Gaussian and Plane Waves Method (GPW)

Jürg Hutter

Department of Chemistry, University of Zurich
References


• Lippert, G; Hutter, J; Parrinello, M. Molecular Physics, 92, 477-487 (1997). *A hybrid Gaussian and plane wave density functional scheme.* http://dx.doi.org/10.1080/002689797170220


Kohn–Sham (KS) Density Functional Theory (DFT)

- **Hohenberg–Kohn Theorem I**
  *The total energy can be calculated from a universal functional of the density.*
  Existence theorem

- **Hohenberg–Kohn Theorem II**
  *The correct density minimizes the energy functional.*
  Variational principle.

- **Kohn–Sham Theorem**
  *For each density exists a system of non-interacting particles in a local external potential with the same density.*
  Orbital picture.
Kohn–Sham DFT (II)
Minimization with constraints

Min \( E(\{\Phi_i\}) \)
\[
\int \Phi_i(r)\Phi_j(r)\,dr = \delta_{ij} \quad \int \rho(r)\,dr = N
\]

Energy functional

\[
E(\{\Phi_i\}) = -\sum_i f_i \int \Phi_i \nabla_i \Phi_i \,dr + \int V_{\text{ext}}(r)\rho(r)\,dr
\]
\[
+ \frac{1}{2} \int \int \frac{\rho(r)\rho(r')}{|r - r'|} \,dr \,dr' + E_{xc}[\rho]
\]

\[
\rho(r) = \sum_i f_i |\Phi_i(r)|^2
\]
Kohn–Sham DFT (III)

- **Kinetic energy**
  \[ E_{\text{kin}} = -\frac{1}{2} \sum_i f_i(\Phi_i|\nabla^2|\Phi_i) \]

- **External energy (electron-nuclei interaction)**
  \[ E_{\text{ext}} = \int V_{\text{ext}}(r)\rho(r) \, dr \]

- **Hartree (classical Coulomb) energy**
  \[ E_{\text{H}} = \int \int \frac{\rho(r)\rho(r')}{|r-r'|} \, dr' \, dr \]

- **Exchange-correlation (non-classical Coulomb) (XC) energy**
  \[ E_{\text{xc}} = \int F[\rho] \, dr \]

- **Orbital orthogonality constraint**
  \[ (\Phi_i|\Phi_j) = \delta_{ij} \]

- **Electron number constraint**
  \[ \sum_i f_i(\Phi_i|\Phi_i) = N \]
Linear Combination of Atomic Orbitals (LCAO)

Basis set

\[ \Phi_i(r) = \sum_\alpha c_{\alpha i} \varphi_\alpha(r) \]

Overlap

\[ S_{\alpha\beta} = \int dr \varphi_\alpha^*(r) \varphi_\beta(r) \]

Orthogonality

\[ \int dr \Phi_i^*(r) \Phi_j(r) = \sum_{\alpha\beta} c_{\alpha i}^* S_{\alpha\beta} c_{\beta j} = \delta_{ij} \]

Density matrix

\[ P_{\alpha\beta} = \sum_i f_i c_{\alpha i} c_{\beta i}^* \]

Density

\[ \rho(r) = \sum_{\alpha\beta} P_{\alpha\beta} \varphi_\alpha(r) \varphi_\beta^*(r) \]

Energy

\[ E = \text{Min}_c[E_{\text{kin}}(c) + E_{\text{ext}}(\rho) + E_H(\rho) + E_{\text{xc}}(\rho)] \]
Gaussian Type Orbitals (GTO): General

- **Primitive function**

\[
\varphi(r) = r^l Y_{lm}(\hat{r}) \exp[-\alpha(r - A)^2]
\]

- **Contracted GTO**

\[
\chi(r) = \sum_k d_k \varphi_k(r)
\]

Contraction coefficients \(d_k\) and exponents are fixed. Contraction over functions with same angular momentum.
Gaussian Type Orbitals: Advantages

- GTO’s are "atomic orbital-like"

- Compact basis set (approx. 15-25 functions per atom)

- Analytic integration possible for many operators.

- Optimal for regular grids. Fourier transform is again a Gaussian.

- Compact support (finite extend)
Gaussian Type Orbitals: Disadvantages

- Non-orthogonal basis
- Linear dependencies for larger basis sets
- Complicated to generate and no easy way to improve
- Basis set superposition error (BSSE)
- Molecules (wavefunction tails) and solids have different requirements
KS-DFT with GTO Basis

- Kinetic energy integrals: analytic
- External potential integrals: analytic
- Coulomb: 4 center electron repulsion integrals (ERI) Mulliken notation \((\alpha\beta|\gamma\delta)\): analytic
  CPU and memory bottleneck
- XC energy and integrals: numerical integration
- Overlap integrals: analytic
Hartree Energy

- **Goal:** Avoid calculation of ERI
- Combine all electrostatic energy terms

**Electrostatic Energy**

\[
E_{\text{Coulomb}} = \frac{1}{2} \int \int \frac{\rho(r)\rho(r')}{|r - r'|} \, dr \, dr' \quad \text{electron-electron interaction}
\]

\[- \sum_A Z_A \int \frac{\rho(r)}{|r - R_A|} \, dr \quad \text{electron-core interaction}\]

\[+ \sum_{A<B} \frac{Z_A Z_B}{|R_A - R_B|} \quad \text{ion-ion interaction}\]
Hartree Energy

Total charge density: electronic charge + Gaussian atomic charges:

\[ \rho_{\text{tot}}(r) = \rho_e(r) + \sum_A \rho_A(r) \]

\[ \rho_A(r) = Z_A \left( \frac{\alpha}{\pi} \right)^{3/2} \exp(-\alpha(r - A)^2) \]

\[ E_{\text{Coulomb}} = \frac{1}{2} \int \int \frac{\rho_{\text{tot}}(r) \rho_{\text{tot}}(r')}{|r - r'|} \, dr \, dr' \]

long range interaction

\[ - \sum_A Z_A \int \frac{\text{erfc}(\alpha(r - R_A))}{|r - R_A|} \rho_e(r) \, dr \]

short range interaction

\[ + \sum_{A < B} E_{\text{pair}}(R_A - R_B) \]

short range pair interaction

\[ - \sum_A E_{\text{self}} \]

self interaction correction
Periodic Boundary Condition (PBC)

- Optimal for condensed phase systems (avoids interface effects)
- Bloch states, Brillouin zone sample, see k-points
- Energy per simulation cell
- $\Gamma$ point simulation (Integration with single point at (0,0,0))
- $(\alpha|\mathcal{O}|\beta) \rightarrow \sum_L (\alpha(0)|\mathcal{O}|\beta(L))$
Plane Waves (PW)

Definition

\[ \varphi(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \exp[i\mathbf{G} \cdot \mathbf{r}] \]

+ orthogonal

+ independent of atomic positions

± naturally periodic

– many functions needed
Computational Box

- Box matrix: \( \mathbf{h} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3] \)

- Box volume: \( \Omega = \det \mathbf{h} \)
Lattice Vectors

- **Direct lattice** \( h = [a_1, a_2, a_3] \)
- Direct lattice vectors: \( L = i \cdot a_1 + j \cdot a_2 + k \cdot a_3 \)
- **Reciprocal lattice** \( 2\pi (h^t)^{-1} = [b_1, b_2, b_3] \)

\[ b_i \cdot a_j = 2\pi \delta_{ij} \]

- Reciprocal lattice vectors: \( G = i \cdot b_1 + j \cdot b_2 + k \cdot b_3 \)

Direct and reciprocal space are connected by Fourier transforms.

The expansion of the periodic part of the functions defined in real space includes only the \( G \) that satisfy the PBC: Fourier decomposition.
Properties of Plane Waves

\[ \varphi_G(r) = \frac{1}{\sqrt{\Omega}} \exp[iG \cdot r] \]

- Plane waves are periodic wrt. box \( h \)
- Plane waves are orthonormal
  \[ \langle \varphi_{G'} | \varphi_G \rangle = \delta_{G',G} \]
- Plane waves are complete

\[ \psi(r) = \psi(r + L) = \frac{1}{\sqrt{\Omega}} \sum_G \psi(G) \exp[iG \cdot r] \]
Cutoff: Finite Basis Set

\[ \frac{1}{2} G^2 \leq E_{\text{cut}} \]

\[ N_{\text{PW}} \approx \frac{1}{2\pi^2} \Omega E_{\text{cut}}^{3/2} [\text{a.u.}] \]

Basis set size depends on volume of box and cutoff only
Real Space Grid

**Sampling Theorem:** Interval $\Delta = \frac{L}{N}$; Nyquist critical frequency $f_c = \frac{1}{2\Delta}$

For a given plane wave cutoff (frequency) there is a minimum number of equidistant real space grid points needed for the same accuracy.

Real space grid: $R_i = (i - 1)\Delta$

**Fast Fourier Transform (FFT)**

$$\psi(G) \longleftrightarrow \psi(R)$$

Information contained in $\psi(G)$ and $\psi(R)$ are equivalent.

- Fourier transform: $N^2$ operations
- Fast Fourier transform: $N \log[N]$ operations
Integrals

\[ I = \int_{\Omega} A^*(\mathbf{r}) B(\mathbf{r}) d\mathbf{r} \]

\[ = \sum_{\mathbf{G}G'} A^*(\mathbf{G}) B(\mathbf{G}') \int \exp[-i\mathbf{G} \cdot \mathbf{r}] \exp[i\mathbf{G}' \cdot \mathbf{r}] d\mathbf{r} \]

\[ = \sum_{\mathbf{G}G'} A^*(\mathbf{G}) B(\mathbf{G}') \Omega \delta_{\mathbf{G}\mathbf{G}'} = \Omega \sum_{\mathbf{G}} A^*(\mathbf{G}) B(\mathbf{G}) \]

Parseval’s theorem

\[ \Omega \sum_{\mathbf{G}} A^*(\mathbf{G}) B(\mathbf{G}) = \frac{\Omega}{N} \sum_{i} A^*(\mathbf{R}_i) B(\mathbf{R}_i) \]

Integrals in real space and in reciprocal space are equivalent.
Long Range Term in Coulomb Energy

\[ E_{LRT} = \frac{1}{2} \int \int \frac{\rho_{tot}(r)\rho_{tot}(r')}{|r - r'|} \, dr \, dr' = \int V_H(r) \rho_{tot}(r) \, dr \]

where \( V_H(r) \) is the solution of Poisson equation

\[ \nabla^2 V_H(r) = -4\pi \rho_{tot}(r) \]

Plane wave expansion of total charge density

\[ \rho_{tot}(r) = \sum_G \rho_{tot}(G) e^{iG\cdot r} \quad V_H(G) = 4\pi \frac{\rho_{tot}(G)}{G^2} \]

\[ E_{LRT} = \frac{2\pi}{\Omega} \sum_G \frac{\rho_{tot}^*(G) \rho_{tot}(G)}{G^2} \]
Exchange-Correlation (XC) Functionals

\[ E_{xc} = \int dr \, \varepsilon_{xc}(r) \rho(r) = \Omega \sum_G \varepsilon_{xc}(G) \rho^*(G) \]

\( \varepsilon_{xc}(G) \) is not local in G space. Calculation in real space requires very accurate integration scheme.

New definition of \( E_{xc} \)

\[ E_{xc} = \frac{\Omega}{N_x N_y N_z} \sum_R \varepsilon_{xc}(R) \rho(R) = \Omega \sum_G \tilde{\varepsilon}_{xc}(G) n(G) \]

where \( \tilde{\varepsilon}_{xc}(G) \) is the finite Fourier transform of \( \varepsilon_{xc}(R) \).
Energy and Force of He Atom

Only translations by a multiple of the grid spacing do not change the total energy. This introduces a small modulation of the energy hyper surface, known as "ripples".
KS-DFT GPW Energy

\[ E_{\text{GPW}}^{\text{KS}} = E_{\text{kin}}(P) + \delta E_{\text{ext}}(P) + E_{\text{xc}}(\tilde{\rho}) + E_{\text{H}}(\tilde{\rho}) + E_{\text{ovrl}} - E_{\text{self}} \]

Gaussian orbital part:
\[ \Phi_i(r) = \sum_\alpha c_{\alpha i} \varphi_\alpha \]
\[ P_{\alpha \beta} = \sum_i f_i c_{\alpha i} c_{\beta i} \]

PW part:
\[ \tilde{\rho}(G) = \sum_{\alpha \beta} (\varphi_\alpha \cdot \varphi_\beta)(G) \]
\[ \tilde{\rho}_{\text{tot}}(G) = \tilde{\rho}(G) + \sum_A \rho_A(G) \]
\[ (\varphi_\alpha \cdot \varphi_\beta)(G) = \varphi_{\alpha \beta}(G) \]

- \( E_{\text{GPW}}^{\text{KS}} \) is variational in the GTO coefficients \( c_{\alpha i} \) alone.
- \( \tilde{\rho}(G) \) is a function of \( c_{\alpha i} \) and the auxiliary PW basis
Efficient Calculation of GPW Energy

Screening

- Always work with primitive Gaussians
- Analytic integrals $\rightarrow$ distance screening with $R = A - B$
  
  Overlap $S_{\alpha\beta}$
  $\varphi_{\alpha}(r - A) \leftrightarrow \varphi_{\beta}(r - B)$
  
  $\downarrow$ sparsity pattern
  $T_{\alpha\beta}$

- Density on the real space grid
  
  $\sum_{\alpha\beta} P_{\alpha\beta} \varphi_{\alpha}(R) \varphi_{\beta}(R) \xrightarrow{\text{FFT}} \tilde{\rho}(G)$
  
  $\downarrow$ overlap screening
  $P_{\alpha\beta}$ is only needed with $S_{\alpha\beta}$ sparsity pattern

- $\varphi_{\alpha\beta}(R) \neq 0$ distance (radial) screening
Screening

• All individual screening thresholds can be controlled by
  \texttt{EPS_DEFAULT}

  \texttt{CP2K_INPUT / FORCE_EVAL / DFT / QS}

• Problems associated with thresholds
  
  • Failure in Cholesky decomposition of overlap matrix
  
  • Combination of basis set condition number and too big \texttt{EPS_DEFAULT}

  • Inaccurate charge on real space grid

  • Too low PW cutoff and/or too big \texttt{EPS_DEFAULT} (extend of $\varphi_{\alpha\beta}$)
Finite cutoff and computational box define a real space grid $\{\mathbb{R}\}$
Gaussians and Plane Waves

\[ \sqrt{\frac{\alpha}{\pi}} \exp\left(-\alpha r^2\right) \xrightarrow{\text{FFT}} \exp\left(-\frac{G^2}{4\alpha}\right) \]

- Efficient screening in R space
- Exponential convergence for integration
For the integration of a Gaussian function with exponent 1 an accuracy of $10^{-10}$ requires an integration range of 10 bohr, a cutoff of 25 Rydberg, resulting in 22 integration points.

$\approx 5000$ integration points/integral batch
All grids are commensurable in G-space
PW Cutoff

• Density expansion

PW cutoff and multigrid settings determine accuracy and efficiency of density expansion

&MGRID . . . &END MGRID section

• CUTOFF: Maximal cutoff used in the calculation (default: 280 Ry)

• REL_CUTOFF: Minimal cutoff used for Gaussian with exponent of 1 (default: 40 Ry)

• NGRIDS: Total number of real space grids (cutoffs) used (default: 4)

• PROGRESSION_FACTOR: factor used for cutoff reduction in multigrids (default: 3)

• MULTIGRID_SET: T/F set multigrid cutoff from input (default: F)

• MULTIGRID_CUTOFF: list of cutoffs for N grids
PW Cutoff

- XC functional

Accuracy of density expansion and total PW cutoff determine XC energy accuracy

CP2K calculates gradient of density from plane wave expansion of density

Fourier interpolation may lead to negative densities in low density regions (problem is enhanced by multigrids!)

LDA vs. GGA vs. Meta-functionals

$\rho$ vs. $(\nabla \rho)^2 / \rho^{4/3}$ vs. $\tau$

DENSITY_CUTOFF, GRADIENT_CUTOFF, TAU_CUTOFF
in FORCE_EVAL / DFT / XC

See more advanced options
in FORCE_EVAL / DFT / XC / XC_GRID
Coulomb Potential

\[
P \rightarrow \rho(R) \xrightarrow{\text{FFT}} \rho(G) \rightarrow V_H(G) = \frac{\rho(G)}{G^2} \xrightarrow{\text{FFT}} V_H(R) \rightarrow V
\]

\[O(n \log n)\]

\[\rho(R) = \sum_{\mu\nu} P_{\mu\nu} \chi_\mu(R) \chi_\nu(R) = \sum_{\mu\nu} P_{\mu\nu} \bar{\chi}_{\mu\nu}(R)\]

\[V_{\mu\nu} = \sum_R V(R) \chi_\mu(R) \chi_\nu(R) = \sum_R V(R) \bar{\chi}_{\mu\nu}(R)\]

Efficient screening of sums using \(\bar{\chi}_{\mu\nu}(R)\).
Accuracy of Plane Wave Expansion

Coulomb Energy

Grid spacing [Å]

Plane wave cutoff [Ry]
Accuracy of Plane Wave Expansion

XC Energy

PBE functional, Bulk Silicon

![Graph showing the accuracy of plane wave expansion for XC Energy with a PBE functional using Bulk Silicon.]
Accuracy and Numerical Errors

64 water, 2560 basis functions, LDA functional, 24 cores

<table>
<thead>
<tr>
<th>EPS_DEFAULT</th>
<th>cutoff(Ry)</th>
<th>ngrids</th>
<th>time(s)</th>
<th>Energy(Ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-12</td>
<td>280(30)</td>
<td>4</td>
<td>1.5</td>
<td>x.0377660911</td>
</tr>
<tr>
<td>-12</td>
<td>400(60)</td>
<td>4</td>
<td>2.7</td>
<td>x.0368292349</td>
</tr>
<tr>
<td>-12</td>
<td>400(60)</td>
<td>1</td>
<td>21.9</td>
<td>x.0368292282</td>
</tr>
<tr>
<td>-12</td>
<td>800(60)</td>
<td>6</td>
<td>3.0</td>
<td>x.0371244786</td>
</tr>
<tr>
<td>-12</td>
<td>800(60)</td>
<td>4</td>
<td>3.0</td>
<td>x.0371244689</td>
</tr>
<tr>
<td>-12</td>
<td>800(60)</td>
<td>1</td>
<td>76.5</td>
<td>x.0371244096</td>
</tr>
<tr>
<td>-8</td>
<td>1600(60)</td>
<td>6</td>
<td>3.7</td>
<td>x.0371421086</td>
</tr>
<tr>
<td>-10</td>
<td>1600(60)</td>
<td>6</td>
<td>4.7</td>
<td>x.0371296795</td>
</tr>
<tr>
<td>-12</td>
<td>1600(60)</td>
<td>6</td>
<td>4.7</td>
<td>x.0371288794</td>
</tr>
<tr>
<td>-14</td>
<td>1600(60)</td>
<td>6</td>
<td>4.9</td>
<td>x.0371287675</td>
</tr>
</tbody>
</table>
Basis Set Superposition Error (BSSE)

- Localized non-orthogonal AO basis sets are incomplete

- Local 'completeness' of basis depends on position of atoms

- Basis set is more complete for molecular clusters than single molecules
  → overstabilization of bound clusters

- Overestimation of binding energies (dimers, molecules on surfaces, etc.)

- Counterpoise Correction: Estimation of BSSE

\[ E_{BSSE} \approx E_A(A + B) + E_B(A + B) - E_A(A) - E_B(B) \]

with fragments \( A, B \) at their cluster geometry.
BSSE in liquid water

Binding energy in water (BSSE)

![Graph showing BSSE in liquid water with different basis sets: DZVP, TZV2P, QZV2P, and QZV3P(f,d).]
Non-Periodic Calculations using PW

Solving Poisson equation for non-periodic boundary conditions

- **Analytic** for spherical cutoff or cylindrical or 1-d cutoff
  Marx and Hutter, Ab initio molecular dynamics, NIC Series

- **Wavelet** solvers
  Genovese et al, JCP 2006, 125 074105

- Solvers by **Martyna and Tuckerman**
  Martyna & Tuckerman, JCP 1999, 110 2810-2821
Pseudopotentials

Why Pseudopotentials?

- **Reduction of basis set size**
  effective speedup of calculation

- **Reduction of number of electrons**
  reduces the number of degrees of freedom

- **Inclusion of relativistic effects**
  relativistic effects can be included "partially" into effective potentials
Frozen Core Approximation

- Replace inactive electronic degrees of freedom in the Hamiltonian by an effective potential
- The potential should be additive and transferable
  - additive most general choice: atomic pseudopotentials
  - transferable remove only core electrons
- Core electrons are chemically inert
- Core/Valence separation is often not clear in plane wave calculations: core = all filled shells
- Core wavefunctions are transferred from atomic reference calculation
- Core electrons of different atoms do not overlap
Remaining Problems

- Valence wavefunctions have to be orthogonalized to core states
  → nodal structures → high plane wave cutoff
- Pseudopotential should produce node-less functions and include Pauli repulsion
- Pseudopotential replaces Hartree and XC potential due to the core electrons
- XC functionals are not linear: approximation

\[ E_{XC}(\rho_c + \rho_v) = E_{XC}(\rho_c) + E_{XC}(\rho_v) \]

This assumes that core and valence electrons do not overlap. This restriction can be overcome with the "non-linear core correction".
General Recipe

1. Atomic all–electron calculation (reference state) ⇒ \( \Phi_i^Y(r) \) and \( \epsilon_i \).

2. Pseudize \( \Phi_i^Y \Rightarrow \Phi_i^{PS} \)

3. Calculate potential from

\[
(T + V_i(r)) \Phi_i^{PS}(r) = \epsilon_i \Phi_i^{PS}(r)
\]

4. Calculate pseudopotential by unscreening of \( V_i(r) \)

\[
V_i^{PS}(r) = V_i(r) - V_H(n_{PS}) - V_{XC}(n_{PS})
\]

\( V_i^{PS} \) is state dependent!
Pseudization of Valence Wavefunctions

Smooth continuation of orbitals inside a cutoff radius

\[ u_p(r) \]
\[ u(r) \]
\[ V_p(r) \]
\[ V(r) \]
Semi-local Pseudopotentials

\[ V_{PS}(r, r') = \sum_{L=0}^{\infty} V_{PS}^{L}(r) |Y_{L}\rangle \langle Y_{L}| \]

Separation of Local and Nonlocal Parts

Approximation: all potentials with \( L > L_{\text{max}} \) are equal to \( V_{\text{loc}}^{PS} \)

\[ V_{PS}(r, r') = \sum_{L=0}^{L_{\text{max}}} (V_{PS}^{L}(r) - V_{\text{loc}}^{PS}(r)) |Y_{L}\rangle \langle Y_{L}| + V_{\text{loc}}^{PS}(r) \]

Final Form

\[ V_{PS}(r, r') = V_{\text{loc}}^{PS}(r) + \sum_{L=0}^{L_{\text{max}}} \Delta V_{PS}^{L}(r) |Y_{L}\rangle \langle Y_{L}| \]

- Local pseudopotential \( V_{\text{loc}}^{PS}(r) \)
- Non-local pseudopotential \( \Delta V_{PS}^{L}(r) \)
Silicon: Radial densities

\[ 4_p r^2 n(r) \text{ (bohr}^{-3}) \]

- Blue solid line: pseudo valence
- Red dashed line: true core

\( r \) (bohr) from 0 to 6.
Silicon: Ionic pseudo potentials

\[ V_{ps}^{\text{ps}}(r) \] (hartree)

- \( 0 r_c = 1.703 \)
- \( 1 r_c = 1.878 \)
- \( 2 r_c = 2.021 \)

\[ r \) (bohr) \]
Kleinman–Bylander Form

Basis set expansion with the following approximation for the identity:

\[
1 = \sum_L \frac{|\varphi_L\rangle\langle \Delta V_L \varphi_L |}{\langle \varphi_L \Delta V_L \varphi_L \rangle}
\]

where \( \varphi_L \) is the pseudo–atomic wavefunction from the reference calculation.

\[
E_{PS} = \sum_L \sum_i f_i \langle \Phi_i | \Delta V_L \varphi_L \rangle \omega_L \langle \Delta V_L \varphi_L | \Phi_i \rangle
\]

where

\[
\omega_L = \langle \varphi_L \Delta V_L \varphi_L \rangle
\]

For an atom with s and p non-local potential this requires the calculation of 4 times number of states integrals \( \langle \Delta V_L \varphi_L | \Phi_i \rangle \).
Dual-Space PP

- Goedecker et al, PRB, (1996), 54, 1703
  Hartwigsen et al, PRB, (1998), 58, 3641
  M. Krack, TCA, (2005), 114, 145

- Fully non-local: easy analytic integrals and FFTs

\[
V_{pp}(r) = V_{loc}(r) + \sum_L \sum_{ij} |p^L_i\rangle h^L_{ij} \langle p^L_j| 
\]

- Gaussian form with few adjustable parameters: \( \bar{r} = \frac{r}{r_c} \)

\[
V_{loc}(r) = -\frac{Z_{ion}}{r} \text{erf} \left[ \frac{\bar{r}}{\sqrt{2}} \right] + \exp \left[ -\frac{\bar{r}^2}{2} \right] \left[ C_1 + C_2 \bar{r}^2 + C_3 \bar{r}^4 + C_4 \bar{r}^6 \right]
\]

\[
p^L_i(r) = N_{iL}(r_i) \exp \left[ -\frac{r^2}{2r_i^2} \right]
\]

- Global optimization of all parameters to fit atomic orbital energies of occupied and virtual orbitals.
Non-Linear Core Correction (NLCC)

For many atoms (e.g. alkali atoms, transition metals) core states overlap with valence states. Linearization assumption for XC energy breaks down.

- Add additional states to valence (semi-core)
  - adds more electrons
  - needs higher cutoff
- Add core charge to valence charge in XC energy \(\Rightarrow\) non-linear core correction (NLCC)
  
Non-Linear Core Correction (NLCC)

\[ E_{xc} = E_{xc}(n + \tilde{n}_{\text{core}}) \quad \text{where} \quad \tilde{n}_{\text{core}}(r) = n_{\text{core}}(r) \quad \text{if} \quad r > r_0 \]
Basis Sets

Molecular Optimized Basis Sets (MOLOPT)

Goals

• Suitable for gas and condensed phase, interfaces

• Systematic increase in accuracy

• Suitable for large scale simulations
  Optimal for small number of functions
  Well conditioned

• Low BSSE for weak interactions
Use generally contracted Gaussian basis sets, including diffuse primitives, fully optimized on molecular calculations.

- generally contracted $\rightarrow$ no lone diffuse functions, well conditioned
- diffuse primitives $\rightarrow$ reduced BSSE
- molecularly optimized $\rightarrow$ small but accurate
MOLOPT

- Generally contracted family basis, all exponents used for all angular momenta (including polarization)
- 6/7 primitive functions (pseudopotentials, valence only)
- Larger sets extend smaller sets
- Currently available for H-Rn

<table>
<thead>
<tr>
<th>basis</th>
<th>1st/2nd row</th>
<th>Hydrogen</th>
</tr>
</thead>
<tbody>
<tr>
<td>m-SZV</td>
<td>1s1p</td>
<td>1s</td>
</tr>
<tr>
<td>m-DZVP</td>
<td>2s2p1d</td>
<td>2s1p</td>
</tr>
<tr>
<td>m-TZVP</td>
<td>3s3p1d</td>
<td>3s1p</td>
</tr>
<tr>
<td>m-TZV2P</td>
<td>3s3p2d</td>
<td>3s2p</td>
</tr>
<tr>
<td>m-TZV2PX</td>
<td>3s3p2d1f</td>
<td>3s2p1d</td>
</tr>
</tbody>
</table>
## Condition Numbers (Liquids)

<table>
<thead>
<tr>
<th></th>
<th>SZV</th>
<th>DZVP</th>
<th>TZV2P</th>
<th>QZV3P</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>1.00</td>
<td>2.97</td>
<td>4.46</td>
<td>5.64</td>
</tr>
<tr>
<td>BQ/MeOH</td>
<td>1.30</td>
<td>5.11</td>
<td>6.89</td>
<td>8.66</td>
</tr>
<tr>
<td>acetonitrile</td>
<td>1.34</td>
<td>4.15</td>
<td>5.69</td>
<td>7.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>aug-DZVP</th>
<th>aug-TZV2P</th>
<th>aug-QZV3P</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>10.11</td>
<td>12.54</td>
<td>15.11</td>
</tr>
<tr>
<td>BQ/MeOH</td>
<td>11.00</td>
<td>13.52</td>
<td>13.94</td>
</tr>
<tr>
<td>acetonitrile</td>
<td>9.89</td>
<td>14.58</td>
<td>14.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>m-SZV</th>
<th>m-DZVP</th>
<th>m-TZV2P</th>
<th>m-TZV2PX</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>0.83</td>
<td>3.20</td>
<td>4.18</td>
<td>4.27</td>
</tr>
<tr>
<td>BQ/MeOH</td>
<td>1.04</td>
<td>3.34</td>
<td>4.46</td>
<td>4.66</td>
</tr>
<tr>
<td>acetonitrile</td>
<td>1.11</td>
<td>3.23</td>
<td>4.18</td>
<td>4.36</td>
</tr>
</tbody>
</table>

\[ \log \kappa = \log \frac{\sigma_{\text{max}}}{\sigma_{\text{min}}} \]
Solving the KS Equations

Fix Point Methods

1. initial guess $n^{in}(r)$
2. calculate potential $V(r)$
3. diagonalize KS matrix, get $c^{out}$
4. calculate new density $n^{out}$
5. if $|n^{in} - n^{out}| \leq \epsilon$ stop
6. calculate new density from $n^{in}$ and $n^{out}$ (mixing)
7. go back to 2
Direct Minimization Methods

Minimum $[E_{KS}(c)]$ with the constraint $\sum_{\alpha} c^*_\alpha c_{\alpha j} = \delta_{ij}$

Lagrange function

$$\tilde{E}_{KS}[c, \Lambda] = E_{KS}(c) - \text{Tr} \left\{ \Lambda (c^\dagger c - 1) \right\}$$

Gradient

$$\frac{\partial \tilde{E}_{KS}}{\partial c^*_\alpha n} = \sum_{\beta} F_{\alpha \beta} c_{\beta n} - \sum_{m} c_{\alpha m} \left( \sum_{\beta \gamma} c^*_\beta m F_{\beta \gamma} c_{\gamma n} \right)$$
Orbital Transformation (OT) Method


- Direct optimization technique
- Similar to orbital rotation method
- Constraint is only linear

- Memory $MN$  
  M Number of basis functions

- Scaling $MN^2$  
  N Number of occupied orbitals
• Set of reference occupied orbitals: \( C_0 \)

• New variables \( X \)

\[
C(X) = C_0 \cos U + XU^{-1} \sin U \\
U = \left( X^T S X \right)^{1/2}
\]

• Linear constraint \( X^T S C_0 = 0 \)

• Standard optimization with line search and preconditioning
OT versus Diag-DIIS

256 H2O TZV(2d,2p) 10240 BF on 4 CPUs SUN ultrasparc
**Direct Inversion in Iterative Subspace (DIIS)**

**DIIS**: Acceleration method for iterative sequences.

**Basic idea**: From a series of steps in an optimization procedure, try to guess a better trial vector.

In DIIS we solve exactly (by direct inversion) an optimality condition within the subspace of the parameter vectors generated by the iterations.

Assume we have generated a sequence of $M$ parameter vectors $\{x_m\}_{1}^{M}$ and that we are able to guess for each of the vectors its difference $e_m$ to the stationary point.

**Ansatz**: Find the best linear combination of vectors $x_{M+1} = \sum_{i=1}^{M} c_i x_i$ with the constraint $\sum_{i=1}^{M} c_i = 1$. 
DIIS

**Ansatz**

\[
\text{Min} \left[ \left< \sum_{i=1}^{M} c_i e_i \right| \sum_{j=1}^{M} c_j e_j \right] \quad \text{with} \quad \sum_{i=1}^{M} c_i = 1
\]

where \( <.|> \) is a suitably defined scalar product. This leads to a system of linear equations with \( b_{ij} = <e_i|e_j> \)

\[
\begin{pmatrix}
    b_{11} & b_{12} & \ldots & b_{1m} & -1 \\
    b_{21} & b_{22} & \ldots & b_{2m} & -1 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    b_{m1} & b_{m2} & \ldots & b_{mm} & -1 \\
    -1 & -1 & \ldots & -1 & 0
\end{pmatrix}
\begin{pmatrix}
    c_1 \\
    c_2 \\
    \vdots \\
    c_m \\
    \lambda
\end{pmatrix}
= \begin{pmatrix}
    0 \\
    0 \\
    \vdots \\
    0 \\
    -1
\end{pmatrix}
\]
What to take for the error vectors?

Any measure for the distance from the stationary point.

**AO basis sets**
Hartree–Fock and Kohn–Sham Methods (Pulay)

\[
\{ \mathbf{e} \}_{ij} = \sum_{kl} \left( F_{ik} P_{kl} S_{lj} - S_{ik} P_{kl} F_{lj} \right)
\]

**GDIIS (based on Newton–Raphson)**

\[
e_i = -P g(x_i) \quad \text{Preconditioner } P
\]
Preconditioner

- Preconditioner is essential for good convergence in direct minimization

- Preconditioner matrix $P = M^{-1}$ is applied to gradient: $Pg(x_i)$

- Preconditioner is state dependent (FULL_ALL), single state approximations are better in memory usage and CPU time

- Preconditioner depends on Hamiltonian and should not be updated during optimizations.

- Restart of optimization is needed if Precodnitionner is too bad (outer SCF procedure in CP2K)
Scaling of GPW Calculations

- **N**: Number of occupied orbitals, number of electrons
- **M**: Number of basis functions

- Kohn-Sham matrix: $O(M \log N)$
- Density matrix (incomplete sparse): $O(MN)$
- OT optimization: $O(MN^2)$
System Size Scaling

![Graph showing system size scaling with core count and time per MD step]
Efficiency: GGA Functionals

![Graph showing performance of GGA functionals across different numbers of cores.](image-url)
Linear Scaling KS-DFT

Avoid Matrix Diagonalization

\[ P = \text{sign} \left( S^{-1} H - \mu I \right) S^{-1} \]

Calculate \( S^{-1} \) and sign\((A)\) using Newton-Schultz iterations

\[ A_{i+1} = \frac{1}{2} A_i \left( 3I - A_i^2 \right) \]

Only matrix multiplications required.

**DBCSR**: Sparse matrix-matrix multiplication library
Linear Scaling

Liquid Water

Wall time [min]

$10^0$ $10^1$ $10^2$ $10^3$

Number of atoms

$10^4$ $10^5$ $10^6$

KS-DFT (DZVP)
46,656 cores

NDDO
9216 cores

DFTB
9216 cores
Linear Scaling

Liquid Water

- KS-DFT (DZVP) 46,656 cores
- KS-DFT (SZV) 46,656 cores
- NDDO 9216 cores
- DFTB 9216 cores

Wall time [min] vs. Number of atoms

10^0 to 10^3 on the y-axis
10^4 to 10^7 on the x-axis
100x vertical line at 10^5 atoms
Table 1. Timings (seconds) for the Complete CP2K Energy Calculation (Full) and the Matrix Multiplication Part (mult) on a System Consisting of $\sim$20000 Atoms, As Described in the Text$^a$

<table>
<thead>
<tr>
<th>nodes</th>
<th>64</th>
<th>100</th>
<th>169</th>
<th>256</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PAO-ML</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>full</td>
<td>87</td>
<td>58</td>
<td>41</td>
<td>33</td>
<td>24</td>
</tr>
<tr>
<td>mult</td>
<td>23</td>
<td>17</td>
<td>13</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td><strong>DZVP-MOLOPT-GTH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>full</td>
<td>5215</td>
<td>2765</td>
<td>1996</td>
<td>1840</td>
<td>1201</td>
</tr>
<tr>
<td>mult</td>
<td>5036</td>
<td>2655</td>
<td>1922</td>
<td>1779</td>
<td>1165</td>
</tr>
</tbody>
</table>

$^a$The PAO-ML method outperforms a standard DFT run with a DZVP-MOLOPT-GTH basis by a factor of at least 50X.
www.cp2k.org