Chapter 8

Density Functional Theory

8.1 Introduction

In the previous lecture we discussed the Hartree-Fock method. This is a powerful *ab initio* method for calculating the electronic structure of molecules, but has only limited success with condensed phases in general, and copes particularly poorly with metallic solids in particular. An alternative method, that is particularly suited to bulk materials, and especially metals, is that of Density Functional Theory. Whereas the fundamental object in HF is the many-body wavefunction which is made up of one-electron wavefunctions (orbitals), the fundamental quantity in DFT is the electronic charge density which is made up of one-electron densities. In this lecture we shall give a brief outline of DFT, discuss a few practical details about common implementations and typical strengths and weaknesses.

8.2 A fundamental theorem

Density Functional Theory is based upon the *Hohenberg-Kohn theorem* - which is that the ground state energy of a system of electrons is solely determined by the ground state charge density and that the charge density which gives the minimum energy is unique. This is a tremendous simplification compared to more traditional approaches such as Hartree-Fock. For a system with N electrons, the many-body wavefunction is $3N$-dimensional, whereas the corresponding charge density is a 3D scalar field! Furthermore, the theorem shows that this ground state energy can be found by minimizing an *energy functional* w.r.t. the charge density. That is, the energy is given by an integral of an (as yet) unknown function of the density over all space, and we minimize the value of this integral to find the optimum density! The problem then is to determine what this function of the density is before we can start to minimize anything . . .

8.3   The search for a universal functional

A year after the publication of the Hohenberg-Kohn theorem (which was viewed at the time as interesting but not very useful) came the publication of the *Kohn-Sham equations*. It was this which turned the Hohenberg-Kohn theorem into a practical calculational scheme. The basis of their work was to write down a form of this unknown energy functional, and show how it could be used to derive a set of effective one-electron Schrödinger-like equations.

Once again, we start from the variational principle, but now in the form of a functional of the electron density (using square brackets, $E[n]$ to show that this is a functional and not a function), that is:

$$E[n] = \min_{\Psi|n} \int \Psi^{\star} \widehat{H} \Psi d^{3N} r \qquad (8.1)$$

where by $\min_{\Psi|n}$ we mean minimizing w.r.t. the set of many-body wavefunctions $\{\Psi\}$ which are consistent with the density $n(\mathbf{r})$. This procedure is equivalent to minimizing w.r.t. the density, which must be subject to the constraint

$$\int n(\mathbf{r}) d^3 r = N \qquad (8.2)$$

where N is the number of electrons, which therefore determines the normalization of Ψ.

We use the Born-Oppenheimer approximation, and consider the system as a set of interacting electrons and ions. We proceed by separating the Hamiltonian into two parts - $\widehat{H}_0$ due to the *homogeneous electron gas* (i.e. jellium - a system of electrons in a uniform, neutralizing, background of positive charge) and $V_{ext}(\mathbf{r})$ being the *external potential* (e.g. the electron-ion Coulomb attraction). We then have

$$E[n] = \min_{\Psi|n} \int \Psi^{\star} \widehat{H_0} \Psi d^{3N} r + \int V_{ext}(\mathbf{r}) n(\mathbf{r}) d^3 r \qquad (8.3)$$

and the first term can be written as $F[n]$, i.e. as a functional of the electrons alone. This functional $F[n]$ does not depend upon the ions or any other external influence, and so must be a universal functional of the electrons. Now all we have to do is find this functional!

We know from our experience with Hartree-Fock theory, that $F[n]$ must contain contributions due to the electronic kinetic energy and the electron-electron Coulomb repulsion, as well as the effects of electron exchange and dynamical correlations. At this point, Kohn-Sham broke this unknown functional into 3 parts:

$$F[n] = T[n] + \frac{1}{2} \int \int n(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} n(\mathbf{r}) d^3 r' d^3 r + E_{xc}[n] \qquad (8.4)$$

where $T[n]$ is a kinetic energy functional, the second term is the electron-electron Coulomb repulsion that we have seen before, and all the many-body effects (including exchange and correlation) are lumped together into the third term - the *exchange-correlation functional $E_{xc}[n]$*.

Unfortunately, it is not known how to calculate the kinetic energy of a charge density. So, they then introduced a fictitious set of non-interacting one-electron wavefunctions, $\{\psi_i\}$, which *do not* represent atomic orbitals but instead are merely chosen to reproduce the ground state density:

$$n\left(\mathbf{r}\right) = \sum_{j=1}^{N} \left|\psi_j\left(\mathbf{r}\right)\right|^2 \qquad (8.5)$$

and so we can now write down the kinetic energy in terms of these functions:

$$T = -\frac{1}{2}\sum_{j=1}^{N} \nabla^2 \psi_j \qquad (8.6)$$

which is actually the kinetic energy of a set of non-interacting electrons that have the same density as the interacting system, and neglects any effects of interaction on kinetic energy.

Note that at this stage, everything in $F[n]$ is exact and universal for any electron charge density, and all the dependence on ions has been put into $V_{ext}\left(\mathbf{r}\right)$. The only unknown quantity remaining is the exchange-correlation functional which by definition therefore contains all the contributions not accounted for by the other terms!

We can now perform the minimization of equation 8.1 w.r.t. the density subject to the constraint of equation 8.2, which then results in the *Kohn-Sham equations*:

$$\left(-\frac{1}{2}\nabla^2 + V_{eff}\right)\psi_j = \varepsilon_j \psi_j \qquad (8.7)$$

which look like a set of one-particle Schrödinger equations for the Kohn-Sham eigenfunctions in terms of an *effective potential V_{eff}* where

$$V_{eff}\left(\mathbf{r}\right) = \int n\left(\mathbf{r}'\right)\frac{1}{|\mathbf{r}-\mathbf{r}'|}d^3r' + V_{xc}\left(n\left(\mathbf{r}\right)\right) + V_{ext}\left(\mathbf{r}\right) \qquad (8.8)$$

and all the unknown terms are contained in the exchange-correlation potential V_{xc}.

At this point, we have all the ingredients for an efficient and exact scheme bar one - we don't know what to do for V_{xc}. So, in true physicist style, we approximate it by things we do know!

8.4   The Exchange-Correlation functional

The exchange-correlation functional is the key ingredient of DFT, and whilst the Hohenberg-Kohn theorem assures us that it exists and is a universal functional that will apply to all electron densities, regardless of how the electrons are interacting, it does not tell us what the form of this functional is!

Kohn-Sham made the following apparently drastic approximation in the spirit of the mean-field treatment of everything else in DFT, that has come to be known as the *Local Density Approximation* (LDA). That is, they assumed that an electron at a point with a given local electron density experiences the same many-body response by the surrounding electrons as if the whole material had this same density. Of course, if the entire system had the same density at all points, then this would correspond to "jellium", i.e. the uniform homogeneous electron gas. This approximation might be expected to work well for metallic crystals with relatively uniform electron densities - which is true. It has also been found, somewhat more surprisingly, to be surprisingly accurate for many other materials, including insulators and semiconductors (both solid and liquid), surfaces, and even molecules.

Formally, we write:

$$E_{xc} = \int \varepsilon_{xc} \left[n\left(\mathbf{r}\right) \right] n\left(\mathbf{r}\right) d^3 r \tag{8.9}$$

where $\varepsilon_{xc}\left[n\right]$ is the exchange-correlation energy per particle of a homogeneous electron gas at density n which has been calculated using another very accurate technique - the Quantum Monte Carlo method - and has been conveniently parameterized for use in LDA calculations.

This approach (and indeed the whole of DFT) can also be extended to spin-polarized systems, where there is a different number of spin-up and spin-down electrons, and the exchange-correlation functional is split into two parts for the spin-up density and the spin-down density. This results in the *Local Spin Density Approximation* (LSDA). New exchange-correlation functionals have recently been developed that include both the local density and also the gradients of the local density, which are collectively known as the *Generalized Gradient Approximation* (GGA).

All forms of the exchange-correlation functional lead to the formation of the exchange-correlation hole. That is, there is a region surrounding each electron where other electrons are excluded - see figure 8.1. This hole in the electron density integrates to exactly one electron, and therefore reveals a single positive charge from the jellium background. It is the electron and this hole that then move together as a single neutral entity. The LDA therefore represents this hole by a sphere, whilst GGAs put in more elaborate shape effects. The key to the effectiveness of the LDA is that it correctly integrates to give exactly one excluded electron in the exchange-correlation hole.

$$\rho(r)$$
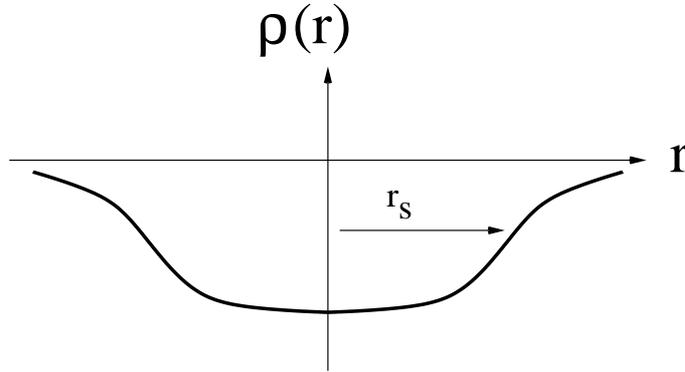
Figure 8.1: Simple view of the exchange-correlation hole, with a typical size r_s being dependent on the local electron density.

8.5 Solving the Kohn-Sham equations

The traditional approach to solving the Kohn-Sham equations is to expand the Kohn-Sham eigenfunctions in terms of plane-waves:

$$\psi_j\left(\mathbf{r}\right) = \sum_{\mathbf{g}} c_{j,\mathbf{g}} e^{i\mathbf{g}\cdot\mathbf{r}} \tag{8.10}$$

where $\mathbf{g}$ is a *reciprocal lattice vector,* as the one-electron Kohn-Sham functions do not have any connection with atomic orbitals, but are merely a convenience for calculating the density and the kinetic energy. If we then substitute equation 8.10 into equation 8.7 and integrate, we get the following *secular equation*:

$$\sum_{\mathbf{g}'} \left(\mathbf{g}^2 \delta_{\mathbf{g},\mathbf{g}'} + V_{Hartree}\left(\mathbf{g}\right) + V_{xc}\left(\mathbf{g}-\mathbf{g}'\right) + V_{ext}\left(\mathbf{g}-\mathbf{g}'\right)\right) c_{j,\mathbf{g}} = \varepsilon_j c_{j,\mathbf{g}} \tag{8.11}$$

which can be trivially cast into matrix form and the Kohn-Sham eigenvalues found by matrix diagonalization as seen in previous lectures. This can be a computationally expensive procedure - a system with N electrons and M basis functions will require the construction and storage of an $M \times M$ Hamiltonian matrix, which will require $O(M^3)$ operations to diagonalize per iteration towards self-consistency. The use of efficient pseudopotential schemes (as discussed in earlier lectures) can help to reduce M but exact diagonalization is still an expensive operation. This was for a long time the major bottle-neck in the method, and prevented the application of the method to systems with any more than a few tens of atoms on even the fastest supercomputers.

8.5.1   Car-Parrinello approach

A major breakthrough came in 1985 when Car and Parrinello put forward
an alternative solution scheme. They considered treating each coefficient
$c_{j,\mathbf{g}}$ in the plane-wave expansion as a separate dynamical variable by asso-
ciating a fictitious mass with each coefficient. They therefore viewed the
problem as being analogous to classical mechanics, with M particles moving
in some potential V_{eff} which could then be studied using traditional molec-
ular dynamics techniques. If a fictitious damping was added, then the net
result was that the coefficients, initially set to some random values, would
evolve to the lowest energy configuration, and hence generate the ground
state wavefunction of the system! As well as being a novel way of approach-
ing the problem, it was also significantly faster than matrix diagonalization,
requiring $O(N^2 M)$ operations to reach the ground state. At a stroke, this
increased by an order of magnitude the number of atoms that could be rea-
sonably studied.

  Car and Parrinello also showed how it was possible to include the motion
of the real ions at the same time, and hence do *ab initio* molecular dynamics,
i.e. move the atoms around using forces derived from the full QM treatment
of the electrons. If damping was added to the ions as well, then this was
now an efficient scheme for generating the optimal configuration of all the
atoms in the system. It could also be used, for the first time, to study real
dynamical processes with an *ab initio* method.

8.5.2   Conjugate-gradients approach

A few years later, another breakthrough was achieved by Teter and Payne,
when they realized that the Car-Parrinello method was effectively treating
the problem as one of *function minimization*, and that the minimizing al-
gorithm they were using was equivalent to the standard *steepest descents*
algorithm. Whilst this is a simple and robust algorithm, it is well know to
be far from the most efficient minimizer available. Therefore, a new approach
was introduced, of iterative minimization using the *conjugate-gradients* al-
gorithm. A schematic of the difference between these two methods is shown
in figure 8.2. This new algorithm produced another order-of-magnitude im-
provement in speed (although the scaling was still $O(N^2 M)$). The new
method could be used both to minimize the energy of the electrons and also
to move the atoms around, either in molecular dynamics fashion or to find
the equilibrium geometry.

  These algorithmic improvements, coupled with developments in the con-
struction of optimal pseudo-potentials, mean that it is now routine to calcu-
late the ground state properties of systems with around a hundred atoms on
a workstation, and several thousand atoms on a supercomputer. In my own
research, I have used a single processor workstation to calculate the optimal
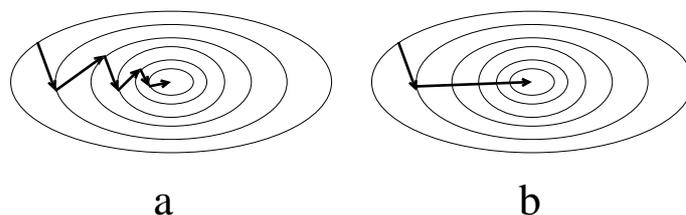
Figure 8.2: Schematic showing the search strategy and consequent performance of two different function minimization algorithms, in a 2D anisotropic harmonic potential. The equipotentials are therefore ellipses as shown. (a) show the behaviour of the steepest descents minimizer which takes many steps to find the minimum, whilst (b) shows the conjugate-gradients minimizer which requires just two steps.

geometry of a system of 864 atoms using these methods (although it took many days to complete and the calculation required 4GB RAM).

8.6   Alternatives to Kohn-Sham

Recently, there has been a lot of effort in finding new formulations of DFT that do not require the introduction of the Kohn-Sham functions, but rather using a different formulation of quantum mechanics based upon *density matrices*. The advantage of such a technique is that it would scale as linearly with the number of atoms - that is, it would be an $O(N)$ *method*. This is something of a Holy Grail for electronic structure calculations, as it would be much more efficient than any other technique for large systems, and would enable the study of very large systems (e.g. 10,000+ atoms). For example, it would then be possible to do first-principles calculations of biological molecules (proteins, etc) for the first time. This is an area of intense ongoing research, with some some of the best results coming from the Cambridge group in recent months.

There has also been much interest recently in generalised Kohn-Sham schemes, where a variety of different approaches to the unknown universal functional have been explored, for example, incorporating recent insights from Many Body Perturbation Theory as in recent work from Prof. Godby's group.

8.7   Results

Like Hartree-Fock, DFT calculations calculate the total energy of a system of atoms, from which many structural, electronic and dynamical properties can be derived. However unlike HF, it is also routinely applied to condensed phases, which opens up new possibilities, such as the calculation of lattice

constants, structural phase transitions, elastic constants, phonon spectra, and more.

Using the LDA, it is usually possible to get lattice constants and inter-atomic distances correct to at least ± 0.02 Å , bond-angles to within $\pm 3°$, relative energies are typically correct to at least 0.005 eV/atom, and vibrational frequencies to within $10 - 50 \, cm^{-1}$. Similarly, it gives a good description of elastic constants, dipole moments, magnetic moments, etc.

However, the LDA tends to overbind atoms which results in it over-estimating the metallic character of a material. For example, one notable failure is that it predicts NiO to be a metal, whereas it is experimentally known to be an anti-ferromagnetic insulator. Similarly, it overestimates very weak bonds, such as hydrogen bonds.

It is possible, although not strictly justified, to consider the Kohn-Sham eigenvalues as electron energies. This often gives a good description of the band-structure of the material, as well as quantities like the work function, optical and UV spectra. However, the LDA tends to underestimate the band-gap in insulators and semiconductors. This is not too surprising as DFT is intended to be a ground-state only theory!

Using either the Car-Parrinello or conjugate-gradients techniques, it is possible to do *ab initio* molecular dynamics, using a typical time step of 1 fsec, and simulating dynamical processes for up to 10 psec on a supercomputer. This makes it possible to directly study dynamical phenomena such as structural phase transitions, diffusion, chemical reactions and catalysis.

The use of GGA improves certain areas, and makes it possible to get absolute binding energies as good as 0.05 eV/atom, and gives a much better description of subtle effects, such as transition states and hydrogen bonds. Unfortunately, it does not give a uniform improvement for all quantities for all systems! The GGA is not a magic bullet. . .

However, there are some more unusual materials for which DFT as described is not an appropriate theory. These are known as *strongly-correlated* materials and require an approach which goes beyond the mean-field approach of DFT. (Surprisingly, HF often works quite well for these materials). Such materials include high-temperature superconductors and colossal magneto-resistance materials. Recent theoretical advances in treating such materials involve developing a time-dependent DFT which is giving very promising results.

Note that the precision of DFT results tends to be less than HF because of the use of numerical integration on grids rather than analytic integration. However, it is simple to make the precision better than the accuracy due to approximate treatment of exchange and correlation, which is therefore good enough.

8.8   Final comments

A few final points to highlight:

- DFT is most often used to study condensed phases, using periodic boundary conditions, involving any atom in the periodic table.

- It is an exact theory in principle, but in practice uses an approximate treatment for exchange and correlation.

- It is most often used with a plane-wave basis set and pseudo-potentials.

- There is no systematic way to improve the accuracy of the calculation (unlike HF with multiple determinants).

- It is simple to parallelize the method using different sets of plane-waves on different nodes of a parallel computer. This makes it simple to apply to very large systems - up to a few thousand atoms.

- It can also be used to generate *ab initio* molecular dynamics, which further widens the range of phenomena that can be studied.

8.9   Further reading

- DFT in "Methods of Electronic Structure Calculations" by M. Springborg, Chapter 15

- DFT in "Computational Physics" by J.M. Thijssen, chapter 5