

Lecture X : Self-consistent solution of the Kohn-Sham equation

I. SELF-CONSISTENT NATURE OF THE DENSITY IN THE KOHN-SHAM SYSTEM

In this lecture and the following, the topics that we shall discuss are going to be technical rather than physical. The methods that we shall look at do not bear relevance as far as density functional theory is concerned but they are central to the efficient execution of the calculations.

As shown in previous lectures, the Kohn-Sham equations have each the form of a single-particle Schrödinger equation with the difference that the effective potential operator, which yields the orbitals that determine the density, depends on the density itself. The Kohn-Sham system then must be solved self-consistently. This is to say that taking as input an initial density, a new density is found. This new density is then fed into the equation system which yields an output density and so on. This simplistic approach however usually results in poor convergence properties. In general, we resort to a set of algorithm which are called *mixing algorithms*. These carefully designed algorithms prescribe ways of updating the density by mixing old and new densities in the best possible way. In this lecture, we shall see why the simple-minded approximation fails to converge and how the more sophisticated algorithms work. For a very accessible description of the problem see the paper by Dederichs and Zeller (*Physical Review B*, **28**, 5462, 1983).

At the N th step of the self-consistent iterations, the output density (which is to be read in in the $N + 1$ th step) may be expressed as a functional of the input density, i.e.

$$n_{N+1} = F[n_N] \quad (1)$$

Let's now assume that we are rather close to the minimum. This assumption is necessary for the rest of the treatment since we are going to assume linear response. To linear order, the difference between the current density and ground state density (the one we are trying to reach) denoted by $\delta n(\vec{r}) = n(\vec{r}) - n_0(\vec{r})$ may be written in terms of the difference in the potential

$$\delta n(\vec{r}) = \int d\vec{r}' \chi(\vec{r}, \vec{r}') \delta V(\vec{r}'). \quad (2)$$

The response function, $\chi(\vec{r}, \vec{r}')$ in Eq. 2 may be shown to be related to the dielectric tensor of the system. Due to the self-consistent nature of the system on the other hand, the potential may just as well be regarded as the independent variable and the difference in the potential may be expressed in terms of the difference in the density

$$\delta V(\vec{r}) = \int d\vec{r}' U(\vec{r}, \vec{r}') \delta n(\vec{r}') \quad (3)$$

where by definition $U(\vec{r}, \vec{r}')$ is the functional derivative of the potential $V(\vec{r})$. In other words it is the second functional derivative of the total energy. It can easily be evaluated to yield

$$U(\vec{r}, \vec{r}') = \frac{1}{|\vec{r} - \vec{r}'|} + \left. \frac{d\epsilon_{xc}}{dn} \right|_{n=n_0} \quad (4)$$

Substituting Eq. 3 into Eq. 2, we obtain (paying careful attention to the integration variables)

$$\begin{aligned} \delta \rho(\vec{r}) &= \int d\vec{r}' \int d\vec{r}'' \chi(\vec{r}, \vec{r}') U(\vec{r}', \vec{r}'') \delta n(\vec{r}'') \\ &= \int d\vec{r}'' \underbrace{\int d\vec{r}' \chi(\vec{r}, \vec{r}') U(\vec{r}', \vec{r}'') \delta n(\vec{r}'')}_{f(\vec{r}, \vec{r}'')} \\ &= \int d\vec{r}' f(\vec{r}, \vec{r}') n(\vec{r}') \end{aligned} \quad (5)$$

where to obtain the last line dummy variables have been exchanged. Eq. 5 displays plainly the self-consistent nature of the density. The iterative version of Eq. 5 is

$$\delta n_{N+1} = \underline{f} \delta n_N \quad (6)$$

where the integral operator in Eq. 5 has been denoted by the symbol \underline{f} . By recursive application of Eq. 6, we arrive at

$$\delta n_{N+1} = \underline{f} n_N = \underline{f}^N \delta n_1 \quad (7)$$

Just like any operator, \underline{f} can be written in the spectral representation with eigenvalues $\{\lambda_i\}$ and eigenvectors $\{|i\rangle\}$. Then Eq. 7 becomes

$$\delta n_{N+1} = \sum_i \lambda_i^N |i\rangle\langle i| \delta n_1 \quad (8)$$

As is evident from the form in Eq. 8 convergence is achieved in this simple-minded approach only if all the eigenvalues of the operator \underline{f} is between -1 and 1. Since this cannot be guaranteed in general such an iterative scheme diverges for most (if not all) systems and a mixing scheme must be applied.

II. MIXING SCHEMES

A. Simple mixing

The simplest algorithm used for mixing is the addition of a certain fraction of the density in the previous step to that of the current step, that is to say

$$n_{N+1}(\vec{r}) = \alpha F[n_N] + (1 - \alpha)n_N \quad (9)$$

where F , defined in Eq. 1 is the functional which yields the output density given an input density and α is a constant determined at the beginning of the calculation. In terms of the linear differences Eq. 9 may be written as

$$\begin{aligned} \delta n_{N+1} &= \alpha \underline{f} \delta n_N + (1 - \alpha) \delta n_N = [1 - \alpha(1 - \underline{f})] \delta n_N \\ &= \sum_i (1 - \alpha(1 - \lambda_i))^N |i\rangle\langle i| \delta n_1 \equiv \sum_i (1 - \alpha\mu_i)^N |i\rangle\langle i| \delta n_1 \end{aligned} \quad (10)$$

where the spectral representation of the operator has once again been employed and the definition $\mu_i \equiv 1 - \lambda_i$ has been made.

For convergence the inequality $|1 - \alpha\mu_i|$ must be satisfied. In other words if we manage to choose α in the *convergence region*

$$0 < \alpha < 2/\mu_{\max} \quad (11)$$

then converge is achieved in the self-consistent iterations. The value α needs to be chosen differently for different systems. While certain systems require very small α (meaning slow convergence), other systems are more stable and converge more rapidly with larger values of α .

B. Anderson mixing

Anderson mixing is an accelerated mixing scheme where both the input and the output densities are mixed with suitable mixing parameters.

$$\begin{aligned} \tilde{n}_N &= \alpha_N n_N + (1 - \alpha_N) n_{N-1} \\ \tilde{F}_N &= \alpha_N F[n_N] + (1 - \alpha_N) F[n_{N-1}] \end{aligned} \quad (12)$$

In contrast to the simple mixing scheme illustrated in Eq. 10, α is not taken to be a constant for the entire duration of the calculation but updated at each iteration step in such a way that

$$\int d\vec{r} [\tilde{F}_N(\vec{r}) - \tilde{n}_N(\vec{r})]^2 \quad (13)$$

is minimized. This is the second best thing to use an \vec{r} -dependent α . Defining

$$r_n \equiv F_N - f_N \quad \text{and} \quad \int f(\vec{r}) g(\vec{r}) d\vec{r} \equiv (f, g) \quad (14)$$

we expand out the terms in Eq. 13 using Eq. 12 to get

$$\begin{aligned} \int d\vec{r} [\tilde{F}_N(\vec{r}) - \tilde{n}_N(\vec{r})]^2 &= \int d\vec{r} [\alpha_N^2 (F_N - n_N)^2 + (1 - \alpha_N)^2 (F_{N-1} - n_{N-1})^2 + 2\alpha_N(1 - \alpha_N)(F_N - n_N)(F_{N-1} - n_{N-1})] \\ &= \alpha_N^2 (r_N, r_N) + (1 - \alpha_N)^2 (r_{N-1}, r_{N-1}) + 2\alpha_N(1 - \alpha_N)(r_N, r_{N-1}). \end{aligned} \quad (15)$$

Differentiating Eq. 15 and solving the resulting equation for α_N yields the optimum mixing parameter for each iteration as

$$\alpha_N = -\frac{(r_{N-1}, r_N - r_{N-1})}{(r_N - r_{N-1}, r_N - r_{N-1})}. \quad (16)$$

C. Broyden mixing

Broyden mixing treats the density update as an approximate inversion procedure. The goal is to minimize the difference

$$R[n] = F[n] - n \quad (17)$$

where R is referred to as the *residual* and F is defined in Eq. 1. We shall return to the concept of the residual in the next lecture. The exact minimization is, of course, not possible as this would amount to the solution of the original problem. However, an alternative and approximate solution may be devised if we are sufficiently close to the true solution. In order to facilitate this, we make use of the *Jacobian* defined as

$$J \equiv \frac{\delta R}{\delta n} \quad (18)$$

The value of the Jacobian that would result in a zero residual at the $i + 1$ th step is

$$n_{i+1} = n_i - J^{-1}R_i \quad (19)$$

which we could see easily by writing Eq. 18 as

$$\delta n_i = J_i^{-1}R_i \quad (20)$$

where $\delta n_i = n_i - n_0$ measures the difference between the density at the current step and the true ground state density. Substituting Eq. 20 into Eq. 19, we have

$$n_{i+1} = n_i - \delta n_i = n_i - (n_i - n_0) = n_0 \quad (21)$$

which means that convergence has been reached.

Of course, the Jacobian can neither be known exactly nor inverted easily. However, the Broyden method proposes a way of making an initial guess for the Jacobian (which is more easily diagonalizable than the original) and iteratively improving it as the iterations progress. To be more precise we don't actually invert at any point during the process the Jacobian itself but we improve the inverse Jacobian progressively instead. At each step i where

$$\delta n_i = n_i - n_{i-1} = -J_{i-1}^{-1}R_{i-1} \quad (22)$$

the new, improved J_i^{-1} is chosen such that :

1. The result of the iteration just completed is reproduced, i.e.,

$$0 = \delta n_i - J_i^{-1}(R_i - R_{i-1}). \quad (23)$$

2. The norm of the change in the Jacobian matrix

$$Q = |J_i^{-1} - J_{i-1}^{-1}| \quad (24)$$

is minimized.

The result is

$$J_i^{-1} = J_{i-1}^{-1} \frac{(\delta n_i - J_{i-1}^{-1}R_i)\delta R_i}{\langle \delta R_i | \delta R_i \rangle}, \quad (25)$$

which is suitable for use in small systems but not large systems where the Jacobian from the previous step needs to be stored in its entirety. At this point the *modified Broyden method* comes to rescue where Q is replaced by a weighted norm

$$Q_{mod} = \sum_{j=1}^i w_j |\delta n_j - J_j^{-1}R_j|^2 + w_0 |J_i^{-1} - J_0^{-1}| \quad (26)$$

III. HELLMAN-FEYNMAN THEOREM

Once the ground state density has been reached through self-consistent iterations, the Kohn-Sham orbitals corresponding to this density may be used to calculate several properties of the system, one of which is the force on the atoms. This section is somewhat unrelated to the rest of this chapter but it has been placed here for lack of a better home.

Forces on the atoms are a central part of any DFT calculation since these forces, calculated at the end of the relaxation of the electronic coordinates of the are used for Born-Oppenheimer based geometric optimization methods such as static minimization or molecular dynamics.

The Hellman-Feynman theorem is really a general statement about derivatives of expectation values at the ground state but we discuss the specific case of atomic forces. As is known from elementary physics, the force on the i th atom is given by

$$\vec{F}_i = -\frac{\partial E}{\partial \vec{R}_i} \quad (27)$$

where the electronic energy for the ground state is clearly given by

$$E_e = \langle \Psi | \hat{H} | \Psi \rangle. \quad (28)$$

Substituting Eq. 28 into Eq. 27, we have, for the electronic part of the forces,

$$\vec{F}_{i,e} = -\left\langle \frac{\partial \Psi}{\partial \vec{R}_i} | \hat{H} | \Psi \right\rangle - \langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle - \langle \Psi | \hat{H} | \frac{\partial \Psi}{\partial \vec{R}_i} \rangle \quad (29)$$

Grouping the first and the third terms together and making use of the relation $\hat{H}|\Psi\rangle = E|\Psi\rangle$ in Eq. 29 yields

$$\vec{F}_{i,e} = -E \left[\left\langle \frac{\partial \Psi}{\partial \vec{R}_i} | \Psi \right\rangle + \left\langle \Psi | \frac{\partial \Psi}{\partial \vec{R}_i} \right\rangle \right] - \langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle = -E \frac{\partial}{\partial \vec{R}_i} \langle \Psi | \Psi \rangle - \langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle = -\langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle \quad (30)$$

where in the second term the derivative of the inner product cancels because this inner product yields 1. This is a very important and useful result which states that while differentiating the energy to obtain the forces, we may entirely ignore the change in the wavefunction and just calculate the change in the Hamiltonian operator. In the case of Coulombic DFT, this is further simplified by the fact that the only term that depends on atomic coordinates is the external potential due to atomic nuclei, which luckily is a single-particle operator whose expectation value may be expressed in terms of the single-particle density. As the finishing touch we add to the electronic contribution to the force the ionic contribution to find the total force on the i th atom. We then finally arrive at

$$\vec{F}_i = - \int d\vec{r} n(\vec{r}) \frac{\partial V_{ext}}{\partial \vec{R}_i} - \frac{\partial E_{II}}{\partial \vec{R}_i} \quad (31)$$