The e-e cusp and DFT

Alston J. Misquitta

Centre for Condensed Matter and Materials Physics
Queen Mary, University of London

March 21, 2014

Post-HF methods I

We have looked at three kinds of post-Hartree–Fock methods:

- Configuration Interaction (CI): Variational. Full CI is exact but scales exponentially. Truncated CI (e.g. CISD) not size-consistent.
- Coupled-cluster (CC): Also potentially exact. Truncated methods are size-consistent. CCSD(T) is the method of choice if you can afford the $\mathcal{O}(N^7)$ computational cost.
- Møller–Plesset Perturbation Theory (MPn): Perturbation theory starting from HF reference state. Usually only used at second-order: MP2. Known to diverge. MP2 is the first term to include correlation (MP0 and MP1 are parts of the HF energy).

Post-HF methods II

Practical considerations: CI, CC and MP2 are expensive!

Method	Cost	Description
HF	N^4	Starting point for correlated methods.
MP2	N^5	OK. Useful for optimizations.
CISD	N^6	Not size-consistent.
CCSD(T)	N^7	Very accurate.

All correlated methods require large basis sets with high angular functions.

The electron–electron cusp I

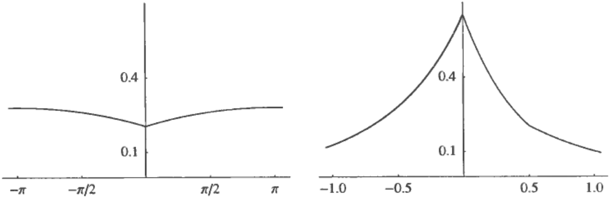


Fig. 7.5. The electronic and nuclear cusps of the ground-state helium atom with one electron fixed at a position $0.5a_0$ from the nucleus (atomic units). On the left, the wave function is plotted on a circle of radius $0.5a_0$ centred at the nucleus; on the right, the wave function is plotted along the axis through the nucleus and the fixed electron.

Correlation is the description of this kink.

All figures from “Molecular Electronic Structure Theory” by Helgaker, Jorgensen and Olsen

The electron–electron cusp II

Principle and partial wave expansions.

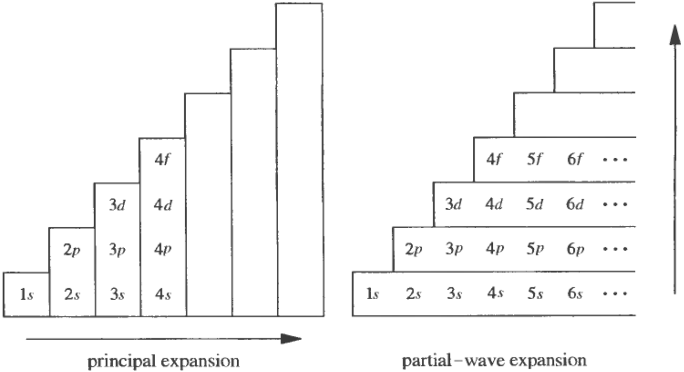
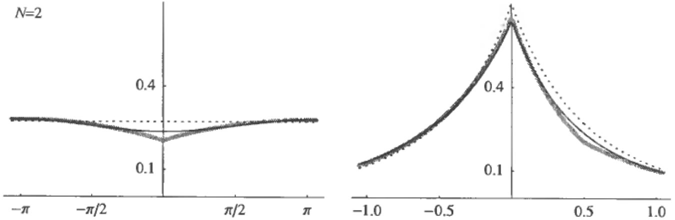


Fig. 7.9. The principal and partial-wave expansions.

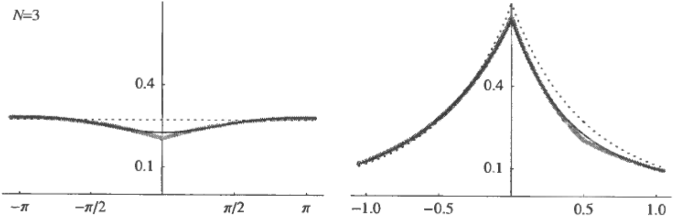
The electron–electron cusp III

He_2: Principle wave expansion with one electron fixed at 0.5Å: $n = 2$



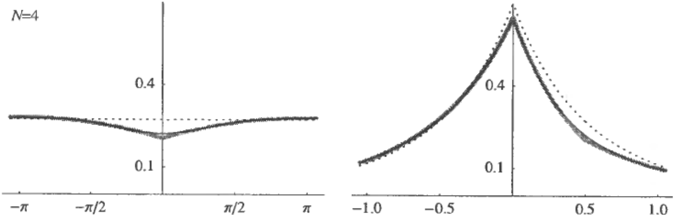
The electron–electron cusp IV

He_2: Principle wave expansion with one electron fixed at 0.5Å:
$n = 3$



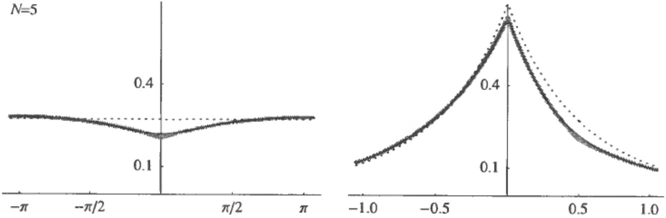
The electron–electron cusp V

He_2: Principle wave expansion with one electron fixed at 0.5Å:
$n = 4$



The electron–electron cusp VI

He_2: Principle wave expansion with one electron fixed at 0.5Å:
$n = 5$



These cusps make correlation hard. Not only do the methods scale poorly with size, but we need rather large basis sets to get sensible results.

Density-Functional Theory I

Hohenberg & Kohn (Phys. Rev. B, **136**, 864 (1964)):

Theorem

H–K Theorem 1 There is a one-to-one mapping between the electronic density and the external potential, and hence, the Hamiltonian:

$$\rho(\mathrm{r}) \iff \mathcal{H}$$

Proof is by *reductio ad absurdum*.

Density-Functional Theory II

Consider the Hamiltonian with the electron–nuclear potential $v_{\rm ext}$ (this is sometimes called the *external* potential):

$$\mathcal{H} = -\frac{1}{2}\sum_{i}^{N}\nabla_i^2 + \sum_{i<j}\frac{1}{r_{ij}} + \sum_{i}^{N} v_{\rm ext}(\rm r_i) \qquad (1)$$

- Let $v_{\rm ext}^1$ and $v_{\rm ext}^2$ arise from the same density.
- We therefore have two Hamiltonians $\mathcal{H}_1$ and $\mathcal{H}_2$ with the same ground state density but with *different* ground state wavefunctions, Ψ_1 and Ψ_2.

Density-Functional Theory III

- Consider $\mathcal{H}_1$: The variational principle states that

$$\begin{aligned} E_0^1 < \langle\Psi_2|\mathcal{H}_1|\Psi_2\rangle &= \langle\Psi_2|\mathcal{H}_2|\Psi_2\rangle + \langle\Psi_2|\mathcal{H}_1 - \mathcal{H}_2|\Psi_2\rangle \\ &= E_0^2 + \langle\Psi_2|v_{\mathrm{ext}}^1 - v_{\mathrm{ext}}^2|\Psi_2\rangle \\ &= E_0^2 + \int \rho(\mathrm{r}) \left[v_{\mathrm{ext}}^1(\mathrm{r}) - \mathrm{v}_{\mathrm{ext}}^2(\mathrm{r})\right] d\mathrm{r} \end{aligned}$$

- Similarly $E_0^2 < E_0^1 + \int \rho(\mathrm{r}) \left[v_{\mathrm{ext}}^2(\mathrm{r}) - \mathrm{v}_{\mathrm{ext}}^1(\mathrm{r})\right] d\mathrm{r}$
- Adding the inequalities:

$$E_0^1 + E_0^2 < E_0^2 + E_0^1 \quad \rightarrow\leftarrow$$

- Hence $\rho(\mathrm{r}) \iff \mathcal{H}$

Density-Functional Theory IV

E. Bright Wilson's observation: To know the Hamiltonian we need to know the number of electrons and position and charge of the nuclei. These can be obtained from the density:

- $N = \int \rho(\mathrm{r})\mathrm{dr}$
- Position and charge of nuclei can be obtained from the cusps:

$$\left.\frac{\partial}{\partial r}\langle\rho(r)\rangle_{\mathrm{sph}}\right|_{r=0} = -2Z\langle\rho(0)\rangle_{\mathrm{sph}}$$

So ρ completely determines the Hamiltonian and therefore the ground-state energy (and also all excited state energies!).

Density-Functional Theory V

What the Hohenberg–Kohn theorem allows us to do is (formally) write the energy as a *functional* of the density.

$$\begin{aligned} E = E[\rho] &= T[\rho] + V_{\mathrm{ee}}[\rho] + V_{\mathrm{en}}[\rho] \\ &= F[\rho] + \int \rho(\mathrm{r})\mathrm{v}_{\mathrm{ext}}(\mathrm{r})\mathrm{dr} \end{aligned}$$

This leads to the second Hohenberg–Kohn theorem:

Theorem

H–K Theorem 2 If $\tilde{\rho}$ is an approximate density then

$$E[\rho] \leq E[\tilde{\rho}]$$

Density-Functional Theory VI

This is how the theorem arises:

- From the previous theorem, we know that $\tilde{\rho}$ determines its own unique potential $\tilde{v}_{\rm ext}$, the Hamiltonian $\tilde{\mathcal{H}}$, and the ground state wavefunction $\tilde{\Psi}$.
- We can now use this wavefunction $\tilde{\Psi}$ as a trial wavefunction for a Hamiltonian $\mathcal{H}$ with external potential $v_{\rm ext}$:
- Using the variational principle

$$E_0 \equiv E[\rho] \leq \langle\tilde{\Psi}|\mathcal{H}|\tilde{\Psi}\rangle = F[\tilde{\rho}] + \int \tilde{\rho}(\mathrm{r})\mathrm{v}_{\rm ext}(\mathrm{r})\mathrm{dr} = \mathrm{E}[\tilde{\rho}]$$

That is for any trial density $\tilde{\rho}$, $E[\rho] \leq E[\tilde{\rho}]$.

So *in principle* we can search over all N-electron densities to find the one that leads to the lowest energy.

Density-Functional Theory VII

This is all simpler than it may look. Consider the minimization procedure in two steps:

$$
\begin{aligned}
E_0 &= \min_{\Psi} \langle \Psi | \hat{T} + \hat{V}_{\mathrm{ee}} + \sum_i^N v(\mathrm{r_i}) | \Psi \rangle \\
&= \min_{\rho} \left(\min_{\Psi \to \rho} \langle \Psi | \hat{T} + \hat{V}_{\mathrm{ee}} + \sum_i^N v(\mathrm{r_i}) | \Psi \rangle \right) \\
&= \min_{\rho} \left(\min_{\Psi \to \rho} \left[\langle \Psi | \hat{T} + \hat{V}_{\mathrm{ee}} | \Psi \rangle + \int v(\mathrm{r}) \rho(\mathrm{r}) \mathrm{dr} \right] \right)
\end{aligned}
$$

In the second line the inner minimization is constraint to all wavefunctions that give $\rho(\mathrm{r})$, while in the outer minimization this condition is removed by searching all $\rho(\mathrm{r})$.

Density-Functional Theory VIII

The first term in the square brackets is the functional $F[\rho]$, therefore

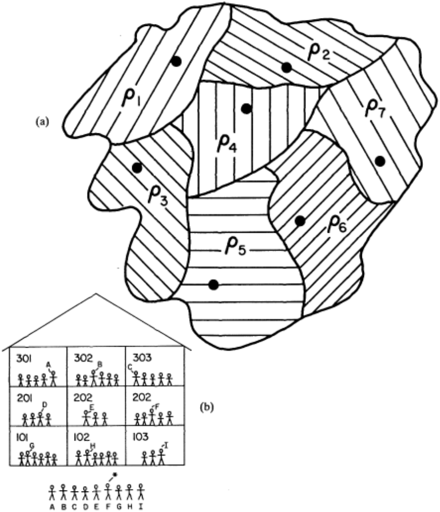
$$\begin{aligned} E_0 &= \min_{\rho} \left(F[\rho] + \int v(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} \right) \\ &= \min_{\rho} E[\rho] \end{aligned}$$

where

$$E[\rho] = F[\rho] + \int v(\mathrm{r})\rho(\mathrm{r})\mathrm{dr}$$

This double-step minimization is illustrated in the following figure (from Parr and Yang, *Density-Functional Theory*).

Density-Functional Theory IX



Density-Functional Theory X

In practice searching over all N-electron densities is next to impossible. How do we do this? If we are given a trial N-electron density $\tilde{\rho}$, how can we obtain the corresponding external potential $\tilde{v}_{\rm ext}$? This can be done for a one or two electron system (Ex. How?) but not in general.
All this would have been a curiosity had it not been for a paper by Kohn & Sham published in 1965 (Phys. Rev. A **140**, 1133) which gave us what we now know as Kohn–Sham DFT.
But before getting to Kohn–Sham DFT, let's look at a couple of other attempts at formulating a density-functional theory...

Thomas–Fermi Theory I

In 1927 Enrico Fermi and Llewellyn Thomas developed the first true density functional for the energy. In 1928, P. A. M. Dirac modified the TF model to include the effects of exchange. This TFD model turned out to be poorer than the original model. As we shall see, the beauty of the TF/TFD models is their simplicity. However, they proved to be not so accurate, and, in particular, could not describe chemical bonding. So what's their use? These models, as precursors to modern DFT, contain ideas that will form the basis of the density functionals we use today.

We will develop a quick and elegant derivation of the TFD model here, at the expense of mathematical rigour.

First we re-visit **uniform scaling of the electronic coordinates...**

Thomas–Fermi Theory II

Theorem

Under uniform scaling of the electronic coordinates:

$$
\begin{aligned}
r_i &\to \alpha r_i, \\
\psi(r_i) &\to \psi_\alpha(r_i) = \alpha^{3N/2}\psi(\alpha r_i).
\end{aligned}
$$

This form of the scaling is needed to ensure normalization.

Q: Show that $\rho_\alpha(\mathrm{r}) = \alpha^3\rho(\alpha\mathrm{r})$.

We had proved the identities:

- $\langle\psi_\alpha|T|\psi_\alpha\rangle = \alpha^2\langle\psi|T|\psi\rangle$
- $\langle\psi_\alpha|V(R)|\psi_\alpha\rangle = \alpha\langle\psi|V(\alpha R)|\psi\rangle$

Thomas–Fermi Theory III

Anticipating a DFT-terminology we will define

- $T[\Psi] = \langle\Psi|\hat{T}|\Psi\rangle \equiv T[\rho]$
- $V_{\text{ee}}[\Psi] = \langle\Psi|\hat{V}_{\text{ee}}|\Psi\rangle \equiv V_{\text{ee}}[\rho]$

Q: How do we justify terming these functionals of the density?

With this notation, the scaling identities can be written as

- $T[\Psi_\alpha] = \alpha^2 T[\Psi]$
- $V_{\text{ee}}[\Psi_\alpha] = \alpha V_{\text{ee}}[\Psi]$

Alternatively we could write these as functionals of ρ.

Thomas–Fermi Theory IV

Our task now is to derive and expression for $T[\rho]$. To do this we assume that the kinetic energy functional can be written in the form

$$T[\rho] = \int t(\rho) d\mathrm{r}$$

where $t(\rho)$ is a homogeneous function of the density. That is, t is of the form $t(x) = Ax^b$ where b is a real number. Why do we assume this can be done? Well, we could state it as an *ansatz*. But there is another reason why it is possible: it is possible to make this derivation using ideas from the free-electron gas. We won't do this here.

Thomas–Fermi Theory V

Consider

$$\begin{aligned}
T[\rho_\alpha] &= \int t(\rho_\alpha(\mathrm{r}))\mathrm{dr} \\
&= \int t(\alpha^3\rho(\alpha\mathrm{r}))\mathrm{dr} \\
&= \alpha^{-3}\int t(\alpha^3\rho(\alpha\mathrm{r}))\mathrm{d}(\alpha\mathrm{r}) \\
&= \alpha^{-3}\int t(\alpha^3\rho(\mathrm{r}))\mathrm{dr}
\end{aligned}$$

where, in the last step we have made a change of variables. But we also have $T[\rho_\alpha] = \alpha^2 T[\rho]$, therefore

$$\alpha^{-3}\int t(\alpha^3\rho(\mathrm{r}))\mathrm{dr} = \alpha^2\int \mathrm{t}(\rho(\mathrm{r}))\mathrm{dr}$$

Thomas–Fermi Theory VI

Therefore

$$t(\alpha^3\rho(\mathrm{r})) = \alpha^5 t(\rho)$$

or, by substituting $\alpha' = \alpha^3$ and then dropping the primes,

$$t(\alpha\rho(\mathrm{r})) = \alpha^{5/3} t(\rho)$$

That is, t is a homogeneous function of degree $5/3$ in ρ: $t(\rho) = A\rho^{5/3}$.

Therefore, using $A = C_F$,

$$T[\rho] = C_F \int \rho^{5/3}(\mathrm{r})\mathrm{dr}$$

where $C_F = \frac{3}{10}(3\pi^2)^{2/3} = 2.871\cdots$.

Thomas–Fermi Theory VII

What about the functional $V_{\mathrm{ee}}[\rho]$?
It consists of two parts: $V_{\mathrm{ee}}[\rho] = J[\rho] + K[\rho]$. We already know the exact form of the Coulomb part:

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathrm{r}_1)\rho(\mathrm{r}_2)}{|\mathrm{r}_1 - \mathrm{r}_2|} d\mathrm{r}_1 \mathrm{d}\mathrm{r}_2$$

The exchange energy functional $K[\rho]$, as part of the full electron–electron energy $V_{\mathrm{ee}}[\rho]$, also involves the operator $1/r$, so the two share the same scaling relation, i.e.,

- $K[\rho_\alpha] = \alpha K[\rho]$

Thomas–Fermi Theory VIII

We need to find a density functional for this and we do it the same way as we did the kinetic energy functional: Assume

$$K[\rho] = \int k(\rho(\mathrm{r}))\mathrm{dr}$$

where k is a homogeneous function of ρ and use the scaling relation to find the form of k.

Q: Show that

$$K[\rho] = -C_X \int \rho^{4/3}(\mathrm{r})\mathrm{dr},$$

where the constant $C_X = \frac{3}{4}(\frac{3}{\pi})^{1/3}$.

Why is this energy defined to be negative?

Thomas–Fermi Theory IX

Now we define the Thomas–Fermi functional (no exchange in this one):

$$\begin{aligned} E_{\mathrm{TF}}[\rho] &= T[\rho] + J[\rho] + \int \rho(\mathrm{r})\mathrm{v}_{\mathrm{ext}}(\mathrm{r})\mathrm{dr} \\ &= C_F \int \rho^{5/3}(\mathrm{r})\mathrm{dr} + \frac{1}{2}\iint \frac{\rho(\mathrm{r}_1)\rho(\mathrm{r}_2)}{|\mathrm{r}_1 - \mathrm{r}_2|}\mathrm{dr}_1\mathrm{dr}_2 + \int \rho(\mathrm{r})\mathrm{v}_{\mathrm{ext}}(\mathrm{r})\mathrm{dr} \end{aligned}$$

We minimize $E_{\mathrm{TF}}[\rho]$ subject to the condition $N = \int \rho(\mathrm{r})\mathrm{dr}$ using the Lagrange multiplier μ_{TF}:

$$\delta\left[E_{\mathrm{TF}}[\rho] - \mu_{\mathrm{TF}}\left(\int \rho(\mathrm{r})\mathrm{dr} - \mathrm{N}\right)\right] = 0$$

Thomas–Fermi Theory X

Q: Show that we get

$$\mu_{\mathrm{TF}} = \frac{\delta E_{\mathrm{TF}}}{\delta \rho(\mathrm{r})} = \frac{5}{3} C_F \rho^{2/3}(\mathrm{r}) - \phi(\mathrm{r})$$

where $\phi(\mathrm{r}) = \mathrm{v}_{\mathrm{ext}}(\mathrm{r}) - \int \frac{\rho(\mathrm{r}')}{|\mathrm{r}-\mathrm{r}'|} \mathrm{dr}'$ is the electrostatic potential.

This can be solved (we will not do it) quite easily for atoms.
All Dirac did was to modify the TF functional to include the exchange functional $K[\rho]$. This turned out to result in somewhat worse agreement with the more accurate Hartree–Fock energies.

Thomas–Fermi Theory XI

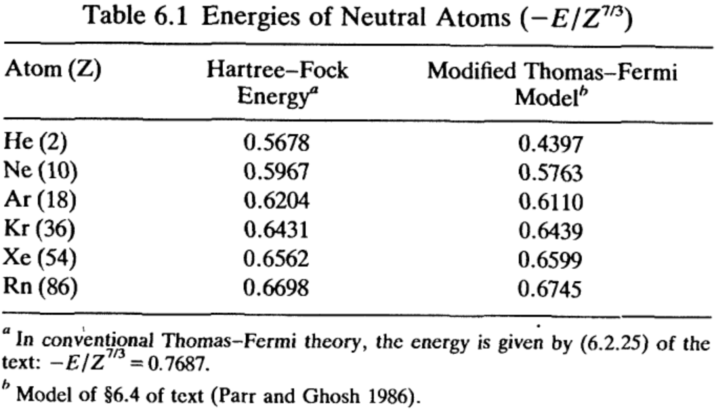


Table 6.1 Energies of Neutral Atoms ($-E/Z^{7/3}$)

Atom (Z)	Hartree–Fock Energy[a]	Modified Thomas–Fermi Model[b]
He (2)	0.5678	0.4397
Ne (10)	0.5967	0.5763
Ar (18)	0.6204	0.6110
Kr (36)	0.6431	0.6439
Xe (54)	0.6562	0.6599
Rn (86)	0.6698	0.6745

[a] In conventional Thomas–Fermi theory, the energy is given by (6.2.25) of the text: $-E/Z^{7/3} = 0.7687$.

[b] Model of §6.4 of text (Parr and Ghosh 1986).

Thomas–Fermi Theory XII

Another problem with TF/TFD theory is that they do not result in any shell structure. Here's the density of Argon obtained using Hartree–Fock and various modifications of the TFD model:

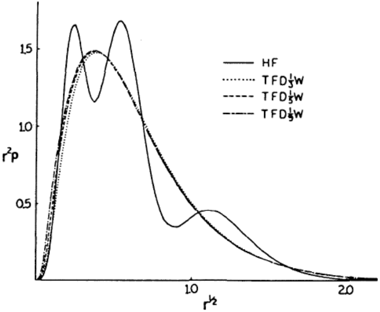


Figure 6.3 Electron density for argon from various models (after Yang 1986.)

Hartree–Fock again I

Before getting on with DFT, let's have another look at Hartree–Fock, but this time, from a slightly different angle. The HF energy can be written as

$$\begin{aligned} E_{\mathrm{HF}} &= \min_{|\Psi_{\mathrm{SD}}\rangle \to N} \langle \Psi_{\mathrm{SD}} | \hat{T} + \hat{V}_{\mathrm{ne}} + \hat{V}_{\mathrm{ee}} | \Psi_{\mathrm{SD}} \rangle \\ &= \min_{|\Psi_{\mathrm{SD}}\rangle \to N} \langle \Psi_{\mathrm{SD}} | \hat{T} + v_{\mathrm{ext}} + \hat{V}_{\mathrm{ee}} | \Psi_{\mathrm{SD}} \rangle \end{aligned}$$

Here $|\Psi_{\mathrm{SD}}\rangle$ is our Slater determinant that yields an N electron density. In the second line I have used our notation for the electron–nuclear potential: the external potential.

Hartree–Fock again II

This minimization procedure gives rise to an effective Hamiltonian — the Fock operator:

$$f^{\mathrm{HF}}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{\alpha} \frac{Z_\alpha}{r_{i\alpha}} + v^{\mathrm{HF}}(i) \tag{2}$$

The HF potential is an effective potential that contains the effective electron–electron Coulomb and exchange interactions (no correlation!). To get us ready for Kohn–Sham theory, we will make a few changes to the above equation.

Hartree–Fock again III

- First of all recognise that the second term in the Fock operator is just the external potential: $v_{\rm ext}$
- Next, split the HF operator into its Coulomb and Exchange parts. These are usually labeled by 'J' and 'X', respectively:

$$v^{\rm HF} = v_{\rm J} + v_{\rm X}$$

So our Fock operator is now written as

$$f^{\rm HF}(i) = -\frac{1}{2}\nabla_i^2 + v_{\rm ext}(i) + v_{\rm J}(i) + v_{\rm X}(i) \tag{3}$$

Hartree–Fock again IV

In some sense we can say that the HF ground state energy is a functional of the density and can be written as

$$E_{\mathrm{HF}}[\rho^{\mathrm{HF}}] = T_{\mathrm{S}}[\rho^{\mathrm{HF}}] + J[\rho^{\mathrm{HF}}] + E_{\mathrm{x}}^{\mathrm{HF}}[\rho^{\mathrm{HF}}] + \int \rho^{\mathrm{HF}}(\mathrm{r})\mathrm{v}_{\mathrm{ext}}(\mathrm{r})\mathrm{dr} \quad (4)$$

where the *non-interacting* Kinetic energy functional is

$$T_{\mathrm{S}}[\rho^{\mathrm{HF}}] = -\frac{1}{2}\sum_{i=1}^{N}\langle\chi_i|\nabla_i^2|\chi_i\rangle \quad (5)$$

which is not strictly a functional of the density, but since the density is implicitly a functional of the orbitals we can still think of the KE functional as a density functional.

Hartree–Fock again V

The Coulomb energy functional is defined as

$$J[\rho^{\mathrm{HF}}] = \frac{1}{2} \iint \frac{\rho^{\mathrm{HF}}(\mathrm{r}_1)\rho^{\mathrm{HF}}(\mathrm{r}_2)}{r_{12}} d\mathrm{r}_1 \mathrm{d}\mathrm{r}_2 \tag{6}$$

The exchange energy functional is non-local and is, like the KE functional, dependent on the HF orbitals:

$$E_{\mathrm{x}}^{\mathrm{HF}}[\rho^{\mathrm{HF}}] = -\sum_{a\in\mathrm{occ}} \sum_{b\in\mathrm{occ}} \iint dx_1 dx_2 \chi_a^*(1) \frac{\chi_b^*(2)\mathcal{P}_{12}\chi_b(2)}{r_{12}} \chi_a(1) \tag{7}$$

where $\mathcal{P}_{12}$ is the permutation operator.

Hartree–Fock again VI

Looked at in this way HF theory is a kind of density functional theory, but one that is not, even in principle, exact (except for 1-electron systems).

In this formulation of HF theory, we define the HF energy as

$$E^{\rm HF} = \min_{\rho \to |\Psi_{\rm SD}\rangle, N} E_{\rm HF}[\rho]$$

That is, we minimize the functional $E_{\rm HF}[\rho]$ over all N-electron densities that arise from a Slater determinant. We need to impose the N-electron constraint and this is done using Lagrange multipliers. We minimize the functional:

$$\Omega[\rho] = E_{\rm HF}[\rho] - \sum_{i,j}^{N} \epsilon_{ij}(\langle \chi_i | \chi_j \rangle - \delta_{ij})$$

Hartree–Fock again VII

Since $\Omega[\rho]$ really is a functional of the orbitals, we need to perform the variation over the orbitals and use the chain rule:

$$\frac{\delta}{\delta\chi_i(\mathrm{r})} = \frac{\delta}{\delta\rho(\mathrm{r})}\frac{\delta\rho(\mathrm{r})}{\delta\chi_i(\mathrm{r})} \tag{8}$$

I do not want to get is bogged down with functional derivatives, but one example of how its done could be illuminating:
First of all, since $\rho(\mathrm{r}) = \sum_{\mathrm{i}=1}^{\mathrm{N}} \chi_{\mathrm{i}}^*(\mathrm{r})\chi_{\mathrm{i}}(\mathrm{r})$ we get

$$\frac{\delta\rho(\mathrm{r})}{\delta\chi_i^*(\mathrm{r})} = \chi_i(\mathrm{r})$$

Hartree–Fock again VIII

Now consider the functional derivative of $J[\rho]$:

$$\begin{aligned}
\frac{\delta J[\rho]}{\delta \chi_i^*(\mathrm{r})} &= \frac{\delta J[\rho]}{\delta \rho(\mathrm{r})} \frac{\delta \rho(\mathrm{r})}{\delta \chi_i^*(\mathrm{r})} \\
&= \frac{\delta}{\delta \rho(\mathrm{r})} \left[\frac{1}{2} \iint \frac{\rho(\mathrm{r})\rho(\mathrm{r}')}{|\mathrm{r}-\mathrm{r}'|} d\mathrm{r}d\mathrm{r}' \right] \times \chi_i(\mathrm{r}) \\
&= \int \frac{\rho(\mathrm{r}')}{|\mathrm{r}-\mathrm{r}'|} d\mathrm{r}' \chi_{\mathrm{i}}(\mathrm{r}) \\
&= v_{\mathrm{J}}(\mathrm{r}) \chi_i(\mathrm{r})
\end{aligned}$$

Do this for all terms and we get back our Fock equation (in a generalised form).

Hartree–Fock again IX

Prove that the variation of

$$\Omega[\rho] = E_{\mathrm{HF}}[\rho] - \sum_{i,j}^{N} \epsilon_{ij}(\langle \chi_i | \chi_j \rangle - \delta_{ij})$$

with respect to χ_i^* does lead to the Fock equations:

Q:

$$f|\chi_i\rangle = \sum_j \epsilon_{ji} |\chi_j\rangle$$

This is the non-canonical form of the Fock equations. As we saw in the Hartree–Fock lectures, with a suitable unitary transformation we can recover the canonical form:

$$f|\chi_i\rangle = \epsilon_i |\chi_i\rangle$$

Kohn–Sham DFT I

From the two Hohenberg–Kohn theorems we know we can write the ground-state energy as a functional of the density:

$$E[\rho] = T[\rho] + V_{\mathrm{ee}}[\rho] + \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{d}\mathrm{r}$$

The problem is that we do not know how to define the first two functionals.

We have seen that Thomas–Fermi theory gives us an expression for $T[\rho]$, but we also saw that this expression was too approximate to be of use for atomic/molecular systems as it does not give rise to the shell structure or chemical bonding.

Kohn–Sham DFT II

In 1965 Walter Kohn & Lu Sham put DFT on a practical footing through what is now known as Kohn–Sham DFT. The began by *postulating* the existence of a non-interacting system with external potential $v_{\rm S}$ that yields the *exact density*. That is

$$\mathcal{H} = \sum_i \left(-\frac{1}{2}\nabla_i^2 + v_{\rm S}(i) \right)$$

has a ground state single determinant solution with density ρ. This is a non-interating system so we can solve it just as we solved the Fock Hamiltonian. The results will be a set of orbitals $\{\chi_i\}$ and orbital eigenvalues $\{\epsilon_i\}$.

Kohn–Sham DFT III

Just as we wrote the kinetic energy in Hartree–Foct, we write the kinetic energy of this Kohn–Sham non-interacting system:

$$T_{\mathrm{S}}[\rho] = \sum_{i}^{N} -\frac{1}{2}\langle \chi_i | \nabla^2 | \chi_i \rangle$$

What Kohn & Sham did was to state that this non-interacting kinetic energy functional $T_{\mathrm{S}}[\rho]$ could be considered a good approximation to the true functional $T[\rho]$.
Impotantly, they knew (from Hartree–Fock theory) that $T_{\mathrm{S}}[\rho]$, as it depended on the orbitals, could describe the molecular shell structure.

Kohn–Sham DFT IV

Using $T_{\rm S}[\rho]$, we can write the Hohenberg–Kohn functional as

$$\begin{aligned} F[\rho] &= T[\rho] + V_{\rm ee}[\rho] \\ &= T_{\rm S}[\rho] + J[\rho] + E_{\rm xc}[\rho] \end{aligned}$$

where this eXchange-Correlation functional is defined as

$$E_{\rm xc}[\rho] = (T[\rho] - T_{\rm S}[\rho]) - (V_{\rm ee}[\rho] - J[\rho])$$

The idea here is that we have defined as much as we could define ($T_{\rm S}[\rho]$ and $J[\rho]$), and have swept the rest of the energy into the unknown functional $E_{\rm xc}[\rho]$.

Kohn–Sham DFT V

So we get the following functional for the ground state energy:

$$E[\rho] = T_{\mathrm{S}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} \qquad (9)$$

Compare this to what we had for the Hartree–Fock functional:

$$E_{\mathrm{HF}}[\rho^{\mathrm{HF}}] = T_{\mathrm{S}}[\rho^{\mathrm{HF}}] + J[\rho^{\mathrm{HF}}] + E_{\mathrm{x}}^{\mathrm{HF}}[\rho^{\mathrm{HF}}] + \int \rho^{\mathrm{HF}}(\mathrm{r})\mathrm{v}_{\mathrm{ext}}(\mathrm{r})\mathrm{dr}$$

The difference between the two is that $E_{\mathrm{HF}}[\rho^{\mathrm{HF}}]$ is necessarily approximate while $E[\rho]$ is exact, albeit, *in principal*.

Kohn–Sham DFT VI

$$\boxed{E[\rho] = T_{\mathrm{S}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{d}\mathrm{r} \qquad (10)}$$

must be minimized subject to the orthonormality constraints

$$\langle \chi_i | \chi_j \rangle = \delta_{ij}.$$

As before, we include these constraints using Lagrange multipliers and minimize

$$\Omega[\rho] = E[\rho] - \sum_{i,j}^{N} \epsilon_{ij}(\langle \chi_i | \chi_j \rangle - \delta_{ij})$$

Kohn–Sham DFT VII

This gives us the Kohn–Sham equations after the usual occupied orbital rotation to make the eigenvalue matrix ϵ_{ij} diagonal:

$$\left(-\frac{1}{2}\nabla_i^2 + v_{\mathrm{S}}(\mathrm{r})\right)\chi_i = \epsilon_i \chi_i$$

where the effective potential is defined as

$$v_{\mathrm{S}}(\mathrm{r}) = \mathrm{v}_{\mathrm{J}}(\mathrm{r}) + \mathrm{v}_{\mathrm{ext}}(\mathrm{r}) + \mathrm{v}_{\mathrm{xc}}(\mathrm{r})$$

Kohn–Sham DFT VIII

The various potentials that enter this expression are:

- v_{J}: The Coulomb potential defined as:

$$v_{\mathrm{J}}(\mathrm{r}) = \int \frac{\rho(\mathrm{r}')}{|\mathrm{r} - \mathrm{r}'|} \mathrm{dr}'$$

- v_{ext}: The external potential, i.e., the electron-nuclear potential:

$$v_{\mathrm{ext}}(\mathrm{r}) = -\sum_{\alpha} \frac{\mathrm{Z}_{\alpha}}{|\mathrm{r} - \mathrm{R}_{\alpha}|}$$

Kohn–Sham DFT IX

- $v_{\rm xc}$: The exchange-correlation potential which is defined through the exchange-correlation energy $E_{\rm xc}[\rho]$ as:

$$v_{\rm xc}({\rm r}) = \frac{\delta {\rm E}_{\rm xc}[\rho]}{\delta \rho}$$

Let's see how these potentials arise.

Kohn–Sham DFT X

Since $\Omega[\rho]$ really is a functional of the orbitals, we need to perform the variation over the orbitals and use the chain rule:

$$\frac{\delta}{\delta\chi_i(\mathrm{r})} = \frac{\delta}{\delta\rho(\mathrm{r})}\frac{\delta\rho(\mathrm{r})}{\delta\chi_i(\mathrm{r})} \tag{11}$$

First of all, since $\rho(\mathrm{r}) = \sum_{\mathrm{i=1}}^{\mathrm{N}} \chi_{\mathrm{i}}^*(\mathrm{r})\chi_{\mathrm{i}}(\mathrm{r})$ we get

$$\frac{\delta\rho(\mathrm{r})}{\delta\chi_i^*(\mathrm{r})} = \chi_i(\mathrm{r})$$

Kohn–Sham DFT XI

Now consider the functional derivative of $J[\rho]$:

$$\begin{aligned}\frac{\delta J[\rho]}{\delta \chi_i^*(\mathrm{r})} &= \frac{\delta J[\rho]}{\delta \rho(\mathrm{r})}\frac{\delta \rho(\mathrm{r})}{\delta \chi_i^*(\mathrm{r})}\\ &= \frac{\delta}{\delta \rho(\mathrm{r})}\left[\frac{1}{2}\iint \frac{\rho(\mathrm{r})\rho(\mathrm{r}')}{|\mathrm{r}-\mathrm{r}'|}d\mathrm{r}\mathrm{d}\mathrm{r}'\right] \times \chi_i(\mathrm{r})\\ &= \int \frac{\rho(\mathrm{r}')}{|\mathrm{r}-\mathrm{r}'|}d\mathrm{r}'\chi_{\mathrm{i}}(\mathrm{r})\\ &= v_{\mathrm{J}}(\mathrm{r})\chi_i(\mathrm{r})\end{aligned}$$

Kohn–Sham DFT XII

Now consider the functional derivative of $E_{\rm ext}[\rho] = \int v_{\rm ext}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr}$:

$$\begin{aligned}\frac{\delta E_{\rm ext}[\rho]}{\delta\chi_i^*(\mathrm{r})} &= \frac{\delta E_{\rm ext}[\rho]}{\delta\rho(\mathrm{r})}\frac{\delta\rho(\mathrm{r})}{\delta\chi_i^*(\mathrm{r})} \\ &= \frac{\delta}{\delta\rho(\mathrm{r})}\left[\int v_{\rm ext}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr}\right] \times \chi_i(\mathrm{r}) \\ &= v_{\rm ext}(\mathrm{r})\chi_i(\mathrm{r})\end{aligned}$$

We cannot evaluate the functional derivative of $E_{\rm xc}[\rho]$ as we still do not know the form of this functional. So we simply define:

$$v_{\rm xc}(\mathrm{r}) = \frac{\delta \mathrm{E}_{\rm xc}[\rho]}{\delta\rho}$$

Kohn–Sham DFT XIII

So the Kohn–Sham non-interacting potential is defined as:

$$\boxed{\begin{aligned} v_{\mathrm{S}}(\mathrm{r}) &= v_{\mathrm{J}}(\mathrm{r}) + \mathrm{v}_{\mathrm{ext}}(\mathrm{r}) + \mathrm{v}_{\mathrm{xc}}(\mathrm{r}) \\ &= \int \frac{\rho(\mathrm{r}')}{|\mathrm{r}-\mathrm{r}'|} d\mathrm{r}' + \mathrm{v}_{\mathrm{ext}}(\mathrm{r}) + \frac{\delta \mathrm{E}_{\mathrm{xc}}[\rho]}{\delta \rho} \end{aligned}}$$

We solve the 1-electron Kohn–Sham equations self-consistently:

$$\boxed{k(1)\chi_i(1) = \left(-\frac{1}{2}\nabla_1^2 + v_{\mathrm{S}}(1)\right)\chi_i(1) = \epsilon_i \chi_i(1)}$$

where we have defined the Kohn–Sham operator $k(1)$.

Kohn–Sham DFT XIV

Once we have the orbitals χ_i, we can evaluate the energy using

$$\boxed{E[\rho] = T_{\mathrm{S}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr}}$$

Kohn–Sham DFT XV

or, equivalently, we use the results we derived when studying Hartree–Fock to realise that

$$\begin{aligned}
\sum_i \epsilon_i &= \langle \Psi_0 | \sum_i k(i) | \Psi_0 \rangle \\
&= \langle \Psi_0 | \sum_i \left(-\frac{1}{2}\nabla_i^2 + v_{\mathrm{S}}(i) \right) | \Psi_0 \rangle \\
&= \langle \Psi_0 | \sum_i -\frac{1}{2}\nabla_i^2 | \Psi_0 \rangle + \langle \Psi_0 | \sum_i v_{\mathrm{S}}(i) | \Psi_0 \rangle \\
&= T_{\mathrm{S}}[\rho] + \langle \Psi_0 | \sum_i v_{\mathrm{J}}(i) + v_{\mathrm{ext}}(i) + v_{\mathrm{xc}}(i) | \Psi_0 \rangle \\
&= T_{\mathrm{S}}[\rho] + 2J[\rho] + \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} + \int \mathrm{v}_{\mathrm{xc}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr}
\end{aligned}$$

Kohn–Sham DFT XVI

Using this, we can write an expression for the kinetic energy functional:

$$\boxed{T_{\mathrm{S}}[\rho] = \sum_i \epsilon_i - 2J[\rho] - \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} - \int \mathrm{v}_{\mathrm{xc}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr}}$$

And then we can write the total energy:

$$\boxed{\begin{aligned} E[\rho] &= T_{\mathrm{S}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} \\ &= \sum_i \epsilon_i - J[\rho] + E_{\mathrm{xc}}[\rho] - \int v_{\mathrm{xc}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} \end{aligned}}$$

Kohn–Sham DFT XVII

- In a sense, what Kohn & Sham did was to sweep all the unknowns under the rug. But they did this intelligently as they had a good idea of how to approximate the unknown bits: the exchange-correlation energy and its functional derivative.
- Also, they knew that their formalism was *in principle* exact as they had proved various theorems to that effect.
- It turned out that their proofs were not mathematically sound, but this was fixed by others.
- The rest of the DFT story is how we find the exchange-correlation functional $E_{\mathrm{xc}}[\rho]$.

Density Functionals I

Exchange correlation functionals are usually written in the form

$$E_{\mathrm{xc}}[\rho] = \int \rho(\mathrm{r})\epsilon_{\mathrm{xc}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots)\mathrm{dr} \tag{12}$$

where $\epsilon_{\mathrm{xc}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots)$ can be regarded as the exchange-correlation density.
We usually split the exchange-correlation density into its exchange and correlation parts:

$$\epsilon_{\mathrm{xc}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots) = \epsilon_{\mathrm{x}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots) + \epsilon_{\mathrm{c}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots) \tag{13}$$

This separation is convenient for we can then think of using well-understood approximations for each of these.

LDA I

The first of the many functionals is the local density approximation or LDA. In this approximation the exchange-correlation density depends on the electron density alone (no dependence on gradients etc.):

$$E_{\mathrm{xc}}[\rho] = \int \rho(\mathrm{r})\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho(\mathrm{r}))\mathrm{dr} \tag{14}$$

The Slater approximation is used for the exchange-energy density:

$$\epsilon_{\mathrm{x}}^{\mathrm{S}}(\rho(\mathrm{r})) = -\frac{3}{4}\left(\frac{3\rho(\mathrm{r})}{\pi}\right)^{3/2}$$

Using this we get the Slater exchange functional:

$$E_{\mathrm{x}}^{\mathrm{S}}[\rho] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{3/2}\int \rho^{4/3}(\mathrm{r})\mathrm{dr} = -\mathrm{C_x}\int \rho^{4/3}(\mathrm{r})\mathrm{dr}$$

LDA II

Paired with this is a correlation functional parameterized on very accurate quantum Monte-Carlo (QMC) calculations of the energy of the homogeneous free electron gas as a function of density.

There are a variety of correlation parameterizations. These differ by the choice of QMC energies used or by the interpolation scheme used in the parameterization (the QMC energies are calculated at a set of densities so some scheme is required to interpolate to all densities).

Common choices of the correlation functional are:

- PW91c The Perdew–Wang (1992) parameterization (called `pw91lda` in NWChem.
- VWN The Voski–Wilk–Nusair (1980) parameterization.

So what is called the LDA translates into a combination of the Slater exchange functional and one of these correlation functionals. The actual choice will vary with program.

LDA III

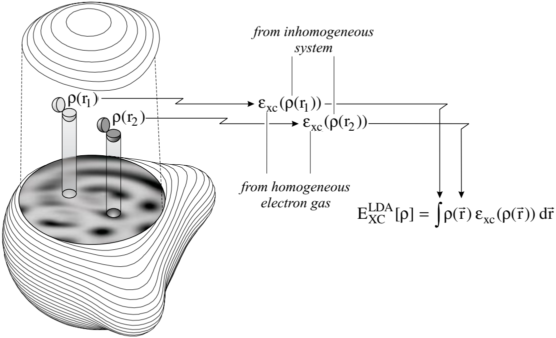


Figure 6-2. The local density approximation.

From Koch & Holthausen *A Chemist's Guide to density Functional Theory* (2001).

LDA IV

Generalization to open-shell systems (local spin-density approximation (LSD)): the exchange-correlation density depends on the spin-up and spin-down densities:

$$E_{\mathrm{xc}}[\rho_\alpha, \rho_\beta] = \int \rho(\mathrm{r})\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho_\alpha(\mathrm{r}), \rho_\beta(\mathrm{r}))\mathrm{dr} \qquad (15)$$

- GOOD LDA is better than HF. Good equilibrium geometries, harmonic frequencies.
- BAD Energetics very poor. Errors in atomization energies 36 kcal/mol. (HF has errors of 78 kcal/mol on same set of molecules)

GGA I

In the generalized gradient approximations the exchange-correlation density is dependent on the density and its gradient.

$$E_{\mathrm{xc}}[\rho] = \int \rho(\mathrm{r})\epsilon_{\mathrm{xc}}^{\mathrm{GGA}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}))\mathrm{dr}$$

As before, we split the exchange-correlation density into its exchange and correlation parts:

$$\epsilon_{\mathrm{xc}}^{\mathrm{GGA}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r})) = \epsilon_{\mathrm{x}}^{\mathrm{GGA}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r})) + \epsilon_{\mathrm{c}}^{\mathrm{GGA}}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}))$$

The exchange part of all GGAs takes the form

$$E_{\mathrm{x}}^{\mathrm{GGA}}[\rho] = \int \rho(\mathrm{r})\epsilon_{\mathrm{x}}^{\mathrm{LDA}}(\rho(\mathrm{r}))\mathrm{F}_{\mathrm{x}}(\mathrm{s})\mathrm{dr}$$

GGA II

$F_{\mathrm{X}}(s)$ is sometimes called the enhancement factor and is written as a function of the reduced density gradient defined as

$$s(\mathrm{r}) = \frac{|\nabla \rho(\mathrm{r})|}{2(3\pi^2)^{1/3}\rho^{4/3}(\mathrm{r})}.$$

Note that in general all quantities will depend on spin.

This definition is used so as to make s dimensionless (Q: Show this!) and means that s will be large when the gradient of the density is large (where the LDA should fail) and also where the density is small (in the region of the density tails).

Two of the common exchange enhancement factors are

GGA III

- Becke, 1988 (B88)

$$F_{\mathrm{x}}^{\mathrm{B88}}(s) = 1 - \frac{\beta s^2}{1 + 6\beta s \sinh^{-1} s}$$

Becke fitted the parameter $\beta = 0.0042$ to reproduce known exchange energies of rare gas atoms. This particular form for the enhancement factor was chosen to obey a few exact relations.

- Perdew, Burke & Ernzerhof, 1996 (PBE)

$$F_{\mathrm{x}}^{\mathrm{PBE}}(s) = 1 + \kappa - \frac{\kappa}{1 - \mu s^2/\kappa}$$

In this functional all parameters were obtained theoretically. $\kappa = 0.804$. Most physcists use this exchange functional.

Hybrid Functionals I

The exchange contribution to the energy is much larger than the correlation energy (which is why HF is not too bad!). So why not use the best exchange energy we have — from HF, usually termed *exact* exchange in this context — and combine it with the best correlation funtional available:

$$E_{\mathrm{xc}}[\rho] = E_{\mathrm{x}}^{\mathrm{HF}}[\rho] + E_{\mathrm{c}}[\rho]$$

This turns out to be better than HF, but much worse than the GGAs.

The precise reason for this somewhat poor behaviour has to do with the nature of the exchange hole. The exact x-hole is local, but the HF x-hole is non-local. So it must be corrected by a non-local correlation hole, but the DFT correlation hole is also local and so cannot correct the HF x-hole. This is all very interesting but also very technical.

Hybrid Functionals II

In 1993, Becke proposed a three-parameter semi-empirical functional that cured this problem. The general idea is to mix some fraction of HF exchange with DFT exchange:

$$E_{\mathrm{xc}}^{\mathrm{hybrid}} = aE_{\mathrm{x}}^{\mathrm{HF}} + (1-a)E_{\mathrm{x}}^{\mathrm{GGA}} + E_{\mathrm{c}}^{\mathrm{GGA}}$$

The B3LYP is the most widely used of these and is a slight modification of Becke's 1993 proposal made the following year by Stephens and others:

$$\begin{aligned} E_{\mathrm{xc}}^{\mathrm{B3LYP}} =& E_{\mathrm{xc}}^{\mathrm{SVWN}} + a_0(E_{\mathrm{x}}^{\mathrm{HF}} - E_{\mathrm{x}}^{\mathrm{S}}) + a_x(E_{\mathrm{x}}^{\mathrm{B88}} - E_{\mathrm{x}}^{\mathrm{S}}) \\ &+ a_c(E_{\mathrm{c}}^{\mathrm{LYP}} - E_{\mathrm{c}}^{\mathrm{VWN}}) \end{aligned}$$

A better choice (in my opinion) is the PBE0 functional (sometimes called PBE1PBE) which mixes PBE with 20% HF exchange.

Jacob's Ladder I

John Perdew has summarised the state of DFT using the Biblical picture of Jacob's Ladder: At the base we have the LDA and at the top, in the heaven of chemical accuracy (interesting concept - what is Physical Accuracy?) we have some unknown functional. Here's the whole Ladder...

Jacob's Ladder II

CHEMICAL HEAVEN

1. corr-hyper-GGA: Modifies the hyper-GGAs by including correlation through range-separation. This is currently done at the RPA level.
2. hyper-GGA: Adds exact exchange using range-separation. Leads to functionals that can fix (part of) the charge-transfer problem of most DFT functionals. CamB3LYP
3. meta-GGA: $\rho, \nabla\rho, \nabla^2\rho, \tau$, here $\tau = \frac{1}{2}\sum_{a\in\mathrm{occ}} |\nabla\chi_a|^2$ is the Kohn–Sham orbital kinetic energy density. TPSS
4. GGA: $\rho, \nabla\rho$. PBE
5. LDA,LSD: ρ

INACCURATE HELL