Third Part

Slide 1

In the final part of this introduction, I will discuss some topics that were left out from the previous two parts, and that you have seen or you are going to see in the CP2K tutorials.

Slide 2

In particular,

- I will give more details about the exchange correlation functionals with the focus on the ones available on CP2K,
- I will discuss the Grimme dispersion corrections to cure a deficiency of such functionals,
- then, I will introduce the concept of pseudopotential,
- and finally, I want to add few words about the basis sets.

Slide 3

We mentioned in a previous lecture that the electronic problem can be solved within the framework of the Density Functional Theory by using the scheme of Kohn and Sham, which leads to solve the Kohn-Sham (KS) equations <u>here</u>. The functional form of all the terms in the KS equations are known, apart from the so-called exchange-correlation energy functional E_xc, which depends on the electron density rho.

The name of this term recalls that it should contain both the exchange interaction, related to the Pauli repulsion when we are dealing with electrons, and the Coulomb correlation, that is a measure of how much the motion of one electron is affected by the presence of all other electrons.

DFT is an exact theory in principle, but to be used in practice, that is to be able to write and solve the KS equations, it requires a guess, an approximation, of the exchange-correlation functional. During the years scientists have built and tested several functionals in order to provide educated guesses for this unknown term.

Slide 4

The simplest approximation, and therefore also the first one to be devised, is the so-called Local Density Approximation (or LDA). This assumes that the exchange-correlation energy at a point r in the space, is simply equal to the exchange-correlation energy of a uniform electron gas that has the same density at that point.

The analytical form of the exchange term is easy to retrieve in this peculiar case, while the most common parametrisations for the correlation part are obtained by interpolating the accurate values coming from quantum Monte Carlo simulations on the homogeneous electron gas at various densities.

The Local Density Approximation ignores corrections to the exchange-correlation energy due to inhomogeneities in the electron density around a point r. Considering this inexact nature of this approximation, it may at first seem somewhat surprising that it is so successful as it is. This can be partially attributed to the fact that LDA gives the correct sum rule to the so-called exchange-correlation hole. That is, there is a total electronic charge of one electron excluded from the neighbourhood of the electron at the point r.

However, the Local Density Approximation is known to overbind, particularly in molecules.

For this reason, in chemistry more sophisticated approximations are commonly employed, as for example the Generalised Gradient Approximation (or GGA). This approximation attempts to incorporate the effects of inhomogeneities by including the gradient of the electron density; as such one refers to this kind of approaches as a semi-local method.

Of course, there is no unique form for the GGA, and indeed many variants have been proposed in the years. Most of them are available on CP2K. An example is the very popular Becke and Lee, Yang, Parr approximation (in short BLYP) or the Perdew-Burke-Ernzerhof (in short PBE) that you use in the tutorials and...

Slide 5

...whose analytic expressions I have sketched in this scary slide for your reference. Of course, I have not time to describe them in detail here. Just summarizing, BLYP finds some parameters by fitting experimental data while PBE was built mainly via analytical and theoretical arguments.

The Generalised Gradient Approximation significantly succeeds in reducing the effects of LDA overbinding but it is still problematic in some contexts as for example the estimation of static properties, such as the atomization and dissociation energies, the bond lengths and the vibrational frequencies but also the estimation of dynamical properties like the diffusion coefficients in liquids like in water, due, ultimately, to a poor description of the covalent O-H bond stretching of the water molecule.

Historically, one of the first way to go beyond the Generalised Gradient Approximation brought to the development of the so-called hybrid functionals.

Slide 6

These are a class of approximations to the exchange–correlation energy functional that incorporates a portion of exact exchange from Hartree–Fock theory with the rest of the exchange–correlation energy from other sources, both ab initio and empirical, and in particular from the Generalised Gradient Approximation.

Here I have reported two popular examples of hybrid functionals of this kind, that is the B3LYP functional, derived combining the BLYP correlation with the exact exchange in the way shown here, and the PBE0 functional, whose correlation part comes from the PBE functional.

Hybrid functionals significantly improve the accuracy of all the molecular properties mentioned before.

A drawback of hybrid functionals is that the exact exchange term is computational expensive to calculate within the framework of plane wave basis set. Moreover, even using hybrid functionals, there are still difficulties in using DFT to properly describe intermolecular interactions, which are of critical importance to understanding for example chemical reactions, especially van der Waals forces, and in particular dispersion forces. The incomplete treatment of this kind of forces can adversely affect the accuracy of DFT in the treatment of systems which are dominated by dispersion (e.g. interacting noble gas atoms) or where dispersion competes significantly with other effects: this is the case of the biomolecules or large systems in general.

How can we overcome this deficiency? Let's go a bit more in depth of this problem.

Slide 7

The London dispersion forces are a type of forces acting between atoms and molecules that are normally electrically neutral and symmetric, that is, the electrons are symmetrically distributed with respect for example to the nucleus in the atoms, and therefore there are no net charges but also no permanent dipoles.

In fact, the London dispersion can be considered a long-ranged electron correlation effect, whose energy term can be approximately described to asymptotically scale as 1 over R to the power of 6, where R is the separation distance between the two interacting neutral objects. This should bring to mind one of the two terms in the Lennard-Jones potential, the potential that you find in any molecular dynamics force field to phenomenologically describe the van der Waals interactions.

Now, this London part of the correlation is not included in standard KS-DFT. Why? The technical reason can be traced back in the absence in DFT of a correct description of the quantum fluctuations, that is the excitations to virtual orbitals. The quantum fluctuations become important when and where the electron density is almost zero. This is the case for large separation distances in between

the neutral objects, where therefore the density rho does not contain significant dispersion signature.

Slide 8

Various approaches are currently in use and under development to accurately model the London dispersion interactions within the DFT. In this slide I have collected the names of the most important ones. However, here I would like to give some details about one of the most popular, the DFT-D method, also known as the Grimme dispersion corrections, because this is the approach that you use in the tutorials.

Slide 9

Really, with the term Grimme dispersion corrections one can refer to at least three different models developed in the years and with increasing complexity. The most recent ones, called DFT-D2 and DFT-D3 are implemented in CP2K.
In general, in the DFT-D schemes the total energy is calculated as the sum of the usual self-consistent KS energy as obtained from the chosen density functional and a dispersion correction, which is in turn a sum of two- and three-body energies.
In the lower part of the slide, I reported the details for the DFT-D2 scheme, which contains only two-body terms and this is the one employed in the tutorials.
In this dispersion correction scheme, the sum is over all the pairs of atoms; R_ij is the interatomic distance of atoms i and j; s_6 is a global scaling parameter depending on the choice of the employed functional the C6 values here are calculated from the empirical atomic dispersion coefficients according to this expression; and finally, f_damp is a function that damps the dispersion correction for shorter interatomic distances in order to avoid near singularities at small distances but also mid-range double-counting effects of correlation at intermediate distances.

Slide 10

In the first lecture we discussed the concept of basis set and we mentioned the two main classes of basis sets, the localized and the nonlocal or plane wave basis set, and that CP2K takes somehow advantage of both of them.

A drawback of the PW basis set is that to describe atomic wavefunctions a large number of basis functions are needed, much larger than the number needed by using a localized basis set in order to reach the same accuracy.

In this table there is a basis set size comparison in representing a simple 1s Slater-type function.

Slide 11

In particular, the problem is to accurately describe the wavefunctions for those electrons, that are closer to the nucleus, because their shapes show more oscillations near the nucleus than in the wavefunctions associated to the outermost electrons.

Therefore, an idea to overcome this issue is to replace the electronic degrees of freedom that are more problematic to represent with plane waves, by effective potentials to add to the Hamiltonian in order to correct the dynamics of the remaining electrons and compensate for the missing interactions with the removed electrons.

Of course, we would like that these potentials or pseudopotentials, as they are commonly called because they do not represent any real interaction, be additive and transferable, which impose to choose only atomic pseudopotentials, that is a pseudopotential per each atomic species and to remove only core electrons, that is the chemically inert ones.

Slide 12

To sum up, we replace the full-potential in the KS equations, that is the all-electron potential, with the interaction potential between the valence electrons plus the pseudopotentials associated to each atom of the system. The core states are eliminated and the valence electrons are described by so-called pseudo-wavefunctions with significantly fewer nodes, that is oscillations. This allows the pseudo-wavefunctions to be described with far fewer basis functions, making the PW basis set practical to use. In this approach only the chemically active valence electrons are therefore treated explicitly, while the core electrons are 'frozen', being considered together with the nuclei as rigid non-polarizable ion cores.

Once chosen the level of theory to use, that is once chosen the exchange-correlation functional, a pseudopotential for each atomic species can be derived from an atomic reference state, by requiring that the pseudo- and real all-electron valence wavefunctions have to have the same energies and amplitude (and thus the same density) outside a chosen core cut-off radius r_c.

However, this condition is not sufficient to uniquely determine the set of atomic pseudopotentials. Different alternatives and conditions are possible, and many different pseudopotential recipes have been devised in the years with different features, pros and cons.

Slide 13

One of the most widely used classes of pseudopotentials is the so-called norm-conserving pseudopotentials, which requires the four conditions here listed to be met.

Even these additional conditions are, though, not sufficient to uniquely determine the analytic form of the pseudopotential for any atomic species.

Therefore, during the years many different kinds of norm-conserving pseudopotentials have been developed and here I listed some related references.

Most of the available pseudopotentials in CP2K are of this kind,...

Slide 14

... including the Goedecker, Teter, Hutter pseudopotentials or GTH pseudopotentials, sometimes called Gaussian and dual space pseudopotentials, which are the ones employed in the practicals of this course.

As for many other norm-conserving pseudopotentials the GTH pseudopotentials are formed by a local part, which does not depend on the angular momentum, and a nonlocal part, which does. Some details are reported in this slide for your reference but are not relevant now. What it is important to emphasize here is the reason why GTH pseudopotentials are so popular within the CP2K community and for us. These pseudopotentials

- are separable, a nice property for efficient computation;
- give in fact optimal efficiency in numerical calculations using plane waves as a basis set;
- at most, only seven coefficients are necessary to specify its analytic form;
- but above all they have optimal decay properties in both real and Fourier space. From this the name of dual space pseudopotentials. Because of this property, the application of the nonlocal part of the pseudopotential to a wave function, usually the most computationally expensive part of the calculations involving pseudopotentials, can be done efficiently on a grid in real space. And real space integration is much faster for large systems like the biomolecules, than ordinary multiplication in Fourier space, since the scaling of these operations in the real space is quadratic with respect to the size of the system, while in the Fourier space is cubic.

Therefore, the use of GTH pseudopotentials significantly contributes to the CP2K capability to scale very efficiently with respect to the size of the system, a feature this for which CP2K stands out from many other quantum codes.

Slide 15

If you use pseudopotentials, as you do in the practical by employing a GPW scheme, that is a Gaussian and Plane Wave approach, the choice of the localized basis set has to be made in combination with the specific chosen pseudopotential class.

In the case of the GTH pseudopotentials, CP2K offers many possible basis sets among which to choose. The first choice is about the type of basis functions to consider.

We mentioned before the Slater functions, which are functions very similar to the orbital solutions of the Schrödinger equation for an atom. However, nowadays a more common and computationally efficient choice is to use the Gaussian type functions. These primitive functions resemble less an orbital solution of the atomic Schrödinger equation, but combining some of them together and building the so-called contracted functions, we can have similar results of the Slater function set but computationally more efficient, because computing integrals of Gaussians is much easier for a computer than computing integrals of Slater functions.

The second choice in selecting the basis set is about its accuracy, that is in this case the number of basis functions you want to use per atomic orbital. The smallest basis set employs only enough functions for a minimum description of the occupied orbitals of the neutral atoms and is called a minimum or Single Zeta (SZ) basis set. For hydrogen and helium atoms, for example, a Single Zeta basis set has only a single s-function. For the elements in the second row in the periodic system it means two s-functions (1s and 2s) and one set of p-functions (2p_x, 2p_y and 2p_z).

The next improvement of the basis set is a doubling of all basis functions used for each atomic orbital, producing a Double Zeta (DZ)-type basis. Then, the next step up in basis set size is a Triple Zeta (TZ) and so on.

The "V" in the name of the basis sets here, which correspond to the names you can find in CP2K, means "valence" and refers to the fact that by using pseudopotentials our electronic degrees of freedom will be only the valence electrons and not the core ones.

Finally, as you can see in the table, the size of the basis sets are usually increased with the so-called polarisation functions, identified with the final "P" in the name of the basis set. In fact, to improve the accuracy in representing the molecular orbitals, functions with higher angular momentum than the valence orbitals have been shown to became important, in particular for better descriptions of bonding but also for taking into account polarization effects, and these functions are called for this reason polarization functions.

In the tutorials, a Double Zeta Valence basis set with a single set of polarization functions is selected. In fact, this basis set is the smallest basis set commonly considered suitable for production runs.

Moreover, among the different basis sets offered by CP2K, you will probably focus on the so-called MOLOPT subclass, that is the basis sets optimized for molecular calculations. This is done by fitting the parameters, like the zeta exponents of the primitive functions and the coefficients of the contracted functions, with respect to a training set of small molecules formed with different elements and with different coordination environments.

The lecture ends here. If you have questions or doubts, you can ask them in the Q&A session.