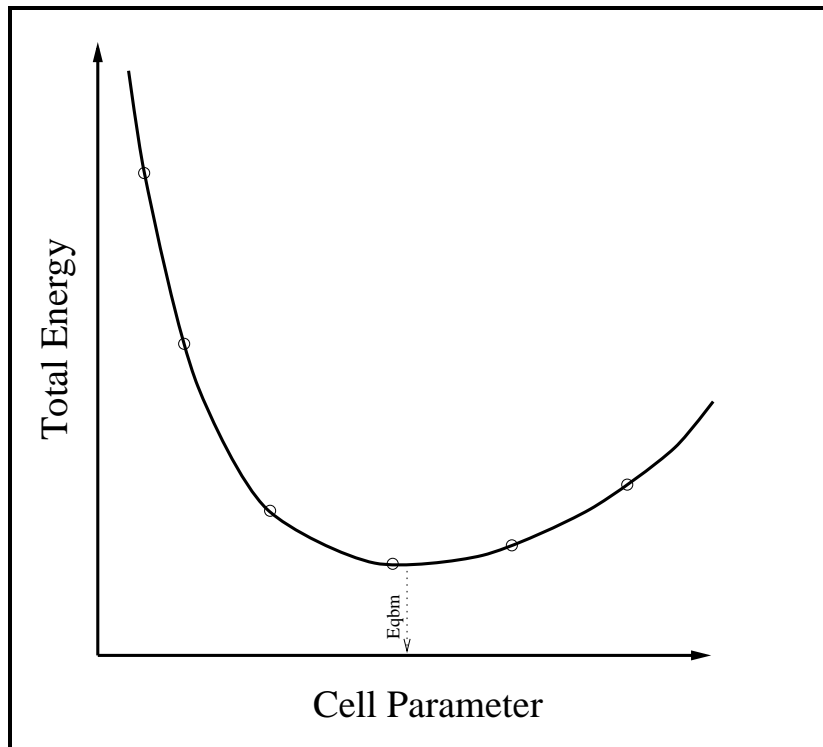


FORCES, STRESSES AND STRUCTURAL OPTIMISATIONS

Chris J. Pickard

Lecture Four

TOTAL ENERGY CALCULATIONS



- Many properties depend on the **total energy** of a system
 - equilibrium lattice constants (density)
 - bulk moduli
 - phonons
 - elastic constants
 - phase transitions
 - chemistry, bonding etc.

DERIVATIVES OF TOTAL ENERGIES

- Many properties depend more on the derivatives of the energies, than the total energies themselves
- This can be computationally costly, and is susceptible to numerical noise
- We could get the derivatives by calculating the total energy at several points around each point, and do a numerical derivative
- Another approach is to use perturbation theory – in the form of the Hellman-Feynman theorem

STRESS AND STRAIN

- The concept of forces is straightforward, but you can also take derivatives with respect to the crystal unit cell

$$\mathbf{h}' = (\mathbf{I} + \boldsymbol{\epsilon})\mathbf{h}$$

- The application of a *strain* changes the shape of the unit cell

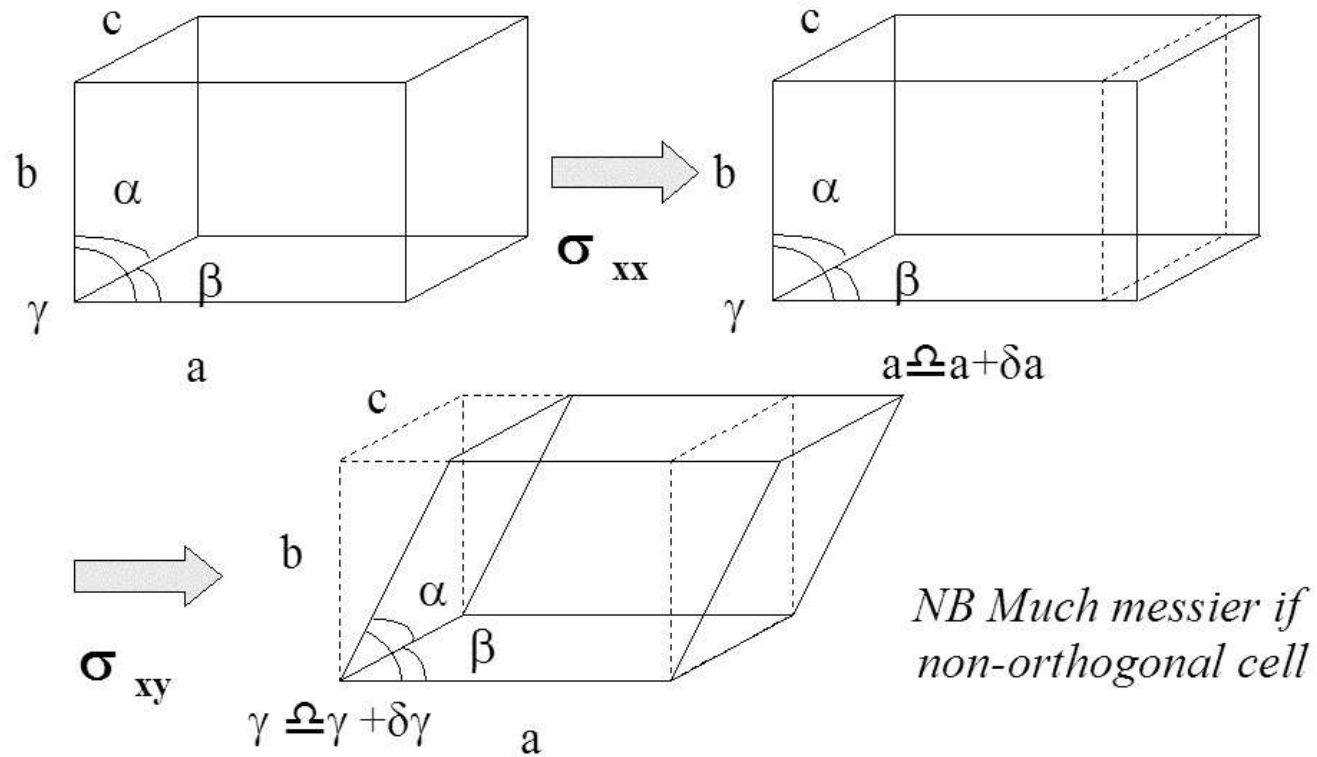
- The *stress* tensor $\boldsymbol{\sigma}$ is related to the strain tensor $\boldsymbol{\epsilon}$:

$$\sigma_{\alpha\beta} = \frac{1}{\Omega} \frac{\partial E}{\partial \epsilon_{\alpha\beta}}$$

- If we write the three unit cell vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ as columns of a matrix \mathbf{h} the shape change is described by:

where $\Omega = \mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ is the volume of the unit cell

STRESS AND STRAIN



THE HELLMAN-FEYNMAN THEOREM

- Classically, the force \mathbf{F} acting on a particle at \mathbf{R} is given by the derivative of the potential energy:
- The Hellman-Feynman Theorem:

$$\mathbf{F} = -\nabla_{\mathbf{R}} U(\mathbf{R})$$

- We might expect the quantum mechanical equivalent to be:

$$\mathbf{F} = -\nabla_{\mathbf{R}} \langle E \rangle \text{ where}$$

$$\langle E \rangle = \langle \Psi | H | \Psi \rangle, \text{ if } |\Psi\rangle \text{ are normalised}$$

$$\frac{\partial E}{\partial \lambda} = \langle \frac{\partial \Psi}{\partial \lambda} | H | \Psi \rangle + \langle \Psi | \frac{\partial H}{\partial \lambda} | \Psi \rangle + \langle \Psi | H | \frac{\partial \Psi}{\partial \lambda} \rangle$$

$$H|\Psi\rangle = E|\Psi\rangle \text{ and so}$$

$$\frac{\partial E}{\partial \lambda} = \langle \Psi | \frac{\partial H}{\partial \lambda} | \Psi \rangle + E \frac{\partial}{\partial \lambda} \langle \Psi | \Psi \rangle$$

$$\boxed{\frac{\partial E}{\partial \lambda} = \langle \Psi | \frac{\partial H}{\partial \lambda} | \Psi \rangle}$$

THE HELLMAN-FEYNMAN THEOREM

- We can calculate the forces (or stresses) from the ground state wavefunctions (which we have to calculate anyway for the total energy)
- If our basis set depends on the ionic positions, such as localised atomic orbitals or gaussians, there are other derivatives that will contribute: these are called *Pulay forces/stresses*
- These Pulay forces/stresses disappear in the limit of a complete basis (never achieved) or with basis functions which are position independent (such as plane waves)

THE HELLMAN-FEYNMAN THEOREM AND DENSITY FUNCTIONAL THEORY

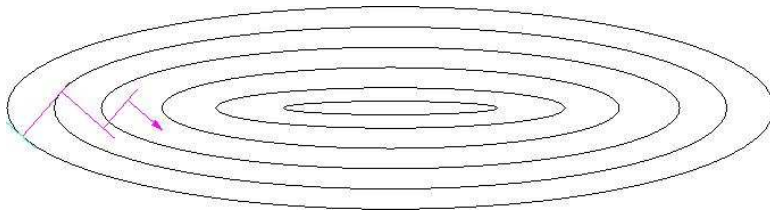
$$H_{\mathbf{R}} = -\frac{1}{2}\nabla_{\mathbf{r}}^2 + V_{\text{e-e}}(\mathbf{r}) + V_{\text{ion-e}}(\mathbf{r}, \mathbf{R}) + V_{\text{xc}}(\mathbf{r}) + V_{\text{ion-ion}}(\mathbf{R})$$

- The DFT Hamiltonian is parameterised in \mathbf{R}
- For the stresses, there are contributions from the kinetic energy and the Hartree terms
- We get a contribution from the ion-electron (pseudo)potential
- We get a contribution from the ion-ion Coulomb interaction (from the Ewald sum)
- While the total energy is correct to second order in the errors, the forces are only good to first order

STRUCTURAL OPTIMISATION

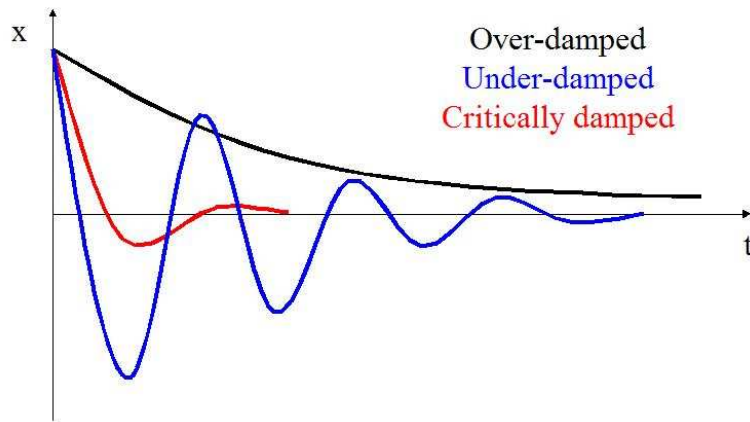
- Assume that electrons adapt a classic optimisation problem instantaneously to the ionic configuration
 - this is the Born-Oppenheimer approximation
- The ionic positions can be considered to be parameters, the total energy is a function of the ionic coordinates
- Finding the ground-state structures is
- The *phase space* grows with the system size
- There are many local minima (corresponding to meta-stable structures)
- In general, we must start with a good guess for the structure

STEEPEST DESCENTS



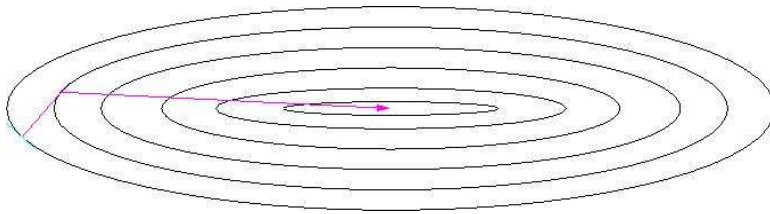
- This is the simplest approach:
 - take a downhill step along the local steepest gradient, and a trial step length
 - use line minimisation to find the optimal step length
- Advantages:
 - simple to implement, and robust
 - reliable – will find the minima eventually
- Disadvantages:
 - very slow to converge
 - can get stuck in a local minima

DAMPED MOLECULAR DYNAMICS



- This improves on steepest descents:
 - use velocities as well as forces
 - start with $\mathbf{v} = \mathbf{0}$ and add damping term to forces $-\gamma\mathbf{v}$
 - adjust γ and time step to obtain optimal convergence
- Advantages:
 - simple to implement, robust and more efficient than SD
 - can use wavefunction extrapolation (or Car-Parinello)
- Disadvantages:
 - convergence rate depends on damping factor γ
 - can get stuck in a local minima

CONJUGATE GRADIENTS



- This improves on steepest descents:
 - the gradient is constructed to be conjugate to all previous directions
 - does not undo previous minimisation
 - a line minimisation is performed
- Advantages:
 - rapid convergence – in a quadratic energy landscape, one dof per iteration
 - low storage requirements
- Disadvantages:
 - more complex to implement than SD
 - Hessian not explicitly calculated
 - can get stuck in a local minima

BROYDEN-FLETCHER-GOLDFARD-SHANNO – BFGS

$$\mathbf{A} = \begin{pmatrix} \frac{\partial^2 E}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 E}{\partial x_1 \partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial x_N \partial x_1} & \cdots & \frac{\partial^2 E}{\partial x_N \partial x_N} \end{pmatrix}$$

$$\delta E = \frac{1}{2}(\mathbf{x} - \mathbf{x}_{\min})^T \cdot \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}_{\min})$$

- If we know the Hessian \mathbf{A} we can move from nearby the minimum straight to it

- we don't know it, so we build up a guess using the BFGS algorithm

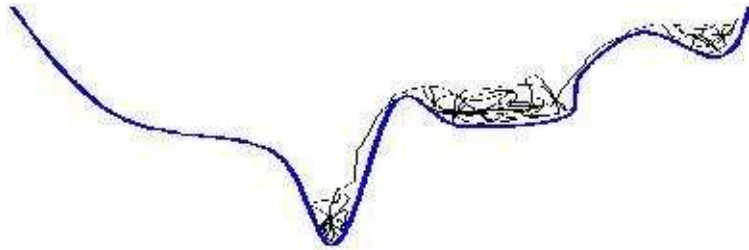
- Advantages:

- convergence similar or better than CG
- extra physical information is contained in the Hessian

- Disadvantages:

- complex to code
- Hessian must be stored ($\# \text{ dof}^2$)
- can get stuck in a local minima

SIMULATED ANNEALING



- Advantages:
 - very robust and reliable
 - reasonably immune to getting stuck in local minima
- Disadvantages:
 - incredibly slow convergence
 - the cooling rate must be carefully adjusted to avoid quenching into local minima
 - no guarantee that the true global minima will be found
- This is a stochastic method:
 - always accept steps that lower the energy, and sometimes accept upward steps, using a Boltzman distribution
 - slowly reduce temperature, and iterate to the goundstate

SYMMETRY AND CONSTRAINTS

- If the symmetry of the system is known, it can be enforced (if desired) applied
- The symmetry operations can be calculated, and the calculated forces can be *symmetrised*
- The simplest are *linear* constraints, e.g. atom N must only move in the Z direction
- Relaxing the structure using these forces results in a system with the desired symmetry
- More difficult are *non-linear* constraints such as enforcing distances between atoms
- Other kinds of *constraints* can be
- External pressures can also be applied

MOLECULAR DYNAMICS

- Can do classical dynamics of ions at various temperatures using the forces/stresses derived from the first principles total energy calculations
- Can generate thermodynamic information from ensemble averaging
- Can cope with bond-breaking, chemical reactions, catalysis, diffusion and so on
- Gives access to time dependent phenomena
- Temperature driven phase transformation can be followed
- Incorporates the effects of finite size