



RESEARCH  
COUNCILS UK



nag<sup>®</sup>



HECTOR

**CSC / NAG Autumn School on  
Core Algorithms in High-Performance  
Scientific Computing**

## **Case Study III**

**Phil Hasnip**

**Electronic energy minimisation, iterative  
diagonalisation**



## Case Study III: Electronic structure

Phil Hasnip

Department of Physics  
University of York

Core Algorithms in High Performance Scientific  
Computing



## What is Electronic Structure?

The properties of materials and chemicals are governed by their electrons. It is the electrons which determine whether atoms are bonded or not, and what the nature of any bonds are.

The term 'electronic structure' is used to describe how the electrons are spread around a crystal or molecule.



## Atomistic Simulations

If we know what the bonding in a material is beforehand, then we can often find good expressions for the forces between atoms, e.g.

- Ionic  $\Rightarrow$  electrostatic potentials
- Covalent  $\Rightarrow$  directional potentials, e.g. Stillinger-Weber
- Metallic  $\Rightarrow$  Bond Order Potentials, EAM...
- Van der Waals  $\Rightarrow$  Born-Mayer, Lennard-Jones...

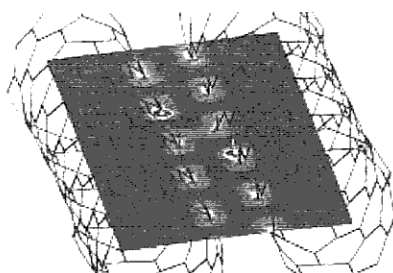
These potentials can do a good job of computing the mechanical properties of materials.



## Atomistic Simulations

These 'forcefield' potentials do have problems:

- Usually parameterised for bulk equilibrium behaviour
- Parameterised for particular class of materials
- Cannot handle bond breaking or formation
- Cannot provide electronic information, so no predictions of
  - Resistance
  - Thermal conductivity
  - Colour



## Electronic Structure Simulations

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

Unfortunately, quantum mechanics is difficult!



## Electronic Structure Simulations

We want to solve the Schrödinger equation. For 1-particle it's:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \hat{V}\right)\Psi(\mathbf{r}, t) = i\hbar\frac{\partial\Psi(\mathbf{r}, t)}{\partial t}$$

Now we just have to solve it for a real material.

How hard can it be?



## The horror.... the horror...

For  $M$  nuclei and  $N$  electrons:

$$\left\{-\frac{\hbar^2}{2}\left(\frac{\nabla_{n1}^2}{m_1} + \dots + \frac{\nabla_{nM}^2}{m_M}, \frac{\nabla_{e1}^2}{m} + \dots + \frac{\nabla_{eN}^2}{m}\right) + \hat{V}(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t)\right\}\Psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t) = i\hbar\frac{\partial\Psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t)}{\partial t}$$

How big are  $M$  and  $N$ ?



## The horror.... the horror...

For  $M$  nuclei and  $N$  electrons:

$$\left\{-\frac{\hbar^2}{2}\left(\frac{\nabla_{n1}^2}{m_1} + \dots + \frac{\nabla_{nM}^2}{m_M}, \frac{\nabla_{e1}^2}{m} + \dots + \frac{\nabla_{eN}^2}{m}\right) + \hat{V}(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t)\right\}\Psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t) = i\hbar\frac{\partial\Psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t)}{\partial t}$$

How big are  $M$  and  $N$ ? For a few grams of material:

- $M \sim 100,000,000,000,000,000,000,000$
- $N \sim 1000,000,000,000,000,000,000,000,000$



## The horror.... the horror...

For  $M$  nuclei and  $N$  electrons:

$$\left\{-\frac{\hbar^2}{2}\left(\frac{\nabla_{n1}^2}{m_1} + \dots + \frac{\nabla_{nM}^2}{m_M}, \frac{\nabla_{e1}^2}{m} + \dots + \frac{\nabla_{eN}^2}{m}\right) + \hat{V}(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t)\right\}\Psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t) = i\hbar\frac{\partial\Psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t)}{\partial t}$$

How big are  $M$  and  $N$ ? For a few grams of material:

- $M \sim 100,000,000,000,000,000,000,000$
- $N \sim 1000,000,000,000,000,000,000,000,000$

Oh dear.



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons
- Concentrate on the groundstate



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons
- Concentrate on the groundstate
- Exploit periodicity of crystals



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons
- Concentrate on the groundstate
- Exploit periodicity of crystals
- Get a computer to do it



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons
- Concentrate on the groundstate
- Exploit periodicity of crystals
- Get a computer to do it



## QM for the electrons

- Nuclei are heavy, slow and boring  
→ no QM for nuclei
- Electrons are light, quick and interesting  
→ Nuclei appear static  
→ No explicit time-dependence for electrons

This is the Born-Oppenheimer approximation.

$$\Psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N, t) \longrightarrow \psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

Now we can solve the time-independent Schrödinger equation.



## Electronic Structure Simulations

$$\left\{ -\frac{\hbar^2}{2m} (\nabla_1^2 + \dots + \nabla_N^2) + \hat{V}(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N) \right\} \psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E \psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

Hmm, looks a bit like an eigenvalue problem...  
Still have  $N \sim 10^{23}$  though.



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons✓
- Concentrate on the groundstate
- Exploit periodicity of crystals
- Get a computer to do it



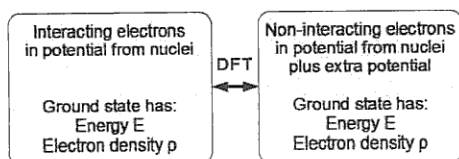
## Concentrate on the groundstate

- Materials' properties are dominated by the groundstate
- For the groundstate we can use a different form of QM  
Density Functional Theory



## Density Functional Theory

The groundstate energy  $E$  and density  $\rho(\mathbf{r})$  of electrons are exactly the same as those of *non-interacting* particles in a specially modified potential.



## Density Functional Theory

Solve  $N$  non-interacting Schrödinger equations:

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + V[\rho](\mathbf{r}) \right\} \psi_j(\mathbf{r}) = \epsilon_j \psi_j(\mathbf{r})$$

$V[\rho]$  is a *density functional*.

$$\rho(\mathbf{r}) = \sum_{j=1}^N |\psi_j(\mathbf{r})|^2$$

We don't know  $V[\rho]$  exactly, but there are reasonable approximations available.



## The story so far....

- Started with time-dependent  $10^{23}$ -electron-nuclear wavefunction
- Born-Oppenheimer approximation  
→ time-independent,  $10^{23}$ -electron wavefunction
- DFT →  $10^{23}$  1-electron wavefunctions



## Electronic Structure Simulations

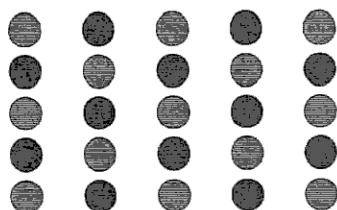
How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons✓
- Concentrate on the groundstate✓
- Exploit periodicity of crystals
- Get a computer to do it



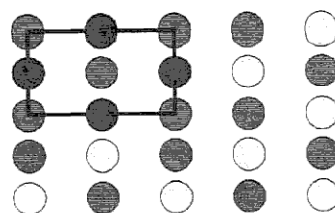
## Crystals and Unit Cells

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?



## 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?

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

i.e. if  $\rho(\mathbf{r})$  is periodic, 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':

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

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

For today's purposes we're going to assume that  $\mathbf{k} = (0, 0, 0)$ , which is sufficient for large simulations.



## Exploiting periodicity

- Only compute  $\psi_j(\mathbf{r})$  and  $\rho(\mathbf{r})$  in a single unit cell
- No longer have  $10^{23}$  wavefunctions, more like  $10^2$
- Now computing energy *per unit cell*

In fact because we've exploited the periodicity, in some ways we're now calculating for an *infinite* number of electrons!



## Electronic Structure Simulations

How can we make quantum mechanics easier for ourselves?

- Only use QM for the electrons✓
- Concentrate on the groundstate✓
- Exploit periodicity of crystals✓
- Get a computer to do it



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

$$\psi(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}}$$

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

Each  $\psi$  is now a vector of coefficients  $c_{\mathbf{G}}$ .

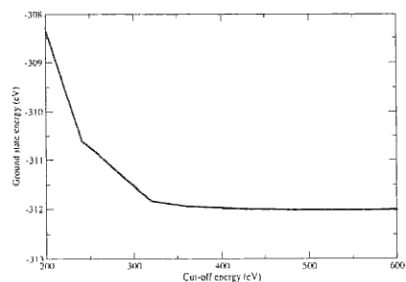


## Cut-off Energy

- $e^{i\mathbf{G} \cdot \mathbf{r}}$  is a plane-wave travelling perpendicular to  $\mathbf{G}$
- There are an infinite number of allowed  $\mathbf{G}$
- As  $|\mathbf{G}| \rightarrow \infty$ ,  $|c_{\mathbf{G}}| \rightarrow 0$   
 $\Rightarrow$  can truncate the Fourier expansion safely



## Cut-off Energy





## An Eigenvalue Problem

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + V[\rho](\mathbf{r}) \right\} \psi_j(\mathbf{r}) = \epsilon_j \psi_j(\mathbf{r})$$

- $\psi_j$  is a vector of num\_pw Fourier coefficients
- The {...} is a num\_pw x num\_pw matrix
- $\rightarrow$  just an eigenvalue problem!

$$H\psi_j = \epsilon_j \psi_j$$

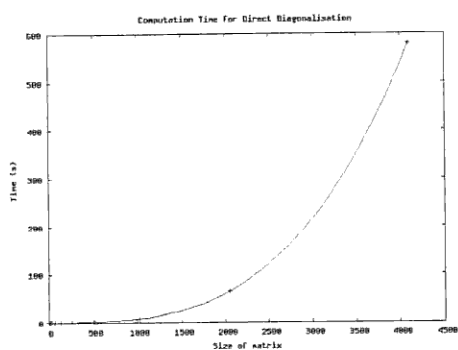


## A Large Eigenvalue Problem

- This is the theory behind the CASTEP DFT program
- Unfortunately need quite a lot of plane-waves
- One of the smallest standard benchmarks has 11,000
- A 11,000 x 11,000 matrix  $\approx$  1.8GB in size
- Diagonalisation time scales as num\_pw<sup>3</sup>



## A Large Eigenvalue Problem



## Iterative Diagonalisation

- Diagonalising 11,000 x 11,000 matrix  $\rightarrow$  11,000 states
- Only need 1 per electron: 164
- Can we just compute the eigenstates we need?



## An Eigenvalue Problem

Recall our Schrödinger equation:

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + V[\rho](\mathbf{r}) \right\} \psi_j(\mathbf{r}) = \epsilon_j \psi_j(\mathbf{r})$$

The {...} part is called the Hamiltonian,  $H$ .

Let's look at each part of this...



## Kinetic energy

$$\begin{aligned} \psi(\mathbf{r}) &= \sum_{\mathbf{G}} c_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}} \\ -\frac{\hbar^2}{2m} \nabla^2 \psi_j &= -\frac{\hbar^2}{2m} \nabla^2 \left( \sum_{\mathbf{G}} c_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}} \right) \\ &= -\frac{\hbar^2}{2m} \sum_{\mathbf{G}} c_{\mathbf{G}} \nabla^2 e^{i\mathbf{G} \cdot \mathbf{r}} \\ &= \sum_{\mathbf{G}} \frac{\hbar^2}{2m} c_{\mathbf{G}} |\mathbf{G}|^2 e^{i\mathbf{G} \cdot \mathbf{r}} \end{aligned}$$

i.e. diagonal in  $\mathbf{G}$ .





## Potential energy

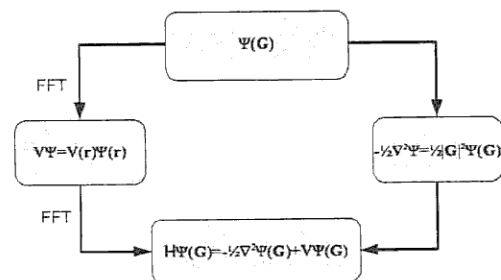
$$V[\rho](\mathbf{r})\psi_j(\mathbf{r})$$

i.e. diagonal in  $\mathbf{r}$ .

- We don't have  $\psi_j(\mathbf{r})$  directly, just  $c_G$
- Fourier transform
- Multiply by the potential
- Fourier transform back



## Applying the Hamiltonian



## Iterative Diagonalisation

- The groundstate energy  $E_0$  is the lowest possible energy
- Any wavefunction has energy  $E \geq E_0$ 
  - guess a wavefunction  $\psi(c_G)$
  - compute  $E = \psi^\dagger H \psi$
  - tweak  $c_G$  to lower  $E$
- When we can't lower  $E$  any more,  $\psi$  is the groundstate!



## The Gradient

What's the best way to fiddle with the Fourier coefficients?

We can compute the change in  $c_G$  that *raises* the energy quickest by differentiating:

$$\begin{aligned} \epsilon_j &= \frac{\psi_j^\dagger H \psi_j}{\psi_j^\dagger \psi_j} \\ \Rightarrow \frac{\delta \epsilon_j}{\delta \psi_j^\dagger} &= \frac{(\psi_j^\dagger \psi_j) H \psi_j - (\psi_j^\dagger H \psi_j) \psi_j}{(\psi_j^\dagger \psi_j)^2} \\ &= H \psi_j - \epsilon_j \psi_j \end{aligned}$$

where on the last line we've assumed  $\psi_j$  is normalised.



## Steepest Descent Diagonalisation

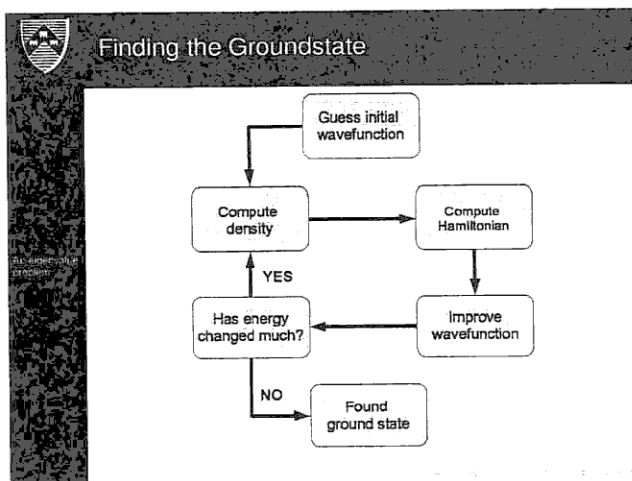
- Starting with  $\psi_j$ , compute  $\epsilon_j$
- Compute the gradient  $\frac{\delta \epsilon_j}{\delta \psi_j^\dagger}$
- This is the change to  $c_G$  that *increases*  $\epsilon_j$  quickest
- We want to *decrease*  $\epsilon_j$ , so use  $-\frac{\delta \epsilon_j}{\delta \psi_j^\dagger}$ .
- Make a new guess eigenstate,  $\psi_j^{\text{new}} = \psi_j - \lambda \frac{\delta \epsilon_j}{\delta \psi_j^\dagger}$
- Vary  $\lambda$  until we've found the lowest  $\epsilon_j$  in this direction. This procedure is often called a *line search*.

There are better methods than steepest descent, e.g. conjugate gradients.



## Iterative Diagonalisation

- Only need to apply  $H$
- Never need to store  $H$  explicitly
- Terms of  $H$  are diagonal in either  $\mathbf{G}$  or  $\mathbf{r}$
- Don't need  $\text{num\_pw} \times \text{num\_pw}$  matrix
- Only need a few  $\text{num\_pw}$ -long vectors



### Multiple Eigenstates

- Can find lowest eigenstate  $\psi_1$
- What about next lowest  $\psi_2$ ?

### Multiple Eigenstates

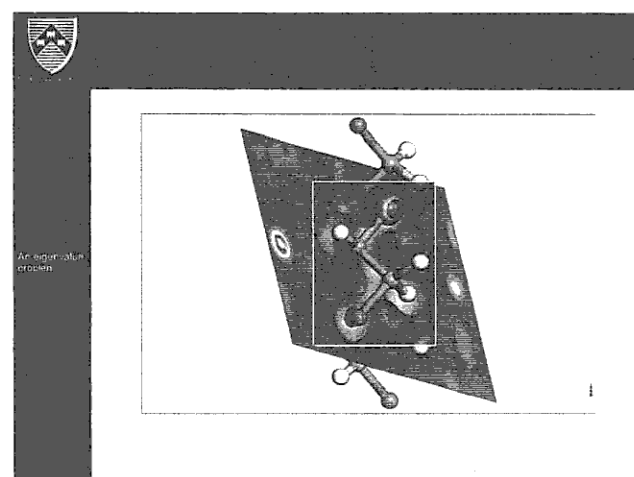
- Can find lowest eigenstate  $\psi_1$
- What about next lowest  $\psi_2$ ?
- If we just repeat method, we'll find lowest again!

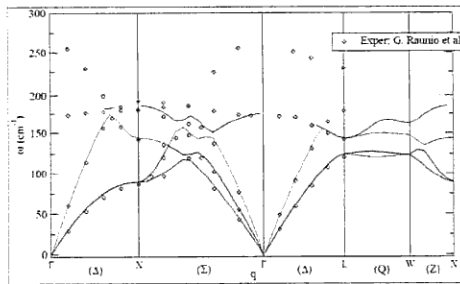
### Multiple Eigenstates

- Can find lowest eigenstate  $\psi_1$
- What about next lowest  $\psi_2$ ?
- If we just repeat method, we'll find lowest again!
- Need to explicitly orthogonalise so  $\psi_2^\dagger \psi_1 = 0$

### Fame and fortune

- Now have all the tools we need
- Can find solve DFT eq. for lowest  $N$  eigenstates  $\rightarrow$  get energy and electron density
- Can now predict:
  - Crystal and molecular structure
  - Bond breaking and formation
  - IR spectra
  - Reactivities
  - Colour, X-ray absorption, NMR spectra
  - ... and much more!





AlN superlattice  
phonon dispersion



## Summary

- Predictive materials simulations need QM calculations
- DFT reduces QM to a large, simple eigenvalue problem
- Iterative diagonalisation can solve large eigenvalue problems quickly for a few states
- Combining these two we can predict a vast range of properties for all kinds of materials and chemicals