

The Plane-Wave Pseudopotential Method

CECAM Tutorial on *Simulating Matter at the Nanoscale...*, 13 Nov. 2006

Outline

- Solution of the DFT problem: Self-consistency, global minimization
- Crystals: periodicity, direct and reciprocal lattice, unit cell, Brillouin Zone
- Plane waves (PW) basis set
- Pseudopotentials (PP)
- PW+PP technicalities

How to find the DFT ground state of a system ?

- Solving the Kohn-Sham equations self-consistently

$$(T + \hat{V} + V_H(n) + V_{xc}(n))\psi_i = \epsilon_i\psi_i$$

with charge density $n(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2$, $i =$ combined \mathbf{k} and band index,
 $f_i =$ occupancy of states, and orthonormality constraints $\langle \psi_i | \psi_j \rangle = \delta_{ij}$

- By constrained minimization of the energy functional

$$E[\{\psi_i\}] = \sum_i f_i \langle \psi_i | T + \hat{V} | \psi_i \rangle + E_H + E_{xc}(n) + E_{ion-ion}$$

with charge density and orthonormality constraints as above, i.e. minimize:

$$\tilde{E}[\{\psi_i\}, \{\Lambda_{ij}\}] = E[\{\psi_i\}] - \Lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij})$$

where the Λ_{ij} are Lagrange multipliers

Self-consistency

Starting from some guess of the input charge density $n^{in}(\mathbf{r})$:

$$n^{in} \longrightarrow (V_H + V_{xc})(n^{in}) \longrightarrow \psi_i(\mathbf{r}) \longrightarrow n^{out}(\mathbf{r}) = \sum_i f_i \psi_i^*(\mathbf{r}) \psi_i(\mathbf{r})$$

Simply re-inserting n^{out} as n^{in} will almost invariably not converge. Reason: In order to converge, such procedure must reduce the error at each step, but the low-frequency (small- \mathbf{G}) components of the error will likely not be reduced in typical condensed-matter systems.

Simple mixing algorithm:

$$n^{new} = \alpha n^{out} + (1 - \alpha) n^{in}, \quad 0 < \alpha < 1$$

Guaranteed to converge if α is small enough. More sophisticated algorithms (Anderson, Broyden) use the input and output of several preceding steps to determine the next optimal input combination.

Exchange-correlation potential

The exchange-correlation potential and energy in LDA are relatively simple functions of the local charge density at point \mathbf{r} :

$$V_{xc}(\mathbf{r}) = v_{xc}(n(\mathbf{r})), \quad E_{xc}[n] = \int e_{xc}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r}$$

They are easily calculated in \mathbf{r} -space. *Gradient-corrected* functionals depend on the gradient of the charge density as well. Slightly more complex to calculate.

Diagonalization of the Kohn-Sham Hamiltonian

Whatever approach one chooses, the basic building block of a DFT calculation is the evaluation of the Kohn-Sham Hamiltonian H_{KS} . How to solve $H\psi = \epsilon\psi$? Expand ψ in some suitable *basis set* $\{\phi_i\}$ as

$$\psi(\mathbf{r}) = \sum_i c_i \phi_i(\mathbf{r}).$$

For an orthonormal basis set, solve

$$\sum_j (H_{ij} - \epsilon \delta_{ij}) c_j = 0$$

where the *matrix elements* $H_{ij} = \langle \phi_i | H | \phi_j \rangle$. For a non-orthonormal basis set, solve:

$$\sum_j (H_{ij} - \epsilon S_{ij}) c_j$$

where $S_{ij} = \langle \phi_i | \phi_j \rangle$ (*overlap matrix*).

Total energy

Once convergence is reached, the total energy of the system may be calculated:

$$E = \sum_i f_i \langle \psi_i | T + \hat{V} | \psi_i \rangle + E_H + E_{xc}[n] + E_{ion-ion}$$

where

$$E_H = \frac{e^2}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'$$

$$E_{xc}[n] = \int e_{xc}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r}$$

$$E_{ion-ion} = \frac{e^2}{2} \sum'_{\mu,\nu} \frac{Z_\mu Z_\nu}{|\mathbf{R}_\mu - \mathbf{R}_\nu|}$$

(the primed sum excludes terms with $\mathbf{R}_\mu - \mathbf{R}_\nu = 0$)

Equivalent expression for the energy, using Kohn-Sham eigenvalues:

$$E = \sum_i f_i \epsilon_i - E_H + \int (e_{xc}(n) - V_{xc}(n))n(\mathbf{r})d\mathbf{r} + E_{ion-ion}$$

Forces, Structural Optimization, Dynamics etc

The total energy depends on atomic positions \mathbf{R}_μ into the unit cell and on the lattice. *Forces* acting on atoms are the derivatives of the total energy wrt atomic positions:

$$\mathbf{F}_\mu = -\frac{\partial E}{\partial \mathbf{R}_\mu} = \sum_i f_i \langle \psi_i | \frac{\partial \hat{V}}{\partial \mathbf{R}_\mu} | \psi_i \rangle$$

The *Hellmann-Feynman theorem* holds in DFT due to its variational character!

Once forces are calculated, one can perform structural optimization, molecular dynamics, etc.

Periodicity

Let us consider the case of the *infinite perfect crystals*, having translation symmetry. A perfect crystal is described in terms of

- a **unit cell** that is periodically repeated
- a **basis** of atomic positions \mathbf{d}_i in the unit cell
- a **lattice** of translation vectors $\mathbf{R} = n_1\mathbf{R}_1 + n_2\mathbf{R}_2 + n_3\mathbf{R}_3$
- a **reciprocal lattice** of vectors \mathbf{G} such that $\mathbf{G} \cdot \mathbf{R} = 2\pi l$, with l integer.
Such conditions hold if $\mathbf{G} = m_1\mathbf{G}_1 + m_2\mathbf{G}_2 + m_3\mathbf{G}_3$ with $\mathbf{G}_i \cdot \mathbf{R}_j = 2\pi\delta_{ij}$.

Non periodic systems ?

What about e.g. defects in crystals, surfaces, alloys, amorphous materials, liquids, molecules, clusters? none of these has perfect periodicity. One can use **supercells**, introducing an artificial periodicity.

The supercell geometry is dictated by the type of system under investigation:

- Molecules, clusters:
the supercell must allow a minimum distance of at least a few Å (~ 6) between the closest atoms in different periodic replica.

- Defects in crystals:
the supercell is commensurate with the perfect crystal cell. The distance between periodic replica of the defect must be “big enough” to minimize spurious defect-defect interactions.
- Surfaces:
slab geometry. The number of layers of the materials must be “big enough” to have “bulk behaviour” in the furthest layer from the surface. The number of empty layers must be “big enough” to have minimal interactions between layers in different regions.
- Alloys, amorphous materials, liquids:
the supercell must be “big enough” to give a reasonable description of physical properties.

Band Structure

The one-electron states $\psi(\mathbf{r})$ of a perfect crystal Hamiltonian $H = T + V$ are described by a **band index** i and a **wave vector** \mathbf{k} .

It is convenient to consider the *thermodynamic limit*: a slab of crystal composed of $\mathcal{N} = \mathcal{N}_1\mathcal{N}_2\mathcal{N}_3$ unit cells, $\mathcal{N} \rightarrow \infty$, obeying Periodic Boundary Conditions:

$$\psi(\mathbf{r} + \mathcal{N}_1\mathbf{R}_1) = \psi(\mathbf{r} + \mathcal{N}_2\mathbf{R}_2) = \psi(\mathbf{r} + \mathcal{N}_3\mathbf{R}_3) = \psi(\mathbf{r}).$$

There are \mathcal{N} wave vectors \mathbf{k} in the unit cell of the reciprocal lattice, called the **Brillouin Zone**. The one-electron states (energy bands) can be written as

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{i,\mathbf{k}}(\mathbf{r})$$

where $u_{i,\mathbf{k}}(\mathbf{r})$ is translationally invariant:

$$u_{i,\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{i,\mathbf{k}}(\mathbf{r}).$$

Most popular basis sets:

- *Localized* basis sets:
Bloch sums $\phi_{\mathbf{k}} = \sum_{\mathbf{R}} \exp(-i\mathbf{k} \cdot \mathbf{R}) \phi(\mathbf{r} - \mathbf{R})$ of functions centered on atoms
 - Linear Combinations of Atomic Orbitals (LCAO)
 - Gaussian-type Orbitals (GTO)
 - Linearized Muffin-Tin Orbitals (LMTO)
- *Delocalized* basis sets:
 - Plane Waves (PW)
- *Mixed* basis sets:
 - Linearized Augmented Plane Waves (LAPW)
 - Projector Augmented Plane Waves (PAW)

PW basis set

A PW basis set for states of wave vector \mathbf{k} is defined as

$$\langle \mathbf{r} | \mathbf{k} + \mathbf{G} \rangle = \frac{1}{\sqrt{\mathcal{N}\Omega}} e^{i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}}, \quad \frac{\hbar^2}{2m} |\mathbf{k} + \mathbf{G}|^2 \leq E_{cut}$$

$\Omega =$ unit cell volume, $\mathcal{N}\Omega =$ crystal volume, $E_{cut} =$ cutoff on the kinetic energy of PWs (in order to have a finite number of PWs!). The PW basis set is *complete* for $E_{cut} \rightarrow \infty$ and *orthonormal*: $\langle \mathbf{k} + \mathbf{G} | \mathbf{k} + \mathbf{G}' \rangle = \delta_{\mathbf{G}\mathbf{G}'}$

The components on a PW basis set are the *Fourier transform*:

$$|\psi_i\rangle = \sum_{\mathbf{G}} c_{i,\mathbf{k}+\mathbf{G}} |\mathbf{k} + \mathbf{G}\rangle$$

$$c_{i,\mathbf{k}+\mathbf{G}} = \langle \mathbf{k} + \mathbf{G} | \psi_i \rangle = \frac{1}{\sqrt{\mathcal{N}\Omega}} \int \psi_i(\mathbf{r}) e^{-i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}} d\mathbf{r} = \tilde{\psi}_i(\mathbf{k} + \mathbf{G}).$$

Hamiltonian diagonalization

The solution of $H_{KS}\psi_i = \epsilon_i\psi_i$ at fixed potential:

$$\sum_{\mathbf{G}'} (H_{KS}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') - \epsilon_i \delta_{\mathbf{G}, \mathbf{G}'}) c_{i, \mathbf{k} + \mathbf{G}'} = 0$$

requires the *diagonalization* of the matrix H_{KS} , whose matrix elements are:

$$H_{KS}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') = \langle \mathbf{k} + \mathbf{G} | H_{KS} | \mathbf{k} + \mathbf{G}' \rangle$$

Matrix element of the Hamiltonian in PWs

- Kinetic energy term:

$$\langle \mathbf{k} + \mathbf{G} | T | \mathbf{k} + \mathbf{G}' \rangle = \frac{\hbar^2}{2m} (\mathbf{k} + \mathbf{G})^2 \delta_{\mathbf{G}\mathbf{G}'}$$

- Hartree and exchange correlation terms:

$$\langle \mathbf{k} + \mathbf{G} | V_H | \mathbf{k} + \mathbf{G}' \rangle = V_H(\mathbf{G} - \mathbf{G}') = 4\pi e^2 \frac{n(\mathbf{G} - \mathbf{G}')}{|\mathbf{G} - \mathbf{G}'|^2} \quad (\mathbf{G} \neq \mathbf{G}')$$

and

$$\langle \mathbf{k} + \mathbf{G} | V_{xc} | \mathbf{k} + \mathbf{G}' \rangle = V_{xc}(\mathbf{G} - \mathbf{G}')$$

where

$$f(\mathbf{G}) = \frac{1}{\Omega} \int f(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} d\mathbf{r}$$

is the *Fourier Transform* of $f(\mathbf{r})$.

- External ionic potential:

$$\hat{V} \equiv V(\mathbf{r}) = \sum_{\mu \in cell} \sum_{\mathbf{R}} V_{\mu}(\mathbf{r} - \mathbf{d}_{\mu} - \mathbf{R})$$

$$\begin{aligned} \langle \mathbf{k} + \mathbf{G} | V | \mathbf{k} + \mathbf{G}' \rangle &= \frac{1}{\mathcal{N}\Omega} \int e^{-i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}} \sum_{\mu} \sum_{\mathbf{R}} V_{\mu}(\mathbf{r} - \mathbf{d}_{\mu} - \mathbf{R}) e^{i(\mathbf{k}+\mathbf{G}') \cdot \mathbf{r}} d\mathbf{r} \\ &= \sum_{\mu} S_{\mu}(\mathbf{G} - \mathbf{G}') V_{\mu}(\mathbf{G} - \mathbf{G}') \end{aligned}$$

where $S_{\mu}(\mathbf{G})$ is the *structure factor*, containing geometrical information:

$$S_{\mu}(\mathbf{G}) = \sum_{\substack{\text{type } \mu \\ \mathbf{d}_i \in cell}} e^{-i\mathbf{G} \cdot \mathbf{d}_i}$$

If the potential has spherical symmetry, $V(\mathbf{r}) = V(r)$, then:

$$V(\mathbf{G}) = V(G) = \frac{4\pi}{\Omega} \int r^2 V(r) \frac{\sin(Gr)}{G} dr.$$

Advantages and disadvantages of various basis sets

- Localized basis sets:
 - + fast convergence with respect to basis set size (just a few functions per atom needed)
 - + no problem with finite systems
 - difficult to evaluate convergence quality (no systematic way to improve convergence)
 - difficult to use (two- and three-center integrals)
 - difficult to calculate forces (*Pulay forces*)
- Plane Waves:
 - slow convergence with respect to basis set size (many more PWs than localized functions needed)
 - require supercells with finite systems
 - + easy to evaluate convergence quality (just increase cutoff)
 - + easy to use (Fourier transform)
 - + easy to calculate forces (no Pulay forces)

Charge density

The calculation of the charge density

$$n(\mathbf{r}) = \sum_{i,\mathbf{k}} f_{i,\mathbf{k}} \psi_{i,\mathbf{k}}^*(\mathbf{r}) \psi_{i,\mathbf{k}}(\mathbf{r})$$

requires in principle a sum over an infinite (\mathcal{N}) number of \mathbf{k} -points in the Brillouin Zone. As a matter of fact, for insulators just a few “special” points (i.e. a uniform grid in most cases) are required. For metals, a much denser grid is usually required.

Charge density in reciprocal space:

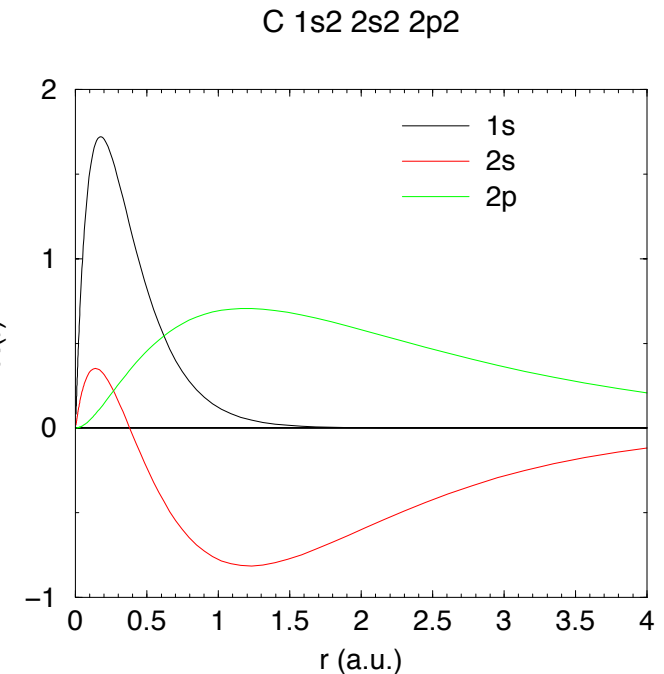
$$n(\mathbf{G}) = \frac{1}{\Omega} \int n(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} d\mathbf{r} = \sum_{\mathbf{G}'} \sum_{i,\mathbf{k}} f_{i,\mathbf{k}} c_{i,\mathbf{k}+\mathbf{G}'}^* c_{i,\mathbf{k}+\mathbf{G}+\mathbf{G}'}$$

This is NOT how $n(\mathbf{G})$ is calculated, in practical calculations

The need for Pseudopotentials

Are PWs a practical basis set for electronic structure calculations? Not really! From elementary Fourier analysis: length scale $\delta \longrightarrow$ Fourier components up to $q \sim 2\pi/\delta$. In a solid, this means $\sim 4\pi(2\pi/\delta)^3/3\Omega_{BZ}$ PWs (Ω_{BZ} = volume of the Brillouin Zone).

Estimate for diamond: $1s$ wavefunction has $\delta \simeq 0.1$ a.u., $\Omega = (2\pi)^3/(a_0^3/4)$ with lattice parameter $a_0 = 6.74$ a.u. \longrightarrow 250,000 PWs!



Need to:

- get rid of core states
- get rid of orthogonality wiggles close to the nucleus

Solution: **Pseudopotentials** (PP). A smooth effective potential that reproduces the effect of the nucleus plus core electrons on valence electrons.