



Tutorial 3: Molecules with DFT

Lab 3. Gaussian

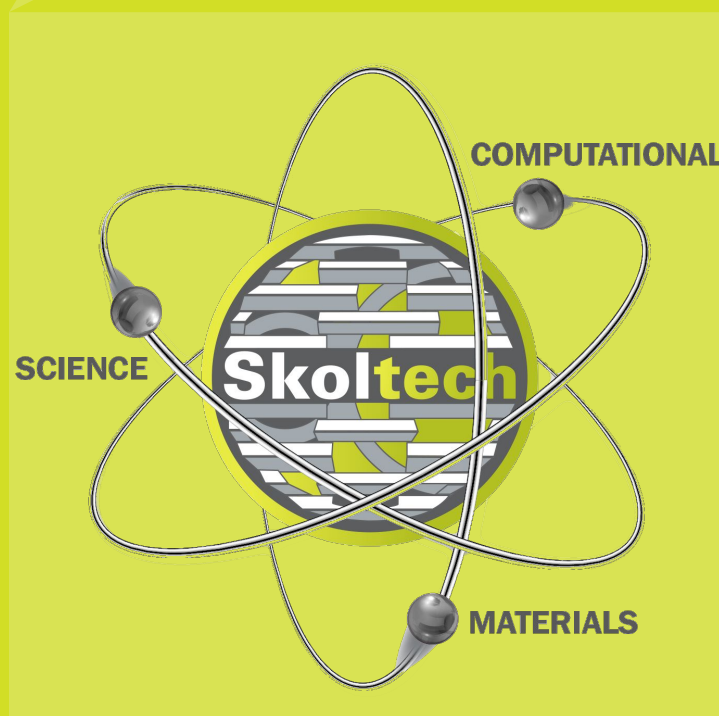
The authors:

PhD Dmitry Govorov

PhD Arseniy Burov

November, 2023

Skoltech



Tutorial 3 agenda

1. Prerequisites for Lab3

- a. DFT for molecules
- b. Set-up your environment

2. Lab 3. Gaussian

- a. How to set-up your lab
- b. Units
- c. Basic commands and functions
- d. Lab tasks

Gaussian-type orbitals (GTO)

Idea: Fit Slater-type atomic orbitals with a superposition of Gaussian orbitals

$$\varphi(x, y, z; \{\alpha\}, i, j, k) = \sum_{a=1}^M c_a \phi(x, y, z; \alpha_a, i, j, k)$$

- GTO's are "atomic orbital-like".
- Compact basis set (approx. 15-25 functions per atom).
- Analytic integration possible for many operators. Optimal for regular grids.
- Fourier transform is again a Gaussian.
- Compact support (finite extend)
- Non-orthogonal basis.
- Linear dependencies for larger basis sets.
- Complicated to generate and no easy way to improve.
- Basis set superposition error (BSSE).
- Molecules (wavefunction tails) and solids have different requirements.

Gaussian-type orbitals (GTO)

Radial part

Spherical harmonic

$$\Phi(\mathbf{r}) = R_l(r) Y_{lm}(\theta, \phi)$$

In Cartesian coordinates:
$$\Phi(x, y, z; \alpha, i, j, k) = \left(\frac{2\alpha}{\pi}\right)^{3/4} \left[\frac{(8\alpha)^{i+j+k} i! j! k!}{(2i)!(2j)!(2k)!}\right]^{1/2} x^i y^j z^k e^{-\alpha(x^2+y^2+z^2)}$$

Gaussian-type orbitals (GTO)

$$\Phi_{\text{GTO}} = N x^l y^m z^n e^{-\alpha r^2}$$

- x, y, z – Cartesian coordinates of nucleus
- l, m, n – angular momenta
- $l + m + n = L$ – total angular momentum
- $L = 0$: s-orbital
- $L = 1$: p_x, p_y, p_z - orbitals
- $L = 2$: five d -orbitals
- $L = 3$: seven f -orbitals

Ideas behind the Lab 3

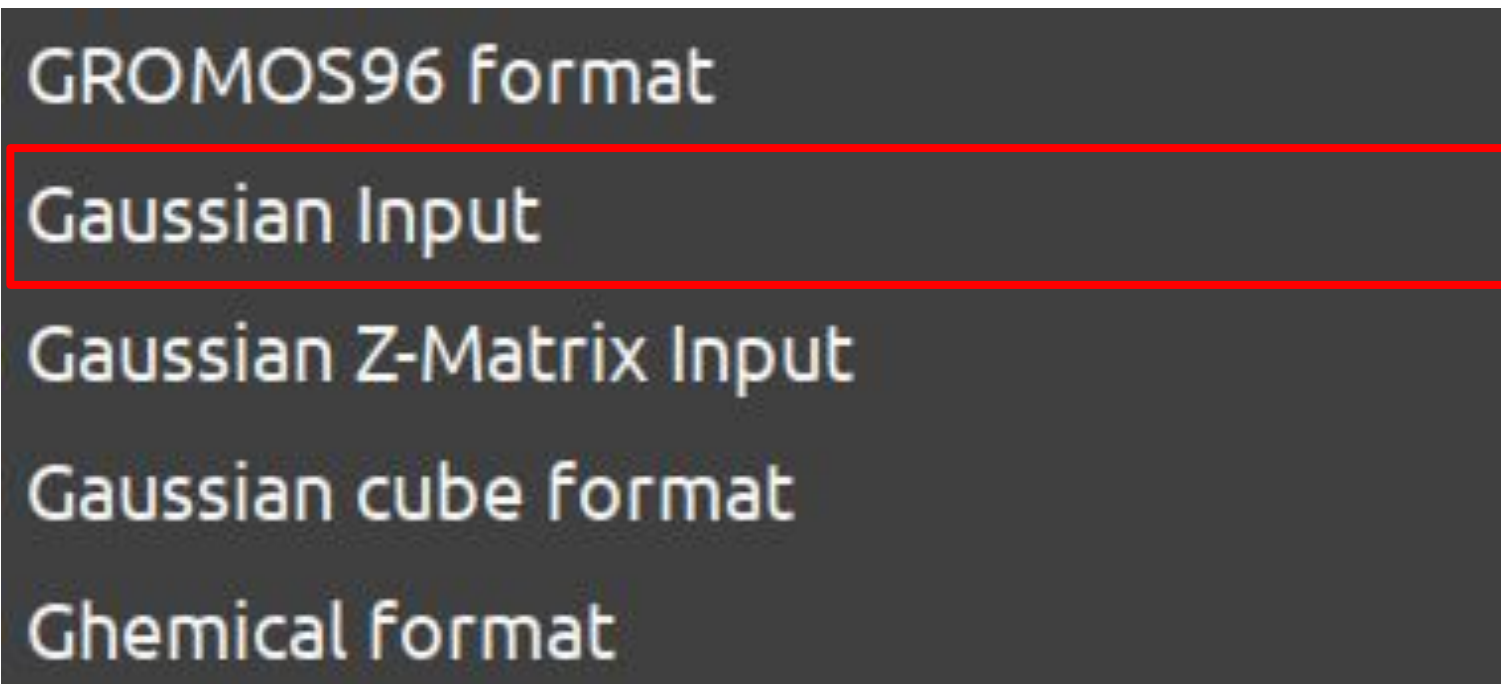
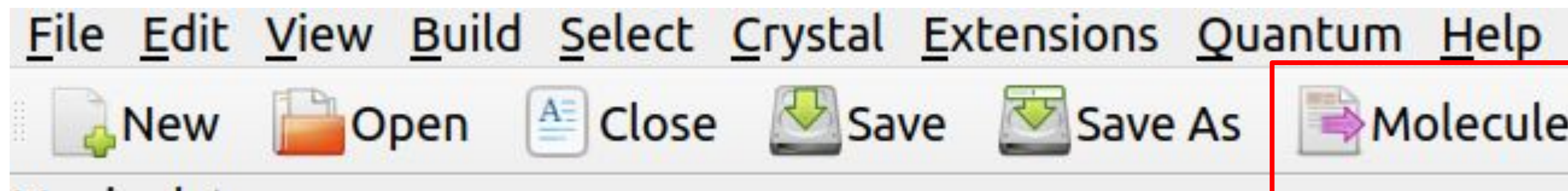
- **Learn** how to perform DFT calculations for molecules
- **Get familiar** with localized basis set (Gaussian-type orbitals)
- **Get familiar** with calculations, using hybrid functionals
- **Compare** the results of different-level methods: Quantum chemistry (DFT) calculations vs semi-empirical

Construction of molecules for Gaussian



Avogadro editor

Avogadro is an advanced molecule editor and visualizer designed for cross-platform use in computational chemistry, molecular modeling, bioinformatics, materials science, and related areas. It offers flexible high quality rendering and a powerful plugin architecture.



Export structure in the format of Gaussian. We recommend to use OpenBabel to *open* your structure and to *export* it.

Lab 3. Set up Gaussian and files



Settings the Lab on the Virtual Machine

Credentials for the Virtual Machine:

'your_login'@10.30.16.178

'your_password'

Download archive with [Lab files for benzene](#).

Upload the archive to the Virtual Machine:

```
scp lab3_benzene.zip a.burov@10.30.16.178:
```

Login to the Virtual Machine and unzip the archive:

```
unzip lab3_benzene.zip
```

Lab 3. Gaussian basics



Gaussian units

Mass	–	atomic mass units
Temp	–	K
Energy	–	atomic units (Hartree), cm^{-1} (for IR, Raman) eV and nm for TDDFT
Distance	–	Å by default , converted to Bohr for calculations
Time	–	very seldom used, 25 attoseconds in BOMD
Force	–	atomic units
Pressure	–	atm
Charge	–	elementary charge
Dipole	–	Debye

How to use Gaussian

To find input files for your calculations, you need go to *lab3_benzene/sampleinputs*

To run:

`g16 <input_file.gau> or <input_file.gjf> or <input file.com>`

.log file is automatically created

Windows version has a GUI

Structure of input file

How many cores to use?
All available are used by default

How much RAM to use?
The default isn't enough for large calculations

Checkpoint file.
Can be used to restart jobs

`%CPU=0-15`

`%Mem=8GB`

`%Chk=filename.chk`

`#n <method>/<basis set> Opt`

Calculation settings

Comment

System charge

0 1

Multiplicity

Cartesian coordinates

O 0.82554

-0.00914

0.00000

H 1.79554

-0.00914

0.00000

H 0.50222

0.31771

0.85412

Two blank lines

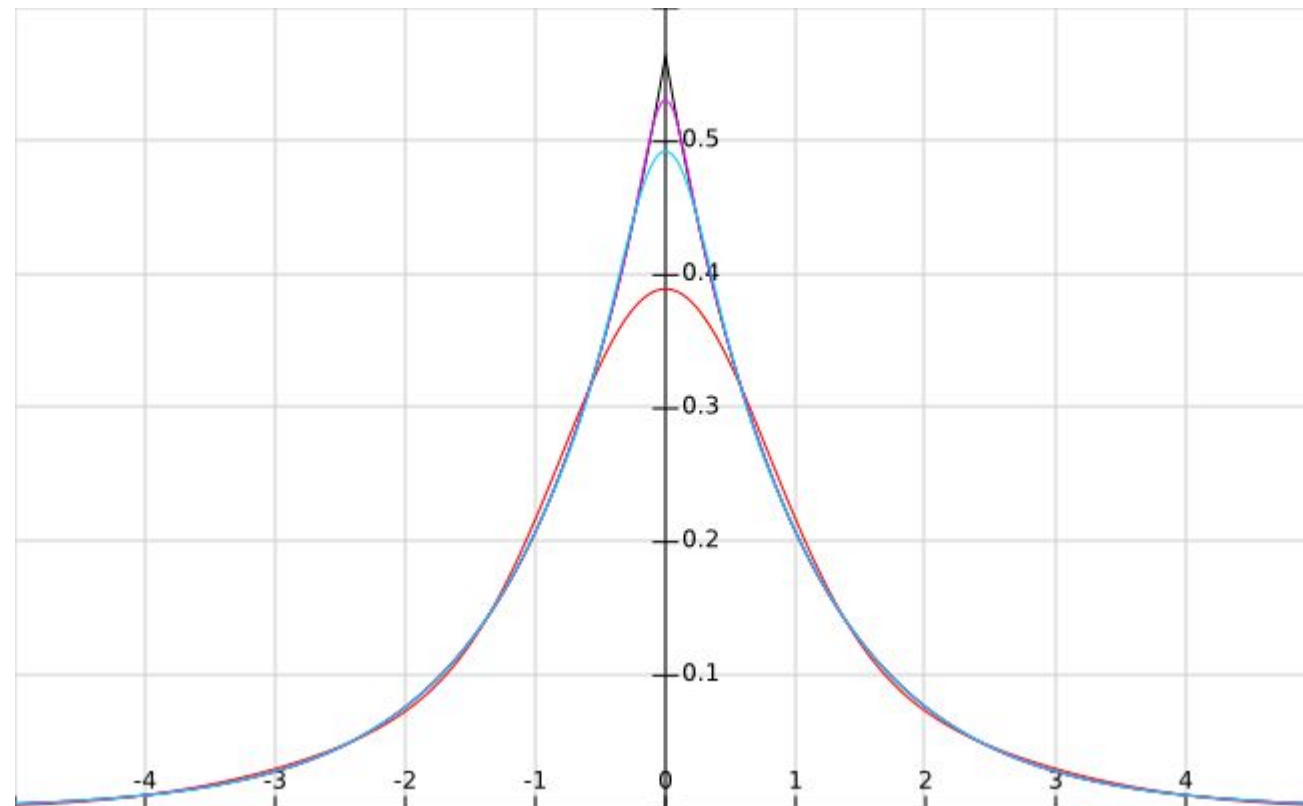
Basis sets

Gaussian is called so after Gaussian basis sets it uses

Main families of Gaussian basis sets:

1. *Pople's* basis sets: 3-21G, 4-31G, 6-31G, 6-311G - all are split-valence, quite outdated
2. *Dunning's*: cc-pVxZ, x = D, T, Q, 5 etc - very useful for post-HF because convenient for CBS
3. *Ahlrich's*: def2-SVP, def2-TZVP, def2-QZVP, more optimal for DFT
4. *Jensen's*: pc-n, n = 0, 1, 2, etc; pc-seg-n, n = 0, 1, 2
5. *Atomic natural orbitals* (ANO), very large and deeply contracted basis sets

Basis sets

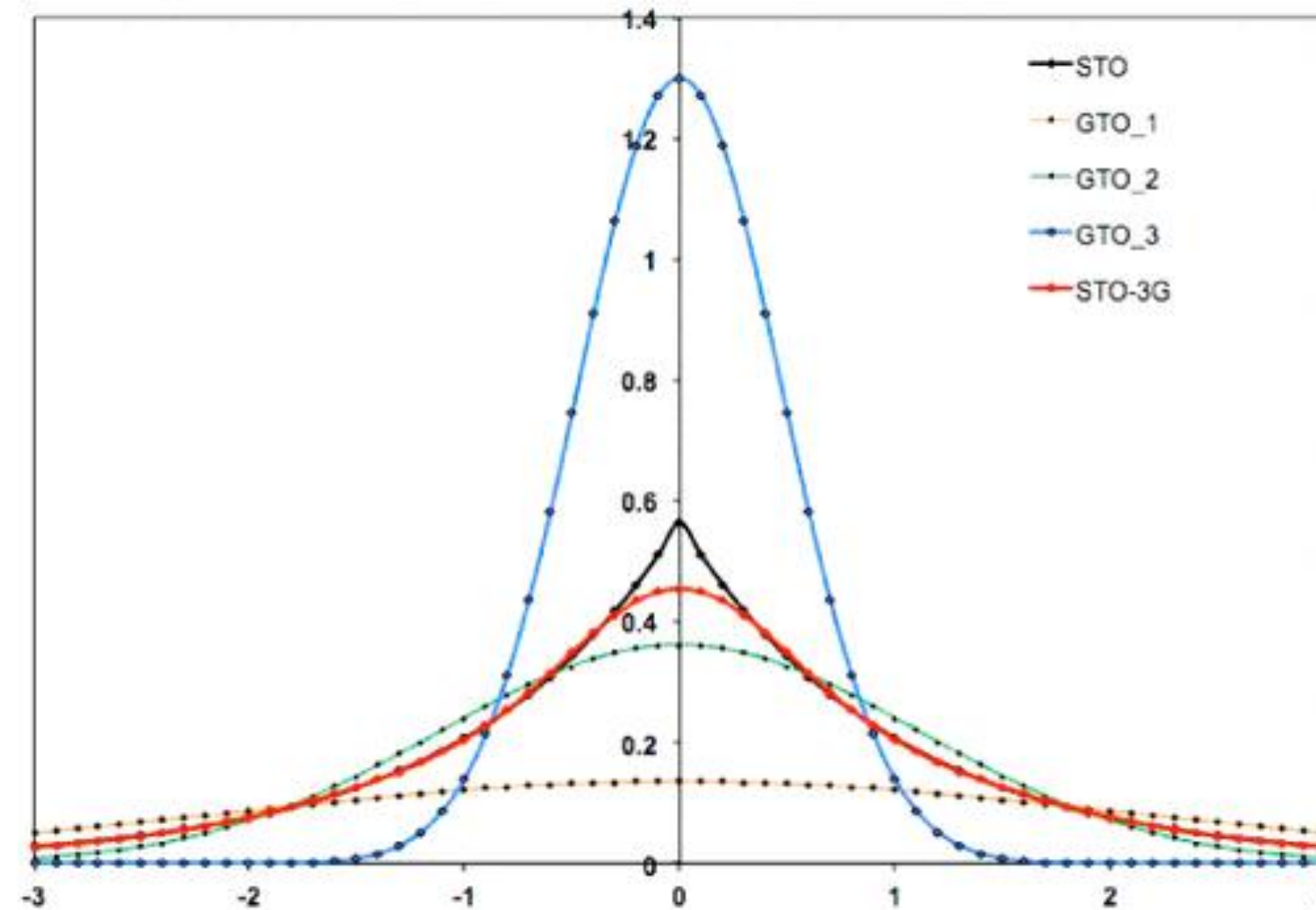


Ideally, a Slater-type basis set is needed because of its shape: slow enough decay and a sharp cusp (tip):

$$R(r) = N r^{n-1} e^{-\zeta r}$$

Using them wasn't practical because integrating such functions is hard

Convenient alternative: r^2



Gaussian basis functions are stacked in a linear combination to make a contracted basis function:

$$\varphi_{\text{CGTO}} = \sum c_i \varphi_{\text{GTO}}$$

STO-*n*G basis set

Minimal basis sets, where n primitive Gaussian orbitals are fitted to a single Slater-type orbital. It is a *split-valence* basis set, the number of basis functions that is assigned to core orbitals differs from the one for the valence orbitals.

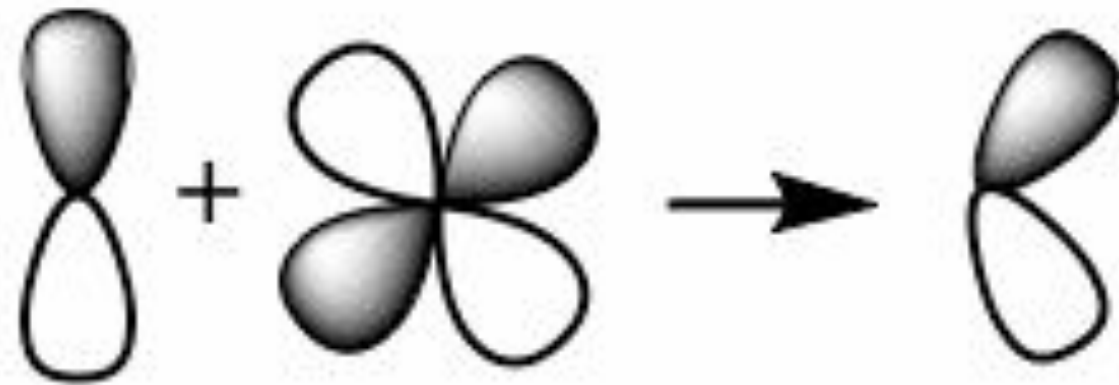
Number of contracting gaussians in the function describing a core orbital

Number of contracting gaussians forming the second valence function

3-21G

Number of contracting gaussians forming the first valence function

Basis sets



Polarization basis functions are linear combinations of basis functions with $AM = x$ and basis functions with $AM = x + 1$

Used for describing the hybridization effect

Marked as * (d) or ** (d,p) in *Pople's* basis sets: **6-31G***, **6-31G****

Already included in *Dunning's* basis sets

Diffuse basis functions are simply basis functions with a small exponent value. They are used to describe electron density far from nuclei.

Useful for more accurate energies, dipole moments, polarizabilities. Sometimes may cause the SCF to not converge. Almost mandatory for anions.

Marked as + or ++ in *Pople's* basis sets: **6-311++G****

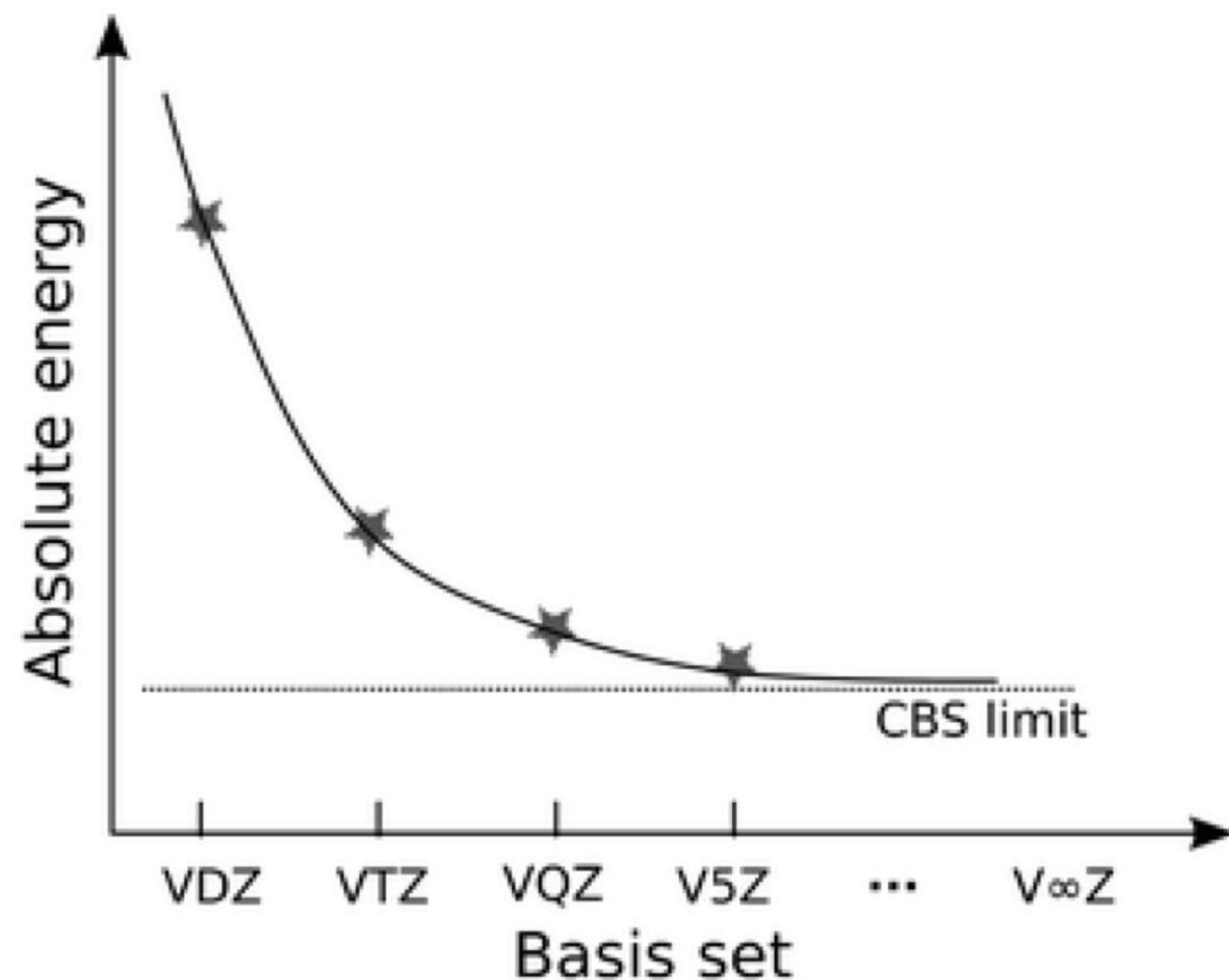
Marked as aug- for *Dunning's* basis sets: **aug-cc-pVTZ**

Marked as D for *Ahlrich's*: **def2-TZVPD**

How to choose basis sets

The rule is simple - the larger, the better

An ideal basis set is infinite (complete)



The larger the basis set - the lower E you get and this applies to forces and other values.

In practice:

6-311+G** is good enough for DFT geometry optimizations and frequencies, 6-31+G**, def2-SVP, def2-TZVP too
While aug-cc-pVTZ, aug-cc-pVQZ are appropriate for MP2, CCSD, CCSD(T) and such
STO-nG, 3-21G, 4-31G are just toys

What is Gaussian good for?

1. Semiempirical calculations for molecules
2. HF, some post-HF
3. DFT calculations
4. TDDFT calculations
5. Potential energy surveys (bond/angle scanning)
6. QM/MM

What is Gaussian not good for?



What is Gaussian not good for?

1. Calculations in periodic boundary conditions
2. Multireference calculations
3. Large-scale post-HF calculations

How to run single-point energy calculations

```
%CPU=0-15
%Mem=8GB
%Chk=filename.chk
#n <method>/<basis set> (Semiempirical don't need basis sets)
Comment
0 1
O      0.82554      -0.00914      0.00000
H      1.79554      -0.00914      0.00000
H      0.50222      0.31771      0.85412
```

Dipole and multipole moments will also be printed
Molecular orbitals are contained in [.chk](#) files
They are converted to .fchk using [formchk](#) program

How to run geometry optimizations

```
%CPU=0-15
```

```
%Mem=8GB
```

```
%Chk=filename.chk
```

```
#n <method>/<basis set> Opt Freq
```

```
Comment
```

```
0 1
```

```
O          0.82554          -0.00914          0.00000
```

```
H          1.79554          -0.00914          0.00000
```

```
H          0.50222          0.31771          0.85412
```

Freq command calculates harmonic frequencies after the Opt is over. Normal mode analysis and thermochemistry will be printed out

Output file

```
SCF Done: E(UB3LYP) = -931.873862428
A.U. after 26 cycles
NFock= 26 Conv=0.49D-08 -V/T= 2.0042
```

Info about individual SCF cycles can be printed using #p

```
Alpha occ. eigenvalues -- -14.50217 -14.41731 -14.41000 -10.26267 -10.21463
Alpha occ. eigenvalues -- -10.21385 -10.20766 -10.20189 -10.20176 -10.19819
```

Orbital energies

Mulliken charges and spin densities:

Mulliken population analysis can be done, spin charges are calculated when multiplicity $\neq 1$

```
-----
Center      Atomic      Forces (Hartrees/Bohr)
Number      Number          X          Y          Z
-----
```

Forces or energy gradients

Output file

	Item	Value	Threshold	Converged?
Maximum	Force	0.002860	0.000015	NO
RMS	Force	0.000336	0.000010	NO
Maximum	Displacement	0.054766	0.000060	NO
RMS	Displacement	0.008174	0.000040	NO

Useful to look at when doing
geometry optimizations

	1	2	3
	A	A	A
Frequencies --	45.7127	112.1230	120.7811
Red. masses --	8.3448	9.8275	7.4890
Frc consts --	0.0103	0.0728	0.0644
IR Inten --	0.5534	1.2086	1.0928

This is how vibrational
frequencies are
printed out

Zero-point correction=	0.228965 (Hartree/Particle)
Thermal correction to Energy=	0.244086
Thermal correction to Enthalpy=	0.245030
Thermal correction to Gibbs Free Energy=	0.186565
Sum of electronic and zero-point Energies=	-931.644927
Sum of electronic and thermal Energies=	-931.629806
Sum of electronic and thermal Enthalpies=	-931.628862
Sum of electronic and thermal Free Energies=	-931.687327

Zero-point energy, H,
G, S are calculated

Output file

Excitation energies and oscillator strengths:

```
Excited State   1:   3.168-A      1.4209 eV   872.57 nm   f=0.0117   <S**2>=2.259
  76A -> 77A           0.93970
  76A -> 78A          -0.10662
  76A -> 81A           0.12822
  72B -> 75B           0.19949
```

This state for optimization and/or second-order correction.

Total Energy, E(TD-HF/TD-DFT) = -931.245631434

Copying the excited state density for this state as the 1-particle RhoCI density.

TDDFT. Electronic state energy is provided in eV, nm
f - oscillator strength (absorption intensity)

How to find a transition state for a reaction

```
%CPU=0-15
```

```
%Mem=8GB
```

```
%Chk=filename.chk
```

```
#n <method>/<basis set> Opt(TS,CalcFC,NoEigenCheck) Freq
```

```
Comment
```

```
0 1
```

```
O          0.82554          -0.00914          0.00000
```

```
H          1.79554          -0.00914          0.00000
```

```
H          0.50222          0.31771          0.85412
```

Calculate Force Constants
= Calculate Hessian
= Calculate harmonic frequencies

Requests optimization
to a transition

For TS you need to guess the starting geometry accurately and
calculate harmonic frequencies beforehand and afterwards
TS searches are to be followed with IRC

How to account for solvation

```
%CPU=0-15
```

```
%Mem=8GB
```

```
%Chk=filename.chk
```

```
#n <method>/<basis set> SCRF(PCM,Solvent=<solvent name>)
```

```
Comment
```

```
0 1
```

```
O          0.82554          -0.00914          0.00000
```

```
H          1.79554          -0.00914          0.00000
```

```
H          0.50222          0.31771          0.85412
```

SCRF – a calculation to be performed in the presence of a solvent

Google “gaussian solvents” to find the list of solvents
CPCM and SMD can be used in place of PCM

How to do TDDFT calculations

```
%CPU=0-15
```

```
%Mem=8GB
```

```
%Chk=filename.chk
```

```
#n <XC functional>/<basis set> TD(Root=,NState=)
```

```
Comment
```

```
0 1
```

```
O          0.82554          -0.00914          0.00000
```

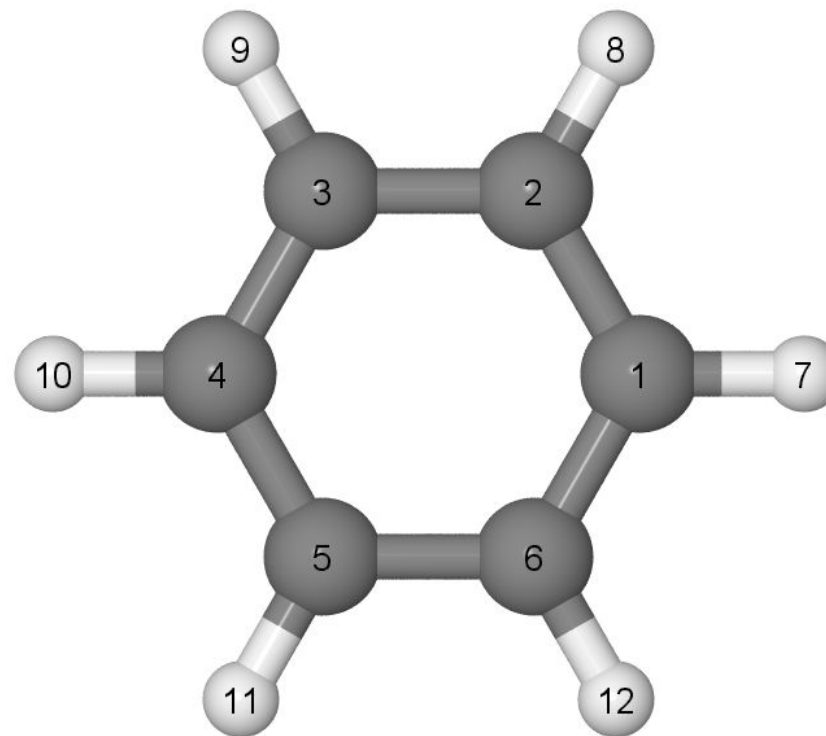
```
H          1.79554          -0.00914          0.00000
```

```
H          0.50222          0.31771          0.85412
```

Root - state of interest, NState = how many ES to solve for
Opt and Freq can be added as well so ES geometries can be located

Without XC functional supplied CIS calculation will be done

Lab 3. Benzene



1. Geometry optimization

Type of functional

Basis set

Optimization settings

```
%nprocshared=8
%mem=8GB
# B3LYP 6-31G* opt(tight,maxcycles=44) int(ultrafine)
```

two-electron integrals

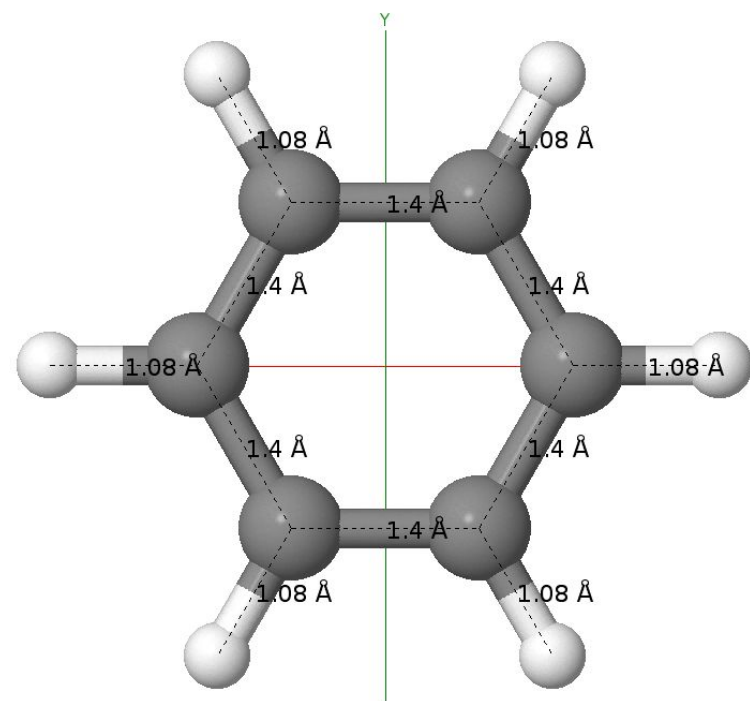
0	1			
C	1.39700000	0.00000000	0.00000000	
C	0.69850000	1.20983749	0.00000000	
C	-0.69850000	1.20983749	0.00000000	
C	-1.39700000	0.00000000	0.00000000	
C	-0.69850000	-1.20983749	0.00000000	
C	0.69850000	-1.20983749	0.00000000	
H	2.48100000	0.00000000	0.00000000	
H	1.24050000	2.14860903	0.00000000	
H	-1.24050000	2.14860903	0.00000000	
H	-2.48100000	0.00000000	0.00000000	
H	-1.24050000	-2.14860903	0.00000000	
H	1.24050000	-2.14860903	0.00000000	

molec_B3LYPp2p_em.gau .gau extension

Skoltech

1. Geometry optimization

- Information on the input data (functional, basis set, structure, etc.).
- Relaxed structure, its symmetry, bond lengths.
- Occupation of orbitals.
- Technical parameters.



Stoichiometry	C6H6					
Framework group	D6H[3C2'(HC.CH)]					
Deg. of freedom	2					
Full point group	D6H	NOp	24			
Largest Abelian subgroup	D2H	NOp	8			
Largest concise Abelian subgroup	D2	NOp	4			
Standard orientation:						

Center	Atomic	Atomic	Coordinates (Angstroms)			
Number	Number	Type	X	Y	Z	

1	6	0	0.000000	1.397000	0.000000	
2	6	0	1.209837	0.698500	-0.000000	
3	6	0	1.209837	-0.698500	-0.000000	
4	6	0	-0.000000	-1.397000	0.000000	
5	6	0	-1.209837	-0.698500	-0.000000	
6	6	0	-1.209837	0.698500	-0.000000	
7	1	0	0.000000	2.481000	0.000000	
8	1	0	2.148609	1.240500	-0.000000	
9	1	0	2.148609	-1.240500	-0.000000	
10	1	0	-0.000000	-2.481000	0.000000	
11	1	0	-2.148609	-1.240500	-0.000000	
12	1	0	-2.148609	1.240500	-0.000000	

Rotational constants (GHZ):			5.6868180	5.6868180	2.8434090	

molec_B3LYPp2p_em.out .out extension

2. Frequency calculations

Checkpoint
file

```
%nprocshared=8
%mem=8GB
%chk=benzene_molec_B3LYPp2p_freq.chk
# B3LYP 6-31G* int(ultrafine) Freq(Raman,SaveNM)

0 1
C      1.39661728      0.00000000      0.00000000
C      0.69830864      1.20950605      0.00000000
C     -0.69830864      1.20950605      0.00000000
C     -1.39661728      0.00000000      0.00000000
C     -0.69830864     -1.20950605      0.00000000
C      0.69830864     -1.20950605      0.00000000
H      2.48363019      0.00000000      0.00000000
H      1.24181509      2.15088684      0.00000000
H     -1.24181509      2.15088684      0.00000000
H     -2.48363019      0.00000000      0.00000000
H     -1.24181509     -2.15088684      0.00000000
H      1.24181509     -2.15088684      0.00000000
```

Compute force constants
and the resulting
vibrational frequencies

2. Frequency calculations

Convert output file to format,
which is readable

```
formchk benzene_molec_B3LYPp2p_freq.chk
```

```
Freq      RB3LYP                                     6-31G(d)
Number of atoms                                I      12
Info1-9                                I  N=      9
      22      20      0      0      0      100
      6      18     -502
Route                                C  N=      4
# B3LYP 6-31G* int(ultrafine) Freq(Raman,SaveNM)
Charge                                I      0
Multiplicity                            I      1
Number of electrons                      I     42
Number of alpha electrons                I     21
Number of beta electrons                 I     21
Number of basis functions                I    102
Number of independent functions           I    102
Number of point charges in /Mol/         I      0
Number of translation vectors            I      0
Atomic numbers                          I  N=     12
      6      6      6      6      6      6
      1      1      1      1      1      1
Nuclear charges                          R  N=     12
  6.000000000E+00  6.000000000E+00  6.000000000E+00  6.000000000E+00  6.000000000E+00
  6.000000000E+00  1.000000000E+00  1.000000000E+00  1.000000000E+00  1.000000000E+00
  1.000000000E+00  1.000000000E+00
```


3. Calculation of Localized Molecular Orbitals

Put basis in output file in format for generalized input

Printing of MOs

Prints the current basis set and density fitting basis set in tabular form

```
%nprocshared=8
%mem=8GB
#T B3LYP 6-31G* int(ultrafine) GFINPUT IOP(6/7=3) GFPRINT IOP(5/33=1) pop(NBO,SaveNLMOs) Density

0 1
C      1.39661728      0.00000000      0.00000000
C      0.69830864      1.20950605      0.00000000
C     -0.69830864      1.20950605      0.00000000
C     -1.39661728      0.00000000      0.00000000
C     -0.69830864     -1.20950605      0.00000000
C      0.69830864     -1.20950605      0.00000000
H      2.48363019      0.00000000      0.00000000
H      1.24181509      2.15088684      0.00000000
H     -1.24181509      2.15088684      0.00000000
H     -2.48363019      0.00000000      0.00000000
H     -1.24181509     -2.15088684      0.00000000
H      1.24181509     -2.15088684      0.00000000
```

Printing of molecular orbitals and several types of population analysis and atomic charge assignments

Select which density to analyze

3. Visualization of LMOs

Number of molecular
orbitals to generate .
Starts from low-energy MOs

Convert output file to format,
which can be visualized

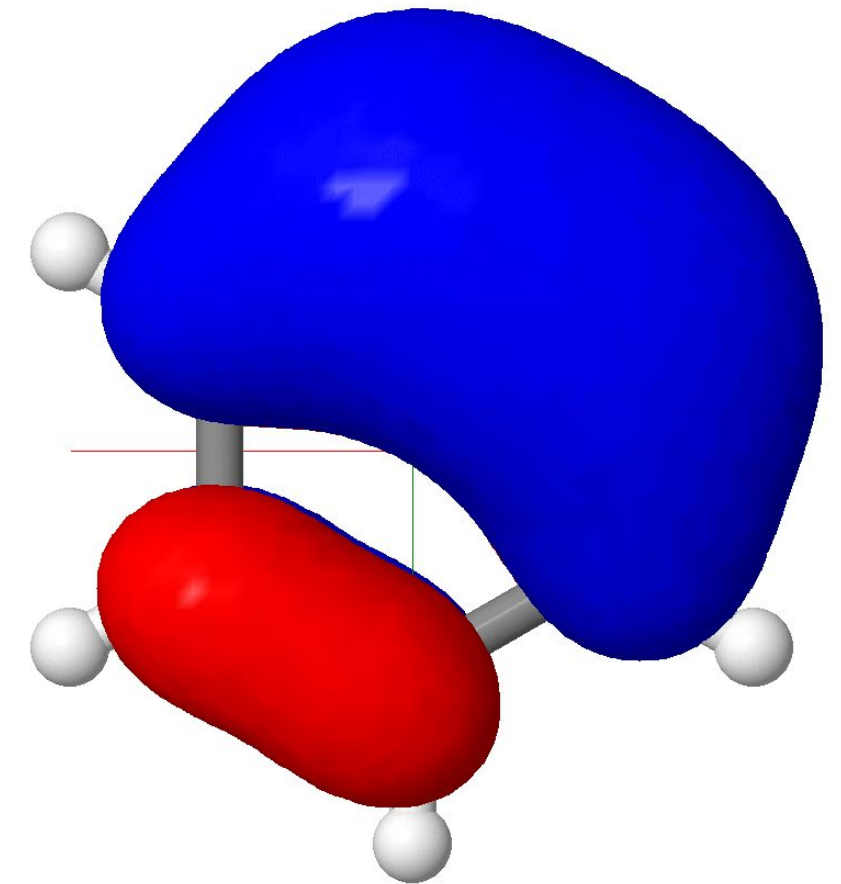
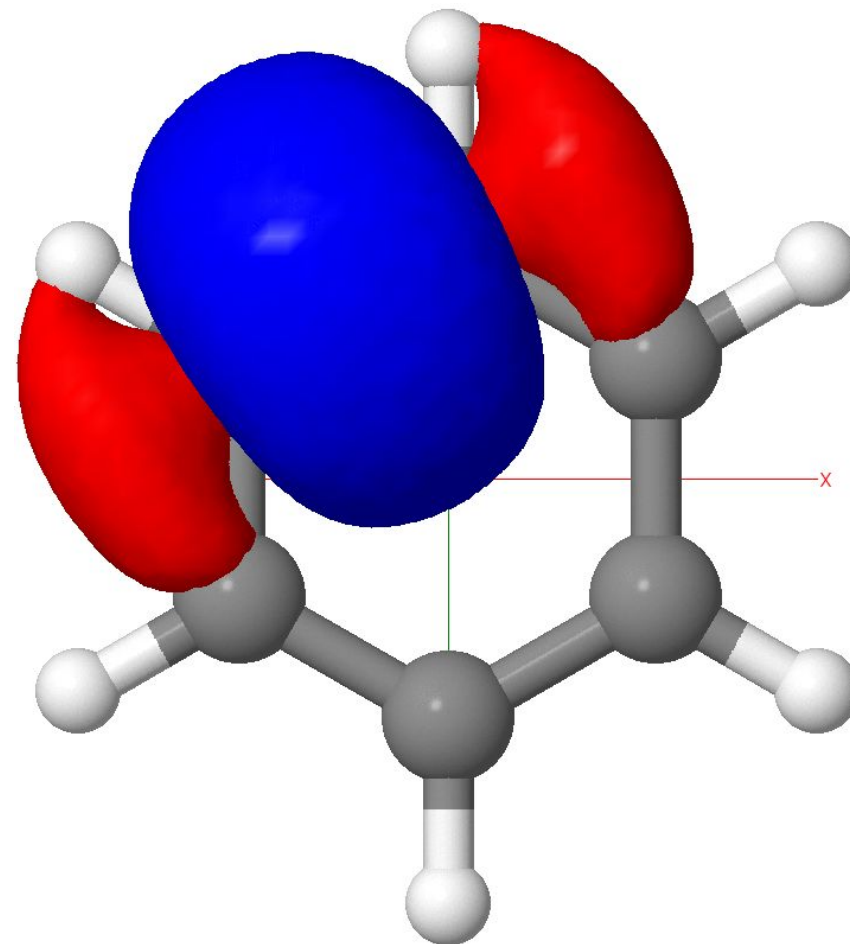
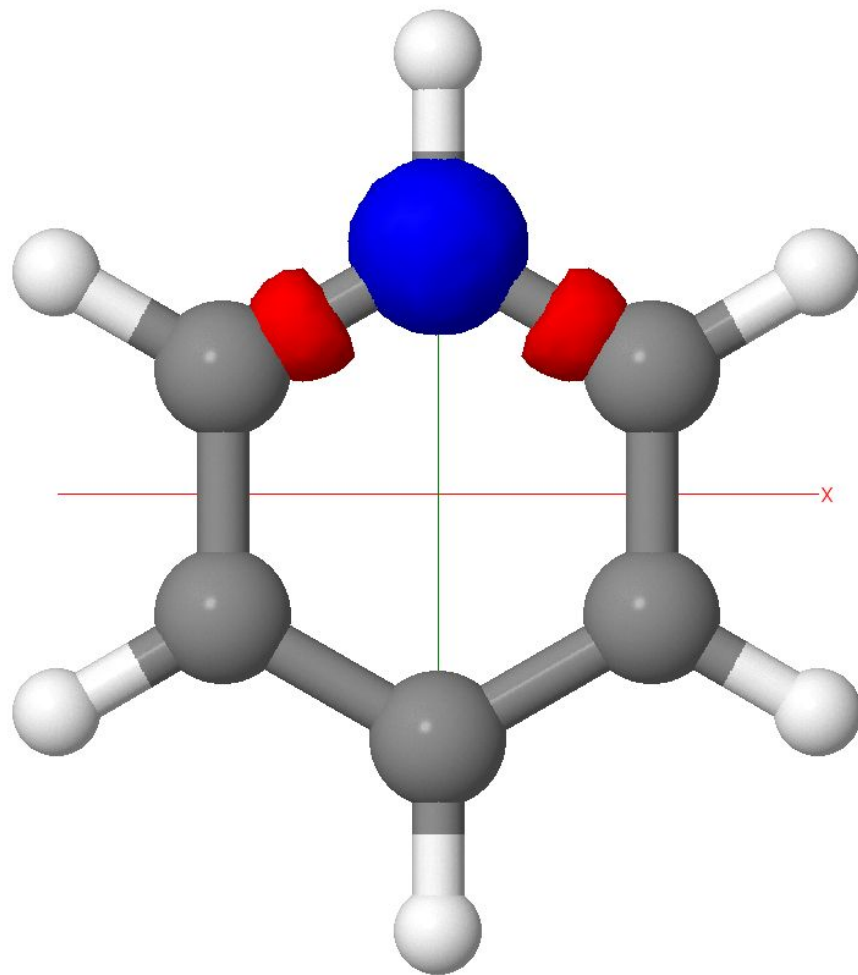
```
./chk2cube.sh benzene_molec_B3LYPp2p_1mo.chk 20
```

Checkfile from your
calculations

The generated files with MOs will be in **orbs.zip** archive.

3. Visualization of LMO

LMOs can be visualized with Jmol software ([orbs.zip](#) archive). However, the LMOs energies and occupation can be found only in [.out](#) file.



Compare results with the semi-empirical methods

Compare the results of Gaussian with results of MOPAC, which uses semi-empirical methods.
What are the difference in HOMO-LUMO gap, MOs, LMOs, etc?



Extra materials

- **Google gaussian + scf/opt/freq/nmr/td/scrf**

This will give you documentation for all Gaussian commands and references to literature.

- [Full list of Gaussian keywords.](#)
- [Chemcraft](#) program can be used for visualization and has a long trial period.
- [Population analysis.](#)
- Official [tutorials](#) from Gaussian team.
- [IR SPect](#): A Python tool to extract frequencies and IR/Raman.
- [basissetexchange.org](#) to look at different basis sets.

Thx