

# Chimica computazionale

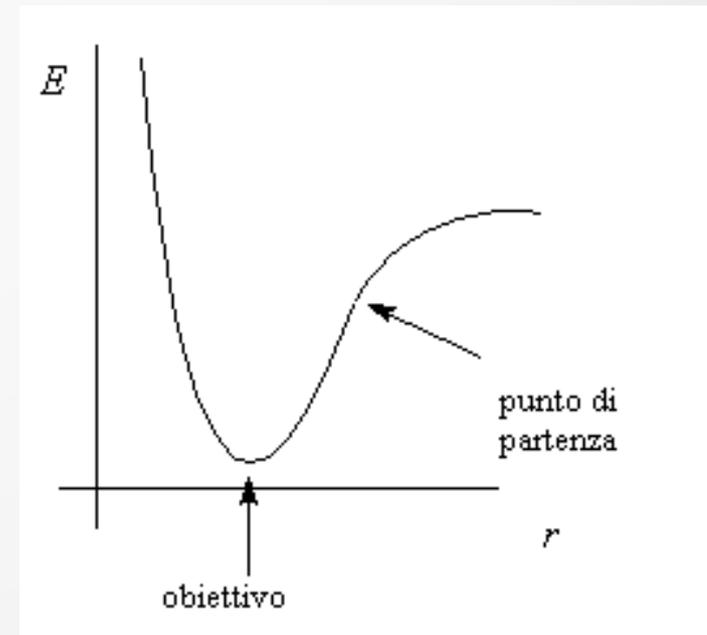
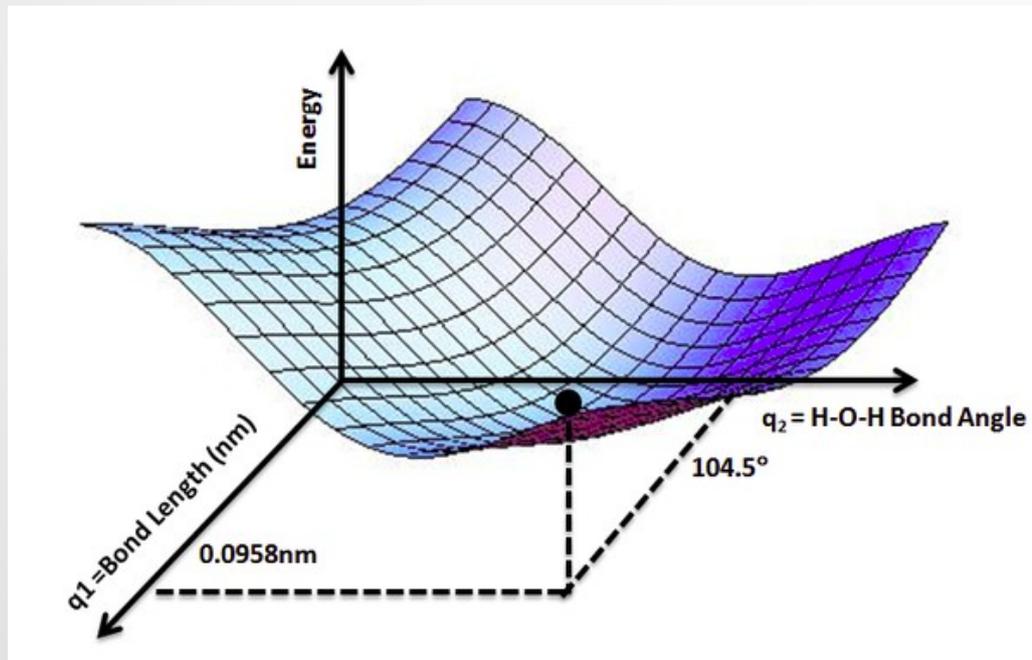
Loriano Storchi

[loriano@storchi.org](mailto:loriano@storchi.org)

<http://www.storchi.org/>

# PES

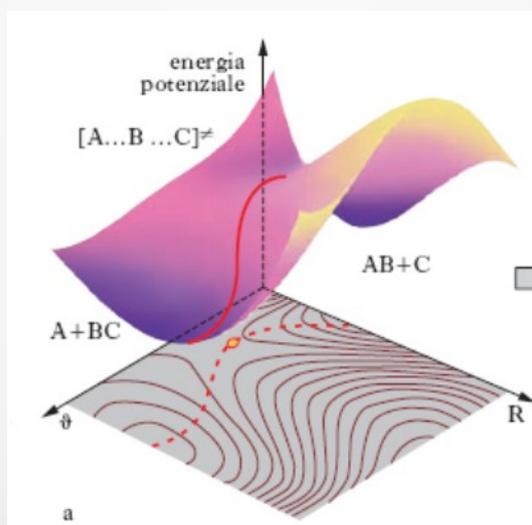
- **PES (Potential Energy Surface)** superficie di energia potenziale . E' in generale una ipersuperficie data dal valore dell'energia potenziale  $U$  di un sistema molecolare in funzione della posizione relativa dei vari atomi.



Molecola biatomica

# PES – punti importanti

- **Minimi locali:** corrispondono a geometrie molecolari ottimali
- **Punti sella:** dove la superficie ha una concavità verso il basso per una delle coordinate e verso l'alto per tutte le altre. Essi rappresentano le barriere di minima energia in un percorso che porta da un minimo locale all'altro.



# PES – punti importanti

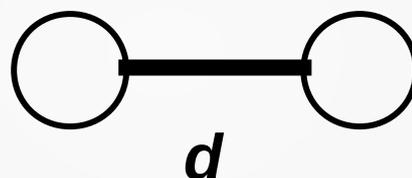
- Sia minimi locale che punti sella hanno la derivata prima di  $U$  rispetto alle coordinate uguale a 0

$$\frac{\partial U}{\partial q_1} = \frac{\partial U}{\partial q_2} = \dots = \frac{\partial U}{\partial q_n} = 0$$

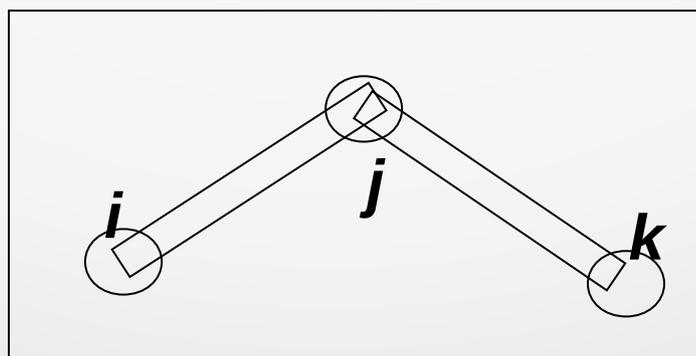
- Coordinate cartesiane: le abbiamo incontrate, e banalmente si specifica per ogni atomo  $X, Y, Z$ , quindi  $3N$  valori

# PES – coordinate interne

- **Coordinate interne (matrice  $z$ )** : in pratica le coordinate di ogni atomo sono specificate in funzione di un qualche atomo di riferimento
- Per due atomi ad esempio basta specificare la distanza di legame

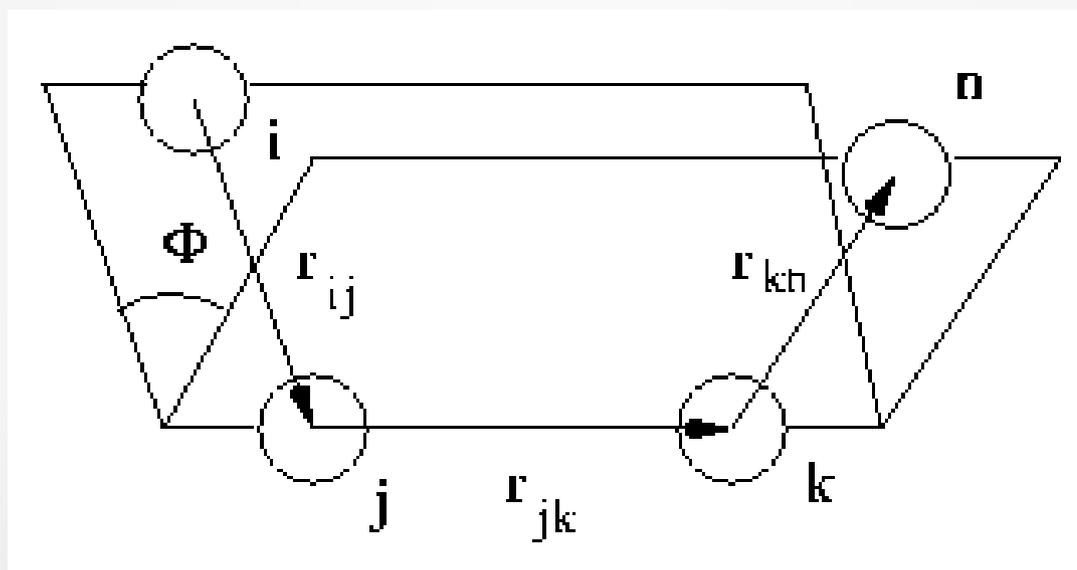


- Nel caso di tre atomi occorre specificare anche l'angolo di legame



# PES – coordinate interne

- Nel caso di quattro atomi o piu' serve specificare anche l'angolo diedro (o di torsione), cioe' nel caso di 4 atomi i-j-k-n è definito come l'angolo tra i due piani, contenenti, uno gli atomi i, j e k e l'altro gli atomi j, k e n.

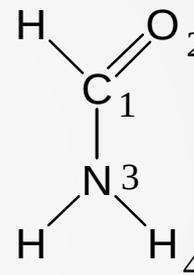


# PES – punti importanti

- Le coordinate interne sono in genere scritte in quella che è definita **matrice Z**. La matrice Z contiene una linea per ogni atomo del sistema.

```
1 C
2 O 1 rCO
3 N 1 rCN  2 aNCO
4 H 3 rNHa 1 aCNHa 2 0.0
5 H 3 rNHb 1 aCNHb 2 180.0
6 H 1 rCH  2 aHCO  4 180.0
```

```
rCO=1.1962565
rCN=1.3534065
rNHa=0.9948420
rNHb=0.9921367
rCH=1.0918368
aNCO=124.93384
aCNHa=119.16000
aCNHb=121.22477
aHCO=122.30822
```



- Esclusi i primi tre atomi tutti hanno 3 coordinate (3N-6) ogni atomo ha quindi una distanza e due angoli riferiti rispetto ai tre atomi precedenti

# PES – ottimizzazione di geometria

- Dato un sistema di N atomi e date le loro coordinate, il calcolo del valore di U (ad esempio mediante il metodo HF) viene chiamato **single point**
- Uno dei primi passaggi in molte procedure computazionali invece è proprio la **determinazione della geometria ottimale**, cioè di quella corrispondente al **punto di minimo della PES**
- Quindi la **geometry optimization** è proprio la procedura che porta alla determinazione della geometria ottimale, e quindi del punto di minimo della PES. Sarà poi possibile ad esempio operare un calcolo di alcune proprietà molecolari oppure ...

# Ottimizzazione di geometria

- Nell'ottimizzazione di geometria si usa molto spesso il metodo **Newton-Raphson**
- **Dato un set di coordinate  $q^K$  questo ci permette di calcolare  $U(q^K)$  e di stimare, mediante l'espansione di Taylor  $U$  al punto  $q^{K+1}$**
- Espansione di Taylor ci permette di esprimere una funzione nell'intorno di un punto come un polinomio dipendente dalla derivate della funzione stessa.

Sia  $f : [a, b] \rightarrow \mathbb{R}$  una funzione, sia  $x_0 \in (a, b)$  e supponiamo che esistano le derivate  $f^{(1)}(x_0), f^{(2)}(x_0), \dots, f^{(n-1)}(x_0)$ . Preso  $h$  tale che  $f$  sia definita in  $[x_0 - h, x_0 + h]$  (intorno chiuso di centro  $x_0$  e raggio  $h$ ), vale la formula

$$f(x) = \sum_{i=0}^{n-1} \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i + R_n(x)$$

prendendo  $x = x_0 + h$

# Ottimizzazione di geometria

- Chiaramente possiamo usare l'espansione in serie di Taylor anche per  $U(q)$ . In particolare se ci fermiamo al termine quadratico:

$$U(q^{K+1}) = U(q^K) + (q^{K+1} - q^K) g^K + \frac{1}{2} (q^{K+1} - q^K) H^K (q^{K+1} - q^K)$$

$$g^K = \begin{bmatrix} \delta U / \delta q_1 \\ \vdots \\ \delta U / \delta q_n \end{bmatrix}_{q=q^K} \quad H^K = \begin{bmatrix} \delta^2 U / \delta q_1 \delta q_1 & \delta^2 U / \delta q_1 \delta q_2 & \cdots & \delta^2 U / \delta q_1 \delta q_n \\ \vdots & \vdots & \vdots & \vdots \\ \delta^2 U / \delta q_n \delta q_1 & \delta^2 U / \delta q_n \delta q_2 & \cdots & \delta^2 U / \delta q_n \delta q_n \end{bmatrix}_{q=q^K}$$

# Ottimizzazione di geometria

- Quindi l'**energia potenziale** in  $q^{K+1}$  dipende da:
  - Ovviamente **dall'energia potenziale** in  $q^K$
  - **Dalla matrice delle derivate prime**  $g^K$ , sempre calcolata in  $q^K$
  - **dall'Hessiana**, dalla **matrice delle derivate seconde** , sempre calcolata in  $q^K$
- Visto che cerchiamo un minimo di  $U(q)$  **imponendo che in  $q^{K+1}$  la derivata prima sia nulla** otteniamo:

$$q^{K+1} = q^K - (H^K)^{-1} g^K$$

- Dove  $q^{K+1}$  e' il punto di minimo

# Ottimizzazione di geometria

- Visto che cerchiamo un minimo di  $U(q)$  imponendo che in  $q^{K+1}$  la derivata prima sia nulla otteniamo:

$$f_T(x) = f_T(x_n + \Delta x) \approx f(x_n) + f'(x_n)\Delta x + \frac{1}{2}f''(x_n)\Delta x^2.$$

$$0 = \frac{d}{d\Delta x} \left( f(x_n) + f'(x_n)\Delta x + \frac{1}{2}f''(x_n)\Delta x^2 \right) = f'(x_n) + f''(x_n)\Delta x$$

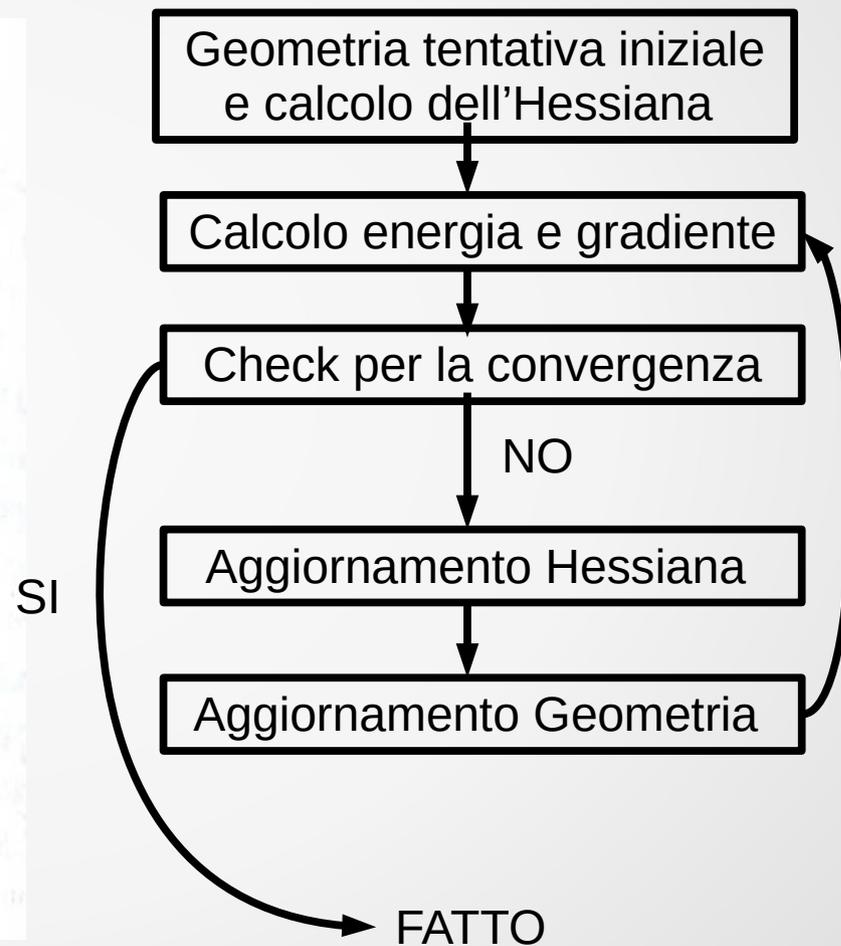
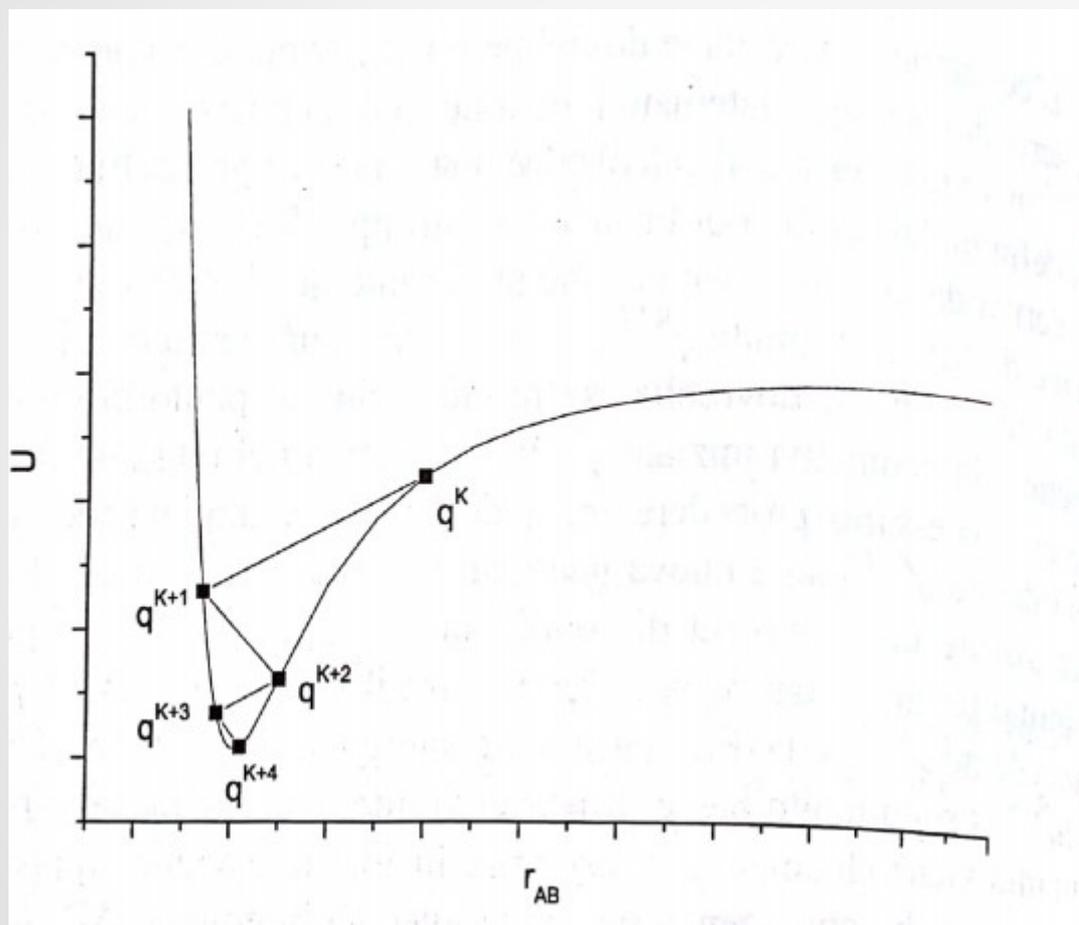
$$\Delta x = -\frac{f'(x_n)}{f''(x_n)}.$$

# Ottimizzazione di geometria

$$q^{K+1} = q^K - (H^K)^{-1} g^K$$

- Se questa equazione non derivasse da **un'approssimazione fatta “tagliando” la serie di Taylor** al termine quadratico, ci permetterebbe di ottenere il punto di minimo data una geometria iniziale  $q^K$  ci permetterebbe di conoscere la geometria di minimo
- **In pratica anche qui arriviamo ad una procedura iterativa. Partendo da un set di coordinate calcolo quelle nuove. Calcolo la nuova energia potenziale e cosi' via fino che la differenza fra le energia e le coordinate ad un passo  $n$  e' poco dirrente rispetto a quelle al passo precedente  $n-1$**

# Ottimizzazione di geometria



# Metodo Newton-Raphson alcune considerazioni

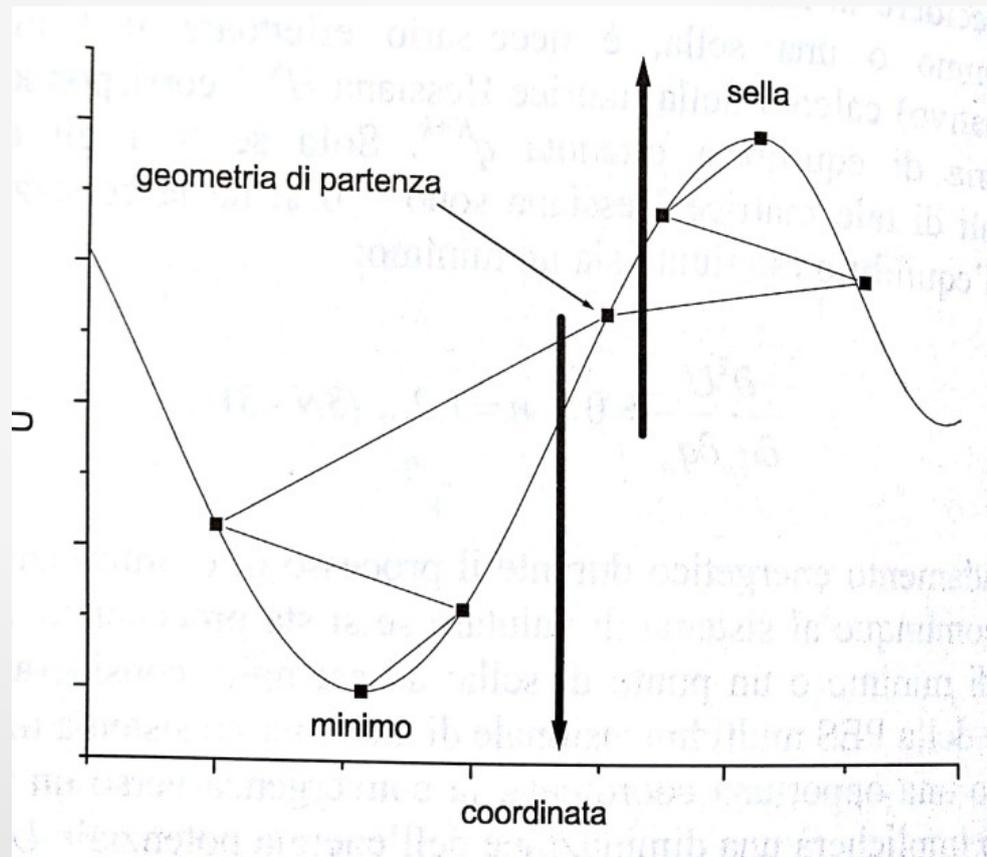
- Una delle **parti piu' onerose della procedura e' la valutazione dell'Hessiana**. Si possono pero' usare alcune approssimazioni (pseudo-Newton-Raphson) non al primo ciclo.
- La **matrice Hessiana contiene  $n^2$  elementi può essere tuttavia approssimata usando la matrice delle derivate prime al punto  $q^K$  e  $q^{K-1}$**

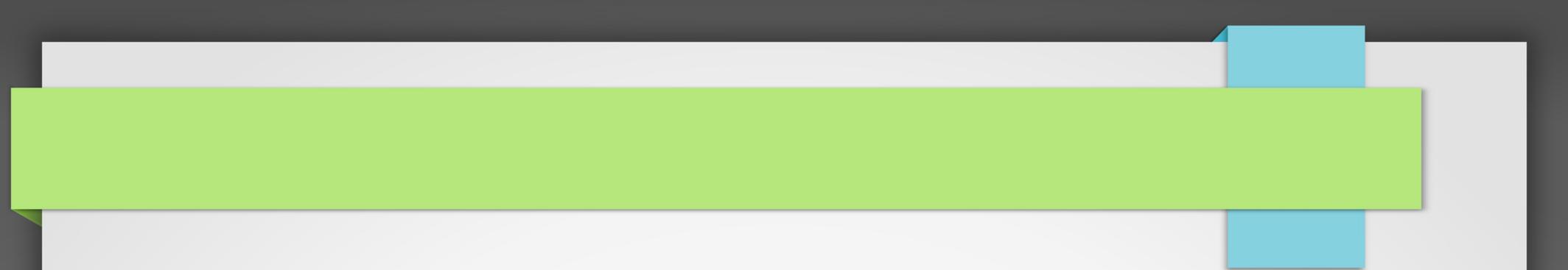
# Metodo Newton-Raphson alcune considerazioni

- Visto che anche i punti sella hanno derivate nulle, per essere certi di aver raggiunto un minimo si devono valutare gli elementi diagonali dell'Hessiana (minimi o punti sella ?). **Minimo se tutti gli elementi diagonali sono  $> 0$ .**
- Ovviamente per capire se ci si sta muovendo o meno verso un minimo si deve **vedere l'andamento dell'energia**

# Metodo Newton-Raphson alcune considerazioni

- Non e' detto che il metodo ci porti verso un minimo assoluto (minimi locali) dipendentemente dalla geometria di partenza





DERIVATE

# Le derivate

- Data una funzione  $f(x)$  come calcolo la derivata ? Parto dalla definizione:

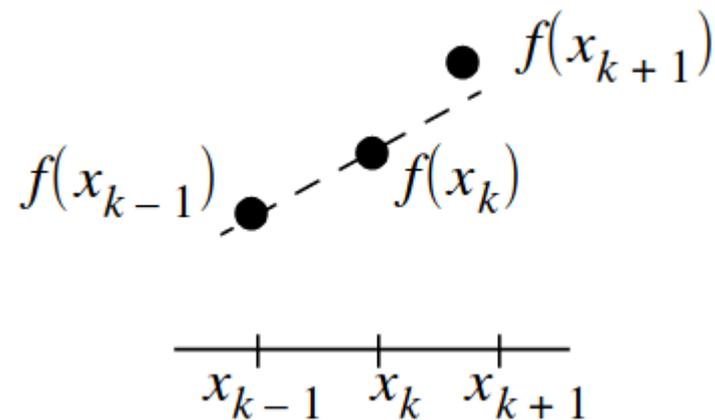
$$f'(x) = \frac{df(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{(x + \Delta x) - x}$$

- Numericamente quindi possiamo approssimare una derivata (pendenza della curva) semplicemente determinando i valori della funzione in punti vicini

# Le derivate

- Backward difference

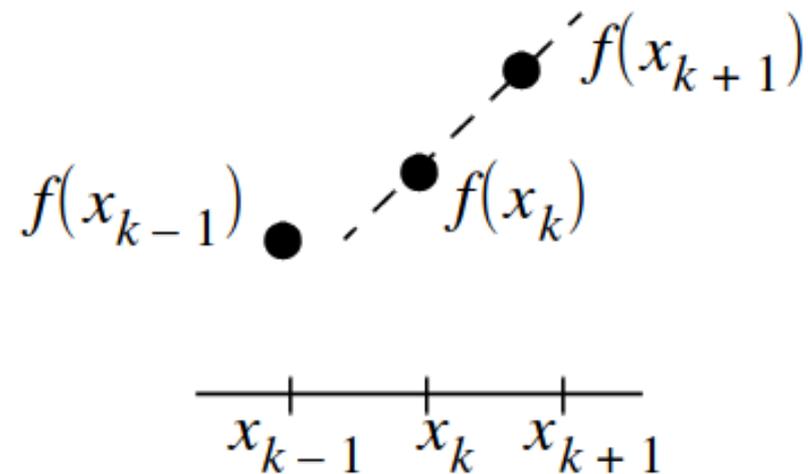
$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$



# Le derivate

- Forward difference

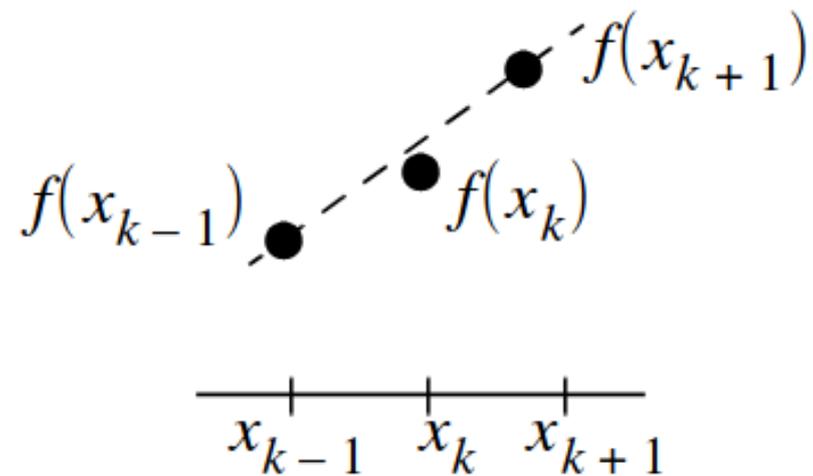
$$f'(x_k) \approx \frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k}$$



# Le derivate

- Central difference

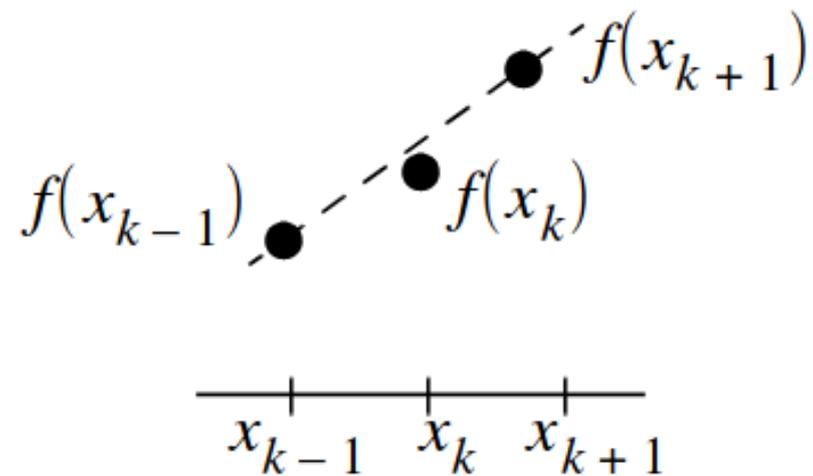
$$f'(x_k) \approx \frac{f(x_{k+1}) - f(x_{k-1}))}{x_{k+1} - x_{k-1}}$$



# Le derivate

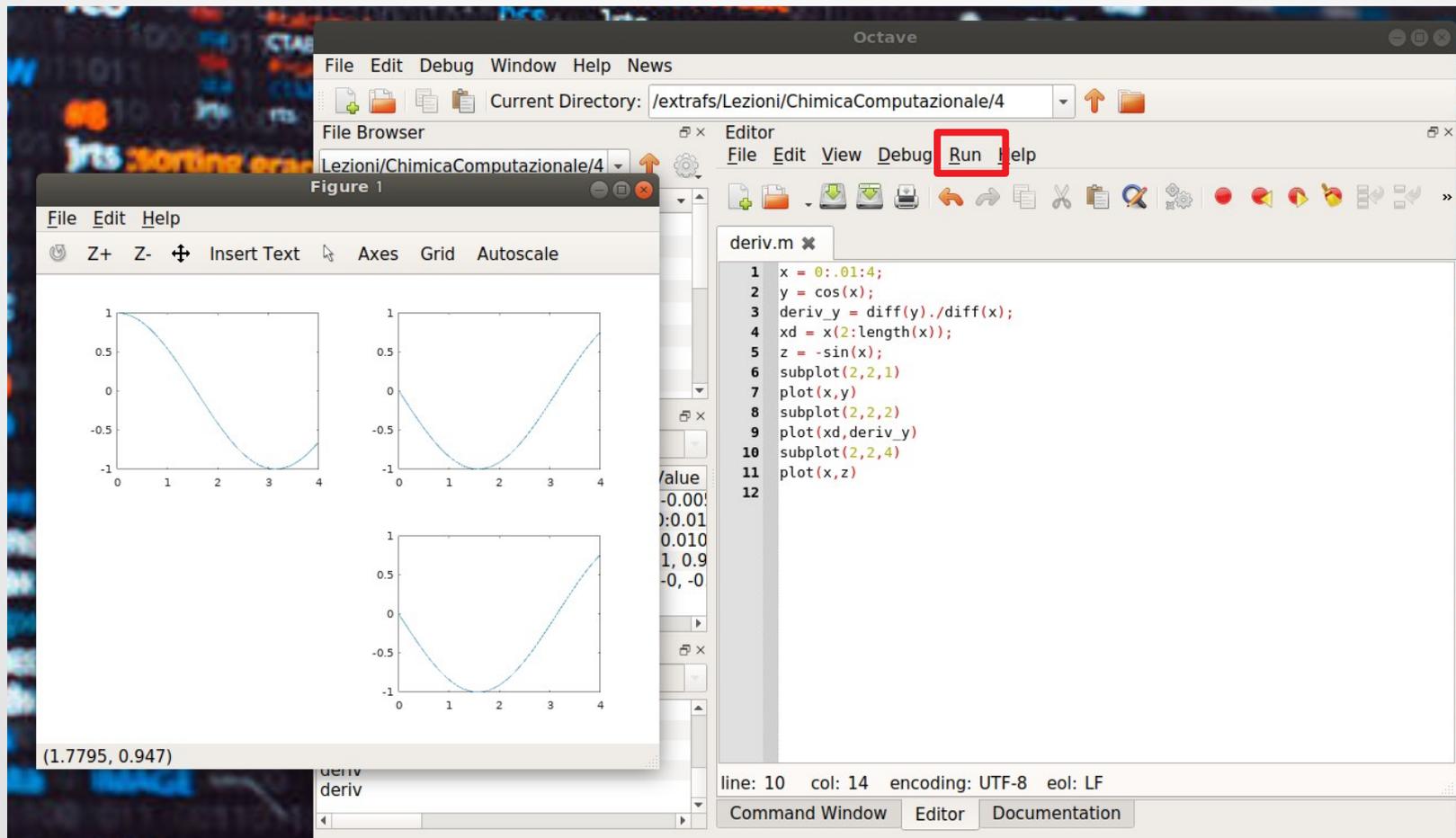
- Central difference

$$f'(x_k) \approx \frac{f(x_{k+1}) - f(x_{k-1}))}{x_{k+1} - x_{k-1}}$$



# Le derivate

- Ad esempio la derivata della funzione  $\cos(x)$  fatta con octave:



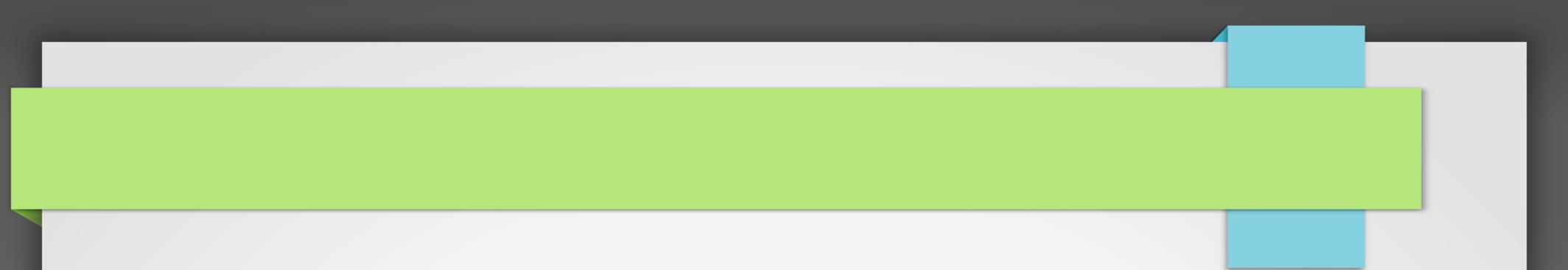
# Le derivate

- La derivata seconda chiaramente:

$$\frac{d^2f}{dx^2} = \lim_{\Delta x \rightarrow 0} \frac{f'(x + \Delta x) - f'(x)}{\Delta x}$$

- E sostituendo le differenze forward e backward per stimare  $f'(x+\Delta x)$  ed  $f'(x)$  avremo:

$$f''(x) \approx \frac{D_+(h) - D_-(h)}{h} = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$



# SIMMETRIA MOLECOLARE

# Simmetria molecolare

- La **classificazione delle molecole in base al gruppo di simmetria** di appartenenza e' un aspetto estremamente importante in molti aspetti e metodo della chimica computazionale.
- **Fra cui come vedremo anche nel caso dell'ottimizzazione di geometria**, che nel caso di sistemi grandi e quindi molti gradi di liberta' puo' diventare un problema computazionalmente oneroso
- La simmetria molecolare descrive le simmetrie strutturali presenti in una molecola ed e' importante in tanti aspetti

# Simmetria molecolare

- Si possono dunque classificare le molecole in base ai loro elementi di simmetria usando la **teoria dei gruppi**
- **Teoria dei gruppi** : una parte della matematica che si occupa dei gruppi. Una **struttura algebrica (insieme S piu' alcune operazioni)** caratterizzata da:
  - una operazione binaria associativa
  - dotata di elemento neutro
  - Ogni elemento e' dotato del suo inversoe' un gruppo
- **Esempio numeri interi con l'addizione**

# Simmetria molecolare

- Esempio

- una operazione binaria associativa:  $(x * y) * z = x * (y * z)$  per ogni  $x, y, z$  in  $S$ .
- dotata di elemento neutro:  $a * u = a$
- Ogni elemento  $e$  è dotato del suo inverso:  $h$  è elemento inverso di  $g$  se  $g * h = h * g = u$  dove  $u$  è l'elemento neutro

# Simmetria molecolare

- Le **applicazioni** della simmetria molecolare (teoria dei gruppi) in ambito chimico sono notevoli:
  - **Previsioni di proprietà chimiche** come la chiralità ad esempio
  - La simmetria influenza ovviamente il **numero di isomeri**
  - Determinazione del **momento di dipolo**
  - Prevedere quali **transizioni spettroscopiche** siano consentite o meno
  - **Trattamento quanto meccanico della struttura elettronica e del legame chimico**

# Simmetria molecolare

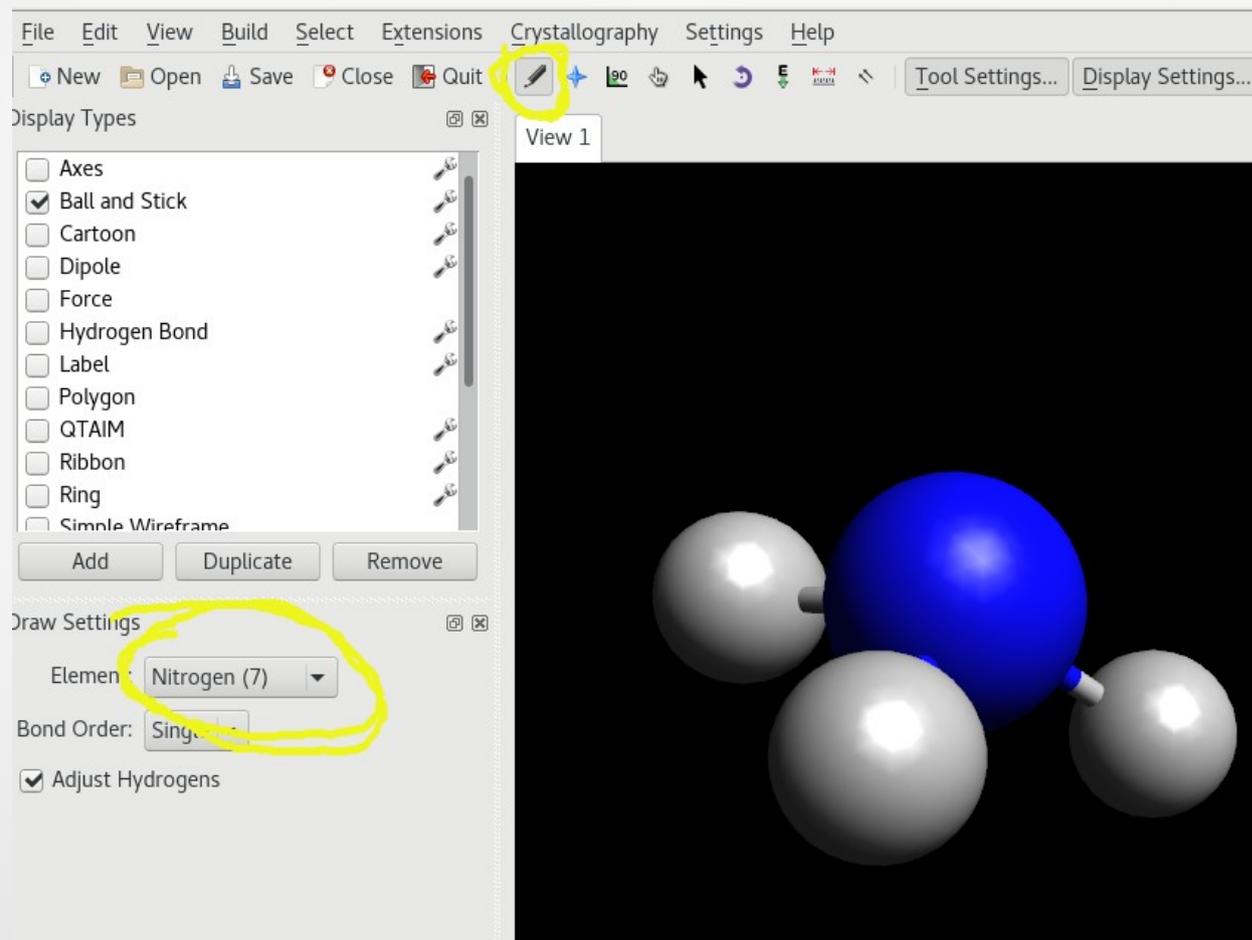
- Esistono **5 operazioni di simmetria fondamentali**:
  - **Identità' (E)**: trasforma ogni atomo in se stesso
  - **Rotazione propria ( $C_n$ )** : rotazione attorno ad un dato asse di angoli  $2\pi/n$
  - **Riflessione ( $\sigma$ )** : riflessione rispetto ad un piano di simmetria
  - **Inversione (i)** : inversione rispetto ad un punto centrale
  - **Rotoriflessione (rotazione impropria) ( $S_n$ )**: In pratica una rotazione propria  $C_n$  seguita da una riflessione rispetto ad un piano ortogonale all'asse (asse della rotazione)

# Simmetria molecolare

- Dunque come classifichiamo una molecola:
  - **Una molecola e' simmetrica rispetto ad una data trasformazione se applicata tale operazione di simmetria il trasformato e' identico alla molecola di partenza**
  - **Una molecola appartiene ad un gruppo se questo e' il gruppo di ordine massimo fra i gruppi di trasformazioni di simmetria per cui la molecola e' simmetrica**

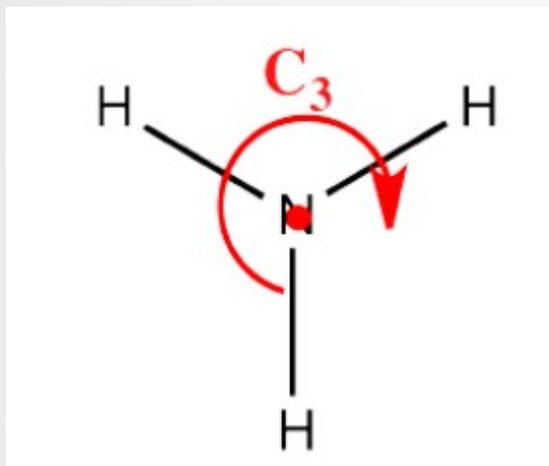
# Simmetria molecolare

- Esempio ammoniaca (disegnatela con avogadro):



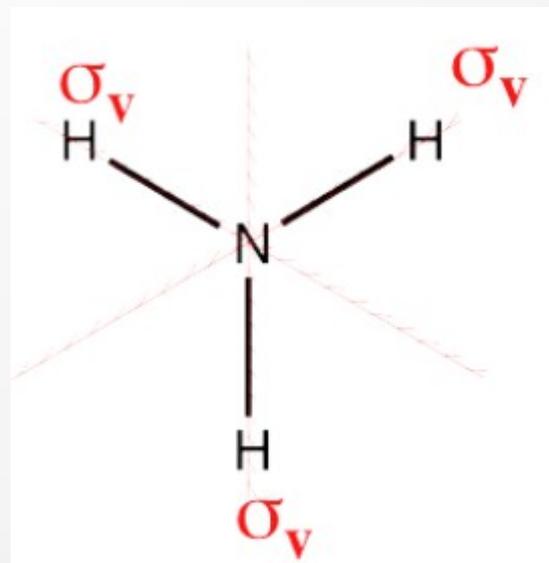
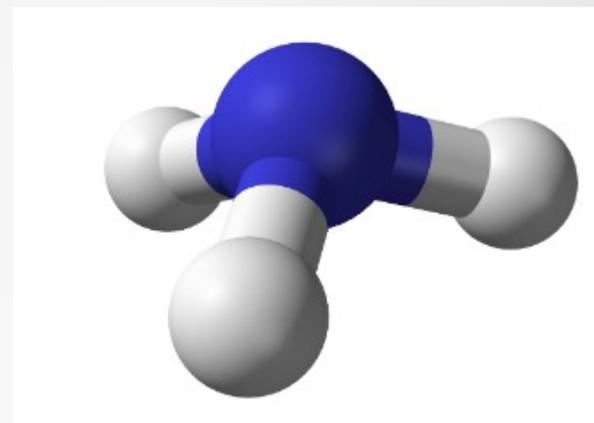
# Simmetria molecolare

- Esempio ammoniaca (identita' E):



Due rotazioni C<sub>3</sub> (angoli di 120°)  
e C<sub>3</sub><sup>2</sup> (angoli di 240°)

E quindi gruppo di simmetria  
C<sub>3v</sub> dato da: E 2C<sub>3</sub> 3σ<sub>v</sub>

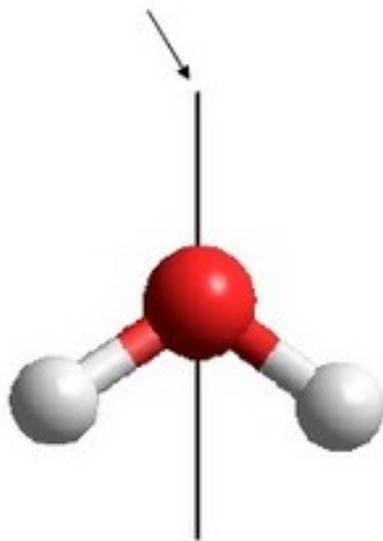


Tre piani di  
Simmetria  
σ<sub>v</sub>

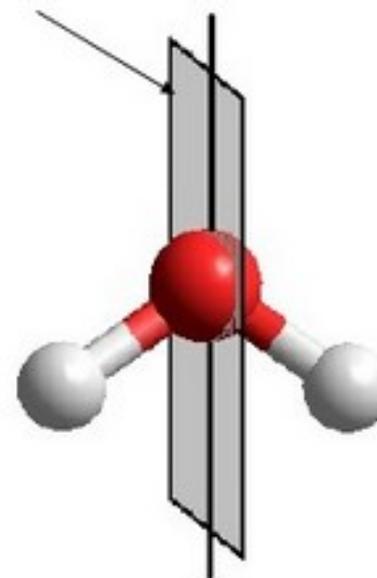
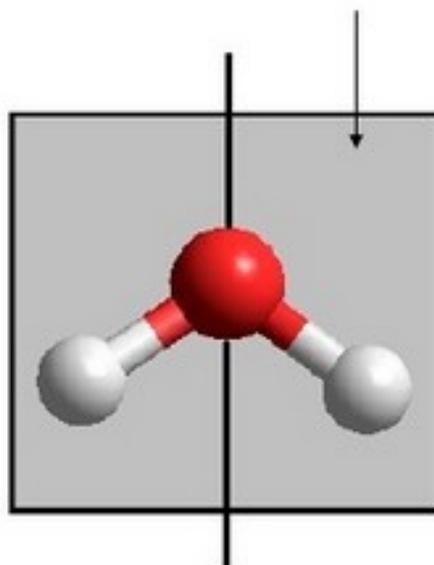
# Simmetria molecolare

- Esempio acqua (identita' E):

only one rotational  
axis =  $C_2$

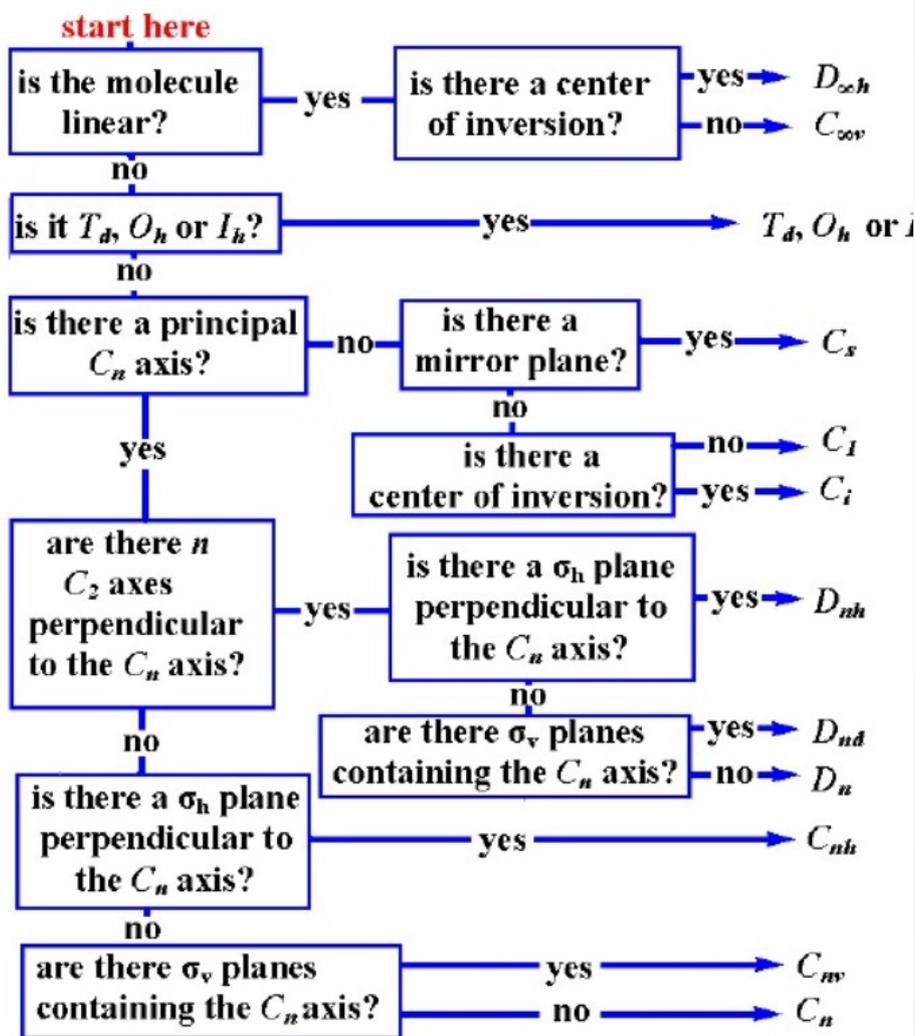


two  $\sigma_v$  but no  $\sigma_h$  mirror planes means  
point group is  $C_{2v}$



The point group of the water molecule is  $C_{2v}$

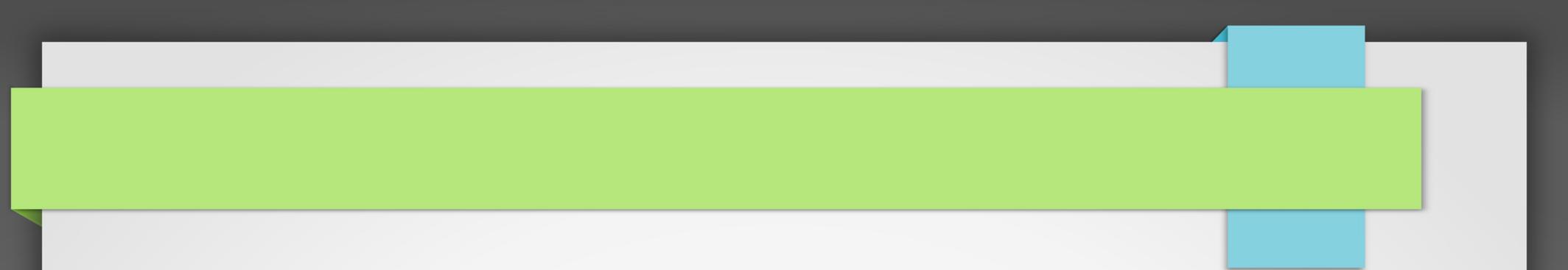
# Simmetria molecolare



Flowchart utilizzabile nella determinazione del gruppo di simmetria delle molecole

Ad esempio imponendo (vincolo) che l'ammoniaca appartenga al Gruppo di simmetria  $C_{3v}$

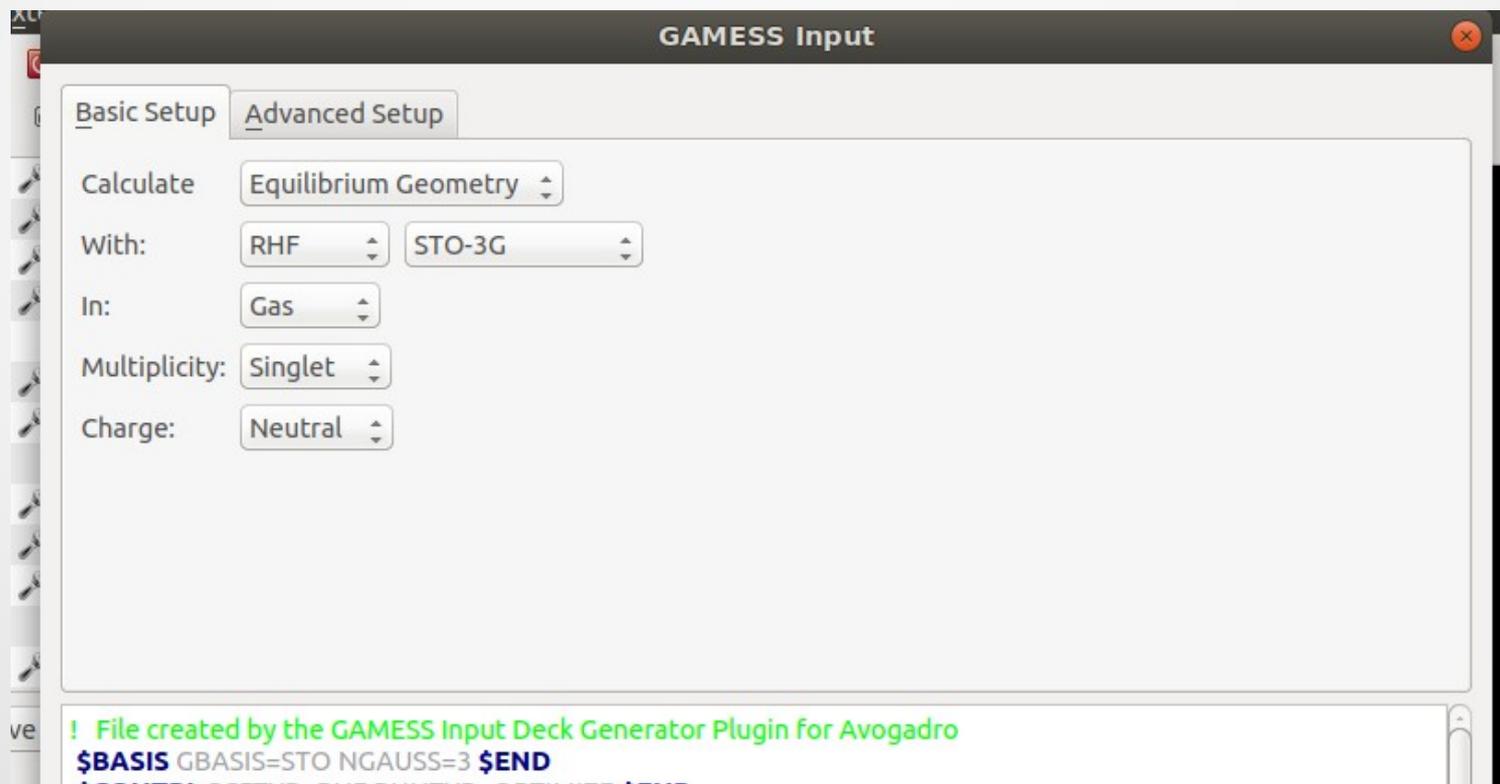
nell'ottimizzazione di geometria si possono usare due gradi di liberta' invece che  $3N - 6 = 6$ . Una sola distanza (N-H) ed un solo angolo (H-N-H)



# ESERCITAZIONE

# Minimizzazione geometria

- Usiamo avogadro e gamess-us per la minimizzazione della geometria usando RHF e due set di base differenti (para.mol2)

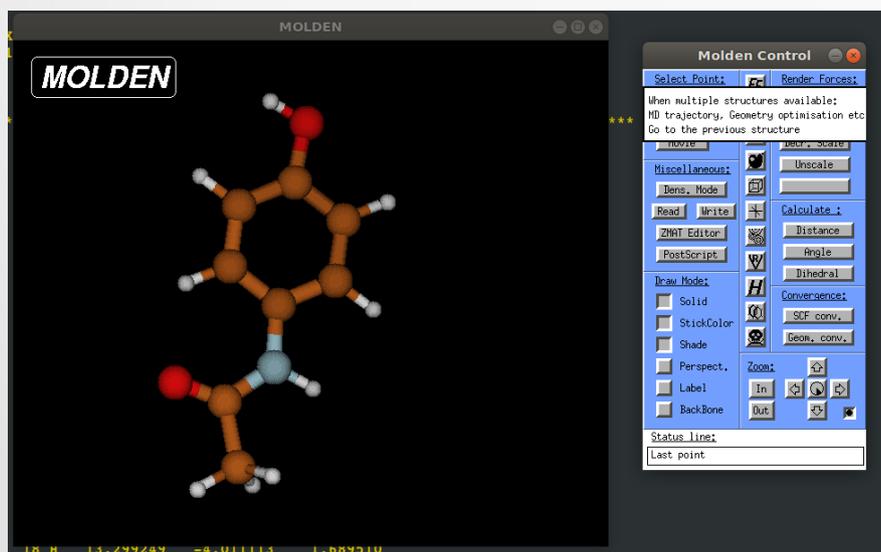


# Minimizzazione geometria

- Facciamo il run e vediamo poi i risultati con molder:

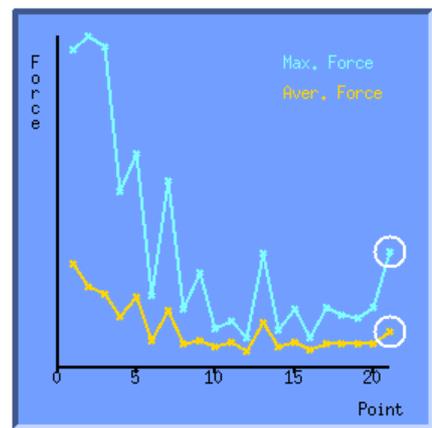
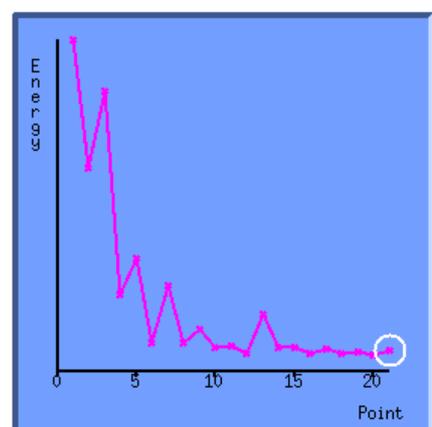
```
[redo@buchner 4]$ time /home/rede/Lezioni/ChimicaComputazionale/gamess/rungms para_opt_631gdp > para_opt_631gdp.out
cp para_opt_631gdp.inp /scr/rede/para_opt_631gdp.F05
unset echo
/extrafs/Lezioni/ChimicaComputazionale/gamess/ddikick.x /extrafs/Lezioni/ChimicaComputazionale/gamess/gamess.00.x pa
l 1 buchner -scr /scr/rede
```

```
-FW-F--F-- 1 Redo Redo 810044 set 3 18:46 para_opt_631gdp.out
[redo@buchner 4]$ time /home/rede/Lezioni/ChimicaComputazionale/gamess/rungms para_opt_sto3g > para_opt_sto3g.out
cp para_opt_sto3g.inp /scr/rede/para_opt_sto3g.F05
unset echo
/extrafs/Lezioni/ChimicaComputazionale/gamess/ddikick.x /extrafs/Lezioni/ChimicaComputazionale/gamess/gamess.00.x par
l 1 buchner -scr /scr/rede
```



`/usr/local/molder/molder file.out`

# Minimizzazione geometria



ote: The following floating-point  
nset echo  
s: No match.  
s: No match.  
s: No match.

redo@banquo:~/Sources/libpka\_builder/libtaut/bin x redo@buchner:~/Lezioni/ChimicaComputazionale/4

**MOLDEN**

The MOLDEN interface displays a ball-and-stick model of a molecule. The structure consists of a central ring of atoms (orange and blue) with several hydrogen atoms (white) and two hydroxyl groups (red and white). The interface includes a 'MOLDEN Control' panel on the right with various options for rendering and calculation.

**Molden Control**

Select Point:  
First  
Next Prev  
Movie

Miscellaneous:  
Dens. Mode  
Read Write  
ZMAT Editor  
PostScript

Draw Mode:  
Solid  
StickColor  
Shade  
Perspect.  
Label  
BackBone

Render Forces:  
Forces  
Incr. Scale  
Decr. Scale  
Unscale

Calculate:  
Distance  
Angle  
Dihedral

Convergence:  
SCF conv.  
Geom. conv.

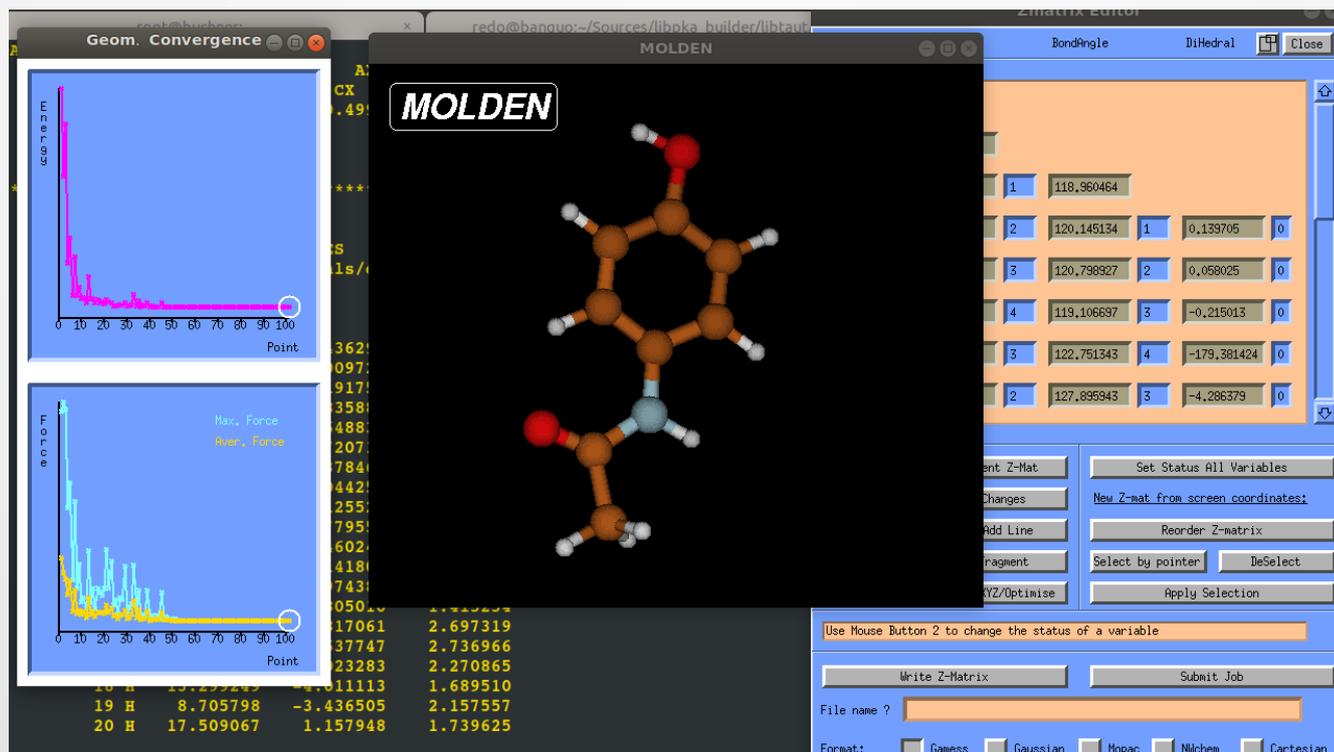
Zoom:  
In  
Out

Status line:  
Last point

# Minimizzazione geometria

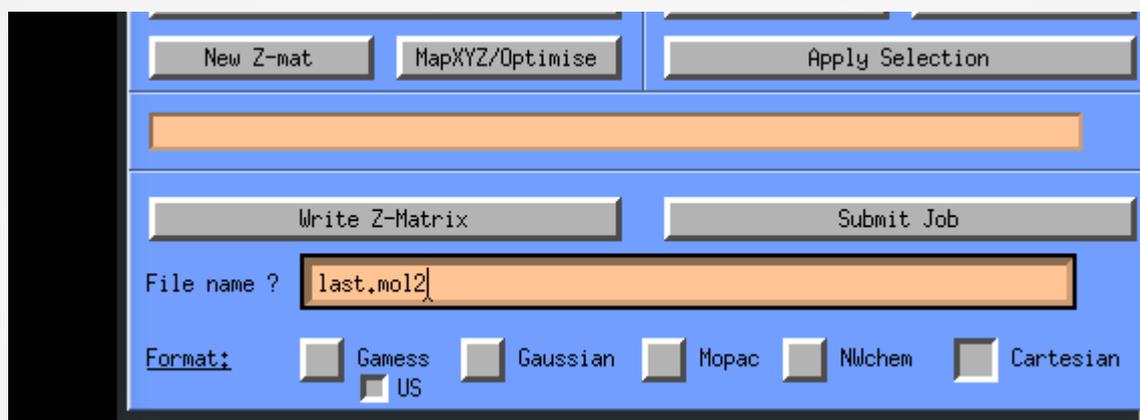
Proviamo ad incrementare il numero di step. Oppure a ripartire dall'ultimo step per andare oltre

```
$ cp para_opt_sto3g.inp para_opt_sto3g_morestep.inp
$ vi para_opt_sto3g_morestep.inp
$ /usr/local/gamess/rungms para_opt_sto3g_morestep > para_opt_sto3g_morestep.out &
```



# Minimizzazione geometria

Dopo aver riaperto l'output con molden ed aver selezionato l'ultimo frame



Ed apriamo punto di partenza e la geometria a convergenza con pymol

```
[redo@buchner 4]$ pymol para.mol2 last.mol2
PyMOL(TM) Molecular Graphics System, Version 1.8.4
Copyright (c) Schrodinger, LLC.
```

# Minimizzazione geometria

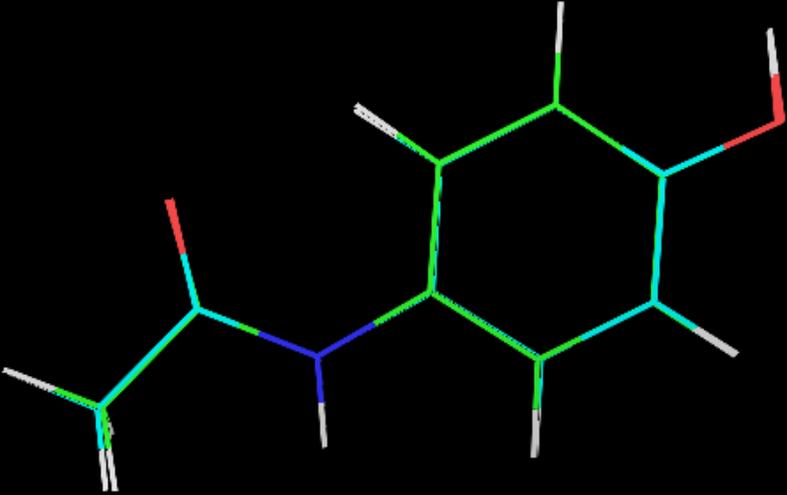
The PyMOL Molecular Graphics System

File Edit Build Movie Display Setting Scene Mouse Wizard Plugin Help

```
GL_RENDERER: Mesa DRI Intel(R) HD Graphics 5500 (Broadwell GT2)
GL_VERSION: 3.0 Mesa 18.0.5
Detected 4 CPU cores. Enabled multithreaded rendering.
CmdLoad: "para.mol2" loaded as "para".
Executive-Warning: name "last" collides with a selection language keyword.
CmdLoad: "last.mol2" loaded as "last".
PyMOL>
```

Reset Zoom Orient Draw Ray  
Unpick Deselect Rock Get View  
< < Stop Play > >| MClear  
Command Builder Volume  
Rebuild Abort

PyMOL Viewer



all	A	S	H	L	C
para ***** 1/1	A	S	H	L	C
last <i>Holden gene</i>	A	S	H	L	C

Mouse Mode 3-Button Viewing  
Buttons L M R Wheel  
& Keys Rota Move MovZ Slab  
Shft +Box -Box Clip MovS  
Ctrl +/- PkAt Pk1 MvSZ  
CtSh Sele Orig Clip MovZ  
SnglClk +/- Cent Menu  
Db1Clk Menu - PkAt  
Selecting Residues  
State 1/ 1

# Minimizzazione geometria

Calcoliamo l'RMSD: Root-mean-square deviation of atomic positions cioè la misura media della distanza fra gli atomi delle due strutture (una misura della differenza fra le due strutture)

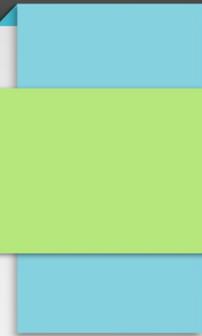
$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2}$$

```
PyMOL>align last , para_
( ( last ) in ( para ) ) ~--
PyMOL>align last , para
Match: read scoring matrix.
Match: assigning 1 x 1 pairwise scores.
MatchAlign: aligning residues (1 vs 1)...
MatchAlign: score 0.000
ExecutiveAlign: 20 atoms aligned.
Executive: RMS = 0.052 (20 to 20 atoms)
```

# Minimizzazione geometria

- Adesso fate la stessa cosa usando pero' questa molecola (minimizzate con due set di base e calcoliamo l'RMDS oltre che riportare i valori di energia e nome chemspider ad esempio)

CN1CCC23C4C1CC5=C2C(=C(C=C5)O)OC3C(C=C4)O



-----