

Lecture on First-principles Computations (15): Algorithms for solving self-consistency & optimization problems.

任新国 (Xinguo Ren)

中国科学技术大学
量子信息重点实验室

Key Laboratory of Quantum Information, USTC

Hefei, 2016.11.16

Solving the Kohn-Sham equation iteratively

$$n_{in}(\mathbf{r}) = n_{trial}(\mathbf{r})$$

$$V_{eff}(\mathbf{r}) = V_{ext}(\mathbf{r}) + \int d^3 r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}$$

$$\left(-\frac{\nabla^2}{2m} + V_{eff}(\mathbf{r}) \right) \phi_l(\mathbf{r}) = \epsilon_l \phi_l(\mathbf{r})$$

$$n_{out}(\mathbf{r}) = \sum_{l=1}^N |\phi_l(\mathbf{r})|^2$$

$$n_{out}(\mathbf{r}) = n_{in}(\mathbf{r})?$$

no

$$n_{in}^{new} = n_{in}^{new}(n_{out}, n_{in})$$

yes

exit

How to determine the optimal n_{in} ?

The mathematical problem can be formulated as:

$$n_{out} = f[n_{in}] \quad \longrightarrow$$

Find the fixed point of the nonlinear equation $n - f[n] = 0$!

At each iteration, n_{out} is a functional of n_{in}

$$n_{out} = n_{out}[n_{in}]$$

Define the charge density residual:

$$R[n_{in}] = n_{out}[n_{in}] - n_{in}$$

and the norm of the residual:

$$\langle R[n_{in}] | R[n_{in}] \rangle = \int d^3 r (R[n_{in}])^2$$

$\langle R[n_{in}] | R[n_{in}] \rangle \rightarrow 0$ at self-consistency.

Linear mixing and the Kerker preconditioner

Simple linear mixing:

$$n_{in}^{m+1} = \alpha n_{out}[n_{in}^m] + (1 - \alpha)n_{in}^m = n_{in}^m + \alpha R[n_{in}^m]$$

The Kerker preconditioner (momentum-resolved linear mixing)

G. P. Kerker Phys. Rev. B **23**, 3082 (1981)

$$n(\mathbf{q}) = \frac{1}{\Omega} \int d^3 r n(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}}$$

The mixing parameter is \mathbf{q} -dependent

$$n_{in}^{m+1}(\mathbf{q}) = n_{in}^m(\mathbf{q}) + G^1(\mathbf{q})R[n_{in}^m(\mathbf{q})]$$

$$G^1(\mathbf{q}) = \alpha \frac{q^2}{q_0^2 + q^2}$$

To prevent “charge sloshing” in metal slabs,
arising from diverging dielectronic function at $q \rightarrow 0$.

Pulay mixer

P. Pulay, *Chem. Phys. Lett.* **73**, 393 (1980)

$$n_{in}^{opt} = \sum_{i=1}^m \alpha_i n_{in}^i$$

Assumption: the residual is linear with respect to the input charge density.

$$R[n_{in}^{opt}] = R\left[\sum_{i=1}^m \alpha_i n_{in}^i\right] = \sum_{i=1}^m \alpha_i R[n_{in}^i]$$

Minimizing the norm $\langle R[n_{in}^{opt}] | R[n_{in}^{opt}] \rangle$ with respect α_i to
under the constraint $\sum_{i=1}^m \alpha_i = 1$

$$\Rightarrow \alpha_i = \frac{\sum_j A_{ij}^{-1}}{\sum_{i,j} A_{ij}^{-1}}, \text{ with } A_{ij} = \langle R[n_{in}^i] | R[n_{in}^j] \rangle$$

$$n_{in}^{m+1} = n_{in}^m + G R[n_{in}^{opt}] \quad 0 < G < 1: \text{ length of step}$$

(Pulay method = “Direct inversion of the iterative sub-space (DIIS) method”)

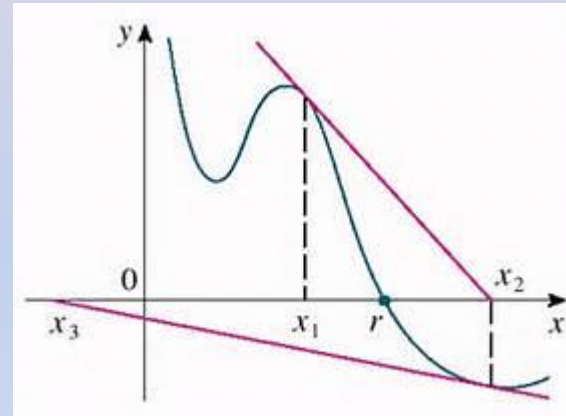
Newton-Raphson method

A method to find successively better approximation to the zeros (roots) of a real-valued function.

$$x: f(x) = 0$$

The Newton-Raphson algorithm:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$



$(x_{n+1}, 0)$ is the intersection of the x -axis and the tangent of the curve at $(x_n, f(x_n))$

The algorithm can be extended to complex numbers and multi-valued functions.

Quasi Newton-Raphson method

Consider the problem

$$F(\mathbf{x}) = \mathbf{x}, \text{ e.g., } n_{out}[n_{in}] = n_{in}$$

The residual: $R(\mathbf{x}) = F(\mathbf{x}) - \mathbf{x}$

Solving the problem amounts to minimizing the norm $\langle R(\mathbf{x}) | R(\mathbf{x}) \rangle$

Quasi Newton-Raphson method:

$$\mathbf{x}_{m+1} = \mathbf{x}_m - J_m^{-1} \cdot R_m$$

Jacobian: $J = \frac{\partial R}{\partial \mathbf{x}}$ ($J_{mn} = \partial R_m / \partial x_n$)

Broyden mixing (BFGS method)

Approximate way of obtaining successive Jacobians J' 's

Start with a reasonable guess

$$J_0^{-1}, \quad \text{e.g. } J_0^{-1} = \alpha I$$

The prediction from J_{m-1}^{-1} : $\delta \mathbf{x}_m = \mathbf{x}_m - \mathbf{x}_{m-1} = -J_{m-1}^{-1} \cdot \mathbf{R}_{m-1}$

The change of the residual from actual calculations:

$$\delta \mathbf{R}_m = \mathbf{R}_m - \mathbf{R}_{m-1}$$

The updated Jacobian satisfy: $\delta \mathbf{x}_m - J_m^{-1} \cdot \delta \mathbf{R}_m = 0$

Further requiring $\|J_m^{-1} - J_{m-1}^{-1}\|$ to be minimized

$$\Rightarrow J_m^{-1} = J_{m-1}^{-1} + \frac{(\delta \mathbf{x}_m - J_{m-1}^{-1} \cdot \delta \mathbf{R}_m) \delta \mathbf{R}_m}{\langle \delta \mathbf{R}_m | \delta \mathbf{R}_m \rangle}$$

The minimization problem

Minimizing a function

$F(x_1, x_2, \dots, x_N)$ with respect to $x_i, i = 1, \dots, N$

- Steepest Descend (SD) method

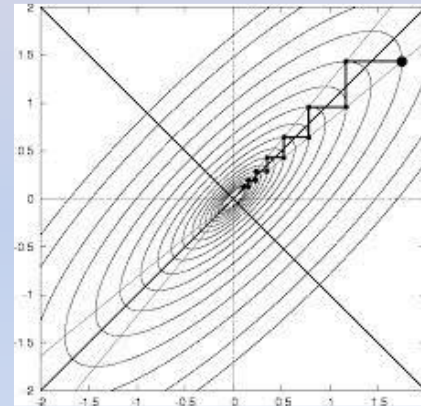
At a starting point in space $\{x_i^0\}$

“The steepest direction”:

$$g_i^0 = -\frac{\partial F}{\partial x_i} \Big|_{x_i = x_i^0}$$

Move along the direction: $g^0 = \{g_i^0\}$

by a distance α , $x_i^1 = x_i^0 + \alpha g_i^0$



Now $F = F(\alpha)$, the optimal $\alpha = \alpha^1$ is found by “line search”.

Repeat the steps :

$$g_i^m = -\frac{\partial F}{\partial x_i} \Big|_{x_i = x_i^m}, \quad x_i^{m+1} = x_i^m + \alpha^{m+1} g_i^m$$

The Conjugate Gradient (CG) method

$$F = F(\mathbf{x}) = \frac{1}{2} \mathbf{x} \cdot \mathbf{H} \cdot \mathbf{x}, \quad \mathbf{g}(\mathbf{x}) = -\frac{\partial F}{\partial \mathbf{x}} = -\mathbf{H} \cdot \mathbf{x}, \quad \mathbf{x} = \{x_i\}$$

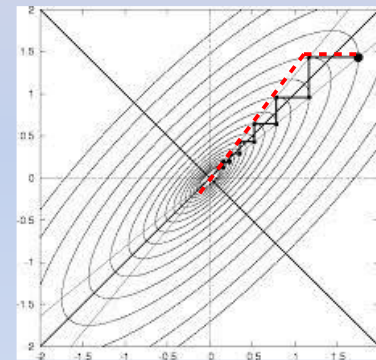
Starting point: $\mathbf{x}^0, \mathbf{d}^0 = \mathbf{g}^0$

First step: $\mathbf{x}^1 = \mathbf{x}^0 + \alpha^1 \mathbf{d}^0$ (same as SD)

From now on: $\mathbf{d}^m \cdot \mathbf{H} \cdot \mathbf{d}^{m+1} = 0$

$$\mathbf{d}^{m+1} = \mathbf{g}^{m+1} + \gamma^{m+1} \mathbf{d}^m$$

$$\gamma^{m+1} = -\frac{\mathbf{d}^m \cdot \mathbf{H} \cdot \mathbf{g}^{m+1}}{\mathbf{d}^m \cdot \mathbf{H} \cdot \mathbf{d}^m}$$



For quadratic functions, the minimum can be reached exactly with N steps for N -variable functions!

Direct minimization of the KS functional

The Kohn-Sham functional

$$E_{KS}[\{\phi_n\}] = \sum_n f_n \langle \phi_n | \hat{T} | \phi_n \rangle + E_{ext}[n] + E_H[n] + E_{xc}[n]$$

Minimizing E_{KS} with respect to ϕ_n , under the condition

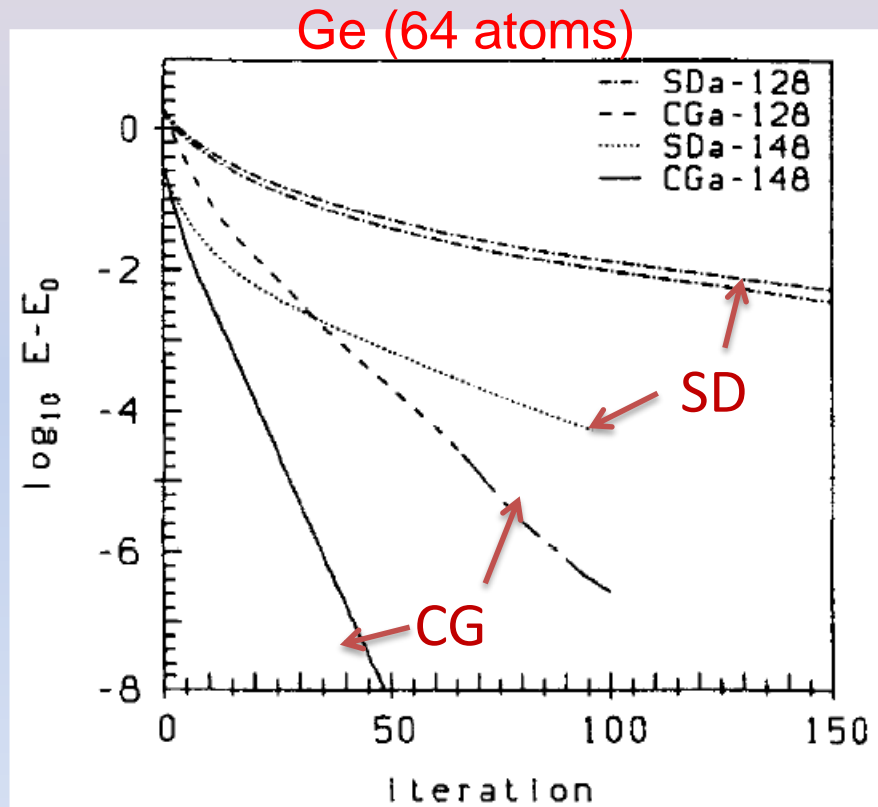
$$\langle \phi_n | \phi_{n'} \rangle = \delta_{n,n'} \quad \rightarrow$$

$$\left(-\frac{\nabla^2}{2m_e} + V_{eff}(\mathbf{r}) \right) |\phi_n\rangle = \epsilon_n |\phi_n\rangle$$

Directly minimizing $E_{KS}[\{\phi_n\}]$ amounts to solving the KS equation !

CG is the method to use!

Example



G. Kresse, Furthmüller, CMS (1996)