

CP2K: the gaussian plane wave (GPW) method

Basis sets and Kohn-Sham energy calculation

R. Vuilleumier

Département de chimie
Ecole normale supérieure – Paris

Tutorial CPMD-CP2K

CPMD and CP2K

CPMD

- ▶ <http://www.cpmd.org>
- ▶ Licensed by IBM and MPI Stuttgart
- ▶ Fortran 77 with extension
- ▶ Plane wave basis set
- ▶ heavy use of FFT
- ▶ possibility of k -point sampling

CP2K

- ▶ <http://cp2k.berlios.de>
- ▶ GNU license
- ▶ high level fortran 90
- ▶ Mixed gaussians and plane waves
- ▶ order N construction of the KS Hamiltonian

The problem

We want to solve **numerically** the self-consistent Kohn-Sham equations:

$$v_{KS}(\vec{r}) = v(\vec{r}) + \int d^3\vec{r}' \frac{n_0(\vec{r}')}{|\vec{r} - \vec{r}'|} + v_{xc}[n_0](\vec{r})$$

$$\text{with } \begin{cases} n_0(\vec{r}) = \sum_{i=1}^N |\varphi_i^0(\vec{r})|^2 \\ -\frac{1}{2}\nabla^2\varphi_i^0(\vec{r}) + v_{KS}(\vec{r})\varphi_i^0(\vec{r}) = \epsilon_i\varphi_i^0(\vec{r}) \end{cases} \quad (1)$$

We need

1. a representation of $\phi_i(\vec{r})$ and $n(\vec{r})$,
2. to be able to compute the total Kohn-Sham energy and its derivatives
3. a method to solve the non-linear the Kohn-Sham equation.

Plane wave vs gaussian basis sets: plane waves pros and cons

Advantages

- ▶ independent of the nuclei position (good for forces)
- ▶ no BSSE
- ▶ one parameter controls the basis set size
- ▶ orthogonal
- ▶ numerical efficiency through use of FFT

Disadvantages

- ▶ large number of basis set elements needed
- ▶ **Necessary use of pseudo-potentials**
- ▶ loss of chemical insight

Border line

- ▶ fill the whole simulation box
- ▶ naturally periodic

Plane wave vs gaussian basis sets: gaussians pros and cons

Advantages

- ▶ Good already for small basis set sizes
- ▶ correspond to chemical insight
- ▶ Computationally efficient (multi-centre integrals)
- ▶ Possibility to perform all-electrons calculations

Disadvantages

- ▶ non-orthogonal
- ▶ atomic position dependent (Pulay forces)
- ▶ Basis set superposition error (BSSE)
- ▶ Systematic improvement not straightforward
- ▶ linear dependencies, over-completeness
- ▶ wrong asymptotic behaviour

Border line

- ▶ no implicit periodicity
- ▶ can be tuned for each application

Plane wave basis set: Definition

Basis set $\{\varphi_\alpha\}$ of plane waves: (Ω is the volume of the box)

$$\varphi_\alpha(\vec{r}) = \frac{1}{\sqrt{\Omega}} e^{i\vec{G}_\alpha \cdot \vec{r}} \quad (2)$$

The wavefunctions are expanded as

$$\phi_i(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} c_i(\vec{G}) e^{i\vec{G} \cdot \vec{r}} \quad (3)$$

For an orthorhombic box with lengths L_x , L_y and L_z , the wavevectors \vec{G} are

$$\vec{G} = i \cdot \frac{2\pi}{L_x} \cdot \vec{x} + j \cdot \frac{2\pi}{L_y} \cdot \vec{y} + k \cdot \frac{2\pi}{L_z} \cdot \vec{z}; \text{ with } i, j, k \in \mathbb{Z}$$

real wavefunctions \Rightarrow only half space is needed

Representation of the electronic density $n(\vec{r})$

$$n(\vec{r}) = \sum_i |\phi_i(\vec{r})|^2 = \frac{1}{\Omega} \sum_i \sum_{\vec{G}\vec{G}'} c_i^*(\vec{G}') c_i(\vec{G}) e^{i(\vec{G}-\vec{G}')\cdot\vec{r}}$$

$$n(\vec{r}) = \frac{1}{\Omega} \sum_{\|\vec{G}\| < 2G_{max}} \tilde{n}(\vec{G}) e^{i\vec{G}\cdot\vec{r}} \quad (4)$$

$$\Rightarrow E_{cut}^n = 4E_{cut}^\phi$$

Auxiliary real space grid

$$\text{Sampling interval } \Delta = \frac{L}{N}$$

$$\rightarrow \text{Nyquist critical frequency } f_c = \frac{2\pi}{2\Delta} \gtrsim \|\vec{G}\|_{\max}$$

$\Rightarrow \Delta$ related to E_{cut}

$X_i = (i - 1)\Delta$ on the real space grid

- ▶ We can go from the real-space grid to the reciprocal-space grid and back using Fourier transform techniques
- ▶ If N is a product of small prime numbers one can use Fast Fourier Transform (FFT) techniques which are very efficient ($N \log N$ instead of N^2)

Density from wavefunction coefficients

Real grid spacing adapted to E_{cut}^n

$$\tilde{n}(\vec{G}) = \frac{\Omega}{N} \sum_{\vec{R}_k} n(\vec{R}_k) e^{-i\vec{G} \cdot \vec{R}_k} = \frac{\Omega}{N} \sum_{\vec{R}_k} \sum_i |\phi_i(\vec{R}_k)|^2 e^{-i\vec{G} \cdot \vec{R}_k} \quad (5)$$

To calculate $\tilde{n}(\vec{G})$ we heavily use the FFT:

$$c_i(\vec{G}) \xrightarrow{(S)INV-FFT} \phi_i(\vec{R}_k) \rightarrow n(\vec{R}_k) = \sum_i |\phi_i(\vec{R}_k)|^2 \xrightarrow{FFT} \tilde{n}(\vec{G})$$

Density derivatives

An important quantity will be $\frac{\partial n(\vec{R}_k)}{\partial c_i(\vec{G})}$

This is

$$\begin{aligned}\frac{\partial n(\vec{R}_k)}{\partial c_i^*(\vec{G})} &= \sum_{\vec{R}_{k'}} \frac{\partial n(\vec{R}_k)}{\partial \phi_i^*(\vec{R}_{k'})} \times \frac{\partial \phi_i^*(\vec{R}_{k'})}{\partial c_i^*(\vec{G})} \\ &= \frac{1}{\sqrt{\Omega}} \phi_i(\vec{R}_k) e^{-i\vec{G} \cdot \vec{R}_k}\end{aligned}$$

Gaussian basis set

Primitive cartesian gaussian functions:

$$g(\vec{r}, \eta, \vec{l}, \vec{A}) = N_c (x - A_x)^{l_x} (y - A_y)^{l_y} (z - A_z)^{l_z} e^{-\eta(\vec{r} - \vec{A})^2} \quad (6)$$

\vec{r} : Electron coordinate

\vec{A} : Atomic coordinate

\vec{l} : Angular momentum

Problems:

- ▶ No cusp at the origin
- ▶ wrong asymptotic behaviour

Contracted cartesian gaussian functions:

$$\varphi_\mu(\vec{r}) = \sum_i d_{\mu i} g_{\mu i}(\vec{r}) \quad (7)$$

CP2K: the GPW method



Joost VandeVondele et al.

Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach.

Computer Physics Communications, 167:103 – 128, 2005.

Two representations of the density:

local basis set:

$$n(\vec{r}) = \sum_{\mu\nu} P^{\mu\nu} \varphi_{\mu}(\vec{r}) \varphi_{\nu}(\vec{r}), \quad (8)$$

$P_{\mu\nu}$: density matrix

and plane waves:

$$n(\vec{r}) = \frac{1}{\Omega} \sum_{|\vec{G}| < G_{max}} \tilde{n}(\vec{G}) \exp(i\vec{G} \cdot \vec{r}) \quad (9)$$

$P^{\mu\nu}$ is the central object to be determined

$\frac{\partial E_{KS}}{\partial P^{\mu\nu}} = K_{\mu\nu}$ is the Kohn-Sham matrix

We need to solve the generalized eigenvalue problem

$$\sum_{\nu} K_{\mu\nu} c_i^{\nu} = \epsilon_i \sum_{\nu} S_{\mu\nu} c_i^{\nu}$$

where $S_{\mu\nu} = \langle \varphi_{\mu} | \varphi_{\nu} \rangle$ is the overlap matrix, with the constraint

$$\sum_{\mu, \nu} c_j^{\mu} S_{\mu\nu} c_i^{\nu} = \delta_{ij}$$

and $P^{\mu\nu} = \sum_{i=1}^{n_{occ}} c_i^{\mu} c_i^{\nu}$

Going from local basis set to plane waves: auxiliary grid

Knowledge of $\tilde{n}(\vec{G})$ is determined by the knowledge of $n(\vec{R}_k)$ on a regular mesh \vec{R}_k :

$$n(\vec{R}_k) = \sum_{\mu\nu} P^{\mu\nu} \varphi_\mu(\vec{R}_k) \varphi_\nu(\vec{R}_k), \quad (10)$$

Here too, heavy use of FFT

Order N construction

Product of two gaussians is a gaussian centered at:

$$\vec{P} = \frac{\eta_a \vec{A} + \eta_b \vec{B}}{\eta_a + \eta_b} \quad (11)$$

with prefactor

$$\exp - \frac{\eta_a \cdot \eta_b}{\eta_a + \eta_b} |\vec{A} - \vec{B}|^2$$

- ▶ mesh points far from \vec{P} are put to zero
- ▶ (μ, ν) pairs far apart are not considered
- ▶ order N achieved using linked-cell method

density derivatives

We will need derivatives for the density: $\frac{\partial n(\vec{R}_k)}{\partial P^{\mu,\nu}}$

$$\frac{\partial n(\vec{R}_k)}{\partial P^{\mu,\nu}} = \varphi_\mu(\vec{R}_k)\varphi_\nu(\vec{R}_k) \quad (12)$$

at the points $\vec{R}_{\mu,\nu}$ where $\varphi_\mu(\vec{R}_k)\varphi_\nu(\vec{R}_k)$ is calculated

Multi-grids

For some products, a lower cutoff the G_{max} can be used. In principle every pair could have a different cutoff; in practice 3 to 5.

$$n(\vec{r}) = \frac{1}{\Omega} \left(\sum_{|\vec{G}| < G_1} \tilde{n}_1(\vec{G}) \exp(i\vec{G} \cdot \vec{r}) + \sum_{|\vec{G}| < G_2} \tilde{n}_2(\vec{G}) \exp(i\vec{G} \cdot \vec{r}) \right) \quad (13)$$

Products with a lower cutoff are colocated on a coarser grid adapted to the lower cutoff. The coarser grid is map on to the finer grid by Fourier transforms (FFT's)

derivatives: chain rule

$$\begin{aligned}\frac{\partial n(\vec{R}_k)}{\partial P^{\mu,\nu}} &= \sum_{\vec{R}'_k} \frac{\partial n(\vec{R}_k)}{\partial n(\vec{R}'_k)} \frac{\partial n(\vec{R}'_k)}{\partial P^{\mu,\nu}} \\ &= \frac{1}{N} \sum_{\vec{R}'_k} \sum_{|\vec{G}| < G_2} \exp i\vec{G} \cdot (\vec{R}_k - \vec{R}'_k) \varphi_\mu(\vec{R}'_k) \varphi_\nu(\vec{R}'_k) \\ &= \frac{1}{\Omega} \sum_{|\vec{G}| < G_2} \exp i\vec{G} \cdot \vec{R}_k \left(\frac{\Omega}{N} \sum_{\vec{R}'_k} \exp -i\vec{G} \cdot \vec{R}'_k \varphi_\mu(\vec{R}'_k) \varphi_\nu(\vec{R}'_k) \right)\end{aligned}$$

Use of FFT

Calculation of the Kohn-Sham energy and its derivatives

$$E[n] = T_s[n] + \int d^3\vec{r} n(\vec{r})v(\vec{r}) + J[n] + E_{xc}[n] + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{\|\vec{R}_I - \vec{R}_J\|}$$

Kinetic energy

$\langle \varphi_\mu | \frac{1}{2} \nabla^2 | \varphi_\nu \rangle$ is analytic

Goedecker-Tetter-Hutter (GTH) pseudopotentials

Local part

$$\begin{aligned}v_{loc}^{PP}(r) &= v_{loc}^{LR}(r) + v_{loc}^{SR}(r) \\ &= -\frac{Z}{r} \operatorname{erf}(\alpha r) + \sum_{i=1}^4 C_i^{PP} (\sqrt{2}\alpha r)^{2i-2} \times \exp[-(\alpha r)^2]\end{aligned}$$

Non-local part

$$v_{nl}^{SR} = \sum_{lm} \sum_{ij} \langle \vec{r} | p_i^{lm} \rangle h_{ij}^l \langle p_j^{lm} | \vec{r} \rangle$$

with the projectors

$$\langle \vec{r} | p_i^{lm} \rangle = N_i^l Y^{lm}(\hat{r}) r^{l+2i-2} \exp\left[-\frac{1}{2} \left(\frac{r}{r_l}\right)^2\right]$$

all short range terms are analytic in the local basis set representation

Electrostatic contribution

$$E_{ES} = \frac{1}{2} \int_{\Omega} \int d^3\vec{r} d^3\vec{r}' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \sum_I \int_{\Omega} d^3\vec{r} v_I^c(\vec{r})n(\vec{r}) \\ + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|}$$

Using the Fourier transformed representation and with PBC: just like previously

$$\frac{1}{2} \int_{\Omega} \int d^3\vec{r} d^3\vec{r}' \frac{n_{tot}(\vec{r})n_{tot}(\vec{r}')}{|\vec{r} - \vec{r}'|} = 2\pi\Omega \sum_{\vec{G} \neq \vec{0}} \frac{\tilde{n}_{tot}(\vec{G})^2}{G^2}$$

$$\tilde{n}_{tot}(\vec{G}) = \tilde{n}(\vec{G}) + \sum_I \tilde{n}_I^c(\vec{G}) \exp(-i\vec{G} \cdot \vec{R}_I)$$

Exchange-correlation energy

$$E_{xc}[n] = \int_{\Omega} d^3\vec{r} \epsilon_{xc}(n(\vec{r}))n(\vec{r})$$

is, just like in CPMD, **approximated** by

$$E_{xc}[n] = \frac{\Omega}{N} \sum_{\vec{R}_k} \epsilon_{xc}(n(\vec{R}_k))n(\vec{R}_k)$$

Differences with cpmd

- ▶ Derivatives of the energy by Fourier transformed or smoothed finite difference
- ▶ Density is first smoothed using nearest-neighbours (reducing the "ripples")

Smoothing operator

$$(S^q f)_{i,j,k} = \frac{q^3}{q^3 + 6q^2 + 12q + 8} \times \sum_{l=-1}^1 \sum_{m=-1}^1 \sum_{n=-1}^1 q^{-|l|-|m|-|n|} f_{i+l,j+m,k+n}$$

- ▶ $q = 10$ to 50
- ▶ Chain rule used again for derivatives \Rightarrow smoothing of the xc potential

Derivatives of the electrostatic and xc contributions

$$\frac{\partial(E_{ES} + E_{xc})}{\partial n(\vec{R}_k)} = \frac{\Omega}{N} (V_H(\vec{R}_k) + V_{xc}(\vec{R}_k))$$

$V_H(\vec{R}_k)$ obtained by FFT from $\tilde{V}_H(\vec{G})$

From chain rule we obtain

$$\frac{\partial(E_{ES} + E_{xc})}{\partial P^{\mu\nu}} = \frac{\Omega}{N} \sum_{\vec{R}_k} (V_H(\vec{R}_k) + V_{xc}(\vec{R}_k)) \varphi_\mu(\vec{R}_k) \varphi_\nu(\vec{R}_k)$$

Optimization

Two ways of solving the Kohn-Sham equation:

1. Reach for self-consistency (traditional scheme)
 - ▶ Obtain lowest occupied orbitals from diagonalization of the KS matrix for a test density
 - ▶ Density from occupied orbitals
 - ▶ Update test density until self-consistency
2. Minimization of the energy functional
 - ▶ Use a non-linear minimization scheme to minimize the KS energy

CP2K (1): traditional scheme – Diagonalization

Given the Kohn-Sham Hamiltonian $K_{\mu,\nu}$ constructed previously, we need to solve

$$\mathbf{K}\mathbf{c} = \mathbf{S}\mathbf{c}\epsilon \quad (14)$$

where, \mathbf{c} are the wavefunction coefficient in the local basis set and \mathbf{S} the overlap matrix. ϵ is the diagonal matrix of eigenvalues.

\mathbf{S} in the local basis set can be calculated analytically.

Direct diagonalization

Since the number of local basis set elements is not prohibitively large (up to a few thousands) the diagonalization can be done directly using linear algebra packages (lapack or scalapack for parallelized versions)

The generalized eigenvalue problem is first transformed to a normal eigenvalue problem introducing $\mathbf{U}^T \mathbf{U} = \mathbf{S}$:

$$\mathbf{K}' \mathbf{c}' = \mathbf{c}' \epsilon \quad (15)$$

with

$$\begin{cases} \mathbf{K}' = (\mathbf{U}^T)^{-1} \mathbf{K} (\mathbf{U})^{-1} \\ \mathbf{c}' = \mathbf{U} \mathbf{c} \end{cases} \quad (16)$$

Cholesky decomposition or symmetric decomposition $\mathbf{U} = \mathbf{S}^{\frac{1}{2}}$

Update of the density matrix \mathbf{P}

A new density matrix is obtained: $(\mathbf{P}_{out})_{\mu,\nu} = \sum_{n=1}^{N_{occ}} c_{n,\mu}^* c_{n,\nu}$

The resulting density is used to update the test density matrix.

Mixing scheme

$$\mathbf{P}_{in}^{i+1} = \alpha \mathbf{P}_{out}^i + (1 - \alpha) \mathbf{P}_{in}^i \quad (17)$$

Orbital transformation: direct minimization in CP2K

Use of an orbital transformation:

$$\mathbf{c}(\mathbf{x}) = \mathbf{c}_0 \cos U + \mathbf{x} U^{-1} \sin U \quad (18)$$

with $\mathbf{c}_0^T \mathbf{S} \mathbf{c}_0 = \mathbf{I}$ and the linear constraint $\mathbf{c}_0^T \mathbf{S} \mathbf{x} = 0$ and $U = \mathbf{x}^T \mathbf{S} \mathbf{x}$

Forces

$$E[n] = T_s[n] + \int d^3\vec{r} n(\vec{r})v(\vec{r}) + J[n] + E_{xc}[n] + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{\|\vec{R}_I - \vec{R}_J\|}$$

We need to calculate $\nabla_I E[n]$ at the optimum density n .

Three types of terms:

1. v depends explicitly on R_I : term $\int d^3\vec{r} n(\vec{r})\nabla_I v(\vec{r})$
2. the basis set elements φ_μ depend on R_I : term $2 \sum_{\mu,\nu} P^{\mu,\nu} \langle \nabla_I \varphi_\mu | \frac{\partial E}{\partial n} | \varphi_\nu \rangle$
3. $P^{\mu,\nu}$ itself depends on R_I : term $\sum_{\mu,\nu} K_{\mu,\nu} \nabla_I P^{\mu,\nu}$

If the overlap matrix is independent of R_I (e.g. plane waves), the last term vanishes. Otherwise, extra force: **Pulay forces**