

# Electronic Structure Theories (ab initio, DFT) and Software

Jemmy Hu  
SHARCNET HPC Consultant

at Trent University  
June 23, 2010

[https://www.sharcnet.ca/~jemmyhu/Trent\\_talk.ppt](https://www.sharcnet.ca/~jemmyhu/Trent_talk.ppt)

# Introduction

- Theoretical background
  - ab initio methods
  - density functional theory (DFT)
  - basis set
  - pseudopotential
- Software
  - software in category
  - software in detail: Gaussian, NWChem, Abinit, PWSCF
- Others: software in Molecular Dynamics

# ab initio methods

- Schrödinger equation

$$\hat{H}\Psi = [\hat{T} + \hat{V} + \hat{U}] \Psi = \left[ \sum_i^N -\frac{\hbar^2}{2m} \nabla_i^2 + \sum_i^N V(\vec{r}_i) + \sum_{i<j}^N U(\vec{r}_i, \vec{r}_j) \right] \Psi = E\Psi$$

it is in general still not possible to find the exact solution,  $\Psi$ , of the electronic Schrödinger equation. Therefore,  $\Psi$  is expanded in known functions and solved numerically.

- Hartree–Fock method (Self-Consistent Field, SCF)

- exact, N-body wave function of the system can be approximated by a single [Slater determinant](#) (fermions) or by a single permanent (bosons) of N [spin-orbitals](#).
- variational principle to get HF wave function and energy for ground state of the system
- each electron is subjected to the mean field created by all other electrons
- electron correlation is completely neglected for the electrons of opposite spin

- Methods needed to address “**electron correlation**”

# Post-Hartree–Fock methods

- Configuration interaction (CIS, CISD, QCI, QCISD)
- Møller–Plesset perturbation theory (MP2, MP3, MP4)
- Coupled cluster (CC, CCSD, CCSD(T))
- Quantum chemistry composite methods (G2, G3, CBS, etc.)

## Related methods:

- Multi-configurational self-consistent field (MCSCF)
- Multireference single and double configuration interaction (MRDCI)
- N-electron valence state perturbation theory (NEVPT)

# Density Functional Theory (DFT)

DFT provides an alternative way to address the many-body problem, the key variable is the particle density  $n(\vec{r})$ , which for a normalized  $\Psi$  is given by

$$n(\vec{r}) = N \int d^3r_2 \int d^3r_3 \cdots \int d^3r_N \Psi^*(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) \Psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N)$$

This relation can be reversed, i.e. for a given ground-state density  $n_0(\vec{r})$  it is possible, in principle, to calculate the corresponding ground-state wave function  $\Psi_0(\vec{r}_1, \dots, \vec{r}_N)$ .

In other words,  $\Psi_0$  is a unique **functional** of  $n_0$

$$\Psi_0 = \Psi[n_0]$$

The ground-state energy is a functional of  $n_0$

$$E_0 = E[n_0] = \langle \Psi[n_0] | \hat{T} + \hat{V} + \hat{U} | \Psi[n_0] \rangle$$

(Pierre Hohenberg, Walter Kohn, in 1964)

# DFT, Kohn-Sham Equation (1965)

$$\left( -\frac{\hbar^2}{2m} \nabla^2 + v_{\text{eff}}(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}).$$

- variational approach to minimize the energy functional

$$E = \sum_i^N \varepsilon_i - V_H[\rho] + E_{\text{xc}}[\rho] - \int \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \rho(\mathbf{r}) d\mathbf{r}$$

- The key issue is the **exchange-correlation** potential.
- The major problem with DFT is that the exact functionals for **exchange and correlation** are not known except for the free electron gas
- But it turns out that even relatively crude approximations can give excellent results
- Approximations: **local-density approximation (LDA)**  
**Generalized gradient approximations (GGA,**  
take into account the **gradient** of the density  
at the same coordinate)

# Local-density approximation (LDA)

- **A few popular LDA's**

Vosko-Wilk-Nusair (VWN)

Perdew-Zunger (PZ81)

Cole-Perdew (CP)

- **Drawbacks**

- In finite systems, the LDA potential decays asymptotically with an exponential form. This is in error; the true exchange-correlation potential decays much slower in a Coulombic manner.

- LDA potential can not support a Rydberg series and those states it does bind are too high in energy: HOMO too high, predict ionization potential based on Koopman's theorem are poor

- LDA provides a poor description of electron-rich species such as anions where it is often unable to bind an additional electron

- **GGA:** PW92, PBE, .....

# Hybrid Functional

- Difficulties in expressing the exchange part of the energy can be relieved by including a component of the exact exchange energy calculated from **Hartree-Fock** theory. Functionals of this type are known as hybrid functionals.
- **B3LYP**: Becke, three-parameter, Lee-Yang-Parr exchange-correlation functional is:

$$E_{xc}^{\text{B3LYP}} = E_{xc}^{\text{LDA}} + a_0(E_x^{\text{HF}} - E_x^{\text{LDA}}) + a_x(E_x^{\text{GGA}} - E_x^{\text{LDA}}) + a_c(E_c^{\text{GGA}} - E_c^{\text{LDA}}),$$

B3P86, O3LYP, X3LYP

## DFT methods on Gaussian

[http://www.gaussian.com/g\\_tech/g\\_ur/k\\_dft.htm](http://www.gaussian.com/g_tech/g_ur/k_dft.htm)

## DFT methods on NWChem

[http://www.emsl.pnl.gov/capabilities/computing/nwchem/releasenotes/dft\\_functionals.jsp](http://www.emsl.pnl.gov/capabilities/computing/nwchem/releasenotes/dft_functionals.jsp)

# Basis set I

- A **basis set** in chemistry is a set of functions used to create the molecular orbitals, which are expanded as a linear combination of atomic orbitals with the weights or coefficients to be determined.
- **Categories: localized/finite basis sets**
  - 1) Minimal (Slater-type) basis sets  
STO-3G, STO-6G
  - 2) Pople (Gaussian-type) basis sets  
6-31G, 6-31G\*, 6-31+G\*, 6-31G(3df, 3pd), 6-311G, 6-311G\*, ...
  - 3) Correlation-consistent basis sets  
cc-pVDZ - Double-zeta  
cc-pVTZ - Triple-zeta  
cc-pVQZ - Quadruple-zeta  
aug-cc-pVDZ, etc. - Augmented versions of the preceding basis sets with added diffuse functions

# Plane-wave basis sets

- PW Basis sets composed of sets of plane waves down to a cutoff wavelength are often used, especially in calculations involving systems with **periodic boundary conditions**
- PW basis sets can also be applied to non-periodic systems such as isolated molecules. To accomplish this the molecule is placed at the centre of a periodic supercell. If the supercell is large enough the interactions between the molecules in neighbouring cells becomes negligible
- Advantage of PW basis sets:
  - the same basis set can be used for all atomic species,
  - convergence toward completeness can easily be tested,
  - plane waves do not depend on nuclear positions so, unlike localized basis sets, correction terms are not needed for the calculation of forces

# Basis Set II

- **How to select the basis sets?**

theoretically the bigger the better, however the bigger the more expensive.

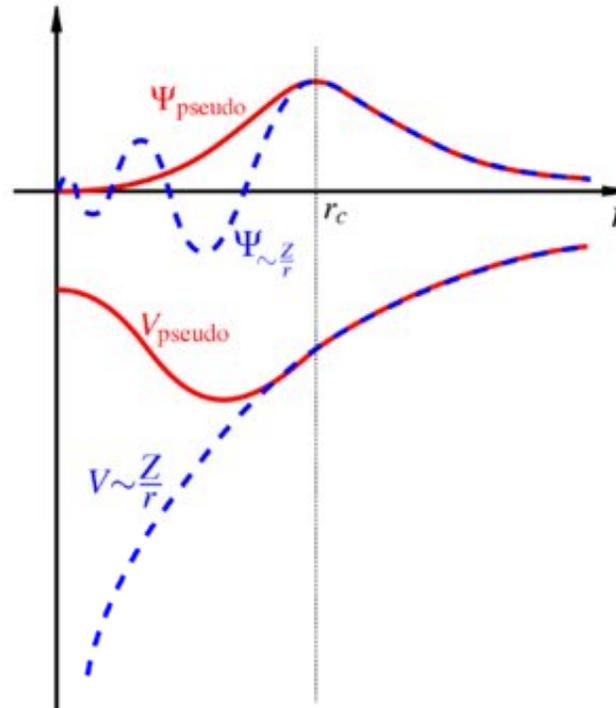
basically, geometry is less sensitive to basis set than molecular property: smaller basis set for geom. optimization, larger basis set for property (frequency, etc) computation.

For finite basis sets, like GTO, check for basis sets convergence.

- **Plane-wave basis set: ?**

# Pseudopotential I

The pseudopotential is an **effective potential** constructed to replace the atomic all-electron potential such that core states are eliminated and the valence electrons are described by nodeless pseudo-wavefunctions. In this approach **only the chemically active valence electrons are dealt with explicitly, while the core electrons are 'frozen'**.



Comparison of a wavefunction in the Coulomb potential of the nucleus (blue) to the one in the pseudopotential (red). The real and the pseudo wavefunction and potentials match above a certain cutoff radius  $r_c$ .

# Pseudopotential II

- **Types, related to basis set**
  - 1) full potential: all electron basis sets
  - 2) small core pseudopotential: valence + next to valence basis sets
  - 3) large core pseudopotential: valence electron only basis sets
- **Plan-wave pseudopotential: PSPW**
- **Usage**

when core electron is not important in the properties to be computed, save computing time

In GTO basis sets

- pseudopotential for middle size elements
- all electrons for second-and third-rows elements
- all electrons for heavy elements where relativistic (such as spin-orbit) effects are important (e.g., NMR chemical shift).

# Applications

- Geometry optimization
- Total Energy
- Reaction path, potential energy surface
- Predicting and Interpreting Spectra  
NMR, IR and Roman, Optical Spectra
- Thermochemistry, Photochemistry and Excited States
- Other Molecular properties

# List of electronic structure theory software I

- In Quantum Chemistry

Package	License	Basis	Periodic*	Mol. Mech.	HF	Post-HF	DFT
<b>Gaussian</b>	comm.	GTO	any	Y	Y	Y	Y
<b>NWChem</b>	acad	GTO, PW	any	Y	Y	Y	Y
<b>ADF/ BAND</b>	comm	STO	any	Y	Y	N	Y
<b>Firefly/ Gamess</b>	acad	GTO	N	Y	Y	Y	Y
<b>MolPro</b>	comm.	GTO	N	N	Y	Y	Y
<b>Q-Chem</b>	comm.	GTO	N	n	Y	Y	Y

\*Support for periodic systems (3d-crystals, 2d-slabs, 1d-rods and isolated molecules)  
 acad: free for academic, license agreement is required. comm.: commercial

# List of electronic structure theory software II

- DFT PW-based, for condense matter physics and chemistry

Package	License	Basis	Periodic*	Mol. Mech.	HF	Post-HF	DFT
<b>Abinit</b>	GPL	PW	3d	Y	N	N	Y
<b>PWSCF/ Espresso</b>	GPL	PW	3d	N	Y	N	Y
<b>DACAPO/ ASE</b>	GPL	PW	3d	Y	N	N	Y
<b>Wien2k</b>	Comm.	PW	3d	Y	N	N	Y
<b>VASP</b>	Comm.	PW	Any	Y	N	N	Y

[http://en.wikipedia.org/wiki/List\\_of\\_quantum\\_chemistry\\_and\\_solid\\_state\\_physics\\_software](http://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid_state_physics_software)

# Gaussian

- The most popular software in computational chemistry and research since 1980
- licensed (academic site license, commercial license, etc.)
- Current version: Gaussian 2009  
Gaussian 09 Features at a Glance  
[http://www.gaussian.com/g\\_prod/g09\\_glance.htm](http://www.gaussian.com/g_prod/g09_glance.htm)
- Parallel solution: OpenMP-SMP (scale to 16 cpus for DFT and 8 cpus for HF and post-HF methods)  
[distributed version via Linda (not available on SHARCNET)]
- SHARCNET license has G09-A.02 installed on bull and goblin.  
Campus site licenses run on the local systems.
- Benchmark tests shows g09 is ~2x faster than g03  
[https://www.sharcnet.ca/~jemmyhu/Gaussian\\_09\\_Benchmarks.pdf](https://www.sharcnet.ca/~jemmyhu/Gaussian_09_Benchmarks.pdf)

# Gaussian: run job at SHARCNET

- 1)        executable: g09                    - installed
- run script: g09                 - ready for user
- Input file: \*.com

## Job submit

Submit command on **bull** is

```
sqsub -q gaussian -n 4 -r 7.0d -o CO-6311-4.out g09 CO-6311-4.com
```

Submit command on **goblin** is

```
sqsub -q threaded -n 4 -r 7.0d -o CO-6311-4.out g09 CO-6311-4.com
```

Submit a serial job (%nproc=1):

```
sqsub -q serial -r 7.0d -o CO-6311-1.out g09 CO-6311-1.com
```

Gaussian output file is **CO-6311-4.log**,

CO-6311-4.out is an output file for error or LSF job message

# G09: simple input file example

```
%rwf=/scratch/username/CO-6311-4
```

```
%nosave
```

```
%mem=3200MB
```

```
%chk=CO-6311-4
```

```
%nproc=4
```

```
#p b3lyp/6-311++g(2df,p) opt freq
```

```
<< CO-6311++g(2df,p)—B3LYP-Opt-Freq >>
```

```
0 1
```

```
c o 1 co
```

```
co=1.1281
```

# G09: multi-step job in a single input file

```
%chk=h2o_link  
%nproc=4  
#p B3LYP/6-31G(d,p) opt freq
```

Gaussian Test B3LYP opt freq

```
0 1  
h  
o 1 d  
h 2 d 1 hoh
```

```
d      0.9815  
hoh    102.3095
```

--Link1--

```
%chk=h2o_link  
%nproc=4  
#p mp2/6-31G(d,p) guess=read geom=check opt
```

Gaussian Test MP2 opt based on B3LYP results

```
0 1
```

# Gaussian: Checkpointing

1. .chk can be used to restart an Optimization job, or used for property (freq, etc) calculation afterwards.
2. In G09, some frequency calculation can be restarted, please refer to G09 user guides for details.
3. Normally, the run time per job is 7 days (1 week) on SHARCNET clusters.  
In case some properties may be one step process which cannot be restarted from .chk file, you can use our [BLCR](#) to do checkpointing/restart.

Details in: <https://www.sharcnet.ca/my/software/show/21>

GaussView: GUI for Gaussian

Free GUI: Molden, Jmol, WebMO

# NWChem

- Developed at **EMSL**: Environmental Molecular Sciences Laboratory at Pacific Northwest National Laboratory
- Free licensed for academic, need to sign a license agreement
- It aims to be scalable both in its ability to treat large problems efficiently and in its usage of available parallel computing resources
- Parallel solution: Global Arrays
- Current version: 5.1.1
- **Capabilities**  
<http://www.emsl.pnl.gov/capabilities/computing/nwchem/capabilities.jsp>

# NWChem: input file structure

- Input is composed of directives to define data (such as basis set, geometries, etc) and actions (task)
- Free format, default extension as .nw (name.nw)
- The start-up directives are
  - START
  - RESTART
  - SCRATCH\_DIR
  - PERMANENT\_DIR
  - MEMORY
  - ECHO

<http://www.emsl.pnl.gov/capabilities/computing/nwchem/docs/user/user.html>

## Input-1: Single point SCF energy

```
start h2o

title "Water in 6-31g basis set"

geometry units au
  O 0.00000000 0.00000000 0.00000000
  H 0.00000000 1.43042809 -1.10715266
  H 0.00000000 -1.43042809 -1.10715266
end

basis
  H library 6-31g
  O library 6-31g
end

task scf
```

## Input-2: Restart and perform a geom opt

```
restart h2o

title "Water geometry optimization"

task scf optimize
```

## Input-3: multi tasks, MP2 opt and CCSD(T) on nitrogen

```
start n2
```

```
geometry
```

```
  symmetry d2h
```

```
  n 0 0 0.542
```

```
end
```

```
basis spherical
```

```
  n library cc-pvtz
```

```
end
```

```
mp2
```

```
  freeze core
```

```
end
```

```
task mp2 optimize
```

```
ccsd
```

```
  freeze core
```

```
end
```

```
task ccsd(t)
```

lots of input examples in

`/opt/sharcnet/nwchem/5.1.1/QA/tests`

# NWChem: run job at SHARCNET

- Things to prepare:

1) Create a .nwchemrc file in your /home

```
nwchem_basis_library /opt/sharcnet/nwchem/current/src/basis/libraries/  
nwchem_nwppw_library /opt/sharcnet/nwchem/current/src/nwppw/libraryps/  
ffield amber
```

.....

2) prepare input file: \*.nw

- sqsub command:

```
sqsub -q mpi -n 4 -o h2o_hu.log nwchem h2o_hu.nw
```

- tests on saw in

/work/jemmyhu/nwchem\_tests

SHARCNET NWChem software page:

<https://www.sharcnet.ca/my/software/show/12>

# Abinit

- Free open source software under GNU General Public License
- A package whose main program allows one to find the total energy, charge density and electronic structure of systems made of electrons and nuclei (molecules and periodic solids) within Density Functional Theory (DFT), using pseudopotentials and a plane-wave basis
- Options to
  - optimize the geometry according to the DFT forces and stresses
  - perform molecular dynamics simulations using these forces
  - generate dynamical matrices, Born effective charges, and dielectric tensors
  - excited states can be computed within the Time-Dependent DFT (for molecules), or within Many-Body Perturbation Theory (the GW approximation)
  - different utility programs for post-processes
- Features described in <http://www.abinit.org/about/presentation.pdf>
- Current version: abinit-6.0.4

# Abinit: things needed to run

- To run abinit you need four things:
  - (1) Access to the executable, abinit
  - (2) An input file.
  - (3) A files file (list of file names in a file).
  - (4) A pseudopotential input file for each kind of element in the unit cell.

With these items a job can be run. The full list of input variables, all of which are provided in the single input file, is given in

[http://www.abinit.org/documentation/helpfiles/for-v6.0/users/new\\_user\\_guide.html](http://www.abinit.org/documentation/helpfiles/for-v6.0/users/new_user_guide.html)

# ABINIT: a files file

The files file, called **t3x.files**, could look like:

```
t31.in          - input file (sample in /work/jemmyhu/abinit/tests on saw)
t3x.out        - output file
t3xi           - root name for input wavefunctions if any
t3xo           - root name for output wavefunctions, density, etc
t3x            - root name for tmp files
/opt/sharcnet/abinit/6.0.4/share/abinit/Psps_for_tests/14si.pspnc
```

Run command on SHARCNET clusters:

```
parallel job: sqsub -q mpi -n 4 -r 1.0h -o t31.log -i t3x.files abinit
```

Example to compute total energy of Crystalline silicon on saw at  
/work/jemmyhu/abinit/tests

SHARCNET ABINIT software usage page:

<https://www.sharcnet.ca/my/software/show/60>

# Abinit: links

- Abinit FAQ

<http://www.abinit.org/support/help-center/faq>

- New User Guide

[http://www.abinit.org/documentation/helpfiles/for-v6.0/users/new\\_user\\_guide.html](http://www.abinit.org/documentation/helpfiles/for-v6.0/users/new_user_guide.html)

- Abinit tutorials

<http://www.abinit.org/documentation/helpfiles/for-v6.0/tutorial/welcome.html>

- Forum

<http://forum.abinit.org/>

- Website

<http://www.abinit.org/>

# Quantum Espresso/PWSCF

- Description:

Quantum ESPRESSO is an integrated suite of computer codes for electronic-structure calculations and materials modeling at the nanoscale. It is based on density-functional theory, plane waves, and pseudopotentials (both norm-conserving and ultrasoft)

- Capabilities:

<http://www.pwscf.org/whatcangedo.php>

- Free open source software under GNU General Public License

- Tools:

PWGui – prepare input data

XCrySDen – visualization of results

- Current version: 4.2

- SHARCNET version: 4.0.4

# PWSCF: run job at SHARCNET

- Things to do:
  - 1) run script: run\_pwscf\_XC sample for user
  - 2) input file
  - 3) sqsub command

```
sqsub -q mpi -n 4 --nompirun -o run_pwscf_XC.log ./run_pwscf_XC
```

- Examples in  
/home/jemmyhu/tests/test\_mpiPWSCF/on-XC
- SHARCNET PWSCF software page  
<https://www.sharcnet.ca/my/software/show/42>

# Others: Software for Molecular Dynamics

Available at SHARCNET for public users:

- AMBER
- Gromacs
- NAMD
- CPMD

User owned:

- CHARMM
- Cosmos
- MOE

[http://en.wikipedia.org/wiki/List\\_of\\_software\\_for\\_molecular\\_mechanics\\_modeling](http://en.wikipedia.org/wiki/List_of_software_for_molecular_mechanics_modeling)

# General policy for software

- Software requested by 2 or more groups – add to SHARCNET software stack with software usage instruction/example webpage online

<https://www.sharcnet.ca/my/software>

- Software requested by one group – user installs in his/her own space with help from SHARCNET if necessary
- SHARCNET owned commercial software – free
- Self owned software – user's response for license issues

Others?

Questions for SHARCNET?

Thank you.