Geometry Optimisation

Matt Probert
Condensed Matter Dynamics Group
Department of Physics,
University of York, UK

http://www.cmt.york.ac.uk/cmd
http://www.castep.org
Overview of Talk

• Motivation

• Background Theory

• CASTEP details

• Example
Why bother?

• Want to find ground state of system
  – Minimise energy of electrons at fixed ionic positions
  – Then optimise the ionic positions and/or the unit cell shape and size
    • particularly if external pressure applied
  – Theoretical minimum often not exactly the same as experiment!
    • depends on choice of pseudopotential, plane-wave cut-off energy, choice of XC functional, etc.
  – Target agreement should be better than 1%
What Can It Tell You?

- Equilibrium bond lengths and angles
- Equilibrium cell parameters
- Discriminating between competing structures
- Elastic constants
- Surface reconstructions
- Pressure-driven phase transitions
- Starting point for many advanced investigations …
Background Theory
Forces and Stresses

• Minimum energy configuration corresponds to zero force
  – Using forces to find the minimum is much more efficient than just using energy alone

• Minimum enthalpy corresponds to zero force and stress
  – Can therefore minimise enthalpy w.r.t. supercell shape due to internal stress and external pressure
  – Pressure-driven phase transitions
QM Forces

• Classically we have the force $\mathbf{F}$ at position $\mathbf{R}$ is determined from the potential energy as

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

• Quantum mechanically we therefore expect

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

where

$$\langle E \rangle = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$
QM Stresses

• If we write the three unit cell vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ as the columns of a matrix $\mathbf{h}$ then the effect of an applied strain $\varepsilon$ is to change the shape of the unit cell:

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

• We then have the stress tensor $\sigma$ related to the strain tensor $\varepsilon$ by:

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

where $\Omega = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is the volume of the unit cell.
Stress and strain in action

\[ \sigma_{xx} \]

\[ \sigma_{xy} \]

NB Much messier if non-orthogonal cell
Hellman-Feynman Theorem (I)

• For any perturbation $\lambda$ we have

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

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

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

which obviously includes forces and stresses.

• We have assumed that the wavefunction is properly normalised and is an exact eigenstate of $H$. 
Hellman-Feynman Theorem (II)

- If we have an approximate eigenstate $\Psi$, for example from using an incomplete basis set, then we must keep all 3 terms in the general expression.

- If our basis set depends upon the ionic positions, such as atomic centred Gaussians, then the other derivatives in the general expression will contribute so-called Pulay forces (stresses).

- Note that Pulay forces (stresses) will vanish in the limit of a complete basis set (but this is never realized in practice), or if position independent basis functions, such as plane-waves, are used.
Density Functional Theory (I)

• In DFT we have the Kohn-Sham Hamiltonian:

\[
\hat{H}(\mathbf{r}, \mathbf{R}) = -\frac{1}{2} \nabla_r^2 + V_{e-e}(\mathbf{r}) + V_{\text{ion-e}}(\mathbf{r}, \mathbf{R}) + V_{\text{XC}}(\mathbf{r}) + V_{\text{ion-ion}}(\mathbf{R})
\]

• Therefore we only get contributions to the forces from the electron-ion (pseudo)potential and the ion-ion Coulomb interaction (the Ewald sum).

• Also contribution from exchange-correlation potential if using non-linear core corrections.

• However, for the stresses, we also get a contribution from the kinetic energy and Hartree terms.
Density Functional Theory (II)

• As we do not have a complete basis, the wavefunction will not be exact even within DFT.

• If we use a variational method to minimize the total energy, then we know that the energy and hence the wavefunction will be correct to second order errors.

• However, the forces will only be correct to first order → need a larger basis set for accurate forces than for energies.
Density Functional Theory (III)

• However, if we use a non-variational minimisation technique, such as *density mixing*, then such statements cannot be made.

• We can no longer guarantee that the energy found is an upper-bound on the true ground state energy.

• This complicates the application of the Hellman-Feynman theorem.

• Consequently, non-variational forces and stresses are less reliable.
Force Convergence using Different Minimizers

unrelaxed Si7, USP, MP=2

|F| (eV/Ang) vs. Ecut (eV)

Legend:
- F1 (AB)
- F2 (AB)
- F3 (AB)
- F1 (DM)
- F2 (DM)
- F3 (DM)

RMS F vs. Ecut

Ecut range: 50 to 250 eV
How Does Geometry Optimisation Work?

- Electrons adjust instantly to position of ions so have a multi-dimensional potential energy surface and want to find global minimum
  - Treat as an optimisation problem
  - Simplest approach is steepest descents
  - More physical approach is damped MD
  - More sophisticated approaches are conjugate gradients or BFGS
  - All of these can get stuck in local minima
  - Hence recent research in Genetic Algorithms to find global minimum (Luke Abraham poster)
Steepest Descents

Enlargement of a single step showing the line minimisation in action – the step continues until the local energy starts to rise again whereupon a new direction is selected which is orthogonal to the previous one.
Damped MD

- Move ions using velocities and forces
- Need to add damping term ‘-\(\gamma v\)’ to forces
- Algorithm to set ‘optimal damping’ factor
- Can also adjust \(\delta t\) for max. convergence
- More efficient than Steepest Descents
Conjugate Gradients

- The initial search direction is given by steepest descents.
- Subsequent search directions are constructed to be orthogonal to the previous and all prior search directions.
- No explicit Hessian required or generated.
BFGS (I)

• Basic idea:
  – Energy surface around a minima is quadratic in small displacements and so is totally determined by the Hessian matrix $A$ (the matrix of second derivatives of the energy):

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

  $\delta E = \frac{1}{2} (X - X_{\text{min}})^T \cdot A \cdot (X - X_{\text{min}})$

  – so if we knew $A$ then could move from any nearby point to the minimum in 1 step!
BFGS (II)

• The Problem
  – We do not know $A$ a priori
  – Therefore we build up a progressively improving approximation to $A$ (or the inverse Hessian $H=A^{-1}$) as we move the ions according to the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm.
  – Also known as a quasi-Newton method.

• Positions updated according to:
  $$X_{i+1} = X_i + \lambda \Delta X_i$$
  $$\Delta X_i = H_i F_i$$
CASTEP details
How to do it …

• Just put
  \texttt{task = Geometry Optimisation}
  
in your \texttt{.param} file
• Is that all?
• What is going on behind the scenes?
• What might go wrong?
• What can you control?
Choice of Scheme

• `geom_method = BFGS` (default)
  – variable ions and/or cell
  – uses fractional ionic coordinates and strains as basic variables
  – use `geom_frequency_est` for improved initialisation of ionic part of $H$ (default=50 THz)
  – use `geom_modulus_est` for improved initialisation of cell part of $H$ (default=500 GPa)
  – improved estimates based analysis of approx. $H$ built up over convergence path printed at end
Alternative Fixed-Cell Schemes

• **geom_method = DELOCALISED**
  - uses delocalised internal coordinates in BFGS
  - hence can use non-linear constraints such as fixed bond lengths/angles or torsion angles

• **geom_method = DAMPED MD**
  - uses Cartesian coordinates
    - `md_damping_scheme` = independent/coupled/steepest descents
  - damping parameters recalculated every `md_damping_reset` DMD steps (default=30)
  - can use estimated optimal DMD timestep if `md_opt_damped_delta_t = true` (default=false)
BFGS in CASTEP

\[ \mathbf{X}_{i+1} = \mathbf{X}_i + \lambda \Delta \mathbf{X}_i \]
\[ \Delta \mathbf{X}_i = \mathbf{H}_i \mathbf{F}_i \]

- Trial step always performed
  - A quadratic system has optimal \( \lambda = 1 \)
  - Exact line minimisation has \( \mathbf{F} \cdot \Delta \mathbf{X} = 0 \)

- Line step
  - only performed if predicted \( |\lambda_{\text{trial}} - \lambda_{\text{line}}| > \text{geom}_\text{linmin_tol} \)

No restriction on the value of \( \lambda \) as long as does not cause too large an ionic displacement or change in lattice parameters
Geometry Convergence

• Will continue for \texttt{geom\_max\_iter} steps (default=30) unless converged:
  \texttt{geom\_energy\_tol} (default=2\times10^5 \text{ eV/atom}) for \texttt{geom\_convergence\_win} steps

Changing ions:
  \texttt{geom\_force\_tol} (default =0.05 \text{ eV/Å})
  \texttt{geom\_disp\_tol} (default=0.001 \text{ Å})

Changing cell:
  \texttt{geom\_stress\_tol} (default=0.1 \text{ GPa})
Quality of Forces/Stresses

• Can only find a good structure if have reliable forces/stresses

• Forces converge more slowly than energies as increase number of plane waves
  – CONVERGENCE!

• Should also check degree of SCF convergence
  \texttt{elec\_energy\_tol} (default=$10^{-5}$ eV/atom)
  \texttt{elec\_force\_tol} (off by default, might want on)

• Stresses converge slower than forces!
Finite Basis Set Correction

• Change unit cell at constant cut_off_energy :
  – changes number of plane-waves
  – change in total energy due to variational principle
  – hence difficult to compare results at different cell sizes.

• Use the Finite Basis Set Correction
  – finite_basis_corr = automatic/manual/none
  – calculates the total energy finite_basis_npoints times with finite_basis_spacing change to cut_off_energy at fixed cell
  – hence calculates and prints basis_de_dloge
  – CASTEP can then use this to correct the total energy and stress at nearby cell sizes
Finite–Basis Set Correction

Perfect Si8, USP, Ecut=120 eV, MP=2

-866.4
-866.6
-866.8
-867
-867.2
-867.4

Lattice parameter a0 (Ang)

-866.4
-866.6
-866.8
-867
-867.2
-867.4

E_tot (eV)

Uncorrected
Corrected
Changing Cell and Plane Waves

• Plane wave basis is independent of ionic positions but NOT of the unit cell size/shape

• **fix_NPW=true**
  – the number of plane waves is fixed
  – hence the variational principle applies and we can search for the minimum of enthalpy as cell changes
  – but the effective cut-off energy varies!

• **fix_NPW=false**
  – number of plane waves varies to keep Ecut constant
  – breaks variational principle
  – search for zero force & stress not minimum enthalpy
.cell file keywords

fix_all_ions (default = false)
fix_all_cell (default = false)
fix_com (default = NOT fix_all_ions)
symmetry_generate

%block external_pressure
  [units]
  P_{xx}  P_{xy}  P_{xz}
  P_{yy}  P_{yz}
  P_{zz}
%endblock external_pressure

hence \( P_{xx} = P_{yy} = P_{zz} = P \) and \( P_{xy} = P_{xz} = P_{yz} = 0 \) is hydrostatic \( P \)
Cell Constraints

%block cell_constraints
|a| |b| |c|
α β γ
%endblock cell_constraints

• Any length (angle) can be held constant (=0), or tied to one or both of the others (=same)
  – e.g. cell optimisation of 2D structures or keeping subset of symmetry etc
  – also fix_vol (=false by default)
Simple Hexagonal cell with symmetry on.

CASTEP constrained-cell optimisation with fix_vol=true

c/a values in good agreement with expt.
Miscellaneous .param keywords

- **geom_spin_fix** (default=0)
  >0 fix spin for this number of steps, <0 always fix

- **num_backup_iter** (default=5 steps) or **backup_interval** (default=0 seconds)
  – interval to write checkpoint file

- **continuation = default**
  – perform a restart from default checkpoint file

- **opt_strategy_bias** (-3 … +3, default=0)
  >0 bias code towards speed
  <0 bias code towards reduced memory
The .geom file

• Records the final configuration after each step:

  10
  -7.93316351E+000 -7.85316331E+000 <-- E
  0.00000000E+000  5.13126785E+000  5.13126785E+000 <-- h
  5.13126785E+000  0.00000000E+000  5.13126785E+000 <-- h
  5.13126785E+000  5.13126785E+000 0.00000000E+000 <-- h
  -3.56997760E-003  0.00000000E+000 -3.33783917E-013 <-- S
  0.00000000E+000 -3.56997760E-003  8.32597229E-013 <-- S
  -3.33783917E-013  8.32597229E-013  5.93008591E-004 <-- S
  Si 1  0.00000000E+000  0.00000000E+000  0.00000000E+000 <-- R
  Si 2  7.56877069E+000  2.52292356E+000  7.56877069E+000 <-- R
  Si 1 -5.22300739E-003  6.43530285E-003 -1.71774942E-003 <-- F
  Si 2  5.22300739E-003 -6.43530285E-003  1.71774942E-003 <-- F

– Uses Cartesian coordinates and atomic units throughout
– Designed for other analysis programs, visualisation tools, etc.
Ionic Constraints

• Can specify any arbitrary number of *linear constraints* on the atomic coordinates, up to the number of degrees of freedom.
  – E.g. fixing an atom, constraining an atom to move in a line or plane, fixing the relative positions of pairs of atoms, fixing the centre of mass of the system, etc.

• Non-linear constraints
  – Much more difficult to apply and specify in general
  – Delocalised Internals will gain some non-linear constraints in CASTEP v4.1
Specifying Linear Constraints

%block ionic_constraints
con spec atom Ax Ay Az

where each constraint fixes 1 degree of freedom
- con is the number of the constraint (can have multiple atoms/constraint) and multiple constraints/calculation
- spec is the species label, atom is the number of the atom for this species
- Ax, Ay, Az specify the constraint coefficient such that
\[ \sum A_r = const \]
Example Linear Constraints

```plaintext
%block ionic_constraints
1 Si 1 0 0 1
%endblock ionic_constraints
fixes the z-coordinate of the 1st Si atom

%block ionic_constraints
1 Si 1 1 0 0
1 Si 2 -1 0 0
%endblock ionic_constraints
fixes the relative x-coord of 1st and 2nd Si atoms, etc.
```
Example:
Si(100) Surface Reconstruction
Si(100) supercell with vacuum gap

... with added hydrogen passivation

9 layers of Silicon and 7 Å vacuum
Convergence Tests

• Converge cut-off energy $\rightarrow$ 370 eV
• Converge k-point sampling $\rightarrow$ 9 k-points
• Converge number of bulk layers $\rightarrow$ 9 layers
• Converge vacuum gap $\rightarrow$ 9 Å
• only then see asymmetric dimerisation:

$()$=best lit. value
Si(100) – The Movie
Si(100) – BFGS Progress

- Initial slow decrease in energy due to surface layer compression.
- Then small barrier to dimer formation overcome around iteration 14.
- Then rapid energy drop due to dimerisation.
- Final barrier to asymmetric dimerisation overcome around iteration 24.
Conclusions

• Need accurate forces and stresses
• Pre-requisite for many property calculations
  – e.g. phonons, NMR, etc.
• Can do optimisation of ions and/or cell
  – BFGS with Cartesian coordinates
• Can also for fixed cell use
  – BFGS with delocalised internal coordinates, or
  – damped MD with Cartesian coordinates