# Ab-initio simulations

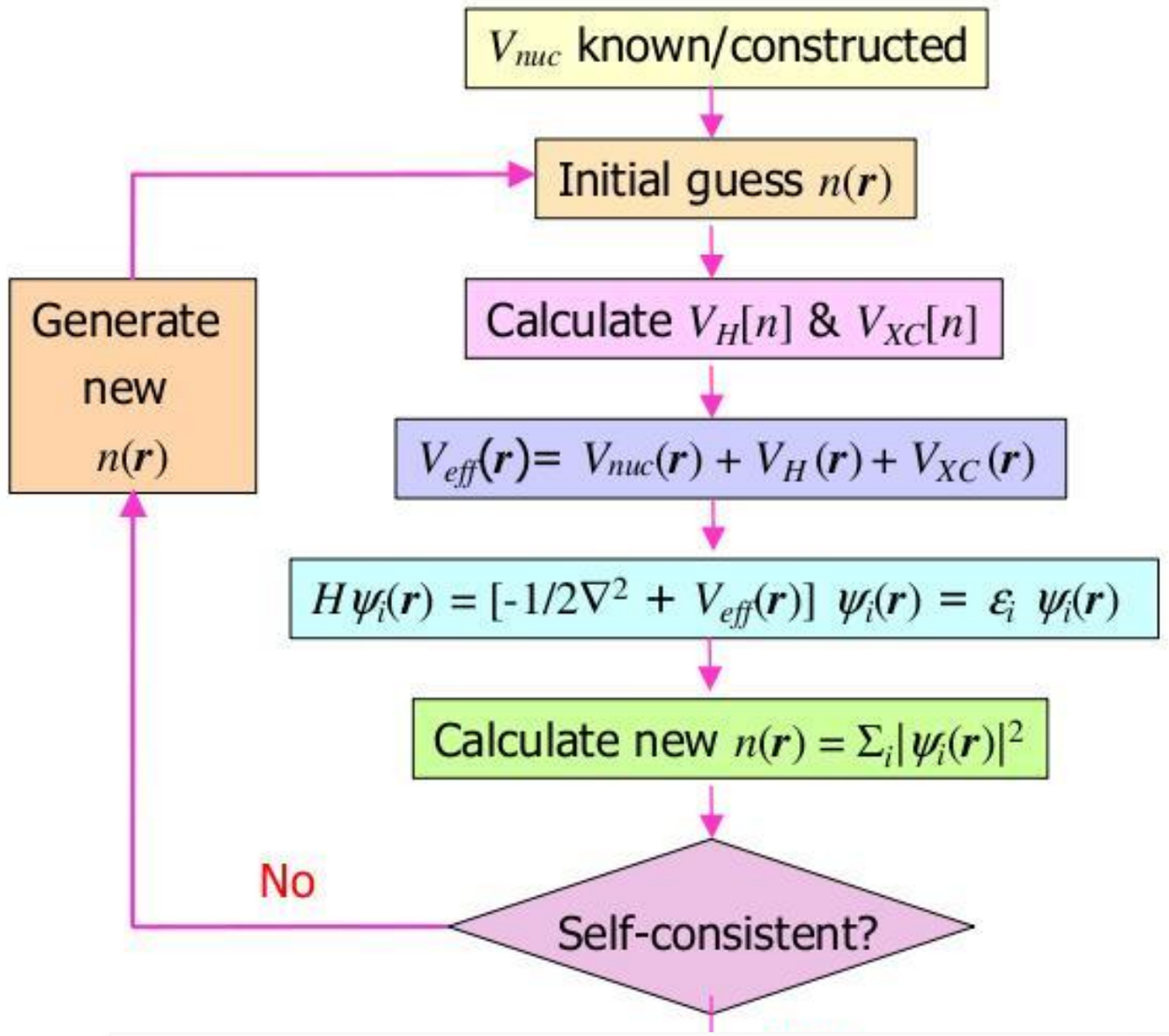
Mini course on DFT for PhD
March 2024

## lecture IV

(many slides are adapted by M. Peressi from P. Giannozzi and S. de Gironcoli)

# DFT and SCF cycle: what's next?

- Supercell geometry: lattice vectors + atoms in the unit cell
- Plane-wave basis set, determined by the lattice and by a single parameter (*cutoff*)
- Atom-based pseudopotentials representing the electron-nuclei potential ($V_{nuc}$ in the figure)
- Charge density computed with valence electrons only, on a suitable grid of **k**-points.




**problem solved!**
**calculate energy, bands (done)**
**but also stress, forces, etc etc**

# What can I do after the SCF cycle?

## Structural optimization

- calculation of **stress** and **forces** acting on atoms
  (tstress=.true. , tprnfor=.true. - default: .false.)
- from forces => **optimization of ionic positions**
  (calculation='**relax**' :
  several SCF steps, followed by calculation of forces and generation of new atomic positions)
- from stress => **optimization of the cell**
  (calculation='vc-relax' : variable-cell relaxation)
- calculations of energy barrier for reactions:

NEB method: saddle points of elementary chemical reactions

- molecular dynamics

https://www.quantum-espresso.org/Doc/INPUT_PW.html

# Input File Description

## Program: pw.x / PWscf / Quantum ESPRESSO (version: 7.2)

### TABLE OF CONTENTS

# Stress

# STRAIN, STRESS (and PRESSURE)

homogeneous deformation

$$r \longrightarrow r' = (1+\epsilon)r$$

strain

stress

$$\sigma = -\frac{1}{\Omega}\frac{\partial E}{\partial \epsilon}$$

It's a first order derivative (Hellman-Feynman)

-> NO NEED to know how electrons re-adjust

-> NO NEED to know how internal coordinates change

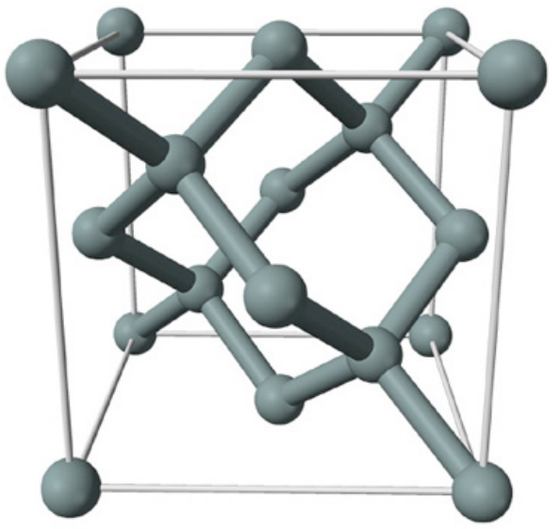
If we are interested just in homogeneous deformation (scalar ε), we can obtain σ directly from the GS wavefunction (no need of doing dE/dε by hand…)

Nielsen and Martin  Phys.Rev.Lett. 50, 697 (1983),
Nielsen and Martin  Phys.Rev. B 32, 3780 & 3792 (1985).

# STRAIN, STRESS (and PRESSURE)

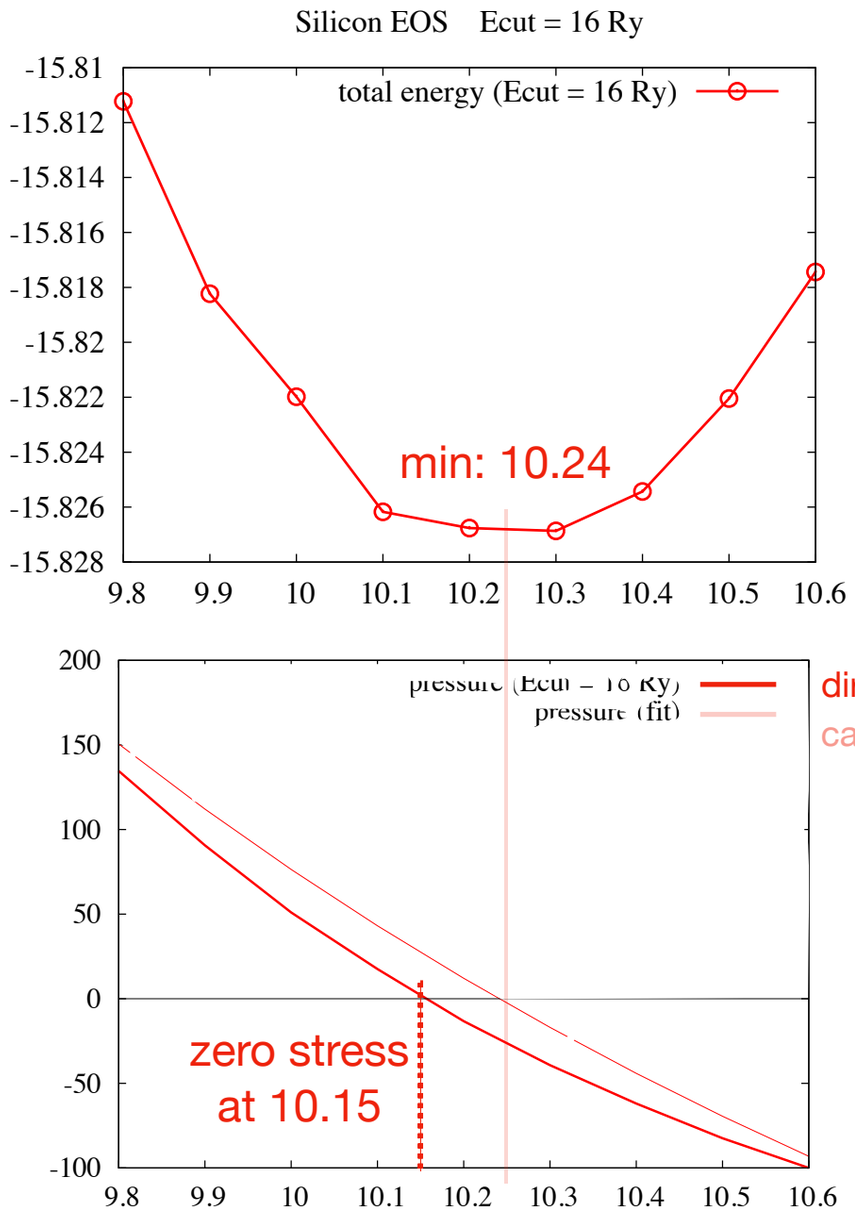
**e.g.: Si** (done)

```
... ... ...

!    total energy              =     -15.82676466 Ry
     Harris-Foulkes estimate   =     -15.82676466 Ry
     estimated scf accuracy    <          4.5E-09 Ry

     The total energy is the sum of the following terms:

     one-electron contribution =       4.80030732 Ry
     hartree contribution      =       1.09369711 Ry
     xc contribution           =      -4.82101051 Ry
     ewald contribution        =     -16.89975858 Ry

     convergence has been achieved in   6 iterations
```



```
     entering subroutine stress ...

          total   stress  (Ry/bohr**3)                   (kbar)     P=  -13.29
-0.00009036   0.00000000   0.00000000    -13.29      0.00      0.00
 0.00000000  -0.00009036   0.00000000      0.00    -13.29      0.00
 0.00000000   0.00000000  -0.00009036      0.00      0.00    -13.29
```

# STRESS (and PRESSURE)
# Minimum energy is not exactly zero pressure

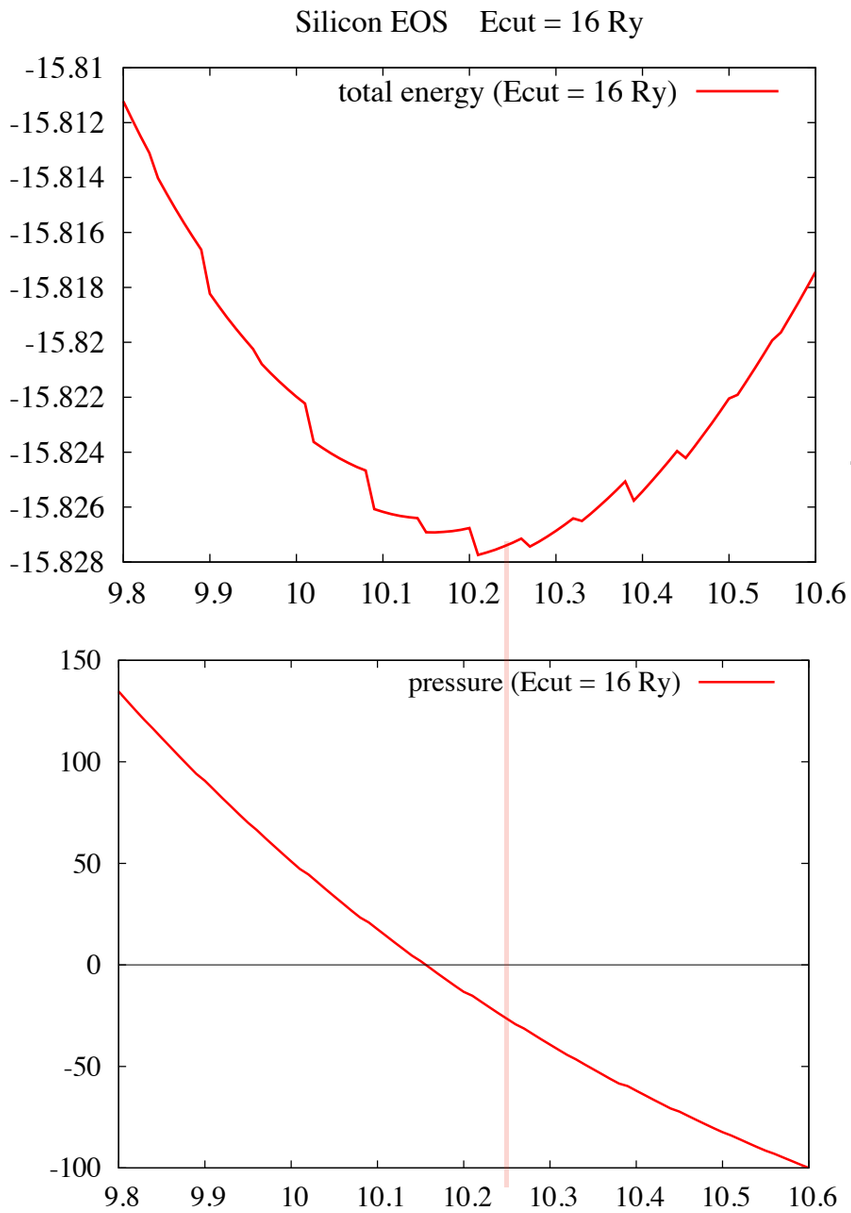



directly calculated from QE (Hellman-Feynman)

calculated as derivative of the energy
calculated from QE

Is there something wrong ?

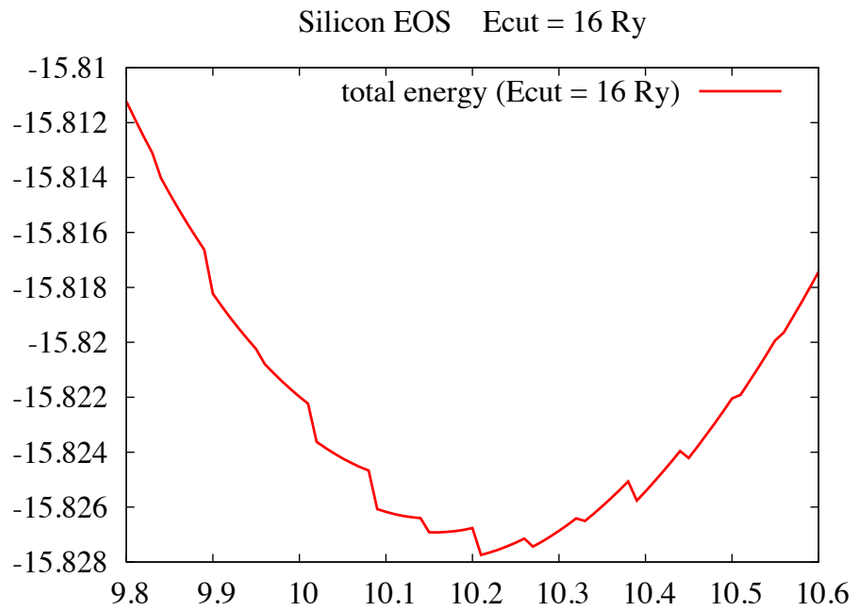
# STRESS (and PRESSURE)
# … and constant energy cutoff does not mean constant NPW

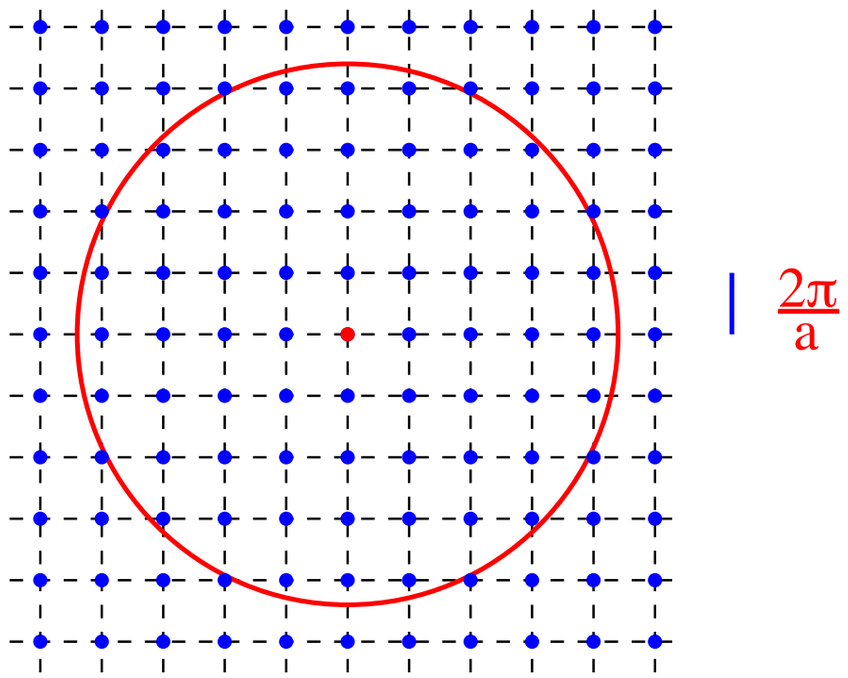



finer sampling in the lattice parameter range
emphasizes discontinuities

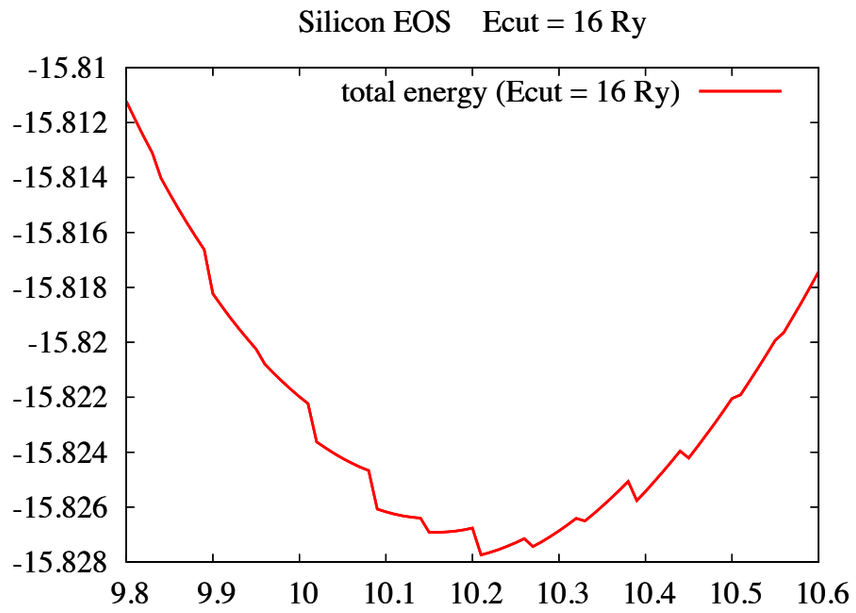
results for a denser grid

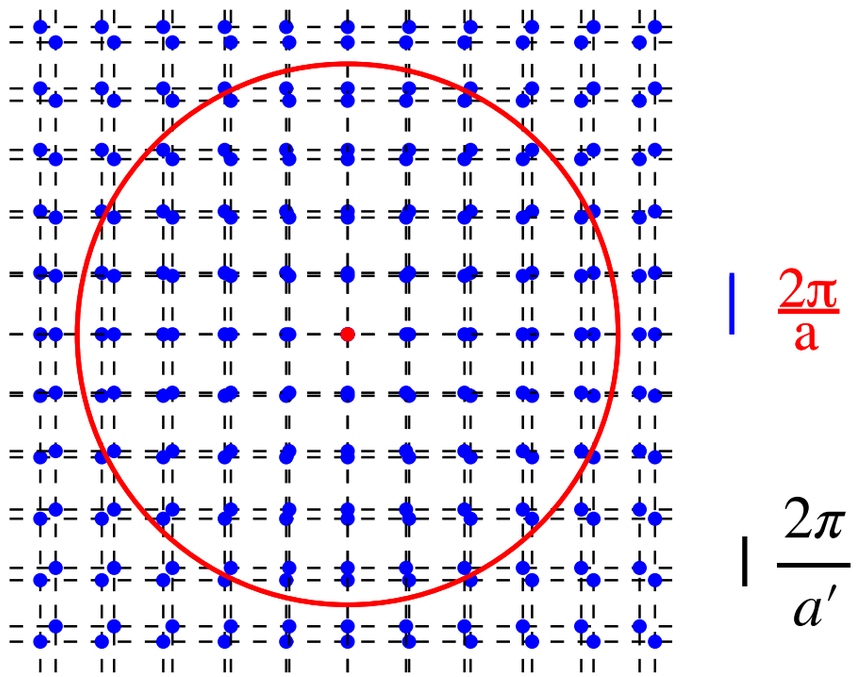
# STRESS (and PRESSURE)
# … and constant energy cutoff does not mean constant NPW



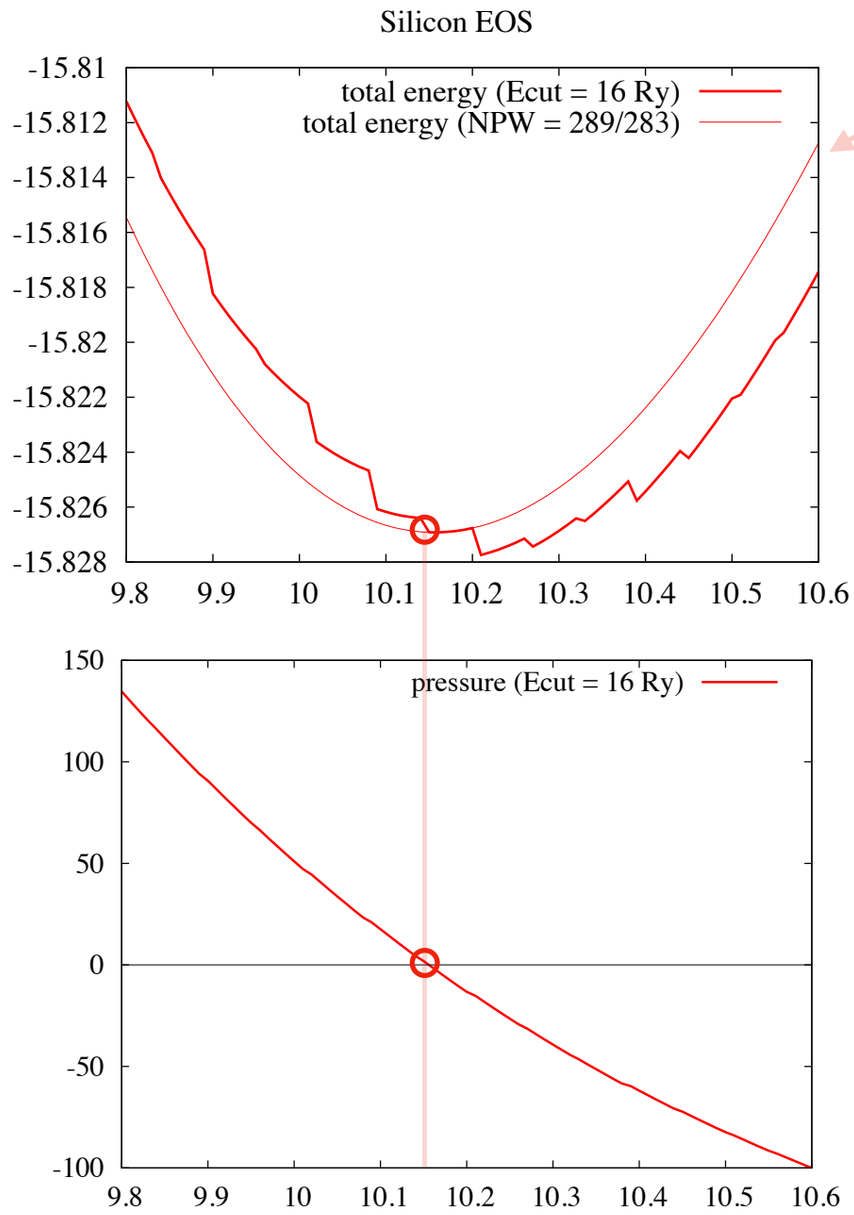
# STRESS (and PRESSURE)
# … and constant energy cutoff does not mean constant NPW




(larger a' => smaller k spacing => more NPW)

# STRESS (and PRESSURE)
# … and constant energy cutoff does not mean constant NPW



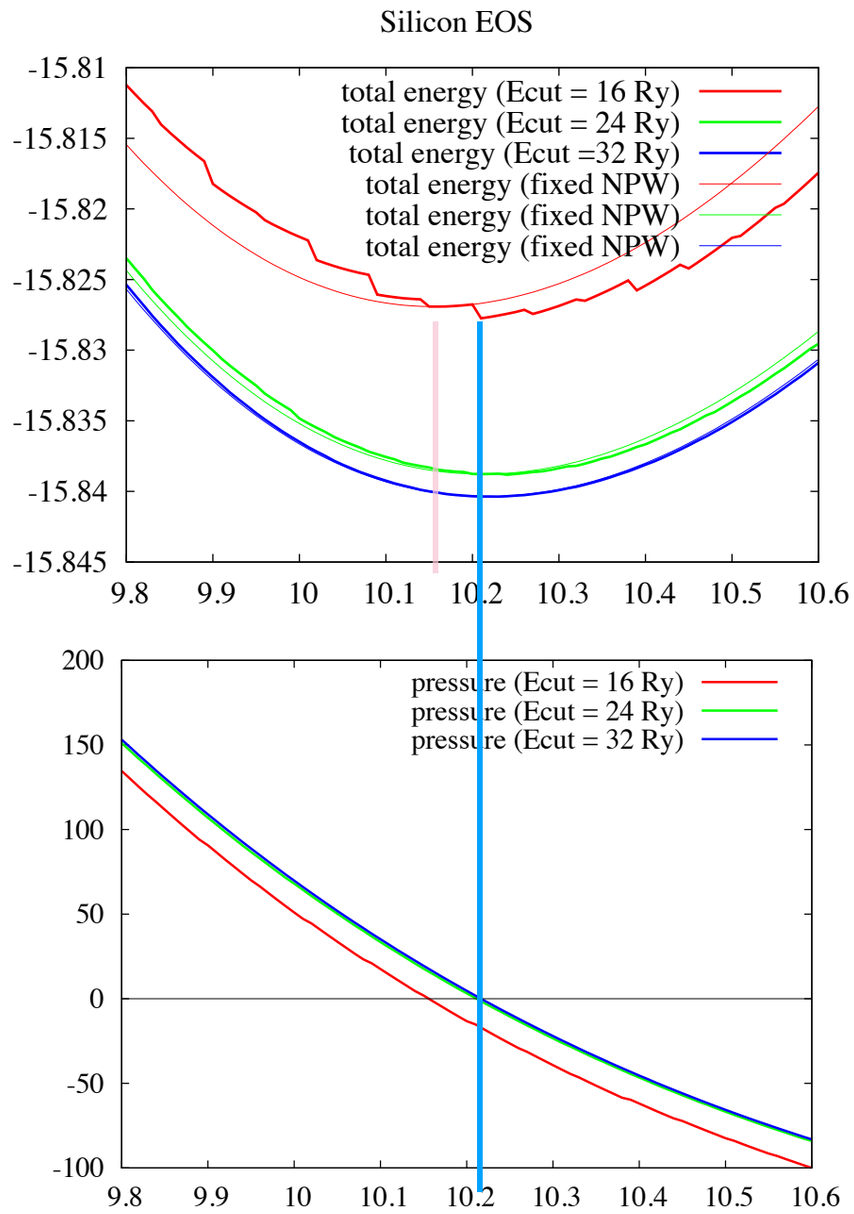

3) keep NPW (almost) fixed when changing celldm(1)

2) get NPW

1) find *a* (celldm(1)) corresponding to zero pressure

fixed Cutoff *vs* fixed Number of Plane waves

# STRESS (and PRESSURE)
# increasing the energy cutoff is better




fixed Ecut ≈ fixed NPW

$\min(E_{tot}) \approx$ pressure=0

fixed Cutoff *vs* fixed Number of Plane waves

Calculations at FIXED CUTOFF need some intepolation in order to extract structural parameters but converge more rapidly to the accurate structural properties than the smoothly varying calculations at FIXED NUMBER OF PW.

Complete convergence is needed for accurate calculation of stress

... or rather ...

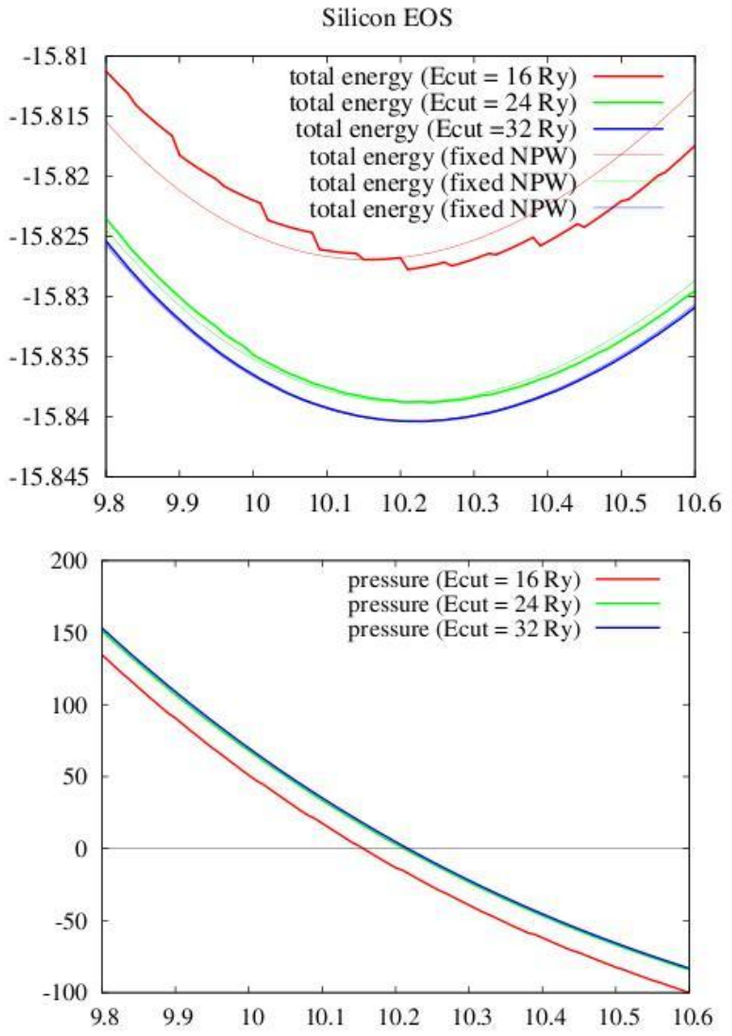
so that the calculation of stress (that assumes a fixed number of PW) agrees with the smooth interpolated result obtained using a fixed cutoff

# STRESS (and PRESSURE)

## Minimum energy is not exactly zero pressure

## Note on incompleteness of the PW basis set

Practical calculations are invariable performed with a cutoff "as low as possible", sometimes quite far from convergence. The consequence are especially visible when comparing the $E(V)$ curve at fixed cutoff and the same at fixed number of plane waves: the curve at low cutoff is "rigged", the pressures calculated from the stress and from the equation of state do not match. This is a manifestation of the *Pulay* incomplete-basis-set error on the stress: plane waves *depend upon the strain* via the unit cell.




The strategy "fixed cutoff + fit to an EOS" converges faster than fixing the number of PWs. The latter is equivalent to locating the zero of the computed pressure (via the stress). **conclusions**

The incompleteness of basis set is not as serious as it may look: energy *differences* between different structures, structural parameters such as lattice parameters and bond lengths, converge much quicker than absolute energies.

# NON HOMOGENEOUS STRAIN (for elastic constants)

We can set a given strain by using celldm(i)

from ibrav=1 (SC)
with celldm(1)=…

⇨

to ibrav=6 (Tetragonal P)
with celldm(1)(=celldm(2))=xxx
and celldm(3)(=c/a)=yyy

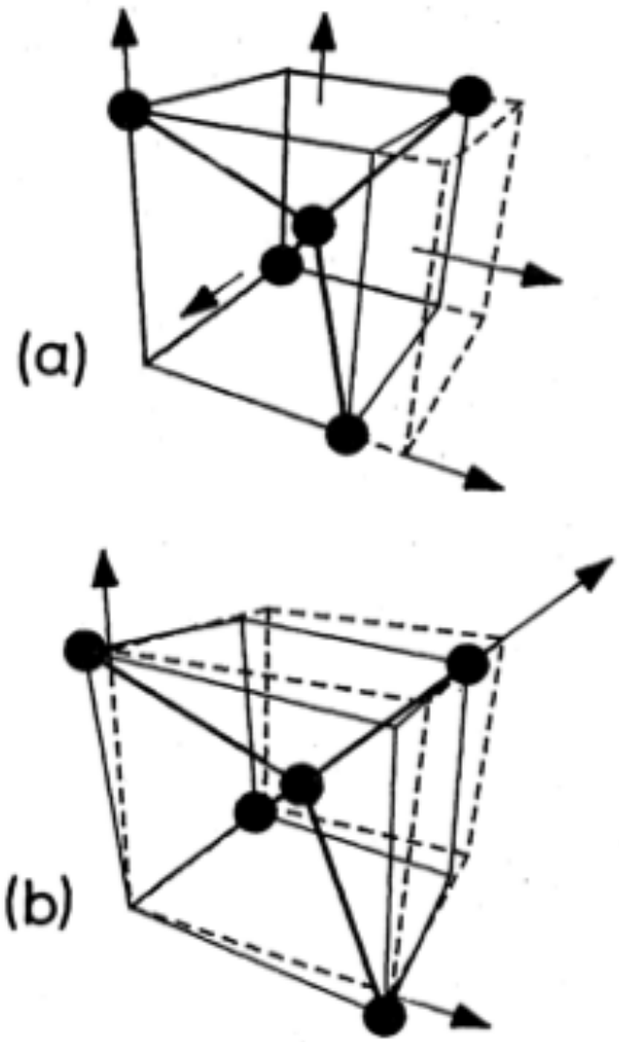



FIG. 2. Perspective view of cubes (dashed lines) deformed by strains to take new shapes (solid lines). Thick arrows indicate the resulting directions of stress exerted by the solid. (a) A strain $\epsilon_1 < 0$ along [100] resulting in $\sigma_1$ and $\sigma_2 = \sigma_3$ stresses. (b) A strain $\epsilon_4 = \epsilon_5 = \epsilon_6 < 0$ along [111] resulting in a stress in the same direction.

To easily enter arbitrary strain (arbitrary lattice vectors), use:

**ibrav = 0**
AND
**CELL_PARAMETERS** card

E.g., for unstrained FCC:

```
CELL_PARAMETERS alat
  0.0 0.5 0.5
  0.5 0.0 0.5
  0.5 0.5 0.0
```

applying the strain to each basis vector:

$$\mathbf{a}_i' = (1 + \epsilon)\mathbf{a}_i$$

Therefore: apply strain (it could be +/- ~5%) => compute the stress => determine the elastic constant (…)

# FORCES

## Hellmann-Feynman Forces

Forces on atoms are the derivatives of the total energy wrt atomic positions. The *Hellmann-Feynman theorem* tells us that forces are the expectation value of the derivative of the external potential only:

$$\mathbf{F}_\mu = -\frac{\partial E}{\partial \mathbf{R}_\mu} = -\sum_i f_i \langle \psi_i | \frac{\partial V}{\partial \mathbf{R}_\mu} | \psi_i \rangle = -\int n(\mathbf{r}) \frac{\partial V}{\partial \mathbf{R}_\mu} d\mathbf{r}$$

the rightmost expression being valid only for *local* potentials, $V \equiv V(\mathbf{r})$ (the one at the left is more general, being valid also for nonlocal potentials $V \equiv V(\mathbf{r}, \mathbf{r}')$).

*Demonstration* (simplified). In addition to the explicit derivative of the external potential (first term), there is an implicit dependency via the derivative of the charge density:

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

The red term cancels due to the *variational character* of DFT: $\delta E / \delta n(\mathbf{r}) = \mu$, constant.

The calculation of the Hellmann-Feynman forces is straightforward (in principle, not necessarily in practice!) once the self-consistent electronic structure is calculated.

**forces are derivatives of the energy => convergence is more difficult!**

(in general: convergence of integrated quantities is much easier)

# Forces

**Forces = 0**

(and stress or pressure = 0 if the calculations are well converged w.r.t. kinetic energy cutoff)

**are enough to identify the equilibrium structure?**

Note that in simple crystalline structures *the force on atoms can be zero* by symmetry, even if the system is *not at equilibrium*!

different crystalline structures for elemental Si:

the minimum of each structure
is characterized by zero forces

but this is not enough to determine
the equilibrium structure

equilibrium structure =
minimum energy + zero forces

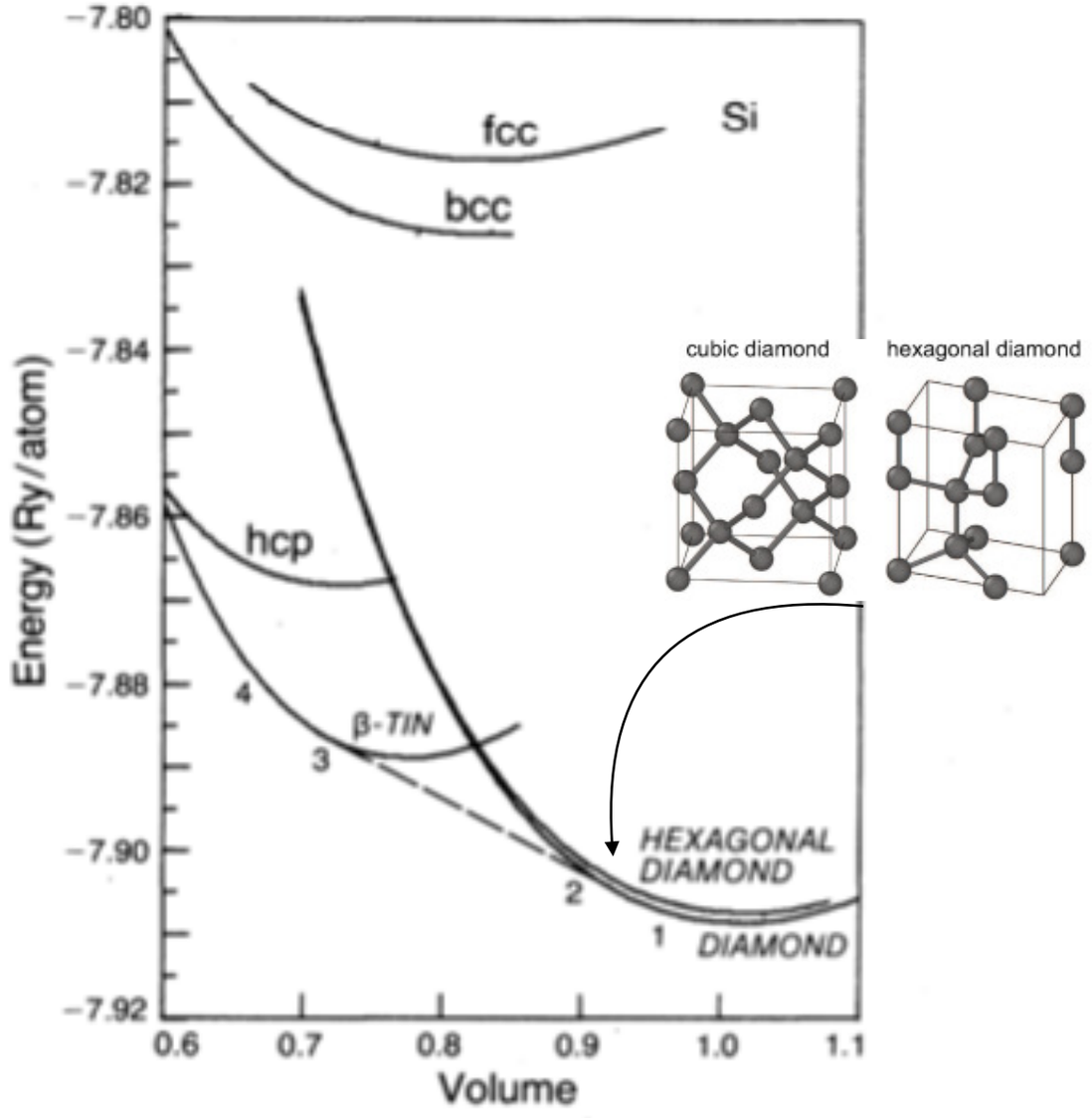



FIG. 1. The diamond, hexagonal diamond, and $\beta$-tin, hcp, bcc, and fcc structural energies (in units of Ry/atom) as a function of the atomic volume [normalized to the measured free volume (Ref. 16)] for Si. The dashed line is the common tangent of the energy curves for the diamond and the $\beta$-tin structures.

Yin and Cohen, PRL 45, 1004 (1980)

# What can I do after the SCF cycle?

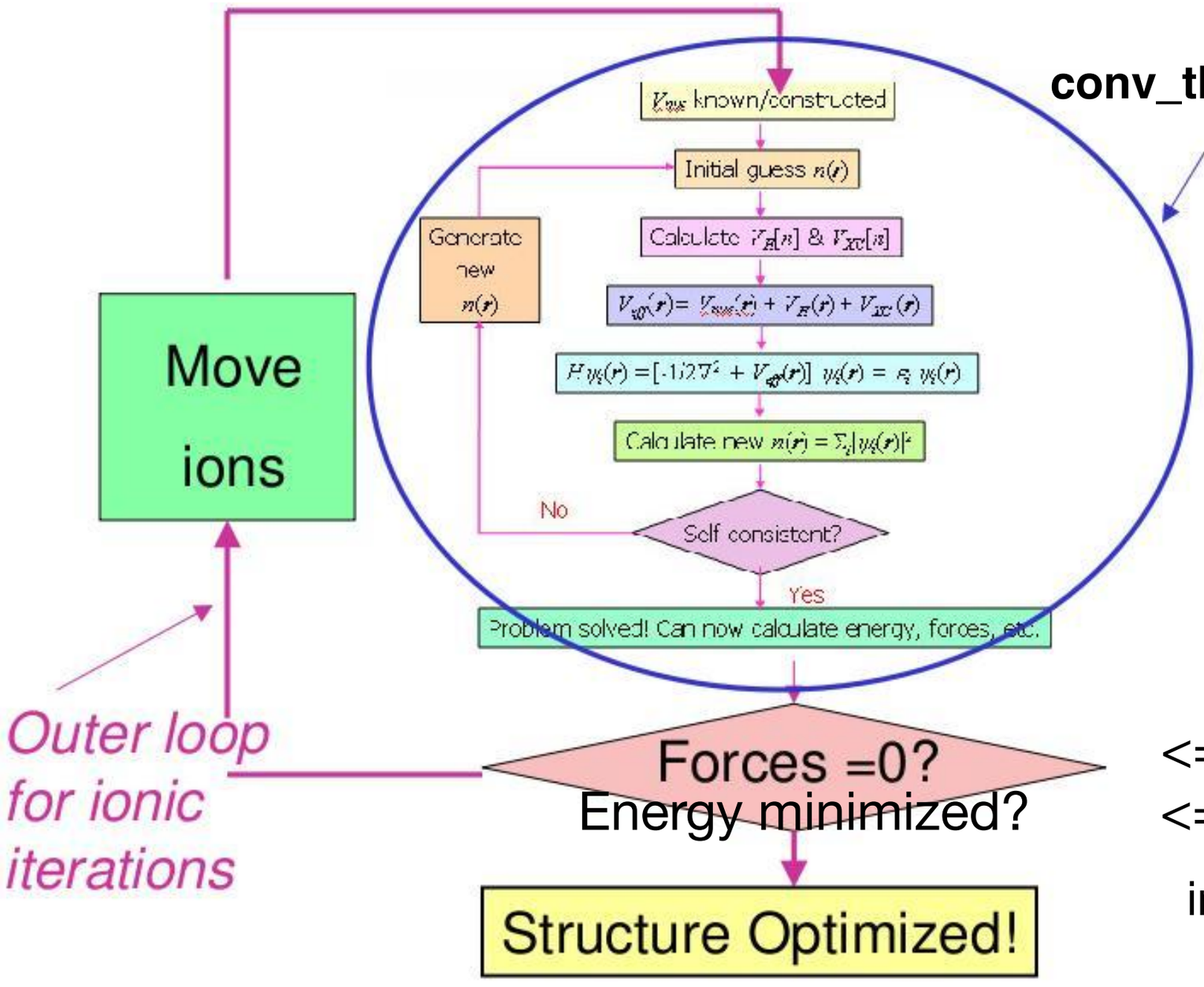
## Structural optimization

- calculation of **stress** and **forces** acting on atoms
  (tstress=.true. , tprnfor=.true.)
- from forces => **optimization of ionic positions**
  (calculation='**relax**' :
  several SCF steps, followed by calculation of forces and generation of new atomic positions)
- from stress => **optimization of the cell**
  (calculation='vc-relax' : variable-cell relaxation)
- calculations of energy barrier for reactions:

NEB method: saddle points of elementary chemical reactions

- molecular dynamics

# Structural optimization




**conv_thr** in &ELECTRONS card

<= forc_conv_thr
<= etot_conv_thr

in &CONTROL card
(active when
calculation='relax')

https://www.quantum-espresso.org/Doc/INPUT_PW.html

# Input File Description

**Program: pw.x / PWscf / Quantum ESPRESSO** (version: 7.2)

**TABLE OF CONTENTS**

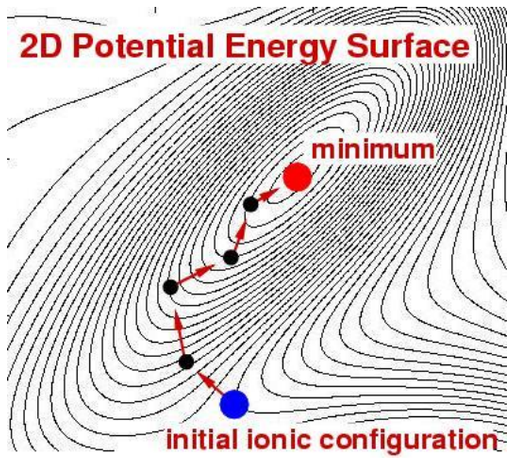
='relax'

extensive!

# Structural Optimization and Molecular Dynamics

Within the *Born-Oppenheimer*, or *adiabatic* approximation, the total energy as a function of atomic positions, or *Potential Energy Surface* (PES), determines the behaviour of nuclei.

The *global* ground state can be found by minimizing the function $E(\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_N)$, depending upon the $3N$ atomic coordinates for a system of $N$ atoms. This is a "standard" mathematical problem: finding the minimum of a function, knowing its derivatives, that is, the Hellmann-Feynman forces (in the picture, a cartoon of a PES in two dimensions with the path to the minimum).




Once forces are calculated, one can perform not only structural optimization (also known as "relaxation"), but also *molecular dynamics*. If a classical behaviour of the nuclei is assumed, all the machinery of classical MD can be recycled, with forces calculated from *first principles*.

# Quasi-Newton algorithms for structural optimization

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):

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

X: atomic positions
g: forces on atoms

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 $\vec{X}_i$

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

incremental displacements according to the forces

where $T_k^L$ is called "trust radius".

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 fixed cell:
  - $\vec{X} = (\vec{d_1}, \ldots, \vec{d_N})$, atomic positions
  - $\vec{g} = -(f_1, \ldots, f_N)$, Hellmann-Feynman forces on atoms
- With variable cell:
  - $\vec{X} = (\vec{d_1}, \ldots, \vec{d_N}, \epsilon_{\alpha\beta})$, atomic positions and cell strains
  - $\vec{g} = -(\vec{f_1}, \ldots, \vec{f_N}, \sigma_{\alpha\beta})$, Hellmann-Feynman forces on atoms and stresses

## Summary (optimization of ionic positions only)

## 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 symetry).

- 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.

# Variable cell

# What can I do after the SCF cycle?

## Structural optimization

- calculation of **stress** and **forces** acting on atoms
  (tstress=.true. , tprnfor=.true.)
- from forces => **optimization of ionic positions**
  (calculation='**relax**' :
  several SCF steps, followed by calculation of forces and generation of new atomic positions)
- from stress => **optimization of the cell**
  (calculation='vc-relax' : variable-cell relaxation)
- calculations of energy barrier for reactions:

NEB method: saddle points of elementary chemical reactions

- molecular dynamics

# How to perform variable-cell relaxation?

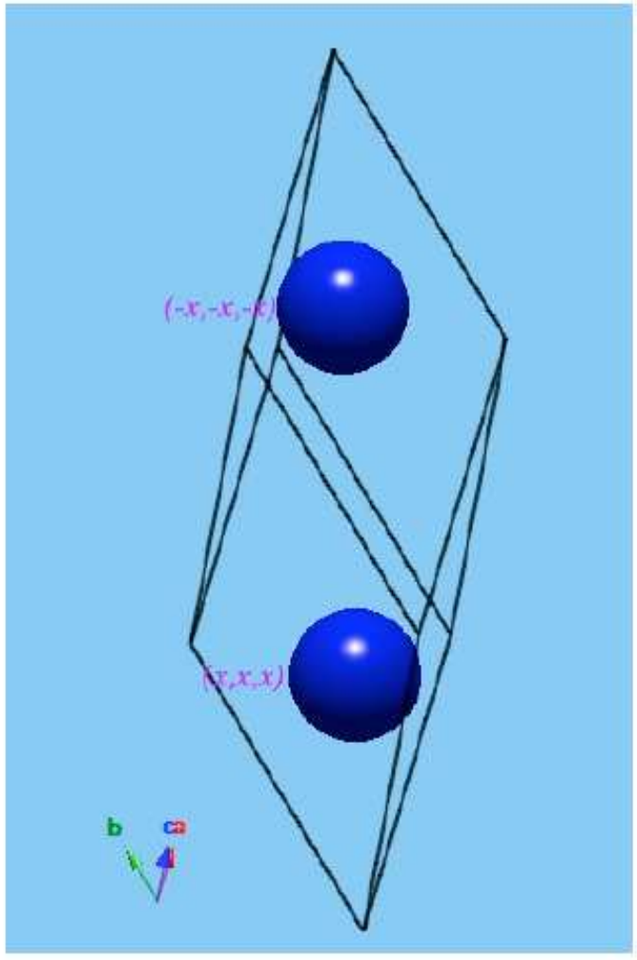
“by hands”: suppose we have a cell with 2 parameters and we need to optimize the a and c lattice parameters => one need to perform a 2D scan over the two parameters.

Variable-cell relaxation: in some cases this can be a more convenient option

- calculation = ’vc-relax’

# A7 to sc transition in As

- Unit Cell: (*a=b=c*, *cosAB=cosAC=cosBC*)
- Guessing  *a*=3.85 Å, *x*=0.275, *cosAB*=0.49517470
- Energy Cut-Off=30 Ry.
- 2 As  per unit cell
- 2As at ±(x, x, x,);
- When *x*=0.25, *cosAB*=0.5
  => Simple Cubic



## Variable Cell-Shape Relaxation

BFGS relaxation

Crystal configuration is defined by 3*NAT + 9 variables

$$\tau_\alpha^s = \sum_k a_\alpha^k x_k^s$$

$x_k^s$ = 3*NAT internal (crystal) coordinates;

$a_\alpha^k$ = Bravais lattice vectors (9 variables)

The algorithm can be applied as usual paying attention to use the appropriate generalized forces

$$-\frac{\partial H}{\partial x_k^s} = \sum_\alpha F_\alpha^s a_\alpha^k; \qquad -\frac{\partial H}{\partial a_\alpha^k} = \Omega \sum_k (a^{-1})_k^\beta \left(\sigma_{\beta\alpha} - P\delta_{\beta\alpha}\right)$$

and to start from an inverse Hessian that respects the symmetry of the crystal.

# nudged elastic bands

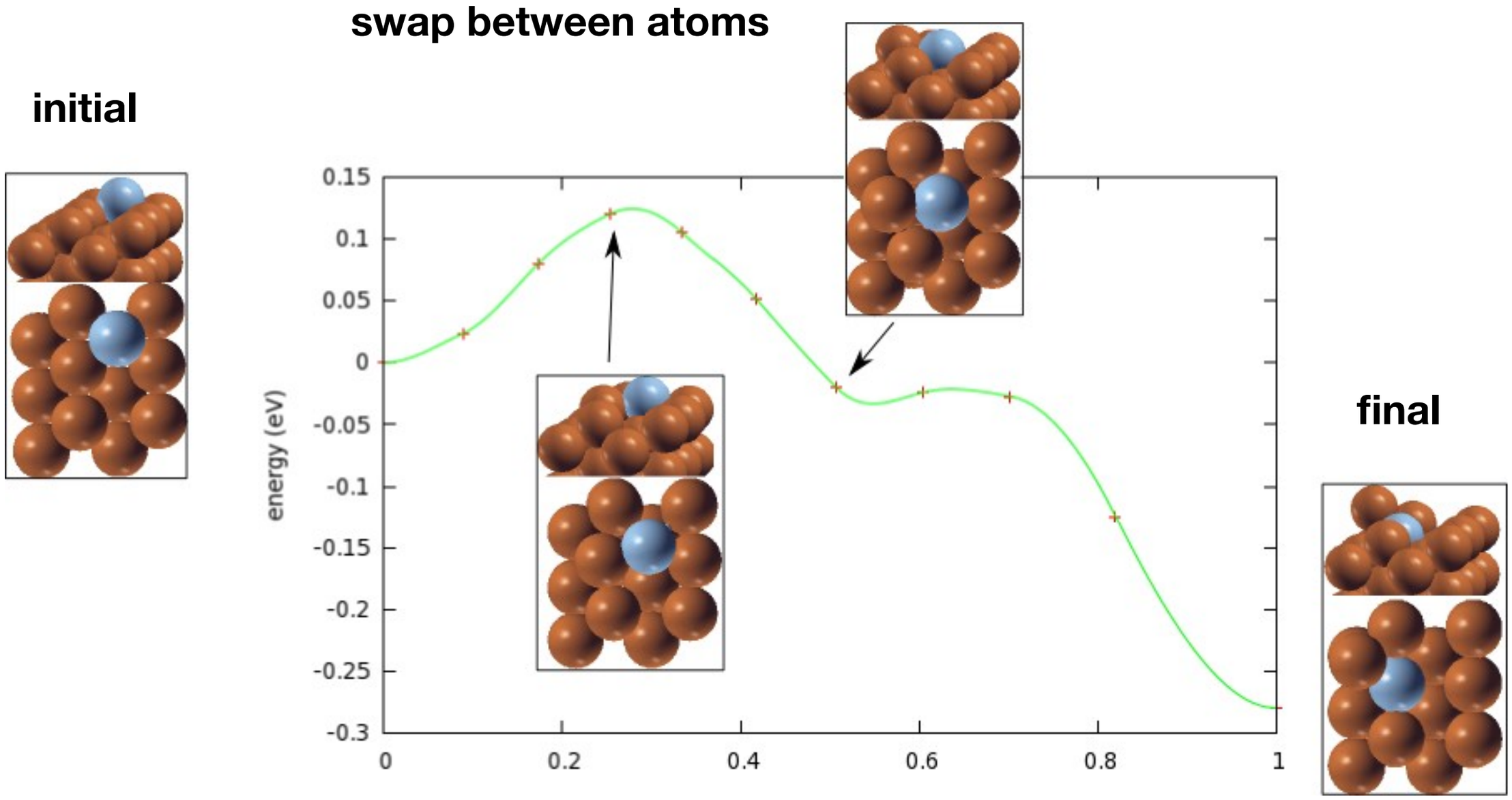
# nudged elastic bands

calculation of:
energy barriers, transition states, reaction pathway

We have to know at least TWO IMAGES :
first (**initial state)** and last (**final state)**
corresponding to “**reactants**” and “**products**”

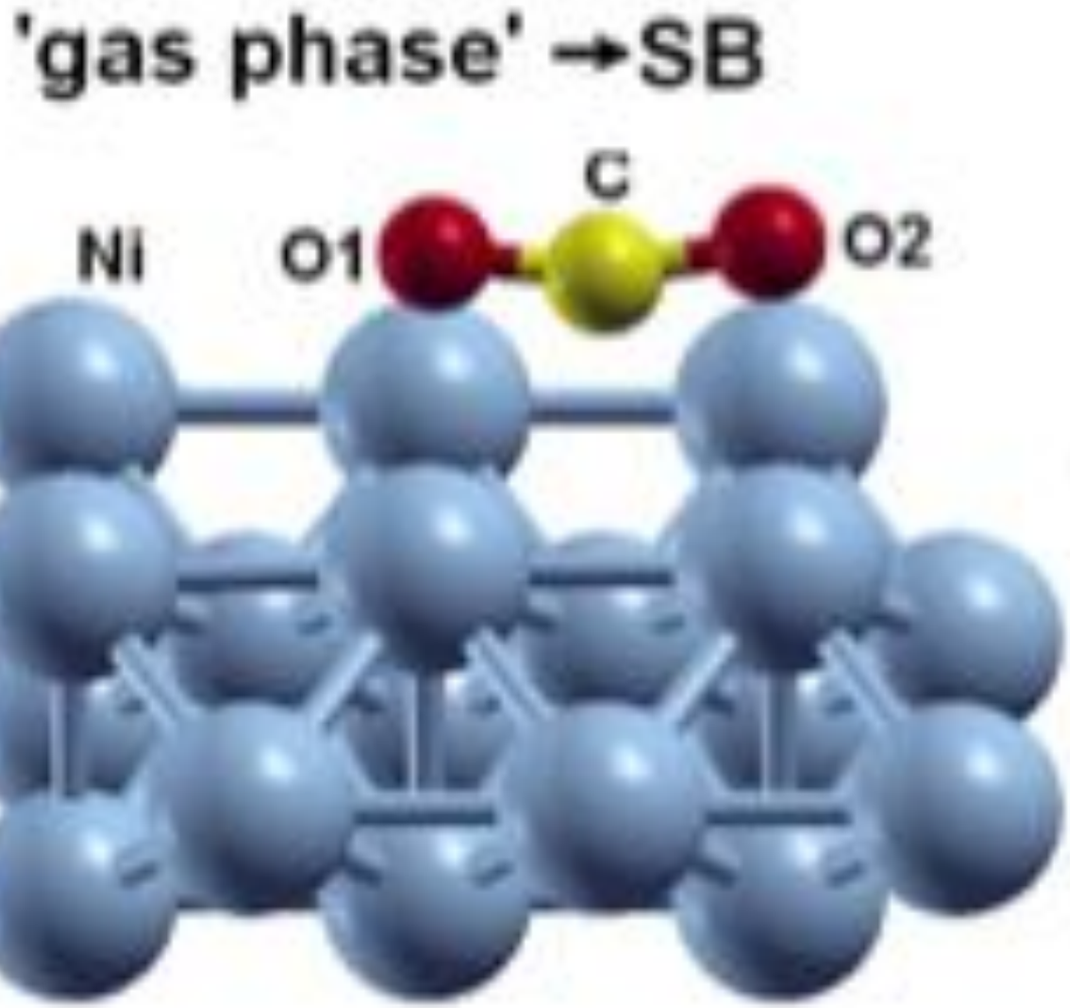
and we want to know what happens in between…

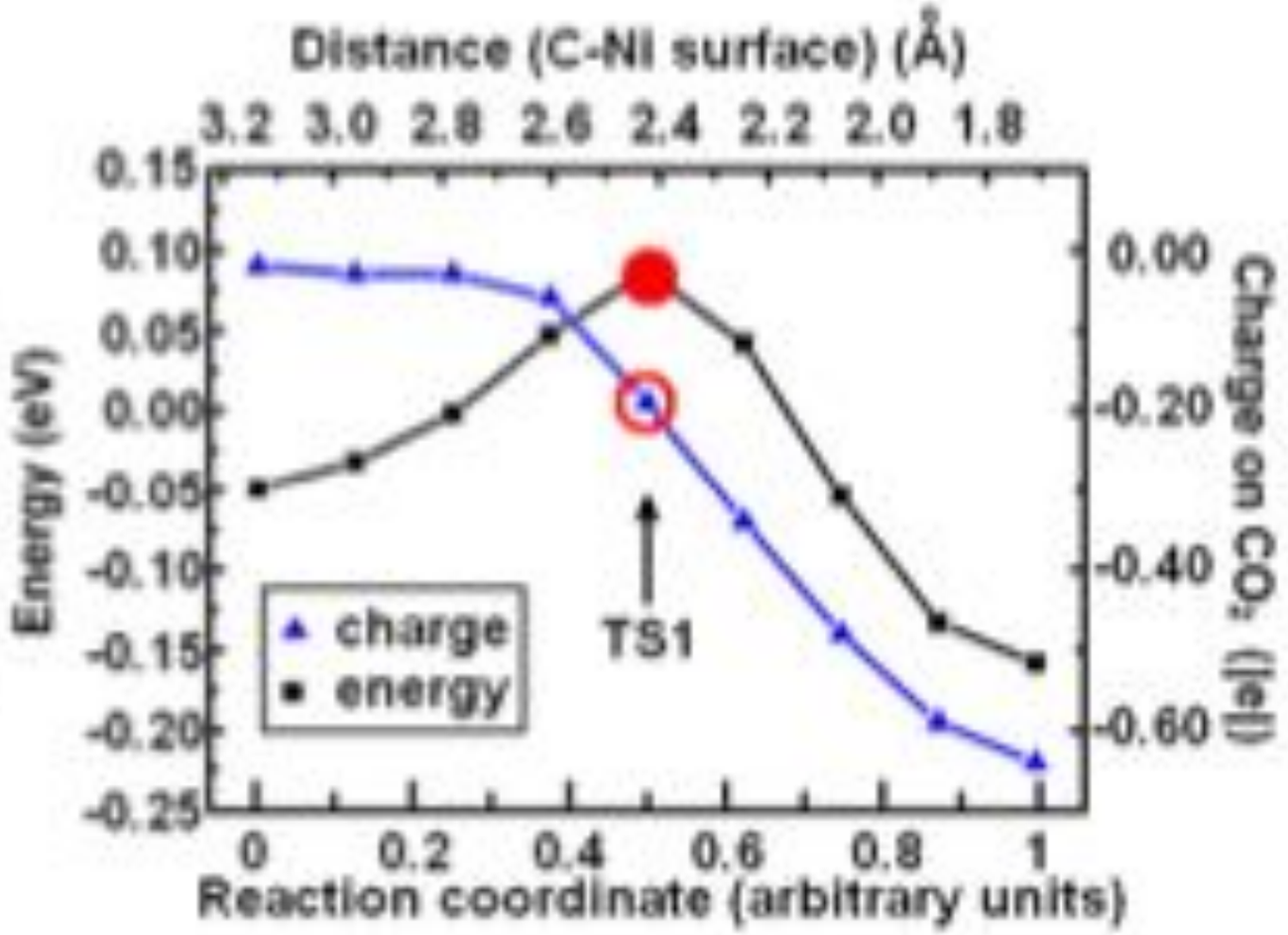
# a few examples



# a few examples

**activated adsorption of a molecule ($CO_2$) on a surface Ni(110)**

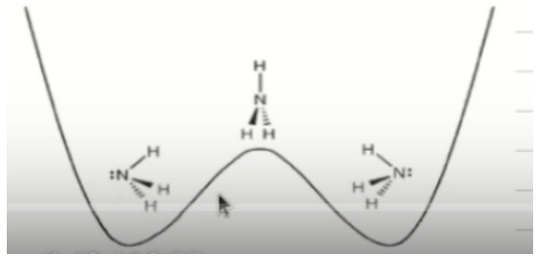


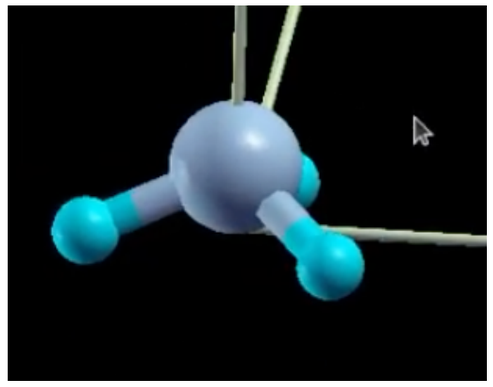



# a few examples

(other from literature)

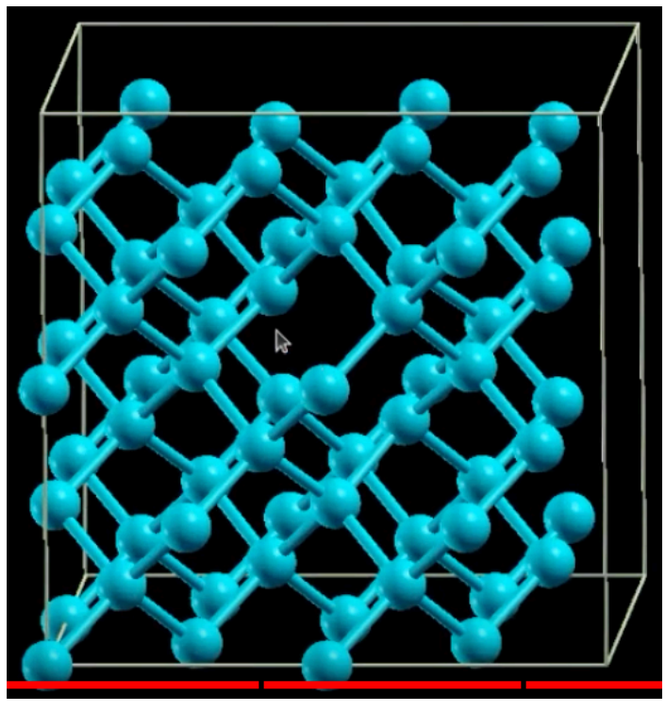
In some cases the TS is well defined by symmetry and barriers can be calculated “by hands” following physical intuition and studying directly the TS configuration (e.g. in this case: N in-plane with H3)



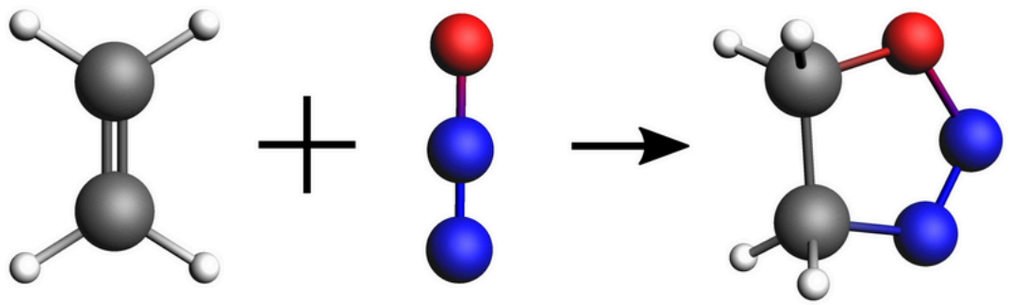


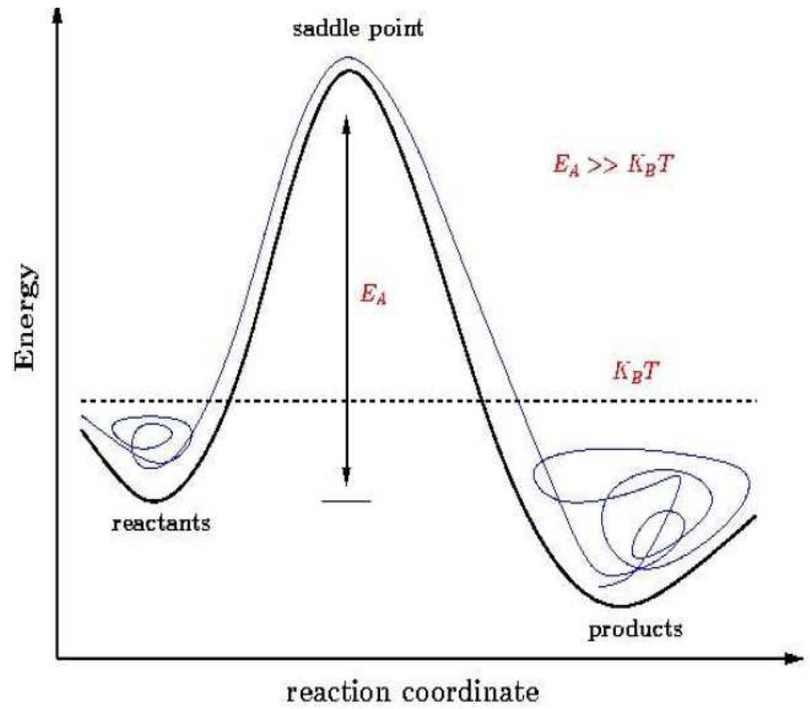
**$NH_3$ umbrella inversion barrier**

**diffusion of vacancies in silicon**



**chemical reactions**





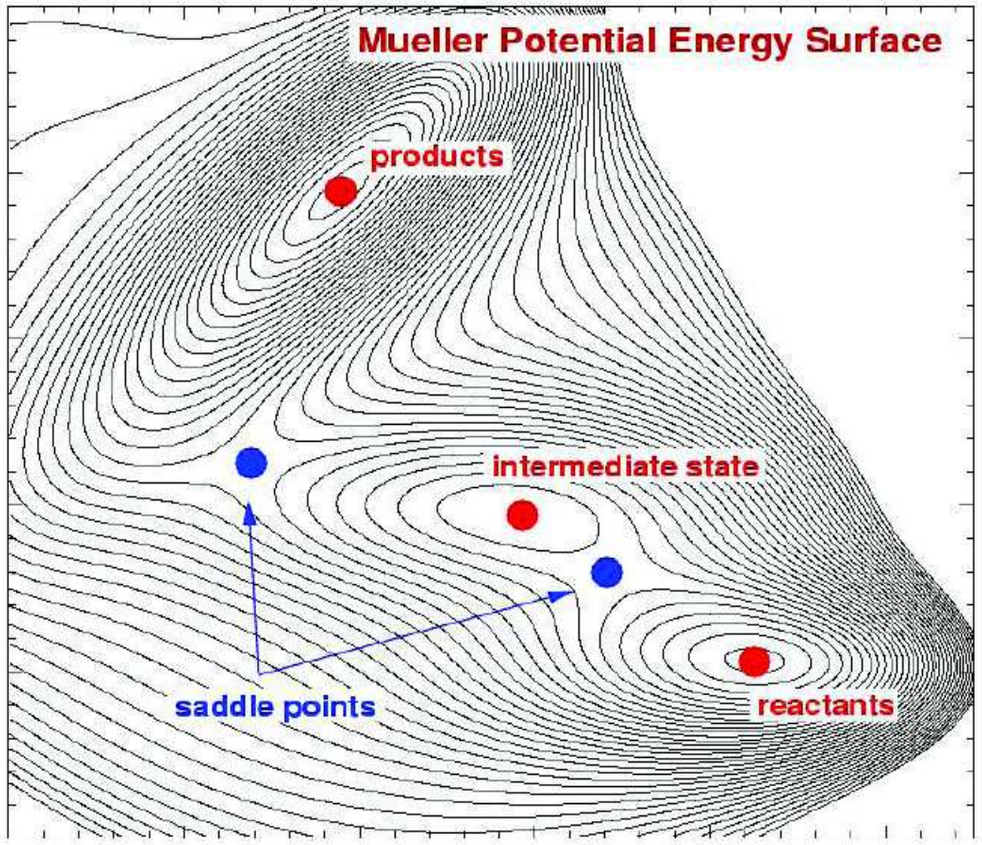

The transition probability can be estimated using equilibrium statistical mechanics.

Once the saddle point has been located we can use harmonic Transition State Theory (hTST) to calculate the rate constants:

$$K_{reactants \longrightarrow products} = \mathcal{A} \times e^{-\frac{E_A}{K_B T}}$$

Saddle points are unstable configurations and their location is a difficult task

# Saddle points in multidimensions: the Mueller Potential

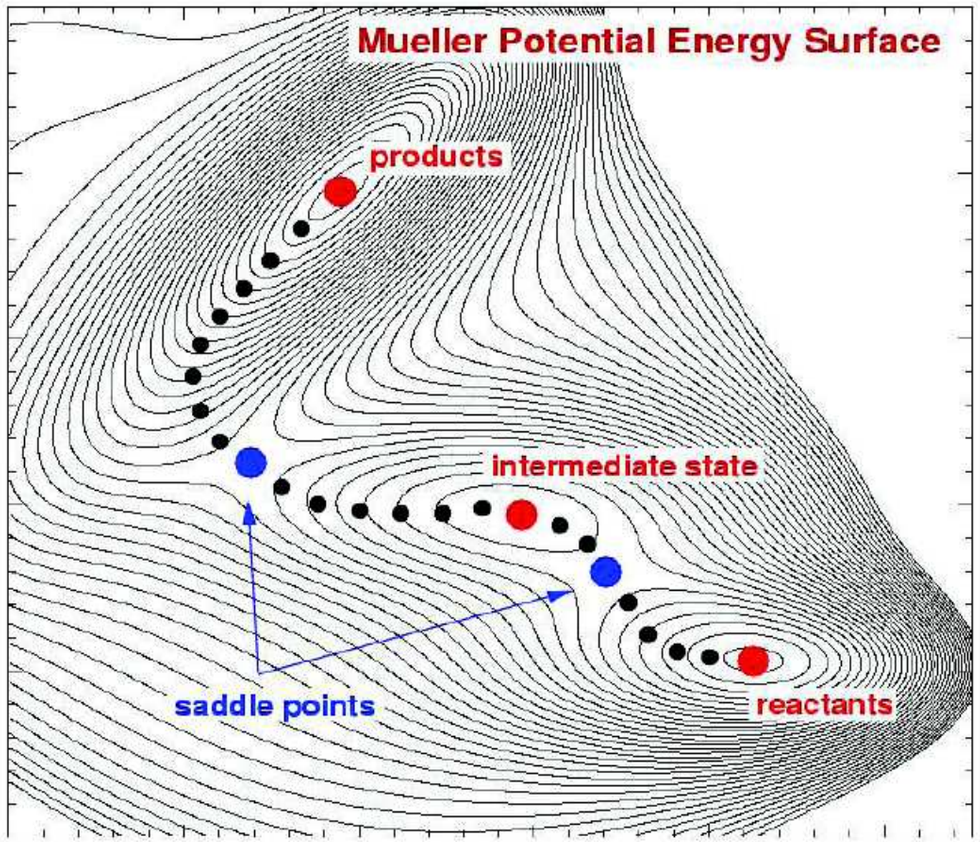



The path with the "highest" transition probability is the Minimum Energy Path.

MEP: the components of the force orthogonal to the path are zero.

$$-\left(\nabla V(x(s)) - \tau(s)\langle\tau(s)|\nabla V(x(s))\rangle\right) = 0$$

# Saddle points in multidimensions: the Mueller Potential



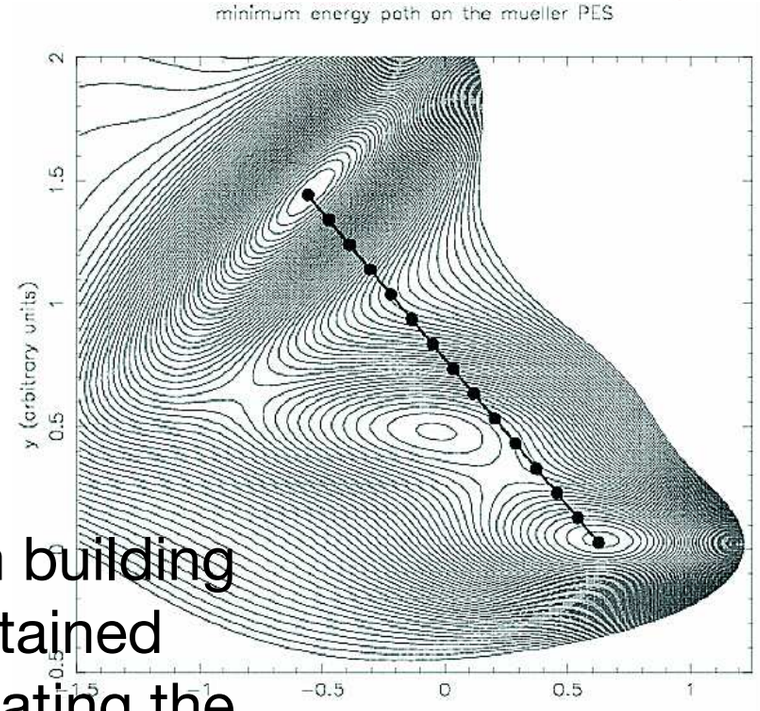

The path with the "highest" transition probability is the Minimum Energy Path.

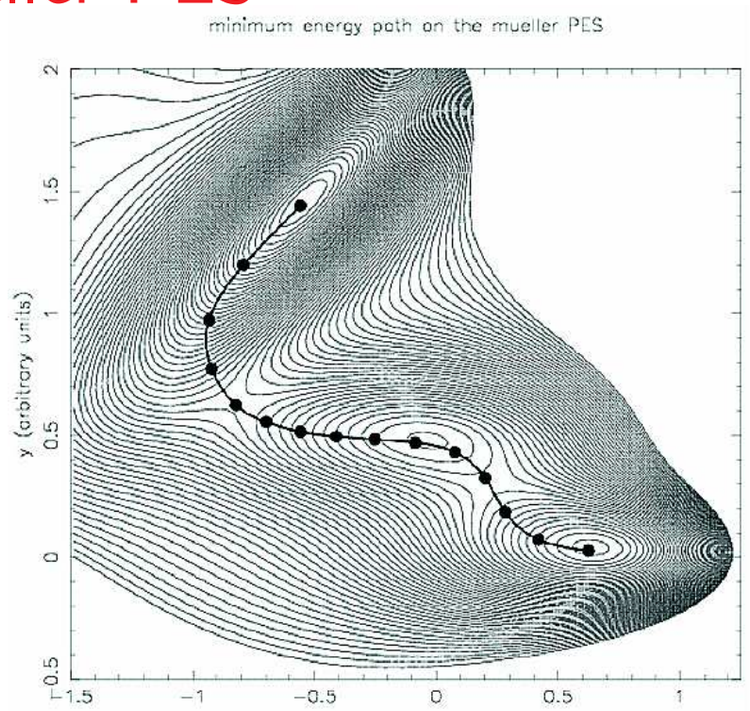
MEP: the components of the force orthogonal to the path are zero.

The MEP goes through the saddle point

Let us connect subsequent images by springs that only operate along the path

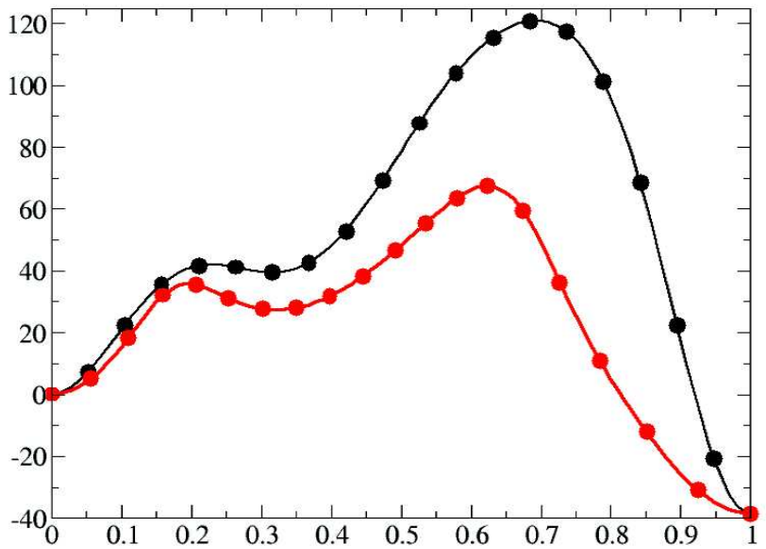
# NEB on the Mueller PES




starts from building images obtained by interpolating the first and the last

but we can supply also some intermediate images to build a better initial reaction path (if our physical/chemical intuition guides us…)




optimize the path on the PES by using “springs”

# nudged elastic bands

PWneb is part of the *Quantum ESPRESSO* distribution and uses the *PWscf* package as electronic-structure computing tools (“engine”)

executable : **neb.x**

# A few basic advices for NEB calculations

1. Don't use Climbing Image (CI) from the beginning. Converge your calculation, then restart from the last configuration with CI option enabled (the barrier will increase!)

2. Carefully choose the initial path. The code can make a linear interpolation of the atomic positions between consecutive input images

3. Try to start the NEB process with most atomic positions fixed, in order to converge the more "problematic" ones, before leaving all atoms move.

4. Especially for larger systems, you can start NEB with lower accuracy (less k-points, lower cutoff) and then increase it when it has converged to refine your calculation.

This example shows how to use neb.x to calculate the minimum energy path (MEP) of a simple activated reaction i.e. the collinear proton transfer reaction :

H2 + H  <==> H + H2

The MEP is obtained by means of the Climbing-Image Nudged Elastic Band (CI-NEB) method and two different climbing image algorithms are used ("auto" and "manual"). (for the meaning of the cited input variables see the Doc/INPUT_NEB* files)

The symmetric reaction path H2 + H  <==> H + H2 is calculated in three different ways.

1) The path connecting the initial and the final configurations is discretized with an odd number of images (7) so that the standard CI_scheme (“auto") will give rise to a symmetric MEP (3 images on the left of the saddle point and 3 images on the right). Note that in this system the use of the climbing image is not necessary. Indeed using CI_scheme = "no-CI" the result is the same.

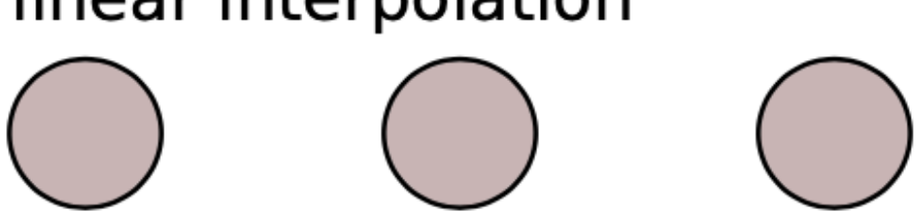
2) The path connecting the initial and the final configurations is discretized with an even number of images (8) and no climbing image is used. The resulting path is symmetric, but no image is at the saddle point.

3) The path connecting the initial and the final configurations is again discretized with an even number of images (8), but the "manual" CI_scheme is used so that the resulting path is asymmetric. The image 5 now is at the saddle point.

**Example 1a**

FIRST_IMAGE

linear interpolation

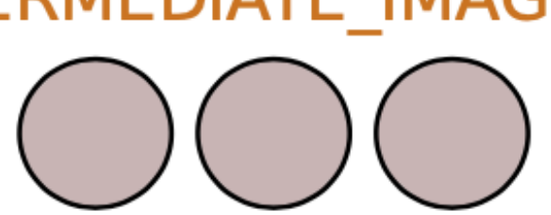


LAST_IMAGE

**Example 1b**

FIRST_IMAGE

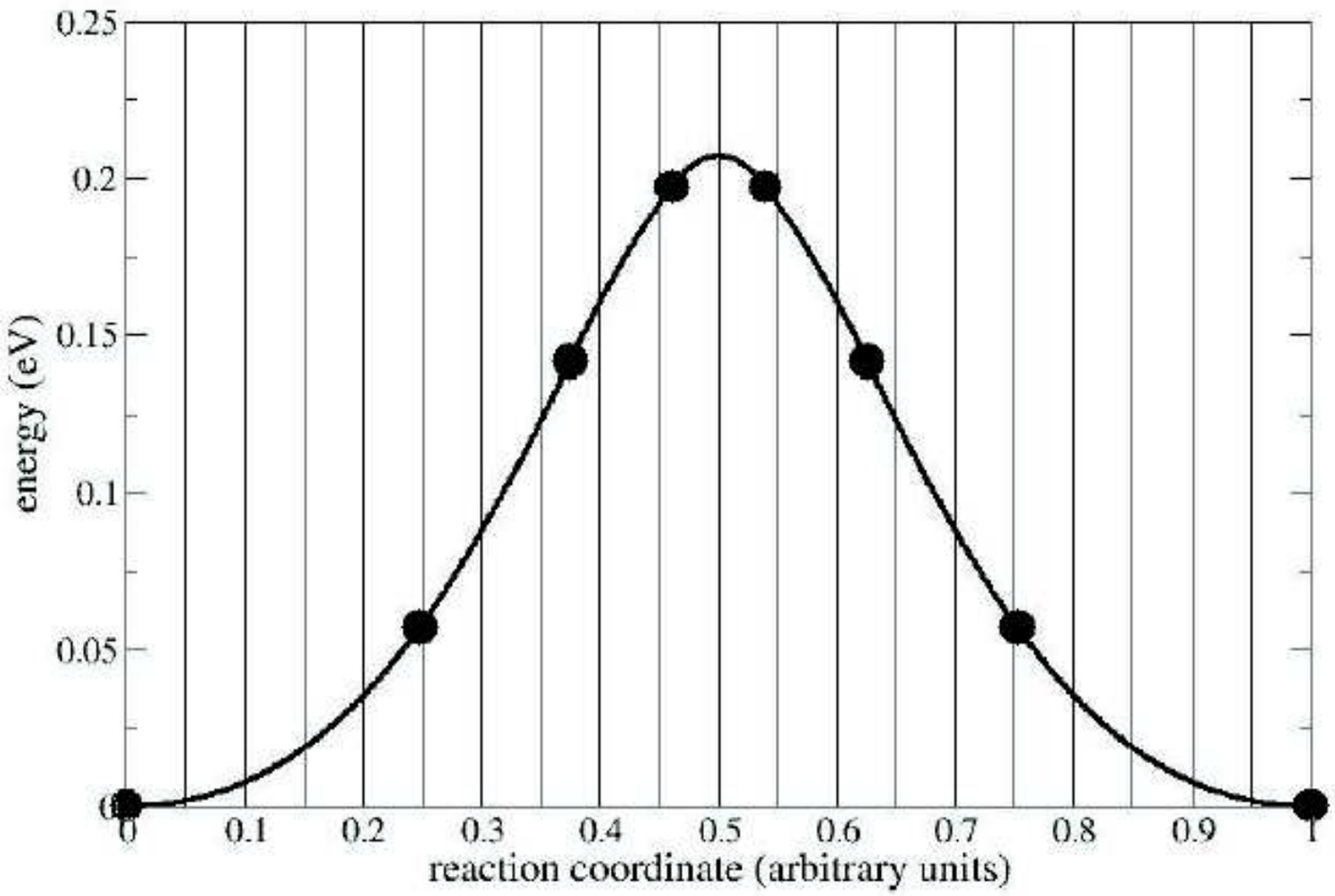
INTERMEDIATE_IMAGE



LAST_IMAGE

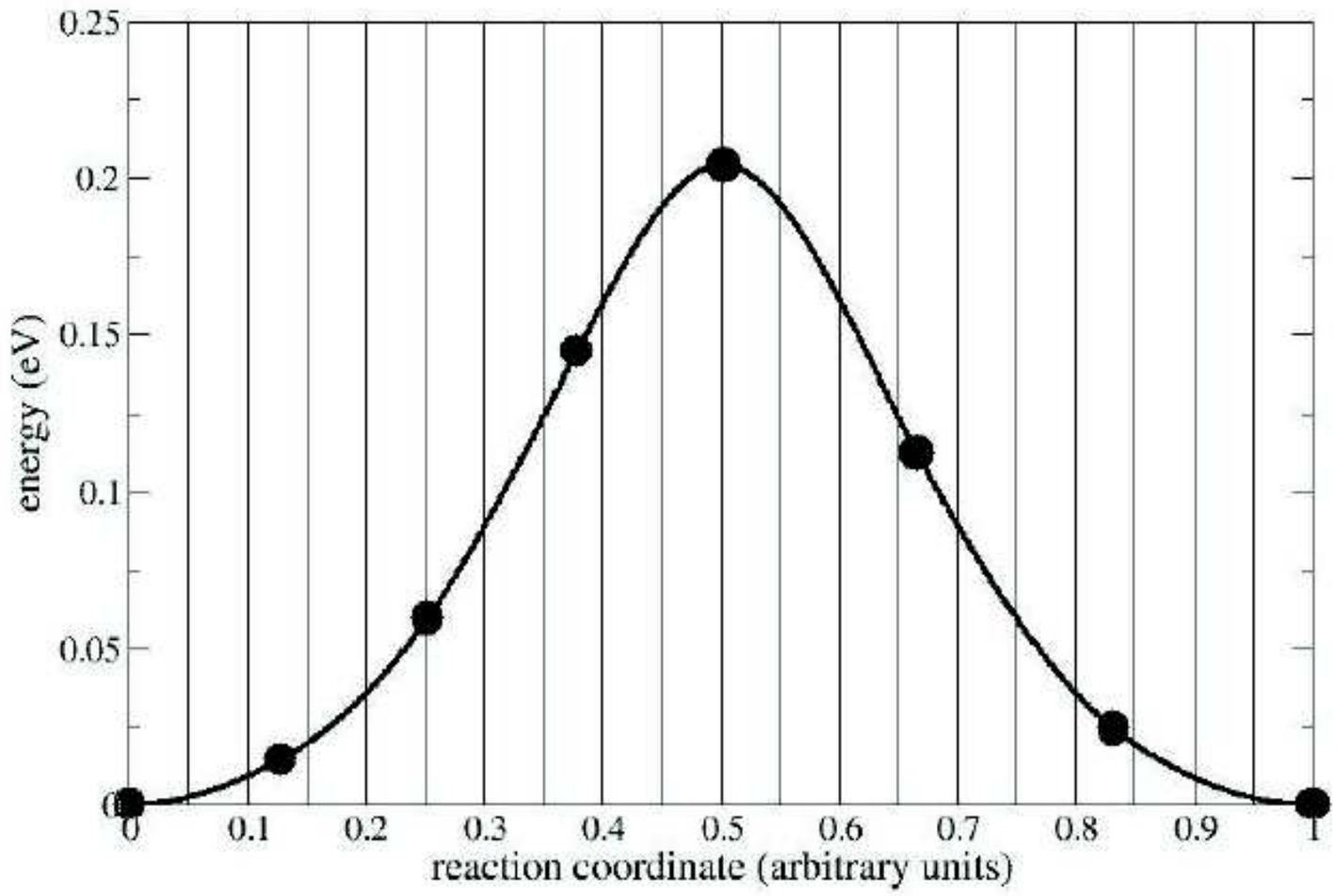
# Example : collinear proton transfer

## variable elastic constants

# Example : collinear proton transfer

## climbing image (manual on image 5)

# Born-Oppenheimer Molecular Dynamics

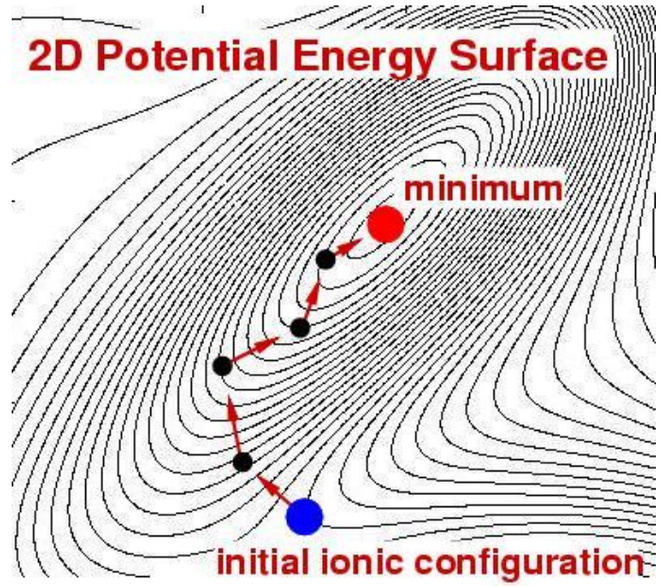
(but we don’t afford it here)

# Something else with forces...

## Structural Optimization and Molecular Dynamics

Within the *Born-Oppenheimer*, or *adiabatic* approximation, the total energy as a function of atomic positions, or *Potential Energy Surface* (PES), determines the behaviour of nuclei.

The *global* ground state can be found by minimizing the function $E(\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_N)$, depending upon the $3N$ atomic coordinates for a system of $N$ atoms. This is a "standard" mathematical problem: finding the minimum of a function, knowing its derivatives, that is, the Hellmann-Feynman forces (in the picture, a cartoon of a PES in two dimensions with the path to the minimum).




Once forces are calculated, one can perform not only structural optimization, but also *molecular dynamics*. If a classical behaviour of the nuclei is assumed, all the machinery of classical MD can be recycled, with forces calculated from *first principles*.

# Born-Oppenheimer Molecular Dynamics

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

$$L = \frac{1}{2}\sum_{\mu} M_\mu \dot{\mathbf{R}}_\mu^2 - E(\mathbf{R})$$

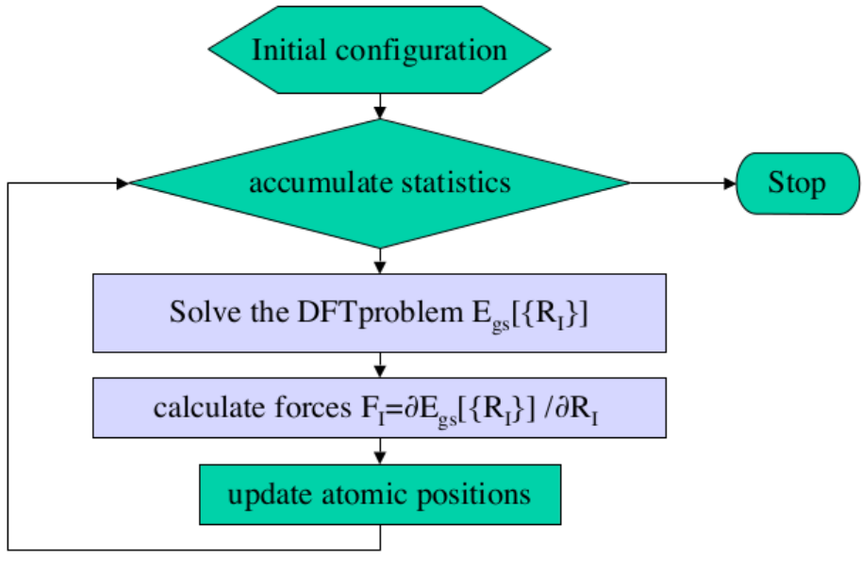
describing the motion of nuclei. The equations of motion:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{R}}_\mu} - \frac{\partial L}{\partial \mathbf{R}_\mu} = 0, \qquad \mathbf{P}_\mu = \frac{\partial L}{\partial \dot{\mathbf{R}}_\mu}$$

are nothing but usual Newton's equations:

$$\mathbf{P}_\mu \equiv M_\mu \mathbf{V}_\mu, \qquad M_\mu \dot{\mathbf{V}}_\mu = \mathbf{F}_\mu,$$

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*:

$$\begin{aligned}
\mathbf{R}_\mu(t+\delta t) &= 2\mathbf{R}_\mu(t) - \mathbf{R}_\mu(t-\delta t) + \frac{\delta t^2}{M_\mu}\mathbf{F}_\mu(t) + \mathcal{O}(\delta t^4) \\
\mathbf{V}_\mu(t) &= \frac{1}{2\delta t}\left[\mathbf{R}_\mu(t+\delta t) - \mathbf{R}_\mu(t-\delta t)\right] + \mathcal{O}(\delta t^3).
\end{aligned}$$

or the *Velocity Verlet*:

$$\begin{aligned}
\mathbf{V}_\mu(t+\delta t) &= \mathbf{V}_\mu(t) + \frac{\delta t}{2M_\mu}\left[\mathbf{F}_\mu(t) + \mathbf{F}_\mu(t+\delta t)\right] \\
\mathbf{R}_\mu(t+\delta t) &= \mathbf{R}_\mu(t) + \delta t\mathbf{V}_\mu(t) + \frac{\delta t^2}{2M_\mu}\mathbf{F}_\mu(t).
\end{aligned}$$

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