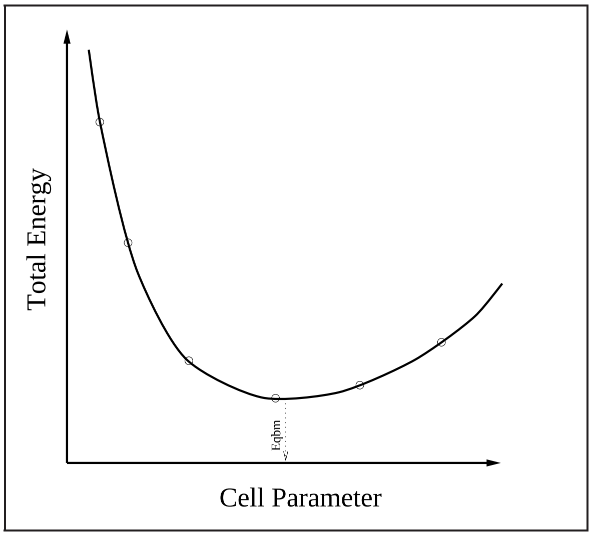
Forces, Stresses and Stuctural Optimisations

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
- We could get the derivatives by calculating the total energy at several points around each point, and do a numerical derivative
- This can be computationally costly, and is susceptible to numerical noise
- 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
- The application of a *strain* changes the shape of the unit cell
- 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:

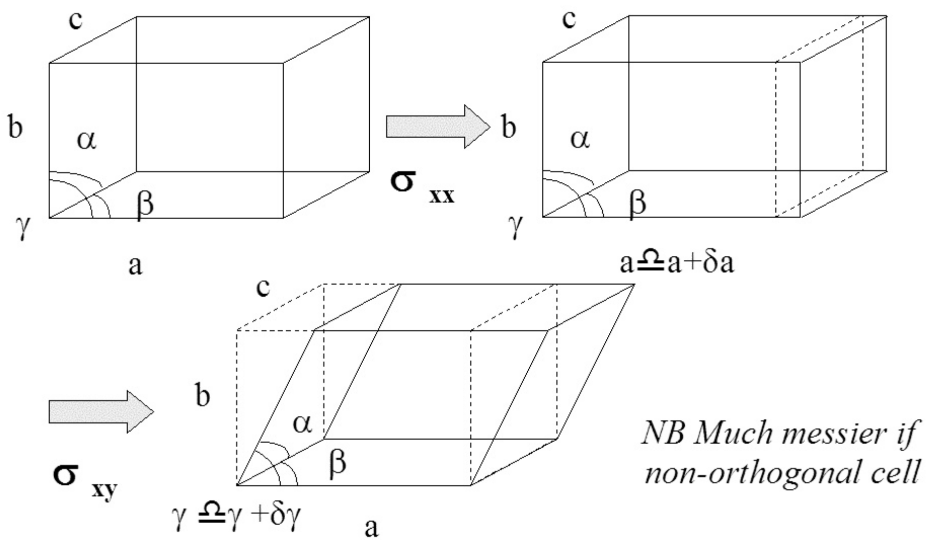
$$\mathbf{h}' = (\mathbf{I} + \boldsymbol{\epsilon})\mathbf{h}$$

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

$$\boldsymbol{\sigma}_{\alpha\beta} = \frac{1}{\Omega}\frac{\partial E}{\partial \epsilon_{\alpha\beta}}$$

where $\boldsymbol{\Omega} = \mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ 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:

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

- We might expect the quantum mechanical equivalent to be:

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

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

- The Hellman-Feynman Theorem:

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

  $H|\Psi\rangle = E|\Psi\rangle$ 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 a 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 acheived) or with basis functions which are position independent (such as plane waves)

The Hellman-Feynman Theorem and Density Functional Theory

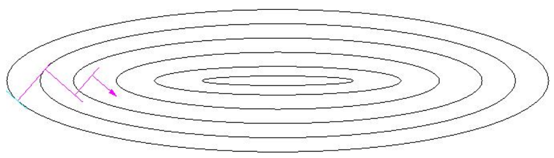
$$\boxed{H_{\mathbf{R}} = -\frac{1}{2}\nabla_{\mathbf{r}}^2 + V_{\mathrm{e-e}}(\mathbf{r}) + V_{\mathrm{ion-e}}(\mathbf{r},\mathbf{R}) + V_{\mathrm{xc}}(\mathbf{r}) + V_{\mathrm{ion-ion}}(\mathbf{R})}$$

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

Structural Optimisation

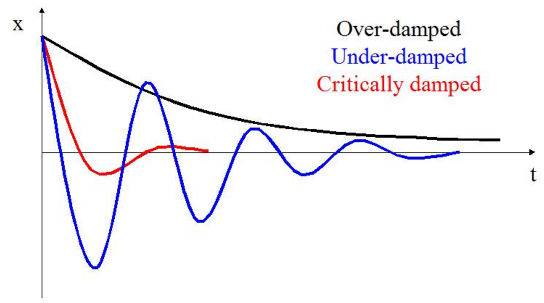
- Assume that electrons adapt instantaneously to the ionic configuration
  - this is the Born-Oppenheimer approximation
- The ionic positions can be considered to be parameters, the total energy is a function of the ionic coordinates
- Finding the ground-state structures is a classic optimisation problem
- 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

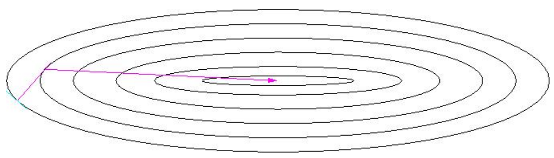


- This improves on steepest descents:
  - use velocities as well as forces
  - start with $\mathbf{v} = \mathbf{0}$ and add damping term to forces $-\gamma\mathbf{v}$
  - adjust γ and time step to obtain optimal convergence

- 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

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

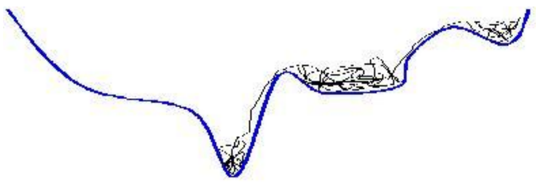
$$\delta E = \tfrac{1}{2}(\mathbf{x} - \mathbf{x}_{\text{min}})^T \cdot \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}_{\text{min}})$$

- 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 (# dof^2)
  - can get stuck in a local minima

Simulated Annealing



- 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

- 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

Symmetry and Constraints

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

Molecular dynamics

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