Tutorial on Using Gaussview and Gaussian 94

Written by Vijay Gupta, with editing by M.L. and S.A.

Overview
     Gaussian 94 takes a text file with a .com extension as an input. In this input file, the molecular configuration of the molecule is described as well as the specific calculations that you want to perform (geometrical optimization, frequency determination, single point energy, etc) and which methods you want to use(HF, DFT, etc.). After running Gaussian, a text output file is generated with the extension .log and the same name as the input file. The output file can be browsed to find selected data.

     Rather than using text files as input/output, life is made easy by a graphical interface, Gaussview. This program lets you draw the molecule of interest and then it will generate the text input file for you. It will also read in the output file, so that you can obtain information about the molecule without having to sort through the long output file (as well as animating the vibrational frequencies, for instance.)

     The objective of using Gaussian in this course is to understand better chemical reactions and transition state theory, in addition to the computations of thermochemistry. With this in mind, our objective is to give you simple, illustrative calculations, and to sacrifice accuracy for speed and simplicity. Additional information can be found at the web sites below:

General Gaussian site:
http://www.gaussian.com/
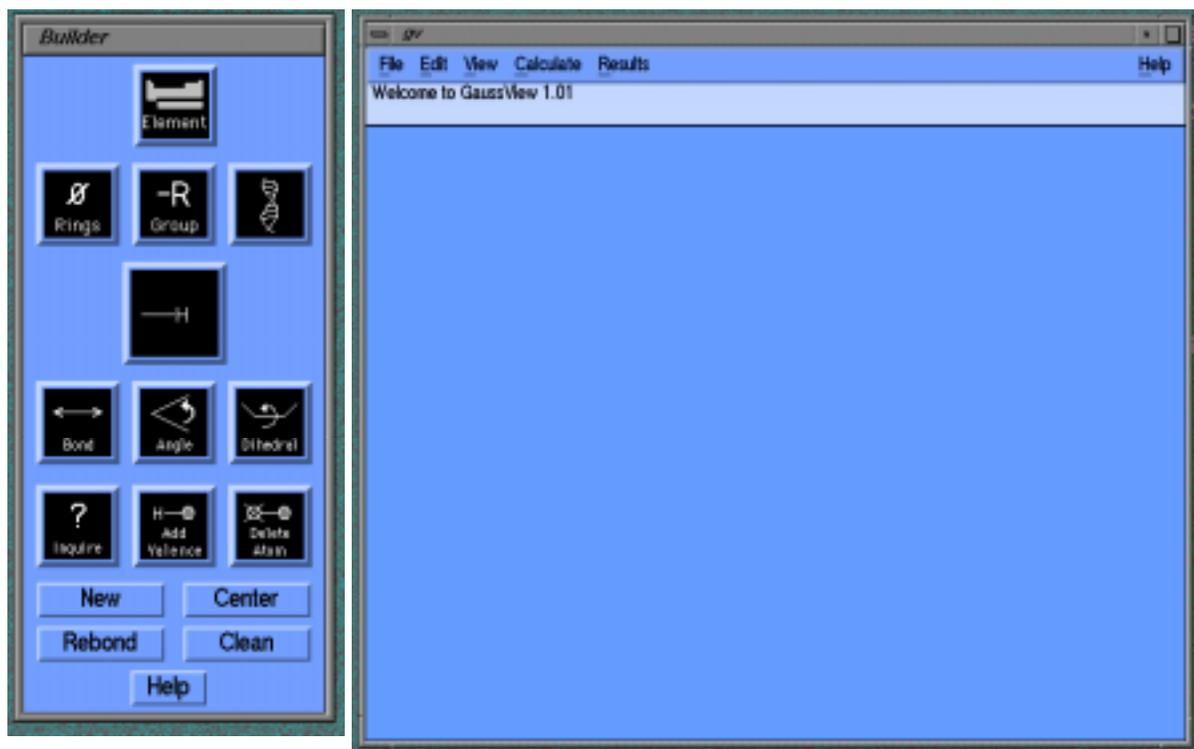
Gaussian98 manual:
http://www.gaussian.com/g98.htm

Also note that Gaussian's units of energy is the Hartree,
1 Hartree = 27.21 eV = 627.46 kcal/mol

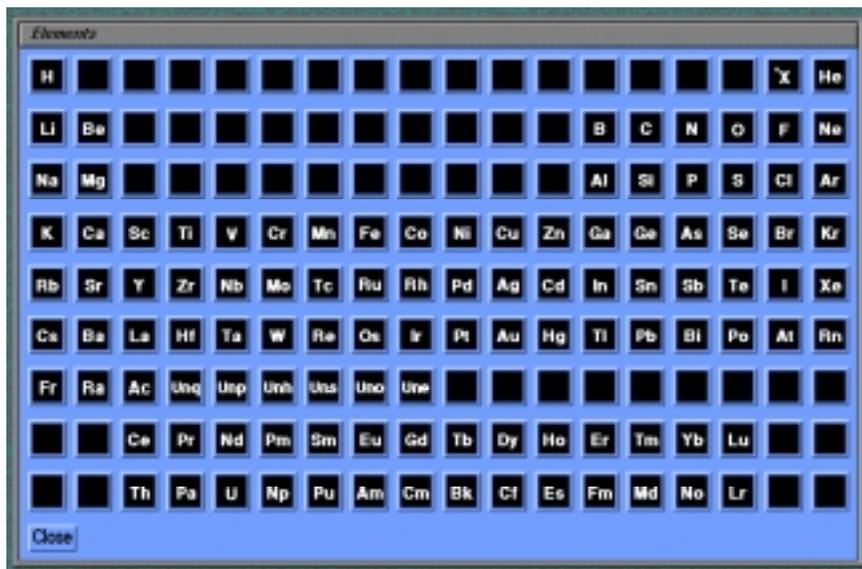# Part I Tutorial on the Graphical Interface

From an SGI:

Start Gaussview

This will open up two windows, the *Builder* Menu and the *GV* window

To build a molecule, click on *Elements* in the *Builder* menu with the left mouse button
This will open up the *Elements* window. (All selections are made with the left mouse button)


For example, to build, C_2H_6:

Select C with the left mouse button and click in the *GV* window. Repeat this procedure to draw another C atom at a different location. Notice that each C atom already has 4 H atoms attached to it.
Click on delete atom option (in *Builder*) and then choose one of the H in the CH_4 molecule.
Click on *Bonds* in *Builder* and then choose the two C atoms one after another. Click on the single bond to connect the two C atoms together. The windows will look something similar to what is shown below. Press OK to complete.
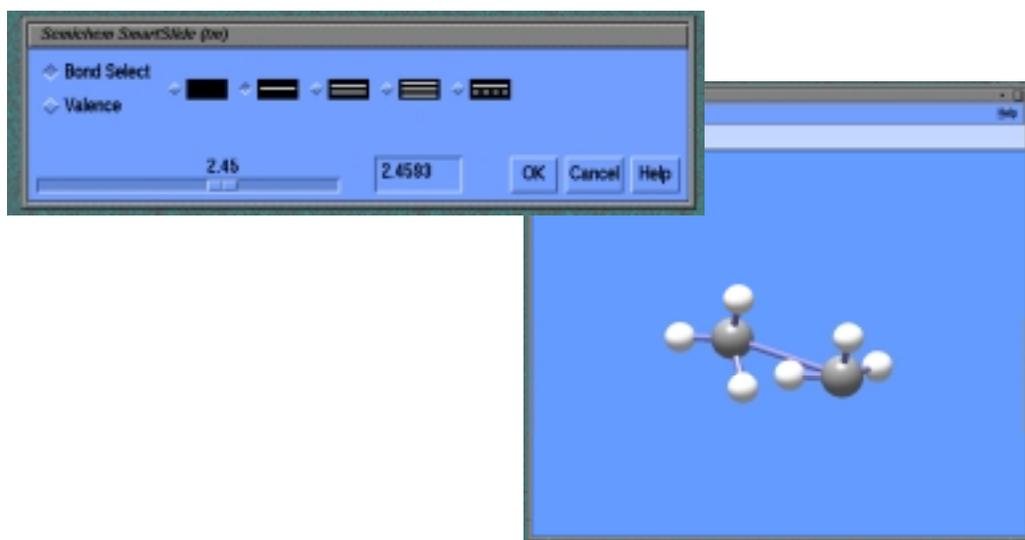
Press Clean in the Builder Menu to clean up the structure. *You will need to press "Clean" before saving your final structure.* It is easy to forget to do this when building a new molecule. Clean places the molecules in initial guess locations based on standard steric information. This initial guess helps the actual computation converge. If you make any changes to the molecule, you have to press "Clean" again.
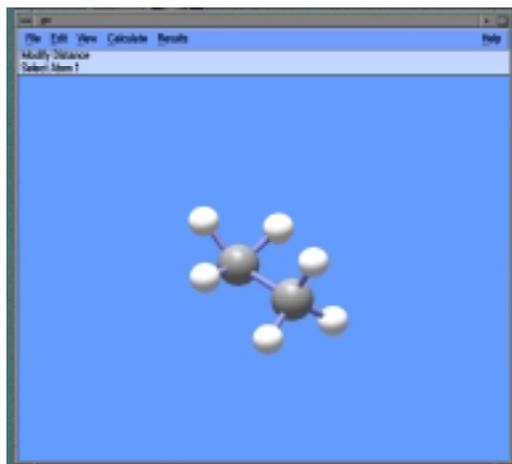
To substitute one atom for another, say C with N, select *Element*, click on N and then click on one of the C atoms in the *GV* window. The number of H atoms also change from 3 on C to 2 on N.
To save the structure, use the pull-down menu *Save* under *Files*.
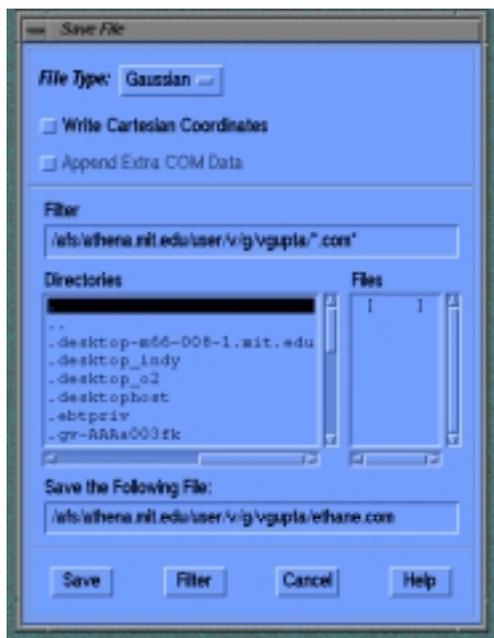
Type in the *filename.com* in the window and click on *Save*.

At this point, you're ready to run Gaussian94. Go to Part II to learn how
to generate the output file. The following section deals with analyzing the generated .log file.

Analyzing Gaussian Results using Gaussview

Gaussview can read in Gaussian output files in the following formats
(a) log files
(b) unformatted checkpoint files
(c) formatted checkpoint files

To open an output file, Use *File – Open* and select the appropriate file.

This will open the molecular structure in the *GV* window.
Bond / Angles / Dihedrals options in the *Builder* can be used to measure the various geometric
parameters be selecting the desired atoms.

To view the vibrational frequencies, select *Result – Vibrations*
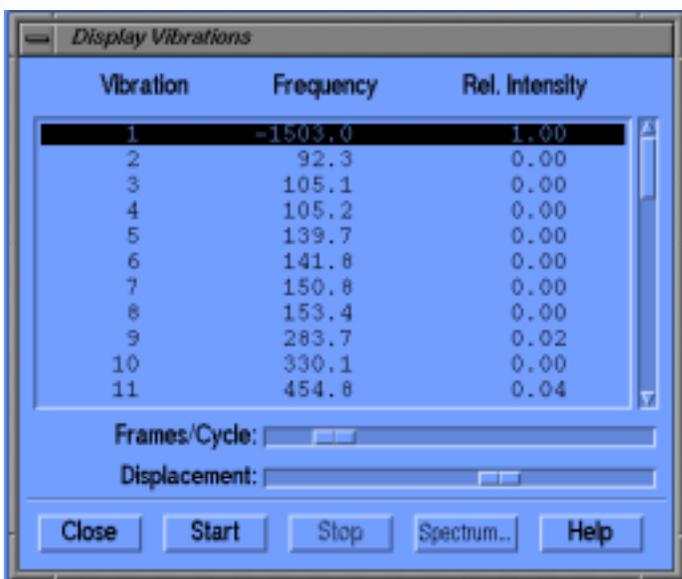
Select the frequency value and press *Start* to animate the frequency.
The vibrational modes corresponding to the selected freq will be animated in the *GV* window.

Part II Running Gaussian94

Submitting Gaussian Runs:

Even though Gaussian jobs can be submitted from the Gaussview interface, the recommended approach is to use Gaussview to build the molecular structure and save it in a file.
Once Gaussview has generated the .com file, you can open it with a text editor and modify the settings for calculations and methods.

Use any text editor to edit the file and modify the parameters such as method, basis set, calculation type etc. and submit the jobs using the command-line.

A few things to note:

Checkpoint files can become really huge and your disc quota may be exceeded.
Do not store checkpoint files unless you need them to do further calculations in the future. Also, if you do have to store them, you may want to zip them.

To zip and unzip a file, use the commands gzip and gunzip
Example: gzip test.chk   (this will create the zipped file test.chk.gz)
          gunzip test.chk.gz

Other useful unix commands are:
`grep, egrep, tail, more, top, jobs, kill`
For more information on these commands, check the manual pages.

EXAMPLES of input files for some common calculations:

```
********************************************************************
%chk=water.chk

# BLYP/3-21G Opt

Water Geom

0 1
O
H    1    R2
H    1    R3   2    A3

R2   0.9797
```

```
R3    0.9797
A3    105.0
```

****

Explanation:

%chk=water.chk	sets the name of the checkpoint file
	blank line
BLYP	Choice of method (other options are HF, B3LYP, MP2 etc.)
3-21G	Choice of basis set (other options are 6-31G, 6-31G** etc)
Opt	Optimize Geometry (other options could be:
	Freq for frequency Analysis
	SP for single point energy calculation
	Another blank line
Water Geom	Comment
	Another blank line
0 1	charge and multiplicity

This is followed by the geometry specification (in z–matrix notation here)
Geometry specification should be terminated by a blank line

The blank lines are required as they signify end of one specification and the beginning of the next.

The format of the input file for other calculations remains the same. The problem specification and directives on the 3d line are the only major changes. For example:

NOTE: You must use the same basis set and functional when running Optimization and Frequency calculations. If you want to start with a smaller basis set (e.g. 3-21g) and work up to a larger basis set (e.g. 6-31g(d)) for a frequency calculation, you will need to rerun the optimization portion again with the larger basis set. However, you can use the smaller basis set results as initial conditions for the larger basis set calculations. This is discussed in the section below.

To do an optimization calculation using a previously computed geometry (optimized earlier at a lower level). Here, the geometry is read in from the checkpoint file along with the Guess for the initial wavefunction.

```
%chk=water.chk

# BLYP/3-21G Opt Geom=Check Guess=Read

Water Geom

0 1
```
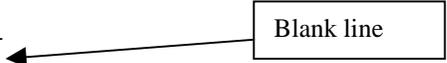

Blank line

Note: There is no geometry specification as it is being read in from the checkpoint file.
Be sure to leave a blank line after the "0 1" line!

To do a geometry optimization followed by a frequency analysis

```
%chk=water.chk

# BLYP/3-21G Opt Freq

Water Geom

0 1
…
```

The Frequency calculations also return Thermochemistry information such as entropy and Internal energies.

To do a transition state search

```
%chk=ts.chk

# B3LYP/6-31G Opt=(TS,NoEigentest)

Transition State Search

0 1
...
```

To restart a geometry optimization (freq jobs cannot be restarted)

```
%chk=water.chk
```

```
# BLYP/3-21G Opt=Restart

Water Geom Restarted

0 1
...
```

To do a single point energy calculation at a fixed geometry

```
%chk=water.chk

#P BLYP/3-21G Geom=Check Guess=Read

Water Geom

0 1
```

The "P" in the third line tells Gaussian to output results with extra details. Gaussian defaults to the SP calculations when nothing is specified. Otherwise, it will perform the SP along with whatever else is requested. SP is always done because minimizing energy is needed when optimizing molecule geometry.

Interpreting the Output

General Notes:
1) The general fate of any Gaussian G94 run can be determined from the return status of the run. If it returns "`Segmentation fault - core dumped`" or something similar, there was an error in your input, something failed to converge, or you exceeded some system resource like memory or disk space. You can look at the very last line of the output for "`Normal termination of Gaussian 94`" which means just that.
2) In the output samples shown below, some features are shown in a **bold face** font which has been added for emphasis but is not present in the original outputs.

I. The Optimization:

Sample input deck:

```
%chk=h2o
#p RHF/6-31g*  Opt

h2o

0   1
o
h 1 r1
h 1 r1 2 a2

r1=1.0
a2=105.0
```

The most interesting result of a geometry optimization is of course the geometry. You can get the geometry from the checkpoint file directly by using the newzmat utility (typing "`newzmat -ichk -ozmat filename.chk`" will produce a new input file, *filename*.`com`, containing the updated geometry.). If you're using it in a subsequent Gaussian calculation, you can read the geometry in directly using the `geom=check` directive--don't forget to leave a blank line after the charge/multiplicity card if you're reading in the geometry.

You can also read the optimized geometry from the output file. Simply look for the phrase: "`Optimization completed`" and a few lines later you'll find a table containing the optimized geometry.

```
Optimization completed.
    -- Stationary point found.
                        ----------------------------
                        !   Optimized Parameters   !
                        ! (Angstroms and Degrees)  !
 ------------------------                          ------------------------
 ! Name   Definition             Value          Derivative Info.           !
 --------------------------------------------------------------------------
 ! R1     R(2,1)                 0.9474         -DE/DX =   -0.0001          !
 ! R2     R(3,1)                 0.9474         -DE/DX =   -0.0001          !
 ! A1     A(2,1,3)             105.4905         -DE/DX =    0.              !
 --------------------------------------------------------------------------
```

Of course you'll also want to know the optimized HF energy, you can find this by finding the final HF calculation in the output. This is most safely done by searching backwards from the end of the file for "SCF Done: E(RHF) =". **The value following the equal sign is the total energy**

of the molecule at the optimized geometry in atomic units (Hartrees, 1 hartree=627.5095 kcal/mol). Note: be sure that you get the energy of the final iteration in the geometry optimization.

```
SCF Done:  E(RHF) =   -76.0107464891      A.U. after    8 cycles
             Convg  =    0.2436D-08              -V/T =   2.0020
             S**2   =    0.0000
 KE= 7.585922282212D+01 PE=-1.990813908029D+02 EE= 3.792390860188D+01
```

Alternatively, out can find the energy in the archive entry printed at the end of the output (unless you requested the test run or a runtype that cannot be archived). The archive entry contains a summary of the entire calculation and is best used with the Browse program, but some information can easily be plucked out of it. The HF energy is shown in **bold** here:

```
1\1\GINC-GAMERA\FOpt\RHF\6-31G(d)\H2O1\MECOLV\10-Aug-1995\0\\#P RHF/6-
31G* OPT\\h2o\\0,1\O,-0.0910031918,0.,-0.0698292051\H,-0.0950588122,0.
,0.8775905504\H,0.8230843465,0.,-0.3189569099\\Version=SGI-G94RevB.1\S
tate=1-A1\HF=-76.0107465\RMSD=2.436e-09\RMSF=7.886e-05\Dipole=0.686419
5,0.,0.5267082\PG=C02V [C2(O1),SGV(H2)]\\@
```

I. The Frequency Calculation:

Sample Input Deck:

```
%chk=h2o
#p rhf/6-31g* freq

h2o HF/6-31G* frequencies

0  1
o
h 1 r1
h 1 r1 2 a2

r1=0.9474
a2=105.4905
```

The goal of the frequency calculation is threefold:

1) To determine if the optimized structure is a minimum, a transition state (1 imaginary frequency), or a saddle point (2 or more imaginary frequencies).
2) To predict the quantum mechanical zero-point energy, the entropy, and the temperature-dependent vibrational, rotational, and translational energy terms.
3) To yield theoretical harmonic vibrational frequencies as well and IR and Raman intensities.

Before checking any of these it's prudent to check that the molecular geometry used in the frequency calculation was in fact a stationary point. This is easy because the frequency calculation includes a single geometry optimization step. To find this, search the frequency output for the word "Converged" (note the "C" is capitalized). Immediately following you'll

find a table indicating which of the convergence criteria are met--all should be "YES". If you have significant residual gradients, double check to make sure you input the correct geometry. If you're sure that's right, check to make sure the frequency calculation is being done at the same level of theory and basis set as the optimization.

```
          Item              Value      Threshold  Converged?
Maximum  Force            0.000176     0.000450      YES
RMS      Force            0.000126     0.000300      YES
Maximum  Displacement     0.000161     0.001800      YES
RMS      Displacement     0.000157     0.001200      YES
Predicted change in Energy=-1.528859D-08
Optimization completed.
    -- Stationary point found.
```

To find out what kind of stationary point this is, search (backward from convergence info) for the phrase: "Low frequencies". Any large negative frequencies on this line indicate that your structure is a transition state or saddle point. If the structure is not a minimum, you'll also see an announcement of the number of imaginary modes immediately following the low frequencies. Note that you should have six frequency values relatively close to zero (five for a linear molecule) which correspond to the rotational and translational degrees of freedom.

```
Low frequencies ---    -0.0028   -0.0026   -0.0024   33.9672   36.6972   41.4994
Low frequencies --- 1826.8582 4069.3547 4187.5090
```

Example of output for a transition state:

```
Low frequencies --- -671.2655   -21.9167    -0.0057    0.0015    0.0019    2.4815
Low frequencies ---    7.0108 1138.9586 1504.3657
 ******    1 imaginary frequencies (negative signs) ******
```

***The zero-point energy and related thermodynamical properties are clustered in a self-explanatory section under the heading "- Thermochemistry -". There you'll find the **Zero-point energy**, **entropy**, **partition function**, **moments of inertia** and other useful thermodynamic data. *Caveat*: Don't forget that these values are based on many approximations and may be inaccurate when the implicit assumptions are violated (e.g. when the molecule has very low-frequency vibrational modes.)

Note that the thermal correction values depicted on the next page already include the zero point energy (ZPE). Also, the above thermochemical values are for 298.15 K. To obtain the thermochemical data at any other temperature, use the freqchk command. Freqchk requires formatted checkpoint files. To generate the formatted checkpoint files use:

athena% formchk test.chk

Creates the file test.fchk

To run freqchk,

athena% freqchk test

```
Zero-point vibrational energy      60314.1 (Joules/Mol)
                                    14.41541 (Kcal/Mol)
VIBRATIONAL TEMPERATURES:   2628.43  5854.86  6024.86
         (KELVIN)

Zero-point correction=                           0.022972 (Hartree/Particle)
Thermal correction to Energy=                    0.025806
Thermal correction to Enthalpy=                  0.026750
Thermal correction to Gibbs Free Energy=         0.005375
Sum of electronic and zero-point Energies=           -75.987774
Sum of electronic and thermal Energies=              -75.984940
Sum of electronic and thermal Enthalpies=            -75.983996
Sum of electronic and thermal Free Energies=         -76.005371


                        E              CV              S
                     KCAL/MOL      CAL/MOL-KELVIN  CAL/MOL-KELVIN
TOTAL                16.194           5.985           44.988
ELECTRONIC            0.000           0.000            0.000
TRANSLATIONAL         0.889           2.981           34.609
ROTATIONAL            0.889           2.981           10.376
VIBRATIONAL          14.416           0.023            0.003
Zero-point vibrational energy      60314.1 (Joules/Mol)
                                    14.41541 (Kcal/Mol)
VIBRATIONAL TEMPERATURES:   2628.43  5854.86  6024.86
         (KELVIN)

Zero-point correction=                           0.022972 (Hartree/Particle)
Thermal correction to Energy=                    0.025806
Thermal correction to Enthalpy=                  0.026750
Thermal correction to Gibbs Free Energy=         0.005375
Sum of electronic and zero-point Energies=           -75.987774
Sum of electronic and thermal Energies=              -75.984940
Sum of electronic and thermal Enthalpies=            -75.983996
Sum of electronic and thermal Free Energies=         -76.005371


                        E              CV              S
                     KCAL/MOL      CAL/MOL-KELVIN  CAL/MOL-KELVIN
TOTAL                16.194           5.985           44.988
ELECTRONIC            0.000           0.000            0.000
TRANSLATIONAL         0.889           2.981           34.609
ROTATIONAL            0.889           2.981           10.376
VIBRATIONAL          14.416           0.023            0.003
```

Finally, the IR and Raman spectral information is in a table of normal modes just before the thermochemical data and is summarized in the form of synthetic spectra immediately following the thermochemical data.

For a more complete and detailed guide see: *Gaussian 94* User's Reference and Exploring Chemistry with Electronic Structure Methods, both available from Gaussian, Inc.

TIPS

To search for a word in a file one can use the command `grep`
```
grep SCF filename.log
```

To search and print out all lines containing one or more expressions in a file, use `egrep`
The following command can be used to follow the optimization process of a job:
```
egrep 'out of|SCF Done|YES|NO ' filename.log
```


If you are not interested in the Raman Intensities of the IR frequencies, you can save ~ 20% of CPU time by using the `Freq=NoRaman` directive frequency runs.

Transition states have one negative eigenvalue. Quite often the initial guess for a transition state will give more than one negative eigenvalues, and crash. To avoid testing for eigenvalues, one can use `Opt(TS,noEigentest)` option.