Potential energy surfaces (PES)

Prepared by Radek Marek group

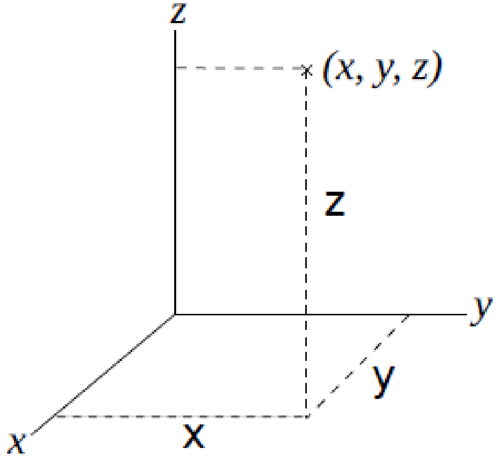
Potential energy surface

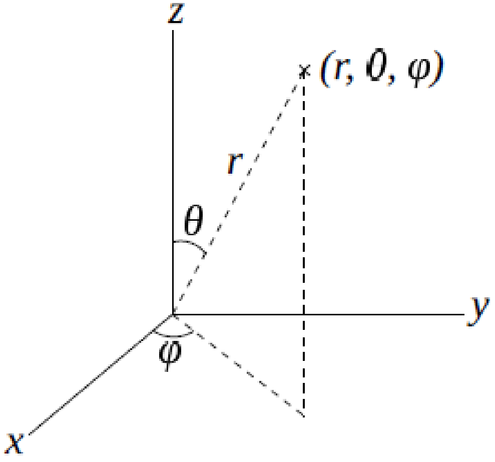
The Potential Energy Surface (PES) is a mathematical function that gives the internal energy of a molecule as a function of geometrical modulations as it stretches, bends, torsions, breaks, etc. The PES of a molecule is obtained by the total interactions of:

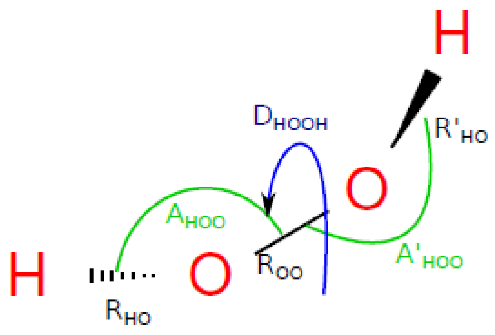
- nuclear-nuclear repulsion
- electron- electron interaction
- electron-nuclear attraction

Geometry description

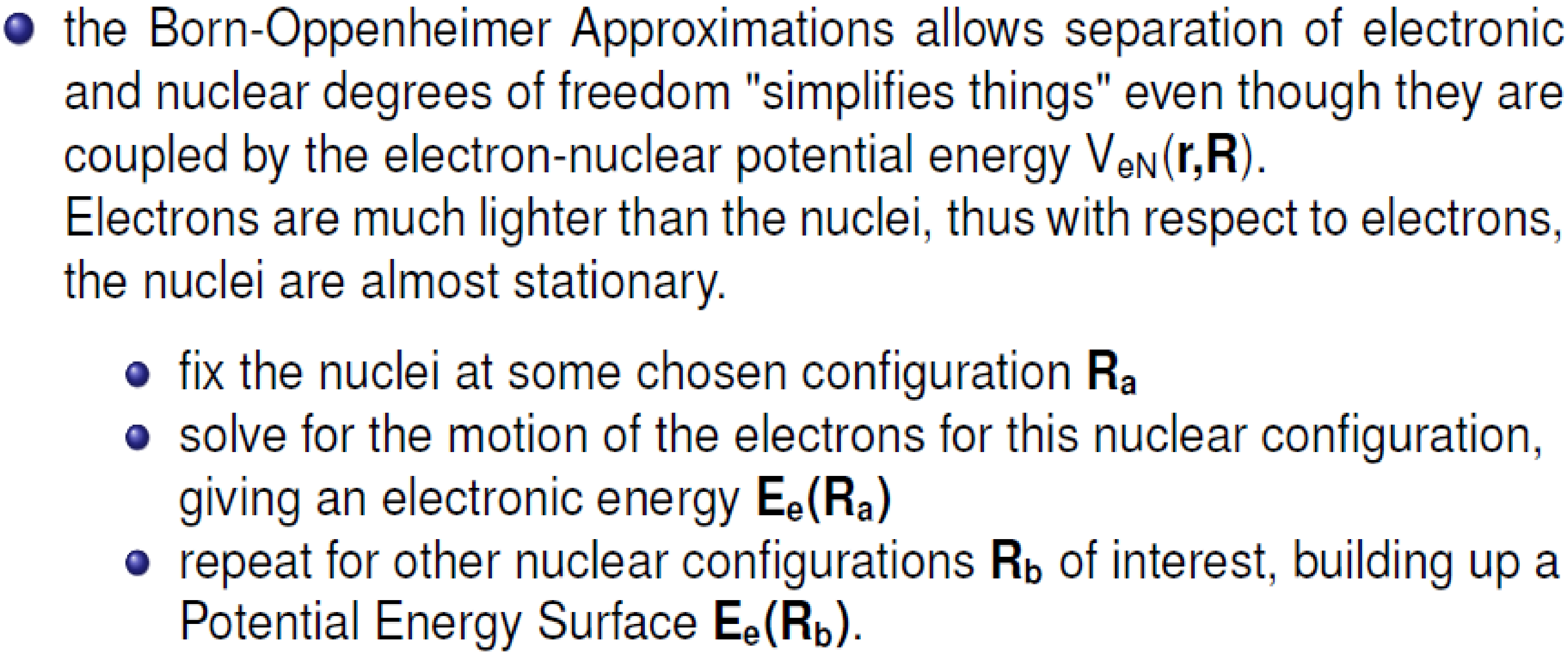
Coordinates	Notation	Degrees of freedom
Cartesian	x, y, z	3N
Spherical	r, θ, ϕ	3N
Internal	R, A, D	3N-6 (3N-5)





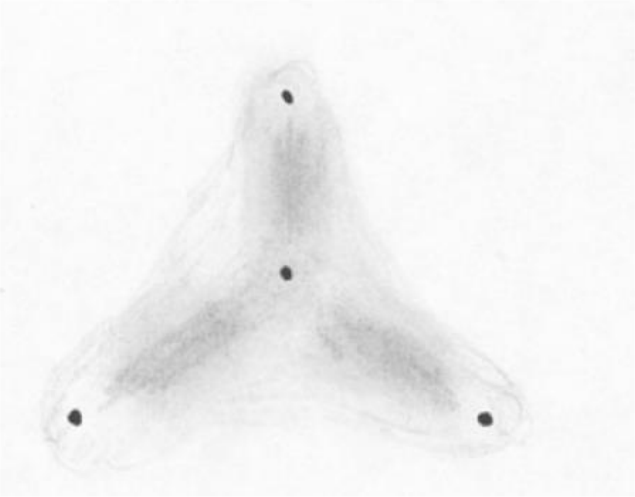


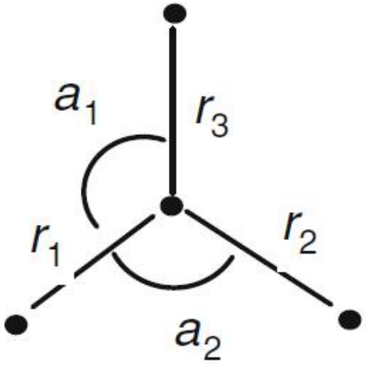
The Born-Oppenheimer Approximation

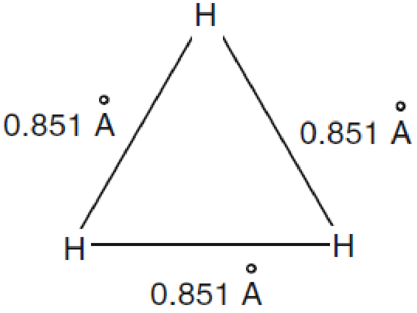


- the Born-Oppenheimer Approximations allows separation of electronic and nuclear degrees of freedom "simplifies things" even though they are coupled by the electron-nuclear potential energy $V_{eN}(\mathbf{r},\mathbf{R})$. Electrons are much lighter than the nuclei, thus with respect to electrons, the nuclei are almost stationary.
  - fix the nuclei at some chosen configuration $\mathbf{R_a}$
  - solve for the motion of the electrons for this nuclear configuration, giving an electronic energy $\mathbf{E_e(R_a)}$
  - repeat for other nuclear configurations $\mathbf{R_b}$ of interest, building up a Potential Energy Surface $\mathbf{E_e(R_b)}$.

The Born-Oppenheimer Approximation



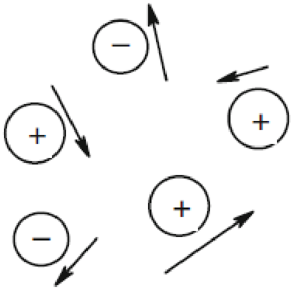




The H_3^+ cation: 3 protons, 2 electrons

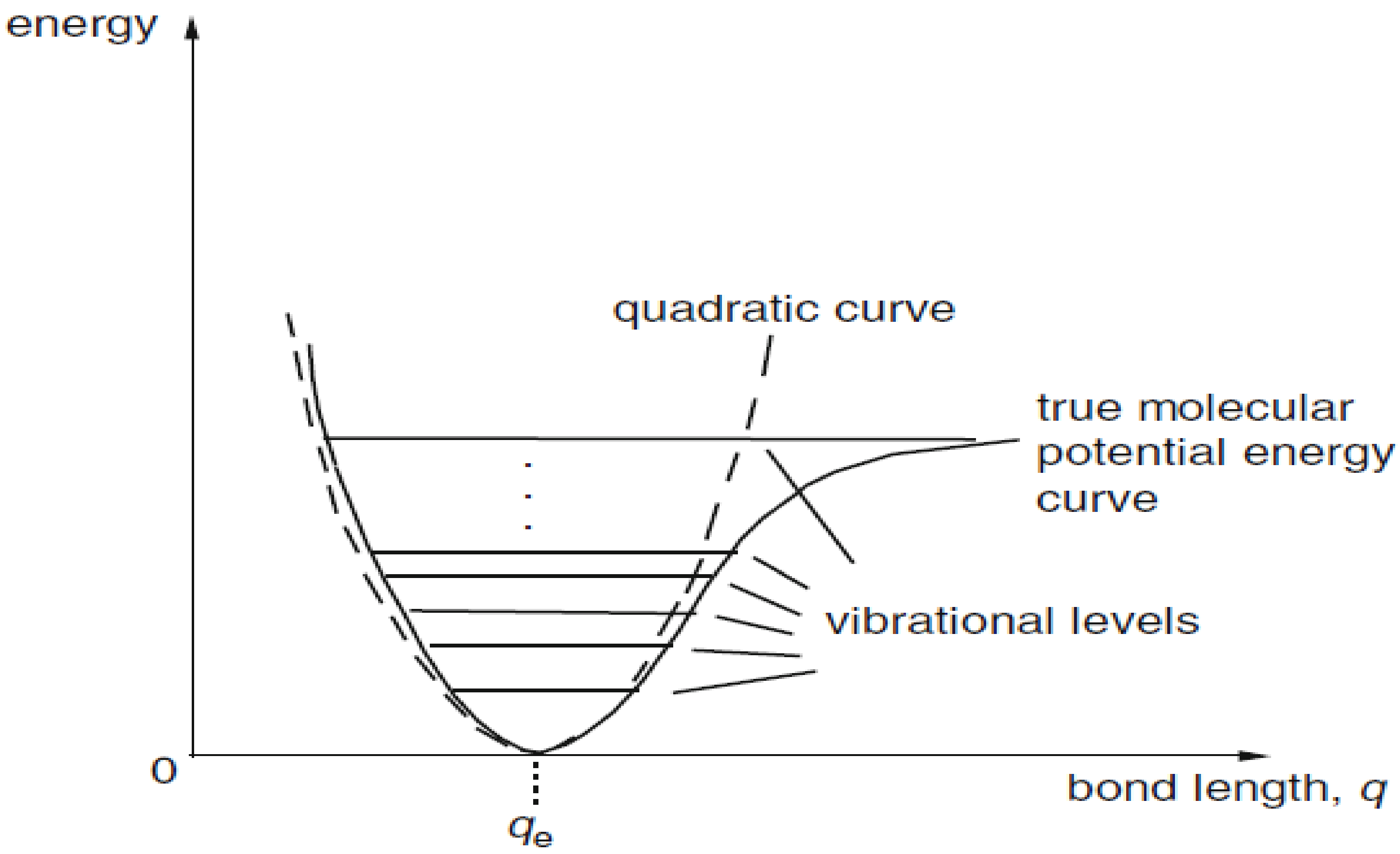
Definite geometry

make the masses of the nuclei and electrons equal →



No definite geometry

PES: a diatomic molecule



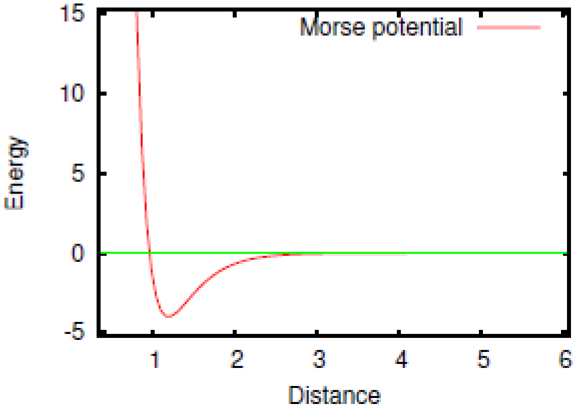
Morse Potential

- Morse potential:

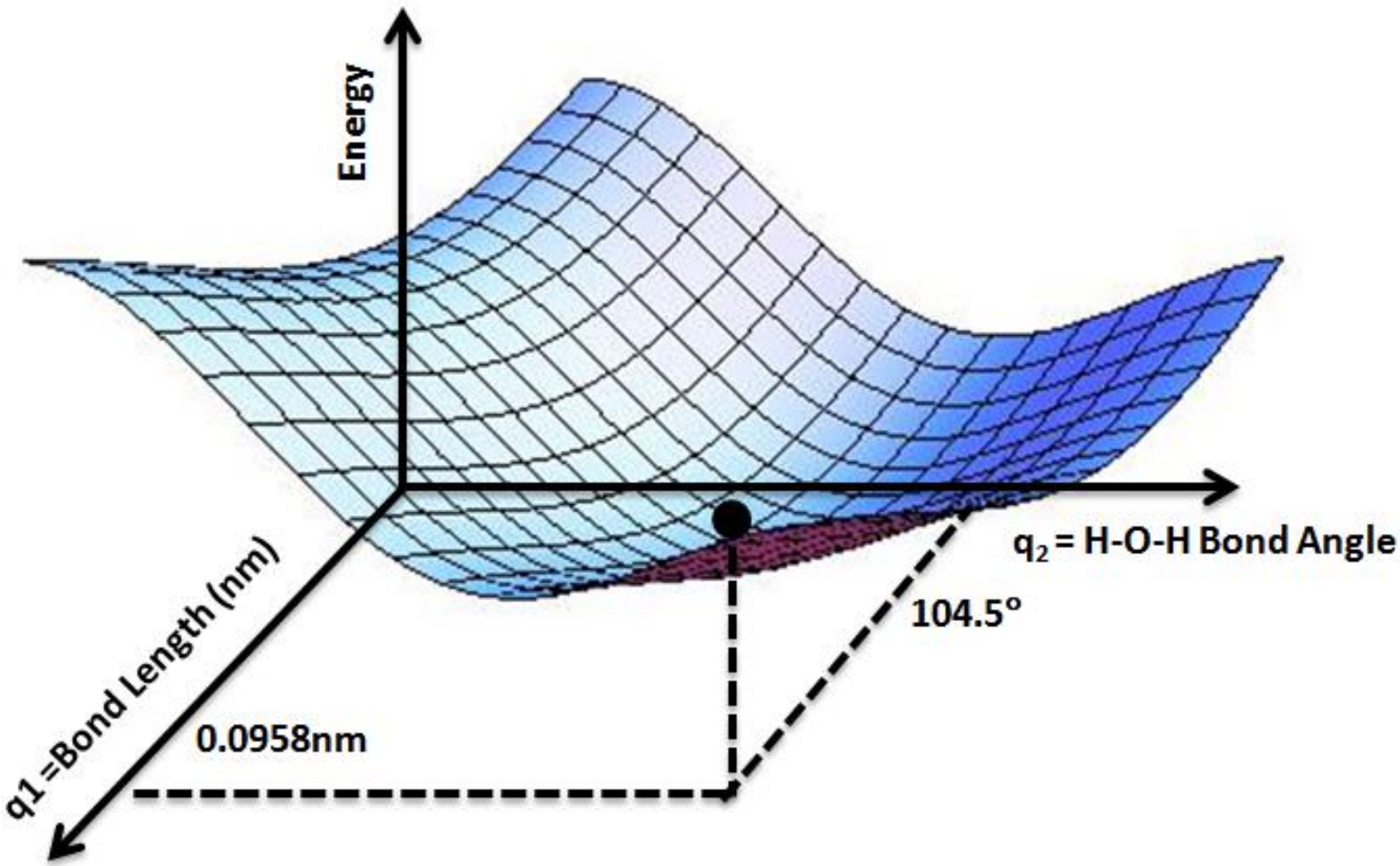
$$V(r) = D_e(1 - e^{-a(r-r_e)})^2 \tag{1}$$

where

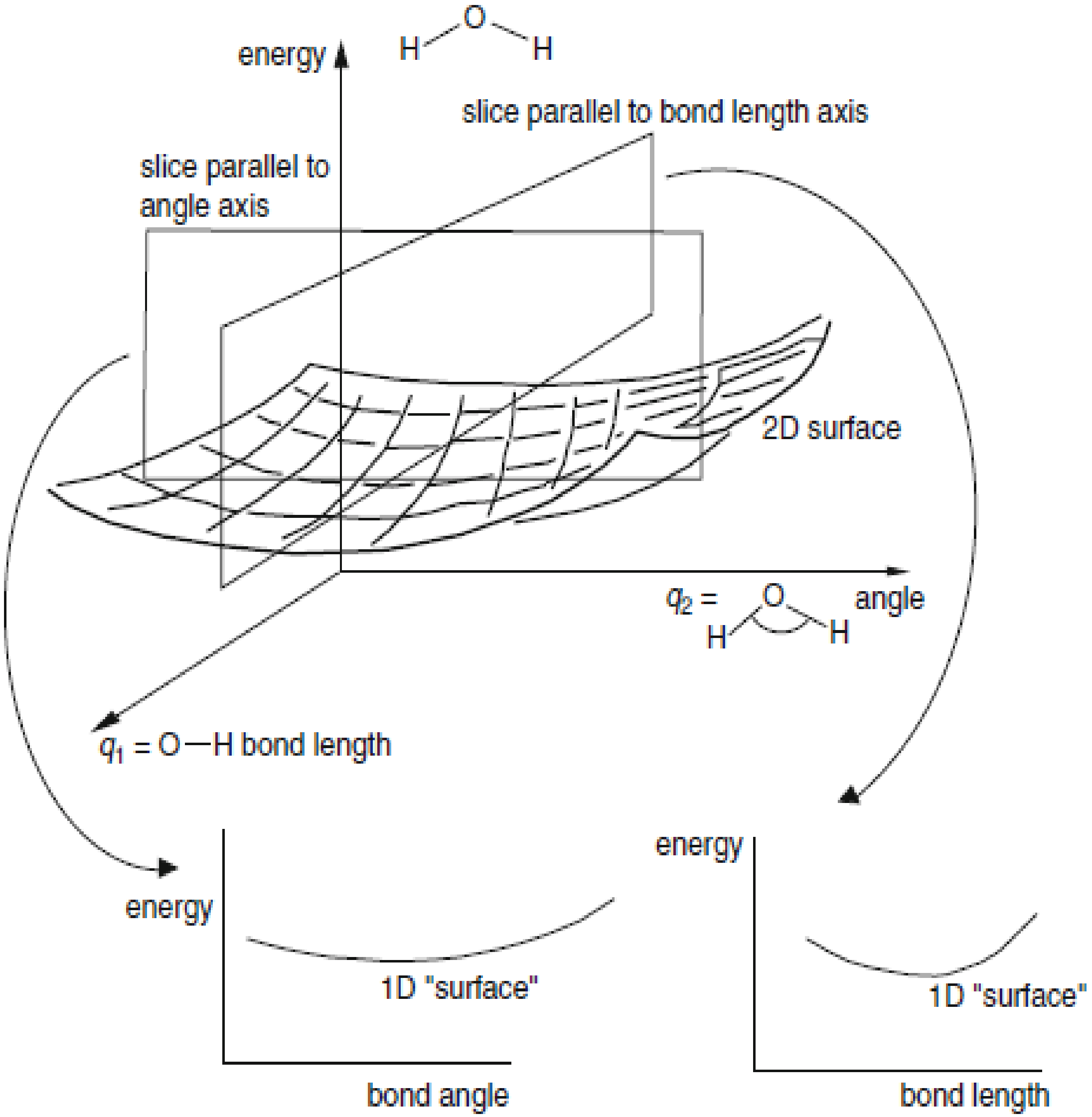
- D_e, is the depth of potential well
- a, controls the width of the potential
- r, is the internuclear separation
- r_e, is the equilibrium distance



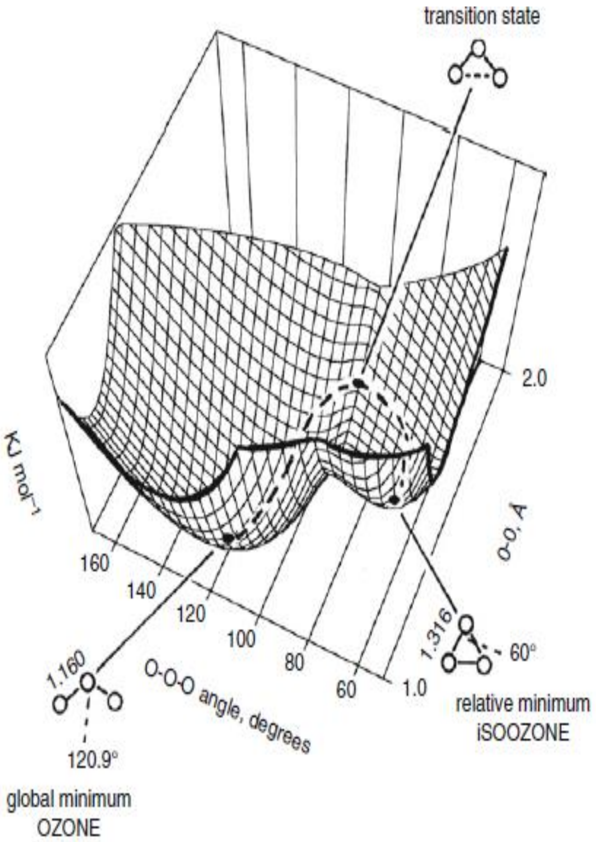
3D PES for a water molecule

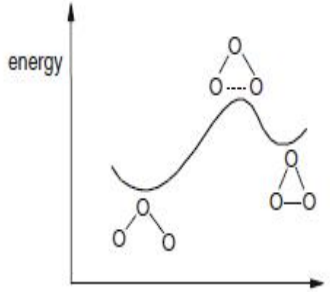


PES slices

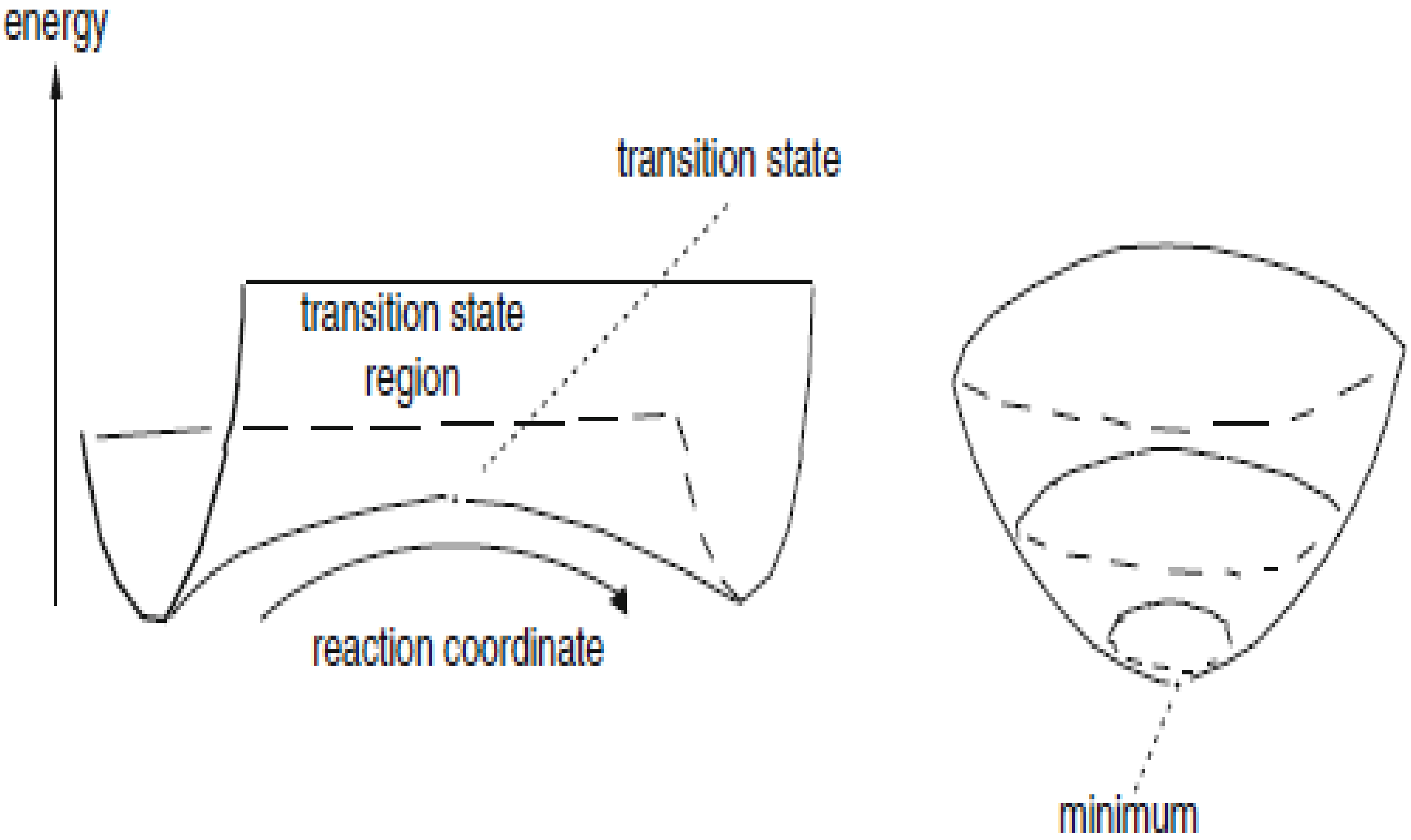

energy
H
O
H
slice parallel to bond length axis
slice parallel to
angle axis
2D surface
q_2 =
O
H
H
angle
q_1 = O—H bond length
energy
1D "surface"
bond angle
energy
1D "surface"
bond length

Intrinsic Reaction Path

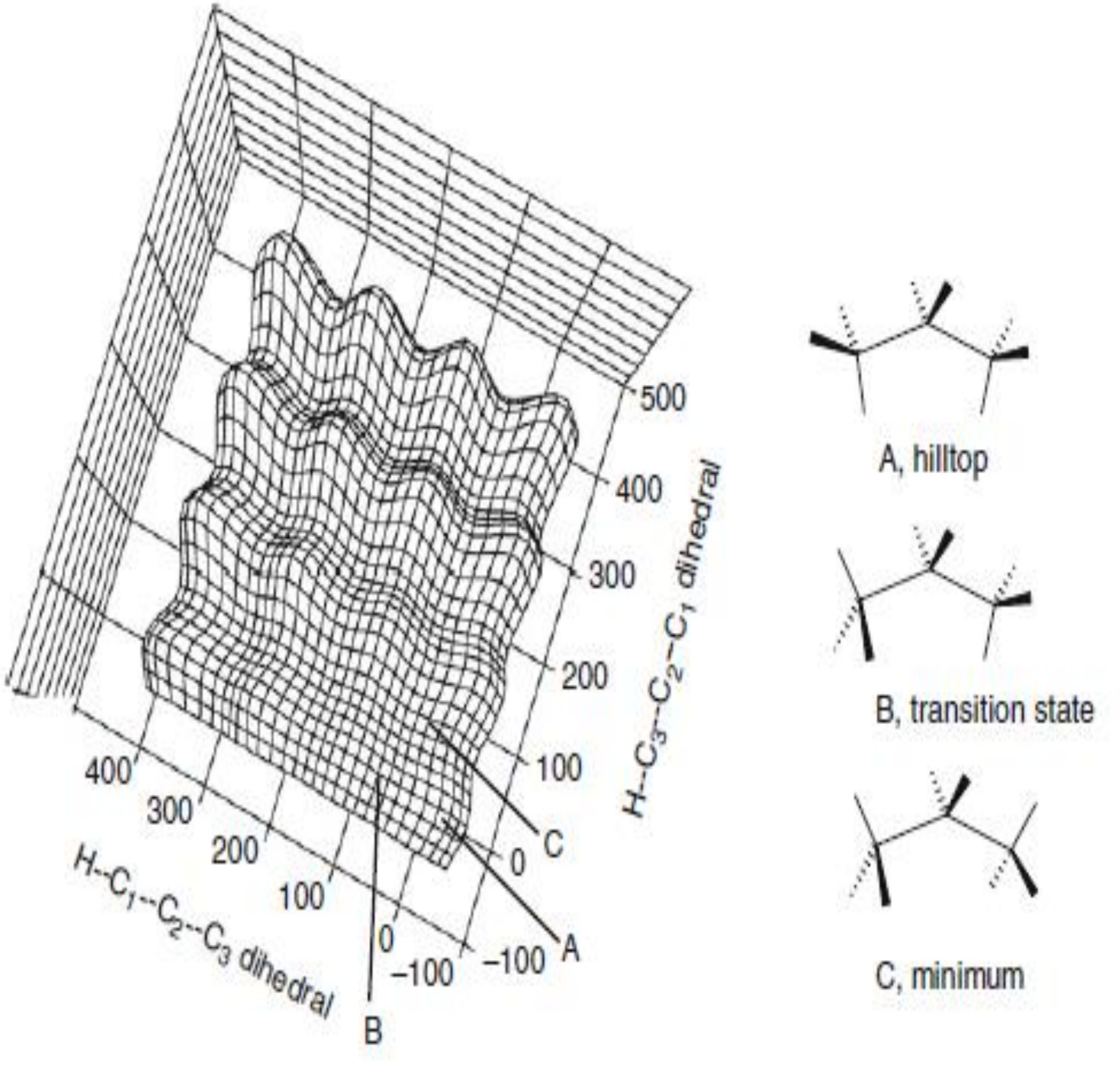




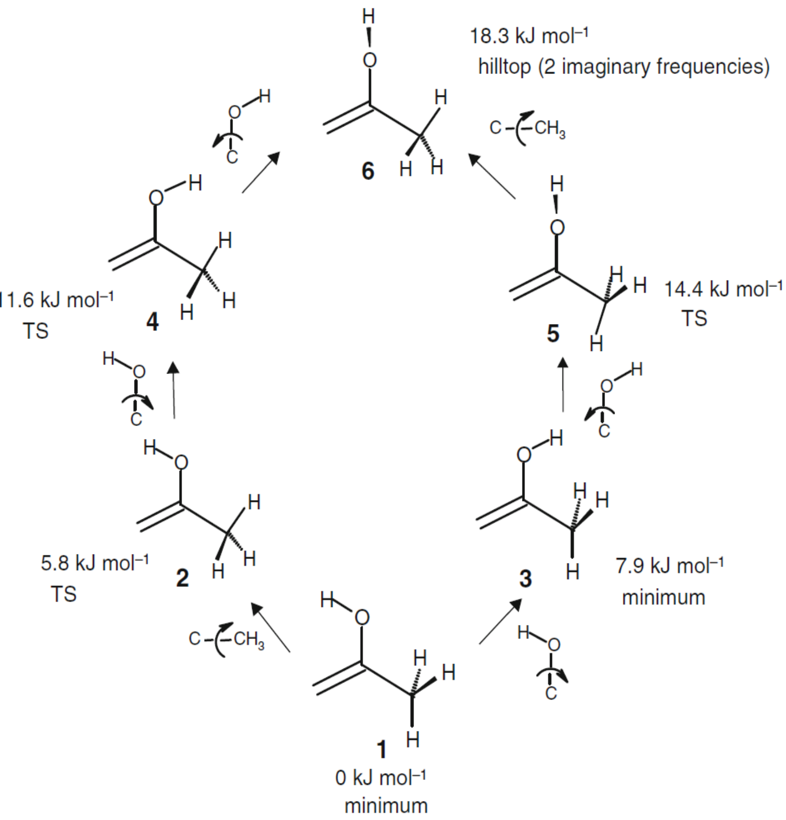
Minima and transition states

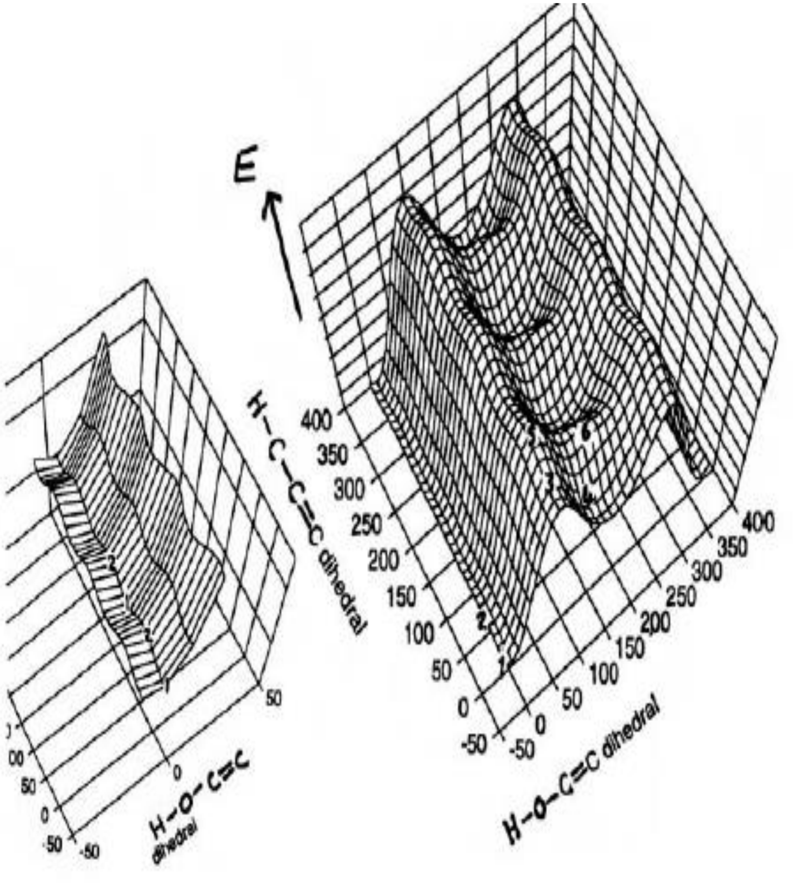


Saddle points and hilltops

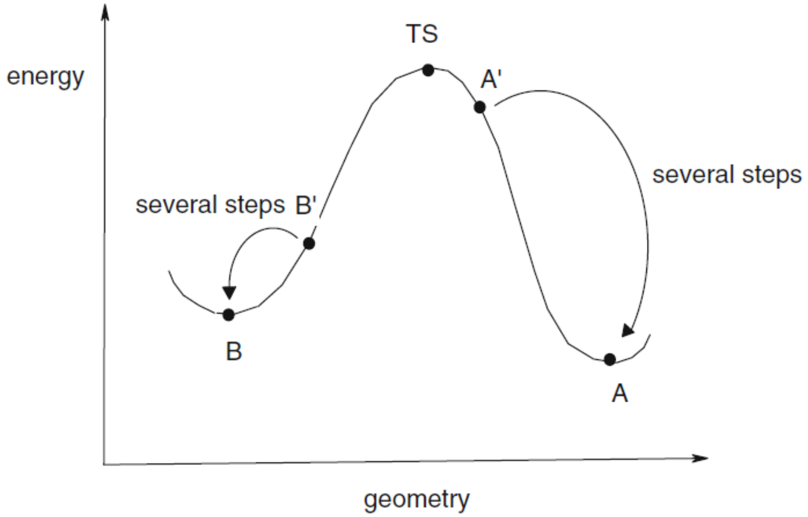


Conformers and transition states





Important points on PES



- stationary points:

$$\frac{\partial E}{\partial q_i} = 0$$

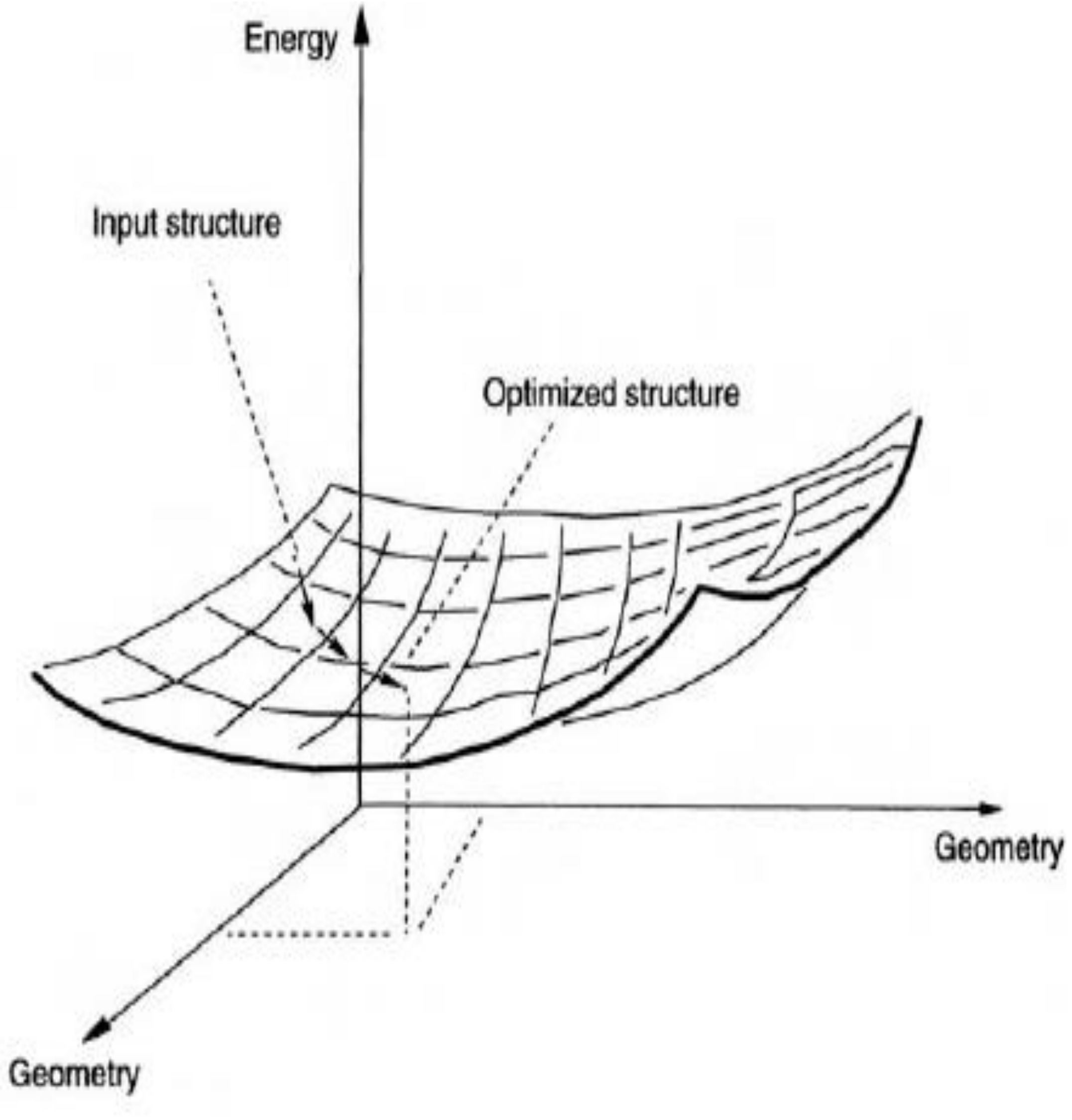
- local minimum:

$$\frac{\partial^2 E}{\partial q_i^2} > 0 \text{ for all degrees of freedom}$$

- n^{th} order saddle point:

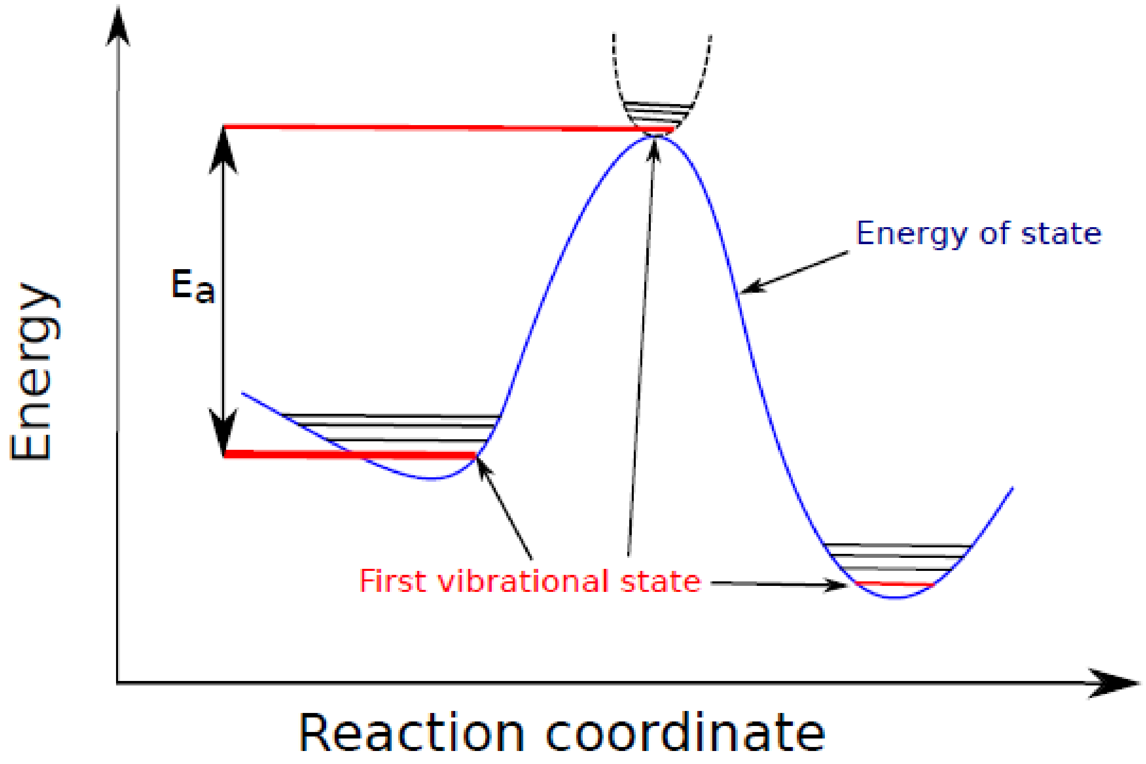
$$\frac{\partial^2 E}{\partial q_i^2} < 0 \text{ for n degrees of freedom}$$

Geometry optimization



Zero-Point Energy (ZPE)

- Vibrational corrections for 0K (ground vibrational state)
- “Cancel out” for energy differences

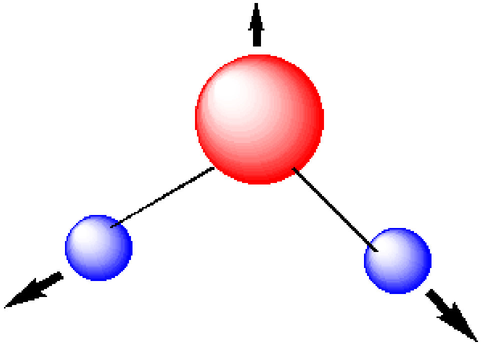


Frequencies, Hessian Index

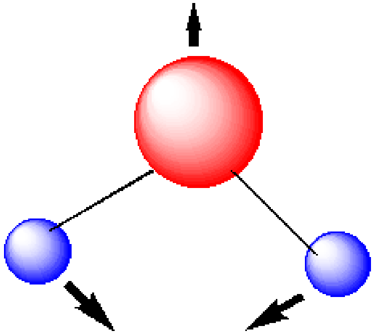
- a calculated optimized geometry needs a vibrational analysis to verify its location in the PES via the Hessian index
- the Hessian index is the number of negative eigenvalues of the force constant matrix (i.e. imaginary frequencies). For a stationary point, this corresponds to the number of internal degrees of freedom along which that point is a potential energy maximum. The Hessian index is:
  - 0 for minima
  - 1 for transition states
  - > 1 for higher-order saddle points

$$\tilde{v} = \frac{1}{2\pi c}\left(\frac{k}{\mu}\right)^{1/2}$$

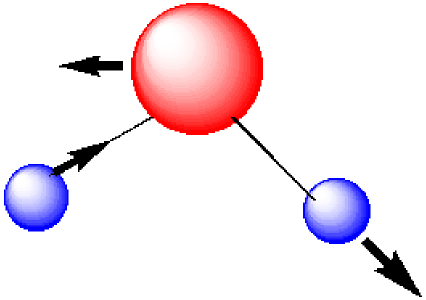
Vibrational modes of water



Symmetric Stretch
3657 cm^{-1}



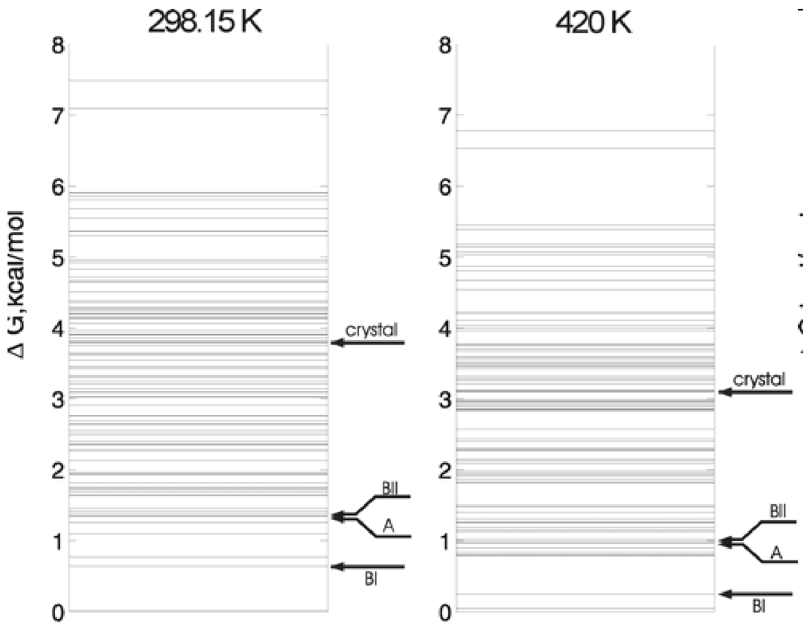
Bend 1595 cm^{-1}



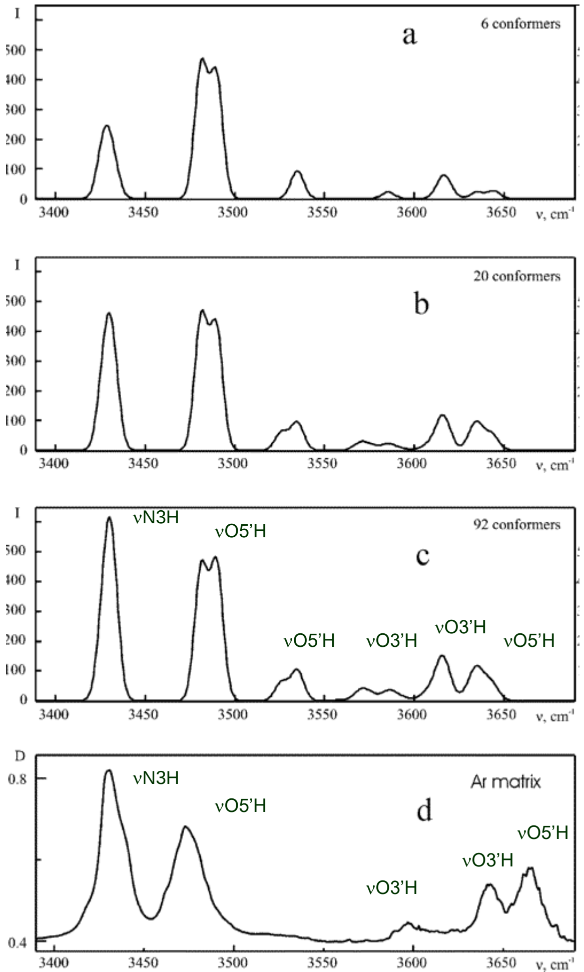
Asymmetric Stretch
3756 cm^{-1}

Ab initio reconstruction of thymidine infrared spectrum

Energetic diagrams of thymidine conformers at 298.15 K and 420 K with indication of DNA-like and crystal-like conformations



Computational experiment allows predicting IR spectra at any temperature including physiological temperatures

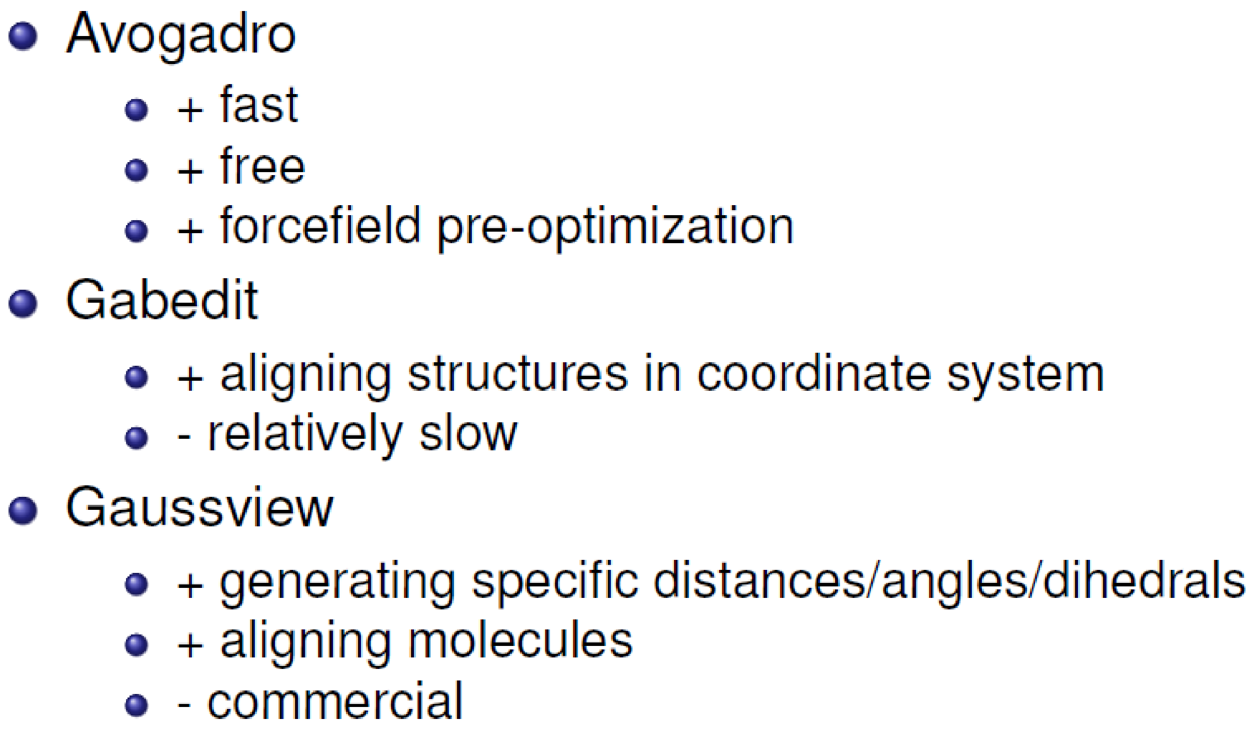


IR calculated spectra of 6, 20 and 92 conformers (MP2/aug-cc-PVTZ// B3LYP/6-31++G(d,p)) and experimental low-temperature matrix spectrum

IR reconstructed spectrum of all 92 conformers MP2/aug-cc-pVTZ/B3LYP/6-31++G(d,p) level of theory

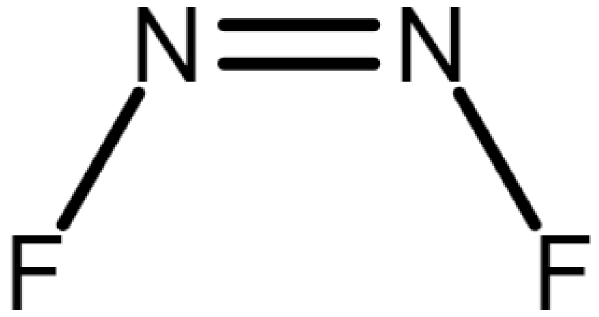
Experimental low-temperature matrix IR spectrum
*A.Yu. Ivanov, G. Sheina, S.A. Krasnokutski, Low. Temp. Phys. 29 (2003), 809

Examples of molecular builders

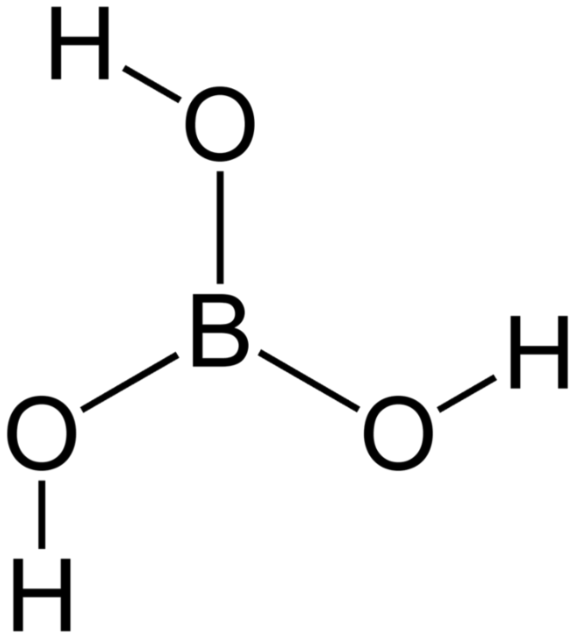


- Avogadro
  - + fast
  - + free
  - + forcefield pre-optimization
- Gabedit
  - + aligning structures in coordinate system
  - - relatively slow
- Gaussview
  - + generating specific distances/angles/dihedrals
  - + aligning molecules
  - - commercial

Building molecules (examples)

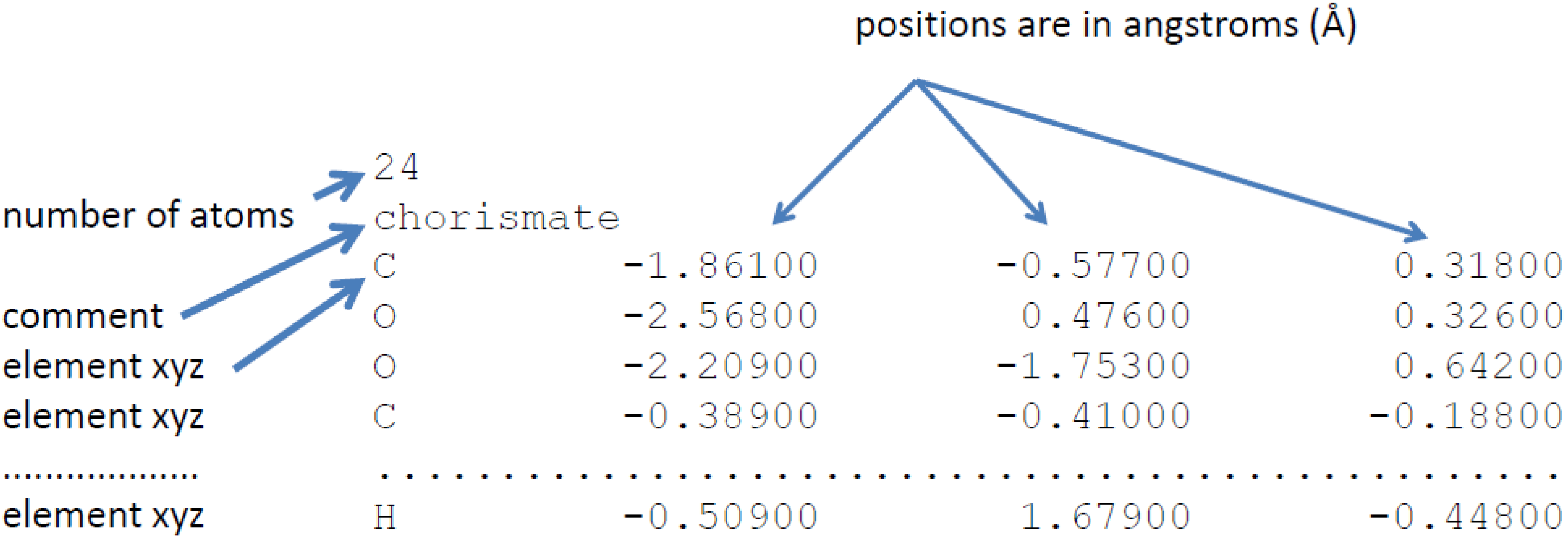


Dinitrogen diflouride



Boric acid

XYZ format



positions are in angstroms (Å)

number of atoms

comment

element xyz

element xyz

....

element xyz

```
24
chorismate
C          -1.86100        -0.57700         0.31800
O          -2.56800         0.47600         0.32600
O          -2.20900        -1.75300         0.64200
C          -0.38900        -0.41000        -0.18800
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
H          -0.50900         1.67900        -0.44800
```

The **xyz** format is a free-formatting text file (values in columns can be separated by any number of spaces or other whitespace).

The format only describes the geometry of the system. It does not contain information about bonds in the system. A program that works with the format must calculate this information (e.g., using atomic radii).

Z-matrix format

Cartesian coordinates

```
O -0.180077 -0.046023 -0.062789
H  0.196208 -0.747659  0.498793
O  0.006537  1.047922  0.877207
H -0.931885  1.299156  0.951390
```

x y z

Number of degrees of freedom:

3N

Internal coordinates (Z-matrix)

```
O
H 1 0.974298
O 1 1.454349 2 96.868054
H 3 0.974298 1 96.868054 2 239.552651
```

bond length bond angle torsion angle

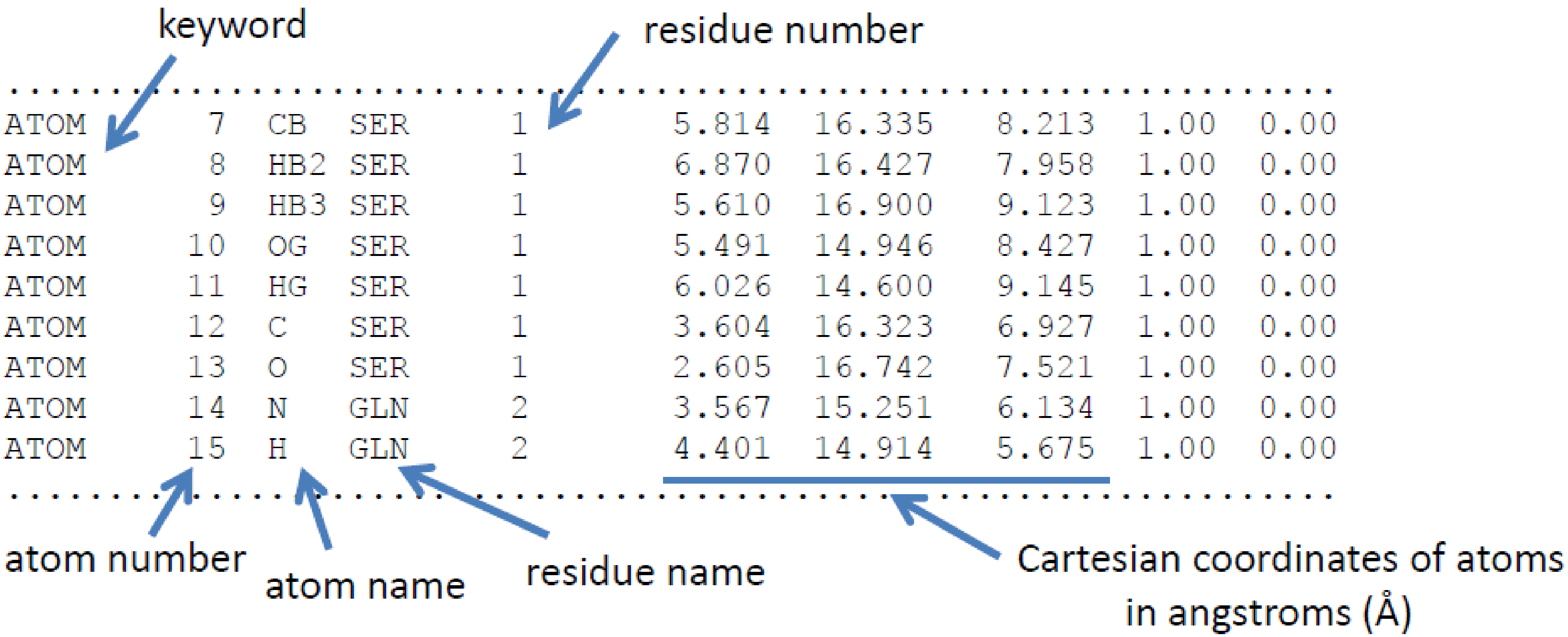
Number of degrees of freedom:

3N-6

3N-5 (linear diatomic molecule)

PDB format

The **pdb** format is employed to store the structures of biomolecules and their complexes. **It is widely used but it has several limitations.** Therefore, it is slowly substituted by more advanced formats such as PDBx/mmCIF and others.



keyword

residue number

```
.......................................................................
ATOM      7  CB   SER     1       5.814  16.335   8.213  1.00   0.00
ATOM      8  HB2  SER     1       6.870  16.427   7.958  1.00   0.00
ATOM      9  HB3  SER     1       5.610  16.900   9.123  1.00   0.00
ATOM     10  OG   SER     1       5.491  14.946   8.427  1.00   0.00
ATOM     11  HG   SER     1       6.026  14.600   9.145  1.00   0.00
ATOM     12  C    SER     1       3.604  16.323   6.927  1.00   0.00
ATOM     13  O    SER     1       2.605  16.742   7.521  1.00   0.00
ATOM     14  N    GLN     2       3.567  15.251   6.134  1.00   0.00
ATOM     15  H    GLN     2       4.401  14.914   5.675  1.00   0.00
.......................................................................
```

atom number

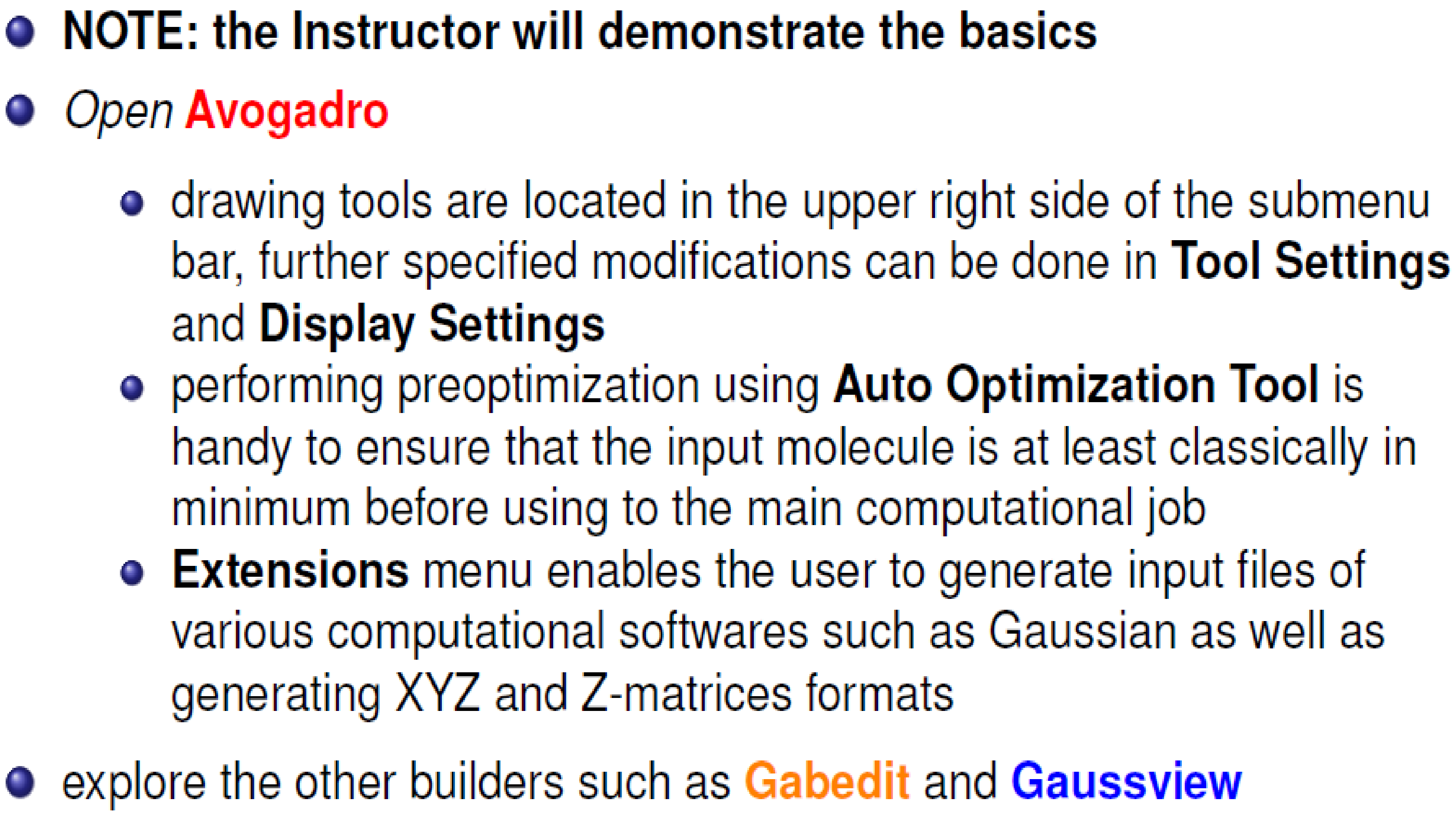
atom name

residue name

Cartesian coordinates of atoms in angstroms (Å)

The pdb format does not usually contain information about bods in the system. The program that works with the format must calculate this information (based on template structures). For non-standard residues, the CONECT keyword can be used.

Practice 1: exploring molecular builders



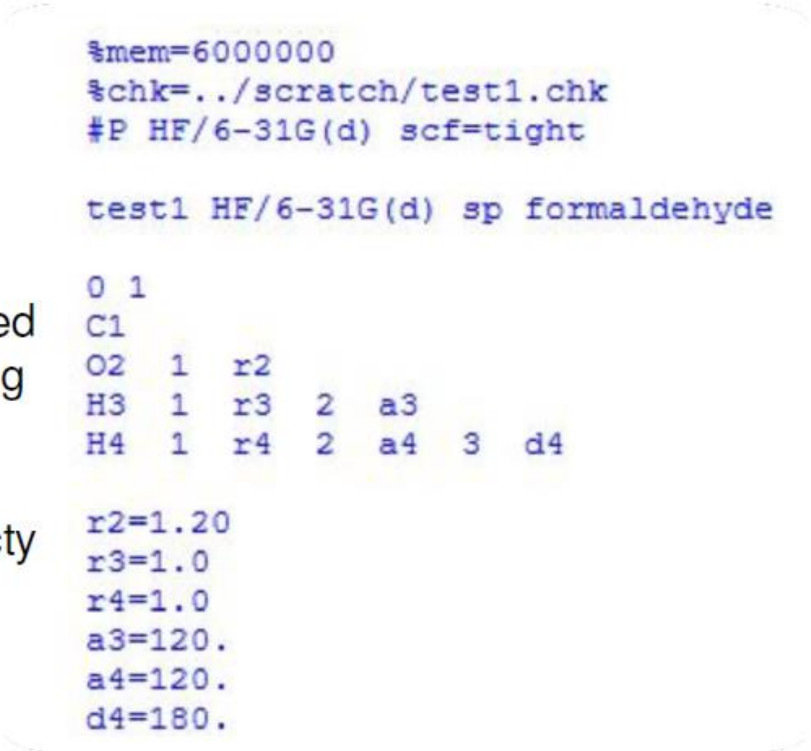
- **NOTE: the Instructor will demonstrate the basics**
- *Open* **Avogadro**
  - drawing tools are located in the upper right side of the submenu bar, further specified modifications can be done in **Tool Settings** and **Display Settings**
  - performing preoptimization using **Auto Optimization Tool** is handy to ensure that the input molecule is at least classically in minimum before using to the main computational job
  - **Extensions** menu enables the user to generate input files of various computational softwares such as Gaussian as well as generating XYZ and Z-matrices formats
- explore the other builders such as **Gabedit** and **Gaussview**

Practice 2: Gaussian calculations

- prepare a Gaussian input file of a **small molecule of your choice** for Single Point Energy and Geometry Optimization calculations:
  - builders can generate the file, it can be written from scratch using **Avogadro**, **Gabedit**, **Gaussview**
  - common suffix is ".com"
  - use, 2 cores and at least 3 GB of memory
- the main Gaussian manual is available at: `https://gaussian.com/man/`, the slight variations of of syntax from g09 to later versions are minimal, it is also specified in the website.

Practice 2: Gaussian calculations

> a sample input file for single point energy calculation of formaldehyde in Z-matrix coordinates
> **#** Route section which described the run type and basis set, adding **OPT** signifies geometry optimization
> **0 1** signify charge and multiplicty
> add Gaussian module then submit
> *g09* input.com

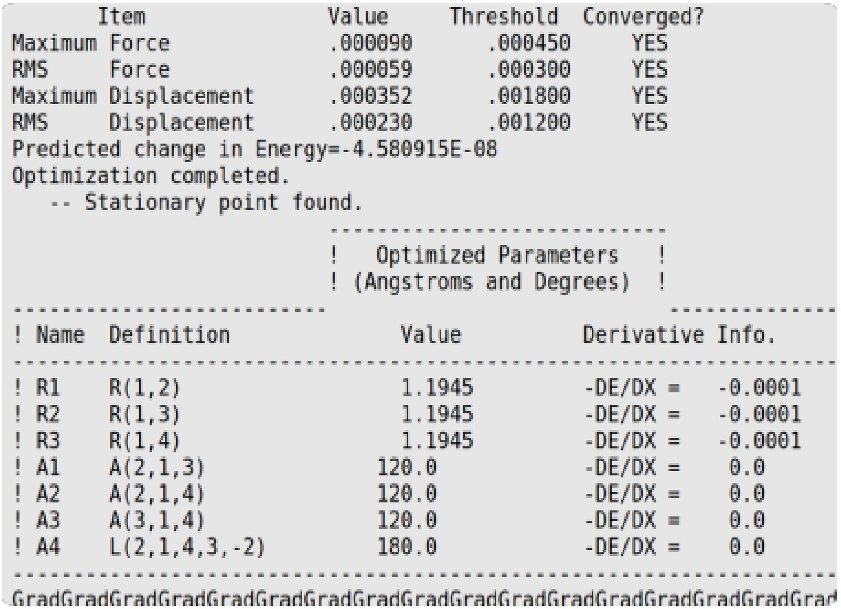


```
%mem=6000000
%chk=../scratch/test1.chk
#P HF/6-31G(d) scf=tight

test1 HF/6-31G(d) sp formaldehyde

0 1
C1
O2   1   r2
H3   1   r3   2   a3
H4   1   r4   2   a4   3   d4

r2=1.20
r3=1.0
r4=1.0
a3=120.
a4=120.
d4=180.
```

- complete details of the input file is available at: `https://gaussian.com/input/`

Gaussian output (log) file

> the logfile of a job in **Normal Termination** (No Error) of Gaussian contains information about calculations cycles, SCF convergence, Energetics, etc
> for an **OPT** job, go to the end of the file and move backwards until you find the final set of forces and displacements, it should look something like this.
NOTE: the Instructor will run through the output file



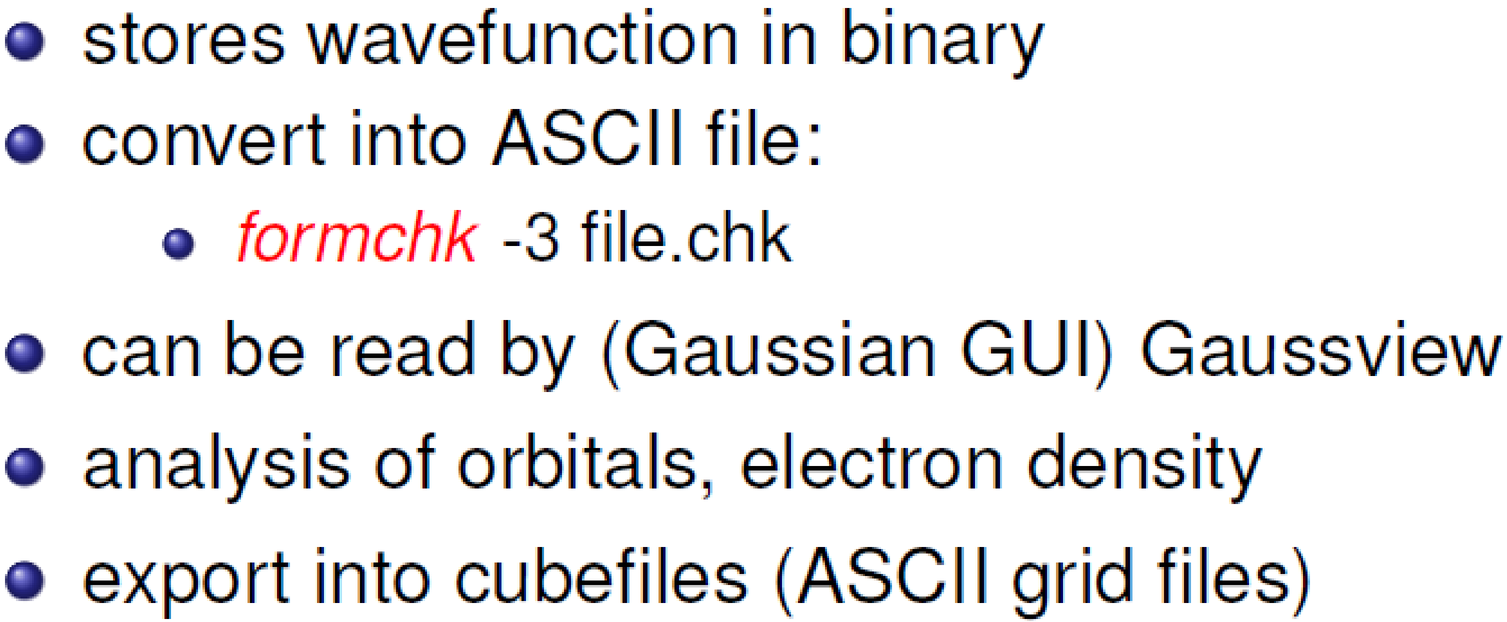
```
       Item               Value     Threshold  Converged?
Maximum Force            .000090     .000450     YES
RMS     Force            .000059     .000300     YES
Maximum Displacement     .000352     .001800     YES
RMS     Displacement     .000230     .001200     YES
Predicted change in Energy=-4.580915E-08
Optimization completed.
   -- Stationary point found.
                          ----------------------------
                          !   Optimized Parameters   !
                          ! (Angstroms and Degrees)  !
--------------------------                            -------------
! Name  Definition              Value          Derivative Info.
--------------------------------------------------------------------
! R1    R(1,2)                  1.1945         -DE/DX =   -0.0001
! R2    R(1,3)                  1.1945         -DE/DX =   -0.0001
! R3    R(1,4)                  1.1945         -DE/DX =   -0.0001
! A1    A(2,1,3)              120.0            -DE/DX =    0.0
! A2    A(2,1,4)              120.0            -DE/DX =    0.0
! A3    A(3,1,4)              120.0            -DE/DX =    0.0
! A4    L(2,1,4,3,-2)         180.0            -DE/DX =    0.0
--------------------------------------------------------------------
GradGradGradGradGradGradGradGradGradGradGradGradGradGradGradGradGrad
```

Gaussian output (log) file

- **Abnormal Termination** (with Errors) of Gaussian happens and troubleshooting can be tedious depending on the source of error. A list of common errors and how to correct them is available at this site: `https://docs.computecanada.ca/wiki/Gaussian_error_messages`
- *qmutil*: nifty module to extract data from gaussian output:
    - *extract-gopt-ene* logfile → Monitoring optimization process
    - *extract-gopt-xyz* logfile → Extract all geometries
    - *extract-gdrv-ene* logfile
    - *extract-gdrv-xyz* logfile
    - *extract-xyz-str* xyzfile framenumber
    - *extract-xyz-numstr* xyzfile

```
$ module add qmutil
```

Gaussian checkpoint file (.chk)

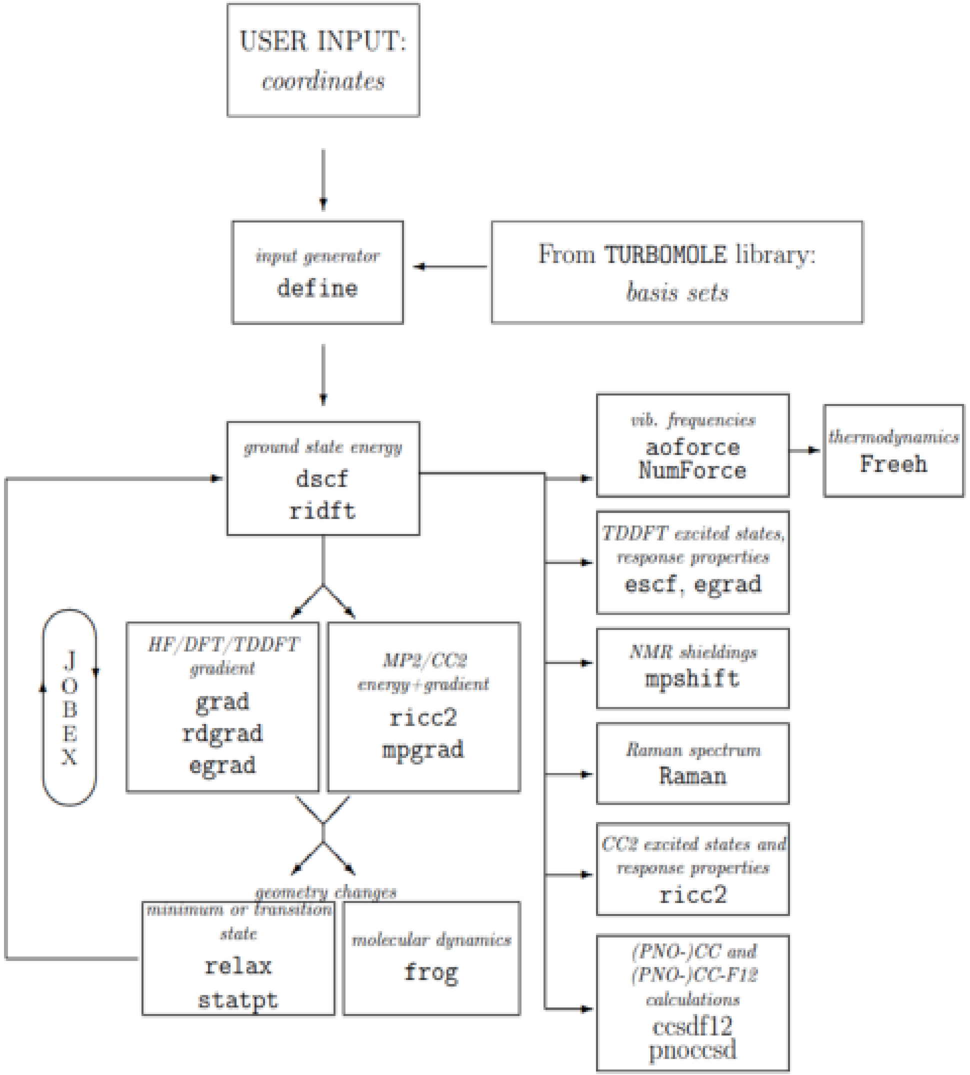


- stores wavefunction in binary
- convert into ASCII file:
  - *formchk* -3 file.chk
- can be read by (Gaussian GUI) Gaussview
- analysis of orbitals, electron density
- export into cubefiles (ASCII grid files)

Turbomole

- Turbomole is probably the “fastest" code available here (for Gaussian-type basis functions)
- Turbomole has interchangeable **Modules**, such that each contains everything necessary to execute only one aspect of the desired functionality.
- interactive preparation of the input file using *define*
- Turbomole manual: `https://www.turbomole.org/turbomole/turbomole-documentation/`

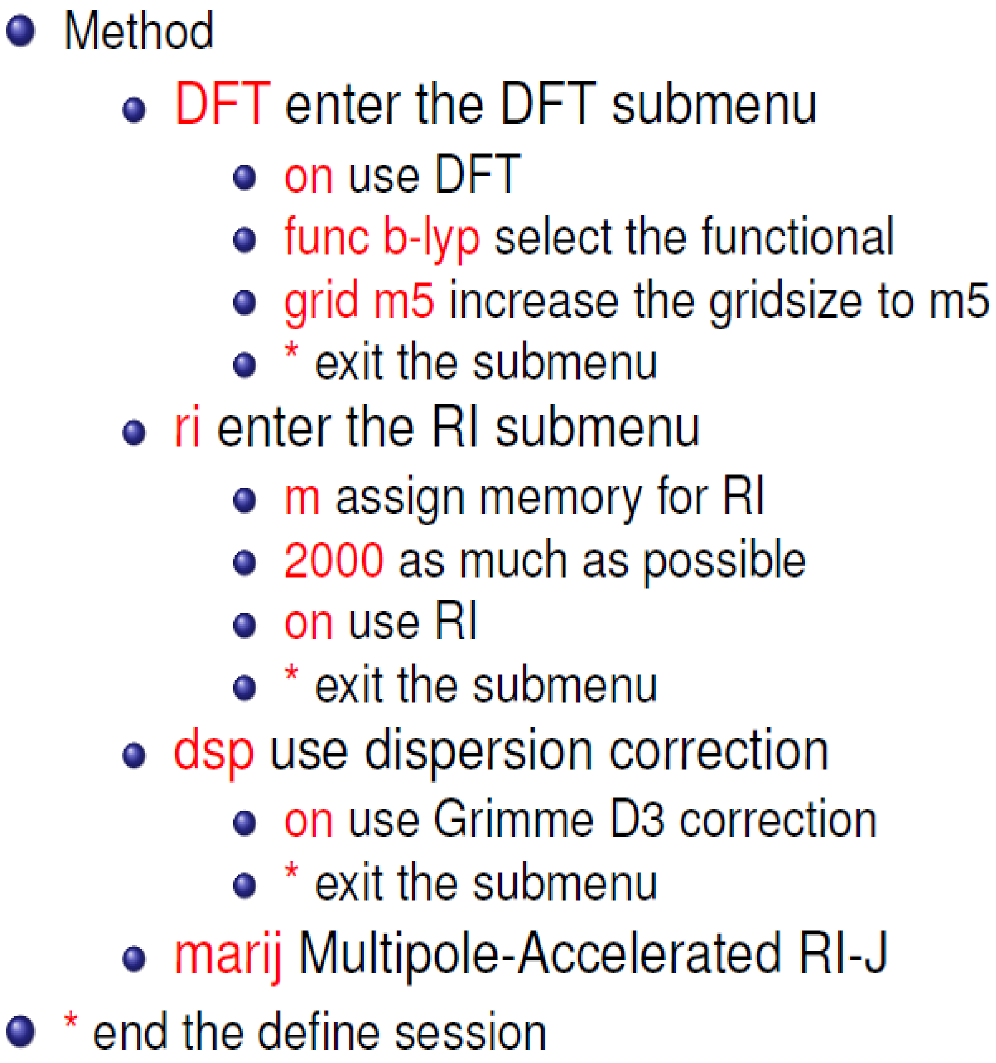
Turbomole Modules



Practice 3: Turbomole

- add Turbomole module, then call *define*
- first two items can be skipped
- molecular geometry:
  - a coord Reads in the geometry
  - ired Generates internal coordinates
  - * Proceed to next stage
- basis set(s):
  - b all def2-SVP Assign this basis set to all atoms
  - * Proceed to next stage
- method
  - eht Perform initial guess from Extended Hückel Theory
  - *Accept all defaults*

Practice 3: Turbomole



- Method
  - DFT enter the DFT submenu
    - on use DFT
    - func b-lyp select the functional
    - grid m5 increase the gridsize to m5
    - * exit the submenu
  - ri enter the RI submenu
    - m assign memory for RI
    - 2000 as much as possible
    - on use RI
    - * exit the submenu
  - dsp use dispersion correction
    - on use Grimme D3 correction
    - * exit the submenu
  - marij Multipole-Accelerated RI-J
- * end the define session

Practice 3: Turbomole

- for running Turbomole in parallel mode use the parallel build
- infinity selects it by default if ncpu > 1

user@wolf

```
#!/bin/bash
module add turbomole:7.02:x86_64:para
dscf -ri -c 1024 > jobname.out
```

- remember **SP = "dscf"**, **OPT = "jobex"**,...
- it is safe to designate the results as a ".out" file

Practice 3: Turbomole

- a successful Turbomole job creates several files

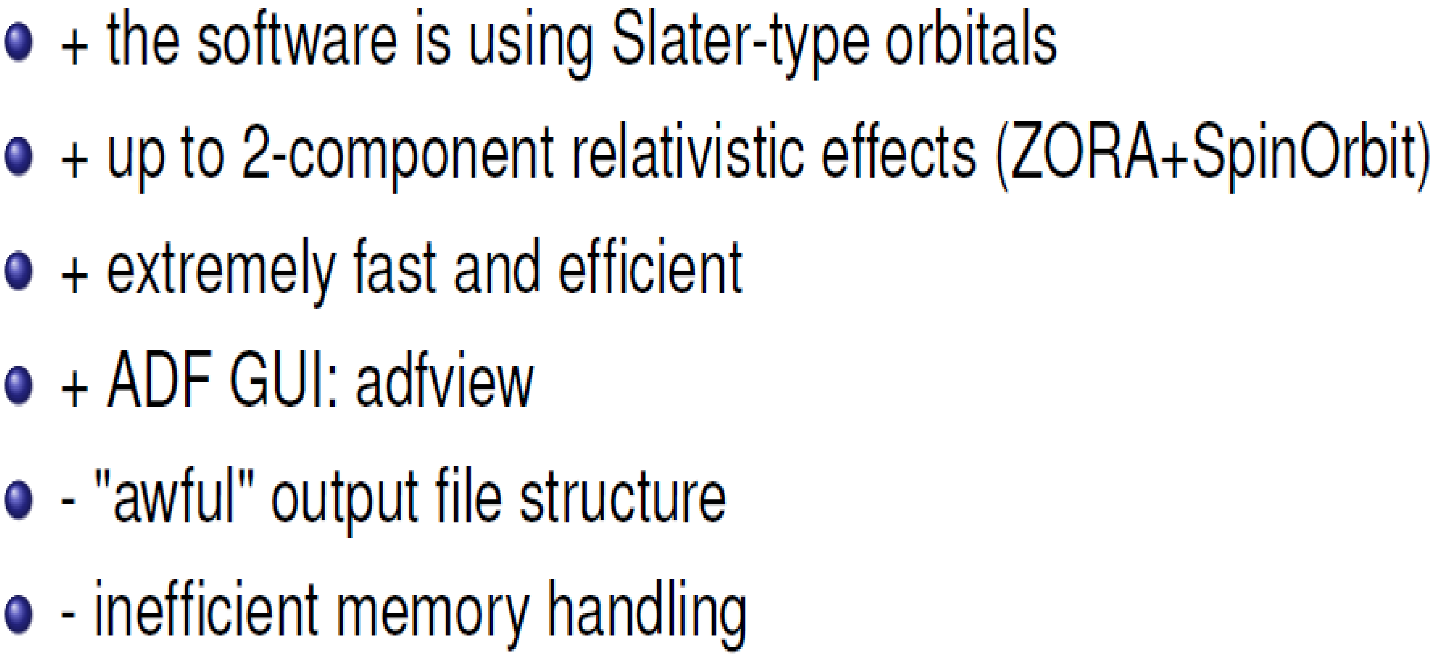
File	Contents
"name.out"	contains data for the chosenprocedure
energy	Energies of steps
gradient	Gradients of steps
mos	Molecular orbitals
freq.out	Output from aoforce program

- to extract data for a geom.optimization run:

user@wolf

```
#!/bin/bash
module add turbomole
t2x > jobname.xyz (generates a series of geometry frames for the entire run)
t2x -c > jobname.xyz (generates the final geometry entire run)
```

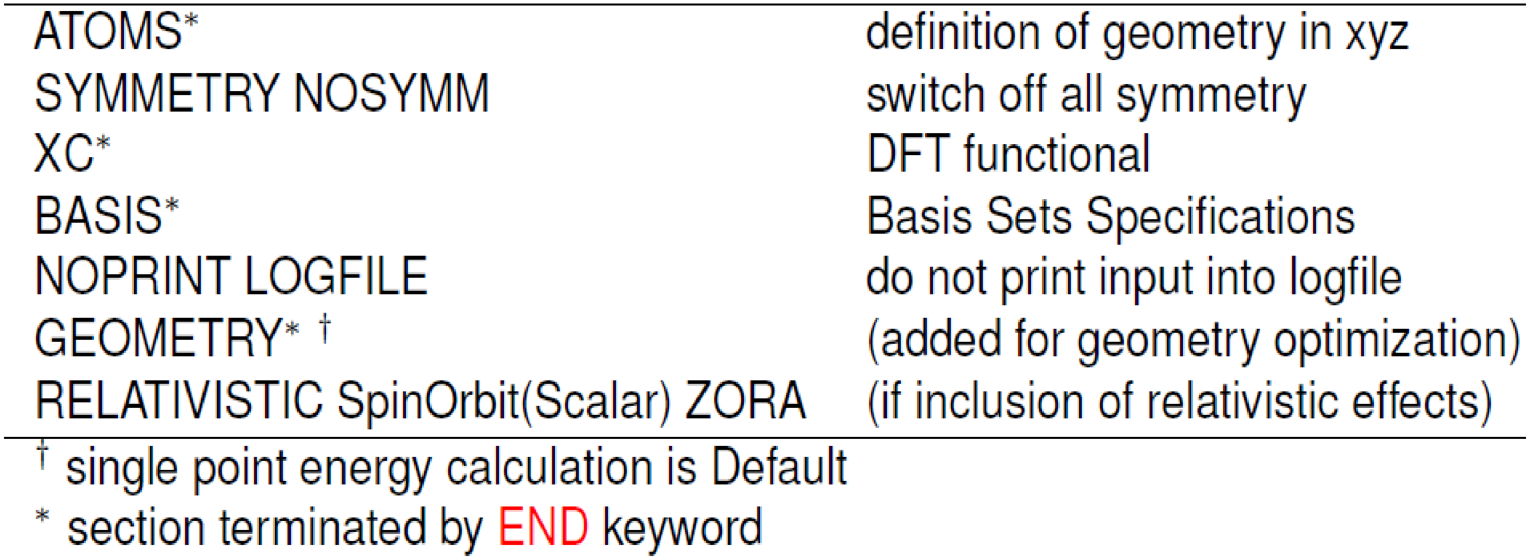
ADF (AMS)



- + the software is using Slater-type orbitals
- + up to 2-component relativistic effects (ZORA+SpinOrbit)
- + extremely fast and efficient
- + ADF GUI: adfview
- - "awful" output file structure
- - inefficient memory handling

ADF (AMS)

- keywords in blocks, case insensitive:



ATOMS*	definition of geometry in xyz
SYMMETRY NOSYMM	switch off all symmetry
XC*	DFT functional
BASIS*	Basis Sets Specifications
NOPRINT LOGFILE	do not print input into logfile
GEOMETRY* †	(added for geometry optimization)
RELATIVISTIC SpinOrbit(Scalar) ZORA	(if inclusion of relativistic effects)

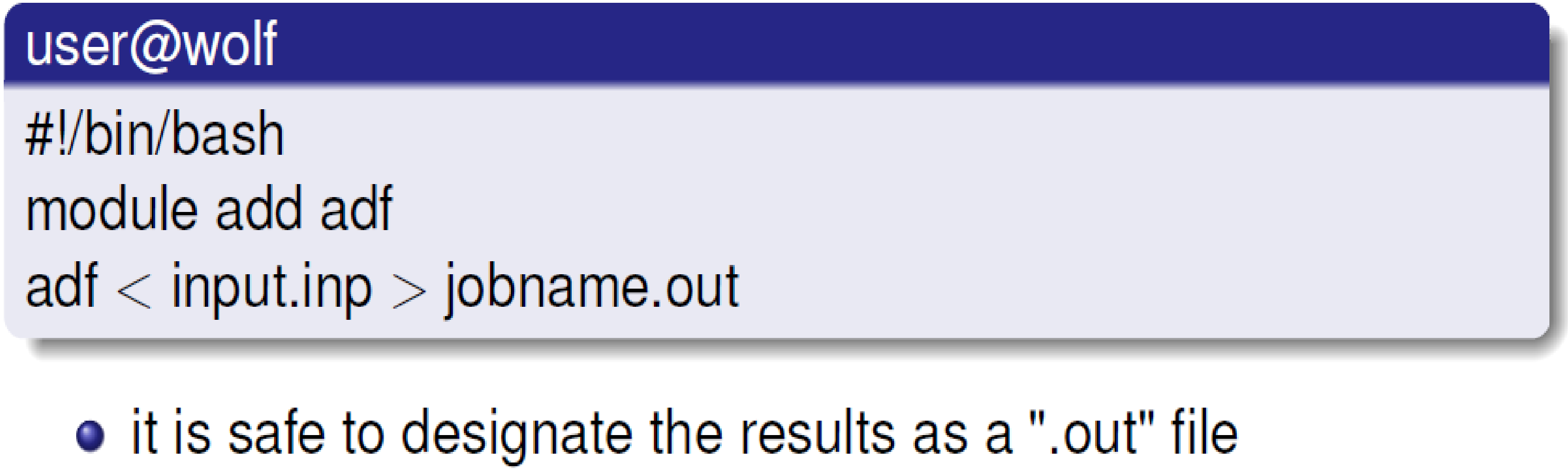
† single point energy calculation is Default

* section terminated by END keyword

- MANUAL:

`https://www.scm.com/doc/Documentation/index.html`

ADF



user@wolf

```
#!/bin/bash
module add adf
adf < input.inp > jobname.out
```

- it is safe to designate the results as a ".out" file

ADF output

- important files of a succesful ADF run

File	Contents
"name.out"	contains data for the chosen procedure
TAPE* files	contains various informations including wavefunction
logfile	brief summary of the job

- using ADFVIEW (ADF GUI) for output visualization

user@wolf

```
#!/bin/bash
module add adf
adfview TAPE*
```