for by option 33 of overlay 4 (see Section 5), and the modified control input is:

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD   1: $      1
CARD   2:        2                                   1
CARD   3:        3                                  111          1
CARD   4:        4                                1 1
CARD   5:        5                                  1
CARD   6:        6                                  1
CARD   7:
CARD   8: (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Example 3A.7 UHF/6-31G* single point calculation

This example introduces some more links and options that are commonly used.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1: $ 1
CARD  2:     2                                   1
CARD  3:     3161                               111        11
CARD  4:     4                                  1
CARD  5:     5                                   1
CARD  6:     6                                  1
CARD  7:
CARD  8: (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Cards 1,2 are as in the previous example

Card 3 specifies some new options and the additional link 312.

IOP(5)=1 indicates a K-31G split valence basis

IOP(6)=6 indicates K=6

IOP(7)=1 indicates standard polarization functions on first row atoms. This implies the 6-31G* basis.

Card 4 calls for an initial guess formed by projection

Card 5 calls for the unrestricted Hartree-Fock solution

Cards 6-8 as in the previous example

Example 3A.8 UHF/6-31G** single point calculation.

A simple extension of the calculation described in example 7 is the addition of a set of p-type polarization functions to each hydrogen atom in the molecule. This is accomplished by setting IOP(7)=2 on the overlay 3 route card. As in the previous example, first row atoms have the standard 6-31G basis supplemented with a set of six second-order gaussian polarization functions

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD   1: $  1
CARD   2:    2                                   1
CARD   3:    3162                               111          11
CARD   4:    4                                  1
CARD   5:    5                                   1
CARD   6:    6                                  1
CARD   7:
CARD   8: (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

This completes the examples of control input for single runs at one nuclear geometry. The reader only concerned with this type of run may wish to skip the rest of this section and proceed directly to 3B for instructions on the construction of data input.

We now turn to geometry optimization. Here the program system can be used to minimize the energy with respect to a series of geometrical parameters. This is accomplished by appropriate use of L202. After the input of the initial geometry and calculation of the corresponding total energy, L202 may be called again with IOP(5)=1. This causes a variable geometrical parameter to be read in (details in Section 3B) together with a stepsize. The energy variation is then investigated by repeating the energy calculation as the parameter is changed. Initially, two additional points are calculated, followed by a third at the minimum predicted by a quadratic fit. If there are several parameters to be optimized, this is repeated for each parameter, except that each later optimization uses the 'final' geometry obtained by optimization of the previous parameter.

Example 3A.9 RHF/STO-3G geometry optimization

The control input for this kind of geometry requires the following steps for execution:

Step	Links required
1. Input starting geometry	202
2. Evaluate integrals	301,302,303,311
3. Initial guess	401
4. Solve SCF equations	501
5. Mulliken population analysis	601
6. Vary geometry	202
7. Return to step 2	

Step 7 introduces a loop into the sequence of overlays. After steps 2-5 are carried out to calculate the wave function and energy at the initial geometry, step 6 changes the geometry and the process is repeated with a varied geometry. When all the requested parameters have been varied, control must cause escape from the loop. The above series of steps can be accomplished by the following route:

```
                 00000000011111111112222222222333333333344444444445555555
                 12345678901234567890123456789012345678901234567890123456
CARD   1: $      1
CARD   2:        2                                   1
CARD   3:        3                                  111        1
CARD   4:        4                                   1
CARD   5:        5                                   1
CARD   6:        6                                   1
CARD   7:        21                                   1                  -4
CARD   8:
CARD   9: (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456
                 00000000011111111112222222222333333333344444444445555555
```

Example 3A.9 (cont'd)

The only new features are on card 7 which performs the 'vary geometry' step and defines the loop over control cards. First IOP(5)=1, indicating that this is a 'vary' instruction rather than an initial geometry. L202 will remember the initial geometry and read data cards indicating a parameter to be optimized. After this, a parameter card specifying the variable to be altered is read in every fourth entry to L202. The JUMP entry -4 in columns 54 and 55 will then transfer control back to overlay 3 (4 cards back) so that the calculation is repeated. The optimization of the energy is continued in this manner by L202 until all the parameters have been varied, at which time the JUMP value is suppressed. Control then passes to the next card (8) which terminates the route.

Although the route in example 9 works correctly and gives a partial geometry optimization, it has several undesirable features. In the first place, it performs a population analysis for each MO calculation, which is usually not necessary. It is normally sufficient to do this only for the initial and final points. Secondly, the dipole integrals play no part in the energy calculations and are not required until the final point. Thirdly, it is usually a good idea to avoid the initial guess after the first point since the converged density matrix for the previous run is normally a better guess than can be produced by L401. In the next example, we give a route which incorporates these time-saving modifications.

Example 3A.10 Modified route for RHF/STO-3G geometry optimization

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1: $  1
CARD  2:    2                                   1
CARD  3:    3                                  111        1
CARD  4:    4                                  1
CARD  5:    5                                  1
CARD  6:    6                                  1
CARD  7:    21                                  1
CARD  8:    3                                  11         1
CARD  9:    5                                  1
CARD 10:    21                                  1                    -2
CARD 11:    3                                  1 1
CARD 12:    6                                  1
CARD 13:
CARD 14: (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Cards 2-6 perform the initial point calculation with population analysis. The resulting density matrix is stored.

Card 7 executes L202 for the second time. Since IOP(5)=1, this link will read a variable parameter from data input and will initiate geometry variation.

Cards 8-9 carry out the next MO calculation. Note that no dipole integrals are needed. Also L401 is not called, as a suitable guess is already available from the previous calculation.

Example 3A.10 (cont'd)

Card 10 is a further call to L202 with IOP(5)=1. Since a variable parameter is now active, no further ones will be read. After readjustment of the variables, control is returned to card 8 by the value JUMP = -2. This constitutes the loop in the program execution. After optimization of the first variable, L202 reads another variable specification for another optimization. When the list of variables is exhausted, the JUMP instruction is suppressed and control passes to the next card (card 11).

Cards 11-12 calculate the dipole integrals and do a population analysis using the geometry of the last MO calculation performed. This completes the full optimization run which is terminated by a blank card.

The structure of the route in example 10 is the standard one for geometry optimizations. It is the route actually generated automatically (section 3A1) by the single card

$ ^^ 1 ^ 3

for an RHF/STO-3G geometry optimization.

We give two additional examples of the full routes for geometry optimizations:

Example 3A.11 RHF/6-31G* geometry optimization

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:  $  1
CARD  2:     2                                   1
CARD  3:     3161                               111        1
CARD  4:     4                                  1
CARD  5:     5                                  1
CARD  6:     6                                  1
CARD  7:     21                                  1
CARD  8:     3161                               11         11
CARD  9:     5                                  1
CARD 10:     21                                 1                        -2
CARD 11:     3161                               1 1
CARD 12:     6                                  1
CARD 13:
CARD 14:  (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

This route includes options relevant to polarization functions as in example 7. It could be generated automatically by the single card:

```
$ 1 3 0 1 6 1
```

Example 3A.12. UHF/4-31G geometry optimization.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:  $  1
CARD  2:     2                                   1
CARD  3:     31                                 111        1
CARD  4:     4                                  1
CARD  5:     5                                   1
CARD  6:     6                                  1
CARD  7:     21                                  1
CARD  8:     31                                 11         1
CARD  9:     5                                   1
CARD 10:     21                                  1                   -2
CARD 11:     31                                 1 1
CARD 12:     6                                  1
CARD 13:
CARD 14:  (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

This control input would be generated automatically by the single card:

```
$ 1 3 0 1 4
 ^^ ^ ^ ^ ^
```

Another feature of GAUSSIAN 76 is its ability to perform a potential surface scan for a series of values of a single geometrical variable. The method of doing this is similar to that for geometry optimization. After the input of the initial geometry (by L202) and the first MO calculation, L202 may be called again with IOP(5)=2. This will request details of the variable to be changed, the number of points to be calculated and the stepsize (see Section 3B for details). A typical route for such a run is as follows:

Example 3A.13. RHF/STO-3G potential surface scan

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:         $  1
CARD  2:            2                                   1
CARD  3:            3                                 111        1
CARD  4:            4                                 1
CARD  5:            5                                 1
CARD  6:            6                                 1
CARD  7:            22                                 1
CARD  8:            3                                 111        1
CARD  9:            5                                 1
CARD 10:            6                                 1
CARD 11:            22                                 1                  -3
CARD 12:
CARD 13:         (FIRST CARD OF DATA INPUT)
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

The structure of the route in example 13 is standard and can be generated automatically by the single card

```
$ 1 4
```

Cards 2-6 perform the initial point calculation.

Card 7 executes L202 for the second time. Since IOP(5)=2, this link will read details about the variable parameter and will initiate geometry variation.

Cards 8-10 carry out the next MO calculation. The density matrix from the previous point is available as a suitable guess.

Card 11 is a further call to L202 with IOP(5)=2. Since the variable parameter specification is now active, no further data is read. The variation is continued, and the JUMP instruction returns control to card 8 for continuation of the computations. When all steps have been completed, the JUMP instruction is suppressed.

Card 12 terminates the control input.

3B. Data Input for Simple Jobs for GAUSSIAN 76

Many of the links in GAUSSIAN 76 may read cards from the data input (which immediately follows the last control input card). These data cards may specify, for instance, the molecular geometry, the basis set, or parameters to be varied in a geometry optimization. This section describes, on a link-by-link basis, the meanings and formats of the various types of data cards which may be required.

3.B.1 Data Input to L202.

The molecular geometry must always be specified on data cards. This can be supplied to L202 in the form of a Z-matrix, which specifies the geometry conveniently in terms of internuclear distances and angles. With this form of geometry input, the first two data cards contain:

1.	(ITITLE(I),I=1,10)	FORMAT(10A6)
2.	ICHARG,MULTIP	FORMAT(2I2)

The first card is simply a title card describing the job. The second card specifies the net charge (ICHARG) and the spin multiplicity (MULTIP). For the latter, we use the '1' for singlets, '2' for doublets, '3' for triplets, etc.

After this comes the Z-matrix which specifies the molecular geometry. The nuclei in the molecule are numbered sequentially, and the nth nucleus is specified by:

IAN,N1,BL,N2,THETA,N3,PHI,J

FORMAT (I3,1X, I3,1X,F6.4,1X,I3,2(1X,F10.4,1X,I3))

Columns 1-3: This is the atomic number of the nucleus in I3 FORMAT.

Columns 4-7: N1 is the sequential number (N1<N) in I4 format of a nucleus for which the internuclear distance R(N,N1) will be given.

Columns 8-14: BL is the magnitude of the internuclear separation R(N,N1). This is a floating point number in units of angstroms.

Columns 15-18: N2(<N,≠N1) is the sequential number of the nucleus for which the internuclear angle θ(N,N1,N2) at N1 will be given.

Columns 19-29: THETA is the magnitude of the internuclear angle θ(N,N1,N2) at N1. The units are degrees.

Columns 30-33: N3 is the sequential number (<N;≠N1;≠N2) of the nucleus for which the internuclear dihedral angle φ(N,N1,N2,N3) will be given.

Columns 34-44: PHI is the magnitude of the internuclear dihedral angle φ(N,N1,N2,N3). This is a floating point number in degrees. The dihedral angle is defined as the angle between the planes (N,N1,N2) and (N1,N2,N3), the sign being positive is the sequence N,N1,N2,N3 is such that movement of the directed vector $\overrightarrow{NN1}$ to the directed vector $\overrightarrow{N2N3}$ involves a righthanded screw motion.

Although it is usually possible to specify the position of nucleus N by a bond length, a bond angle, and a dihedral angle, it is sometimes simpler to replace the dihedral angle by a second bond angle. Thus, we may specify the following alternate items:

Columns 30-33: N3 here is the sequential number (N3<N; ≠N1; ≠N2) of the nucleus for which the second internuclear angle X(N,N1,N3) at N1 will be given.

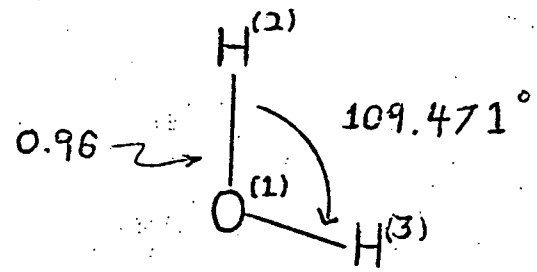
Columns 34-44: The magnitude of the angle X(N,N1,N3), a floating point number in units of degrees.

Columns 47 and 48: Specification of the two angles, θ,X will lead to two possible directions for the line N1N. To fix one of these, the number +1 is specified here if the triple vector product $\vec{N1N}$ ($\vec{N1N2}$X$\vec{N1N3}$) is positive. If the triple vector product is negative, -1 is used. The presence of either ±1 here is used as an indication that the second angle is an internuclear angle and not a dihedral angle.

These seven (or possibly eight) items are generally sufficient to specify the nature and position of nucleas N uniquely if the nuclei with smaller sequential numbers are already defined. At the beginning of the sequence, less information is required. For nucleus 1, only the atomic number is required, and it is positioned at the origin. For nucleus 2, only the first three fields need be specified, and the angles are determined by placing this atom on the Z-axis. For the third nucleus, the first five fields are required, and thus nucleus is placed in XZ-plane. A few examples will illustrate this means of geometry input.

Example 3B.1 RHF/STO-3G single calculation of H_2O.

Consider the standard model geometry for water (1). The



1

superscripts are the sequence numbers assigned to the atoms. The following deck will perform a MO calculation on this molecule.

```
            00000000011111111112222222222333333333344444444445555555555
            12345678901234567890123456789012345678901234567890123456789
CARD   1:   $   1 1
CARD   2:          WATER MOLECULE  RHF/STO-3G
CARD   3:    0 1
CARD   4:     8
CARD   5:     1    1 0.96
CARD   6:     1    1 0.96      2 109.471
CARD   7:
            12345678901234567890123456789012345678901234567890123456789
            00000000011111111112222222222333333333344444444445555555555
```

Example 3B.1 (cont'd)

Card 1 is the system control card, directing the system to perform a single RHF/STO-3G calculation.

Cards 2 and 3 specify (i) the title of the job, and (ii) the net charge and spin multiplicity for the molecule.

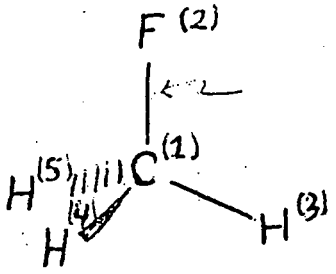
Card 4 specifies "Nucleus 1 is an oxygen (atomic number = 8)". This atom is placed at the origin, and no more values need be specified.

Card 5 specifies "Nucleus 2 is a hydrogen, and is 0.96 Å from Nucleus 1". This is placed on the Z-axis, so that no angle is specified on the card.

Card 6 specifies "Nucleus 3, a hydrogen, is 0.96Å from Nucleus 1, and the H-O-H angle is 109.471°".

Example 3B.2

The use of the dihedral angle is illustrated in the following example of a calculation on the fluoromethane molecule:




```
            00000000011111111112222222222333333333344444444445555555555
            12345678901234567890123456789012345678901234567890123456789
CARD   1:   $  1 1 0 1 6 1
CARD   2:       FLUOROMETHANE MOLECULE   RHF/6-31G*
CARD   3:    0 1
CARD   4:     6
CARD   5:     9    1 1.384
CARD   6:     1    1 1.097     2 110.6
CARD   7:     1    1 1.097     2 110.6          3 120.
CARD   8:     1    1 1.097     2 110.6          3 -120.
CARD   9:
            12345678901234567890123456789012345678901234567890123456789
            00000000011111111112222222222333333333344444444445555555555
```

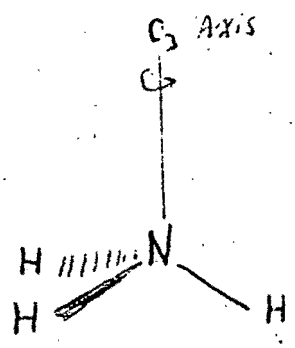
Cards 1-6 have nothing new. The first directs the system to perform a single RHF/6-31G* calculation. The next 2 specify the title, charge, and multiplicity. Cards 4-6 specify the geometry of the $F^2C^1H^3$ fragment. Card 7 specifies "nucleus 4, a hydrogen, is bound to nucleus 1 by 1.097Å, forming an $F^2C^1H^4$ angle of 110.6. The plane defined by nuclei H^4, C^1 and F^2 forms a 120° dihedral angle with the plane defined by nuclei C^1F^2 and H^3. Card 8 specifies a hydrogen nucleus identical to that on card 7 except that the dihedral angle here is -120°. Card 9 terminates the geometry output.

Example 3B.3 RHF/STO-3G Ethane molecule (staggered)

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:         $  1 1
CARD  2:            ETHANE MOLECULE, RHF/STO-3G, STAGGERED CONFORMATION
CARD  3:          0 1
CARD  4:          6
CARD  5;          6    1 1.31
CARD  6:          1    1 1.08      2 109.471
CARD  7:          1    1 1.08      2 109.471       3 120.
CARD  8:          1    1 1.08      2 109.471       3 -120.
CARD  9;          1    2 1.08      1 109.471       3 60.
CARD 10:          1    2 1.08      1 109.471       6 109.471      +1
CARD 11:          1    2 1.08      1 109.471       6 109.471      -1
CARD 12:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

This example illustrates the use of the alternate scheme for specification of the second angle. The first methyl group is specified using dihedral angles and the second using another internuclear angle.

There are two other types of 'atoms' which may be specified in a Z-matrix: those with atomic number zero and those with negative atomic numbers. Atomic number zero specifies a 'ghost atom'. This is a center to which a basis function may be attached, for instance, but has zero charge. Standard internal bases should not be used if a 'ghost atom' is specified; external basis input should be used. An example using a ghost atom is given later, in the description of external basis input. An atom with a negative atomic number is a "dummy atom", and basis functions may not be attached to this type of center. It is used only to help generate coordinates of the other atoms, and is non-existent outside L202. Dummy atoms are useful, for instance, in fixing the orientations of a symmetry axis. Consider the C_{3v} ammonia molecule (3)



Example 3B.4: RHF/STO-3G NH_3 molecule.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD   1:        $  1 1
CARD   2:              AMMONIA MOLECULE, RHF/STO-3G
CARD   3:         0 1
CARD   4:          7
CARD   5:        -1    1 10.
CARD   6:         1    1 1.01      2 115.
CARD   7:         1    1 1.01      2 115.          3 120.
CARD   8:         1    1 1.01      2 115.          3 -120.
CARD   9:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

This matrix specifies a C_{3v} NH_3 molecule with NH bonds of 1.01Å and angles of 115.0° from the C_3 axis to one of the hydrogens. Card 5 specifies a "dummy atom", which defines the three-fold notation axis to coincide with the z-axis of the coordinate system. The hydrogens are then placed around this.

Geometry Optimization and Potential Surface Scan

We now consider geometry optimizations. In this procedure, L202 reads not only the initial geometry, but also specifications and step size for parameters to be varied. However, these are used in different executions of the link; the initial geometry is read when L202 is executed with IOP(5)=0, and the optimization data is read in subsequent entries when L202 is executed with IOP(5)=1. To specify a parameter to be varied, first assign sequence numbers to the floating point numbers in the Z-matrix: first the internuclear distances (BL), then the internuclear angles (THETA) and finally the dihedral angles (PHI), or the second internuclear angles. The parameters to be varied together are specified on a card in the following format:

(IVAR(I),I=1,15), STEPSZ FORMAT(15I2,F10.6)

The IVAR(I) are the sequence numbers of the Z-matrix variables which are to be varied simultaneously using stepsize STEPSZ. A couple of examples should illustrate this.

Example 3B.5 RHF/STO-3G geometry optimization of one bond length in H_2O.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD   1:  $  1 3
CARD   2:        WATER MOLECULE  RHF/STO-3G OPTIMIZATION
CARD   3:   0 1
CARD   4:    8
CARD   5:    1    1 0.96
CARD   6:    1    1 0.96      2 109.471
CARD   7:
CARD   8:   1                                  .01
CARD   9:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Card 1 is the usual system control and specifying RHF/STO-3G geometry optimization.

Cards 2-7 specify the title, charge, multiplicity, and initial geometry for the water molecule. The two bond lengths are variables 1 and 2, while the angle is variable 3.

Card 8 indicates to the optimization program (L202) that the first variable (the first bond length) is to be optimized using a stepsize of .01Å.

Card 9, a blank card, terminates the reading of optimization parameters.

Since only one of the two bond lengths is being optimized, the initial C_{2v} symmetry of this molecule will be lost during the optimization. If this is undesirable, both bond lengths may be varied simultaneously and equally, as is illustrated in the following example.

Example 3B.6. RHF/STO-3G optimization of two variables simultaneously with equal stepsizes.

```
                00000000011111111112222222222333333333344444444445555555555
                12345678901234567890123456789012345678901234567890123456789
CARD   1: $  1 3
CARD   2:     WATER MOLECULE      RHF/STO-3G OPTIMIZATION
CARD   3:   0 1
CARD   4:    8
CARD   5:    1    1 0.96
CARD   6:    1    1 0.96      2 109.471
CARD   7:
CARD   8:   1 2                              0.01
CARD   9:
                12345678901234567890123456789012345678901234567890123456789
                00000000011111111112222222222333333333344444444445555555555
```

Cards 1-7 are the same as in the previous example.

Card 8 specifies that variables 1 and 2 (the two bond lengths) are to be varied simultaneously using the stpsize 0.01Å. Note that although two z-matrix variables are involved, this is a single parameter optimization.

Card 9 terminates the list of parameters.

The next example shows how both bond lengths can be optimized independently.

Example 3B.7.

```
                00000000011111111112222222222333333333344444444445555555555
                12345678901234567890123456789012345678901234567890123456789
CARD   1:  $  1 3
CARD   2:     WATER MOLECULE, RHF/STO-3G OPTIMIZATION
CARD   3:   0 1
CARD   4:    8
CARD   5:    1    1 0.96
CARD   6:    1    1 0.96      2 109.471
CARD   7:
CARD   8:   1                                  0.01
CARD   9:   2                                  0.01
CARD  10:
                12345678901234567890123456789012345678901234567890123456789
                00000000011111111112222222222333333333344444444445555555555
```

Cards 1-7 are the same as before.

Card 8 indicates that variable 1 in the Z-matrix (the first bond length) is to be optimized using the stepsize 0.01Å.

Card 9 indicates that variable 2 is to be optimized (after optimization of the first parameter).

Several variables can be varied simultaneously using different stepsizes in the following way. On the first card, list those to be variables to be varied by the first step size (in the usual format), and terminate the list by '99'. This indicates that the parameter description continues on the next card. On the next card list those variables to be varied by the second stepsize. If there are only two different stepsizes, terminate the list by a blank; otherwise, use '99' to indicate continuation on the following card. A maximum of four continuation cards is allowed. The following example illustrates the use of this feature.

Example 3B.8:

```
           00000000011111111112222222222333333333344444444445555555555
           12345678901234567890123456789012345678901234567890123456789
CARD  1: $  1 3
CARD  2:    WATER MOLECULE, RHF/STO-3G OPTIMIZATION
CARD  3:  0 1
CARD  4:   8
CARD  5:   1    1 0.96
CARD  6:   1    1 0.96      2 109.471
CARD  7:
CARD  8:  1 299                                0.01
CARD  9:  3                                    1.0
CARD 10:
           12345678901234567890123456789012345678901234567890123456789
           00000000011111111112222222222333333333344444444445555555555
```

Cards 1-7 are the same as before.

Card 8 indicates that variables one and two (the two bond lengths) are to be optimized, varying both simultaneously using stepsize 0.01Å. However, the '99' as the last non-zero variable indicates that the parameter description continues on the next card.

Card 9 indicates that variable three (the bond angle) is also to be varied but a stepsize of 1.0° is used.

Card 10 terminates the list of parameters.

Example 3B.8 (cont'd)

After the calculation at the initial geometry, both bond lengths are incremented by 0.01Å, the bond angle is incremented by 1.0°, and the second calculation is performed. Thus, although three variables are being varied, this is a single parameter optimization.

The parameter specifications are read by L202 one at a time. The first time that L202 is executed with IOP(5)=1, the first parameter specification is read in. After that, a new parameter specification is read every fourth entry to L202 until a blank card is encountered.

The input for a potential surface scan is very similar to that for a geometry optimization. Once again, this data is read in the first time that L202 is executed with IOP(5)=2. The specification of the parameter to be varied is the same as for geometry optimizations, except that an additional card preceding the parameter specification must indicate the number of steps for that parameter (Format I2)). The following example illustrates a potential surface scan.

Example 3B.9

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:         $  1 4
CARD  2:              WATER MOLECULE  POTENTIAL SCAN
CARD  3:          0 1
CARD  4:           8
CARD  5:           1    1 0.96
CARD  6:           1    1 0.96      2 109.0
CARD  7:
CARD  8:         4
CARD  9:         3                                  1.0
CARD 10:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Card 1, the system control card, requests as RHF/STO-3G potential surface scan.

Cards 2-7 specify the title, charge, multiplicity, and initial geometry for the H_2O molecule.

Card 8 indicates that 4 points on the potential surface are to be calculated.

Card 9 specifies that for each point, the bond angle (variable 3) is to be incremented by 1.0°.

Card 10 terminates the parameter specification.

This job will perform 4 energy calculations with HOH bond angles of 109.0°, 110.0°, 111.0°, and 112.0°. As in geometry optimizations, parameter variation may involve simultaneous incrementation of several variables with equal or different stepsizes. This is illustrated in the following example.

Example 3B.10

```
           00000000011111111112222222222333333333344444444445555555555
           12345678901234567890123456789012345678901234567890123456789
CARD  1:  $  1 4
CARD  2:      WATER MOLECULE  RHF/STO-3G POTENTIAL SCAN
CARD  3:   0 1
CARD  4:    8
CARD  5:    1    1 0.96
CARD  6:    1    1 0.96      2 109.471
CARD  7:
CARD  8:   4
CARD  9:   1 299                                0.01
CARD 10:   3                                    1.0
           12345678901234567890123456789012345678901234567890123456789
           00000000011111111112222222222333333333344444444445555555555
```

Cards 1-7 are the same as before.

Card 8 requests four points on the potential surface.

Card 9 indicates that variables 1 and 2 are to be varied by 0.01Å each step, and that the parameter specification continues on the next card.

Card 10 specifies that variable 3 also be varied each step, but with stepsize 1.0°.

Card 11 terminates the list of parameters.

3-B.2 Data Input to L301

It is unnecessary to supply data input to link 301 in those calculations employing a standard (internally stored) basis. It is possible, however, for the user to supply input data to L301 to

(a) perform limited modification of internally stored basis sets or

(b) input non-internally stored (i.e. external) basis sets.

Various options, discussed below, control the use of these input features.

Before describing the exact nature of the various input options, it is necessary to fully define the concept of a shell of functions introduced in Section 1. A shell is a set of basis functions φ_μ with shared exponents. Gaussian 76 supports 4 kinds of shells: s-shells, p-shells, d-shells and sp-shells. An s-shell contains a single s-type function consisting of up to 6 primitive gaussian s-type functions. A p-shell contains the three functions p_x, p_y and p_z likewise built up from one to six primitive p-type gaussian functions. Since p_x, p_y and p_z primitive functions share exponents, they constitute a primitive shell. d-shells may be defined to contain the six second-order gaussians d_x^2, d_y^2, d_z^2, d_{xy}, d_{xz}, d_{yz} or the five d-functions $d_z^2-r^2$, $d_x^2-y^2$, d_{xy}, d_{xz}, d_{yz}. Finally, an sp-shell consists of up to 6 primitive sp-shells (a primitive sp-shell contains an s-function and a set of p-functions (p_x, p_y and p_z) with a common gaussian exponent).

The complete specification of a basis function $\varphi_\mu^{(i)}$ appearing in the i-th shell is*

$$\varphi_\mu^{(i)} = \sum_{k=1}^{N} d_{k\mu} \times \exp(-\alpha_k f_i^2 r^2)$$

In the above equation, φ_μ is expanded as a linear combination of N primitive gaussian functions X exp $(\alpha_k f_i^2 r^2)$. N is frequently referred to as the <u>degree of contraction</u>. Note that the contraction coefficients $d_{k\mu}$ depend on both k and μ indicating that functions of different angular momentum in a shell are not constrained to have the same coefficients. α_k is the exponent of the primitive gaussian function and f_i is the <u>scale factor</u> associated with the i-th shell. X provides the angular description of the basis function and may take on the values indicated in Table . Note that a d-shell can be defined to contain the true d-functions (IOP(8)=1) or the six second-order gaussians (IOP(8)=2). Completing the definition of $\varphi_\mu^{(i)}$, r is the distance from the center of the function and x,y,z, etc. appearing in Table 2 are the cartesian coordinates relative to the center of the function.

To further illustrate the features described above, consider the series of basis sets STO-3G, 4-31G and 6-31G* for the carbon atom. The STO-3G basis on any first row atom consists of only 2 shells. One shell consists of a set of 3 gaussian functions least-squares fitted to a Slater 1s orbital with an appropriate scale-factor; for the carbon atom, the 1s orbital scale-factor is 5.67. The other shell is a least-squares fit of

*The super-script i will generally be included only when it is necessary to indicate the shell to which φ_μ belongs.

TABLE 2. Values of X for s-, p- and d-shells.

Shell	X
s	1
p	x, y or z
d	x^2, y^2, z^2, xy, xz or yz (second order gaussians
	$3z^2-r^2$, xz, yz, x^2-y^2 or xy (true d functions)

3 gaussians to a Slater 2s and 2p orbitals with the constraint that the s and p functions have equal exponents. For carbon atom the 2 sp-shell has a scale factor of 1.72. The 4-31G basis on a first row atom has three shells. One shell is a contraction of four primitive s-type gaussians with a scale factor of unity. The second shell is a combination of three primitive sp-shells, again with a scale factor of unity. The third shell consists of a single sp-function with a scale factor of 1.04. The 6-31G* basis has four shells and differs from the 4-31G basis in only two respects. First, the innermost shell is a contraction of six gaussians instead of four; and secondly the last shell is a d-shell (second order gaussian).

With these definitions established, we discuss now the exact nature of L301's input options, considering internal basis set modification first. Once a basis set has been specified by setting IOP(5) to IOP(8), it may be subsequently altered in either or both of two ways: reading in replacement scale-factors and reading in replacement polarization exponents.

Replacement scale factors are read in when IOP(10) = 1 or 3. The format of the required cards is (15I2, 5F10.4). One specifies the desired centers (up to 15) in the numbering established by the order in which the centers were read in in L202 and the scale-factor to be applied to each shell in the set of shells for the particular atoms. Table 4 tabulates the correlation between the floating-point fields on the cards read and the shells to which the new scale-factors are applied. Input is terminated by a blank card.

Example 3B.11 Use of scale-factor input.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789012345678
CARD  1:         $  1
CARD  2:            2                                   1
CARD  3:            3      1                           11          1
CARD  4:            4                                  1
CARD  5:            5                               1 1
CARD  6:
CARD  7:               WATER, RHF/STO-3G, REPLACEMENT OF SCALE FACTORS
CARD  8:         0 1
CARD  9:          8
CARD 10:          1    1 0.96
CARD 11:          1    1 0.96      2 109.471
CARD 12:
CARD 13:         1                                 7.66        2.275
CARD 14:         2 3                               1.28
CARD 15:
                 12345678901234567890123456789012345678901234567890123456789012345678
                 00000000011111111112222222222333333333344444444445555555555
```

In this example, cards 1,2 and 4 to 12 should already be familiar. On card 3, evaluation of the dipole moment integrals has been omitted (L303 is not selected) merely because we have also omitted overlay 6. IOP(10) on card 3 is selected, and this causes cards 13 to 15 to be read by program GINPUT. Card 13 specifies replacement scale factors for the oxygen atom, center 1. We are using the standard value for the 1s orbital scale-factor and the Slater's rule value (2.275) for the valence shell. The standard 1s orbital scale-factor for both hydrogen atoms is replaced by the optimum STO-3G value for the water molecule (1.28). Card 15 terminates scale-factor input.

When IOP(10) = 2 or 3, the program will read in one card (4E20.10) specifying replacement polarization exponents (hydrogen atoms in the N-31G** basis) and (first row N-31G* atoms or second row STO-NG* atoms). The fields on the card are further detailed in Table 5 . If a field is left empty, the program will use the default (i.e. internally stored value for the particular atom/basis).

TABLE 4. Assignment of the Floating-Point Fields on Scale-Factor Replacement Cards.

Field(s)	Columns	Shells Affected		
		STO-NG	N-31G[a]	NN-31G
1	31-40	1s	1s	1s
2	41-50	2sp	$2sp_i$[c]	2sp
3	51-60	3sp[b]	$2sp_o$[c]	$3sp_i$[c]
4	61-70	-	-	$3sp_o$[c]

[a]Also N-31G* and N-31G**

[b]Second row atoms only

[c]Subscript i refers to 'inner' and subscript o refers to 'outer'.

TABLE 5. Definition of Fields on Replacement Polarization Exponent Card.

Field	Columns	Replacement
1	1-20	N-31G** hydrogen p-exponent
2	21-40	N-31G* first row atom d-exponent
3	41-60	STO-NG* d-exponent for Na or Mg
4	61-80	STO-NG* d-exponent for Aℓ to Cℓ.

Example 3B.12 6-31G* calculation with replacement polarization exponents.

```
                 00000000011111111112222222222333333333344444444445555555555666666
                 12345678901234567890123456789012345678901234567890123456789012345
CARD   1:  $  1
CARD   2:     2                                                  1
CARD   3:     3161  2                                          111          11
CARD   4:     4                                                  1
CARD   5:     5                                                  1
CARD   6:     6                                                  1
CARD   7:
CARD   8:     METHANOL, RHF/6-31G*, REPLACEMENT OF POLARIZATION EXPONENTS
CARD   9:   0 1
CARD  10:    6
CARD  11:    8    1 1.43
CARD  12:    1    1 1.09      2 109.471221
CARD  13:    1    2 0.96      1 109.471221    3 180.0
CARD  14:    1    1 1.09      2 109.471221    4 +60.0
CARD  15:    1    1 1.09      2 109.471221    4 -60.0
CARD  16:
CARD  17:                          0.9
                 12345678901234567890123456789012345678901234567890123456789012345
                 00000000011111111112222222222333333333344444444445555555555666666
```

In Example 3B.12, a single RHF/6-31G* calculation is performed. The polarization exponent α_d on both the carbon atom and the oxygen atom is 0.9 instead of its usual value of 0.8. No other modifications are made to the basis. As one further example of basis set modification, consider the vinyl-cation at RHF/6-31G**.

Example 3B.13. Vinyl cation with modified scale-factors and modified polarization exponents.

```
            00000000011111111112222222222333333333344444444445555555555666666
            12345678901234567890123456789012345678901234567890123456789012345
CARD  1:  $  1
CARD  2:     2                                        1
CARD  3:     3162  3                                 111        11
CARD  4:     4                                       1
CARD  5:     5                                       1
CARD  6:     6                                       1
CARD  7:
CARD  8:     NON-CLASSICAL VINYL CATION, RHF/6-31G** (MODIFIED)
CARD  9:   0 1
CARD 10:   -1
CARD 11:    6    1 0.607
CARD 12:    1    2 1.065     1 177.0
CARD 13:    6    1 0.607     2 180.0          3 180.0
CARD 14:    1    4 1.065     1 177.0          3 0.0
CARD 15:    1    1 1.1507    2 90.0           3 180.
CARD 16:
CARD 17:  1.15
CARD 18:   1 3                                 1.0        1.0        1.0
CARD 19:   2 4 5                              1.0        1.0
CARD 20:
            12345678901234567890123456789012345678901234567890123456789012345
            00000000011111111112222222222333333333344444444445555555555666666
```

This example demonstrates how to modify both scale factors and polarization exponents. Note that the polarization exponent card is read before the scale-factor modification cards. Additionally, since the field assigned to α_d is blank, the internally stored value of 0.8 is used. The p-exponent is taken as 1.15 instead of the standard value of 1.1. The atom optimized scale-factor (1.0 throughout) are specified for both the carbon and hydrogen atoms. Note in particular the numbering of the various centers; the dummy atom is not counted.

External basis sets may be read in to Gaussian 76 by setting IOP(5)=7 and IOP(6)=0. Further specification of the desired external basis can be achieved by using IOP(8) (refer to Appendix for a listing of the overlay 3 options). All external basis input is handled by routine GBASIS.

The basic unit of information that routine GBASIS deals with is the <u>shell definition block</u>. A shell definition block, together with IOP(8), contains all necessary information to define a shell of functions. It consists of a shell descriptor card and from zero to six primitive gaussian cards. The shell descriptor card has 5 fields with format (A4, A6, A4, I2, F10.4), and contains IBASIS, IORB, ITYPE, NGAUSS and SC. ITYPE defines the shell type and shell constraint and may take on the values '^^^S', '^^^P', '^^^D', '^^SP' and '^SPD' for respectively an s-shell, p-shell, etc. NGAUSS specifies the number of primitive gaussian shells in the contraction for the shell being defined and must be between 1 and 6. The shell scale-factor is given by SC, IBASIS and IORB control where the exponents and coefficients (α_k and $d_{k\mu}$) of primitive gaussian shells are taken from. When IBASIS = '^^^^', the α_k and $d_{k\mu}$ are read in from cards and the contents of IORB are used only in the printing of the gaussian function table. The format of the cards defining the α_k and $d_{k\mu}$ is 4E20.10; field 1 contains α_k and fields 2, 3 and 4 contain the s-, p- and d-coefficient for the particular primitive shell. NGAUSS such cards are required. The only acceptable non-blank value for IBASIS is '^STO' indicating that the α_k and $d_{k\mu}$ are to be taken from the internally stored STO-NG bases. In this case, IORB determines the type of STO-NG shell that is constructed; the possible values are given in Table 6. ITYPE, although redundant, is required; NGAUSS and SC have the same meanings as outlined above.

One customarily places at least one, and quite often several shells on any given nuclear center. A center definition block consists of a center identifier card, and one shell definition block for each shell desired on the center(s) specified, and is terminated by a card with '****' in columns 1-4. The center identifier card has 35I2 format and simply gives the numbers of the centers on which the basis functions are to be placed.

The description of input to GBASIS is now complete except for the first card and the last card. The first card (80I1) contains the degree of contraction for each shell in the calculation. Centers are delimited by a contraction of 0 and the list is terminated by a 9. If necessary, the program will read additional cards until a 9 is encountered. Overall input to GBASIS is terminated by a blank card.

TABLE 6. STO-NG Shells Defined by IORB.

IORB	Type of STO-NG Shell Defined	Required Type of Shell*
'~~~~1s'	1s-orbital	s
'~~~~2s'	2s-orbital	s
'~~~~2p'	a set of 2p-functions	p
'~~~2sp'	2s-orbital + a set of 2p functions with the S=P constrains	sp
'~~~~3s'	3s-orbital	s
'~~~~3p'	a set of 3p-functions	p
'~~~3sp'	3s-orbital + a set of 3p-functions with the S=P constraing	sp

*To be placed right justified in field 3 of the shell descriptor card.

Example 3B.14 Triplet methylene (5s3p/3s) uncontracted.

```
            00000000011111111112222222222333333333344444444445555555555666666
            12345678901234567890123456789012345678901234567890123456789012345
CARD  1:  $  1 1 0 2 0 0 0 1
CARD  2:      METHYLENE RADICAL, TRIPLET (5S3P/3S) VAN DUIJNEVELDT BASIS
CARD  3:   0 3
CARD  4:    6
CARD  5:    1    1 1.08
CARD  6:    1    1 1.08      2.130.0
CARD  7:
CARD  8:  1111111101110111 9
CARD  9:   1
CARD 10:            S1    S 1 1.0
CARD 11:  456.243                1.0
CARD 12:            S2    S 1 1.0
CARD 13:  68.7971                1.0
CARD 14:            S3    S 1 1.0
CARD 15:  15.3852                1.0
CARD 16:            S4    S 1 1.0
CARD 17:  4.08163                1.0
CARD 18:            S5    S 1 1.0
CARD 19:  0.29683                1.0
CARD 20:            P1    P 1 1.0
CARD 21:  4.18735                0.0                  1.0
CARD 22:            P2    P 1 1.0
CARD 23:  0.85405                0.0                  1.0
CARD 24:            P3    P 1 1.0
CARD 25:  0.19977                0.0                  0.0
CARD 26:  ****
CARD 27:   2 3
CARD 28:            S1    S 1 1.0
CARD 29:  4.5018                 1.0
CARD 30:            S2    S 1 1.0
CARD 31:  0.681444               1.0
CARD 32:            S3    S 1 1.0
CARD 33:  0.151398               1.0
CARD 34:  ****
CARD 35:
            12345678901234567890123456789012345678901234567890123456789012345
            00000000011111111112222222222333333333344444444445555555555666666
```

In this example, a fully uncontracted basis is employed giving rise to 20 atomic orbital basis functions. Card 1 is the system control card, cards 2 to 7 specify the route and cards 8 to 13 specify the nuclear geometry. On card 3, IOP(5)=7 indicating to the program that an external basis is to be read from cards; cards 14 to 41 constitute the requested input. Card 14 specifies the number of shells on each of the three centers and the number of primitive gaussian shells in each shell. In more detail, the first 8 columns all contain one's indicating that the first center contains 8 shells each consisting of one primitive shell. The zero in column 9 indicates that the next string of digits is for the next center, in this case center 2. The one's punched in columns 10, 11 and 12 instruct GBASIS that center 2 will have 3 shells, each containing one primitive shell. The zero in column B again serves as a punctuation mark stepping the program to the next center. The last center receives 8 shells consisting of one function each and the list is terminated by a 9.

Cards 15 to 32 and cards 33 to 40 constitute two different <u>center identification blocks</u>. The functions defined by cards 16 to 31 are applied to center 1 (the carbon atom) and the functions defined by cards 34 to 39 are placed on each of the hydrogen atoms.

Each of the pairs of cards 16-17, 18-19, 20-21, 22-23, 24-25, 26-27, 28-29, 30-31, 34-35, 36-37, 38-39 constitutes a <u>shell definition block</u>. Let's examine cards 26 and 27 more closely. The first field on card 26, the shell descriptor card, is blank, indicating that the gaussian exponents and contraction coefficients for the shell will be read on the following cards. The 1 in the fourth field specifies that the degree of contraction is 1 and indicates to GBASIS that only 1 primitive gaussian card is to be read in following the shell descriptor card. The alphanumerie string '^^^^P1' in

Example 3B.14 (cont'd)

field two of card 26 is merely used as a print variable in the printout of the gaussian function table. The 'P', right-justified in field 3 of card 26 declares this shell to be a p-shell, containing a set of p-functions (px,py and pz). Finally, the last field contains 1.0, supplying the scale-factor for this shell.

Card 27 is an example of a primitive gaussian card. Here, there are 4 fields, all E20.10. The first field supplies the exponent, α=4.18735, for this shell. Fields 2 and 4 (starting in columns 21 and 61 respectively) are blank, indicating that the s- and d-coefficients for this shell are zero. The p-coefficient, in field 2 of card 27, is non-zero (1.0).

The specification of an s-shell, e.g. cards 16 and 17, is quite similar, the only differences being the contents of the third field on the shell descriptor and ('^^^S') and the location of the non-zero coefficient on the primitive gaussian card (column 21).

Example 3B.15. 6-31G** basis, fed in as external basis.

In this example, two input decks will be presented. The end result from either deck is the same, limited of course by the small precision differences encountered in reading the basis in as compared to using the internally stored version.

```
                 00000000011111111112222222222333333333344444444445555555555666666
                 12345678901234567890123456789012345678901234567890123456789012345
CARD  1: S  1.1 0 1 6 2
CARD  2:       WATER MOLECULE, RHF/6-31G**, EXPERIMENTAL GEOMETRY
CARD  3:  0 1
CARD  4:   8
CARD  5:   1    1 0.957
CARD  6:   1    1 0.957    2 104.53
CARD  7:
                 12345678901234567890123456789012345678901234567890123456789012345
                 00000000011111111112222222222333333333344444444445555555555666666
```

The above deck introduces nothing new--the program will perform a 6-31G** (6-31G basis augmented by the addition of a set of second-order gaussians to the oxygen atom and addition of a set of p-functions to each hydrogen atom) calculation on water molecule at the specified geometry. Consider now the following deck:

```
           00000000011111111112222222222333333333344444444445555555555666666
           12345678901234567890123456789012345678901234567890123456789012345
CARD  1: $  1
CARD  2:    2                                       1
CARD  3:    37  2                                  111        11
CARD  4:    4                                      1
CARD  5:    5                                      1
CARD  6:    6                                      1
CARD  7:
CARD  8:      WATER MOLECULE, RHF/6-31G**, EXPERIMENTAL GEOMETRY
CARD  9:  0 1
CARD 10:   8
CARD 11:   1    1 0.957
CARD 12:   1    1 0.957    2 104.53
CARD 13:
CARD 14: 6311031103119
CARD 15:  1
CARD 16:          1S    S 6 1.0
CARD 17: 5484.671631          0.00183107443 4
CARD 18: 825.2349472          0.01395017223
CARD 19: 188.0469589          0.06844507810
CARD 20: 52.96449995          0.2327143364
CARD 21: 16.89757037          0.4761928981
CARD 22: 5.799635351          0.3585208543
CARD 23:        2SPI   SP 3 0.99
CARD 24: 15.85513341          -0.1107775494        0.07087426819
CARD 25: 3.673026830          -0.1480262615        0.3397528280
CARD 26: 1.034345224          1.130767003          0.7271585763
CARD 27:        2SPO   SP 1 0.98
CARD 28: 0.2811389230         1.0                  1.0
CARD 29:           D    D 1 1.0
CARD 30: 0.80                 0.0                  0.0                    1.0
CARD 31: ****
CARD 32:  2 3
CARD 33:          1S    S 3 1.20
CARD 34: 13.007734            0.03349460434
CARD 35: 1.96207942           0.2347269535
CARD 36: 0.444528953          0.8137573262
CARD 37:          1S    S 1 1.5
CARD 38: 0.1219491560         1.0
CARD 39:           P    P 1 1.0
CARD 40: 1.15                 0.0                  1.0
CARD 41: ****
CARD 42:
           12345678901234567890123456789012345678901234567890123456789012345
           00000000011111111112222222222333333333344444444445555555555666666
```

Relative to example 3B.14 , the following new features are employed:

1. Shells containing more than one gaussian. (See card 14 and note that two of the shell definition blocks have 3 gaussians contracted together and one block (the oxygen 1s-orbital) has 6).

2. This example utilizes sp-shells, two of which are present on the oxygen atom, one a contraction of three primitive sp-shells, the other uncontracted. Note in particular the use of '^^SP' in field 3 of the cards 23 and 27.

3. Use of d-shells. Cards 29 and 30 define a d-shell to be placed on the oxygen atom. Because IOP(8)=2 on card 3, the six second-order gaussians (x^2, y^2, z^2, xy, xz and yz) are used.

Note that both decks in example 3B.15 will produce the same final energy. Obviously the second deck requires considerably more effort and presents greater opportunity for error.

Routine GBASIS has a limited capability to utilize some internally stored bases in external basis calculations. Currently, this feature is only programmed for the STO-NG bases.

Example 3B.16. Hydrogen cyanide, STO-3G external basis input.

```
                 00000000011111111112222222222333333333344444444445555555555666666
                 12345678901234567890123456789012345678901234567890123456789012345
CARD  1:  $  1
CARD  2:     2                                          1
CARD  3:     37                                        111        1
CARD  4:     4                                         1
CARD  5:     5                                         1
CARD  6:     6                                         1
CARD  7:
CARD  8:       HYDROGEN CYANIDE, RHF/STO-3G
CARD  9:   0 1
CARD 10:    7
CARD 11:    6    1 1.153
CARD 12:    1    2 1.070     2 180.0
CARD 13:
CARD 14:  33033039
CARD 15:   1
CARD 16:   STO     1S    S 3 6.67
CARD 17:   STO    2SP   SP 3 1.95
CARD 18:  ****
CARD 19:   2
CARD 20:   STO     1S    S 3 5.67
CARD 21:   STO    2SP   SP 3 1.72
CARD 22:  ****
CARD 23:   3
CARD 24:   STO     1S    S 3 1.24
CARD 25:  ****
CARD 26:
                 12345678901234567890123456789012345678901234567890123456789012345
                 00000000011111111112222222222333333333344444444445555555555666666
```

Comparing this example with the previous external basis examples, the obvious difference is the absence of the primitive gaussian cards; the program takes the necessary data from internally stored tables. Note that the contents of field two is no longer arbitrary--it specifies the type of STO-NG functions obtained as outlined in Table 7. The contents of the third field must be chosen from the appropriate entry in column 3 of Table 7. The number of gaussians in the contraction is given by the fourth field. The scale factors, field five, have been chosen in this example to coincide with the internally stored values for carbon, nitrogen and hydrogen.

As a final example of GBASIS input, consider the following example of a geometry optimization utilizing the basis re-use feature.

Example 3B.17 Geometry optimization with general basis.

```
                 00000000011111111112222222222333333333344444444445555555555666666
                 12345678901234567890123456789012345678901234567890123456789012345
CARD  1:  S  1
CARD  2:     2                                   1
CARD  3:     37                                 111        1
CARD  4:     4                                  1
CARD  5:     5                                   1
CARD  6:     6                                  1
CARD  7:     21                                  1
CARD  8:     371                                11         1
CARD  9:     5                                   1
CARD 10:     21                                  1                   -2
CARD 11:     371                                1 1
CARD 12:     6                                  1
CARD 13:
CARD 14:       SULFUR OXIDE, UHF/STO-3G OPTIMIZATION
CARD 15:   0 3
CARD 16:   16
CARD 17:    8    1 1.472
CARD 18:
CARD 19:  33310339
CARD 20:   1
CARD 21:   STO     1S    S 3 15.47
CARD 22:   STO    2SP   SP 3 5.79
CARD 23:   STO    3SP   SP 3 2.05
CARD 24:            D    D 1 1.0
CARD 25:  0.39                  0.0                 0.0                   1.0
CARD 26:  ****
CARD 27:   2
CARD 28:   STO     1S    S 3 7.66
CARD 29:   STO    2SP   SP 3 2.25
CARD 30:  ****
CARD 31:
CARD 32:   1                                0.01
CARD 33:
                 12345678901234567890123456789012345678901234567890123456789012345
                 00000000011111111112222222222333333333344444444445555555555666666
```

Example 3B.17 (cont'd)

In this example, card 1 is the system control card. Cards 2 to 13 is the control input supplying the route information. Cards 14 to 18 specify the initial geometry and are read in when the program executes L202 the first time (card 2). The input to routine GBASIS is cards 19 to 31; note the mixture of internally stored STO-3G shells and an external shell (the polarization function) on the sulfur atom. The five d-functions ($3z^2-r^2$, x^2-y^2, xy, xz and yz) are used because IOP(5)=7 and IOP(8)=0 on the two integral route cards (3 and 8 in this example). When the basis is first processed (card 3), it is stored in the RW-files for possible later use; this is always done for external bases. When link 301 is subsequently re-entered (card 8) the program will recover the previously stored basis (IOP(5)=7 and IOP(6)=1) thus eliminating the need to read in the bases again. (In earlier versions of the program, particularly Gaussian 70, it was necessary to read the basis in for each point calculated, four times in the current example.) Cards 32 and 33 are read in subsequent re-executions of L202 (cards 7 and 10).

4. Control Input for Compound Jobs

4A. External Initial Guesses

A compound job consists of two executions, the second of which uses output from the first. The most common type of compound job is one in which the converged MO coefficients from one job are read as an initial guess to a subsequent job. RHF/STO-3G MO coefficients, for instance, may be used as an initial guess in a RHF/6-31G* calculation. In this case, the first job must save the necessary information, and the second must read it back in.

Output from a job can be saved by using the "punch" options. The information is not really punched onto cards, but rather is written onto FORTRAN unit 8. The contents of this unit can then be saved on a disk file, for instance, to be read later by another job. Usually, for the initial guess, both the basis set and MO coefficients from the first job should be saved, and the following example illustrates how this can be done:

Example 4A.1 Write out basis and MO coefficients from RHF/STO-3G single point calculations of H_2O.

```
                 00000000011111111112222222222333333333344444444445555555555666666
                 12345678901234567890123456789012345678901234567890123456789012345
CARD  1:  $  1
CARD  2:     2                                   1
CARD  3:     3                              1   111        1
CARD  4:     4                                  1
CARD  5:     5                              1   1
CARD  6:     6                                  1
CARD  7:
CARD  8:       WATER MOLECULE   SAVE MO COEFS
CARD  9:   0 1
CARD 10:    8
CARD 11:    1    1 0.96
CARD 12:    1    1 0.96      2 109.471
CARD 13:
                 12345678901234567890123456789012345678901234567890123456789012345
                 00000000011111111112222222222333333333344444444445555555555666666
```

This is just the usual route for a single point RHF/STO-3G calculation, except that IOP(32) is set to "1" in overlays 3 and 5. In overlay 3, this indicates that L301 (GINPUT) should write the basis set onto unit 8, while in overlay 5, it causes the final MO coefficient to be written. The user must insure that unit 8 is saved on permanent mass storage for later use (by use of the appropriate CDC control statements, for instance).

The route above can be generated automatically by specifying the "punch" parameter (column 20) on the "dollar sign" card:

```
Card 1:  $  1 1 0 0 0 0 0 0 1
            ^^ ^ ^ ^ ^ ^ ^ ^
```

This card generates a route identical to that in Example 15. The '1' in column 20 indicates that both the basis set and the final MO coefficients are to be written to unit 8.

If the proper options are set in L401, this data can be read in a subsequent job, and used as an initial guess. The data is read from unit 9. The key option is IOP(5) in L401, which specifies the type of guess:

IOP(5) ... Type of guess.

0 ... Projected Huckel-type

1 ... Diagonalize core Hamiltonian

2 ... Huckel type

3 ... Read density matrix from unit 9.

4 ... Read density from unit 9 and project into space of current basis.

5 ... Read MO coefficients from unit 9.

6 ... Read MO coefficients from unit 9 and project these into the space of the current basis.

If the density matrix or MO coefficients to be read are in a different basis than that of the current job, then projection must be specified (IOP(5) = 4 or 6). In this case, the guess (in the current basis) is formed by choosing those coefficients (or density matrix elements) which give the best (least-squares) fit to the data read.

The previous example produced as output the basis set and final MO coefficients. This data is used as input in the next example.

Example 4A.2 Read converged MO coefficients from unit 9.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1: $  1
CARD  2:    2                                        1
CARD  3:    3                                      111          1
CARD  4:    45                                     1
CARD  5:    5                                      1
CARD  6:    6                                      1
CARD  7:
CARD  8:      WATER MOLECULE    READ MO COEFS
CARD  9:  0 1
CARD 10:   8
CARD 11:   1    1 0.96
CARD 12:   1    1 0.96      2 109.471
CARD 13:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

This is the usual route for a single RHF/STO-3G calculation, except that in overlay 4, IOP(5) is set to 5. This option indicates that the initial guess MO coefficients are to be read in, and the program assumes that these are available on unit 9. The above route can be generated automatically with the single card:

```
Card 1:  $ 1 1 0 0 0 0 5
           ^^ ^ ^ ^ ^ ^
```

The '5' in column 16 specifies the type of guess, and is the same as IOP(5) in overlay 4.

In the following examples, two input decks are given: the first writes MO coefficients to unit 8, and the second reads these from unit 9.

Example 4A3. Water molecule. RHF/STO-3G MO coefficients used as a guess for RHF/6-31G* calculation.

Deck 1: Save RHF/STO-3G basis and coefficients.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD   1:  $  1 1                  1
CARD   2:     WATER MOLECULE, RHF/STO-3G
CARD   3:   0 1
CARD   4:    8
CARD   5:    1    1 0.96
CARD   6:    1    1 0.96      2 109.471
CARD   7:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Deck 2: RHF/6-31G* calculation using projected guess.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD   1:  $  1 1 0 1 6 1 6
CARD   2:     WATER MOLECULE, RHF/6-31G*, READ GUESS
CARD   3:   0 1
CARD   4:    8
CARD   5:    1    1 0.96
CARD   6:    1    1 0.96      2 109.471
CARD   7:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Deck 1 in the above example writes to unit 8 both the basis set and final MO coefficients for RHF/STO-3G water molecule. Deck 2 reads this data from unit 9 and uses it to form an initial guess in the RHF/6-31G* run.

Example 4A.4 RHF/4-31G MO coefficients used as guess in RHF/6-31G* calculation (water molecule).

Deck 1: Save coefficients from RHF/4-31G calculations.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:         $  1 1 0 1 4 0 0 0 1
CARD  2:            WATER MOLECULE, RHF/4-31G, SAVE MO COEFS
CARD  3:          0 1
CARD  4:           8
CARD  5:           1    1 0.96
CARD  6:           1    1 0.96      2 109.471
CARD  7:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Deck 2: RHF/6-31G* calculation. Initial guess is read and projected.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:         $  1 1 0 1 6 1 6
CARD  2:            WATER MOLECULE, RHF/6-31G*, GUESS FROM RHF/4-31G
CARD  3:          0 1
CARD  4:           8
CARD  5:           1    1 0.96
CARD  6:           1    1 0.96      2 109.471
CARD  7:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

4B. Restarts of geometry optimizations

It sometimes happens that an optimization run dies before the optimization is completed (max time exceeded, etc.). In this case, the restart feature may be useful. The following data input is necessary for a restart of a geometry optimization.

1. Title
2. Charge, Multiplicity
3. Specification of initial geometry by a Z-matrix
4. Total energy of initial geometry. FORMAT (4X,E20.10).
5. Variable parameter specification. (Same format as described in Section 3B.)
6. Total energy of incremented geometry. Format (4X, E20.10)
7. Total energy of third geometry. Format (4X, E20.10)
8. A card with stars in columns 1-4
9. Blank card.

All of this data is read by L202 when restart is specified by IOP(6)=1. The usual input to L202 is given by items 1-3. The "initial" geometry specified should be that from the last successful extrapolation. If the total energies (6) and (7) are not available, simply omit these; the card with stars indicates that no more energies are available. If optimization of additional parameters is desired, specification of these should precede the blank card (9). Some examples may clarify this:

Example 4B.1 Restart of RHF/STO-3G optimization of H_2O.

```
                 00000000011111111112222222222333333333344444444445555555555
                 12345678901234567890123456789012345678901234567890123456789
CARD  1:         S  1 S
CARD  2:            WATER MOLECULE, RESTART OF RHF/STO-3G OPTIMIZATION
CARD  3:          0 1
CARD  4:           8
CARD  5:           1    1 0.96
CARD  6:           1    1 0.96      2 109.471
CARD  7:
CARD  8:              -74.960723
CARD  9:          1 2                                0.01
CARD 10:              -74.961642
CARD 11:         ****
CARD 12:          3                                   2.0
CARD 13:
                 12345678901234567890123456789012345678901234567890123456789
                 00000000011111111112222222222333333333344444444445555555555
```

Card 1 specifies restart of an RHF/STO-3G geometry optimization.

Cards 2-7 give the usual input to L202.

Card 8 gives the total energy at the initial geometry.

Card 9 specifies a variable parameter and the stepsize.

Card 10 gives the total energy at the incremented geometry.

Card 11 indicates that no more energies are available, and the normal optimization will pick up here.

Card 12 specifies another variable parameter and stepsize.

Card 13 terminates the parameter list.

The route automatically generated by the system control card (1, above) is identical to that for a normal geometry optimization, except that IOP(6)=1 on the first entry to L202. This option indicates that this is an optimization restart.

5. List of Options

As indicated in section 2, various options (IOP(5) to IOP(34) are available on an overlay basis. These are read in by overlay control cards or are implicit in a selected standard route. A given option is valid for all links within an overlay.

In this section, we list the options available to users of GAUSSIAN 76/QCPE.

5A. Options for Overlay 2

L202 generates the cartesian coordinates of each nucleus in a molecule and stores them in blank common for later use by other links. On a normal run (see IOP(5)), L202 needs a matrix, which specifies the molecular geometry in terms of internuclear distances and angles. This information is then converted into x, y, and z coordinates. However, this link can, if desired, read the nuclear coordinate directly from cards.

During a geometry optimization, L202 is executed several times. The first time (a normal entry), it constructs the initial nuclear coordinates of the molecule from the Z-matrix that it read in. The second execution of L202 is a "continuation of the geometry optimization" (see IOP(5)). At this point, it reads from cards a geometry parameter to be optimized, and increments the coordinates in blank common accordingly. Subsequent entries to L202 are also 'continuations', but a new parameter is read only every fourth entry. By incrementing the specified geometry parameter and remembering the total energies calculated for each geometry, the program can find the minimum energy value for the parameter.

IOP(5) ... Type of entry to L202.

0 ... This is a normal entry to L202, read in Z-matrix.

1 ... This is a continuation of a geometry optimization, using the Z-matrix.

2 ... This is a continuation of a potential surface scan, using the Z-matrix.

3 ... Normal entry, read in coordinates in angstrom units.

4 ... Normal entry, read in coordinates in atomic units.

IOP(6) ... Whether this is a restart

0 ... No. This is a normal run.

2 ... Yes. This is a restart of a geometry optimization.

IOP(7) ... Punching of restart cards.

0 ... Do punch these.

1 ... Suppress punching of the restart cards.

IOP(8) ... Test of second derivative during optimizations.

0 ... Abort optimization run if negative second derivative is found.

1 ... Do not abort. Extrapolate to extremum.

IOP(9) ... Printing of distance matrix.

0 ... Print the distance matrix.

1 ... Suppress printing the distance matrix.

IOP(10) ... Fixing of near-tetrahedral angles.

0 ... Do not look for near-tetrahedral angles

1 ... Will interpret 109.5° as the tetrahedral angle accurate to 17 figures.

IOP(11) ... Print reduction.

0 ... Regular print is produced

1 ... All printing except the title, charge, and multiplicity is suppressed. Printing of distance matrix is controlled by IOP(9).

5B. Options for Overlay 3

The options for overlay 3 are conveniently divided into those options that specify the basis to be used (IOP(5) to IOP(9)) and other miscellaneous infrequently used options (IOP(22) to IOP(34)) (debug, activation of special features, etc.). In what follows, the options are merely summarized. Further information regarding shell definitions and input examples are given in Section 3B. Users are advised to exercise caution in using external bases, non-standard features and the like; problems and conflicts can arise. Use of some options produce voluminous output (e.g. IOP(33) = 3 or 4) while others can cause the program to run slowly (breaking the s=p constraint can easily double the two-electron integral evaluation time).

IOP(5) and IOP(6) ... Specification of basis set.

0 ... Minimal STO-NG basis where N is specified by IOP(6).

1 ... Split valence N-31G basis where N is specified by IOP(6).

7 ... External basis. If IOP(6)=0, the basis is read in from cards. When IOP(6)=1, the basis in the RW-files is merged with the coordinates in blank common to produce the new basis. This feature is useful in geometry optimizations employing an external basis by eliminating the necessity to read the basis in for each energy calculation in the optimization.

IOP(7) ... Addition of polarization functions to <u>internal</u> basis sets.

0 ... No polarization functions are added to the basis set.

1 ... IOP(5)=0, an uncontracted d-shell is added to each second row atom, converting the basis to STO-NG*.

IOP(5)=1, an uncontracted d-shell is added to each first row atom, converting the basis to N-31G*.

2 ... IOP(5)=0, the STO-NG* basis, described above is produced.

IOP(5)=1, in addition to placing a d-shell on first row atoms, the basis is further extended by the addition of an uncontracted p-shell to hydrogen atoms, producing the 6-31G** basis.

IOP(8) ... Selection of second-order gaussians/true d-functions for d-shells.

0 ... The selection is determined by the basis:

N-31G* ... 6 second-order gaussians,

N-31G** ... 6 second-order gaussians,

STO-NG* ... 5 true d-functions,

General basis ... 5 true d-functions.

1 ... All d-shells consist of the 5 true d-functions ($3z^2-r^2, x^2-y^2, xy, xz, yz$)

2 ... All d-shells consist of the 6 second-order gaussians ($x^2, y^2, z^2, xy, xz, yz$)

IOP(9) ... Specification of overall shell constraint.

0 ... s, s=p or d, depending on the angular quantum number.

1 ... s, p or d, depending on the angular quantum number.

2 ... s, s=p or s=p=d, depending on the angular quantum number.

IOP(10) ... Modification of internally stored bases.

0 ... No modification is performed.

1 ... Read in replacement scale factors. Details on the use of this option are provided in the section describing input to L301.

2 ... Read in replacement polarization exponents. Further information on the use of this option is given in the section describing input to L301.

3 ... Read in both replacement scale-factors and replacement polarization exponents. Note that the program will read the replacement polarization exponents before reading the replacement scale-factors.

IOP(11) to IOP(21) ... inactive in G76/QCPE.

IOP(22) ... Partial control of short-cuts and cutoffs in the two-electron d-integral program.

0 ... Only regular cutoffs and short-cuts are taken.

1 ... Special cutoffs, parameterized for the 6-31G* basis, are invoked, substantially reducing the integral evaluation time. Use of this option is <u>not</u> recommended for other bases (at least not without extensively examining the effects).

IOP(23) ... Definition of two-electron integral scale-factor (not to be confused with the scale-factors that appear in the basis set). Two-electron integrals are then stored as integers.

0 ... 10^8 (the normal value) is selected.

N ... 10^{8+n} is used.

IOP(24) ... Control of printing of gaussian function table.

0 ... The table is printed only if non-standard features are employed (e.g. external basis, replacement scale-factors, etc.).

1 ... The table is printed.

2 ... The table is not printed.

IOP(25) and IOP(26) ... Specification of last two-electron integral link.

(0,0) ... The last link on the overlay 3 card is assumed to be the last two-electron link and has the responsibility of closing the two-electron file.

(N,K) ... NK is a two-digit octal number specifying the last two-electron link.

IOP(27) ... Handling of small two-electron integrals.

0 ... Integrals with magnitude less than 10^{-6} are discarded.

N ... Discard integrals with magnitude less than 10^{-N}.

IOP(28) ... Special accuracy option.

0 ... No special features are involved.

1 ... Inactive in G76/QCPE.

2 ... Compute all two-electron integrals in L312, program UNCON. Note that Link 311 must be omitted from the route for this option to work properly.

IOP(29) ... Coordinate rotation.

0 ... Coordinates are not rotated.

1 ... Read one card (3E20.10) specifying the three Euler angles φ, θ and χ.

IOP(30) to IOP(31) ... inactive in G76/QCPE.

IOP(32) ... Punch option.

0 ... No punching is performed.

1 ... Common /B/ is punched (via BINWT).

IOP(33) ... Print option.

0 ... No integrals are printed.

1 ... All one-electron integrals are printed.

3 ... Print two-electron integrals in standard format.

4 ... Print two-electron integrals in debug format.

5 ... Combination of 1 and 3.

6 ... Combination of 1 and 4.

IOP(34) ... Dump option.

0 ... No dumping.

1 ... Various control words are dumped as they are modified or examined, and common /B/ is printed by L301.

2 ... Additionally, common /B/ is printed in each integral link and all the integrals are printed.

5C. Options for Overlay 4

This is a program which produces an initial guess to the solution of the SCF equations. This guess is in the form of molecular orbital coefficients and/or density matrices which are stored on the appropriate read-write files. The steepest descents procedure (Link 503)(G78 only) requires MO coefficients as an initial guess, while the classical SCF procedures (Links 501 and 502) require density matrices. Since a density matrix can be produced from the MO coefficients, but not vice versa, the former is a more constraining requirement.

There are several ways in which this guess may be produced. One easy way is to diagonalize the core Hamiltonian. In general, this is not a very good guess, but it is applicable to any basis set, and is available as an option.

Another type of guess is called the Huckel guess, which is modeled after extended Huckel MO theory. Essentially, the initial guess is formed from internally stored constants (for more details see subroutine Huckel). These constants were determined from studies on internal minimal and split valence basis sets (STO-3G, 4-31G, 6-31G), so the use of this type of guess with bases other than these is not recommended.

This Huckel guess can be applied to other bases in the following ways. The guess MO coefficients are formed from internal data as if there were an STO-3G basis set on the molecule. The guess MO vectors in the desired basis are then formed by choosing the vectors which give the best (least-squares) fit to those described in the STO-3G basis. Since this will usually produce fewer than NBASIS vectors, the MO coefficient matrix must be completed with orthonormal vectors of the proper symmetry if the

full matrix is required (Link 503). This procedure is called a projected Huckel guess, and is applicable to any basis set.

A still better type of guess, usually, is to read the coefficient or density matrix from cards. If the matrix read in is for a basis other than the one used in the current run, the matrix can be projected (by a least-squares fit) into the desired basis. Since the projected MO vectors can be normalized and orthogonalized, and this is not possible for a projected density matrix, projection of MO coefficients usually produces a better guess than projection of the density matrix.

The options which determine the initial guess procedure to be used are described below.

IOP(5)... Type of guess

... 0 Projected Huckel guess

... 1 Huckel guess

... 2 Diagonalize the core Hamiltonian

... 3 Read in the density matrix from cards.

... 4 Read from cards the density matrix in another basis and project it into the space of the current basis (see also IOP(6)).

... 5 Read the MO coefficients from cards.

... 6 Read from cards the MO coefficients in another basis and project these into the space of the current basis set (see also IOP(6)).

IOP(6) ... Input of previous basis set information. This is used only when IOP(5)=4 or 6.

... 0 The basis set information is to be read from cards in compressed form before the MO coefficient or density matrix is read in.

... 1 The MO coefficient or density matrix read in is for a minimal basis. No basis set input is necessary.

IOP(7) ... Removal of degeneracies (only for blocked Huckel guess).

... 0 Scan the diagonal elements and remove coincidental degeneracies.

... 1 Do not scan diagonal elements.

IOP(8) ... Alteration of configuration.

... 0 Do not alter configuration.

... 1 Read in pairs of integers indicating which pairs of MO's are to be interchanged. Pairs are read until a blank card is encountered.

... 2 Read in (80I1) the occupation level of each MO. Normally values of only zero or one are used. Larger integers may be used to partially occupy an MO (OCC=1/N). This option should not be used for the PHF procedure.

Note:

If alteration of configuration is done on an open shell system, the program will expect two sets of data as described above: The first for the alpha MO's and the second for beta.

IOP(9) ... Off-diagonal scale factor used in forming Huckel matrix.

... 0 Default (K = 1.65)

... N (K = N*0.875)

IOP(10) ... Diagonal elements for Huckel matrix

... 0 Determine from internal data

... 1 Read in diagonal elements (8F10.4). For a blocked Huckel guess, the program will read NBASIS values. For a projected Huckel guess, it will read MBASIS values, where MBASIS is the number of basis functions in a minimal basis set for the system.

IOP(11) ... Completion of coefficient matrix

... 0 Do not complete

... 1 After a projected guess, complete the coefficient matrix.

IOP(12) ... Fractional occupation of MO's in partially filled degenerate sets.

... 0 Do not adjust occupation levels.

... 1 Adjust the occupation levels so that each of the orbitals in a degenerate set is equally occupied.

IOP(13) ... Open or closed shell.

... 0 Use multiplicity to determine.

... 1 Closed shell.

... 2 Open shell.

IOP(14) ... Type of basis

... 0 Use ILSW to determine

... 1 Minimal basis

... 2 Extended basis

IOP(15) ... Polarization functions on hydrogen

... 0 Use ILSW to determine

... 1 No p functions on hydrogen

... 2 Put a set of p functions on hydrogen

IOP(16) ... Polarization functions on first row atoms

... 0 Use ILSW to determine

... 1 No d functions

... 2 Yes, a set of d functions on first row atoms

IOP(17) ... Polarization functions on second row atoms

... 0 Use ILSW to determine

... 1 No d functions on second row

... 2 Put a set of d functions on second row

IOP(18) ... Control of size of largest off-diagonal element in EIGEN.

See routine EIGEN for details.

IOP(19) ... Five or six d functions

... 0 Use ILSW to determine

... 1 Five d functions

... 2 six d functions

IOP(20) ... Print option

... 0 Print nothing

... 1 Print only the MO coefficients

... 2 Print the initial guess Fock matrix, the MO coefficients, and the density matrix.

IOP(21) ... Blocking of the overlap matrix (only for blocked Huckel guess).

... 0 Do not block

... 1 Block off the overlap matrix before calculating the transformation matrix. In order to reproduce a run from an earlier system, it is necessary to specify this along with IOP(5)=1.

5D. Options for Overlay 5.

The programs in overlay 5 seek solutions to the SCF equations:

$$\underset{\sim}{F}\,\underset{\sim}{C} = \underset{\sim}{S}\,\underset{\sim}{C}\,\underset{\sim}{E}$$

or their UHF equivalents. Starting with some initial guess at the solutions (IOP(5)) the program iterates either until the root-mean-squared difference in two successive density matrices is below the convergence threshold (IOP(6)), or until the maximum number of cycles is exceeded (IOP(7)). There are also extrapolation options (IOP(11)) which can help speed up the convergence.

IOP(5) ... Initial guess density matrix.

0 ... Already on read-write file. This is either the output from L401, or the result of a previous SCF calculation (during optimizations, for instance).

1 ... Density matrix is calculated from MO coefficients obtained by diagonalization of the core Hamiltonian.

2 ... The density matrix is to be read from unit 9. This must have been produced as the output of a previous job.

IOP(6) ... Convergence on density matrix.

0 ... Terminate iteration when the rms error in the density matrix is less than 5*10**(-5).

N ... Terminate iteration when density matrix has converged to 10**(-N).

IOP(7) ... Maximum number of cycles.

0 ... Terminate iteration after twenty cycles.

N ... Terminate iteration after 2**(N-1) cycles.

IOP(10) ... Re-use of the previously generated set of two-electron integrals. Note that on electron integrals must always be re-evaluated.

0 ... Normal run, no restart.

N ... Copy integrals from unit IUX(N) to IUX(1). See subroutine RESCLO.

IOP(11) ... Three-point extrapolation. This is a procedure for helping to achieve rapid convergence of the density matrix elements.

0 ... When possible, perform a 3-point extrapolation.

1 ... Suppress the extrapolation.

IOP(12) ... Entry mode

0 ... Normal entry, SCF is performed.

1 ... Null mode entry is made only for possible punch.

IOP(32) ... Writing MO coefficients and a density matrix to unit 8.

0 ... Nothing

1 ... Write out final MO coefficients.

2 ... Write out final density matrix.

3 ... Write out both.

4 ... Write out both each cycle of the SCF.

IOP(33) ... Printing of coefficients and density.

0 ... No printing.

1 ... Print final MO coefficients

2 ... Print final density matrix and MO coefficient

3 ... Print both each cycle of the SCF.

IOP(34) ... Dump option.

0 ... No dump

1 ... Turn on all possible printing.

5E. Options for Overlay 6

This program performs a Mulliken population analysis on a wave function. The options are described below.

IOP(5) ... Open or closed shell.

0 ... Use ILSW to determine.

1 ... Forced open shell.

2 ... Forced closed shell.

The remaining options are print/no-print options. If the value of the option is zero, the default value (given below) is assumed. If the option is set to 1, the information is printed, and if it is 2, the printing is suppressed.

0 ... Default

1 ... Print.

2 ... Do not print.

IOP(6) ... Distance matrix. Default: No-print.

IOP(7) ... Molecular orbital coefficients. Default: Print.

IOP(8) ... Density matrix. Default: No-print.

IOP(9) ... Full population analysis. Default: Print.

IOP(10) ... Gross orbital charges. Default: Print.

IOP(11) ... Gross orbital type charges. Default: No-print.

IOP(12) ... Condensed to atoms. Default: Print.

References

1. W. S. Hehre, W. A. Lathan, R. Ditchfield, M. D. Newton and J. A. Pople, "Gaussian 70, a series of programs designed to perform ab initio molecular orbital calculations on organic molecules", Program No. 26, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Indiana.

2. Roothaan, C. C. J., Rev. Mod. Phys. 23, 69 (1951).

3. Pople, J. A. and R. K. Nesbet, J. Chem. Phys. 22, 571 (1954).

4. J. B. Collins, P.v.R. Schleyer, J. S. Binkley, J. A. Pople, J. Chem. Phys. 64, 5142 (1976).

5. Hehre, W. J., R. F. Stewart and J. A. Pople, J. Chem. Phys. 51, 2657 (1969).

6. Ditchfield, R., W. J. Hehre and J. A. Pople, J. Chem. Phys. 54, 724 (1971).

7. Hariharan, P. C. and J. A. Pople, Chem. Phys. Lett. 16, 217 (1972).

8. Mulliken, R. S., J. Chem. Phys. 23, 1833, 1841, 2338, 2343 (1955).

9. Gaussian 78 is a further development of Gaussian 76 and is as yet unpublished.