

Forces, Stress, and Structural Optimization

Pietro Delugas (SISSA)

Outline



-
- Structural parameters
 - Forces
 - Stress
 - Optimization methods



Structural parameters

- A *system* is defined by:
 - Number and species (pseudopotentials) of ions.
 - The spatial relation between the ions. The *structure*
- The *structure* of a periodic infinite system is defined by:
 - the lattice parameters: $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$
 - the atomic positions: \mathbf{r}_I
- To keep the structural parameters distinguished we use fractional *crystal* coordinates:

$$\mathbf{r}_I = \sum_{i=1,2,3} x_i \mathbf{a}_i$$

- In the Born Oppenheimer approximation the potential energy $E(\mathbf{x}_I, \mathbf{a}_i)$ as function of these parameters determines the potential energy *surface* for our system.

Forces



The total force on atoms is given by (minus) the derivatives of the electronic and nuclei-nuclei interaction energy.

$$\mathbf{F}_I = -\mathbf{F}_I^{el} - \frac{\partial}{\partial \mathbf{r}_I} \left(\frac{e^2}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{r}_I - \mathbf{r}_J|} \right)$$

The *Hellman-Feynman* theorem applies also in DFT so the *electronic* forces can be computed as expectation values of derivatives of the external potential.

$$\mathbf{F}_I^{el} = -\frac{\partial E}{\partial \mathbf{r}_I} = -\sum_{\nu} f_{\nu} \left\langle \psi_{\nu} \left| \frac{\partial V}{\partial \mathbf{r}_I} \right| \psi_{\nu} \right\rangle = -\int n(\mathbf{r}) \frac{\partial V}{\partial \mathbf{r}_I} d\mathbf{r}$$



The force theorem

The Hellmann-Feynman theorem holds also in DFT:

- Energy is variational w.r.t. density:

$$\int \frac{\delta E}{\delta n(\mathbf{r})} \delta n(\mathbf{r}) d\mathbf{r} = 0 \implies \frac{\delta E}{\delta n(\mathbf{r})} = \mu$$

- The contribution due to $\frac{\partial n(\mathbf{r})}{\partial \mathbf{r}_I}$ vanishes

$$\frac{\partial E}{\partial \mathbf{r}_I} = \int n(\mathbf{r}) \frac{\partial V}{\partial \mathbf{r}_I} d\mathbf{r} + \int \frac{\delta E}{\delta n(\mathbf{r})} \frac{\partial n(\mathbf{r})}{\partial \mathbf{r}_I} d\mathbf{r}$$

SCF convergence and accuracy of forces



- Forces are accurate only when we are close to self-consistency.
- `pw.x` prints an estimate of the *red* term at the end of self-consistency:
 - It is an approximate correction (Chan *et al.*: Phys. Rev. B **47**,4771–1993)
 - $\frac{\partial n(\mathbf{r})}{\partial \mathbf{r}_I}$ is estimated by the derivatives of the superposition of the atomic charge densities.
 - $\frac{\delta E}{\delta n(\mathbf{r})}$ is estimated as the the difference of $V_{scf}(n(\mathbf{r}))$ at the final iteration.
- During relaxation or molecular dynamic symulations it is always important to monitor that the value of this estimate remains some order of magnitude smaller than the computed forces.



Strain and Stress



- The change of the lattice vectors implies a uniform deformation of the system.
- Deformations are described by the *strain* tensor $e_{\alpha\beta}$

$$r'_\alpha = \sum_{\beta} (\delta_{\alpha\beta} + e_{\alpha\beta}) r_\beta$$

- The stress tensor is defined as the derivative of the Energy per unit volume w.r.t. strain:

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

Note: Only the symmetric part of $e_{\alpha\beta}$ is an effective deformation that affects the energy, the antisymmetric part being a solid rotation

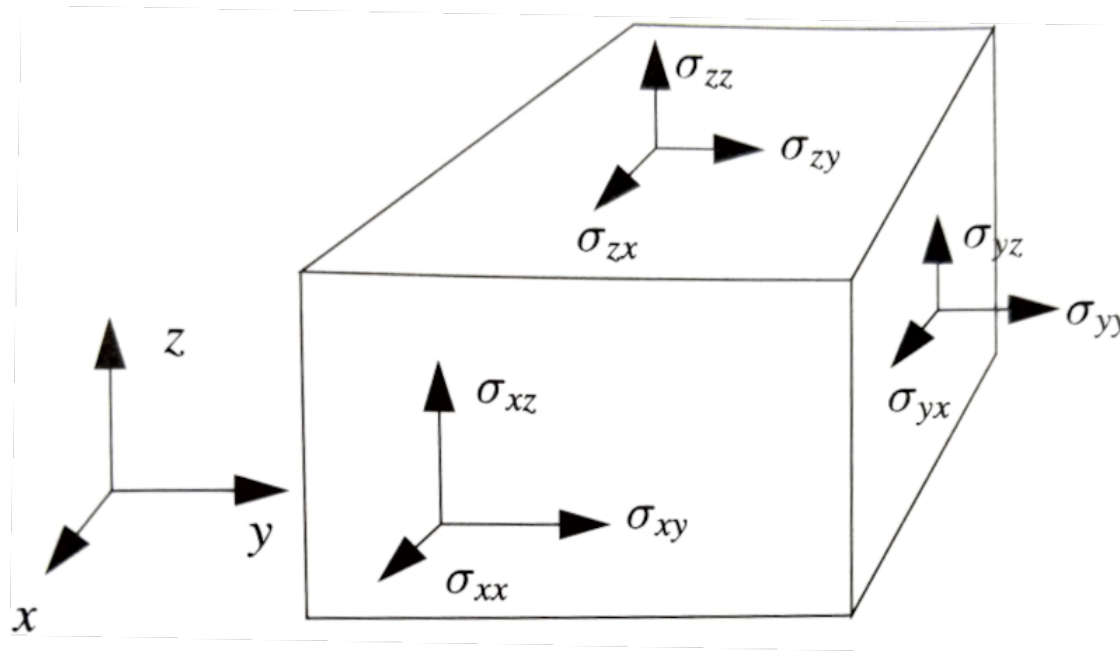


Stress

- Diagonal terms are related to pressure

$$P = \frac{1}{3} \sum_{\alpha} \sigma_{\alpha\alpha}$$

- Non diagonal components are the shear stress



Stress Theorem

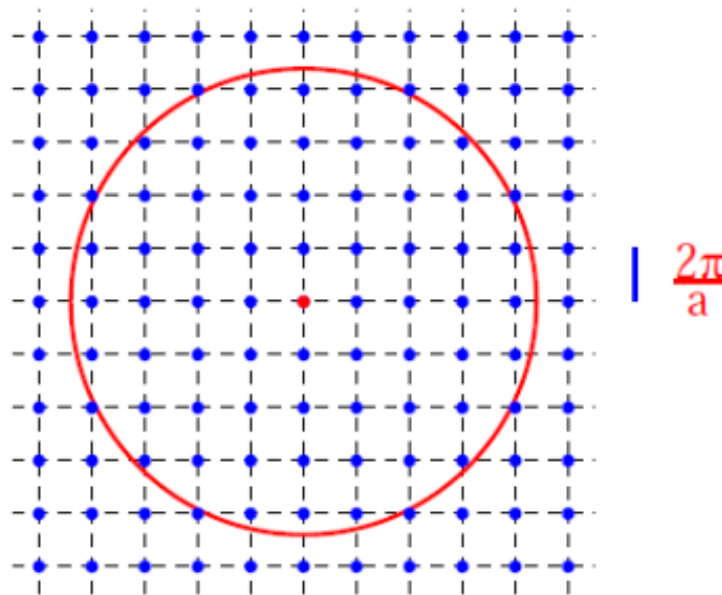


- Stress is a first derivative of external parameter and *Hellman-Feynman* theorem holds.
Nielsen and Martin: Phys.Rev.Lett. **50**, 697 (1983),
Nielsen and Martin: Phys.Rev.B **32**, 3780 (1985)
- Derivative of the Kinetic Energy, Hartree and Exchange-Correlation w.r.t. a uniform deformation.
- Does not depend explicitly from the ion positions.



Pulay Stress and smooth cutoff

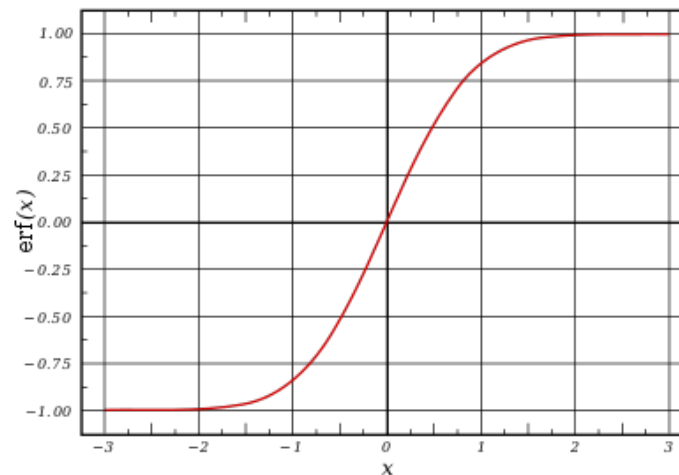
- The plane wave basis set depends on the strain via the length of the reciprocal lattice vectors.
- Evaluations of Energy and Stress can be discontinuous if the cutoff is not at full convergence.



Smooth Cutoff

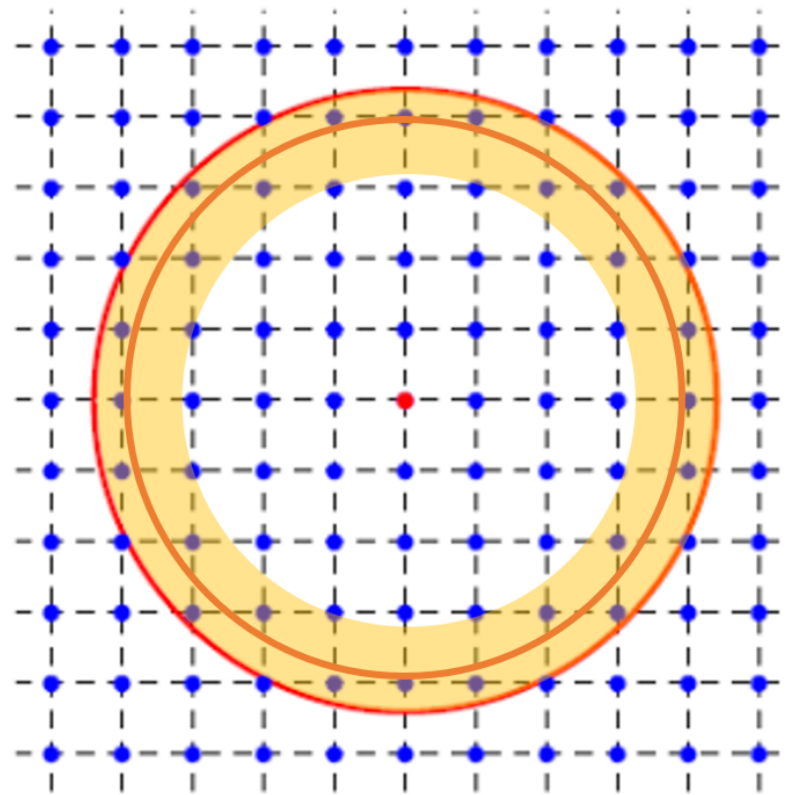
- The smooth cutoff creates a region where plane wave transit in and out from the significant basis set with regularity, being account when evaluating stress and energy.
- It is done changing the Kinetic Energy Functional so that borderline plane waves become energetically *expensive*

$$T(\mathbf{G}) = |\mathbf{G}|^2 + Q_{cut} \left(1 + \operatorname{erf} \left(\frac{|\mathbf{G}|^2 - E_{cfix}}{\sigma} \right) \right)$$



Smooth cutoff

- In `pw.x`:
 - `ecfied` is the energy over which the plane waves reach the full penalty cost
 - `qesigma` is the width the which the penalty start to be added
 - `qcutz` is the value of the energy penalty.



Structural Optimization

Structural optimization in `pw.x` can be done:

- At fixed cell: `calculation='relax'`
 - $\vec{X} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$, atomic position
 - $\vec{g} = -(\mathbf{F}_1, \dots, \mathbf{F}_N)$, forces on atoms
- At variable cell: `calculation='vc-relax'`
 - $\vec{X} = (\mathbf{r}_1, \dots, \mathbf{r}_N, e_{\alpha\beta})$, atomic position, cell strains
 - $\vec{g} = -(\mathbf{F}_1, \dots, \mathbf{F}_N, \sigma_{\alpha\beta})$, forces on atoms, stress tensor
- Algorithms:
 - BFGS: quasi Newton algorithm `ion_dynamics='bfgs'`, `cell_dynamics='bfgs'`
 - quickmin: Verlet Damped dynamic `ion_dynamic='damp'`,
`cell_dynamics='damp-pr'` or `'damp-w'`

Quickmin



Damped molecular dynamics can be used for structural optimization. The damping mechanism drains kinetic energy out of the system while the system falls into the structural minimum.

- quickmin kills any generalized velocity component whose direction is opposite to the corresponding generalized force component.
- pros: it is robust, works well even when the starting point is not close to the minimum.
- cons: requires some experience with molecular dynamics for setting the time step and for variable cell the fictitious mass of the cell. Generally slower than BFGS.



BFGS

The BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm is the workhorse for structural minimization, either at fixed cell or with variable cell.

Close to an equilibrium point $\vec{X}^{(eq)}$, for which $\nabla E(\vec{X}^{(eq)}) = 0$ holds, a quadratic form is assumed for the function $E(\vec{X})$ (H is the Hessian matrix):

$$E(\vec{X}) \simeq E(\vec{X}^{(eq)}) + \frac{1}{2}(\vec{X} - \vec{X}^{(eq)})^T H (\vec{X} - \vec{X}^{(eq)})$$

Given two points \vec{X}_1 and \vec{X}_0 and corresponding gradients $\vec{g} = \nabla E(\vec{X})$, this means $\vec{g}_1 - \vec{g}_0 = H(\vec{X}_1 - \vec{X}_0)$, that is, $\vec{g}_1 = 0$ if $\vec{X}_1 = \vec{X}_0 - H^{-1}\vec{g}_0$ (*Newton-Raphson step*).

Practical algorithm: a sequence of calculations at positions \vec{X}_i

$$\vec{X}_{i+1} = \vec{X}_i + T_k^L \frac{\vec{s}_k^{NR}}{|\vec{s}_k^{NR}|}, \quad \vec{s}_k^{NR} = -H_k^{-1}\vec{g}_k$$

where T_k^L is called "trust radius".

BFGS: update and line search

The inverse Hessian matrix is updated at each step using the BFGS scheme:

$$H_{k+1}^{-1} = H_k^{-1} + \left(1 + \frac{\gamma_k^T H_k^{-1} \gamma_k}{s_k^T \gamma_k}\right) \frac{s_k s_k^T}{s_k^T \gamma_k} - \left(\frac{s_k \gamma_k^T H_k^{-1} + H_k^{-1} \gamma_k s_k^T}{s_k^T \gamma_k}\right)$$

$$\gamma_k = g_{k+1} - g_k$$

- At each step the trust radius is updated so that it satisfies the 2 Wolfe conditions:
 - Sufficient decrease: the decrease of the energy value must be more than an estimated amount otherwise the radius is decreased.

$$E_{new} \leq E_{old} + w_1 T_k^L \vec{g} \cdot \vec{s}_k$$

- Curvature condition: the new slope is checked: if it is still too high the trust radius is increased.

$$\vec{g}_{new} \cdot \vec{s}_k \geq w_2 \vec{g}_{old} \cdot \vec{s}_k$$

Some important aspects of structural optimization



- Structural optimization may find only the closest minimum: it cannot overcome potential barriers, so it may be trapped into a local minimum.
- Structural optimization *does not break crystal symmetry*, at least in principle (numerical noise may occasionally break the symmetry).
- In variable-cell optimization, the PW basis set is kept fixed during optimization. This means that the final result is not exactly equal to what one gets by starting the calculation from scratch with the same cutoff, because the two basis sets are not exactly the same.
- Structural optimization uses both energies and forces to locate the minima along search directions. Discrepancies between those values, due to insufficient scf convergence, will lead to bad convergence of the BFGS algorithm or even to errors. Remember that the error on forces is *linear* in the self-consistency error, while the error on energies is *quadratic* due to its variational character.



Born-Oppenheimer Molecular Dynamics



Let us assume classical behavior for the nuclei and electrons in the ground state. We introduce a classical Lagrangian:

$$L = \frac{1}{2} \sum_I M_I \dot{\mathbf{r}}_I^2 - E(\mathbf{r}_I)$$

describing the motion of nuclei. The equations of motion:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{r}}_I} - \frac{\partial L}{\partial \mathbf{r}_I} = 0, \quad \mathbf{P}_I = \frac{\partial L}{\partial \dot{\mathbf{r}}_I}$$

are nothing but usual Newton's equations:

$$\mathbf{P}_I \equiv M_I \mathbf{v}_I, \quad M_I \dot{\mathbf{v}}_I = \mathbf{F}_I,$$

that can be discretized and solved by integration.

This procedure defines Molecular Dynamics “on the Born-Oppenheimer surface”, with electrons always at their instantaneous ground state.



Discretization of the equation of motion



Like in classical MD, the equation of motions can be discretized using the *Verlet algorithm*:

$$\begin{aligned}\mathbf{r}_I(t + \delta t) &= 2\mathbf{r}_I(t) - \mathbf{r}_I(t - \delta t) + \frac{\delta t^2}{M_I}\mathbf{F}_I(t) + \mathcal{O}(\delta t^4) \\ \mathbf{v}_I(t) &= \frac{1}{2\delta t} [\mathbf{r}_I(t + \delta t) - \mathbf{r}_I(t - \delta t)] + \mathcal{O}(\delta t^3).\end{aligned}$$

or the *Velocity Verlet*:

$$\begin{aligned}\mathbf{v}_I(t + \delta t) &= \mathbf{v}_I(t) + \frac{\delta t}{2M_I} [\mathbf{F}_I(t) + \mathbf{F}_I(t + \delta t)] \\ \mathbf{r}_I(t + \delta t) &= \mathbf{r}_I(t) + \delta t\mathbf{v}_I(t) + \frac{\delta t^2}{2M_I}\mathbf{F}_I(t).\end{aligned}$$

Both sample the *microcanonical ensemble*, or NVE: the energy (mechanical energy: kinetic + potential) is conserved.



Technicalities



- time step as big as possible, but small enough to follow nuclear motion with little loss of accuracy. Rule of thumb: $\delta t \sim 0.01 - 0.1 \delta t_{max}$, where $\delta t_{max} = 1/\omega_{max} =$ period of the fastest phonon (vibrational) mode.
- calculations of forces must be *very well converged* (good self-consistency needed) at each time step or else a systematic *drift* of the conserved energy will appear

Note that:

- the error on DFT *energy* is a *quadratic* function of the self-consistency error of the charge density (because energy has a minimum in correspondence to the self-consistent charge)
- the error for DFT *forces* is a *linear* function of the self-consistency error of the charge density

As a consequence, Born-Oppenheimer MD is usually computationally heavy



THE END