MedeA Gaussian: Expanding the Limits of Computational Chemistry

Contents

The following section is no replacement for the Gaussian manual, but rather shows to the experienced Gaussian user, how Gaussian calculations are performed from within *MedeA*.

1 Installing Gaussian

The Gaussian program should be installed according to the instructions in the Gaussian manual, and the installation directory should be added to the PATH environment variable. *MedeA* will then be able to automatically detect the location of all the executables required to run the Gaussian program.

2 Gaussian: Ab Initio Gausian Code

The user interface to Gaussian is based on flowcharts.

Once you have an active molecular system under periodic or non periodic conditions, bring up the interface by selecting Jobs >> New Job .

In the flowchart interface, clicking Gaussian: Ab Initio Gausian Code under the **Methods** section adds a Gaussian stage, you can double-click on the Gaussian box to open the flowchart. Once you have opened the Gaussian stage, you will find the following sections and stages for its flowchart.

3 Setup

There are three stages under the *Setup* section:

3.1 Set computational resources

This stage sets the **Maximum RAM** and **Maximum disk space** this Gaussian job can use. This stage is optional.

DOCUMENTATION

3.2 Configure basis and pseudopotentials

This stage is used to select a basis set for Gaussian calculations, including the representation of any pseudopotentials and density fitting basis sets used in DFT. This stage is **recommended**. The default basis set is 6-31G.

Selecting a basis set

Gaussian calculations are performed using a basis set of localized functions to represent the molecular orbitals. Compared to periodic basis sets, such as the plane waves that are typically employed in standard DFT calculations, these basis sets are very quick to evaluate and allow space to be included in a system without adding computational cost. However, given that there is a great deal of freedom in the form that these localized functions can take, there are many different varieties of basis set available. Full details of each basis set can be found in the Gaussian manual, but a basic procedure is to perform a quick geometry optimization using the 6-31G basis, followed by a subsequent geometry optimization and other types of calculations using one of the cc-pVnZ family. The 6-31G basis produces reasonable geometries for relatively little computational cost, and the cc-pVnZ family of basis sets allow an easy selection of the size of the basis (n = D, T, Q, 5, 6 in increasing quality and computational expense).

The periodic table in the basis set selection dialog highlights (in yellow) the atoms that can be represented by a specific basis set. As the selected basis is changed, the coverage of the periodic table is updated. It is also possible to select multiple atoms from the periodic table, and the list of available basis sets will contract so that only those basis sets that can be applied to all selected atoms are displayed. This enables you to choose a basis set that can be applied to all the different kinds of atom in the system.

Many basis sets can be augmented by polarization or diffuse functions, the form of which are specific to each basis set family. Broadly speaking, polarization functions represent mixtures of different orbital types, such as 's' and 'p', and diffuse functions enhance the quality of the basis in the outer region of the charge density. Full details of these options can be found in the Gaussian manual.

If the selected basis set is 'Custom', then the text input area at the bottom of the dialog is used to define the functions in the basis. Again, details of the required syntax can be found in the Gaussian manual.

The basis selected using this stage will be applied to all subsequent stages, unless overwritten by a subsequent basis set selection stage.

Selecting pseudopotentials

The use of a pseudopotential approximates the contribution of the innermost electrons with a model potential, reducing the computational complexity of the calculation. As with the basis set selection, choosing 'Custom' allows the form of the pseudopotentials to be specified manually. Consult the Gaussian manual for the required syntax.

Density fitting basis sets

For pure (non-hybrid) DFT calculations, a density fitting approximation can provide significant performance gains for calculations on medium-sized systems. This approach expands the density in a set of atom-centered functions when computing the Coulomb interaction instead of computing all of the two-electron integrals. In addition to predefined basis sets, a 'Custom' basis set can be created, using the same syntax as the regular basis set specification.

Charge/Spin multiplicity

Optionally, the charge and spin multiplicity of the system can be changed to any values after setting a mark in the check box on the left side in this tab.

Note: The values entered for charge and spin multiplicity in this tab are not checked for consistency with the number of electrons of the molecule, contrary to the Charge/Spin tab of the *Molecular Builder*.

3.3 Configure periodic boundary conditions

This stages sets the optional options such as:

- Use the Gamma point only
- Approximate number of k-points to use
- Maximum extent of image cells (N Bohr)
- Minimum number of image cells
- Maximum number of image cells
- Minimum number of image cells for DFT XC quadrature
- Minimum number of image cells for exact change

This stage is also optional. These parameters are only taken into account when the system is periodic.

4 Calculation Types

Gaussian calculations are performed using the Hartree-Fock (HF), density functional theory (DFT), Moller-Plesset (MP), and coupled cluster (CC) stages, each of which uses the computational method indicated by their name. Whilst MP and CC calculations can offer an increased accuracy compared to HF and DFT, the scaling of the computational effort required with the system size is far more dramatic. If N is some approximate measure of system size, the cost of HF and DFT is roughly proportional to N^3, MP is at least N^4 (for MP2), and CC is at least N^6 (for CCSD). For this reason, it may only be possible to use HF or DFT to study systems containing many atoms. Please note that if most dialog boxes in the Gaussian stages possess pulldown submenus with a few choices, you may also enter your text manually (e.g., by giving a numerical value instead of a default or automatic parameter).

4.1 Hartree-Fock (HF)

The HF stage presents a dialog containing the configurable parameters of the HF calculation. These are common to all the other calculation types; DFT calculations add an extra step in the HF self-consistency loop (an integration over the electron density), and MP and CC calculations are post-Hartree-Fock methods that require the computation of an initial HF wave function. This wave function can be one of two forms: restricted or unrestricted.

Restricted HF wave functions are composed of orbitals that are used by both the spin-up and spin-down electrons. Thus, systems containing equal numbers of spin-up and spin-down electrons will have each orbital doubly occupied. Unrestricted HF uses different orbitals for electrons with different spins, making it suitable for open shell molecules or systems with uneven numbers of spin-up and spin-down electrons.

4.2 Density functional theory (DFT)

In addition to the HF parameters, the DFT stage includes options for controlling the integration over the electron density and selecting the form of the exchange-correlation (XC) functional. Here, it is possible to select from a limited subset of XC functionals, classified as either 'pure', 'hybrid', or 'double hybrid'. Pure XC functionals allow the use of a density fitting basis set (see the Gaussian basis section of this documentation), which can greatly improve the efficiency of a calculation, but hybrid functionals can produce results which are more similar to those obtained from the experiment. Double hybrid functionals add an MP2-like correlation

---

to a DFT calculation, with a corresponding computational cost. It is also possible to specify an XC functional manually, which overrides all of the other settings in the XC selection GUI. Here you are free to enter any of the huge numbers of XC specifications that Gaussian can process - see the Gaussian manual for more details.

For some of the XC functionals listed in the GUI, it is also possible to add long-range corrections and empirical dispersion. For the latter, the options are PFD, GD2, GD3, and GD3BJ, corresponding to the Petersson-Frisch dispersion model from the APFD functional, and the D2 and D3 versions of Grimme's dispersion with GD3BJ including Becke-Johnson damping, respectively.

4.3 Moller-Plesset (MP)

The MP2, MP3, MP4, and MP5 methods are supported by Gaussian. For MP4, there are different variants available, corresponding to the types of substitutions included. Select MP4(DQ) to use only the space of double and quadruple substitutions, MP4(SDQ) for single, double, and quadruple substitutions, or MP4(SDTQ) for full MP4 with single, double, triple, and quadruple substitutions. As the order of the MP method and the number of substitutions increases, so does its computational cost. MP5 calculations are only possible using an unrestricted HF wave function.

By default, post-Hartree-Fock methods in Gaussian use a 'frozen core' approximation where the inner-shells are excluded from the correlation calculation. The corresponding tab in the GUI allows you to customize this procedure by selecting a freezing convention - see the Gaussian manual for full details.

4.4 Coupled cluster (CC)

The CC variants available in Gaussian are CCSD, CCSD(T), and CCSD(T, E4T), where T indicates the inclusion of non-iterative triple excitations, and E4T requests the inclusion of triple excitations for the MP4 component of the calculation too.

As for MP calculations, the full set of frozen core options are available.

4.5 Calculation type tab

In **HF**, **DFT**, **MP**, and **CC** stages, the following calculation types are available for all computational methods under the **Calculation type** tab.

Single point energy

This stage calculates the total energy of the system using the selected level of theory. Hartree-Fock and density functional theory calculations also yield the dipole and quadrupole moments.

Geometry optimization

A geometry optimization stage attempts to use the selected level of theory to find an atomic structure at a local energy minimum.

The 'Eigenvalue following' algorithm, which is the only option for the MP4(SDTQ), MP5, CCSD(T), and CCSD(T,E4T) methods, is subject to the following conditions:

- 250,000 or fewer atoms/centers
- 50 or fewer optimized variables

All parameters other than 'Algorithm' or 'Convergence' should be given as a positive integer (if not "Automatic").

DOCUMENTATION

Frequencies

Frequencies calculations enable the determination of many different properties.

Before performing a frequencies calculation, it is important to perform a geometry optimization first, using the same computational method with identical convergence criteria and basis set, as they assume the geometry of the system is in a local energy minimum with respect to the same level of theory. Frequencies calculated using different computational parameters to those used to determine the minimum energy geometry may be meaningless!

Infrared (IR) spectra are produced from the frequencies found in the system, and when frequencies are computed analytically, which may not be possible for Moller-Plesset and coupled cluster methods. The polarizability is also determined. Raman spectra can also be requested by selecting the corresponding option. In all cases, a thermochemical analysis is performed, using the temperature, pressure, and frequency scaling parameters in the associated section.

Selecting normal and anharmonic modes

Normal modes and anharmonic modes can be selected using Gaussian syntax (if not using 'Automatic' or 'All' from the respective menus). Some examples are:

```
2-5
```

Includes modes 2 through 5.

```
atoms=O
```

Includes modes involving oxygen atoms.

```
1-20 atoms=Fe
```

Includes modes 1 through 20 and any modes involving iron atoms.

```
notatoms=H
```

Excludes hydrogen atoms from any normal modes.

Raman optical activity incident light

Incident light can be specified using a space-separated list of floating point numbers, the units of which are set by the following option - one of either wavenumber (1/cm), wavelength (nm), or energy (Hartree).

Force constant matrix diagonalization

This option determines the lowest frequencies in the system, which can be a useful indicator of the reliability of a calculation. If the low lying modes are significantly non-zero compared to the lowest real frequencies, this suggests that the calculation is unstable or not well converged. Translations and rotations may be important modes for this system, a better integration grid may be needed, or the geometry may not be a minimum in energy.

Additional parameters

This free-form input section is placed at the end of the Gaussian input file and allows very fine-grained control over the parameters of the anharmonic vibrational analysis. Refer to the Gaussian documentation for the correct syntax.

Polarizability

Polarizability calculations provide dipole moment vectors and polarizability tensors. For methods other than MP4(SDTQ), MP5, CCSD(T), and CCSD(T, E4T), the quadrupole moment is also produced.

The Hartree-Fock and density functional theory methods produce hyperpolarizabilities by default, and also permit the calculation of frequency-dependent properties, optical rotations, and magnetic susceptibilities.

Frequencies can be specified using a space-separated list of floating-point numbers, the units of which are set by the following option - one of either wavenumber (1/cm), wavelength (nm), or energy (Hartree).

4.6 SCF visualization tab

The SCF visualization tab enables the creation of volumetric (.cube) files that can be viewed using the analysis menu in *MedeA*. This is available for all calculation types, but post-Hartree-Fock methods (Moller-Plesset and coupled cluster), SCF visualization corresponds to the Hartree-Fock orbitals only.

The display grid determines the density of three-dimensional points, where a 'Fine' grid produces a more accurate representation of the selected properties at the expense of much larger file size.

Individual or collections of orbitals can be selected using the 'Select orbitals' input. Here, the keywords 'HOMO' and 'LUMO' select the highest occupied molecular orbital and the lowest unoccupied molecular orbital, respectively, 'all' selects all the orbitals, 'occA' and 'occB' select all spin-alpha and spin-beta orbitals, respectively, 'valence' selects all valence orbitals and 'virtuals' selects all virtual orbitals.

5 Custom Stage

This stage runs a custom Gaussian calculation. If you have no experience running Gaussian outside of *MedeA*, then it is recommended that you read the Gaussian user documentation before attempting to use this stage. The contents of this help dialog give some simple examples of using variables set by *MedeA*.

Link 0 commands

Every type of Gaussian stage (including this one) created by *MedeA* is preceded by the Link 0 commands

```
%Chk=gaussian

%NProcShared=$N

%Mem=$M
```

where *$N* and *$M* is set by the JobServer and the 'Set computational resources' stage respectively. This allows information to be passed between consecutive jobs via the checkpoint file gaussian.chk. Subsequent Link 0 commands can be specified in the first input area of this dialog.

Route section

This part of the dialog lets you specify the desired level of theory, the calculation type, and other calculation keywords. Variables available to use in this section are:

```
%BASIS%
```

This variable is replaced by basis selected in the last 'configure basis' stage.

Title

Here, you can enter a title for this stage which will be printed in the output.

Input

Here, you can enter the remainder of the Gaussian input file. The overall structure of the file is laid out like:

```
%link0command1

%link0command2

%link0command3

# Route section
```

DOCUMENTATION

```
[charge] [spin multiplicity]

$input
```

where *$input* represents the contents of the input section of this dialog and [charge] and [spin multiplicity] are set from the properties of the active system in *MedeA* or from the stage 'Configure basis and pseudopotentials', under 'Charge/Spin multiplicity' if it exists. Variables available to use in this section are:

```
%GEOMETRY%
```

This variable is replaced by the initial geometry of the system, as it is on entry to the Gaussian flowchart. If a Gaussian stage alters the geometry of the system, the contents of this variable will be unaffected. It is therefore advisable to pass the system geometry via the checkpoint file for stages other than the first using the Geom=Checkpoint keyword in the Route section.

Examples

1-A first calculation custom stage where a frequencies calculation using a subset of normal modes is done:

Route section

```
# HF/6-31G Freq=SelectNormalModes
```

Input

```
%GEOMETRY%

3-9
```

2-A subsequent calculation custom stage performing a frequencies calculation using a subset of normal modes, which follows a geometry optimization stage:

Route section

```
# HF/%BASIS% Geom=Checkpoint Freq=SelectNormalModes
```

Input

```
3-9
```

DOCUMENTATION