Introduction to Density Functional Theory in a Plane Wave + Pseudopotential Framework

Shobhana Narasimhan
Theoretical Sciences Unit
JNCASR, Bangalore
shobhana@jncasr.ac.in

School- “From the chemical bond to the chemical plant: Challenges in gas conversion technology”
JNCASR, Bangalore, September 1, 2014
What is DFT?

- A very successful approach to calculating the (ground state) properties of many-electron systems \textit{from first principles}

- \textbf{Features:}
  - \textit{Ab initio} (only I/P = Z, A)
  - Quantum Mechanical
  - Numerical
  - Approximate (could be exact, in principle?)
Nice things about DFT

• *Ab initio*, no empirical information

• Predictive power

• Very *accurate* for most materials (errors typically < few %)

• Can calculate *lots of properties* of scientific / technological interest

• Nobel Prize to Walter Kohn, 1998.
Not-so-nice things about DFT

- Computationally intensive.
- Therefore (currently) limited to small length / time scales.
- For computational reasons, introduce further approximations, might not always work....
- Not very successful for strongly-correlated systems (e.g., high-\(T_c\) superconductors).
Outline of talk

• **Formalism**
  - Hohenberg-Kohn theorem
  - Kohn-Sham equations

• **How to solve the problem**
  - Basis sets, plane wave basis sets
  - Pseudopotentials
  - Self consistent loop
  - Brillouin zone sums & smearing

• **What one gets out of calculations**
I. How to set up the problem
The Basic Problem

• Many-electron system:

\[ N_n \text{ nuclei } \{Z_i, R_i\} + N \text{ electrons } \{r_j\} \]

• Interact via Coulomb interaction

• Examples: atoms, molecules, condensed matter systems.

• Want to solve Many-electron Schrödinger Equation:

\[ H\Psi(r_1, r_2, \ldots r_N) = E \Psi(r_1, r_2, \ldots r_N) \]

– Solutions must also be properly antisymmetrized
(to satisfy Pauli Exclusion Principle)

• Don’t know how to solve!!
How to deal with many-particle Schrödinger Eq.??

Many-particle Schrödinger Equation

Many-body Theory
• Retain many-body nature
• Use model Hamiltonian (e.g., Ising model, Hubbard model)
• Use suitable parametrization
• Solve numerically / analytically

Density Functional Theory
• Map onto 1-particle Schrödinger Eq.
• Exact, in principle
• Approximate, in practice
• No free parameters
• Solve numerically
Hohenberg-Kohn Theorem - I

- [Phys. Rev. 136 B 864 (1964)]
- **Density as basic variable:**
  - Collection of particles in external potential \( v(r) \)
  - Usually:
    - Solve many-particle Schrödinger Eq.
    - \( \nu(r) \rightarrow \Psi_{\text{many-particle}} \rightarrow n(r) = \Psi^*\Psi \)
  - Hohenberg & Kohn showed (simple 3-line proof!)
    - \( n(r) \rightarrow 1-1 \text{ mapping exists} \rightarrow \nu(r) \rightarrow \text{all properties of system} \)
Variational Formulations

• When solving various complicated problems, it might be possible to formulate the problem in terms of the calculus of variations (e.g., Fermat’s principle in optics, Rayleigh-Ritz variational principle in Quantum Mechanics).

• Typically, one guesses a solution of a given form (ansatz), with some free parameters that have to be determined.

• The value of the parameters is determined by minimizing some quantity.

• There are sophisticated numerical techniques available for minimization problems.

• Solution may be exact / approximate (depending on ansatz).
Hohenberg-Kohn Theorem - II

- **Variational Principle:**
  There exists a universal functional $F[n]$, such that:
  - $E_v[n] ≡ F[n] + \int \psi(r) n(r) \, dr$ has its minimum value when $n(r) = n_{GS}(r) →$ density of system in ground state
  - $E_v[n = n_{GS}] = E_{tot}^{GS} →$ total energy in ground state

- Note: only proof of existence of $F[n]$; we don’t know form.

- **If** $F[n]$ were known...
  … then have converted many-electron pbm. to one of minimizing a functional $→ n_{GS}(r), E_{tot}^{GS}$.
What have we achieved? - I

$N$-electron Schrödinger Equation
Solve for $\Psi_{GS}(r_1, r_2, \ldots, r_N)$

maps exactly if $F[n]$ known

Variational Principle
Minimize $E_v[n]$ to find $n_{GS}, E_{GS}$

$d n(r) \left\{ \frac{\delta E_v[n(r)]}{\delta n(r)} \right\} dr = 0$

Shobhana Narasimhan, JNCASR
What would we want now?

• Approximation for $F[n]$

• Ansatz for form of $n(r)$

• Good prescription for minimizing $E_v[n]$ wrt variations in $n(r)$

  OR an alternative to this.
$F[n]$: non-interacting & interacting systems

• Recall:

$$E_v[n] = F[n] + \int \nu(r) \ n(r) \ dr$$

• For a non-interacting system,

$$F[n] = T_0[n]$$

• For an interacting system,

$$F[n] = T_0[n] + \frac{1}{2} \left( \iiint n(r) n(r') \ dr \ dr' \right) + E_{xc}[n]$$

External potential (e.g., due to nuclei)

Hartree [Coulomb]

Exchange-Correlation
$F[n]$: non-interacting & interacting systems

- Recall:
  
  \[ E_v[n] = F[n] + \int v(r) n(r) \, dr \]

- For a non-interacting system,
  
  \[ F[n] = T_0[n] \]

- For an interacting system,
  
  \[ F[n] = T_0[n] + \frac{1}{2} \int \int \frac{n(r) n(r')}{|r-r'|} \, dr \, dr' + E_{xc}[n] \]

External potential (e.g., due to nuclei)
Hartree potential

\[ V_H(r) = \int \frac{n(r')}{|r-r'|} \, dr' \]

Hartree [Coulomb]
Exchange-Correlation
Kohn-Sham Equations

- [Phys. Rev. 140 A 1133 (1965)]

<table>
<thead>
<tr>
<th>Using H-K theorem</th>
<th>Non-interacting system</th>
</tr>
</thead>
</table>
| \[ \int_\delta n(r) \left\{ \frac{\delta T_0}{\delta n} + u(r) \right\} dr = 0 \] | }
Kohn-Sham Equations

- [Phys. Rev. 140 A 1133 (1965)]

<table>
<thead>
<tr>
<th>Using H-K theorem</th>
<th>Non-interacting system</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \int \delta n(r) \left{ \frac{\delta T_0}{\delta n} + \nu(r) \right} , dr = 0 )</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Solving Schrödinger Eq.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( [-\frac{1}{2} \nabla^2 + \nu(r)] \psi_i(r) = \varepsilon_i \psi_i(r) )</td>
<td></td>
</tr>
</tbody>
</table>
### Kohn-Sham Equations

- [Phys. Rev. 140 A 1133 (1965)]

<table>
<thead>
<tr>
<th>Using H-K theorem</th>
<th>Non-interacting system</th>
<th>Interacting system</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \int \delta n(r) \left{ \frac{\delta T_0}{\delta n} + \nu(r) \right} , dr = 0 )</td>
<td>( \int \delta n(r) \left{ \frac{\delta T_0}{\delta n} + V_H(r) + \frac{\delta E_{xc}}{\delta n} + \nu(r) \right} , dr = 0 )</td>
<td></td>
</tr>
<tr>
<td>Solving Schrödinger Eq.</td>
<td>( \left[ -\frac{1}{2} \nabla^2 + \nu(r) \right] \psi_i(r) = \varepsilon_i \psi_i(r) )</td>
<td></td>
</tr>
</tbody>
</table>

Shobhana Narasimhan, JNCASR
## Kohn-Sham Equations

- [Phys. Rev. 140 A 1133 (1965)]

<table>
<thead>
<tr>
<th></th>
<th>Non-interacting system</th>
<th>Interacting system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using H-K theorem</td>
<td>( \int \delta n(r) \left{ \frac{\delta T_0}{\delta n} + \nu(r) \right} dr = 0 )</td>
<td>( \int \delta n(r) \left{ \frac{\delta T_0}{\delta n} + V_H(r) + \frac{\delta E_{xc}}{\delta n} + \nu(r) \right} dr = 0 )</td>
</tr>
<tr>
<td>Solving Schrödinger Eq.</td>
<td>([-\frac{1}{2} \nabla^2 + \nu(r)] \psi_i(r) = \varepsilon_i \psi_i(r))</td>
<td>([-\frac{1}{2} \nabla^2 + \nu_{eff}(r)] \psi_i(r) = \varepsilon_i \psi_i(r))</td>
</tr>
</tbody>
</table>
Kohn-Sham Equations

- [Phys. Rev. 140 A 1133 (1965)]

<table>
<thead>
<tr>
<th></th>
<th>Non-interacting system</th>
<th>Interacting system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Using H-K theorem</strong></td>
<td>( \int \delta n(r) \left{ \frac{\delta T_0}{\delta n} + \nu(r) \right} , dr = 0 )</td>
<td>( \int \delta n(r) \left{ \frac{\delta T_0}{\delta n} + V_H(r) + \frac{\delta E_{xc}}{\delta n} + \nu(r) \right} , dr = 0 )</td>
</tr>
<tr>
<td><strong>Solving Schrödinger Eq.</strong></td>
<td>( [-\frac{1}{2} \nabla^2 + \nu(r)] \psi_i(r) = \varepsilon_i \psi_i(r) )</td>
<td>Kohn-Sham equations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( [-\frac{1}{2} \nabla^2 + \nu_{\text{eff}}(r)] \psi_i(r) = \varepsilon_i \psi_i(r) )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{(1-electron eqns. for interacting system!)} )</td>
</tr>
</tbody>
</table>
What have we achieved? - II

N-electron Schrödinger Equation
Solve for $\Psi_{\text{GS}}(r_1, r_2, \ldots, r_N)$

maps exactly if $F[n]$ known

Variational Principle
Minimize $E_v[n]$ to find $n_{\text{GS}}, E_{\text{GS}}$

if $E_{xc}$ known

Effective one-electron equations
Solve to get $n_{\text{GS}}, E_{\text{GS}}$
What have we achieved? - III
What have we achieved? - III

Shobhana Narasimhan, JNCASR
II. How to Solve the Problem
Determining K-S Hamiltonian

Want to solve:

\[
\left[ -\frac{1}{2} \nabla^2 + v_{\text{eff}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r})
\]

where

\[ v_{\text{eff}}(\mathbf{r}) = V_H \{ n(\mathbf{r}) \} + V_{xc} \{ n(\mathbf{r}) \} + v(\mathbf{r}) \]

- Problem 1: $V_H, V_{xc} \leftarrow n \leftarrow$ solution of problem!
- Problem 2: Form of $V_{xc}[n]$ not known….
- Convention: $e = \hbar = m_e = 1$
Kohn-Sham Equations in a Basis

- Can choose to expand wavefunctions in a basis set:

\[ \psi_i(r) = \sum_{\alpha=1}^{N_b} c_{i\alpha} f_\alpha(r) \]

- Eigenvalue equation then becomes:

\[ \sum_{\beta} H_{\alpha\beta} c_{i\beta} = \varepsilon_i c_{i\alpha} \]

- Solving \( \Leftrightarrow \) Have to diagonalize a matrix of size \( N_b \times N_b \)

\[ \begin{align*}
\text{Matrix element} & \quad \text{Eigenvalue} \\
\text{Eigenvector} & \quad \text{Size of basis}
\end{align*} \]
Advantages of a Plane Wave Basis

- **Simple**: Easy to take derivatives, etc. ⇒ Coding is easy!

- **Orthonormal**: No overlap integrals.

- **Independent of atomic positions** ⇒ No “Pulay forces”; easy to calculate forces for structural relaxation & molecular dynamics.

- **Unbiased**: No assumption about where charge concentrated. (But ∴ also wasteful?)

- **Easy to control convergence** w.r.t. size of basis: only one parameter $E_{cut}$.

- **Can take advantage of FFT’s**: r-space ↔ k-space
Disadvantages of a Plane Wave Basis

- The set of plane waves is discrete only if the system is periodic!
  (Will discuss…solution = introduction of artificial supercell or periodic approximant.)

- Recall:
  - for free electrons, wavefunction = plane wave.
  - for nearly free electrons, wavefunction = superposition of small number of plane waves.
  - for tightly bound electrons, need a HUGE number of plane waves to get an adequate expansion, i.e., $N_b$ very large!
  (Will discuss…solution = introduction of pseudopotentials.)

- Sometimes interpretation harder.
Periodic Systems & Bloch’s Theorem

• For a periodic system, recall Bloch’s Theorem:

\[ \psi_k(r) = e^{i k \cdot r} u_k(r) \]

• \( u_k(r) \) has the periodicity of the system, i.e.,

\[ u_k(r) = u_k(r + R), \quad \text{where } R = \text{lattice vector} \]

• As for all lattice-periodic functions, only certain plane waves will appear in the Fourier expansion of \( u_k(r) \):

\[ u_k(r) = \frac{1}{\Omega} \sum_G c_{k,G} e^{iG \cdot r} \quad \text{where } G = \text{reciprocal lattice vector} \]
Plane Waves & Periodic Systems

• For a periodic system:

\[ \psi_k(r) = \frac{1}{\Omega} \sum_G c_{k,G} e^{i(k+G) \cdot r} \]

where \( G = \) reciprocal lattice vector

• The plane waves that appear in this expansion can be represented as a grid in k-space:

• Only true for periodic systems that grid is discrete.

• In principle, still need infinite number of plane waves.
Truncating the Plane Wave Expansion

- In practice, the contribution from higher Fourier components (large $|k+G|$) is small.
- So truncate the expansion at some value of $|k+G|$.
- Traditional to express this cut-off in energy units:

$$\frac{\hbar^2 |k + G|^2}{2m} \leq E_{\text{cut}}$$

Have to check convergence wrt $E_{\text{cut}}$.
How to Specify the System

- All periodic systems can be specified by a Bravais Lattice and an atomic basis.
What if the system is not periodic?

- **Example 1**: Want to study properties of a system with a surface.
- Presence of surface $\Rightarrow$ No periodicity along $z$.
- Use a supercell: artificial periodicity along $z$ by repeating slabs separated by vacuum.
- Have to check convergence w.r.t. slab thickness & vacuum thickness.
What if the system is not periodic?

- **Example 2**: Want to study properties of a nanowire.

- **Example 3**: Want to study properties of a cluster
What if the system is not periodic?

• **Example 2**: Want to study properties of a nanowire ⇒ introduce artificial periodicity along \( y \) & \( z \).

• **Example 3**: Want to study properties of a cluster ⇒ introduce artificial periodicity along \( x, y \) & \( z \).
What if the system is not periodic?

- **Example 4**: Want to study a system with a defect, e.g., a *vacancy* or *impurity*:
What if the system is not periodic?

- **Example 4:** Want to study a system with a defect, e.g., a *vacancy* or *impurity*:
What if the system is not periodic?

- **Example 5**: Want to study an amorphous or quasicrystalline system.
What if the system is not periodic?

- **Example 5**: Want to study an amorphous or quasicrystalline system: approximate by a periodic system (with large unit cell).
Artificially Periodic Systems ⇒ Large Unit Cells

- **Note**: In all these cases, to minimize the effects of the artificially introduced periodicity, need a large unit cell.

  ↓

- Long \( \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \) (primitive lattice vectors)

  ↓

- Short \( \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \) (primitive reciprocal lattice vectors)

  ↓

- Many \( \mathbf{G} \)'s will fall within \( E_{cut} \) sphere!
Using Supercells Increases Computation Time

Calculation takes longer when unit cell size increased:

**REAL SPACE**

- Unit cell expands

**RECIPIROCAL SPACE**

- 1st Brillouin zone shrinks

Shobhana Narasimhan, JNCASR
Self-Consistent loop

1. Generate new $n(r)$
2. Calculate $V_{H}[n]$ & $V_{XC}[n]$
3. $\psi_{i}(r) = \frac{-1/2 \nabla^2 + v_{eff}(r)}{\varepsilon_{i}} \psi_{i}(r)$
4. Calculate new $n(r) = \sum_{i} |\psi_{i}(r)|^2$
5. Self-consistent?
6. Yes, Problem solved! Can now calculate energy, forces, etc.
7. No, Generate new $n(r)$

Initial guess $n(r)$

$V_{mic}(r)$ known/constructed
Step 1: Obtaining $V_{nuc}$

1. $V_{nuc}$ known/constructed
2. Initial guess $n(r)$
3. Calculate $V_H[n]$ & $V_{XC}[n]$
4. $V_{eff}(r) = V_{nuc}(r) + V_H(r) + V_{XC}(r)$
5. $H\psi_i(r) = [-1/2\nabla^2 + V_{eff}(r)]\psi_i(r) = \epsilon_i \psi_i(r)$
6. Calculate new $n(r) = \Sigma_i |\psi_i(r)|$
7. Self-consistent?
   - Yes: Problem solved! Can now calculate energy, forces, etc.
   - No: Generate new $n(r)$ and repeat from step 2.
Nuclear Potential

• Electrons experience a Coulomb potential due to the nuclei.

• This has a known and simple form:

\[ V_{nuc} = -\frac{Z}{r} \]

• But this leads to computational problems!
Electrons in Atoms

- Electrons in atoms are arranged in shells.
- Quantum numbers:
  - $n$ [principal], $l$ [angular], $m_l$ [magnetic], $m_s$ [spin]
- Rare gas atoms have certain complete subshells (inert configurations):
  - He: $1s^2$  Ne: $[\text{He}], 2s^2, 2p^6$  Ar: $[\text{Ne}] 3s^2, 3p^6$
  - Kr: $[\text{Ar}], 3d^{10}, 4s^2, 4p^6$  Xe: $[\text{Kr}], 4d^{10}, 5s^2, 5p^6$
  - Rn: $[\text{Xe}], 4f^{14}, 5d^{10}, 6s^2, 6p^6$
- Can divide electrons in any atom into core and valence.  
- This division is not always clear-cut, but usually core = rare gas configuration [+ filled d/f subshells]
**Atomic Wavefunctions**

- For hydrogenic atoms, recall:

\[
\psi_{lm}(r) = \psi_l(r) Y_{lm}(\theta, \phi) = r^{-l} \phi_l(r) Y_{lm}(\theta, \phi)
\]

- **Radial part & Angular Part.**

- Being eigenfunctions of a Hermitian operator, \( \psi_{lm} \)'s are orthonormal.

- Wavefunctions with same \( n \), different \( l \) are orthogonal due to the nature of the angular part of the wavefunction.

- Wavefunctions with different \( n \), same \( l \) are orthogonal due to the nature of the radial part of the wavefunction.
Example: Wavefunctions for Ag atom

Ground state configuration: [Kr], 4d^{10}, 5s^{1}, 5p^{0}, 5d^{0}

- Core wavefunctions sharply peaked near nucleus.
- Valence wavefunctions peaked far away from nucleus, lots of wiggles near nucleus.
- 1s, 2p, 3d, 4f,... nodeless.
- Not clear whether 4d should be considered core / valence.
Electrons in molecules/solids

• Chemical bonds between atoms are formed by sharing / transferring electrons.

• Only the valence electrons participate in bonding.

• Wavefunctions of valence electrons can change significantly once the bond is formed.

• e.g., when Ag is a constituent of a solid, the wavefunction may also acquire some $5p$ or $5d$ character?

• Wavefunctions of core electrons change only slightly when the bond is formed.
Problem for Plane-Wave Basis

Core wavefunctions: sharply peaked near nucleus.

Valence wavefunctions: lots of wiggles near nucleus.

High Fourier components present

i.e., need large $E_{\text{cut}}$ 😞
Core wavefunctions: sharply peaked near nucleus.

Valence wavefunctions: lots of wiggles near nucleus.

High Fourier components present
i.e., need large $E_{cut}$ 😞

Don’t solve for the core electrons!

Remove wiggles from valence electrons.
The Pseudopotential Approximation

• **Frozen core**: remove core-electron degrees of freedom i.e., NOT an “All-electron” calculation.

• Valence electrons see a weaker potential than the full Coulomb potential.

\[ V_{\text{nuc}}(r) \rightarrow V_{\text{ion}}(r) \]

• **Further tailor this potential** so that wavefunctions behave ‘properly’ in region of interest, yet computationally cheap.
How the Pseudopotential Helps

(Numerical) Advantages when solving Kohn-Sham eqns.: 
• When solving using a basis (especially plane waves), basis size drastically reduced (smaller matrices to diagonalize).
• Have to solve for fewer eigenvalues.
• No Coulomb singularity (cusp in wavefunction) at origin.

Disadvantages:
• Can lose accuracy.
An analogy!

- “Dummy cops” used by many law-enforcement agencies!
- Don’t care about internal structure as long as it works right!
- But cheaper!!
- Obviously it can’t reproduce all the functions of a real cop, but should be convincing enough to produce desired results....
Hey, we have them in Bangalore, too!
Wish List for a Good Pseudopotential

For accuracy:

• Should reproduce scattering properties of true potential.
• Transferable: Nice to have one pseudopotential per element, to use in variety of chemical environments.
• Norm conserving? (will explain)
• Ab initio? (no fitting to experimental data)

For (computational) cheapness:

• Smooth / Soft: Need smaller basis set (esp. plane waves)
• ‘Separable’? (will skip!) but ‘Ghost free’ (should not introduce spurious states when making separable!)
Generating an *ab initio* pseudopotential
(Note: general outline, schemes differ!)

1) Pick electronic configuration for atom (reference config.)
   [e.g., may want to promote some electrons to excited states]

2) Perform all-electron calculation → \( \phi_{nl}^{AE}(r) \), \( \epsilon_{nl}^{AE} \)

\[
-\frac{1}{2} \frac{d^2}{dr^2} \phi_{nl}^{AE}(r) + \left[ \frac{l(l+1)}{2r^2} + V_{eff}(r) - \epsilon_{nl}^{AE} \right] \phi_{nl}^{AE}(r) = 0
\]

where

\[
V_{eff} = -\frac{Z}{r} + V_H[\rho; r] + V_{XC}[\rho; r]
\]

(contd.)
3) Divide electrons into core and valence.

4) Pick a core radius $r_c$
   - $r_c$ too small $\rightarrow$ hard pseudopotential
   - $r_c$ too large $\rightarrow$ transferability poor
   - $r_c$ should be large enough to avoid overlapping cores
   - $r_c$ can be different for each $l$
   - $r_c$ should be larger than $r$ for outer-most node of radial wavefunction

(contd.)
Generating an \textit{ab initio} pseudopotential

(\textit{contd.})

5) Construct pseudowavefunction (one $l$ at a time):
- Pseudowavefunction & all-electron wavefunction are identical outside the cut-off radius $r_c$: 

\begin{center}
\includegraphics[width=0.5\textwidth]{diagram.png}
\end{center}
5) Construct pseudowavefunction \((one \ l \ at \ a \ time)\):
- Pseudowavefunction & all-electron wavefunction are identical outside the cut-off radius \(r_c\):

\[
\phi_{l,\text{ref}}^{AE}(r) = \phi_{l,\text{ref}}^{PS}(r) \quad r \geq r_c
\]
5) Construct \textit{pseudowavefunction (one } l \textit{ at a time)}:
- Pseudowavefunction \& all-electron wavefunction are identical outside the \textit{cut-off radius } r_c:\n\begin{align*}
\text{Inside } r_c \quad \phi_{l, \text{ref}}^{PS}(r) &= f(r) \\
\text{Lots of freedom for choice of } f (\text{choose for right log derivatives, softness, norm conservation, etc.})
\end{align*}
Generating an *ab initio* pseudopotential

(continued)

6) Invert Schrödinger equation:

\[
V_l^{scr}(r) = \varepsilon_l - \frac{l(l+1)}{2r^2} + \frac{1}{2\phi_l(r)} \frac{d^2[\phi_l(r)]}{dr^2}
\]

- Will get correct (all-electron) eigenvalue.
- “Screened” pseudopotential
  (includes Hartree + XC potentials)

7) “Unscreen”, i.e., remove Hartree and XC contributions.

\[
V_l^{PS}(r) = V_l^{scr}(r) - V_H[\rho^{val}(r)] - V_{XC}[\rho^{val}(r)]
\]
What does a pseudopotential look like?

Example for Mo:

- Weaker than full Coulomb potential
- No singularity at $r=0$
- Different pseudopotential for each $l$ (example of “semilocal” pseudopotential)
- Will be $V_{ion}$ (replacing nuclear potential)

Transferability

• Condition that pseudoatom reproduces behavior of all-electron atom in a wide variety of chemical environments.

Recall, pseudopotential derived for reference configuration (atom with a given occupation of levels), using a reference eigenvalue.

When eigenvalue changes from reference one:
• do scattering properties of potential change correctly? (Look at log derivatives)

When the filling changes:
• - do eigenvalues shift correctly? (look at “chemical hardness”)
• - do scattering properties change correctly?
Scattering

Recall (from a quantum mechanics course?):

- Scattering properties of a potential described by phase shift $\eta_l$.
- Related to logarithmic derivatives: [see, e.g. Eq. J.6, Martin]

$$D_l(\epsilon, r) = r \frac{d}{dr} \ln \psi_l(\epsilon, r) = r \frac{d}{dr} \ln \left( \frac{\phi_l(\epsilon, r)}{r} \right)$$

- Weaker potentials will have fewer bound states.
- In the pseudopotential approximation: want to make the potential weak enough that the valence electron is the lowest bound state (with that $l$), while reproducing log derivatives to the extent possible....
Log derivatives

Eric Walter
Norm Conservation

• By construction, log derivatives satisfy:
  \[ D_l^{AE}(\epsilon, r_c) = D_l^{PS}(\epsilon, r_c) \]

• In addition, if we impose norm conservation:
  \[ \int_0^{r_c} \phi^{*AE}(r)\phi^{AE}(r) dr = \int_0^{r_c} \phi^{*PS}(r)\phi^{PS}(r) dr \]
  then from the identity (see e.g. pg. 214 of Martin for derivation):
  \[ \frac{\partial}{\partial \epsilon} D_l(\epsilon, r_c) = -\frac{r_c}{|\phi_l(r_c)|^2} \int_0^{r_c} dr |\phi_l(r_c)|^2 \]
  we have*
  \[ \frac{\partial}{\partial \epsilon} D_l^{AE}(\epsilon, r_c) = \frac{\partial}{\partial \epsilon} D_l^{PS}(\epsilon, r_c) \]
  i.e., if energy is shifted slightly from that of reference eigenvalue, log derivatives ~ unchanged →
  improved transferability!

*Hamann, Schlüter, & Chiang, PRL 43, 1494 (1979)
Some Popular Pseudopotentials: BHS

- “Pseudopotentials that work: from H to Pu”
- Ab initio, norm conserving, so good transferability (?)
- Semilocal $V_l(r)$ [local in radial coordinates, nonlocal in angular coordinates]
- Parametrized form: chosen to give nice analytical expressions with many basis sets, 9 parameters, tabulated for all elements.
- Non-linear fitting procedure, caution needed!
- Fairly hard pseudopotentials since smoothness not built in explicitly, frequently need high cut-off.
Soft / Smooth Pseudopotentials

• Want to lower $E_{cut}$ (cut-off for plane wave basis).

• Various strategies:
  - Optimize so as to minimize error in KE introduced by truncating basis (Rappe, Rabe, Kaxiras & Joannopoulos, [RRKJ] 1990)
  - Make smooth near origin (Troullier & Martins, 1991)

• Cut-offs lowered considerably, but still higher than we would like, especially for
  > first row elements (1s, 2p nodeless)
  > transition metals (3d nodeless)
  > rare-earths (4f nodeless)
Need lower $E_{\text{cut}}$ with soft pseudopotentials

e.g. Cu: localized d orbitals $\rightarrow$
high cut-off needed with BHS pseudopotential

**Troullier-Martins**

**RRKJ**

---

**FIG. 8.** The calculated total energy of fcc Cu plotted against the cutoff energy of the plane-wave basis set for the four pseudopotentials shown in Fig. 7. The total energy for all four curves are referenced to the total energy calculated at a cutoff energy of 225 Ry. The squares, circles, and triangles are the calculated data points and the curves are obtained from a spline interpolation.

**FIG. 3.** Atomic (solid lines) and fcc solid (dots) total energies as a function of cutoff energy for copper in the HSC and present approaches. The zero of atomic total energy for each pseudopotential was chosen to be the total atomic energy at a cutoff energy of 324 Ry. The zero of solid total energy was chosen for each pseudopotential so that the atomic and solid total energies coincide at a cutoff energy of 80 Ry.
Ultrasoft Pseudopotentials

- Do away with norm conservation!!
- Can make $\psi^{PS}$ extremely soft!
- Drastically reduces $E_{cut}$, especially for “difficult” elements.
- New separable form.
- Choose multiple energy references (to improve transferability).

---

Shobhna Narasimhan, JNCASR

Laasonen, Car, Lee & Vanderbilt
Step 2: Initial Guess for $n(r)$

- $V_{\text{ion}}$ known/constructed
- Initial guess $n(r)$
  - Calculate $V_{H}[n]$ & $V_{XC}[n]$
    - $V_{\text{eff}}(r) = V_{\text{ion}}(r) + V_{H}(r) + V_{XC}(r)$
    - $H\psi_{i}(r) = [-1/2\nabla^{2} + V_{\text{eff}}(r)] \psi_{i}(r) = \varepsilon_{i} \psi_{i}(r)$
  - Calculate new $n(r) = \sum_{i} |\psi_{i}(r)|$
- Self-consistent?
  - Yes
    - Problem solved! Can now calculate energy, forces, etc.
  - No
    - Generate new $n(r)$

Initial Choice of $n(r)$

Various possible choices, e.g.,:

- Converged $n(r)$ from a closely related calculation (e.g., one where ionic positions slightly different).
- Superpositions of atomic densities.
- Approximate $n(r)$, e.g., from solving problem in a smaller/different basis.
- Random numbers.
- Some combination of the above....
Steps 3 & 4: Effective Potential

- \( V_{\text{ion}} \) known/constructed
- Initial guess \( n(r) \) or \( \psi_i(r) \)
- Calculate \( V_H[n] \) & \( V_{XC}[n] \)
- \( V_{\text{eff}}(r) = V_{\text{nuc}}(r) + V_H(r) + V_{XC}(r) \)
- \( H\psi_i(r) = [-1/2\nabla^2 + V_{\text{eff}}(r)] \psi_i(r) = \varepsilon_i \psi_i(r) \)
- Calculate new \( n(r) = \sum_i |\psi_i(r)|^2 \)
- Self-consistent?
  - No
  - Yes
- Problem solved! Can now calculate energy, forces, etc.

Note that type of exchange-correlation chosen while specifying pseudopotential in QE input file.
Exchange-Correlation Potential

• $V_{xc} \equiv \delta E_{xc}/\delta n$ contains all the many-body information.

• Known [numerically, from Quantum Monte Carlo; various analytical approximations] for homogeneous electron gas.

• Local Density Approximation:

$$E_{xc}[n] = \int n(r) V_{xc}^{\text{HOM}}[n(r)] \, dr$$

Surprisingly successful!

• Generalized Gradient Approximation(s): Include $\nabla n(r)$
Step 5: Diagonalization

- **Initial guess** $n(r)$
- **Calculate** $V_H[n]$ & $V_{XC}[n]$
- **Calculate** $V_{eff}(r) = V_{ion}(r) + V_H(r) + V_{XC}(r)$
- **Solve:** $H \psi_i(r) = [-1/2 \nabla^2 + V_{eff}(r)] \psi_i(r) = \epsilon_i \psi_i(r)$
- **Generate new** $n(r)$
- **Calculate new** $n(r) = \Sigma_i |\psi_i(r)|^2$

**Self-consistent?**
- **Expensive!** 😞

**Problem solved! Can now calculate energy, forces, etc.**
Diagonalization

• Need to diagonalize a matrix of size $N_{PW} \times N_{PW}$

• $N_{PW} \gg N_b = \text{number of bands required} = N_e/2$ or a little more (for metals).

• OK to obtain lowest few eigenvalues.

• Exact diagonalization is expensive!

• Use iterative diagonalizers that recast diagonalization as a minimization problem.
Step 6: New Charge Density

1. **Initial guess** $n(r)$
2. **Generate new new** $n(r)$
3. **Calculate** $V_H[n]$ & $V_{XC}[n]$
4. **Calculate** $V_{eff}(r) = V_{ion}(r) + V_H(r) + V_{XC}(r)$
5. **Calculate new** $n(r) = \sum_i |\psi_i(r)|^2$

**Self-consistent?**

- **Yes**: Problem solved! Can now calculate energy, forces, etc.
- **No**: Back to step 1.
Brillouin Zone Sums

- Many quantities (e.g., $n$, $E_{tot}$) involve integrals over $k$.
- In principle, need infinite number of $k$’s.
- In practice, sum over a finite number: BZ “Sampling”.
- Number needed depends on band structure.
- Typically need more $k$’s for metals.
- Need to test convergence wrt k-point sampling.

\[
\langle P \rangle = \frac{\Omega}{(2\pi)^3} \sum_{n \text{ occ}_BZ} \int P_n(k) d^3k
\]

\[
\langle P \rangle = \frac{1}{N_k} \sum_{n \text{ occ}} \sum_{k \in BZ} P_n(k)
\]

- $\mathbf{k}$ (wave-vector) is in the first Brillouin zone
- $n$ (band index) runs over occupied manifold
Using the Irreducible BZ; Weights

• Need not sum over $k$’s in entire BZ; can restrict to Irreducible BZ, with appropriate weights.

*e.g., for FCC:* 

\[
\langle P \rangle = \sum_{k \in IBZ} P_n(k) w(k)
\]

Count this only once.

Count this 8 x $\frac{1}{2} = 4$ times.

Count this 48 times.
Types of k-point meshes

- **Special Points:** [Chadi & Cohen]
  Points designed to give quick convergence for particular crystal structures.
- **Monkhorst-Pack:**
  Equally spaced mesh in reciprocal space.
  May be centred on origin [‘non-shifted’] or not [‘shifted’]

\[ nk1 = nk2 = 4 \]
Why it might be better to use a “shifted” grid

10 pts in IBZ

“Unshifted”

6 pts in IBZ

“Shifted”

Shobhana Narasimhan, JNCASR
Choosing Grid Divisions

- Space grid in a way (approximately) commensurate with length of primitive reciprocal lattice vectors $b$’s.
- Remember that dimensions in reciprocal space are the inverse of the dimensions in real space!
Choosing Grid Divisions

• For artificially periodic supercells, choose only 1 division along the dimensions that have been extended (in real space) by introducing vacuum region.
Problems with Metals

• Recall:

\[
\langle P \rangle = \frac{\Omega}{(2\pi)^3} \sum_{n \text{occ } BZ} \int P_n(k) d^3k
\]

• For metals, at \( T=0 \), this corresponds to (for highest band) an integral over all wave-vectors contained within the Fermi surface, i.e., for highest band, sharp discontinuity in k-space between occupied and unoccupied states… need many k-points to reproduce this accurately.

• Also can lead to scf convergence problems because of band-crossings above/below Fermi level.
Problems with Metals

The basic problem is that anything with sharp edges or features can’t be reproduced well if it is sampled coarsely...

...So smear out the quantity we are sampling into something that can be sampled coarsely...but of course...the procedure of smearing out may lead to errors...
A Smear Campaign!

- Problems arise because of sharp discontinuity at Fermi surface / Fermi energy.
- “Smear” this out using a smooth operator!
- Will now converge faster w.r.t. number of k-points (but not necessarily to the right answer!)
- The larger the smearing, the quicker the convergence w.r.t. number of k-points, but the greater the error introduced.
Fermi-Dirac Smearing

- Recall that the Fermi surface, which is sharply defined at $T=0$, becomes fuzzy as $T$ increased.
- One way of smearing: occupy with Fermi-Dirac distribution for a (fictitious) temperature $T > 0$.

$$f(E) = \frac{1}{e^{\frac{E-\mu}{\sigma}} + 1}$$

$$\sigma = k_B T$$

(schematic figs.)

Shobhana Narasimhan, JNCASR
The Free Energy

• When occupying with a finite $T$ distribution, what is variational (minimal) w.r.t. wavefunctions and occupations is not $E$ but $F=E-TS$

\[
S = -2k_B \sum_i [f_i \ln f_i + (1 - f_i) \ln(1 - f_i)]
\]

• What we actually want is $E(\sigma \to 0)$
• $E(\sigma \to 0) = \frac{1}{2} (F+E)$ (deviation $O(T^3)$)


Shobhana Narasimhan, JNCASR
Gaussian Smearing

- Think of the step function as an integral of $\delta$-fn.
- Replace sharp $\delta$-fn. by smooth gaussian.

$$f(E) = \frac{1}{2} \left[ 1 - \text{erf}\left(\frac{E - \mu}{\sigma}\right) \right]$$

(this is what you get if you integrate a Gaussian)

- Now have a generalized free energy \(E-TS\), where $S$ is a generalized entropy term.
- Converges faster (w.r.t. k-mesh) than Fermi-Dirac.
- Problem: need not converge to the right value, can get errors in forces.
- Want: fast convergence w.r.t. k-mesh to right answer!

Shobhana Narasimhan, JNCASR
Convergence wrt grid & smearing

- Gaussian smearing:

![Graph showing total energy vs k-point mesh for Bulk Al with different Gaussian smearings.]

Madhura Marathe

Shobhana Narasimhan, JNCASR
Better Smearing Functions

- **Methfessel & Paxton:**
  - Can have a successive series of better (but smooth) approximations to the step function.
  - \( E \) converges fast [wrt \( \sigma \)] to \( E(\sigma \to 0) \)

- **Marzari & Vanderbilt:**
  - Unlike Methfessel-Paxton, don’t have negative occupancies.

Convergence wrt grid & smearing

• Gaussian:

• Methfessel-Paxton:

represents an energy difference of 1 mRy

Madhura Marathe

Shobhana Narasimhan, JNCASR
Convergence wrt k-points & smearing width

e.g., for bcc Fe, using $14 \times 14 \times 14$ grid:
Smearing for Molecules

- Consider a molecule where HOMO is multiply degenerate and only partially occupied.
- If we don’t permit fractional occupancies…the code will occupy only one (or some) of the degenerate states, resulting in wrong symmetry.
- Smearing will fix this problem.
Step 7: Check if Convergence Achieved

1. **$V_{ion}$ known/constructed**

2. **Initial guess $n(r)$ or $\psi_i(r)$**

3. **Generate new $n(r)$**

4. **Calculate $V_H[n]$ & $V_{XC}[n]$**

5. **Calculate $V_{eff}(r) = V_{nuc}(r) + V_H(r) + V_{XC}(r)$**

6. **$H\psi_i(r) = [-1/2 \nabla^2 + V_{eff}(r)] \psi_i(r) = \epsilon_i \psi_i(r)$**

7. **Calculate new $n(r) = \Sigma_i |\psi_i(r)|$**

8. **Self-consistent?**

   - No
   - Yes

   **Problem solved! Can now calculate energy, forces, etc.**
Testing for scf convergence

- Compare nth and (n-1)th approximations for density, and see if they are close enough that self-consistency has been achieved.

*and/or*

- Compare values obtained for total energy in nth and (n-1)th scf iterations, and see if they match to within a pre-decided tolerance.
Step 8: Mixing

V_{ion} known/constructed

Initial guess \( n(r) \) or \( \psi_i(r) \)

Calculate \( V_H[n] \) & \( V_{XC}[n] \)

\[ V_{eff}(r) = V_{nuc}(r) + V_H(r) + V_{XC}(r) \]

\[ H\psi_i(r) = [-1/2 \nabla^2 + V_{eff}(r)] \psi_i(r) = \varepsilon_i \psi_i(r) \]

Calculate new \( n(r) = \sum_i |\psi_i(r)|^2 \)

Self-consistent?

No

Yes

Problem solved! Can now calculate energy, forces, etc.

Can take a long time to reach self-consistency!
Mixing

○ Iterations $n$ of self-consistent cycle:

○ Successive approximations to density:

$$n_{in}(n) \rightarrow n_{out}(n) \rightarrow n_{in}(n+1).$$

○ $n_{out}(n)$ fed directly as $n_{in}(n+1)$? No, usually doesn’t converge.

○ Need to mix, take some combination of input and output densities (may include information from several previous iterations).

○ Goal is to achieve self consistency ($n_{out} = n_{in}$) in as few iterations as possible.
III. What can we calculate?
What we can calculate

• Total energy in the ground state
• Density in the ground state
• Forces on nuclei
• Eigenvalues (not rigorously exact)
• Vibrational frequencies (e.g., by Density Functional Perturbation Theory)
• Magnetic properties (e.g., using Local Spin Density Approximation)
• Ferroelectric properties (e.g., using Berry’s phase formulation)
• Excited state properties (using Time Dependent Density Functional Theory)
• etc., etc.!
Output Quantities: Total Energy

- Perhaps the most important output quantity is the TOTAL ENERGY.
- Can use, e.g., to optimize structure.
- e.g., for a cubic crystal, where the structure can be specified by a single parameter (side of cube):

\[
\text{Energy vs. lattice constant or volume is given by "Equation of State"}
\]

Energy vs. lattice constant or volume is given by "Equation of State"
Output Quantities: Total Energy

- Can be used to calculate energy barriers:

![Graph showing the relationship between d(O-O) (Å) and barrier for O₂ dissociation (eV). The graph includes molecular structures of O₂, Au, Al, and MgO.]
Conclusions

Density Functional Theory is a very successful approach to calculating the properties of many-electron systems from first principles.

Bibliography

...and references therein.