

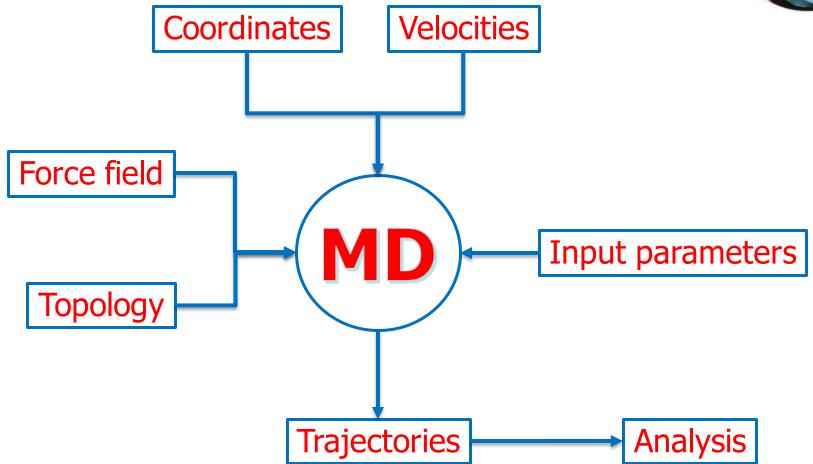
Introduction to Classical Molecular Dynamics

Giovanni Chillemi g.chillemi@cineca.it HPC department, CINECA



MD ingredients











Equation of motion



The equations that describe the temporal evolution of a physical system is called **equation of motion**. There are several different equations of motions, which characterize the motion in different types of physical systems:

- Time-dependent Schrödinger's Equation
 - for quantum-mechanical system
- Newton's Equation
 - for classical-mechanical system
- Langevin's Equation
 - for stochastic system







Newton's Equation of motion



Molecules are quantum-mechanical systems whose motion should be described by Schrödinger's Equation. However, technical difficulties make solving Schrödinger's Equation for large systems impractical.

Therefore the motion of a molecule is usually approximated by the laws of <u>classical mechanics</u> and by Newton's equation of Motion. In its most simplistic form Newton's second law of motion states:

$$\mathbf{f}_{i} = \mathbf{m}_{i} \cdot \mathbf{a}_{i}$$

where m_i is the mass of particle i, a_i is its acceleration. The force f_i is given as the derivative of the potential energy function V:

$$\mathbf{f}_{i} = -\frac{\partial V}{\partial \mathbf{r}_{i}}$$



where r_i is the position of particle i



Potential energy function



$$V(\mathbf{r}_{1}, \mathbf{r}_{2}, ..., \mathbf{r}_{n}) = \sum_{bond} \frac{1}{2} k_{b_{n}} (b_{n} - b_{0_{n}})^{2} + \sum_{angle} \frac{1}{2} k_{\theta_{n}} (\theta_{n} - \theta_{0_{n}})^{2} + \sum_{improper} \frac{1}{2} k_{\xi_{n}} (\xi_{n} - \xi_{0_{n}})^{2} + \sum_{dihedral} k_{\phi_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\xi_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{improper} k_{\theta_{n}} [1 + \cos(m_{n} \phi_{n} - \delta_{n})] + \sum_{im$$

$$+\sum_{\substack{improper\\dih odnol}} \frac{1}{2} k_{\xi_n} \left(\xi_n - \xi_{0n}\right)^2 + \sum_{\substack{dihedral\\dihedral}} k_{\phi_n} \left[1 + \cos\left(m_n \phi_n - \delta_n\right)\right] +$$

$$+ \sum_{\substack{nonbonded \\ pairs (ij)}} \left(\left(\frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}} \right) + \frac{1}{4\pi\varepsilon_{0}} \frac{q_{i}q_{j}}{\varepsilon_{r}r_{ij}} \right)$$



bonded interactions

non bonded interactions









Model	Degree of freedom		Example of predicted properties	
	Considered	Removed		
Quantum mechanic	Nucleus, electrons	Nucleons	Chemistry reaction	
Polarizable atoms	Atoms, dipoles	Electrons	Binding of charged substrates	
Non polarizable atoms	Solute atoms, solvent atoms	Dipoles	Conformational transitions Hydration	
Implicit solvent	Solute atoms	Solvent atoms	Folding topology of macromolecules	

Classical Molecular Dynamics







A Brief History

	E
-	

Year	System
1964	Liquid Argon (Rahman Phys Rev)
1974	Water (Rahman J. Chem Phys.)
1977	Small protein in vacuo (Mc Cammon Karplus Nature)
1988	First Protein in explicit water (Levitt PNAS)
From 1995	Protein-DNA Complexes – Membrane Proteins- Complex Systems







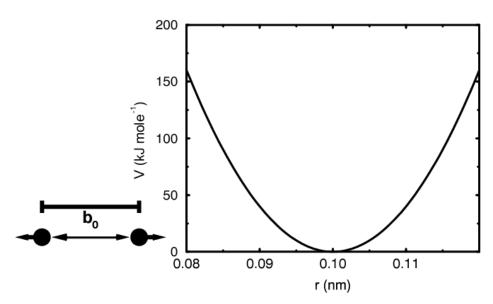
Bond Stretching Energy



$$= \sum_{bond} \frac{1}{2} k_{b_n} (b_n - b_{0n})^2 + \dots$$

k_b is the spring constant of the bond b₀ is the bond length at equilibrium

Unique k_b and b₀ assigned for each bond pair, i.e. C-C, O-H



Principle of bond stretching (left), and the bond stretching potential (right).







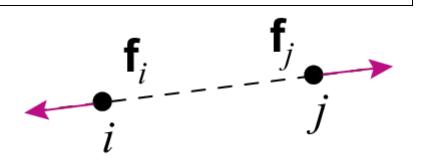
Bond Stretching Force



$$\mathbf{f}_{i} = -\frac{\partial V^{bond}}{\partial \mathbf{r}_{i}} = -\frac{\partial V^{bond}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial \mathbf{r}_{i}} = k_{b} \left(r_{ij} - b_{0} \right) \frac{\mathbf{r}_{ij}}{r_{ij}}$$

$$\mathbf{f}_{j} = -\mathbf{f}_{i}$$

If atom i and j are closer than b_0 , the bond force separates them





If atom i and j are farther than b_0 , the bond force draws them nearer



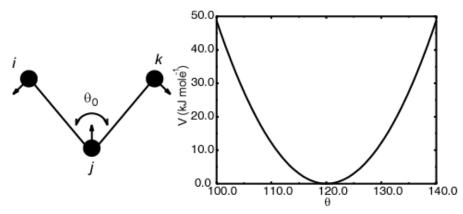


Bending Energy



$$+\sum_{angle}\frac{1}{2}k_{\theta_n}\left(\theta_n-\theta_{0n}\right)^2+$$

 k_{θ} is the spring constant of the bending. θ_0 is the angle bending at equilibrium.



Principle of angle vibration (left) and the bond angle potential (right).

Unique parameters for angle bending are assigned to each bonded triplet of atoms based on their types (e.g. C-C-C, C-O-C, C-C-H, etc.)

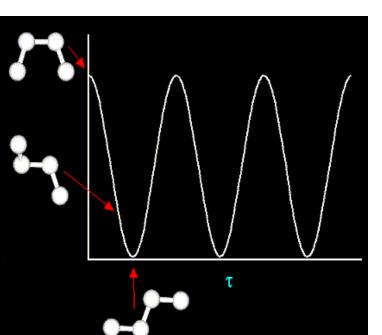


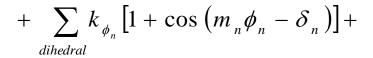




Torsional o Dihedral Energy



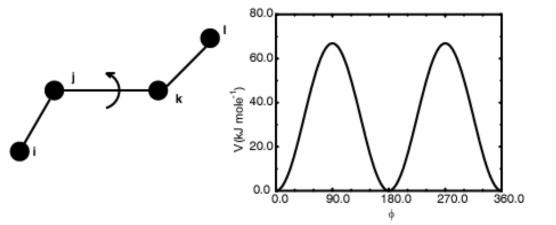




$$\phi$$
 = angle

$$\delta$$
 = phase

m = number of peaks in a full rotation







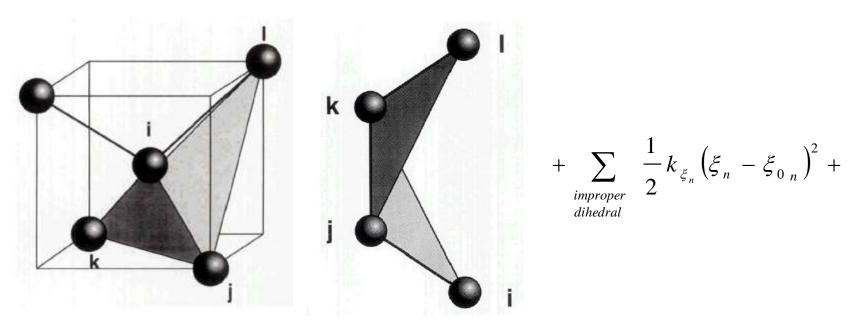
Principle of proper dihedral angle (left, in trans form) and the dihedral angle potential (right).



Improper Dihedral Energy



The energy required to deform a group of atoms from its equilibrium angle, x_0 . Used for tetrahedral or planar groups



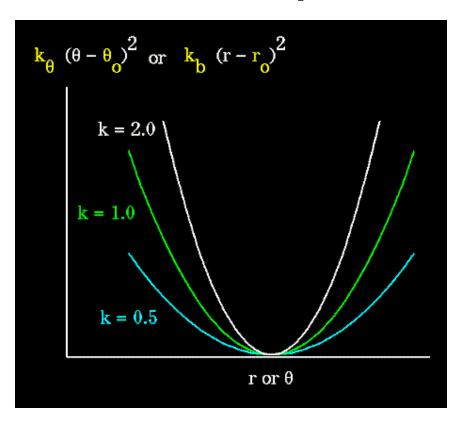
Again this system can be modeled by a spring, and the energy is given by the Hookean potential with respect to the planar angle





The "Hookean" potential





 k_b and k_θ broaden or steepen the slope of the parabola The larger the value of k, the more energy is required to deform an angle (or bond) from its equilibrium value







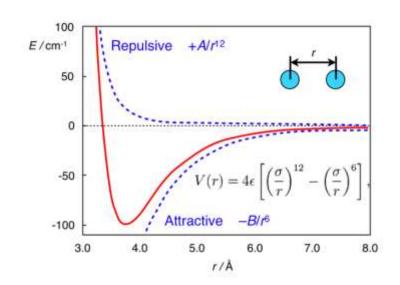
Lennard Jones (Van der Waals) interactions

Computing Applications and Innovation



Sir John Lennard Jones







Johannes Diderik Van der Waals

LJ interactions

$$+ \sum_{\substack{nonbonded \\ pairs \ (ij)}} \left(\left[\frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}} \right] + \frac{1}{4\pi\varepsilon_{0}} \frac{q_{i}q_{j}}{\varepsilon_{r}r_{ij}} \right]$$

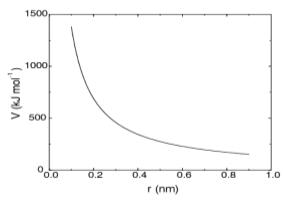






Electrostatic interactions

The q_i and q_j are the partial atomic charges for atoms i and j, separated by a distance r_{ij} . ϵ_r is the relative dielectric constant. For gas phase calculations ϵ_r is normally set to 1. Larger values of ϵ_r are used to approximate the dielectric effect of intervening solute (ϵ ~60-80) or solvent atoms in solution



The Coulomb interaction (for particles with equal signed charge)

Electrostatic

$$+ \sum_{\substack{nonbonded \\ pairs \ (ij)}} \left(\left(\frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}} \right) + \frac{1}{4\pi\varepsilon_{0}} \frac{q_{i}q_{j}}{\varepsilon_{r}r_{ij}} \right)$$



Charles Augustin de Coulomb

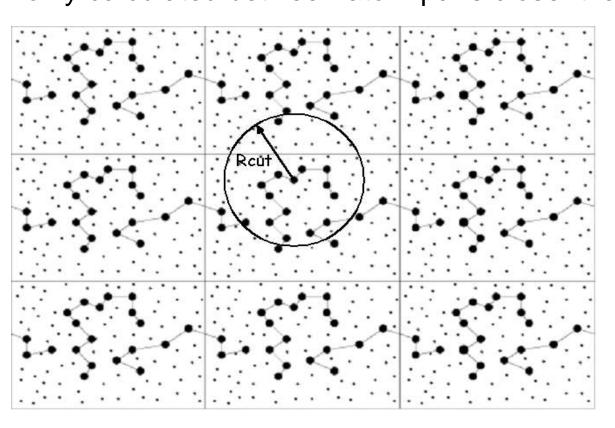






Periodic boundary conditions and cut-off radius

To simulate our finite system in liquid conditions, we apply the pbc: i.e. the system box is virtually surrounded in all directions by copy of itself. An atom close to a box border interacts with the atoms in another pbc image. The non-bonded interactions are only calculated between atom pairs closer than a spherical cut-off



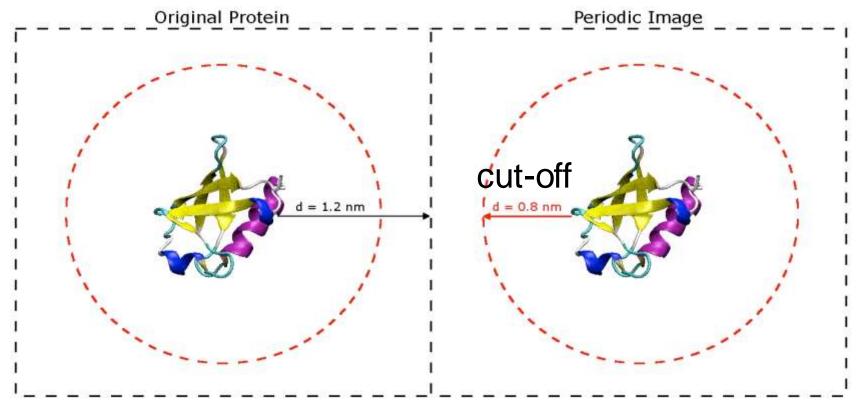


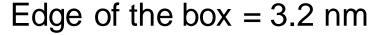


BOX dimension

The edge of cubic box must be large enough to avoid interactions of the solute with itself.

Its minimal dimension therefore depends by the chosen cut-off for the non bonded interactions











Electrostatic interactions: Particle Mesh Ewald (PME)

The cut-off radius method for electrostatic interactions is particularly inaccurate for charged molecules such as DNA of for dipolar groups such as alpha helices

PME corrects these errors and it helps maintaining short the cut-off in the real space: i.e. the number of atom pairs is reduced

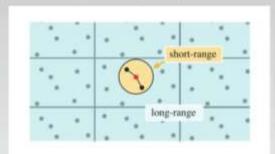
- -Short range in the real space
- -Long range in the Fourier space

Coulomb forces on N particles, charges q_i , positions \mathbf{r}_i , box length L, periodic b.c.

electrostatic potential

$$V = \frac{1}{2} \sum_{i,j=1}^{N} \sum_{\mathbf{n} \in \mathbb{Z}^3}^{\prime} \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|}$$

straightforward summation impracticable

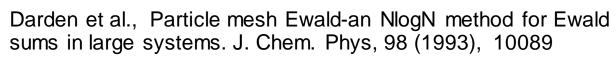


Trick 1: Split problem into 2 parts with help of:

$$\frac{1}{r} = \underbrace{\frac{f(r)}{r}}_{\text{short range}} + \underbrace{\frac{1 - f(r)}{r}}_{\text{long range}}$$









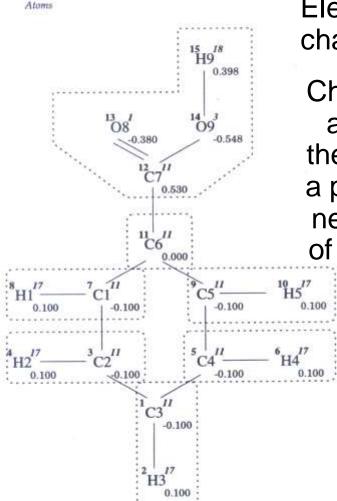




Charge groups and atom types



Name:



A charge group is a neutral charge group composed by several partially charged atoms of a chemical group. Electrostatics can be calculated between charge groups instead that atom pairs

Charge groups were first introduced to reduce artifacts in the electrostatics calculation but they can also speed up the calculations; given a pair of water molecules for instance, we only need to determine one atom distance instead of nine (or sixteen for a four-site water model)

> Note that an atom type is not a physical feature. O8 is defined with a different atom type than O9. In fact, their bond constants with C7 and atomic charges are different



Force field

$$V(\mathbf{r}_{1}, \mathbf{r}_{2}, ..., \mathbf{r}_{n}) = \sum_{bond} \frac{1}{2} k_{b_{n}} (b_{n} - b_{0n})^{2} + \sum_{angle} \frac{1}{2} k_{\theta_{n}} (\theta_{n} - \theta_{0n})^{2} +$$



$$+\sum_{\substack{improper\\dihedral}} \frac{1}{2} k_{\xi_n} \left(\xi_n - \xi_{0_n} \right)^2 + \sum_{\substack{dihedral\\dihedral}} k_{\phi_n} \left[1 + \cos \left(m_n \phi_n - \delta_n \right) \right] +$$

$$+ \sum_{\substack{nonbonded \\ pairs \ (ij)}} \left(\left(\frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}} \right) + \frac{1}{4\pi\varepsilon_{0}} \frac{q_{i}q_{j}}{\varepsilon_{r}r_{ij}} \right)$$

The potential energy function, together with the parameters required to describe the behavior of different kinds of atoms and bonds (k_b , k_{θ} , k_{ξ} , C_{ii} , ...), is called a **force field**.

Several force fields are currently used and the choice depends from the studied system. Some force field are better suited for nucleic acids, for example, while others for membrane proteins







Available forcefield in Gromacs (4.6.5)



- 1. AMBER03 protein, nucleic AMBER94 (Duan et al., J. Comp. Chem. 24, 1999-2012, 2003)
- 2. AMBER94 force field (Cornell et al., JACS 117, 5179-5197, 1995)
- 3. AMBER96 protein, nucleic AMBER94 (Kollman et al., Acc. Chem. Res. 29, 461-469, 1996)
- 4. AMBER99 protein, nucleic AMBER94 (Wang et al., J. Comp. Chem. 21, 1049-1074, 2000)
- 5. AMBER99SB protein, nucleic AMBER94 (Hornak et al., Proteins 65, 712-725, 2006)
- 6. AMBER99SB-ILDN protein, nucleic AMBER94 (Lindorff-Larsen et al., Proteins 78, 1950-58, 2010)
- 7. AMBERGS force field (Garcia & Sanbonmatsu, PNAS 99, 2782-2787, 2002)
- 8. CHARMM27 all-atom force field (with CMAP) version 2.0
- 9. GROMOS96 43a1 force field
- 10. GROMOS96 43a2 force field (improved alkane dihedrals)
- 11. GROMOS96 45a3 force field (Schuler JCC 2001 22 1205)
- 12. GROMOS96 53a5 force field (JCC 2004 vol 25 pag 1656)
- 13. GROMOS96 53a6 force field (JCC 2004 vol 25 pag 1656)
- 14. GROMOS96 54a7 force field (Eur. Biophys. J. (2011), 40,, 843-856)
- 15. OPLS-AA/L all-atom force field (2001 aminoacid dihedrals)
- 16. [DEPRECATED] Encad all-atom force field, using full solvent charges
- 17. [DEPRECATED] Encad all-atom force field, using scaled-down vacuum charges
- 18. [DEPRECATED] Gromacs force field (see manual)
- 19. [DEPRECATED] Gromacs force field with hydrogens for NMR







CA Integration of the equation of motion



Numeric integration of Newton's equation of motion is typically done step by step using methods that are called **Finite Difference** methods.

These methods use the information available at time t to predict the system's coordinates and velocities at a time t + δt , where δt is a short time interval and are based on a Taylor expansion of the position at time t + δt

$$r(t+\delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^{2} + \dots$$



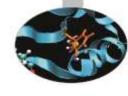


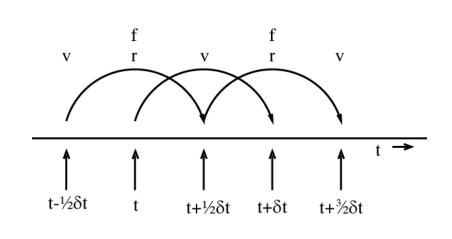


Integration of the equation of motion

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + a(t)\delta t^{2}$$







$$r(t + \delta t) = r(t) - v(t + \frac{1}{2}\delta t)\delta t$$

$$v(t + \frac{1}{2}\delta t) = v(t - \frac{1}{2}\delta t) + a(t)\delta t$$

Leap-frog integrator

$$r(t + \delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^{2}$$
$$v(t + \delta t) = v(t) + \left[a(t) + a(t + \delta t)\right]\frac{\delta t}{2}$$

Velocity Verlet







Choice of the timestep



The length of the timestep must be small compared to the period of the highest frequency motions being simulated

Force characteristics	Relaxation time (fs)	Time step (fs)
High frequency motion: bond	10	0.5
stretching vibrations		
Medium frequency motion: angle		
bending, proper and improper dihedral	40	2
angle deformation, LJ and short range		
Coulombian interactions		
Low frequency motion: long range	1000	20
coulombian interactions		

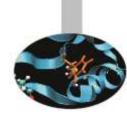
The bond stretching vibrations are generally of minimal interest in the study of biomolecular structure and function. Therefore this degree of freedom is usually kept frozen with constraint algorithms such as Shake, Settle, Lincs and a typical timestep is 2 fs (2 x 10⁻¹⁵ s)

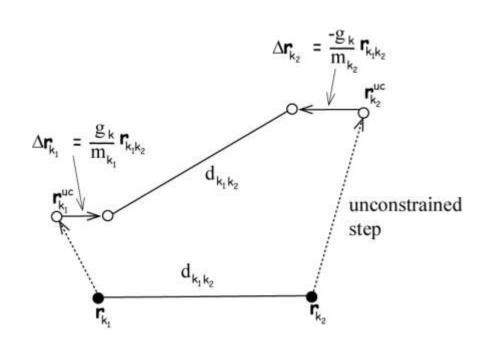






Constraints





L'applicazione di vincoli geometrici per mantenere la lunghezza dei legami covalenti durante la simulazione permette di usare un time step fino a 2 fs

In SHAKE (storicamente il primo metodo per l'applicazione dei vincoli in un codice di dinamica molecolare) l'equazione del moto è integrata soddisfacendo contemporaneamente i vincoli sulla distanza degli atomi legati, usando il metodo dei **moltiplicatori di Lagrange**Metodo iterativo





Constraints



L'algoritmo LINCS (proposto nel 1997, venti anni dopo SHAKE) risolve in maniera non iterativa vincoli tra atomi legati ed angoli isolati

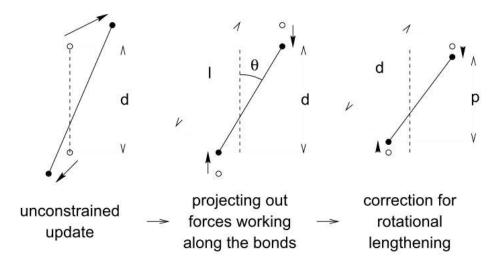


Figure 3.9: The three position updates needed for one time step. The dashed line is the old bond of length d, the solid lines are the new bonds. $l = d\cos\theta$ and $p = (2d^2 - l^2)^{\frac{1}{2}}$.

SETTLE è un algoritmo non iterativo per applicare i vincoli ad una molecola d'acqua







Constraints



constraints:

none

No constraints except for those defined explicitly in the topology, *i.e.* bonds are represented by a harmonic (or other) potential or a Morse potential (depending on the setting of morse) and angles by a harmonic (or other) potential.

h-bonds

Convert the bonds with H-atoms to constraints.

all-bonds

Convert all bonds to constraints.

h-angles

Convert all bonds and additionally the angles that involve H-atoms to bond-constraints.

all-angles

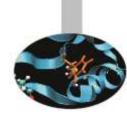
Convert all bonds and angles to bond-constraints.

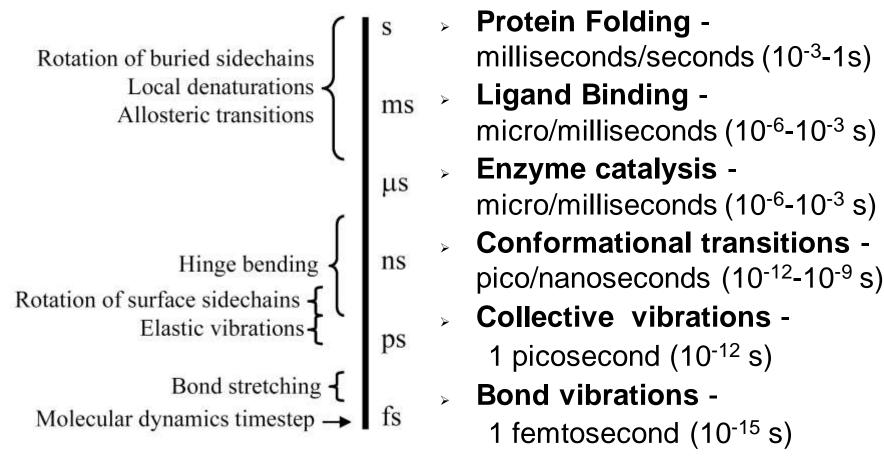






Timescale









Topology



The topology file describes the atoms composing a molecule and their bond connections

Es: flexspc.itp in gromacs

```
[ moleculetype ]
; molname
             nrexcl
SOL
[ atoms ]
; id at type
             res nr res name at name cg nr charge
                                                   mass
   OW_spc
                                        -0.82
                                                15.99940
                    SOL
                            OW
 2 HW_spc
                            HW1
                                                1.00800
              1 SOL
                                         0.41
                            HW2
   HW_spc
                    SOL
                                         0.41
                                                1.00800
[bonds]
         funct length force.c.
                0.1
                     345000 0.1
                                   345000
                     345000 0.1
                0.1
                                   345000
[ angles ]
              funct angle force.c.
                     109.47 383
                                   109.47 383
```







Constraints in Topology



Only in case of water, the constraint algorithm can be selected in the topology file

```
[settles]; OW funct doh dhh SPC.itp
1 0.1 0.16333
```







Topology

Fig. 4.2.70 Molecular topology building block definition

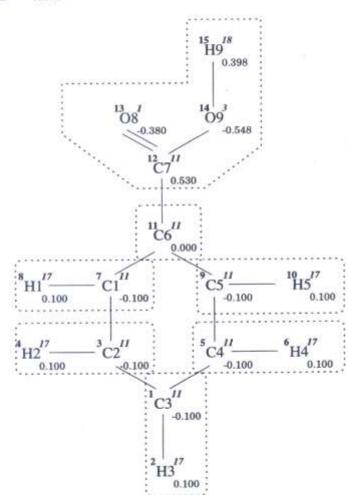
Solute building block:

Benzoic acid (neutral)

Name:

BA

. Atoms





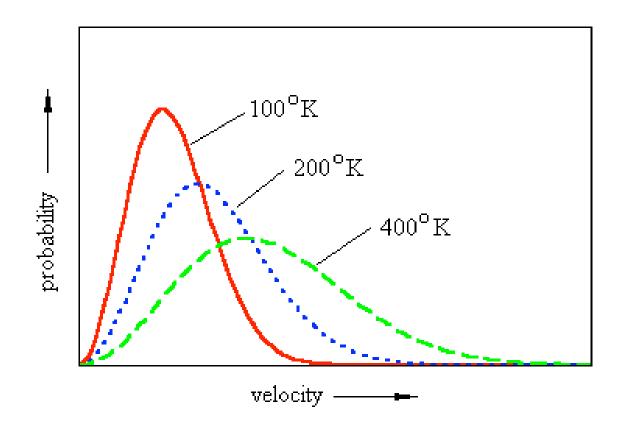






Initial velocities

The initial velocity of each atom is random assigned through a **Maxwell-Boltzmann** distribution that is function of the temperature



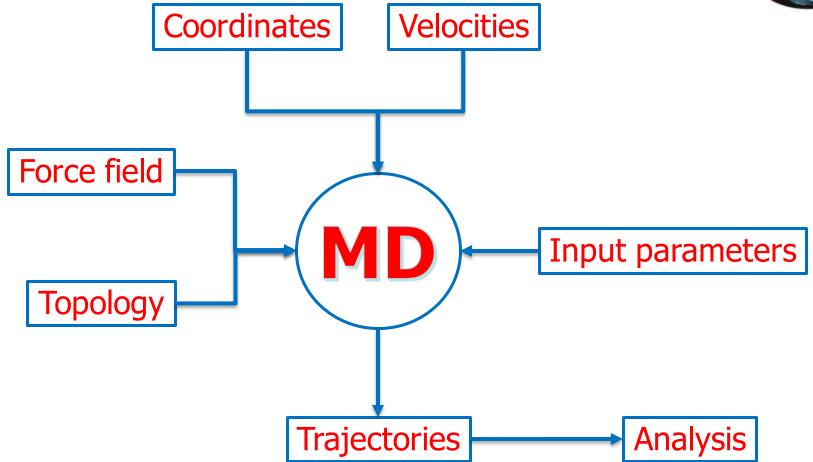






To recapitulate...











THE GLOBAL MD ALGORITHM

1. Input initial conditions

Potential interaction V as a function of atom positions Positions r of all atoms in the system Velocities v of all atoms in the system



repeat 2,3,4 for the required number of steps:

2. Compute forces

The force on any atom

$$F_i = -\frac{\partial V}{\partial r_i}$$

is computed by calculating the force between non-bonded atom pairs:

$$F_i = \sum_j F_{ij}$$

plus the forces due to bonded interactions (which may depend on 1, 2, 3, or 4 atoms), plus restraining and/or external forces.

The potential and kinetic energies and the pressure tensor are computed.



3. Update configuration

The movement of the atoms is simulated by numerically solving Newton's equations of motion

$$\frac{\mathrm{d}^2 r_i}{\mathrm{d}t^2} = \frac{F_i}{m_i}$$
or
$$\frac{\mathrm{d}r_i}{\mathrm{d}t} = v_i; \quad \frac{\mathrm{d}v_i}{\mathrm{d}t} = \frac{F}{m_i}$$

4. if required: Output step

write positions, velocities, energies, temperature, pressure, etc.









Molecular Dynamics ensembles



- The method discussed above is appropriate for the micro-canonical ensemble: constant N (number of particles), V (volume) and E_T (total energy = E + E_{kin})
- Note that if time step is short enough, the system loses/gains no net energy (potential + kinetic) when running MD in the NVE ensemble
- When simulating biological macromolecules, it might be more appropriate to simulate under constant Temperature (T) or constant Pressure (P):
 - Canonical ensemble: NVT
 - Isothermal-isobaric: NPT

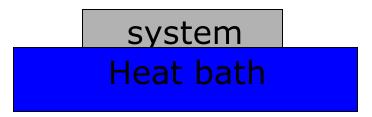






Simulating at constant T: the Berendsen scheme





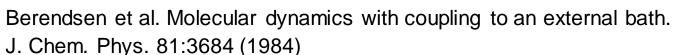
Bath supplies or removes heat from the system as appropriate

$$\frac{dT\left(t\right)}{dt} = \frac{T_0 - T\left(t\right)}{\tau_T}$$

where τ determines how strong the bath influences the system

Exponentially scale the velocities at each time step by the factor λ :

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T(t)} - 1\right)\right]^{\frac{1}{2}}$$
T: "kinetic" temperature









Simulating at constant T: the Berendsen scheme



system

Heat bath

A small τ , close to the timestep (strong thermostat), is useful in the equilibration phase, when the quick decreasing of the potential energy could increase too much the kinetic energy of the protein

A bigger τ , e.g. equal to ten times the timestep (weak thermostat), is useful in the production phase, when we want to keep at minimum the perturbation to the conformational sampling







Simulating at constant P: the Berendsen scheme





Pressure bath

Couple the system to a pressure bath

$$\frac{dP(t)}{dt} = \frac{P_0 - P(t)}{\tau_P}$$

A change in pressure *P* is related to a change in volume *V*

To regulate pressure: exponentially scale the volume of the simulation box at each time step by a factor μ

$$\mu(t) = \left[1 - k_T \frac{\Delta t}{\tau_P} (P_o - P(t))\right]^{\frac{1}{3}}$$

where k_T : isothermal compressibility

 τ_P : coupling constant

Berendsen et al. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81:3684 (1984)







Sample input file of gromacs

http://manual.gromacs.org/current/online/mdp.html



cpp = /lib/cpp

include = -I../top

define =

integrator = md dt = 0.002

nsteps = 500000

nstxout = 5000

nstvout = 5000

nstlog = 5000

nstenergy = 250

nstxout-compressed = 250

compressed-x-grps = Protein

energygrps = Protein SOL

nstlist = 10

ns-type = grid

rlist = 0.8

coulombtype = cut-off

rcoulomb = 1.4

rvdw = 0.8

tcoupl = Berendsen

tc-grps = Protein SOL

tau-t $= 0.1 \ 0.1$

ref-t $= 300 \ 300$

Pcoupl = Berendsen

tau-p = 1.0

compressibility = 4.5e-5

ref-p = 1.0

gen-vel = yes

gen-temp = 300

gen-seed = 173529

constraints = all-bonds







1-4 interactions



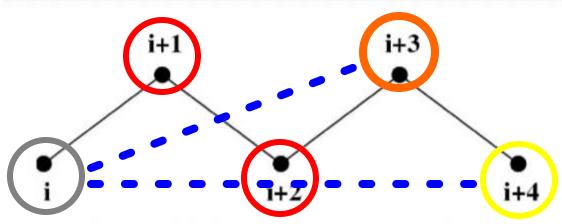


Figure 4.15: Atoms along an alkane chain.

Atoms covalently bound are defined as first neighbours second neighbours and so on....

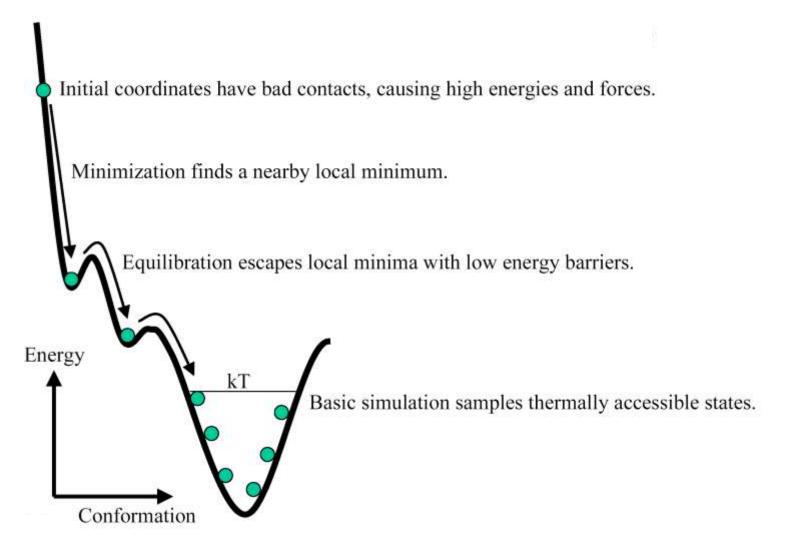
LJ and electrostatic interactions are not calculated among first and second neighbours since they are considered in the stretching (first) or in the bending potential (second)

The standard non-bonding interactions are too strong for the third neighbours and are reduced (interactions 1-4; list 1-4)



Conformational sampling













Energy minimization

La superficie potenziale di una molecola è definita da un gran numero di minimi locali (configurazioni stabili dove tutte le derivate prime della funzione energia potenziale rispetto le coordinate sono nulle e tutte le derivate seconde sono non negative)

o punti di sella (stati di transizione)

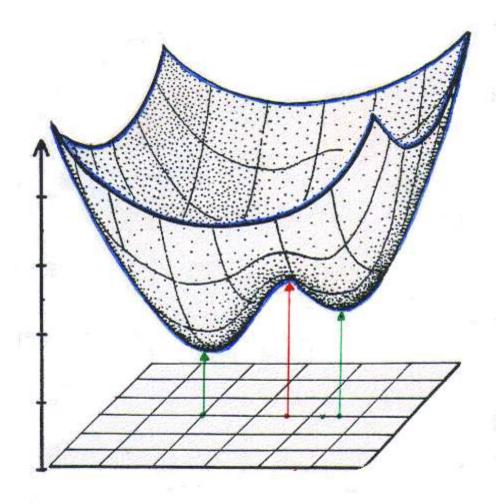






SuperComputing Applications and Innovation











Gli algoritmi di minimizzazione trovano il minimo *locale*della funzione di energia potenziale Cioè quello che viene raggiunto procedendo lungo il gradiente negativo della superficie di energia potenziale

In generale NON trovano il minimo globale

Due algoritmi molto diffusi sono

- steepest descent
- conjugate gradient

entrambi del *primo ordine* utilizzano cioè la derivata prima della funzione potenziale rispetto alle coordinate





Il metodo *steepest descent* utilizza solo il gradiente del potenziale nel punto per calcolare lo spostamento

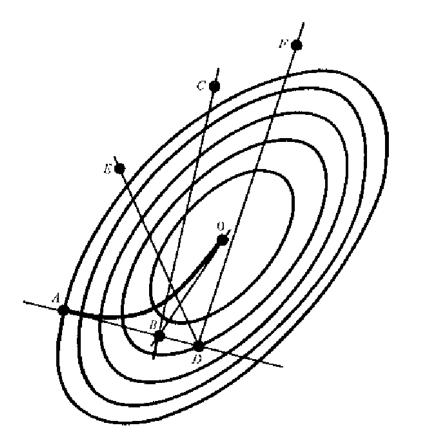


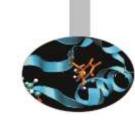
È più rapido nella singola iterazione ma meno preciso nel raggiungere il minimo locale quindi è più adatto ad una prima minimizzazione poco accurata

Il metodo *conjugate gradient* utilizza, oltre al gradiente istantaneo del potenziale, anche quello dello step precedente

È più accurato nel trovare il minimo locale ma è più lento nei primi passi di minimizzazione quindi è spesso usato dopo un ciclo di steepest descent







Partendo da A, lo *steepest descent*s percorre A-B-C (o A-D-F) mentre il *conjugate gradient*, pesando il gradiente A-B con quello B-C riesce a percorrere A-B-O







GROMACS USER MANUAL



steep

A steepest descent algorithm for energy minimization. The maximum step size is **emstep** [nm], the tolerance is **emtol** [kJ mol⁻¹ nm⁻¹].

cg

A conjugate gradient algorithm for energy minimization, the tolerance is **emtol** [kJ $\text{mol}^{-1} \text{ nm}^{-1}$]. CG is more efficient when a steepest descent step is done every once in a while, this is determined by **nstcgsteep**.

emtol: (100.0) [kJ mol⁻¹ nm⁻¹]

the minimization is converged when the maximum force is smaller than this value

emstep: (0.01) [nm]

initial step-size

nstcgsteep: (1000) [steps]

frequency of performing 1 steepest descent step while doing conjugate gradient energy minimization.



