VASP: Plane waves, the PAW method, and the Selfconsistency cycle

Martijn Marsman

Computational Materials Physics, Faculty of Physics, University Vienna, and Center for Computational Materials Science

universität wien

CMS

Computational Materials Science ★ Center for ★

DFT and beyond, 14th July 2011, Berlin, Germany

A system of N electrons

$$\hat{H}\Psi(\mathbf{r}_1, ..., \mathbf{r}_N) = E\Psi(\mathbf{r}_1, ..., \mathbf{r}_N)$$

$$\left(-\frac{1}{2}\sum_i \Delta_i + \sum_i V(\mathbf{r}_i) + \sum_{i \neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\right)\Psi(\mathbf{r}_1, ..., \mathbf{r}_N) = E\Psi(\mathbf{r}_1, ..., \mathbf{r}_N)$$

Many-body WF storage requirements are prohibitive

$$(\#\text{grid points})^N$$

Map onto "one-electron" theory

$$\Psi(\mathbf{r}_1, ..., \mathbf{r}_N) \rightarrow \{\psi_1(\mathbf{r}), \psi_2(\mathbf{r}), ..., \psi_N(\mathbf{r})\}$$

such as Hohenberg-Kohn-Sham density functional theory

Do not need $\Psi(\mathbf{r}_1, ..., \mathbf{r}_N)$, just the density $\rho(\mathbf{r})$:

$$E[\rho] = T_s[\{\psi_i[\rho]\}] + E_H[\rho] + E_{\mathrm{xc}}[\rho] + E_Z[\rho] + U[Z]$$

$$\Psi(\mathbf{r}_1, ..., \mathbf{r}_N) = \prod_i^N \psi_i(\mathbf{r}_i) \qquad \rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \qquad E_H[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'$$

One-electron Kohn-Sham equations

$$\left(-\frac{1}{2}\Delta + V_Z(\mathbf{r}) + V_H[\rho](\mathbf{r}) + V_{\mathrm{xc}}[\rho](\mathbf{r}) \right) \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r})$$

Hartree   Exchange-Correlation

$$V_H[\rho](\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \qquad E_{\mathrm{xc}}[\rho] = ??? \qquad V_{\mathrm{xc}}[\rho](\mathbf{r}) = ???$$

Per definition: $E_{\mathrm{xc}} = E - T_s - E_H - E_{\mathrm{ext}}$

In practice: Exchange-Correlation functionals are modelled on the uniform electron gas (Monte Carlo calculations): e.g., local density approximation (LDA).

- Translational invariance implies the existence of a good quantum number, usually called the Bloch wave vector $\mathbf{k}$. All electronic states can be indexed by this quantum number

$$|\Psi_{\mathbf{k}}\rangle$$

- In a one-electron theory, one can introduce a second index, corresponding to the one-electron band $\mathbf{n}$,
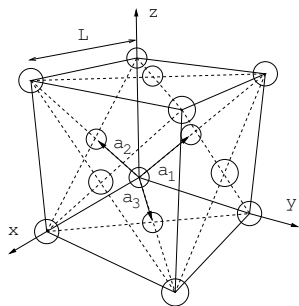
$$|\psi_{n\mathbf{k}}\rangle$$

- The Bloch theorem states that the one-electron wavefunctions obey the equation:
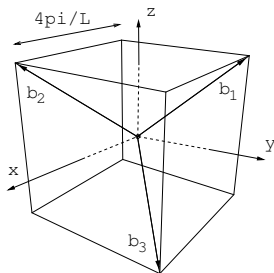
$$\psi_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) = \psi_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\mathbf{R}}$$

where $\mathbf{R}$ is any translational vector leaving the Hamiltonian invariant.

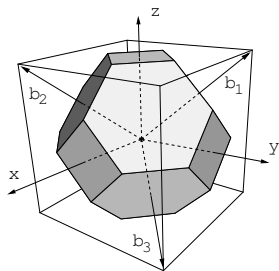- $\mathbf{k}$ is usually constrained to lie within the first Brillouin zone in reciprocal space.

A   B   C

$$\mathbf{b}_1 = \frac{2\pi}{\Omega}\mathbf{a}_2 \times \mathbf{a}_3 \quad \mathbf{b}_2 = \frac{2\pi}{\Omega}\mathbf{a}_3 \times \mathbf{a}_1 \quad \mathbf{b}_3 = \frac{2\pi}{\Omega}\mathbf{a}_1 \times \mathbf{a}_2$$

$$\Omega = \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3 \qquad \mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$$

- The evaluation of many key quantities, e.g. charge density, density-of-states, and total energy) requires integration over the first BZ. The charge density $\rho(\mathbf{r})$, for instance, is given by

$$\rho(\mathbf{r}) = \frac{1}{\Omega_{\mathrm{BZ}}} \sum_n \int_{\mathrm{BZ}} f_{n\mathbf{k}} |\psi_{n\mathbf{k}}(\mathbf{r})|^2 d\mathbf{k}$$

- $f_{n\mathbf{k}}$ are the occupation numbers, i.e., the number of electrons that occupy state $n\mathbf{k}$.

- Exploiting the fact that the wave functions at $\mathbf{k}$-points that are close together will be almost identical, one may approximate the integration over $\mathbf{k}$ by a weighted sum over a discrete set of points

$$\rho(\mathbf{r}) = \sum_n \sum_{\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} |\psi_{n\mathbf{k}}(\mathbf{r})|^2 d\mathbf{k},$$

where the weights $w_{\mathbf{k}}$ sum up to one.

The intractable task of determining $\Psi(\mathbf{r}_1, ..., \mathbf{r}_N)$ (for $N \sim 10^{23}$) has been reduced to calculating $\psi_{n\mathbf{k}}(\mathbf{r})$ at a discrete set of points $\{\mathbf{k}\}$ in the first BZ, for a number of bands that is of the order of the number of electrons *per unit cell*.
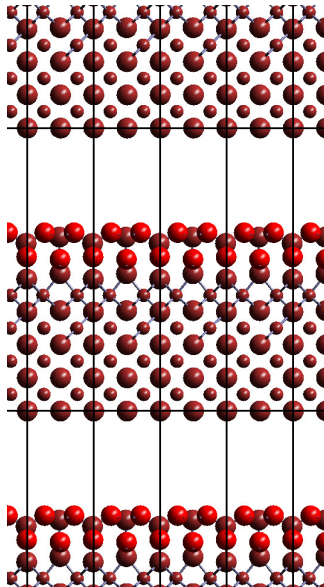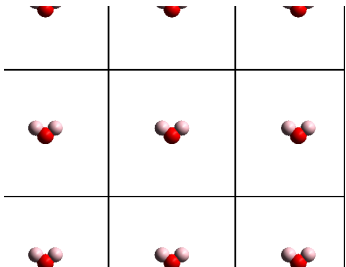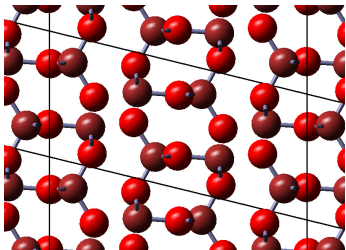
- The total energy

$$E[\rho, \{\mathbf{R}, Z\}] = T_s[\{\psi_{n\mathbf{k}}[\rho]\}] + E_H[\rho, \{\mathbf{R}, Z\}] + E_{xc}[\rho] + U(\{\mathbf{R}, Z\})$$

- The kinetic energy

$$T_s[\{\psi_{n\mathbf{k}}[\rho]\}] = \sum_n \sum_{\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \langle \psi_{n\mathbf{k}}| - \frac{1}{2}\Delta|\psi_{n\mathbf{k}}\rangle$$

- The Hartree energy

$$E_{\mathrm{H}}[\rho, \{\mathbf{R}, Z\}] = \frac{1}{2} \iint \frac{\rho_{eZ}(\mathbf{r})\rho_{eZ}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' d\mathbf{r}$$

where $\rho_{eZ}(\mathbf{r}) = \rho(\mathbf{r}) + \sum_i Z_i \delta(\mathbf{r} - \mathbf{R}_i)$

- The electronic charge density

$$\rho(\mathbf{r}) = \sum_n \sum_{\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} |\psi_{n\mathbf{k}}(\mathbf{r})|^2 d\mathbf{k},$$

- The Kohn-Sham equations

$$\left(-\frac{1}{2}\Delta + V_H[\rho_{eZ}](\mathbf{r}) + V_{xc}[\rho](\mathbf{r})\right)\psi_{n\mathbf{k}}(\mathbf{r}) = \epsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r})$$

- The Hartree potential

$$V_H[\rho_{eZ}](\mathbf{r}) = \int \frac{\rho_{eZ}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$$

- Introduce the cell periodic part $u_{n\mathbf{k}}$ of the wavefunctions

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

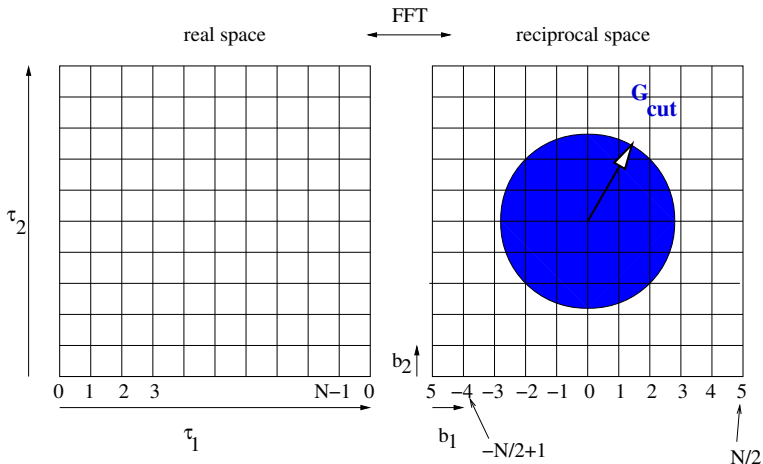 with $u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{n\mathbf{k}}(\mathbf{r})$.

- All cell periodic functions are now written as a sum of plane waves

$$u_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\Omega^{1/2}} \sum_{\mathbf{G}} C_{\mathbf{G}n\mathbf{k}} e^{i\mathbf{G}\mathbf{r}} \quad \psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\Omega^{1/2}} \sum_{\mathbf{G}} C_{\mathbf{G}n\mathbf{k}} e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}}$$

$$\rho(\mathbf{r}) = \sum_{\mathbf{G}} \rho_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}} \qquad V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}$$

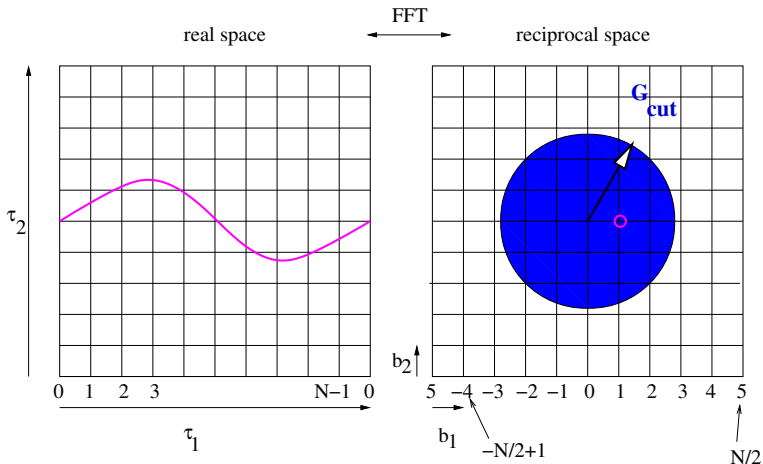- In practice only those plane waves $|\mathbf{G} + \mathbf{k}|$ are included for which

$$\frac{1}{2}|\mathbf{G} + \mathbf{k}|^2 < E_{\text{cutoff}}$$

M. Marsman   VASP: PWs, the PAW method, and the SC cycle

real space $\xleftrightarrow{\text{FFT}}$ reciprocal space

$x_1 = n_1 / N \; \tau_1$

$g_1 = n_1 \, 2\pi / \tau_1$

$$C_{\mathbf{r}n\mathbf{k}} = \sum_{\mathbf{G}} C_{\mathbf{G}n\mathbf{k}} e^{i\mathbf{G}\mathbf{r}} \quad \xleftrightarrow{\text{FFT}} \quad C_{\mathbf{G}n\mathbf{k}} = \frac{1}{N_{\text{FFT}}} \sum_{\mathbf{r}} C_{\mathbf{r}n\mathbf{k}} e^{-i\mathbf{G}\mathbf{r}}$$
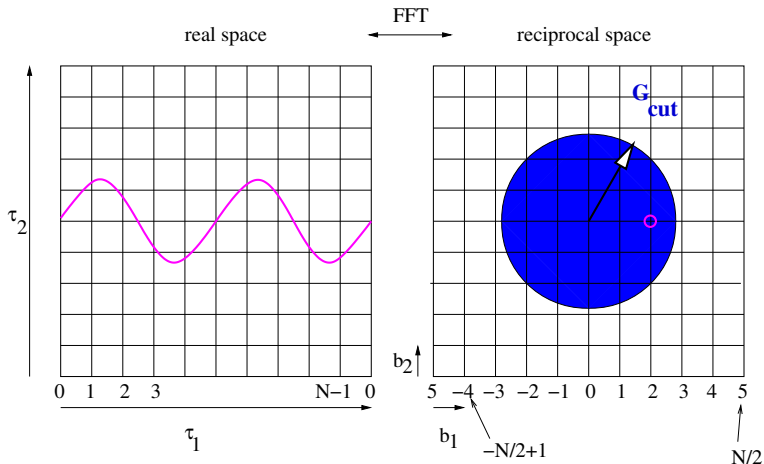
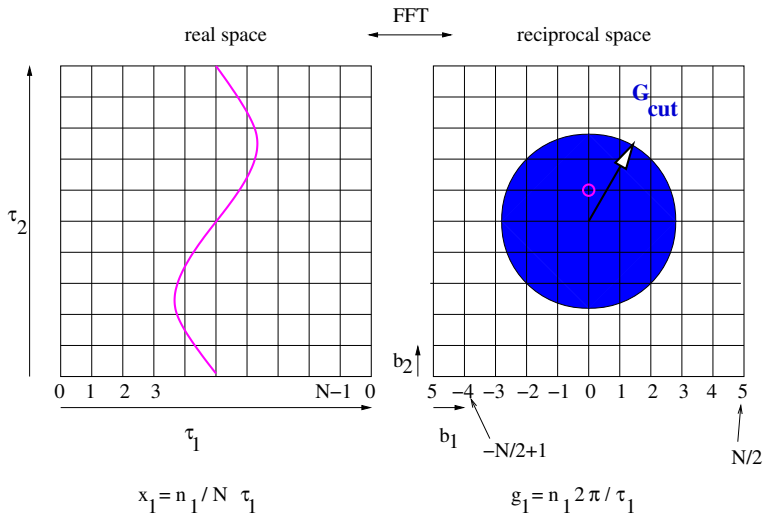$$C_{\mathbf{r}n\mathbf{k}} = \sum_{\mathbf{G}} C_{\mathbf{G}n\mathbf{k}} e^{i\mathbf{G}\mathbf{r}} \quad \xleftrightarrow{\text{FFT}} \quad C_{\mathbf{G}n\mathbf{k}} = \frac{1}{N_{\text{FFT}}} \sum_{\mathbf{r}} C_{\mathbf{r}n\mathbf{k}} e^{-i\mathbf{G}\mathbf{r}}$$

real space   ⟷ FFT ⟷   reciprocal space

$$x_1 = n_1 / N \ \tau_1 \qquad\qquad g_1 = n_1 \, 2\pi / \tau_1$$

$$C_{\mathbf{r}n\mathbf{k}} = \sum_{\mathbf{G}} C_{\mathbf{G}n\mathbf{k}} e^{i\mathbf{G}\mathbf{r}} \quad \overset{\mathrm{FFT}}{\longleftrightarrow} \quad C_{\mathbf{G}n\mathbf{k}} = \frac{1}{N_{\mathrm{FFT}}} \sum_{\mathbf{r}} C_{\mathbf{r}n\mathbf{k}} e^{-i\mathbf{G}\mathbf{r}}$$

FFT

real space ◄───────► reciprocal space

τ_2

τ_1

0 1 2 3 N−1 0

$\mathbf{G_{cut}}$

b_2

b_1

5 −4 −3 −2 −1 0 1 2 3 4 5

−N/2+1

N/2
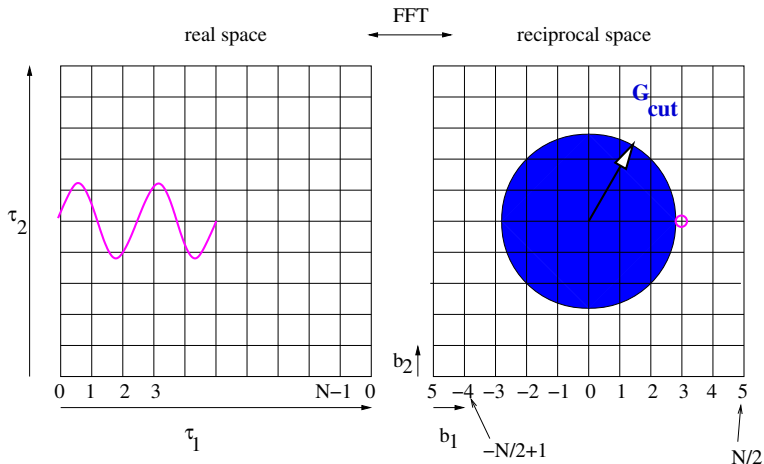
$x_1 = n_1 / N \ \tau_1$

$g_1 = n_1 \, 2\pi / \tau_1$

$$C_{\mathbf{r}n\mathbf{k}} = \sum_{\mathbf{G}} C_{\mathbf{G}n\mathbf{k}} e^{i\mathbf{G}\mathbf{r}} \quad \overset{\text{FFT}}{\longleftrightarrow} \quad C_{\mathbf{G}n\mathbf{k}} = \frac{1}{N_{\text{FFT}}} \sum_{\mathbf{r}} C_{\mathbf{r}n\mathbf{k}} e^{-i\mathbf{G}\mathbf{r}}$$

Why use plane waves?

- Historical reason: Many elements exhibit a band-structure that can be interpreted in a free electron picture (metallic s and p elements). Pseudopotential theory was initially developed to cope with these elements (pseudopotential perturbation theory).

- Practical reason: The total energy expressions and the Hamiltonian $\mathbf{H}$ are easy to implement.

- Computational reason: The action $\mathbf{H}|\psi\rangle$ can be efficiently evaluated using FFT's.

Evaluation of $\mathbf{H}|\psi_{n\mathbf{k}}\rangle$

$$\left(-\frac{1}{2}\Delta + V(\mathbf{r})\right)\psi_{n\mathbf{k}}(\mathbf{r})$$

using the convention

$$\langle\mathbf{r}|\mathbf{G}+\mathbf{k}\rangle = \frac{1}{\Omega^{1/2}}e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}} \rightarrow \langle\mathbf{G}+\mathbf{k}|\psi_{n\mathbf{k}}\rangle = C_{\mathbf{G}n\mathbf{k}}$$

- Kinetic energy:

$$\langle\mathbf{G}+\mathbf{k}| -\frac{1}{2}\Delta|\psi_{n\mathbf{k}}\rangle = \frac{1}{2}|\mathbf{G}+\mathbf{k}|^2 C_{\mathbf{G}n\mathbf{k}} \qquad N_{\mathrm{NPLW}}$$

- Local potential: $V = V_{\mathrm{H}}[\rho] + V_{xc}[\rho] + V_{\mathrm{ext}}$
) Exchange-correlation: easily obtained in real space $V_{\mathrm{xc},\mathbf{r}} = V_{\mathrm{xc}}[\rho_{\mathbf{r}}]$
) FFT to reciprocal space $\{V_{\mathrm{xc},\mathbf{r}}\} \rightarrow \{V_{\mathrm{xc},\mathbf{G}}\}$
) Hartree potential: Poisson equation in reciprocal space $V_{\mathrm{H},\mathbf{G}} = \frac{4\pi}{|\mathbf{G}|^2}\rho_{\mathbf{G}}$
) add all contributions $V_{\mathbf{G}} = V_{\mathrm{H},\mathbf{G}} + V_{\mathrm{xc},\mathbf{G}} + V_{\mathrm{ext},\mathbf{G}}$
) FFT to real space $\{V_{\mathbf{G}}\} \rightarrow \{V_{\mathbf{r}}\}$
  The action

$$\langle\mathbf{G}+\mathbf{k}|V|\psi_{n\mathbf{k}}\rangle = \frac{1}{N_{\mathrm{FFT}}}\sum_{\mathbf{r}} V_{\mathbf{r}}C_{\mathbf{r}n\mathbf{k}}e^{-i\mathbf{G}\mathbf{r}} \qquad N_{\mathrm{FFT}}\log N_{\mathrm{FFT}}$$

The action of the local potential
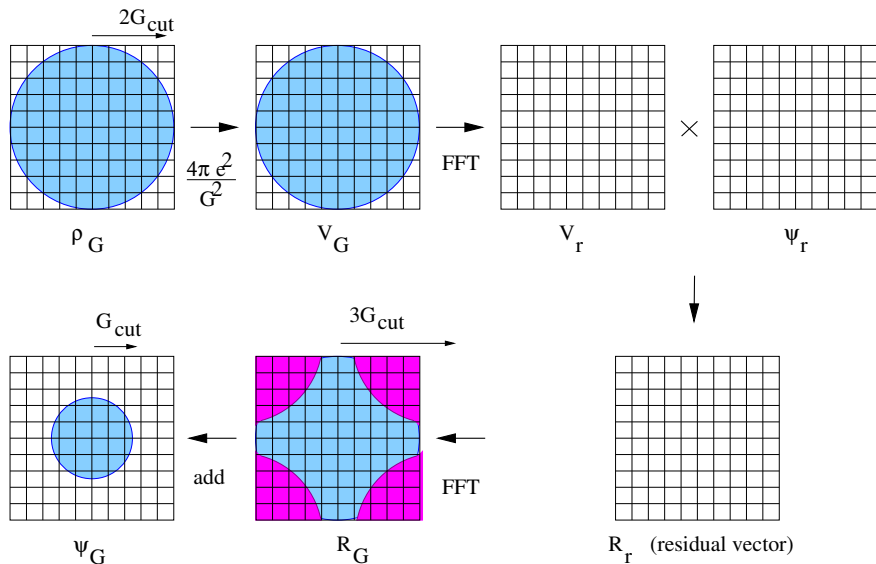
The PAW method

The number of plane waves needed to describe

- tightly bound (spatially strongly localized) states

- the rapid oscillations (nodal features) of the wave functions near the nucleus

exceeds any practical limit, except maybe for Li and H.

The common solution:

- Introduce the frozen core approximation:
  Core electrons are pre-calculated in an atomic environment and kept frozen in the course of the remaining calculations.

- Use pseudopotentials instead of exact potentials:
) Norm-conserving pseudopotentials
) Ultra-soft pseudopotentials
) The Projector-Augmented-Wave (PAW) method
    [P.E. Blöchl, Phys. Rev. B 50, 17953 (1994)]
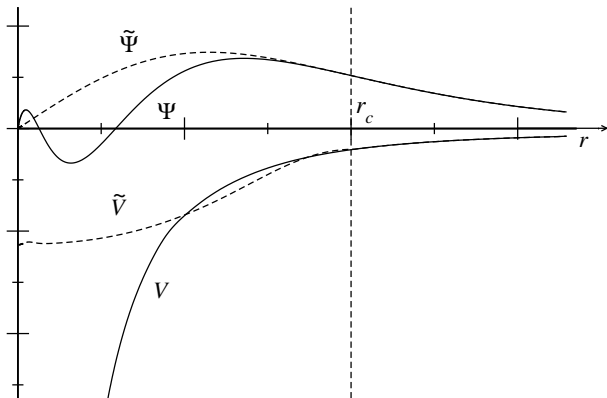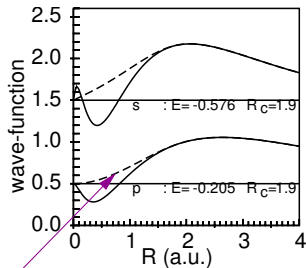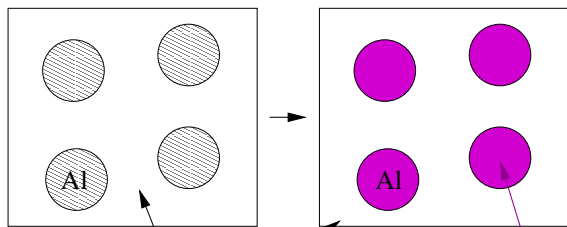
exact potential (interstitial region)   pseudopotential

Al	effective Al atom	PAW Al atom
3p	2p	3p
3s	1s	3s
~~2p~~		
~~2s~~	2p and 1s are	nodal structure
~~1s~~	nodeless !!!!	is retained

$$|\psi_n\rangle = |\widetilde{\psi}_n\rangle + \sum_i (|\phi_i\rangle - |\widetilde{\phi}_i\rangle)\langle\widetilde{p}_i|\widetilde{\psi}_n\rangle$$

- $|\widetilde{\psi}_n\rangle$ is a pseudo wave function expanded in plane waves
- $|\phi_i\rangle$, $|\widetilde{\phi}_i\rangle$, and $|\widetilde{p}_i\rangle$ are atom centered localized functions
- the all-electron partial waves $|\phi_i\rangle$ are obtained as solutions to the radial scalar relativistic Schrödinger equation for the spherical non-spinpolarized atom

$$(-\frac{1}{2}\Delta + v_{\text{eff}})|\phi_i\rangle = \epsilon_i|\phi_i\rangle$$

- a pseudization procedure yields

$$|\phi_i\rangle \rightarrow |\widetilde{\phi}_i\rangle \qquad v_{\text{eff}} \rightarrow \widetilde{v}_{\text{eff}} \qquad \langle\widetilde{p}_i|\widetilde{\phi}_j\rangle = \delta_{ij}$$

- the pseudo partial waves $|\widetilde{\phi}_k\rangle$ obey

$$\left(-\frac{1}{2}\Delta + \widetilde{v}_{\text{eff}} + \sum_{ij}|\widetilde{p}_i\rangle D_{ij}\langle\widetilde{p}_j|\right)|\widetilde{\phi}_k\rangle = \epsilon_k\left(1 + \sum_{ij}|\widetilde{p}_i\rangle Q_{ij}\langle\widetilde{p}_j|\right)|\widetilde{\phi}_k\rangle$$

- with the socalled PAW parameters:

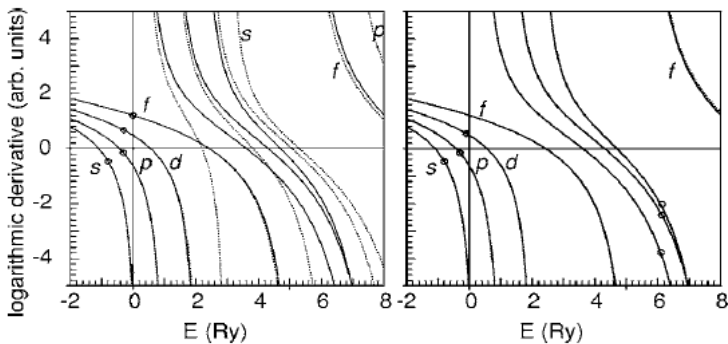$$Q_{ij} = \langle\phi_i|\phi_j\rangle - \langle\widetilde{\phi}_i|\widetilde{\phi}_j\rangle$$

$$D_{ij} = \langle\phi_i| - \frac{1}{2}\Delta + v_{\text{eff}}|\phi_j\rangle - \langle\widetilde{\phi}_i| - \frac{1}{2}\Delta + \widetilde{v}_{\text{eff}}|\widetilde{\phi}_j\rangle$$

The all-electron and pseudo eigenvalue spectrum is identical, all-electron scattering properties are reproduced over a wide energy range.

$$\left(-\frac{1}{2}\Delta + v_{\text{eff}}\right)|\phi_i\rangle = \epsilon_i|\phi_i\rangle$$

$$\left(-\frac{1}{2}\Delta + \widetilde{v}_{\text{eff}} + \sum_{kl}|\widetilde{p}_k\rangle D_{kl}\langle\widetilde{p}_l|\right)|\widetilde{\phi}_i\rangle = \epsilon_i\left(1 + \sum_{kl}|\widetilde{p}_k\rangle Q_{kl}\langle\widetilde{p}_l|\right)|\widetilde{\phi}_i\rangle$$

$$\left.\frac{\partial\widetilde{\phi}_l(r,\epsilon)}{\partial r}\frac{1}{\widetilde{\phi}_l(r,\epsilon)}\right|_{r=r_c} \approx \left.\frac{\partial\phi_l(r,\epsilon)}{\partial r}\frac{1}{\phi_l(r,\epsilon)}\right|_{r=r_c}$$

1st s-channel: ϵ_1
Mn $4s$ "bound" state

2nd s-channel: ϵ_2
Mn s "non-bound" state
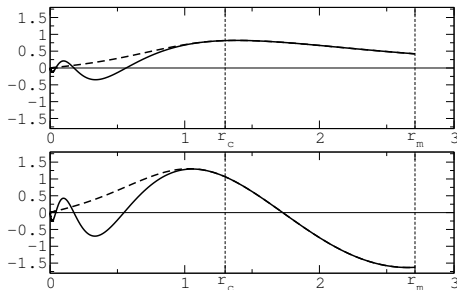
Frozen core approximation:

$$v_{\text{eff}}[\rho_v] = v_H[\rho_v] + v_H[\rho_{Zc}] + v_{xc}[\rho_v + \rho_c] \qquad \rho_v(\mathbf{r}) = \sum_i a_i |\phi_i(\mathbf{r})|^2$$

$$\widetilde{v}_{\text{eff}}[\widetilde{\rho}_v] = v_H[\widetilde{\rho}_v] + v_H[\widetilde{\rho}_{Zc}] + v_{xc}[\widetilde{\rho}_v + \widetilde{\rho}_c] \qquad \widetilde{\rho}_v(\mathbf{r}) = \sum_i a_i |\widetilde{\phi}_i(\mathbf{r})|^2$$

$|\widetilde{\psi}_n\rangle$
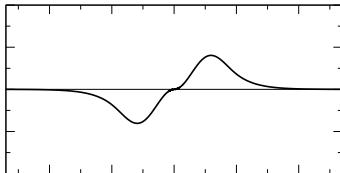
$|\widetilde{\psi}_n\rangle$

$|\widetilde{\psi}_n\rangle - \sum_i |\widetilde{\phi}_i\rangle\langle\widetilde{p}_i|\widetilde{\psi}_n\rangle$
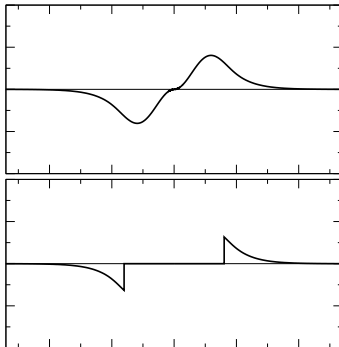
$|\widetilde{\psi}_n\rangle$
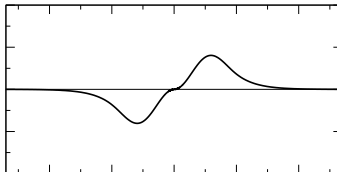
$|\widetilde{\psi}_n\rangle - \sum_i |\widetilde{\phi}_i\rangle\langle\widetilde{p}_i|\widetilde{\psi}_n\rangle$
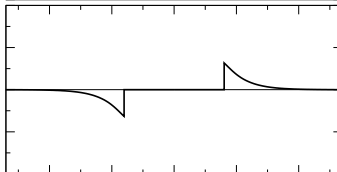
$|\widetilde{\psi}_n\rangle - \sum_i |\widetilde{\phi}_i\rangle\langle\widetilde{p}_i|\widetilde{\psi}_n\rangle + \sum_i |\phi_i\rangle\langle\widetilde{p}_i|\widetilde{\psi}_n\rangle$
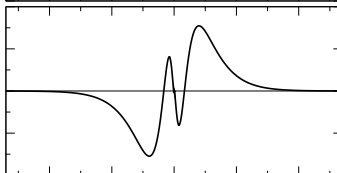
- Character of wavefunction: $c_{lm\epsilon} = \langle \tilde{p}_{lm\epsilon} | \tilde{\psi}_n \rangle$

$$|\psi_n\rangle \quad = \quad |\tilde{\psi}_n\rangle \quad - \sum |\tilde{\phi}_{lm\epsilon}\rangle c_{lm\epsilon} + \sum |\phi_{lm\epsilon}\rangle c_{lm\epsilon}$$

| AE | pseudo | pseudo-onsite | AE-onsite |

- Same trick works for

  - Wavefunctions
  - Charge density
  - Kinetic energy
  - Exchange correlation energy
  - Hartree energy

The kinetic energy

- For instance, the kinetic energy is given by

$$E_{\mathrm{kin}} = \sum_n f_n \langle \psi_n | -\frac{1}{2}\Delta|\psi_n\rangle$$

- By inserting the transformation ($i = lm\epsilon$)

$$|\psi_n\rangle = |\tilde{\psi}_n\rangle + \sum_i (|\phi_i\rangle - |\tilde{\phi}_i\rangle)\langle\tilde{p}_i|\tilde{\psi}_n\rangle$$

into E_{kin} one obtains: $E_{\mathrm{kin}} = \tilde{E} - \tilde{E}^1 + E^1$ (assuming completeness)

$$\underbrace{\sum_n f_n \langle\tilde{\psi}_n| -\frac{1}{2}\Delta|\tilde{\psi}_n\rangle}_{\tilde{E}} - \underbrace{\sum_{\mathrm{site}\ (i,j)} \rho_{ij}\langle\tilde{\phi}_i| -\frac{1}{2}\Delta|\tilde{\phi}_j\rangle}_{\tilde{E}^1} + \underbrace{\sum_{\mathrm{site}\ (i,j)} \rho_{ij}\langle\phi_i| -\frac{1}{2}\Delta|\phi_j\rangle}_{E^1}$$

- ρ_{ij} is an on-site density matrix:

$$\rho_{ij} = \sum_n f_n \langle\tilde{\psi}_n|\tilde{p}_i\rangle\langle\tilde{p}_j|\tilde{\psi}_n\rangle$$

- For any (quasi) local operator A there exists a PS operator

$$\tilde{A} = A + \sum_{ij} |\tilde{p}_i\rangle \left(\langle \phi_i | A | \phi_j \rangle - \langle \tilde{\phi}_i | A | \tilde{\phi}_j \rangle \right) \langle \tilde{p}_j |$$

so that

$$\langle \psi | A | \psi \rangle = \langle \widetilde{\psi} | \tilde{A} | \widetilde{\psi} \rangle$$

- For instance the PS operator that corresponds to the density operator $|\mathbf{r}\rangle\langle\mathbf{r}|$ is given by

$$|\mathbf{r}\rangle\langle\mathbf{r}| + \sum_{ij} |\tilde{p}_i\rangle \left(\langle \phi_i | \mathbf{r} \rangle \langle \mathbf{r} | \phi_j \rangle - \langle \tilde{\phi}_i | \mathbf{r} \rangle \langle \mathbf{r} | \tilde{\phi}_j \rangle \right) \langle \tilde{p}_j |$$

and the density

$$
\begin{aligned}
\langle \psi | \mathbf{r} \rangle \langle \mathbf{r} | \psi \rangle &= \langle \widetilde{\psi} | \mathbf{r} \rangle \langle \mathbf{r} | \widetilde{\psi} \rangle + \sum_{ij} \langle \widetilde{\psi} | \tilde{p}_i \rangle \left(\langle \phi_i | \mathbf{r} \rangle \langle \mathbf{r} | \phi_j \rangle - \langle \tilde{\phi}_i | \mathbf{r} \rangle \langle \mathbf{r} | \tilde{\phi}_j \rangle \right) \langle \tilde{p}_j | \widetilde{\psi} \rangle \\
&= \widetilde{\rho}(\mathbf{r}) - \widetilde{\rho}^1(\mathbf{r}) + \rho^1(\mathbf{r})
\end{aligned}
$$

- Non-local operators are more complicated

The Hartree energy

- The pseudo-wavefunctions do not have the same norm as the AE wavefunctions inside the spheres
- To deal with long range electrostatic interactions between spheres a soft compensation charge $\hat{\rho}$ is introduced (similar to FLAPW).

| AE | pseudo + compens. | pseudo+comp. onsite | AE-onsite |
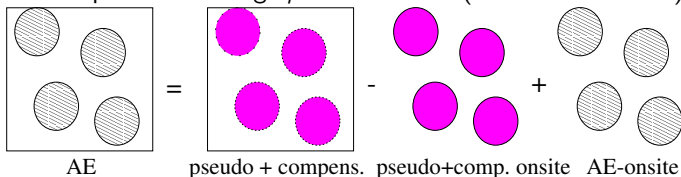
- Hartree energy becomes: $E_H = \tilde{E} - \tilde{E}^1 + E^1$

$$E_H[\tilde{\rho} + \hat{\rho}] - \sum_{\text{sites}} E_H[\tilde{\rho}^1 + \hat{\rho}^1] + \sum_{\text{sites}} E_H[\rho^1]$$

$\tilde{\rho}^1$ one-center pseudo charge charge $\hat{\rho}^1$ one-center compensation

PAW energy functional

Total energy becomes a sum of three terms: $E = \tilde{E} + E^1 - \tilde{E}^1$

$$
\begin{aligned}
\tilde{E} &= \sum_n f_n \langle \tilde{\psi}_n | -\frac{1}{2}\Delta | \tilde{\psi}_n \rangle + E_{xc}[\tilde{\rho} + \hat{\rho} + \tilde{\rho}_c] + \\
& \quad E_H[\tilde{\rho} + \hat{\rho}] + \int v_H[\tilde{\rho}_{Zc}] \left(\tilde{\rho}(\mathbf{r}) + \hat{\rho}(\mathbf{r}) \right) d^3\mathbf{r} + U(\mathbf{R}, Z_{\text{ion}}) \\
\tilde{E}^1 &= \sum_{\text{sites}} \left\{ \sum_{(i,j)} \rho_{ij} \langle \tilde{\phi}_i | -\frac{1}{2}\Delta | \tilde{\phi}_j \rangle + \overline{E_{xc}[\tilde{\rho}^1 + \hat{\rho} + \tilde{\rho}_c]} + \right. \\
& \quad \left. \overline{E_H[\tilde{\rho}^1 + \hat{\rho}]} + \int_{\Omega_r} v_H[\tilde{\rho}_{Zc}] \left(\tilde{\rho}^1(\mathbf{r}) + \hat{\rho}(\mathbf{r}) \right) d^3\mathbf{r} \right\} \\
E^1 &= \sum_{\text{sites}} \left\{ \sum_{(i,j)} \rho_{ij} \langle \phi_i | -\frac{1}{2}\Delta | \phi_j \rangle + \overline{E_{xc}[\rho^1 + \rho_c]} + \right. \\
& \quad \left. \overline{E_H[\rho^1]} + \int_{\Omega_r} v_H[\rho_{Zc}] \rho^1(\mathbf{r}) d^3\mathbf{r} \right\}
\end{aligned}
$$

- $\tilde{E}$ is evaluated on a regular grid

    Kohn-Sham functional evaluated in a plane wave basis set

    with additional compensation charges to account for the incorrect norm of the pseudo-wavefunction (very similar to ultrasoft pseudopotentials).

    $\widetilde{\rho} = \sum_n f_n \widetilde{\psi}_n \widetilde{\psi}_n^*$ pseudo charge density
    $\hat{\rho}$ compensation charge

- E^1 and $\tilde{E}^1$ are evaluated on radial grids centered around each ion.

    Kohn-Sham energy evaluated for basis sets $\{\widetilde{\phi}_i\}$ and $\{\phi_i\}$

    these terms correct for the shape difference between the pseudo and AE wavefunctions.

- No cross-terms between plane wave part and radial grids exist.

- The pseudo wave functions $|\widetilde{\psi}_n\rangle$ (plane waves!) are the self-consistent solutions of

$$\left(-\frac{1}{2}\Delta + \widetilde{V}_{\text{eff}} + \sum_{ij}|\widetilde{p}_i\rangle(D_{ij}+...)\langle\widetilde{p}_j|\right)|\widetilde{\psi}_n\rangle = \epsilon_n\left(1 + \sum_{ij}|\widetilde{p}_i\rangle Q_{ij}\langle\widetilde{p}_j|\right)|\widetilde{\psi}_n\rangle$$

$$D_{ij} = \langle\phi_i| - \frac{1}{2}\Delta + v_{\text{eff}}^1[\rho_v^1]|\phi_j\rangle - \langle\widetilde{\phi}_i| - \frac{1}{2}\Delta + \widetilde{v}_{\text{eff}}^1[\widetilde{\rho}_v^1]|\widetilde{\phi}_j\rangle$$

$$\rho_v^1(\mathbf{r}) = \sum_{ij}\rho_{ij}\langle\phi_i|\mathbf{r}\rangle\langle\mathbf{r}|\phi_j\rangle \qquad \widetilde{\rho}_v^1(\mathbf{r}) = \sum_{ij}\rho_{ij}\langle\widetilde{\phi}_i|\mathbf{r}\rangle\langle\mathbf{r}|\widetilde{\phi}_j\rangle$$

$$\rho_{ij} = \sum_n f_n\langle\widetilde{\psi}_n|\widetilde{p}_i\rangle\langle\widetilde{p}_j|\widetilde{\psi}_n\rangle$$

- If the partial waves form a complete basis within the PAW spheres, then the all-electron wave functions $|\psi_n\rangle$ are orthogonal to the core states!

Accuracy

Subset of G2-1 test set: Deviation PAW w.r.t. GTO, in [kcal/mol].
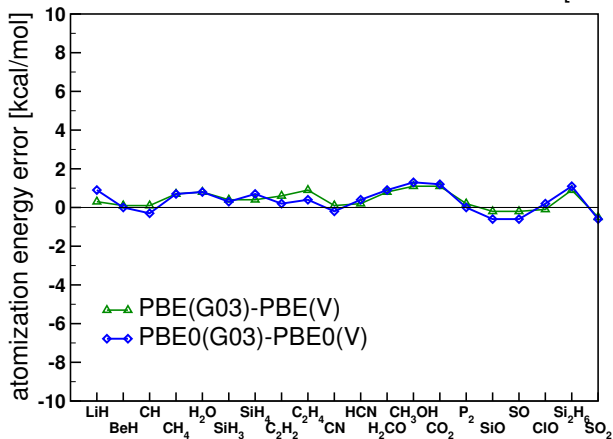
$|\Delta E_{\mathrm{AE}}| < 1$ kcal/mol.

Accuracy

Relative PBE bond lengths of Cl_2, ClF, and HCl for various GTO basis sets specified with respect to plane-wave results:
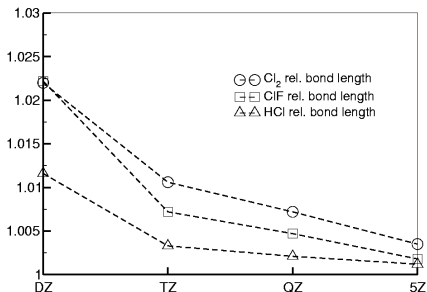
aug-cc-pVXZ (X= D,T,Q,5)
N.B.: aug-cc-pV5Z basis set for Cl contains 200 functions!

$$|\psi_n\rangle = |\tilde{\psi}_n\rangle + \sum_{lm\epsilon} \left(|\phi_{lm\epsilon}\rangle - |\tilde{\phi}_{lm\epsilon}\rangle\right) \langle \tilde{p}_{lm\epsilon}|\tilde{\psi}_n\rangle$$

- $|\tilde{\psi}_n\rangle$ is the variational quantity of the PAW method.
- The PAW method is often referred to as an all-electron method. Not in the sense that all electrons are treated explicitly, but in the sense that the valence electronic wave functions are kept orthogonal to the core states.

AE   pseudo + compens. pseudo+comp. onsite AE-onsite
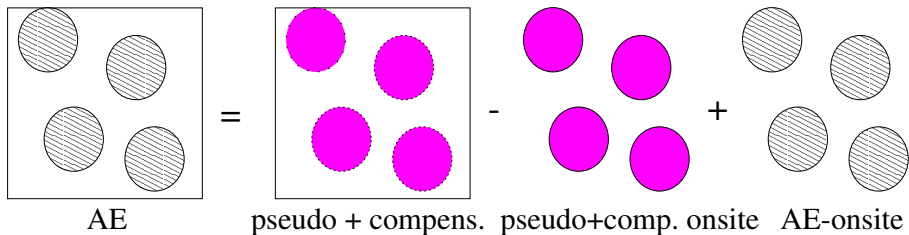
- This general scheme applies to all operators.

- Sometimes one may choose to include only parts of the PAW expressions.

  lazy: only implement plane wave part (GW, ...)
  efficient: physics of localized orbitals; only spheres (LDA+U, DMFT, ...,)

Electronic optimization

Direct minimization of the DFT functional (Car-Parrinello, modern)
Start with a set of wavefunctions $\{\psi_n(\mathbf{r})|n = 1, .., N_e/2\}$ (random numbers) and minimize the value of the functional (iteration)

$$\text{Gradient:} \quad F_n(\mathbf{r}) = \left(-\frac{\hbar^2}{2m_e}\nabla^2 + V^{\text{eff}}(\mathbf{r}, \{\psi_n(\mathbf{r}')\}) - \epsilon_n\right)\psi_n(\mathbf{r})$$

The Self Consistency Cycle (old fashioned)
Start with a trial density ρ, set up the Schrödinger equation, and solve it to obtain wavefunctions $\psi_n(\mathbf{r})$

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2 + V^{\text{eff}}(\mathbf{r}, \{\rho(\mathbf{r}')\})\right)\psi_n(\mathbf{r}) = \epsilon_n\psi_n(\mathbf{r}) \qquad n = 1, ..., N_e/2$$

as a result one obtains a new charge density $\rho(\mathbf{r}) = \sum_n |\psi_n(\mathbf{r})|^2$ and a new Schrödinger equation $\Rightarrow$ iteration

Direct mimimization vs. SCC

disordered diamond, insulator

disordered fcc Fe, metal
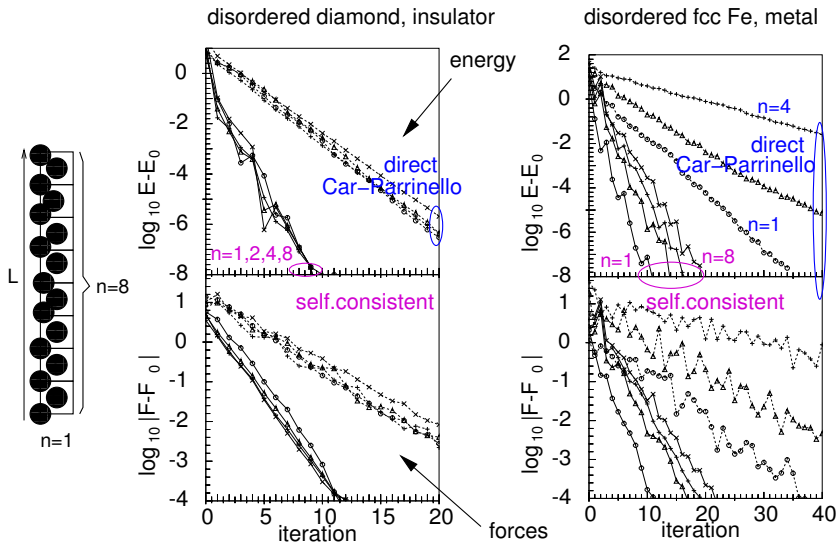
G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11169 (1996).

Direct optimization (charge sloshing)

The derivative of the total energy w.r.t. the wave function $\langle\psi_n|$ is

$$|g_n\rangle = f_n\Big(1 - \sum_m |\psi_m\rangle\langle\psi_m|\Big)\hat{H}|\psi_n\rangle + \sum_m \frac{1}{2}\mathbf{H}_{nm}(f_n - f_m)|\psi_m\rangle$$

where $\mathbf{H}_{nm} = \langle\psi_m|\hat{H}|\psi_n\rangle$.

Consider two states

$\psi_n = e^{i(\mathbf{k}_F - \delta\mathbf{k})\mathbf{r}}$ $\psi_m = e^{i(\mathbf{k}_F + \delta\mathbf{k})\mathbf{r}}$

$\psi_n' = \psi_n + \Delta s\psi_m$ $\psi_m' = \psi_m - \Delta s\psi_n$

slowly varying charge
$4\pi e / G^2$
strong change in potential

$$\delta\rho(\mathbf{r}) = 2\Delta s\,\mathrm{Re}\,e^{i2\delta\mathbf{k}\cdot\mathbf{r}} \qquad \delta V_{\mathrm{H}}(\mathbf{r}) = \frac{2\Delta s\,4\pi e^2}{|2\delta\mathbf{k}|^2}\mathrm{Re}\,e^{i2\delta\mathbf{k}\cdot\mathbf{r}}$$

The smallest $|\delta\mathbf{k}| \propto 1/L$ where L is the largest dimension of the supercell.
$\Rightarrow$ the response of the potential $\delta V_{\mathrm{H}} \propto L^2 \Rightarrow$ stable step size $\Delta s \propto 1/L^2$
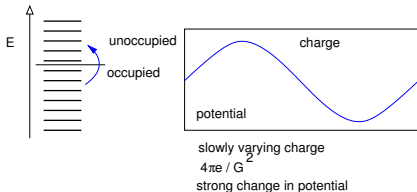
A naively straightforward algorithm

Express the Hamiltonian in the plane wave basis set and diagonalize it

$$\mathbf{H} = \langle \mathbf{G} | \hat{H}[\rho] | \mathbf{G}' \rangle \rightarrow \text{diagonalize } \mathbf{H} \rightarrow \{\psi_i, \epsilon_i\} \ \ i = 1,.., N_{\text{FFT}}$$

Self-consistency

$$\rho_0 \rightarrow \mathbf{H}_0 \rightarrow \rho' \rightarrow \rho_1 = f(\rho_0, \rho') \rightarrow \mathbf{H}_1 \rightarrow ...$$

iterate until $\rho = \rho'$

BUT: we do not need N_{FFT} one-electron orbitals, at a cost of $\mathcal{O}(N_{\text{FFT}}^3)$ ...
we only need the N_b lowest eigenvectors of $\mathbf{H}$

$\Rightarrow$ Iterative diagonalization of $\mathbf{H}$ aimed at finding its N_b lowest eigenvectors
($N_b \approx N_{\text{el}}/\text{cell}$).

Blocked Davidson algorithm, RMM-DIIS, ...

The Self-Consistency Cycle

trial-charge ρ_{in} and trial-wavevectors ψ_n

set up Hamiltonian $H(\rho_{in})$

iterative refinements of wavefunctions $\{\psi_n\}$

new charge density $\rho_{out} = \sum_n f_n |\psi_n(\mathbf{r})|^2$

refinement of density $\rho_{in}, \rho_{out} \Rightarrow$ new ρ_{in}

no $\Delta E < \delta$

calculate forces, update ions
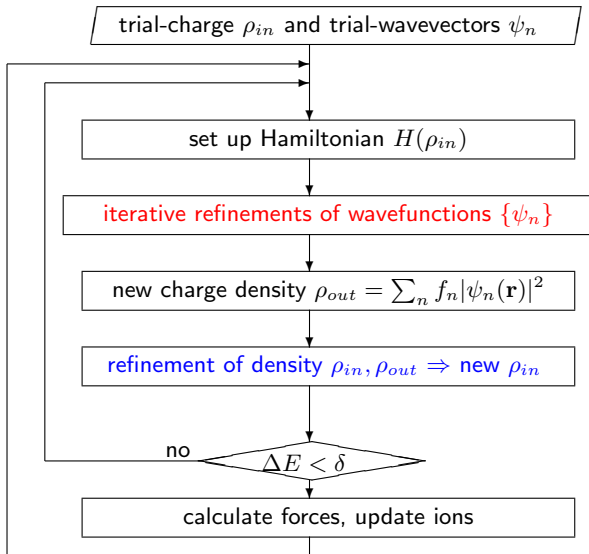
- two subproblems optimization of $\{\psi_n\}$ and ρ_{in}
- refinement of density: DIIS algorithm
  P. Pulay, Chem. Phys. Lett. 73, 393 (1980)
- refinement of wavefunctions: DIIS or Davidson algorithm

Key ingredients: Subspace diagonalization and the Residual

- Rayleigh-Ritz: the diagonalization of the $N_b \times N_b$ subspace

$$\sum_m \bar{H}_{nm} B_{mk} = \sum_m \epsilon_k^{\mathrm{app}} \bar{S}_{nm} B_{mk}$$

with

$$\bar{H}_{nm} = \langle \psi_n | \hat{H} | \psi_m \rangle \qquad \text{and} \qquad \bar{S}_{nm} = \langle \psi_n | \hat{S} | \psi_m \rangle$$

The N_b eigenvalues/eigenvectors $\epsilon_k^{\mathrm{app}}$ and $|\bar{\psi}_k\rangle = \sum_m B_{mk} |\psi_m\rangle$ are the best approximation to the exact N_b lowest eigenvalues of $\hat{H}$ within the subspace spanned by the one-electron orbitals ψ_m.

- The residual vector:

$$|R(\psi_n)\rangle = (\hat{H} - \epsilon_{\mathrm{app}} \hat{S}) |\psi_n\rangle, \qquad \text{with } \epsilon_{\mathrm{app}} = \frac{\langle \psi_n | \hat{H} | \psi_n \rangle}{\langle \psi_n | \hat{S} | \psi_n \rangle}$$

(its norm is a measure of the error in the eigenvector).

The blocked Davidson algorithm

- Take a subset of all bands $\{\psi_n | n = 1, .., N\} \Rightarrow \{\psi_k^1 | k = 1, .., n_1\}$

  - Extend this subset by adding the (preconditioned) residual vectors to the presently considered subspace
    $$\{\psi_k^1 / g_k^1 = \mathbf{K}(\mathbf{H} - \epsilon_{\mathrm{app}}\mathbf{S})\psi_k^1 | k = 1, .., n_1\}$$

  - Raighley-Ritz optimization ("sub-space" rotation) in the $2n_1$ dimensional subspace $\{\psi_k^1 / g_k^1\}$ to determine the n_1 lowest eigenvectors $\{\psi_k^2 | k = 1, .., n_1\}$.
  - Extend the subspace with the residuals from $\{\psi_k^2\}$
    $$\{\psi_k^1 / g_k^1 / g_k^2 = \mathbf{K}(\mathbf{H} - \epsilon_{\mathrm{app}}\mathbf{S})\psi_k^2 | k = 1, .., n_1\}$$

  - Raighley-Ritz optimization $\Rightarrow \{\psi_k^3 | k = 1, .., n_1\}$
  - etc ...
  - The optimized $\{\psi_k^m | k = 1, .., n_1\}$ replace $\{\psi_n | n = 1, .., n_1\}$

- Move on to subset $\{\psi_k^1 | k = n_1 + 1, .., n_2\}$, ..., etc, ...

After treating all bands: Raighley-Ritz optimization of $\{\psi_n | n = 1, .., N\}$

Charge density mixing (RMM-DIIS)

Minimization of the norm of residual vector

$$R[\rho_{\mathrm{in}}] = \rho_{\mathrm{out}}[\rho_{\mathrm{in}}] - \rho_{\mathrm{in}} \qquad |R[\rho_{\mathrm{in}}]| \Rightarrow \mathsf{min}$$

with $\rho_{\mathrm{out}}(\vec{r}) = \sum_{\mathrm{occupied}} w_k f_{nk} |\psi_{nk}(\vec{r})|^2$

DIIS algorithm is used for the optimization of the norm of the residual vector.

Linearization of $R[\rho_{\mathrm{in}}]$ around ρ_{sc} (linear response theory)

$$R[\rho_{\mathrm{in}}] = \mathbf{J}(\rho_{\mathrm{in}} - \rho_{\mathrm{sc}})$$

where $\mathbf{J}$ is the charge dielectric function

$$\mathbf{J} = \mathbf{1} - \mathbf{PU}$$

and $\mathbf{U}$ is the Coulomb kernel

$$\mathbf{U} = \frac{4\pi e^2}{q^2}$$

Provided $\mathbf{J}$ were to be know

$$\rho_{\mathrm{sc}} = \rho_{\mathrm{in}} - \mathbf{J}^{-1} R[\rho_{\mathrm{in}}]$$

Starting from

$$\delta\rho_{\rm tot} = \delta\rho_{\rm ext} + \delta\rho_{\rm ind} \qquad \mathbf{J} = \mathbf{1} - \mathbf{PU}$$

where $\mathbf{P} = \frac{\partial\rho_{\rm ind}}{\partial\nu_{\rm tot}}$ is the irreducible polarizability, we have that

$$\mathbf{J}\delta\rho_{\rm tot} = \delta\rho_{\rm tot} - \frac{\partial\rho_{\rm ind}}{\partial\nu_{\rm tot}}\mathbf{U}\delta\rho_{\rm tot} = \delta\rho_{\rm tot} - \delta\rho_{\rm ind} = \delta\rho_{\rm ext}$$

where $\delta\rho_{\rm ind}$ is limited to the induced charge density due to changes in the Hartree potential only.

From

$$\delta\rho_{\rm tot} = \mathbf{J}^{-1}\delta\rho_{\rm ext}$$

and

$$\delta\rho_{\rm tot} = \rho_{\rm in} - \rho_{\rm sc} = \mathbf{J}^{-1}\delta\rho_{\rm ext}$$

we see that in the linear response regime, where $R[\rho_{\rm in}] \propto (\rho_{\rm in} - \rho_{\rm sc})$,

$$\delta\rho_{\rm ext} = R[\rho_{\rm in}]$$

i.e.,

$$\rho_{\rm sc} = \rho_{\rm in} - \mathbf{J}^{-1}R[\rho_{\rm in}] = \rho_{\rm in} - \mathbf{J}^{-1}(\rho_{\rm out}[\rho_{\rm in}] - \rho_{\rm in})$$

Divergence of the dielectric function

Eigenvalue spectrum of $\mathbf{J}$ determines convergence

$$\mathbf{J} = \mathbf{1} - \mathbf{P} \underbrace{\mathbf{U}}_{\frac{4\pi e^2}{q^2}}$$

"Broader" eigenvalue spectrum $\Rightarrow$ slower convergence

- For insulators and semi-conductors, the width of the eigenvalue spectrum is constant and system size independent (ϵ_∞)!

- For metals the eigenvalue spectrum diverges, its width is proportional to the square of the longest dimension of the cell:

    - Short wavelength limit $\mathbf{J} \approx \mathbf{1}$ (no screening)
    - Long wavelength limit $\mathbf{J} \approx \mathbf{1}/\mathbf{q^2} \propto \mathbf{L^2}$ (metallic screening)

    Complete screening in metals causes charge sloshing

The dielectric matrix

- Use a model dielectric function that is a good initial approximation for most systems

$$\mathbf{J}^{-1} \approx \mathbf{G}_q^1 = \max(\frac{q^2 \texttt{AMIX}}{q^2 + \texttt{BMIX}}, \texttt{AMIN})$$
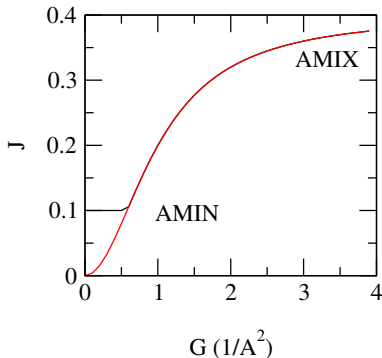
- This is combined with a convergence accelerator.

  The initial guess for the dielectric matrix is improved using information accumulated in each electronic (mixing) step (DIIS).

Return to direct optimization: Why?

Pure DFT functional depends only on the density

$$\left(-\frac{1}{2}\Delta + V_{\text{eff}}[\rho](\mathbf{r}) + V_{\text{ext}}(\mathbf{r})\right)\psi_n(\mathbf{r}) = \epsilon_n\psi_n(\mathbf{r})$$

DFT-Hartree-Fock Hybrid functional depends explicitly on the wave functions

$$\left(-\frac{1}{2}\Delta + V_{\text{eff}}[\rho](\mathbf{r}) + V_{\text{ext}}(\mathbf{r})\right)\psi_n(\mathbf{r}) + C\sum_m^{\text{occ}}\psi_m(\mathbf{r})\int\frac{\psi_m^*(\mathbf{r}')\psi_n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' = \epsilon_n\psi_n(\mathbf{r})$$

so density-mixing will not work (reliably).

Unfortunately we know direct optimization schemes are prone to charge sloshing for metals and small-gap systems.

Mixed scheme

- The gradient of the wave functions is given by

$$|g_n\rangle = f_n\Big(1 - \sum_m |\psi_m\rangle\langle\psi_m|\Big)\hat{H}|\psi_n\rangle + \sum_m \frac{1}{2}\mathbf{H}_{nm}(f_n - f_m)|\psi_m\rangle$$

with $\mathbf{H}_{nm} = \langle\psi_m|\hat{H}|\psi_n\rangle$

- A search direction towards the groundstate w.r.t. unitary transformations between the orbitals within the subspace spanned by wave functions can be found from perturbation theory

$$\mathbf{U}_{nm} = \delta_{nm} - \Delta s \frac{\mathbf{H}_{nm}}{\mathbf{H}_{mm} - \mathbf{H}_{nn}}$$

but this is exactly the term that is prone to charge sloshing!

- Solution: Use density mixing to determine the optimal unitary transformation matrix $\mathbf{U}_{nm}$.

Optimal subspace rotation

- Define a Hamilton matrix $\bar{\mathbf{H}}_{kl} = \langle \psi_l | \bar{H}[\rho] | \psi_k \rangle$
  where $\bar{H}[\rho] = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{eff}}[\rho] + \hat{V}_{\text{X}}^{\text{nl}}[\{\psi\}, \{f\}]$

- Determine the subspace rotation matrix $\mathbf{V}$ that diagonalizes $\bar{\mathbf{H}}_{kl}$

- Recompute the (partial) occupancies $\rightarrow \{f'\}$

- The transformed orbitals $\sum_l \mathbf{V}_{nl} \psi_l$ and partial occupancies $\{f'\}$ define a new charge density ρ'

- mix ρ and ρ'

- and iterate the above until a stable point is found $\rightarrow \rho_{\text{sc}}$

- $\mathbf{H}_{nm}^{\text{sc}} = \langle \psi_m | \bar{H}[\rho_{\text{sc}}] | \psi_n \rangle$ defines the optimal subspace rotation

$$\mathbf{U}_{nm} = \delta_{nm} - \Delta s \frac{\mathbf{H}_{nm}^{\text{sc}}}{\mathbf{H}_{mm}^{\text{sc}} - \mathbf{H}_{nn}^{\text{sc}}}$$

- N.B.: we do not update the orbital dependent part of the Hamiltonian $\hat{V}_{\text{X}}^{\text{nl}}[\{\psi\}, \{f\}]$

The full mixed scheme

The iterative optimization of the wavefunctions cycles through the following steps:

1. construct the Hamiltonian, $\mathbf{H}$, from the current wavefunctions and partial occupancies, and calculate $\mathbf{H}|\psi\rangle$;

2. inner loop: determine the self-consistent Hamiltonian, $\mathbf{H}^{\mathrm{sc}}$, defining the preconditioned direction for the subspace rotation $\mathbf{U}$.

3. minimization along the preconditioned search direction, defined by $(1 - \sum_m |\psi_m\rangle\langle\psi_m|)\hat{H}|\psi_n\rangle$, $\mathbf{U}$, and a gradient acting on the partial occupancies. For instance by means of a conjugate-gradient algorithm.

This loop is repeated until the change in the free energy from one iteration to the next drops below the required convergence threshold ΔE_{thr} (usually 10^{-4} eV).

It works: fcc Fe
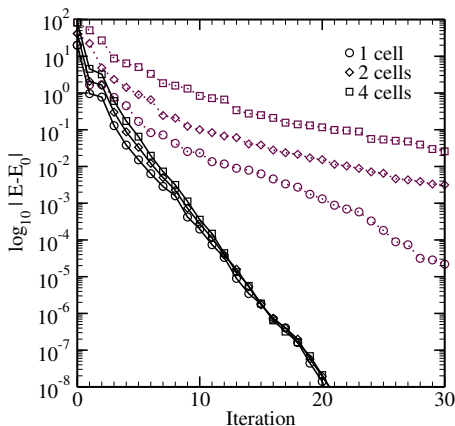
The convergence behaviour of HSE03 calculations using the improved direct minimization procedure (solid lines) and a standard conjugate gradient algorithm (dotted lines). Calculations on single, double, four times, and eight times repeated cells are marked with circles ($\bigcirc$), diamonds ($\diamond$), squares ($\square$), and triangles ($\triangle$), respectively.