Generalised Gradient Approximations (GGA's)

To understand how the GGA came into existence we must first know how the LDA failed. So we first we look at why the LDA works or why it doesn't. Then we try to improve it based on what features it lacks.

LDA (a recap)

As we saw, the LDA was the most basic case that we can think of for approximating the exchange-correlation functional. Here we considered the electron density of inhomogenoues system to be locally homogenous. The exchange-correlation energy, E_{xc} is expressed as:

$$E_{xc}^{LDA} = \int \rho(\mathbf{r})\epsilon_{xc}^{h}(\rho(\mathbf{r}))d\mathbf{r} \tag{1}$$

and

$$E_{xc}^{LSDA} = \int \rho(\mathbf{r})\epsilon_{xc}^{h}(\rho_{\uparrow}(\mathbf{r}), \rho_{\downarrow}(\mathbf{r}))d\mathbf{r} \tag{2}$$

where $\epsilon_{xc}^{h}(\rho(\mathbf{r})) \rightarrow$ exchange-correlation energy density at $\mathbf{r}$, evaluated by taking the E_{xc} of the uniform gas with density $\rho(\mathbf{r})$.
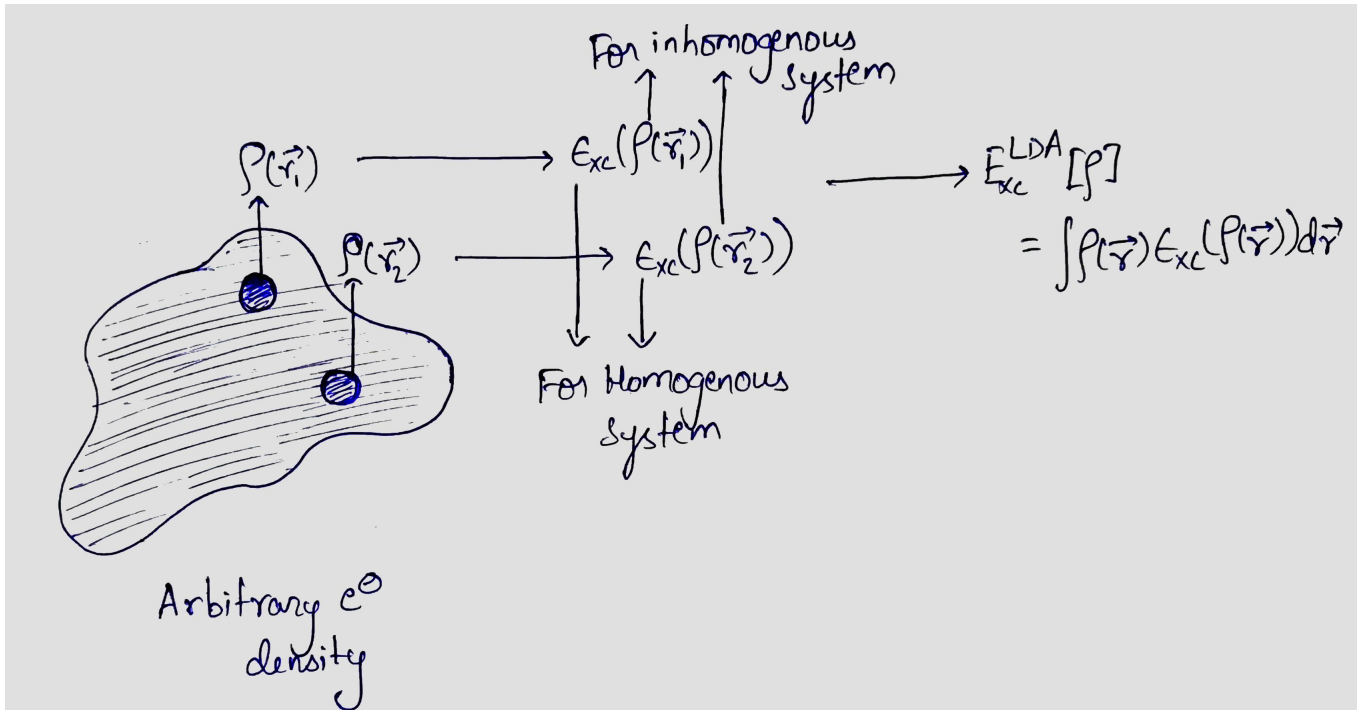

Figure 1: Evaluation of LDA

Based on the approximation considered, we can expect that the LDA should work fine for a totally (or nearly) homogenous system and in regions of large density gradients it should fail to calculate the properties of the system.

Expectations from LDA

Figure 2: LDA should work

Figure 3: LDA should be somewhat useless

Figure 4: LDA should be totally useless

Reality→LDA works considerably well in all the above cases!!!

LDA calculates many properties fairly well (even for inhomogenous systems), some of which are:

1. It works nearly perfect for many properties of metals, as should be expected. This is because, in many metals, e^- behave as nearly free which results in a nearly homogenous e^- density in metals.

2. It predicts molecular properties like equilibrium structures, charge moments etc. In molecules, the e^- density is relatively much more inhomogenous. But still the LDA works fine.

However, the energy details are not so good for the inhomohenous systems. Comparing with the experiments, the unsigned standard deviation in atomisation energies are:
$$\Delta_{LDA} = 36 \, \text{Kcal/mol}(1.56eV) \tag{3}$$
$$\Delta_{HF} = 78 \, \text{Kcal/mol}(3.38eV) \tag{4}$$

Why should the LDA even work at all for the inhomogenous system???
Answer: LDA is for homogenous electron gas →its hole is spherical in nature and it nicely represents some of the exact properties of spherically-averaged holes as discussed earlier:

1. The sum rule is satisfied and $n_{xc}(\mathbf{r}_1, \mathbf{r}_2) \leq 0$ everywhere, which ensures that around any e^- in the density, there is exactly one e^- missing from its surrounding.

2. The sum rule for exchange is also satisfies and $n_x(\mathbf{r}_1, \mathbf{r}_2) \leq 0$ everywhere. This ensures the *Pauli-exclusion principle* and also helps in taking care of the self-interaction due to the Hartree term.

3. The correlation part integrates out to 0, and also nicely generates the e^--e^- cusp condition in the many-e^- Kohn-Sham wavefunction.

4. Other than the hole properties, the LDA functional also satisfies the universal scaling constraints which must be satisfied by any XC-functional under the Kohn-Sham scheme.

What is the problem with LDA?
Answer: LDA is for homogenous electron gas→tends to homogenise the bond for inhomogenous systems also→ error in calculation of energetic quantities. Eg: It leads to overbinding in molecules.

Overbinding in LDA - A step by step guide

1. Atoms have a highly inhomogenous electron density

2. Molecules have a relatively more homogenous electron density compared to atoms as electrons are now more delocalised around two(or more atoms) but still its inhomogenous

3. LDA best approximates homogenous electron density

4. LDA tends to homogenise the electron density more in the molecule

5. More bonding character in molecule than should be present

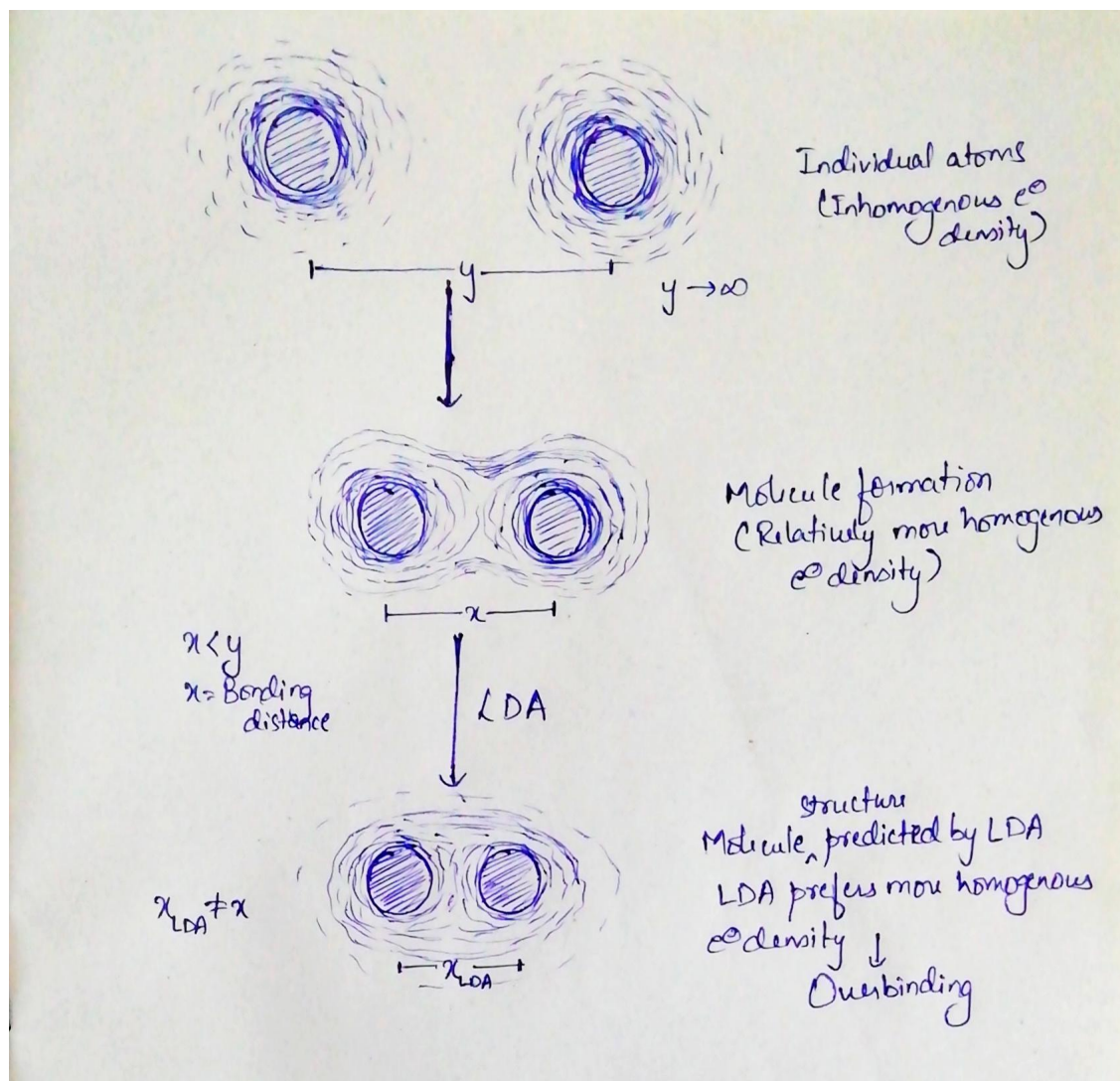6. Exchange energy of the molecule is too negative $\rightarrow$ Overbinding

Figure 5: Overbinding process of LDA

Gradient expansion (GE)

To improve the already existing LDA, **Ma** and **Brueckner** in 1968 tried to attempt an perturbative expansion of the homogenous e^- density. Their reasoning can be eplained via the following steps:

1. Divergence from uniformity of the e^- density is due to perturbation to the homogenous system.

2. We can think of the small-perturbation to the homogenous system as little distortions in the potential which will result in small perturbation to the e^- density making the system slightly inhomogenous.

3. Hence we can try to attempt a correction to LDA by taylor expansion of e^- density around the homogenous electron density

So we can try:

$$\rho^{inh}(\mathbf{r}) \rightarrow \rho^h(\mathbf{r}) \left[1 + \nabla \rho^h(\mathbf{r'})|_{\mathbf{r'}=\mathbf{r}} + \mathcal{O}(\nabla^2 \rho^h(\mathbf{r}))\right] \tag{5}$$

and what's done in GE is:

$$\begin{aligned}
E_{xc}^{GE}[\rho] &= \int \rho \epsilon_{xc}(\rho) d\mathbf{r} + \int C_{xc} \rho \frac{|\nabla \rho|}{\rho^{\frac{4}{3}}} d\mathbf{r} + ... \\
&= E_{xc}^{LDA}[\rho] + \text{Perturbative corrections}
\end{aligned} \tag{6}$$

How good is GE?

Expectations: Since it's a Taylor series expansion around uniform density $\rightarrow$ should perform well for small gradients in density.

Reality: Performance **significantly reduced** compared to LDA

But why?

1. Sum-rules for the total as well as exchange hole are broken.

2. XC-hole is not restricted to be negative for any pair $(\mathbf{r}_1, \mathbf{r}_2)$ which is in strict violation to $n_{xc}(\mathbf{r}_1, \mathbf{r}_2) \leq 0$.

3. Scaling relations are also not satisfied.

Due to the breaking of the above universal conditions of exact holes $\rightarrow$ The relationship between on-top hole and its extension is lost $\rightarrow$ the GE-hole doesn't represent the exact hole in any manner $\rightarrow$ the E_{xc}^{GE} (which represents the attraction between an electron density and its hole) will now have inconsistent behaviour.

Brute Force method - The birth of GGA

With GE, came the correction to homogenous e^- density. But we lost some of the properties of exact holes. So the main concern now is, how to introduce the density correcions as well as keeping the properties of exact holes intact and also to follow the scaling contraints. To tackle this concern, what's now done is $\rightarrow$ keep the gradient expansion, but modify it forcefully to obey the required constraints. This method can be summarised as:

1. Parts in GE which violate $\rho_{xc} \leq 0$ and $\rho_x \leq 0 \rightarrow$ Just set them to 0.

2. To make sure the **sum rules** are obeyed $\rightarrow$ truncate the XC-holes such that h_x and h_c contain 1 and 0 electron charges respectively

3. To help with the above points and also to make the $E_{xc}[\rho]$ obey the scaling relations, a (or some) parameter was introduced into the GE-expansion.

So we have

$$E_{xc}^{GE}[\rho] + \text{XC-properties} + \text{Scaling-relations} \rightarrow E_{xc}^{GGA}[\rho] \tag{7}$$

where E_{xc}^{GGA} is the Generalised Gradient Approximation (GGA)

$$E_{XC}^{GGA}[\rho] = \int \rho(\mathbf{r})\epsilon_{xc}^h(\rho(\mathbf{r}))(1 + \mu s^2 + \mathcal{O}(s^4))d\mathbf{r} \tag{8}$$

where μ is a parameter and s is the dimensionless quantity (can be verified by simple dimensional analysis):

$$s = \frac{|\nabla\rho(\mathbf{r})|}{\rho^{\frac{4}{3}}(\mathbf{r})} \tag{9}$$

The GGA parameter $\rightarrow \mu$

There are two approches to calculate the parameter μ:

1. Semi-Emperical: Eq(8) is derived in a way that μ can be extracted by fitting to the experimental data. Eg: B88 (by Becke 1988, later additions by Lee,Yang and Parr - BLYP) uses exact exchange energies of rare gas atoms He through Rn, to get

$$\mu^{BLYP} = 0.2743 \tag{10}$$

2. Non-emperical: Eq(8) is rigorously derived by putting more universal contraints (like the Leib-oxford bound). Eg: PBE (Perdew,Burke,Ernzerhof 1996) found

$$\mu^{PBE} = 0.2195 \tag{11}$$

Point to be noted $\rightarrow$ GGA's are like "Hit and Trial" and mostly are **NOT** based on any physical model.

Properties of GGA

1. GGA-hole satisfies sum rules and $\rho_{xc} \leq 0$ conditions for total as well as the exchange holes.

2. GGA-hole satiesfies the $e^- $-$e^-$ cusp condition in the many-e^- Kohn-Sham wavefunction as it satisfies the sum rule for correlation hole also.

3. GGA functional satisfies the scaling contraints.

4. the Non-emperical-GGA follow the additional Leib-oxford bound.

5. GGA considers the gradients in density and takes into consideration of inhomogenity of e^- density in molecules.

Because of the above properties, we can expect that the GGA should in-principle perform better than the LDA. This is also the case in reality. GGA corrects many of the problems with LDA. For eg: the overbinding problem is significantly reduced in GGA. If we look at the GGA atomisation energies compared to the value of LDA in eqn(3)

$$\Delta_{GGA} = 8\,\text{Kcal/mol}(0.35eV) \tag{12}$$

This is a significant increment in quantitative performance and we can attribute it to the inclusion of inhomogenity in e^- density. Not only the energetic details but many comparative studies have shown significant improvement in other molecular properties like equilibrium bond-distances, bond-dissociation energies.[1] Although in many cases GGA performs better than LDA, in many solid-state applications LDA is still preferred (as here GGA performs same as LDA) and LDA is comutationally inexpensive.

References

1. CECAM summer school(2017) video lectures by Levy,Perdew,Kieron Burke - `https://www.youtube.com/channel/UCfLssAro7SMxgaeKTNFFeeA`

2. ABC of DFT, Kieron Burke

3. A Chemist's Guide to Density Functional Theory, Wolfram Koch, Max C. Holthausen

4. Electronic Structure Calculations for Solids and Molecules- Theory and Computational Methods,Jorge Kohanoff

5. GGA made simple - John P. Perdew, Kieron Burke, Matthias Ernzerhof,Phys. Rev. Lett. 78, 1396 (1997)

---

[1]You can look at many standard textbooks and papers on DFT for detailed data like refrences 3 and 5