

The following note is based on the Coursera Mooc **Density Functional Theory** by **École Polytechnique** which is available at <https://www.coursera.org/learn/density-functional-theory>

▼ Week 3: Approximation and Strategies

The universal Functional:

We have the total energy functional:

$$E[\rho] = E_k^{non}[\rho] + E_{ext}[\rho] + E_H[\rho] + E_{xc}[\rho]$$

Here, E_k^{non} is the kinetic energy of the non-interacting system that we previously denoted with T_s . Also, $E_{ext}[\rho] = \int U_{ext}(r)\rho(r) dr$. Notice that all the terms here are known exactly except $E_{xc}[\rho]$.

It turns out the *XC energy functional* term is very small compared to other energy terms, but it decides many properties. So, we have to approximate it as better as we can. There are a large collection of such approximations but all of them can be divided into two categories:

- *Empirical approximation:*

These are functionals that have free parameter in them, and those parameters are obtained as a fitting parameter from experiment. They lack the *true first-principles nature*. There are a lot of publicly available data. Also, machine learning is a possibility where we can bypass the KS equation.

- *Non-empirical approximation:*

These are functionals constructed theoretically from some constraint and they do not contain any fitting parameters.

Usually, empirical approximations are preferred by quantum chemists and are more accurate. On the other hand, non-empirical approximations are preferred by physicists and are more universally applicable.

Approximating XC functional:

There is a systematic way of approximating XC functional where we can get better and better approximation by the cost of computational power. This is due to *Perdew and Schmidt* (<https://aip.scitation.org/doi/10.1063/1.1390175>) which is called the Jacob's ladder. Below is the list of classification where as we go from top to bottom, the computational cost increases and the approximation gets better.

1. Local density approximation (LDA)
2. Generalized gradient approximation (GGA)
3. Meta-GGA

4. Hybrid approximation
5. RPA-like approximation

We discuss them below.

LDA:

LDA is a non-empirical approximation. Here, we derive this approximation by using the density of *jellium*, a uniform electron gas. For such simple system, the exchange functional is known exactly and the correlation functional can be approximated. In LDA, we *treat all system locally as a homogeneous electron gas* and then we integrate it over the whole system. Some key points about LDA:

- Lattice constant of solids within 1-2% (typically too short)
- Total molecular ground state energy within 1-5% of the exact result
- Too large atomization energy (over binding)
- No bound negative ions
- Not accurate enough for most solids. Useful for simple solids only

The reason LDA works:

Even though LDA rests on very crude assumption of homogeneous electron gas, this approximation works quite well for inhomogeneous system too! To understand the reason why it works so well, let us consider a general system.

In a general many-body system, the electrons feel three types of tendency to avoid each-other: (i) *Coulomb interaction*, (ii) *Pauli's exclusion principle*, and (iii) *electronic correlation*. All together we define a exchange-correlation hole, that is simply the region around electron where it is less likely to find other electrons. Mathematically, we define the pair density, $P(r, r')$, which is the probability of finding any pair of electrons in the volume element:

$$\text{For classical electrons: } P_{\text{classical}}(r, r') = \rho(r)\rho(r')$$

$$\text{For real electrons: } P(r, r') = \rho(r)\rho(r') + \rho(r)\rho_{xc}(r, r')$$

It turns out that the exchange-correlation density, $\rho_{xc}(r, r') = \rho_x(r, r') + \rho_c(r, r')$ obeys three rules:

$$\begin{aligned} \rho_x(r, r') &\leq 0 \\ \int \rho_x(r, r') dr' &= -1 \\ \int \rho_c(r, r') dr' &= 0 \end{aligned}$$

Now, it also turns out that the exchange-correlation hole density of uniform electron gas (LDA assumption) also satisfies these three constraints of a general system. That's why it works well.

GGA and others:

While LDA assumes only linearly to the density, the GGA incorporates higher non-linear terms. In practice, we use an enhancement factor $F_{xc}(s)$ with the LDA approximation. This is also a non-empirical approximation. *BLYP* and *PBE* are two of the most cited GGA. Detail in the note of Ch. 5 of Lee's book.

While in LDA we saw only a *local* electron density dependence of the XC energy density, in GGA, the XC energy density depends on the electron density and its gradient. That's why it is a *semi-local approach*. There can be completely **non-local* approach where the XC energy density depends on the electron density of everywhere. Meta-GGA is a semi-local approximation like GGA but Hybrid and RPA-like approximations are non-local.

The difference between GGA and meta-GGA is that in meta-GGA, the XC energy density depends on the kinetic energy density of the spin densities, τ^σ . GGA and meta-GGA works when the XC holes are sufficiently localized. It fails for weakly bonded system (van der Waals), strongly correlated system, and band-gap prediction.

Band gap of solids:

There can be several types of band gap in solids such as optical gap, fundamental gap, direct gap, minimum gap, etc. The band gap is an observable which can be obtained from the density. The functional is expressed as $E_g[\rho(r)]$. The band gap is given as:

$$E_g = (E_{n+1, s} - E_n) - (E_n - E_{n-1, s}) \quad ; \quad s = 0, \text{ ground state}$$

Here, $E_{n+1, s}$ means the energy of $n + 1$ electron at the s -state. So, from the density, the flow of calculating the band gap is the following:

$$\rho \xrightarrow[T, s]{} U_{ext} \xrightarrow[s, n]{} \psi \longrightarrow E_g$$

That means we have to specify the temperature ($T = 0$), the ground state ($s = 0$), and the number of electrons, n . That means we need to specify the ensemble we are working on.

Now, the problem is that KS equation correctly give us the density, ρ , but give all other observables incorrectly. If one calculates the band structure from KS eigenvalues, ε_i , this would be mistaken. This is known as *KS band gap problem*. Mathematically, $\Delta\varepsilon_i[\rho] \neq E_g[\rho]$. Therefore, we tweak the effective potential in such a way that the density doesn't change, but it also give us the correct energy. KS equation is:

$$\left(-\frac{\nabla^2}{2} + U_{KS}(r) \right) \phi_i(r) = \varepsilon_i \phi_i(r)$$

At first, we make the potential non-local by assing a projector of all the non-occupied orbitals (c) with a multiplicative factor Δ_g . :

$$U_{KS}(r) \longrightarrow U_{KS}(r) + \Delta_g \sum_c \phi_c(r) \phi_c^*(r') = U_{KS}(r, r')$$

Since, these orbitals are non-occupied, the density doesn't change. We tune Δ_g from empirical data. In this way, by adding extra degree of freedom in the system, we can calculate more observables.

Solving KS equations:

As we mentioned earlier, the KS equations - like Hartree equation or HF equation - should be solved self-consistently.

$$\hat{H}_{KS}\phi_i(r) = \varepsilon_i\phi_i(r)$$

It can be solved by diagonalization of the Hamiltonian but the complexity is cubic in that case. That's why in practice, we use iterative self-consistent approach where at first we guess an initial ρ or ϕ . Then we update U_H and U_{xc} from our choice of XC functional approximation. Then we solve the eigenvalue problem and obtain ϕ_{new} (eigenfunctions), ε (eigenvalues), and ρ_{new} (via the eigenfunctions). Now, we check if $|\rho - \rho_{new}| < \tau$ where τ is a predefined tolerance value. If it converges, we stop and output the converged eigenvalues and eigenfunctions (hence, density). If it doesn't converge, then we update the ρ and start over.

Updating ρ for each new iteration can be done in many ways. We might simply input the previous ρ as the new ρ . In practice, this is more complicated in order to reduce the number of iteration and also to remove algorithmic instability. The input density might depend on previous densities with coefficients and weights! This is how KS solver is implemented in various *ab initio* codes.

Final Remark:

Even though we know that KS eigenvalues should not be used as, we often use them as an approximation when constructing band structure/band gap. While LDA/GGA underestimates band gaps almost $\sim 50\%$, the Hybrid approximations are improved and underestimates around $\sim 12\%$.

Lastly, there are several database dedicated to collect, organize and share DFT results such as *The Materials Project*, *OQMD*, *NOMAD-CoE*, etc. The huge number of available data offers the possibility of applying machine learning.