Electronic structure, plane waves and pseudopotentials

P.J. Hasnip

DFT Spectroscopy Workshop 2009

Introduction

We want to be able to predict what electrons and nuclei will do from *first principles*, without needing to know what they'll do beforehand. We can do this using quantum mechanics.

We could try to solve the Schrödinger equation

$$-i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi.$$

where we can write

$$\hat{H} = \hat{T}_{en} + \hat{V}_{e-e} + \hat{V}_{e-n} + \hat{V}_{n-n}$$

The Many-Body Wavefunction

The full many-body wavefunction is a function of time, and the position of every particle:

$$\Psi = \Psi(\mathbf{r}_1, ..., \mathbf{r}_n, \mathbf{R}_1, ..., \mathbf{R}_N, t).$$

It tells us the probability of any particular configuration of electrons and nuclei occurring at any particular time.

The Born-Oppenheimer Approximation

Compared to electrons, nuclei are massive and slow. This has two consequences:

- Whenever a nucleus moves, the electrons react so quickly that it may as well be instant.
- The wavefunctions for the nuclei are zero except in a very small region – we may as well forget the wavefunction and just say 'there they are'!

The Born-Oppenheimer Approximation

Compared to electrons, nuclei are massive and slow. This has two consequences:

- Whenever a nucleus moves, the electrons react so quickly that it may as well be instant.
- The wavefunctions for the nuclei are zero except in a very small region – we may as well forget the wavefunction and just say 'there they are'!

The Born-Oppenheimer Approximation

Now we only have to worry about quantum mechanics for the electrons. Since they react instantly to any change in the positions of the nuclei we only have to solve the time-*independent* Schrödinger equation:

$$\hat{H}\psi = E\psi$$

where we can write

$$\hat{H} = \hat{T}_e + \hat{V}_{e-e} + \hat{V}_{e-n}$$

Density Functional Theory

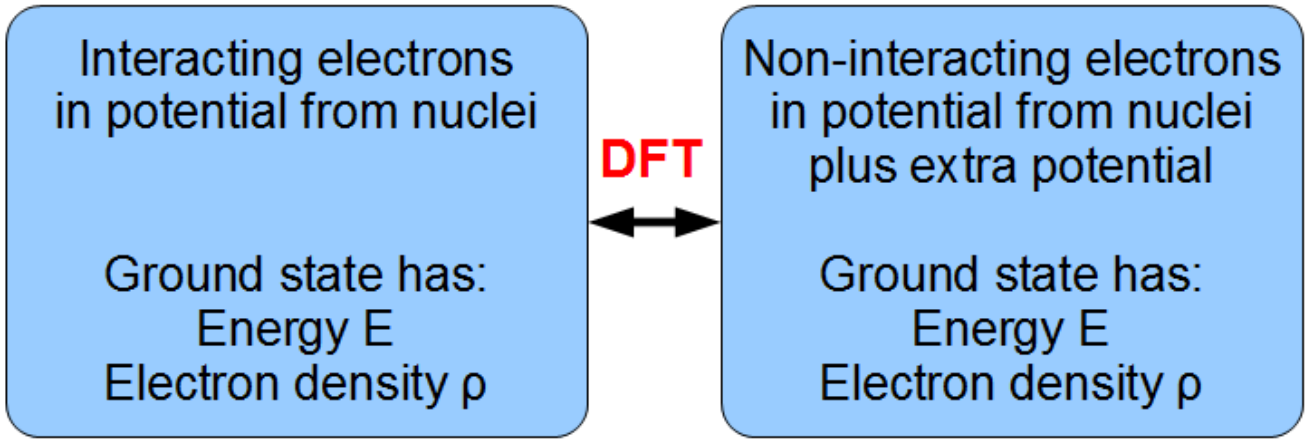
$$\hat{H} = \hat{T}_e + \hat{V}_{e-e} + \hat{V}_{e-n}$$

Unfortunately this is still really difficult – we don't know $\hat{T}_e$ and $\hat{V}_{e-e}$.

In real life the behaviour is usually dominated by the ground state. Can this help us?

Density Functional Theory

In the mid-1960s two papers by Hohenberg and Kohn, and Kohn and Sham gave a solution to this tricky problem. They showed that the ground state energy and charge density of interacting electrons in any external potential were exactly the same as those of *non-interacting* electrons in a specially modified potential.



Density Functional Theory

In the Kohn-Sham scheme, we write

$$E_{kin} + E_{e-e} = E_{kin}^{non-int.} + E_{e-e}^{non-int.}[\rho] + E_{xc}[\rho]$$

The extra term on the right-hand side is a functional of the electron density $\rho(r)$ and is called the exchange-correlation functional.

We can compute $\rho(r)$ from the wavefunctions ψ_m,

$$\rho(\mathbf{r}) = \sum_m |\psi_m(\mathbf{r})|^2$$

Density Functional Theory

Our final Hamiltonian is:

$$\hat{H}[\rho] = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}_{e-e}^{non-int.}[\rho] + \hat{V}_{e-n}[\rho] + \hat{V}_{xc}[\rho].$$

and particle m goes in the mth solution of

$$\hat{H}[\rho]\psi_m = E_m\psi_m.$$

Unfortunately we don't know what $\hat{V}_{xc}[\rho]$ is!

Density Functional Theory

Our final Hamiltonian is:

$$\hat{H}[\rho] = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}_{e-e}^{non-int.}[\rho] + \hat{V}_{e-n}[\rho] + \hat{V}_{xc}[\rho].$$

and particle m goes in the mth solution of

$$\hat{H}[\rho]\psi_m = E_m\psi_m.$$

Unfortunately we don't know what $\hat{V}_{xc}[\rho]$ is!

LDA

This is a pretty good approximation to the energy – only $E_{xc}[\rho]$ is unknown and this is often only about 10% of the total energy.

In the special case that the electrons are evenly spread throughout space, we can actually compute $E_{xc}[\rho]$. We can use this as an approximation to the true E_{xc} in our calculations.

In this approximation the contribution to E_{xc} from any region of space only depends on the density in that region, so this is usually called the *Local Density Approximation*, or just LDA.

DFT Summary

We started with a time-dependent, many-nuclei, many-electron quantum mechanics problem.

We've ended up with a static, single-particle quantum mechanics problem:

$$\hat{H}[\rho]\psi_m = E_m\psi_m$$

Now we just need to calculate enough states to accommodate all of our electrons.

Things to Remember

- Even 'perfect' DFT is only exact for the ground state
- Exchange-Correlation is approximated crudely
- No collective electron-nuclear motion
- No nuclear quantum effects (e.g. zero-point motion)

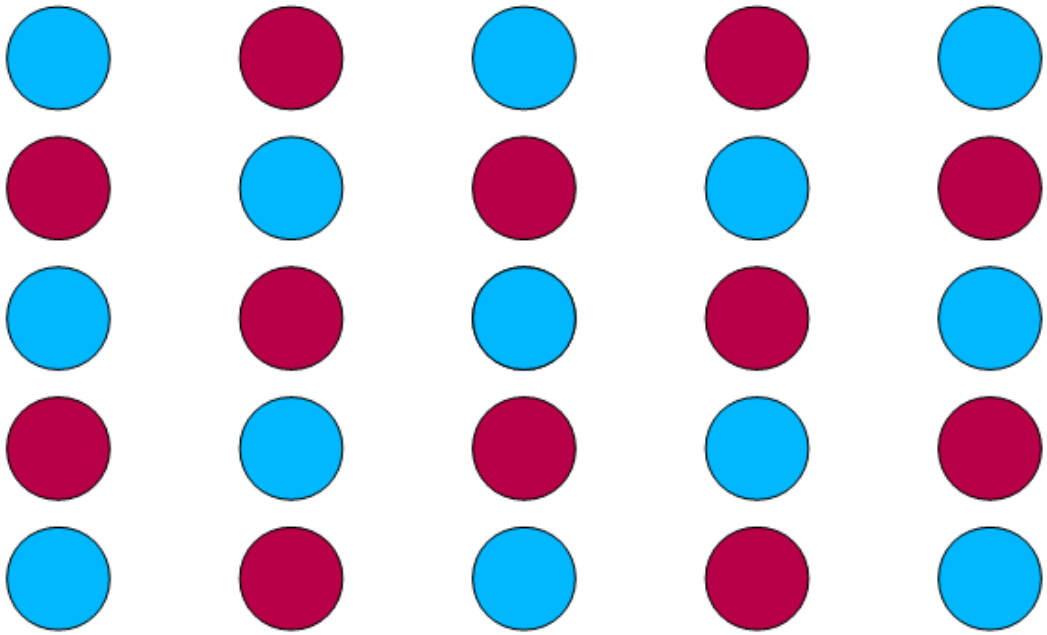
Lots of Atoms

We've simplified our quantum problem so that we just need to solve a single-particle equation, but we need to solve for enough states for every electron.

Even a few grams of material has over 10^{23} electrons, which means we need a *lot* of states! That could take us a while...

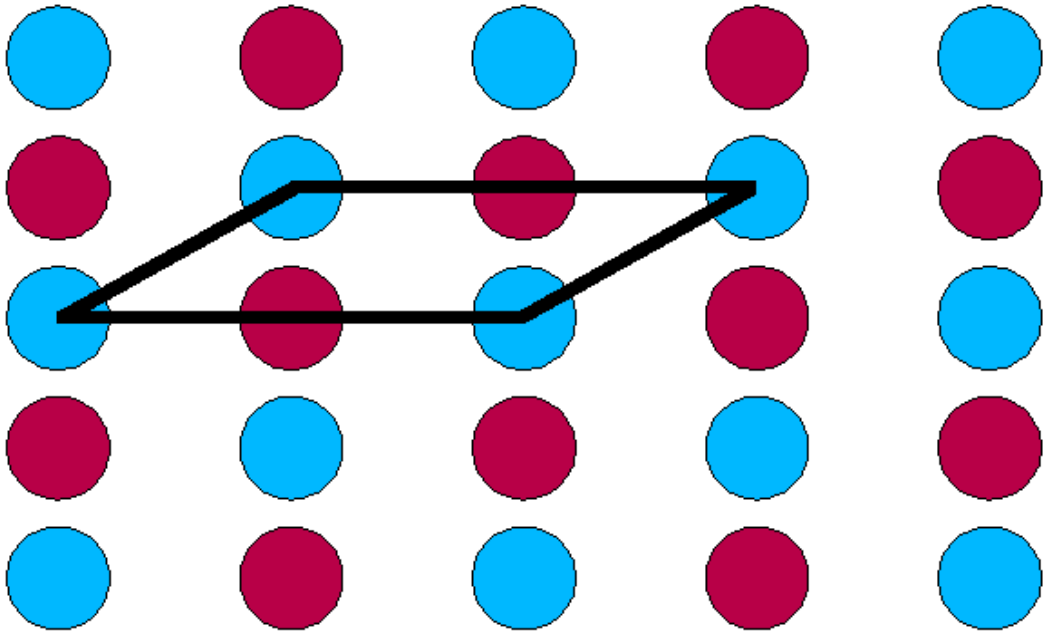
Crystals and Unit Cells

In the solid state, most materials like to have their atoms arranged in some kind of regular, repeating pattern, e.g.



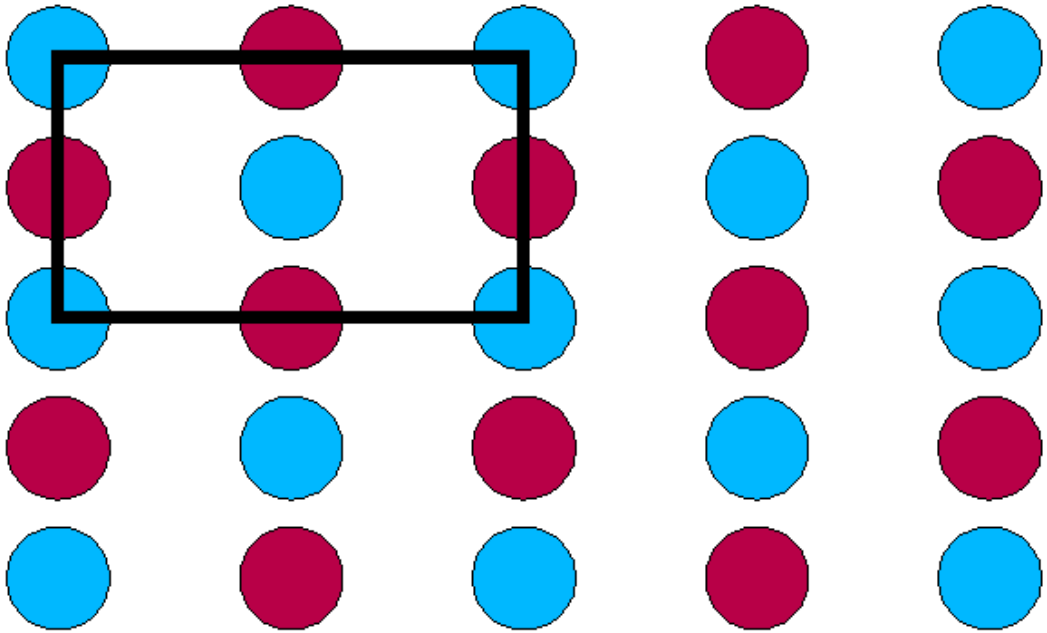
Crystals and Unit Cells

In the solid state, most materials like to have their atoms arranged in some kind of regular, repeating pattern, e.g.



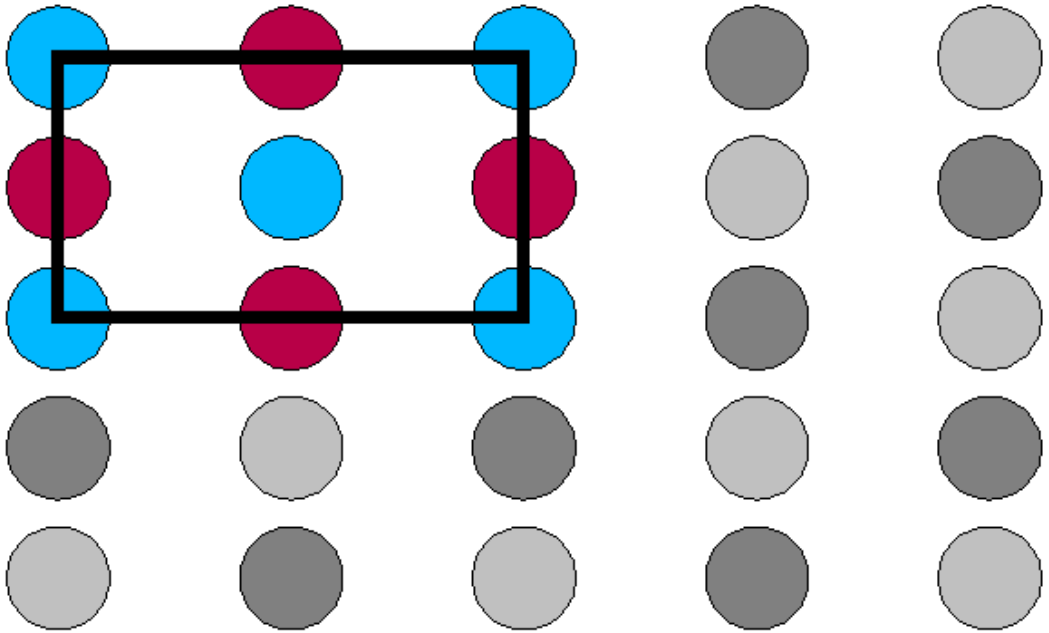
Crystals and Unit Cells

In the solid state, most materials like to have their atoms arranged in some kind of regular, repeating pattern, e.g.



Crystals and Unit Cells

In the solid state, most materials like to have their atoms arranged in some kind of regular, repeating pattern, e.g.



Bloch's Theorem

If the nuclei are arranged in a periodically repeating pattern, their potential acting on the electrons must also be periodic.

$$V(\mathbf{r} + \mathbf{L}) = V(\mathbf{r})$$

where $\mathbf{L}$ is any lattice vector.

What does this mean for the density and wavefunction?

Bloch's Theorem

If the potential is periodic, then so is the density:

$$\rho(\mathbf{r} + \mathbf{L}) = \rho(\mathbf{r})$$

What about the wavefunction?

$$\rho(\mathbf{r}) = |\psi(\mathbf{r})|^2$$

so if $\rho(\mathbf{r})$ is periodic, then so is the magnitude of the wavefunction.

Remember wavefunctions are complex; their magnitude is periodic, but their phase can be anything.

Bloch's Theorem

If the potential is periodic, then so is the density:

$$\rho(\mathbf{r} + \mathbf{L}) = \rho(\mathbf{r})$$

What about the wavefunction?

$$\rho(\mathbf{r}) = |\psi(\mathbf{r})|^2$$

so if $\rho(\mathbf{r})$ is periodic, then so is the magnitude of the wavefunction.

Remember wavefunctions are complex; their magnitude is periodic, but their phase can be anything.

Bloch's Theorem

Bloch's theorem: in a periodic potential, the density has the same periodicity. The possible wavefunctions are all 'quasi-periodic':

$$\psi_k(\mathbf{r}) = e^{i\mathbf{k}.\mathbf{r}} u_k(\mathbf{r}),$$

where $u_k(\mathbf{r}+\mathbf{L}) = u_k(\mathbf{r})$, and $e^{i\mathbf{k}.\mathbf{r}}$ is an arbitrary phase factor.

$$\begin{aligned} \psi_k(\mathbf{r}+\mathbf{L}) &= e^{i\mathbf{k}.(\mathbf{r}+\mathbf{L})} u_k(\mathbf{r}+\mathbf{L}) \\ &= e^{i\mathbf{k}.\mathbf{L}} \psi_k(\mathbf{r}) \end{aligned}$$

k-point sampling

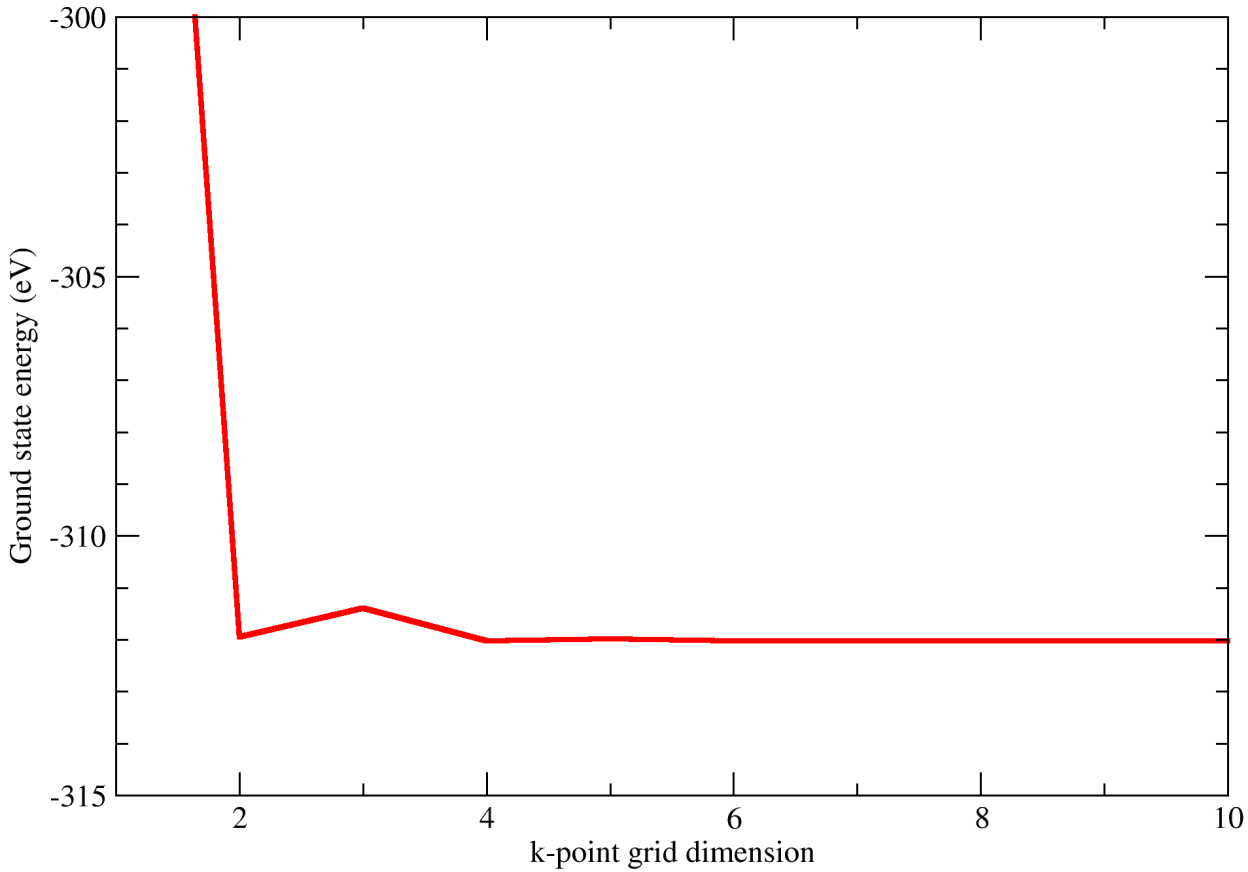
In principle we need to integrate over all possible **k** when constructing the density. Fortunately the wavefunctions change slowly as we vary **k**, so we can approximate the integral with a summation:

$$\begin{aligned} \rho(\mathbf{r}) &= \int |\psi_k(\mathbf{r})|^2 d^3\mathbf{k} \\ &\approx \sum_k |\psi_k(\mathbf{r})|^2 \end{aligned}$$

We need to make sure we use enough **k**-points to get accurate results.

We'll be looking at **k**-points more closely in a later talk.

k-point sampling



Basis sets

We need to choose a suitable basis set to represent our wavefunctions, but what should we choose...

- Points on a grid?
- Polynomials?
- Gaussians?
- Atomic orbitals?

None of these reflect the periodicity of our problem.

Plane-waves

Since $u_k(\mathbf{r})$ is periodic, we express it as a 3D Fourier series

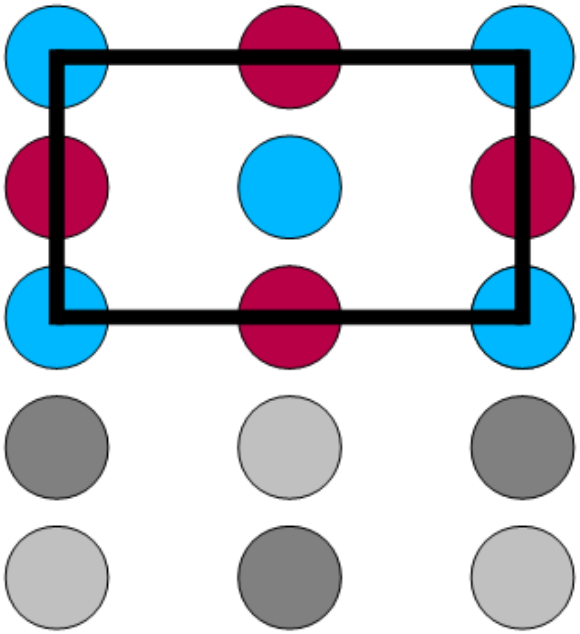
$$u_k(\mathbf{r}) = \sum_G c_{Gk} e^{i\mathbf{G}.\mathbf{r}}$$

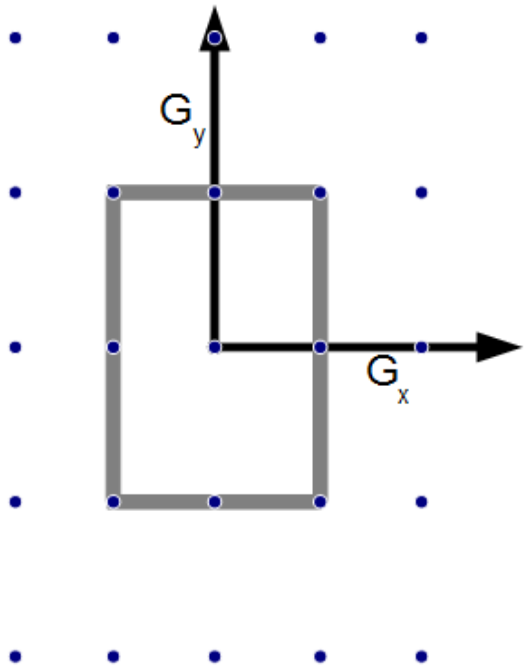
where c_{Gk} are complex Fourier coefficients, and the sum is over all wavevectors (spatial frequencies) with the right periodicity.

Only a discrete set of wavevectors have the right periodicity – these are the reciprocal lattice vectors. If we make the cell longer in one direction, the allowed wavevectors in that direction become shorter.

Reciprocal-space

Since wavevectors are measured in units of 1/length, they describe points in reciprocal-space. We can plot the allowed wavevectors in this space:





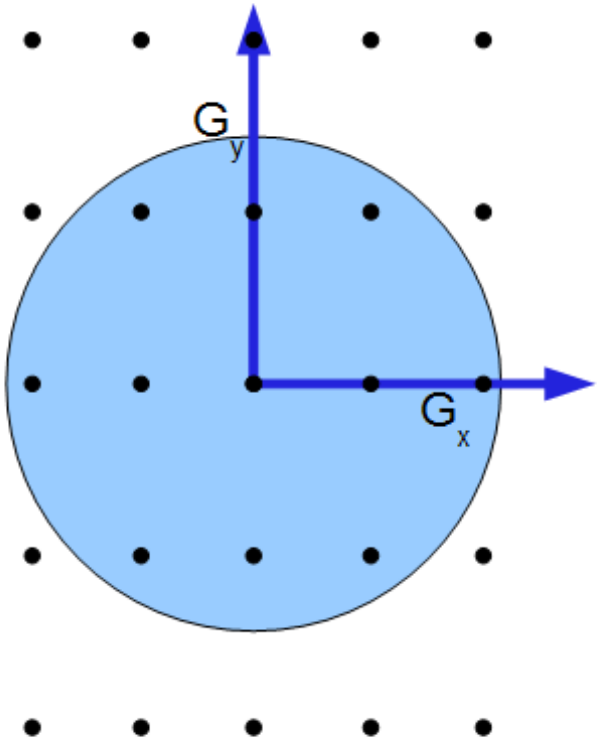
Cut-off Energy

Each of the Fourier basis functions $e^{i\mathbf{G}.\mathbf{r}}$ represents a plane-wave travelling in space, perpendicular to the vector $\mathbf{G}$.

There are an infinite number of allowed $\mathbf{G}$, but the coefficients c_{Gk} become smaller and smaller as $|\mathbf{G}|^2$ becomes larger and larger.

We define a cut-off energy $E^{cut} = \frac{\hbar^2}{2m}|\mathbf{G}|^2$ and only include plane-waves with energies less than this cut-off.

Cut-off Energy

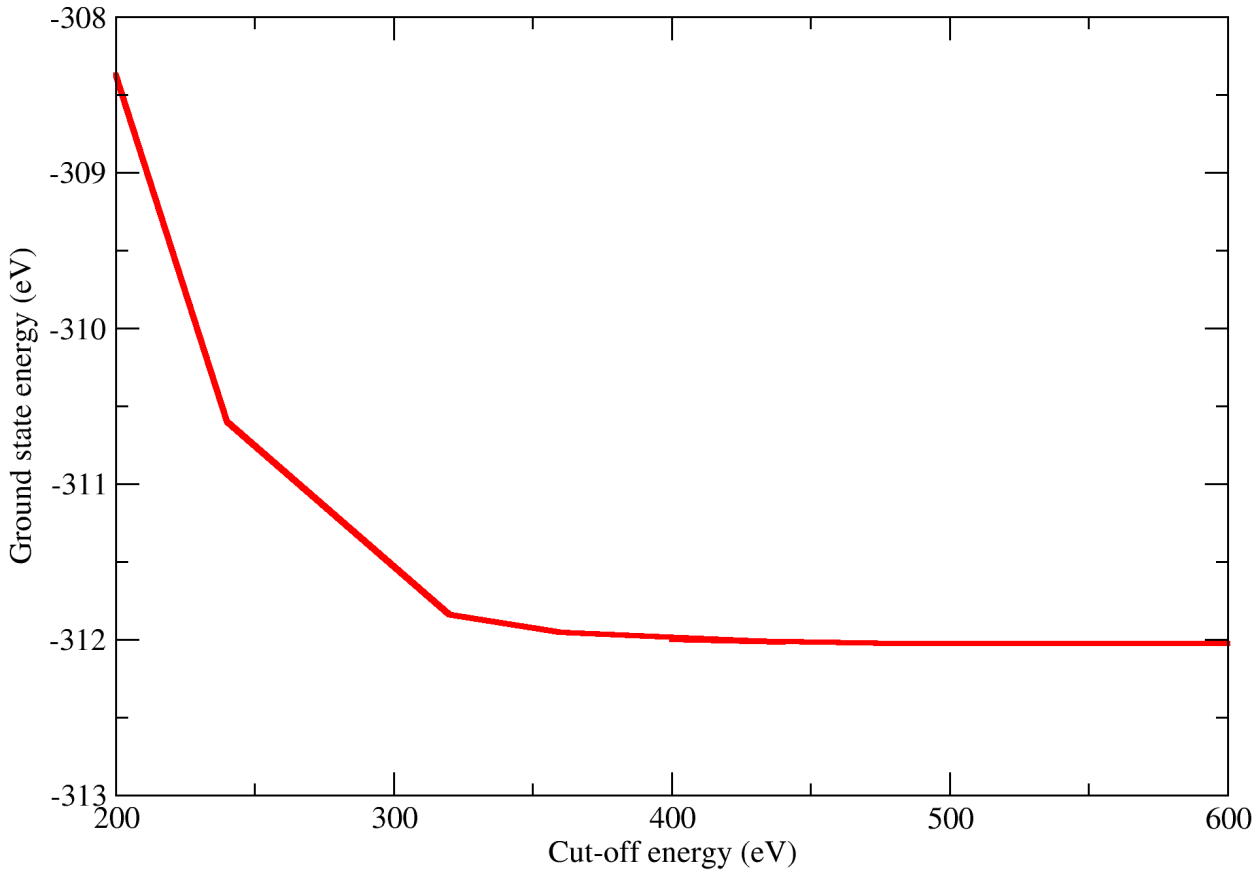


Cut-off Energy

We always have to ensure the cut-off energy is high enough to give accurate results. We repeat the calculations with higher and higher cut-off energies until the properties we're interested in have converged.

How do the results of our calculations depend on the cut-off energy?

Cut-off Energy



Pseudopotentials

The very large **G** components describe the region of space where the wavefunction is varying very quickly. This happens when the potential is very attractive – the strongest potentials are those near the nuclei.

Pseudopotentials

The wavefunctions near the nuclei are not actually very interesting, because they don't affect the chemical, mechanical or electronic properties very much.

We can replace the Coulomb potential near each nucleus with a modified, weaker potential. This modified potential is called a pseudopotential.

Now the wavefunctions don't vary as quickly near the nucleus, so we can use a smaller plane-wave cut-off energy.

Pseudopotentials

The core electrons spend all their time near the nucleus. They repel the outer electrons, so the outer electrons feel a weaker potential from the nucleus, but otherwise they don't affect the chemical properties etc.

Provided we reproduce this screening effect, we can ignore these core electrons altogether! We consider each atom's nucleus and core electrons as an ion, and produce a pseudopotential that has the same effect on the outer electrons.

Not only have pseudopotentials reduced the cut-off energy we need, they've also let us concentrate on the valence electrons, reducing the number of states we need from our Schrödinger equation.

Finding the Ground State

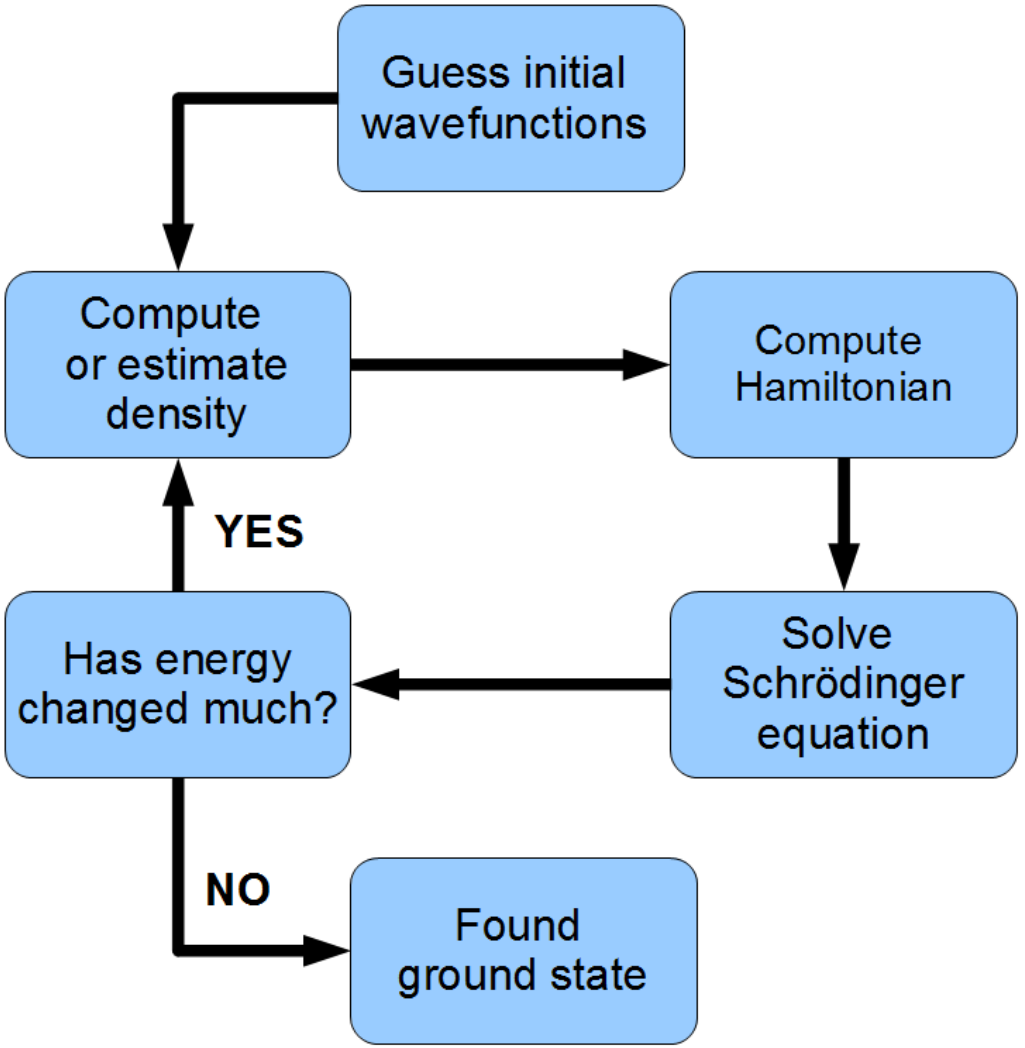
$$\psi_{m,k}(\mathbf{r}) = \sum_G c_{Gmk} e^{i(\mathbf{G}+\mathbf{k}).\mathbf{r}}.$$

Now all we need to do is solve $\hat{H}_k[\rho]\psi_{m,k} = E_{m,k}\psi_{m,k}$ to get the lowest energy states for each **k**. BUT the Hamiltonian depends on the electron density $\rho(\mathbf{r})$, which depends on the wavefunctions:

$$\rho(\mathbf{r}) = \sum_{m=1}^{N_e} \sum_{k=1}^{N_k} \left|\psi_{m,k}(\mathbf{r})\right|^2$$

We have to solve the equations *iteratively*.

Finding the Groundstate



Finding the Groundstate

You need to decide what energy change is acceptable, and this is set by an energy convergence tolerance. The smaller this tolerance, the closer Castep will get to the ground state before finishing – but of course it will take longer.

Castep can also compute the forces, and you can set a convergence tolerance on those as well as the energies. For many spectral properties, the forces are a more accurate measure of how far you are from the true ground state.

Finding the Groundstate

There are two different methods used in Castep to find the groundstate:

- Density Mixing (DM)
  - Estimates density
  - Fast
  - Possibly unstable
  - Energy converges quickly, forces converge slower
- Ensemble Density Functional Theory (EDFT)
  - Computes density
  - Slow
  - Stable
  - Energy and forces converge quickly

Summary

Castep uses a plane-wave basis set to express the wavefunctions, and these are sampled at a discrete number of **k**-points. The ground state is found by iteratively improving an initial guess until the change in energy is small.

You need to make sure your answers are converged with respect to:

- Cut-off energy
- **k**-point sampling