# DFT: Analysis and Failures

Alston J. Misquitta

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

March 20, 2019

# Kohn–Sham DFT: summary I

$$\boxed{E[\rho] = T_{\mathrm{S}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + \int v_{\mathrm{ext}}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} \qquad (1)}$$

must be minimized subject to the orthonormality constraints

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

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

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

# Kohn–Sham DFT: summary II

where the effective potential is defined as

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

# Kohn–Sham DFT: summary III

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: summary IV

Once we have the orbitals $\chi_i$, we can evaluate the energy using

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

or, equivalently,

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

# Kohn–Sham DFT: summary V

Exchange correlation functionals are usually written in the form

$$E_{\rm xc}[\rho] = \int \rho(\mathrm{r})\epsilon_{\rm xc}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots)\mathrm{dr} \tag{2}$$

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

$$\epsilon_{\rm xc}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots) = \epsilon_{\rm x}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots) + \epsilon_{\rm c}(\rho(\mathrm{r}), \nabla\rho(\mathrm{r}), \cdots) \tag{3}$$

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

# Kohn–Sham DFT: summary VI

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.):

$$E_{\mathrm{xc}}[\rho] = \int \rho(\mathrm{r})\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho(\mathrm{r}))\mathrm{dr} \tag{4}$$

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

$$\epsilon_{\mathrm{x}}^{\mathrm{S}}(\rho(\mathrm{r})) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{3/2} \rho^{1/3}(\mathrm{r})$$

Using this we get the Slater exchange functional:

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

# Kohn–Sham DFT: summary VII

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.

# Kohn–Sham DFT: summary VIII

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

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

# Kohn–Sham DFT: summary IX

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

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

Two of the common exchange enhancement factors are

- 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.

# Kohn–Sham DFT: summary X

- 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 physicists use this exchange functional.

# Kohn–Sham DFT: summary XI

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.

# Meaning of the KS orbital energies I

- The Kohn–Sham non-interacting system was initially regarded as no more than a device to facilitate the solution of the Schrödinger equation.
- The orbitals and orbital eigenvalues were not taken to mean anything with one exception:
- $\epsilon_{\rm HOMO} = -I$
  Perdew, Parr, Levy and Balduz (Phys. Rev. Lett. **49**. 1691 (1982)) had shown that the energy of the highest occupied molecular orbital was exactly equal to the negative of the vertical Ionization energy.
- However, there was a lot of empirical evidence that the Kohn–Sham orbital energies were closely related to the experimental ionization energies.
- But they were generally shifted w.r.t. the experimental values.

# Meaning of the KS orbital energies II

- In 2001, Chong, Gritsenko and Baerends (J. Chem. Phys. **116**, 1760) showed that for the exact XC potential (they used a method called SAOP that had many of the properties of the exact XC potential):

  $$I_k \approx -\epsilon_k$$

  With the relation being exact for the HOMO.
- In practice this means that we can use the KS orbital energies as a good approximation to the experimental excitation levels of our system, but with a constant, and possibly large, shift.
- Q: Why are the orbital energies shifted?

Before seeing evidence for the above we will prove that $\epsilon_{\mathrm{HOMO}} = -I$.

# Meaning of the KS orbital energies III

To prove that $\epsilon_{\mathrm{HOMO}} = -I$ we follow the steps:

- In principle, the DFT density is the exact density. So we can use the result we have proved earlier (lecture on Exact Results):

$$\rho(r) \rightarrow e^{-2\sqrt{2E_I}r}$$

Q:

- In Kohn–Sham DFT the density is written as the sum of orbital densities:

$$\rho(\mathrm{r}) = \sum_{i}^{N} |\chi_i(\mathrm{r})|^2 = \sum_{\mathrm{i}}^{\mathrm{N}} \rho_{\mathrm{i}}(\mathrm{r})$$

- Now determine the asymptotic form of the orbital densities $\rho_i(\mathrm{r})$.

continued...

# Meaning of the KS orbital energies IV

To prove that $\epsilon_{\rm HOMO} = -I$ continued....

- The asymptotic form of $\rho_i(\rm r)$ is found using techniques we developed in the lecture on Exact Results. The Kohn–Sham Hamiltonian for orbital $\chi_i$ is

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

Q:

- This is a one-electron Hamiltonian. We will soon show that $v_{\rm S} \to \frac{1}{r}$, so you can write the large-$r$ form of this Hamiltonian exactly as we did in the lecture on Exact Results and hence show that

$$\chi_i(r) \to e^{-\sqrt{-2\epsilon_i}r}$$

Hence $\rho_i(r) \to e^{-2\sqrt{-2\epsilon_i}r}$.

# Meaning of the KS orbital energies V

To prove that $\epsilon_{\rm HOMO} = -I$ continued....

Q:

- Now realise that because $\rho(\rm r)$ is the sum of the $\rho_i$, the asymptotic form of $\rho$ will be determined by the (occupied) orbital with the largest (least negative) energy. This will be the HOMO. Hence we should have, in KS-DFT,

$$\rho(r) \to e^{-2\sqrt{-2\epsilon_{\rm HOMO}}r}$$

- Hence show that $\epsilon_{\rm HOMO} = -I$.

# Meaning of the KS orbital energies VI

Now back to the evidence for the relation:

$$I_k \approx -\epsilon_k$$

Initially the evidence was numerical. Using standard methods for solving the Schrödinger equation (in this case, a technique called multi-reference (i.e., multiple determinant) CCSD, or, MRCCSD), Casida *et al.* showed that density functionals like the LDA (remember, this was the simplest functional we could think of), could, when suitably corrected (more later), produce Kohn–Sham orbital energies that satisfied the above relation.
Have a look at the TDLDA/LB94 results on the next two slides...

# Meaning of the KS orbital energies VII

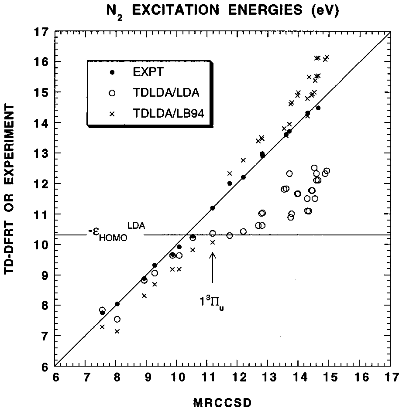


FIG. 1. Correlation plot comparing TD-DFRT results with the multireference coupled cluster singles and doubles (MRCCSD) results of Ref. 53 for the first 35 vertical excitation energies (not counting degeneracies) of $N_2$. Experimental values taken from Ref. 53 are also shown.

Casida, Jamorski, Casida & Salahub, J. Chem. Phys. **108**, 4439 (1998). TD = Time-Dependent. You need to solve the time-dependent Schrödinger equation to get excitation energies in DFT.

IP(expt) = 15.58 eV (S.G. Lias, NIST)

# Meaning of the KS orbital energies VIII

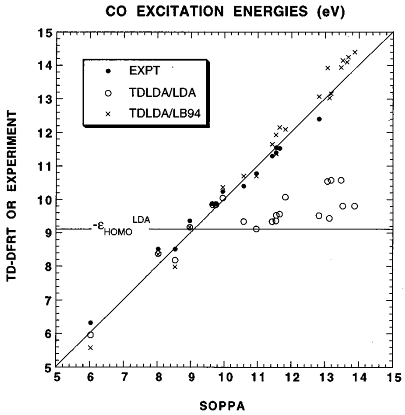


FIG. 2. Correlation plot comparing TD-DFRT results with second-order polarization propagator (SOPPA) ($S\neq 1$ results from Table II of Ref. 54) for the first 23 vertical excitation energies (not counting degeneracies) of CO. Experimental values taken from Ref. 54 are also shown.

Casida, Jamorski, Casida & Salahub, J. Chem. Phys. **108**, 4439 (1998).

IP(expt) = 14.01 eV (S.G. Lias, NIST)

# Meaning of the KS orbital energies IX

What Casida *et al.* found what that

- MRCCSD and a method called SOPPA produced excitation energies in line with experiment.
- Time-dependent LDA was good for the occupied states, but severely underestimated the higher excitation energies.
- However, using the LB94 functional (more on this one) which specifically corrected the long-range problems of functionals like the LDA, they got very good agreement for the higher excitations too.

# Meaning of the KS orbital energies X

Also in 1998, Savin, Umrigar & Gonze published a superb set of results, this time using exact XC potentials. They obtained these exact, or very accurate XC potentials using a method of inversion:

- Calculate a very very accurate density, say using QMC.
- From the first Hohenberg–Kohn theorem there is a one-to-one mapping between this density and the Kohn–Sham potential for a non-interacting system that produces this density.
- Use a convenient method to obtain this potential. Q: How do you do this for the Helium atom density?
- Solve the Kohn–Sham equations using this potential.
- The resulting orbitals and orbital energies are the most accurate you can get.

Here are two sets of tables from their paper in Chem. Phys. Lett. **288**, 391 (1998):

# Meaning of the KS orbital energies XI

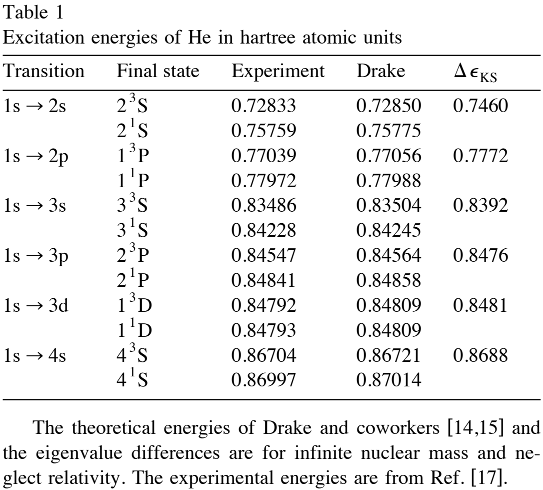


Table 1
Excitation energies of He in hartree atomic units

| Transition | Final state | Experiment | Drake | $\Delta\epsilon_{\mathrm{KS}}$ |
|---|---|---|---|---|
| 1s → 2s | $2\,^3\mathrm{S}$ | 0.72833 | 0.72850 | 0.7460 |
| | $2\,^1\mathrm{S}$ | 0.75759 | 0.75775 | |
| 1s → 2p | $1\,^3\mathrm{P}$ | 0.77039 | 0.77056 | 0.7772 |
| | $1\,^1\mathrm{P}$ | 0.77972 | 0.77988 | |
| 1s → 3s | $3\,^3\mathrm{S}$ | 0.83486 | 0.83504 | 0.8392 |
| | $3\,^1\mathrm{S}$ | 0.84228 | 0.84245 | |
| 1s → 3p | $2\,^3\mathrm{P}$ | 0.84547 | 0.84564 | 0.8476 |
| | $2\,^1\mathrm{P}$ | 0.84841 | 0.84858 | |
| 1s → 3d | $1\,^3\mathrm{D}$ | 0.84792 | 0.84809 | 0.8481 |
| | $1\,^1\mathrm{D}$ | 0.84793 | 0.84809 | |
| 1s → 4s | $4\,^3\mathrm{S}$ | 0.86704 | 0.86721 | 0.8688 |
| | $4\,^1\mathrm{S}$ | 0.86997 | 0.87014 | |

The theoretical energies of Drake and coworkers [14,15] and the eigenvalue differences are for infinite nuclear mass and neglect relativity. The experimental energies are from Ref. [17].

# Meaning of the KS orbital energies XII

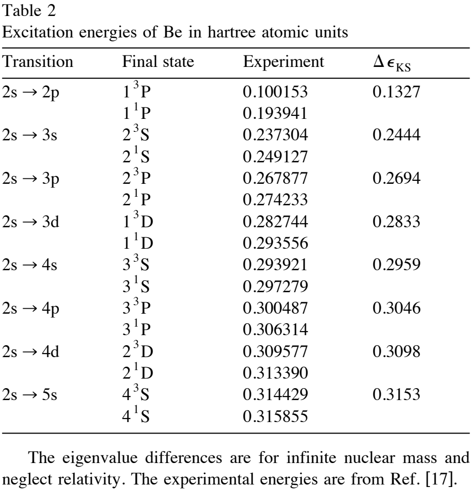


Table 2
Excitation energies of Be in hartree atomic units

| Transition | Final state | Experiment | $\Delta \epsilon_{KS}$ |
|---|---|---|---|
| 2s → 2p | $1\,^3P$ | 0.100153 | 0.1327 |
| | $1\,^1P$ | 0.193941 | |
| 2s → 3s | $2\,^3S$ | 0.237304 | 0.2444 |
| | $2\,^1S$ | 0.249127 | |
| 2s → 3p | $2\,^3P$ | 0.267877 | 0.2694 |
| | $2\,^1P$ | 0.274233 | |
| 2s → 3d | $1\,^3D$ | 0.282744 | 0.2833 |
| | $1\,^1D$ | 0.293556 | |
| 2s → 4s | $3\,^3S$ | 0.293921 | 0.2959 |
| | $3\,^1S$ | 0.297279 | |
| 2s → 4p | $3\,^3P$ | 0.300487 | 0.3046 |
| | $3\,^1P$ | 0.306314 | |
| 2s → 4d | $2\,^3D$ | 0.309577 | 0.3098 |
| | $2\,^1D$ | 0.313390 | |
| 2s → 5s | $4\,^3S$ | 0.314429 | 0.3153 |
| | $4\,^1S$ | 0.315855 | |

The eigenvalue differences are for infinite nuclear mass and neglect relativity. The experimental energies are from Ref. [17].

# Meaning of the KS orbital energies XIII

To summarise:

- In KS-DFT with an *exact* functional, we have $I_k \approx -\epsilon_k$.
- This relation gets better as the excitations involve the higher lying states.
- For the HOMO level we have an exact relation: $\epsilon_{\mathrm{HOMO}} = -I$.
- Contrast these relations with Koopman's theorem from Hartree–Fock theory.
- However, for approximate functionals none of these results hold. Instead the HOMO level is generally shifted closer to the LUMO (the gap closes), and the excitation energies are therefore underestimated.
- The LB94 functional appears to fix the problem (but...see later).

# Self-Interaction I

Q: What are the problems with using an approximate XC functional?
Q: What is the origin of the constant shift of energies mentioned above?
To understand this we will work out how $v_{\rm xc}$ should behave for the hydrogen atom.

- What is the form of the exact Kohn–Sham potential $v_{\rm S}$ for large $r$?

$$\left(-\frac{1}{2}\nabla^2 + v_{\rm S}(\mathrm{r})\right)\chi_k(\mathrm{r}) = \epsilon_k\chi_k(\mathrm{r})$$

This is equivalent to asking what the potential felt by an electron will be as we pull it off the atom/molecule. It will see

# Self-Interaction II

a hole and hence experience a $-1/r$ potential. Therefore we must have as $r \to \infty$

$$v_{\mathrm{S}} \to -\frac{1}{r}.$$

- We know that $v_{\mathrm{S}} = v_{\mathrm{J}} + v_{\mathrm{ext}} + v_{\mathrm{xc}}$. We also know the long-range (asymptotic) forms of $v_{\mathrm{J}}$ and $v_{\mathrm{ext}}$:

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

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

Here $N$ is the number of electrons and $Z$ is the nuclear charge. For a neutral system these are equal. Therefore these two cancel out asymptotically.

# Self-Interaction III

- Hence we must have

$$v_{\mathrm{xc}}(\mathrm{r}) \rightarrow -\frac{1}{\mathrm{r}}$$

- How do common XC potentials behave at asymptotically? Best to use the simplest XC functional: the Slater exchange functional (the VWN correlation part does not change the picture very much). The Slater functional is

$$E_{\mathrm{x}}^{\mathrm{S}}[\rho] = -C_{\mathrm{x}} \int \rho^{4/3}(\mathrm{r})\mathrm{d}\mathrm{r}$$

# Self-Interaction IV

Using $\rho \to e^{-\alpha r}$, this gives us an XC potential:

$$\begin{aligned} v_{\mathrm{xc}}(\mathrm{r}) &= \frac{\delta E^{\mathrm{S}}_{\mathrm{xc}}[\rho]}{\delta\rho} \\ &= -\frac{4}{3} C_{\mathrm{x}} \rho^{1/3}(\mathrm{r}) \\ &\to -e^{-\frac{\alpha}{3} r} \end{aligned}$$

It has the wrong asymptotic form. It decays too quickly with distance.

- This is what leads to a small band-gap in DFT: the unoccupied levels are all shifted down with respect to the occupied orbitals.

# Self-Interaction V

- Self-Interaction: Another way of looking at this problem is to realise that the too rapid decay of $v_{\rm xc}$ with distance is equivalent to the electron 'seeing' itself.

$$v_{\rm xc}(\mathrm{r}) \to -e^{-\frac{\alpha}{3}r} \\ \to -\frac{\left(+re^{-\frac{\alpha}{3}r}\right)}{r}$$

  I.e., rather than see a hole with charge $+1$, it sees a hole with charge $+re^{-\frac{\alpha}{3}r}$ which goes exponentially fast to zero.

- Thus for moderate separation the electron will see little or no attraction to the ion, and will therefore be very weakly bound, or even unbound.

# Self-Interaction VI

- This is the **self-interaction problem**, and is termed thus as it is as if the electron interacts with itself as the hole ‘fills up’ with increasing $r$.
- Any molecular property that depends on the unoccupied levels will there be effected. Examples are: polarizabilities, hyperpolarizabilities, excitations, in particular charge-transfer excitations, NMR shifts.

# Self-Interaction VII

In 1994, van Leeuwen and Baerends (Phys. Rev. A **49**) constructed an XC functional that resulted in an XC potential that decayed as $-1/r$. This functional, termed `LB94` takes the form:

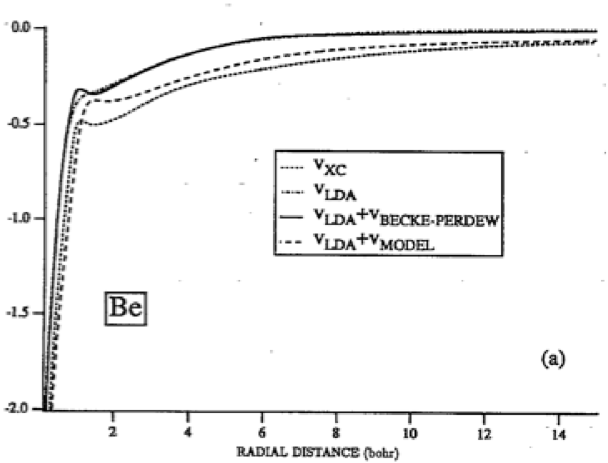
$$v_{\mathrm{xc}}^{\mathrm{LB94}}(\mathrm{r}) = -\beta \mathrm{n}^{1/3}(\mathrm{r}) \frac{\mathrm{x}^2}{1 + 3\beta \mathrm{x} \sinh^{-1}(\mathrm{x})},$$

where $\beta = 0.05$, $n(\mathrm{r})$ is the electronic density, and the reduced-gradient is defined as $x = |\nabla n|/n^{4/3}$.

Q: Using the asymptotic form of the electronic density: $r \to \infty$, $n \to e^{-\alpha r}$, where $\alpha$ is a constant. Show that $v_{\mathrm{xc}}^{\mathrm{LB94}} \to -1/r$. You will need to use an appropriate series expansion for $\sinh^{-1}(x)$.

# Self-Interaction VIII

Here is how the LB94 functional behaves for the beryllium atom:



# Self-Interaction IX

- Here, $v_{\rm xc}$ is the reference XC potential obtained from a very accurate density. $v_{\rm LDA}$ and $v_{\rm LDA} + v_{\rm Becke-Perdew}$ are two variants of the LDA that include the Slater form of the exchange. The latter two clearly decay too quickly with separation in comparison with the reference $v_{\rm xc}$.
- Finally we have $v_{\rm LDA} + v_{\rm model}$ which is XC potential from the LB94 functional. This one agrees with $v_{\rm xc}$ quite well indeed, but *only at long range!*.
- Notice what happens at short range: The LDA model potentials are both far off from the exact potential, but they are *uniformly shifted* from $v_{\rm xc}$. A uniform shift in a potential is never a problem. However the short-range behaviour of LB94 is, by comparison, rather poor.

# Self-Interaction X

- As a consequence, all properties that depend on the short-range part of the XC potential (the density, and all properties that depend on the density) come out poorly with LB94.
- But properties that are more sensitive to the long-range part (higher excitation energies) come out better. We have already seen this above.

It was Casida and Salahub (J. Chem. Phys., **113** (2000)) who noticed this problem with LB94. They argued that rather than decay as $-1/r$, all local and semi-local functionals should decay as $-1/r + \Delta$, where $\Delta$ is a constant shift.
The then demonstrated that rather than use the LDA or LB94 on their own, we should instead splice LB94 to the LDA so as to

# Self-Interaction XI

achieve a much better form of the overall XC potential. But this splicing should include the constant off-set $\Delta$:

$$v_{\rm xc}^{\rm AC-LDA}(\rm r) = \max[v_{\rm xc}^{\rm LDA}(\rm r) - \Delta, \rm v_{xc}^{LB94}(r)].$$

where the shift is given by $\Delta = \epsilon_{\rm HOMO}^{\rm LDA} + I$.

This is the `CS00` asymptotic correction used in NWCHEM. It is one of a handful of possible corrections, but NWCHEM includes just this one.

# Self-Interaction XII

*What is the origin of the shift?*
It can be shown that the exact asymptotic form of the XC potential is

$$v_{\mathrm{xc}} \rightarrow -\frac{1}{r} + (\epsilon_{\mathrm{HOMO}} + I).$$

This is a generalisation of the result we derived earlier. The proof is not essential.

- For an exact functional, we have shown that $\epsilon_{\mathrm{HOMO}} = -I$. So the term in the brackets vanish.
- But for local and semi-local functionals it does not.
- This has to do with what is called the *derivative discontinuity*.

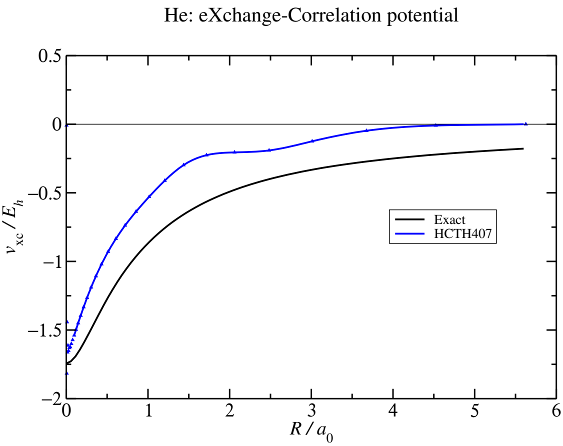
# Self-Interaction XIII

- Put simply, for the exact XC functional, the energy of the HOMO orbital changes discontinuously as the electron number crosses an integer value. Think about it like this: For $N$ electrons our HOMO orbital energy will be $\epsilon_{\mathrm{HOMO}} = \epsilon_N = -I(N)$, but for $N + \delta$ electrons it will be $\epsilon_{\mathrm{HOMO}} = \epsilon_{N+1} = -I(N+1) = A(N)$, where $I(N)$ is the vertical ionisation energy for an $N$-electron system and $A(N)$ is the electron affinity for the $N$-electron system. So the HOMO energy is discontinuous around $N$.
- However, local and semi-local functionals cannot describe this discontinuity, instead they interpolate between the two values.
- This means that we no longer have $\epsilon_{\mathrm{HOMO}} = -I$, and so approximate XC potentials will be offset by the amount $\epsilon^{\mathrm{approx}}_{\mathrm{HOMO}} + I$. This is the origin of the $\Delta$.

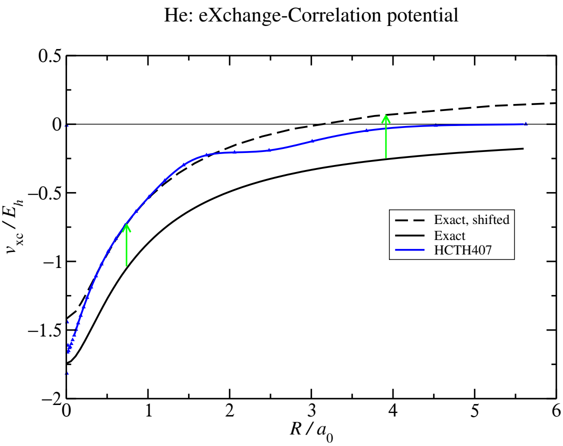
# Self-Interaction XIV

It is best to see this pictorially. In the next few images we will look at the XC potential for Helium calculated using the HCTH407 functional compared with a (nearly) exact XC potential (this was obtained by calculating a very accurate He density and *inverting* it to obtain the potential).

# Self-Interaction XV



# Self-Interaction XVI



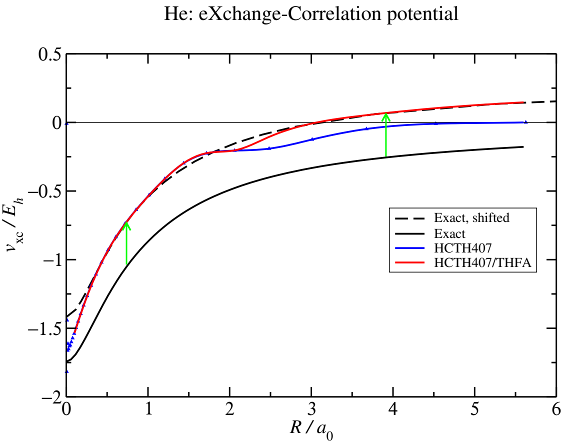
# Asymptotic-correction I

Since we know what the asymptotic form of $v_{\mathrm{xc}}$ should be we can enforce it through an empirical fix known as the asymptotic correction. We need to account for the shift. Tozer and Handy & Casida worked all this out in 1998:

$$v_{\mathrm{xc}}(\mathrm{r}) \rightarrow -\frac{1}{\mathrm{r}} + \mathrm{I} + \epsilon_{\mathrm{HOMO}}$$

So if know (or calculate) $I$, calculate $\epsilon_{\mathrm{HOMO}}$ from a standard DFT calculation, then we will be able to work out the shift and apply this correction. This is known as the asymptotic correction.

# Asymptotic-correction II



# Asymptotic-correction III

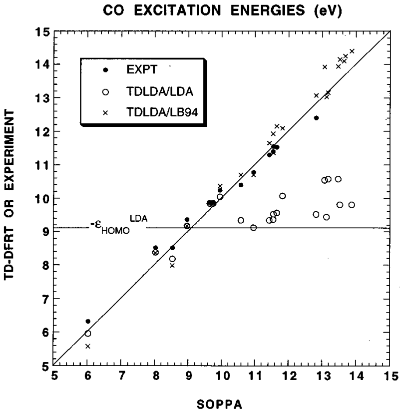


FIG. 2. Correlation plot comparing TD-DFRT results with second-order polarization propagator (SOPPA) ($S \neq 1$ results from Table II of Ref. 54) for the first 23 vertical excitation energies (not counting degeneracies) of CO. Experimental values taken from Ref. 54 are also shown.

The LB94 functional is one route to imposing an asymptotic correction. The effect of this on the excitation energies is quite dramatic. Casida, Jamorski, Casida & Salahub, J. Chem. Phys. **108**, 4439 (1998).

# Asymptotic-correction IV

## Excitation energies

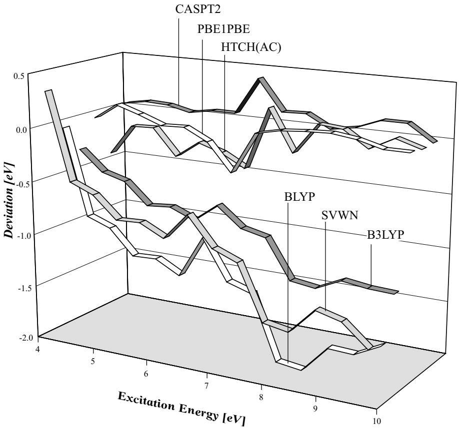
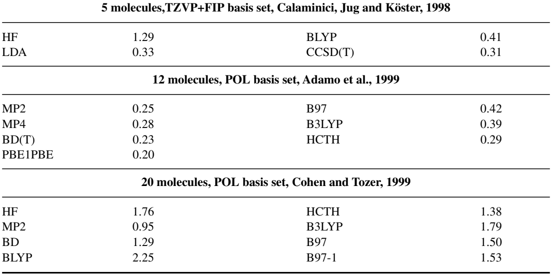


**Figure 9-2.** Performance of various functionals in the framework of time-dependent DFT for excitation energies of ethylene.

# Asymptotic-correction V

## Polarizabilities



| | | | |
|---|---|---|---|
| **5 molecules,TZVP+FIP basis set, Calaminici, Jug and Köster, 1998** | | | |
| HF | 1.29 | BLYP | 0.41 |
| LDA | 0.33 | CCSD(T) | 0.31 |
| **12 molecules, POL basis set, Adamo et al., 1999** | | | |
| MP2 | 0.25 | B97 | 0.42 |
| MP4 | 0.28 | B3LYP | 0.39 |
| BD(T) | 0.23 | HCTH | 0.29 |
| PBE1PBE | 0.20 | | |
| **20 molecules, POL basis set, Cohen and Tozer, 1999** | | | |
| HF | 1.76 | HCTH | 1.38 |
| MP2 | 0.95 | B3LYP | 1.79 |
| BD | 1.29 | B97 | 1.50 |
| BLYP | 2.25 | B97-1 | 1.53 |

# Asymptotic-correction VI

- The asymptotic correction does fix what is called the one-electron self-interaction error.
- But there is no clear way to apply an asymptotic correction in the bulk phase. And the self-interaction error manifests itself there too.
- We know that Hartree–Fock is free of self-interaction, so one solution to the problem is to include more and more Hartree–Fock-type exchange in KS-DFT. But this leads to an overall loss in accuracy.

# Asymptotic-correction VII

- A better solution is to use *range-separation*: Split the e-e interaction operator into a short- and long-range part:

  $$\frac{1}{r_{12}} = \frac{\mathrm{erfc}(\beta r_{12})}{r_{12}} + \frac{(1 - \mathrm{erfc}(\beta r_{12})}{r_{12}}$$

  The complementary error function is chosen as it allows easy integral evaluation. Now use DFT on the short-range part and Hartree–Fock-exchange on the long-range part. In this way you get the best of both worlds.
- The DFT usually takes care of all correlation, and only the local part of the exchange which it is known to get right. HF (or something better) then takes care of the long-range exchange.
- Functionals that use this technique are termed range-separated or Long-range Corrected (LC) functionals.

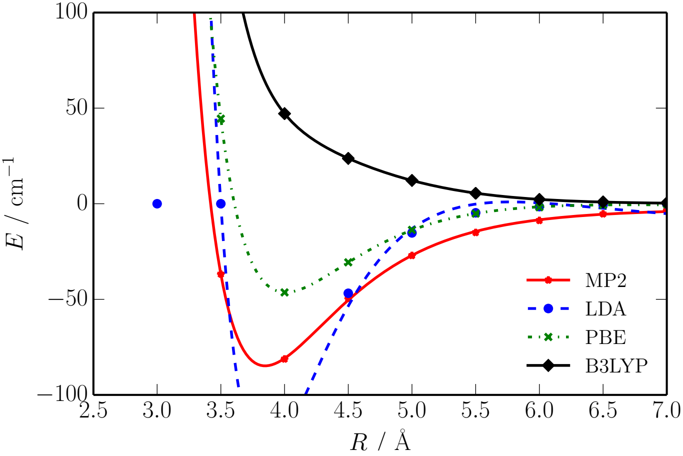
# Asymptotic-correction VIII

- Functionals such as `CamB3LYP`, `LC-PBE`, `LC-PBE0` use this principle.
- There is one free parameter in this model: the extent of the range-separation controlled by $\beta$. Several authors have worked on techniques to determine $\beta$ self-consistently. But issues remain: in a strongly anisotropic system, $\beta$ should probably vary with position, or in direction. Issues like this remain unsolved.
- It is also possible to use post-Hartree–Fock methods on the long-range part. For example, you could use MP2 or the RPA (random phase approximation). This would allow the dispersion interaction to be described by DFT. More on this next.

# Dispersion I

The other problem with DFT is that almost all conventional functionals fail to describe the dispersion (van der Waals) interaction. This is a long-range and non-local interaction that arises from the correlation of quantum mechanical fluctuations on the interaction species.
Consider the argon dimer: this is a dispersion-bound system, that is, the attraction between two argon atoms arises purely from the dispersion interaction. This is typical of the rare-gas atoms. On the next slide we see interaction energies calculated for this system with MP2, LDA, PBE and B3LYP using the `aug-cc-pVTZ` basis set using the counterpoise correction.

# Dispersion II



# Dispersion III

- The reference MP2 energies exhibit the classic interaction energy curve as expected. Recall that MP2 is not perfect for this system, but it will serve as a reference here.
- The density functionals are all over the place. LDA and PBE show some binding but is it from the dispersion?
- The clue is in the long-range behaviour: all density functionals decay to zero much too quickly with $R$.
- On the other hand, the dispersion energy (in the MP2 tail) is more slowly decaying as $R^{-6}$.
- B3LYP is completely repulsive!

# Perturbation Theory I

The dispersion energy first arises at second-order in intermolecular perturbation theory.

Consider a pair of spherical atoms $A$ and $B$ placed along the $z$-axis and separated by a distance $R$. The Hamiltonian for this system may be written as

$$H = H^{(0)} + H^{(1)},$$

where $H^{(0)} = H_A^{(0)} + H_A^{(0)}$ is the sum of the unperturbed Hamiltonians of $A$ and $B$, and, the intermolecular interaction operator takes the leading-order multipole expanded form:

$$H^{(1)} = \frac{1}{R^3}(\hat{x}_A\hat{x}_B + \hat{y}_A\hat{y}_B - 2\hat{z}_A\hat{z}_B),$$

# Perturbation Theory II

where, $\hat{x}_A$ is the position operator (dipole moment operator) along the $x$-axis and centred at $A$.

If $H_A^{(0)}|m\rangle = E_{A,m}^{(0)}|m\rangle$ and $H_B^{(0)}|n\rangle = E_{B,n}^{(0)}|n\rangle$, then a convenient basis for the dimer are the states $|mn\rangle$.

The zeroth order wavefunction for this system is then $|00\rangle$ and the zeroth order energy is $E^{(0)} = E_{A,0}^{(0)} + E_{B,0}^{(0)}$.

Q: Show that the first-order Raleigh–Schrödinger perturbation theory (RSPT) energy correction for this system is zero. That is,

$$E^{(1)} = \langle 00|H^{(1)}|00\rangle = 0.$$

# Perturbation Theory III

The second-order dispersion energy for the two interacting systems takes the standard Raleigh–Schrödinger form

$$E^{(2)} \equiv E^{(2)}_{\mathrm{disp}} = -\sum_{m\neq 0, n\neq 0} \frac{|\langle 00|H^{(1)}|mn\rangle|^2}{E^{(0)}_{A,m} + E^{(0)}_{B,n} - E^{(0)}_{A,0} - E^{(0)}_{B,0}}.$$

This term can be evaluated using ideas borrowed from linear-response time-dependent DFT (To find out more about LR-TDDFT see the book by Carsten), but here we will use the *average energy approximation* to simplify this expression.
In the average energy approximation each energy difference in the denominator is approximated as a constant $E^{(0)}_m - E^{(0)}_0 \approx \Delta$ which will represent an average excitation energy, that will typically be the atomic ionisation energy.

# Perturbation Theory IV

Q: Show that in this approximation the dispersion energy *for two identical atoms* may be written as

$$E^{(2)}_{\text{disp}} \approx -\frac{C_6}{R^6},$$

where

$$C_6 = \frac{3}{\Delta}\langle 0|\hat{z}^2|0\rangle.$$

Hints:

- Use the resolution of the identity to result in an expression that involves terms such as $\langle 0|xy|0\rangle$.
- Use the fact that the state $|0\rangle$ is invariant under reflection in, say, the *yz*-plane to show that such terms are zero.

# Perturbation Theory V

Using the expression for the dynamic polarizability:

$$\alpha_{ij}(\omega) = 2 \sum_{n \neq 0} \frac{\omega_{n0} \langle 0 | r_i | n \rangle \langle n | r_j | 0 \rangle}{\omega_{n0}^2 - \omega^2}$$

it can be shown (try it!) that in the average energy approximation

$$C_6 = \frac{3}{4} \Delta \alpha(0)^2,$$

where $\alpha(0)$ is the static polarizability.

# Perturbation Theory VI

This is the origin of the well-known $-\frac{C_6}{R^6}$ form for the van der Waals, or dispersion interaction. A few comments about this expression:

- It is only the first term in an infinite series. More generally we will have an expansion like $-\frac{C_6}{R^6} - \frac{C_7}{R^7} - \frac{C_8}{R^8} \cdots$.
- The $C_n$ coefficients will generally be orientationally dependent.
- For a spherically symmetric system, you can show (using symmetry arguments) that both the angular dependence and the odd-$n$ terms vanish.
- The expansion diverges as $R \to 0$ so it must be damped using functions that cancel out the offending powers of $1/R$.

# Perturbation Theory VII

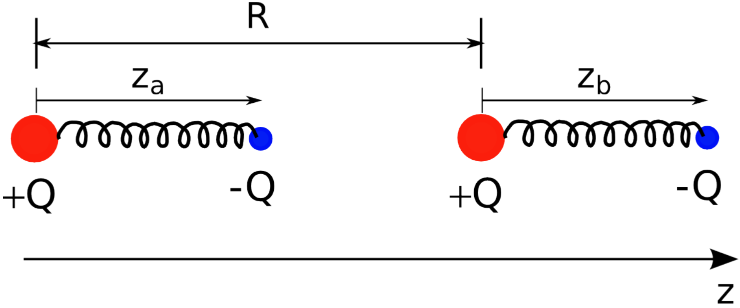
- For a molecular system the expansion is usually generalised to include a double sum over the atoms *a* in molecule A and atoms *b* in molecule B:

$$E^{(2)}_{\rm disp}[AB] = -\sum_{a\in A, b\in B}\sum_{n=6}^{\infty}\frac{C_n^{ab}}{r_{ab}^n}.$$

- For low-dimensional systems with small HOMO–LUMO (band) gaps, this expression is qualitatively wrong as it implicitly assumes that all electron fluctuations (see next topic on Drude oscillators) are *local*. This is not the case in such materials and we get a substantial contribution from the long-range plasmon-like fluctuations. This leads to the presence of terms in the expression that behave like $1/R^2$. (See the Casimir force and papers by Misquitta *et al.* and Tkatchenko *et al.* that have addressed this unusual case.)

# Drude Model I

An alternative picture for the dispersion energy arises from coupled quantum Drude oscillators as follows.
The dispersion energy cannot be described in terms of classical interactions as the electrostatic and induction terms can. A semi-classical picture is required.



# Drude Model II

Model each molecule with a fixed charge $+Q$ at the centre and an oscillating charge $-Q$. In the usual scaled units (i.e. energy in units of $h\nu = \hbar\omega$, length in units of $(\hbar^2/km)^{1/4}$) the Hamiltonian is (assuming infinite separation):

$$\mathcal{H} = -\frac{1}{2}\frac{\partial^2}{\partial {z_A}^2} + \frac{1}{2}{z_A}^2 - \frac{1}{2}\frac{\partial^2}{\partial {z_B}^2} + \frac{1}{2}{z_B}^2. \tag{5}$$

The energy is the sum of the individual energies, i.e., $E_{v_A v_B} = v_A + v_B + 1$. The ground state energy ($v_A = v_B = 0$) is 1 unit, i.e. $h\nu$.

# Drude Model III

If the instantaneous displacements are $z_A(t)$ and $z_B(t)$, the dipole moments on A and B are $\mu_A = -Qz_A(t)$ and $\mu_B = -Qz_B(t)$, respectively.
At a finite separation $R$, these dipoles interact. The general form of the dipole–dipole interaction operator is (Q: How does it relate to the earlier form used in the Perturbation Theory section above?)

$$H_{\mu\mu} = -\frac{\hat{\mu}_A\hat{\mu}_B}{R^3}\left(2\cos\theta_A\cos\theta_B - \sin\theta_A\sin\theta_B\cos\phi\right),$$

Here, $\theta_A = \theta_B = \pi$ and $\phi = 0$ so the Hamiltonian at finite separations has the additional term $cz_Az_B$ where $c = -\frac{2Q^2}{R^3}$.

# Drude Model IV

Using the new variables $Z_1 = \sqrt{\frac{1}{2}}(z_A + z_B)$ and $Z_2 = \sqrt{\frac{1}{2}}(z_A - z_B)$, the potential term in the Hamiltonian becomes

$$V = \frac{1}{2}{z_A}^2 + c z_A z_B + \frac{1}{2}{z_B}^2 = \frac{1}{2}(1+c){Z_1}^2 + \frac{1}{2}(1-c){Z_2}^2,$$

while the kinetic energy is unchanged in form:

$$T = -\frac{1}{2}\frac{\partial^2}{\partial {z_A}^2} - \frac{1}{2}\frac{\partial^2}{\partial {z_B}^2} = -\frac{1}{2}\frac{\partial^2}{\partial {Z_1}^2} - \frac{1}{2}\frac{\partial^2}{\partial {Z_2}^2}.$$

# Drude Model V

So we now have an oscillator $Z_1$ with frequency $\sqrt{1+c}$, and another, $Z_2$, with frequency $\sqrt{1-c}$. The allowed energies (in the original scaled units) are now $(v_1+\frac{1}{2})\sqrt{1+c}+(v_2+\frac{1}{2})\sqrt{1-c}$. In a classical system the coupling doesn't change the minimum energy, which occurs when both oscillators are at rest. That is, $z_A = z_B = 0$, so $Z_1 = Z_2 = 0$ also, and the total energy is zero. A quantum system, however, has zero-point energy: 1 unit in the original uncoupled system.

# Drude Model VI

When $v_1 = v_2 = 0$ the energy of the interacting system is

$$\begin{aligned} E = \frac{1}{2}(\sqrt{1+c} + \sqrt{1-c}) &= \frac{1}{2}[(1 + \frac{1}{2}c - \tfrac{1}{8}c^2 + \cdots) \\ &\quad + (1 - \frac{1}{2}c - \tfrac{1}{8}c^2 - \cdots)] \\ &= 1 - \tfrac{1}{8}c^2 - \cdots . \end{aligned}$$

That is, the zero-point energy is smaller for the correlated oscillators than for the uncoupled ones, whether $c$ is positive or negative. The stabilization energy is the Drude approximation to the dispersion.

# Drude Model VII

The Drude expression for the dispersion energy is $-\frac{1}{8}c^2 = -\frac{Q^4}{2R^6}$ which varies as $\frac{1}{R^6}$. The coefficient of this term is usually labeled $C_6$ and, inserting the energy factors scaled out, is defined as

$$C_6 = \frac{\hbar\omega Q^4}{2(4\pi\epsilon_0)^2 k^2}.$$

We now need to relate $Q$ and $k$ to measurable quantities. This is done using classical ideas.

# Drude Model VIII

If the charge $-Q$ extends by $z$ in an electric field $E$ then balancing forces we must have $kz = -QE$, or $z = -\frac{QE}{k}$. Now, by definition, $\mu = -zQ = \frac{Q^2}{k}E$. But, by definition of the polarizability, $\mu = \alpha E$, therefore

$$\alpha = \frac{Q^2}{k}.$$

This allows us to re-write the $C_6$ as

$$\boxed{C_6 = \frac{\hbar\omega\alpha^2}{2(4\pi\epsilon_0)^2}.}$$

# Drude Model IX

In 3-dimensions this expression becomes

$$C_6 = \frac{3\hbar\omega\alpha^2}{4(4\pi\epsilon_0)^2},$$

Q: Show this!

and taking, as London did, $\hbar\omega = E_I$, the ionization energy, we get

$$\boxed{C_6 = \frac{3E_I\alpha^2}{4(4\pi\epsilon_0)^2}.}$$

This is an approximation, but it contains all the correct physics. The dispersion energy is always attractive (at second-order) and can be interpreted as arising from a correlation in the electronic fluctuations on the molecules.

# Drude Model X

- In this phenomenological model, the dispersion energy arises from the correlations of quantum fluctuations. It is a purely quantum phenomenon and has no classical analogue.
- Further, it as it is a non-local phenomenon, we can now understand why local and semi-local density functional are unable to describe this energy. Functionals need to be explicitly non-local to be able to describe the dispersion energy.
- However many density functionals can be corrected to account for the missing dispersion by adding to the DFT energy a term like:

$$E_{\mathrm{disp}} = -\sum_{b>a} f_{\mathrm{switch}}(\beta r_{ab}) \frac{C_6^{ab}}{r_{ab}^6},$$

# Drude Model XI

where $a$ and $b$ are atomic sites separated by $r_{ab}$ and with a dispersion coefficient $C_6^{ab}$. The switching function $f_{\mathrm{switch}}(\beta r_{ab})$ which typically depends on one or more parameters (here only one is indicated) has to be very carefully chosen to avoid double-counting the dispersion at short range. Also, this switching function needs to be tuned to each density functional.

- This kind of correction was first introduced in 2001 by Wu *et al.* from the Scoles group, and was generalized by Grimme in 2004. Grimme has subsequently improved this correction in methods termed ‘D2’ (2006), ‘D3’ (2010), and a ‘D4’ correction is due to come out soon.

# Drude Model XII

- In many of these models, it is assumed that the dispersion coefficient between atoms is fixed and does not vary with changes in chemical environment. This is often a poor approximation.
- Furthermore, the $C_6$ term is only one term in the dispersion expansion and it is generally angular-dependent.
- Additionally, for semiconductors and metals additional terms (as low as $C_2$) arise from plasmon modes. These are long wavelength fluctuations arising from the conduction electrons. For such systems the above model breaks down.
- There are explicitly non-local functionals which do not need this correction, but in practice, this correction, if well tuned, can be more accurate than many other more sophisticated non-local functionals.

# Drude Model XIII

- There are at present many empirical dispersion correction models and some that are less empirical. One of the latter is the so-called ‘many-body dispersion’ , or MBD, method of Tkatchenko, diStasio and others (2012). Here the dispersion energy is computed through a coupled dipole oscillator model in much the same way as we coupled the two Drude oscillators. This method uses a physical model for the terms and can account for much of the anisotropy in the dispersion. The authors call this a ‘many-body’ method as the atomic polarizabilities of the atoms (which are the bodies) are coupled together to result in a many-atom, or many-body dispersion energy.

# Drude Model XIV

- One problem all empirical dispersion models face is the choice of the switching function. No matter how accurate the long-range dispersion energy may be, if the switching is not done correctly, errors can and will be large. It is quite possible that the main differences in the various models is not so much in the choice of dispersion parameters, but is in the methods used to switch from the dispersion model to the density-functional.
- Finally, just to emphasise an important and often ignored point: none of the dispersion models used in DFT actually compute the dispersion energy as defined through a perturbation theory like SAPT (symmetry-adapted perturbation theory — one of the most accurate methods for intermolecular interactions). Instead what is computed is the leading and perhaps next-leading order contribution only. This

# Drude Model XV

is often sufficient for DFT as the higher order terms come into play mainly at short-range, and this is where the switching function together with the density-functional are meant to account (i.e., fudge) the missing terms. Therefore it is wrong to term these energies as *the dispersion energy*, nevertheless it is often, erroneously, done in the literature.

# DFT: Best usage I

## How do we best use DFT?

- Use a basis set appropriate to the problem!
- Try a range-separated functional, and determine the range-separation needed by enforcing the DFT version of Koopman' theorem: $\epsilon_{\mathrm{HOMO}} = -I$
- The DFT+U scheme (which we have not discussed) may be appropriate when you have transition metals in your system. Here, an empirical parameter (the 'U') is used to fix the self-interaction problem.
- Use a dispersion correction. This should be the default as any correction may prove better than none.

# DFT: Best usage II

**How do we best use DFT?**

- If you are after subtle correlation effects use range-separation with the RPA (random-phase approximation) for the long-range. This account for both exchange and correlation (approximately) at long-range.
- For weak interactions consider the dispersionless density-functional (dlDF) of Pernal *et al.* (2009). To this you need to add an accurate dispersion model. This method has been shown to result in reliably accurate interaction energies for weakly bound complexes.
- For large systems where even DFT becomes too computationally expensive consider the newer of the tight-binding DFT (TB-DFT) methods. These are of course more approximate.