Molecular Dynamics Simulations

Dr. Noelia Faginas Lago
Dipartimento di Chimica,Biologia e Biotecnologie
Università di Perugia

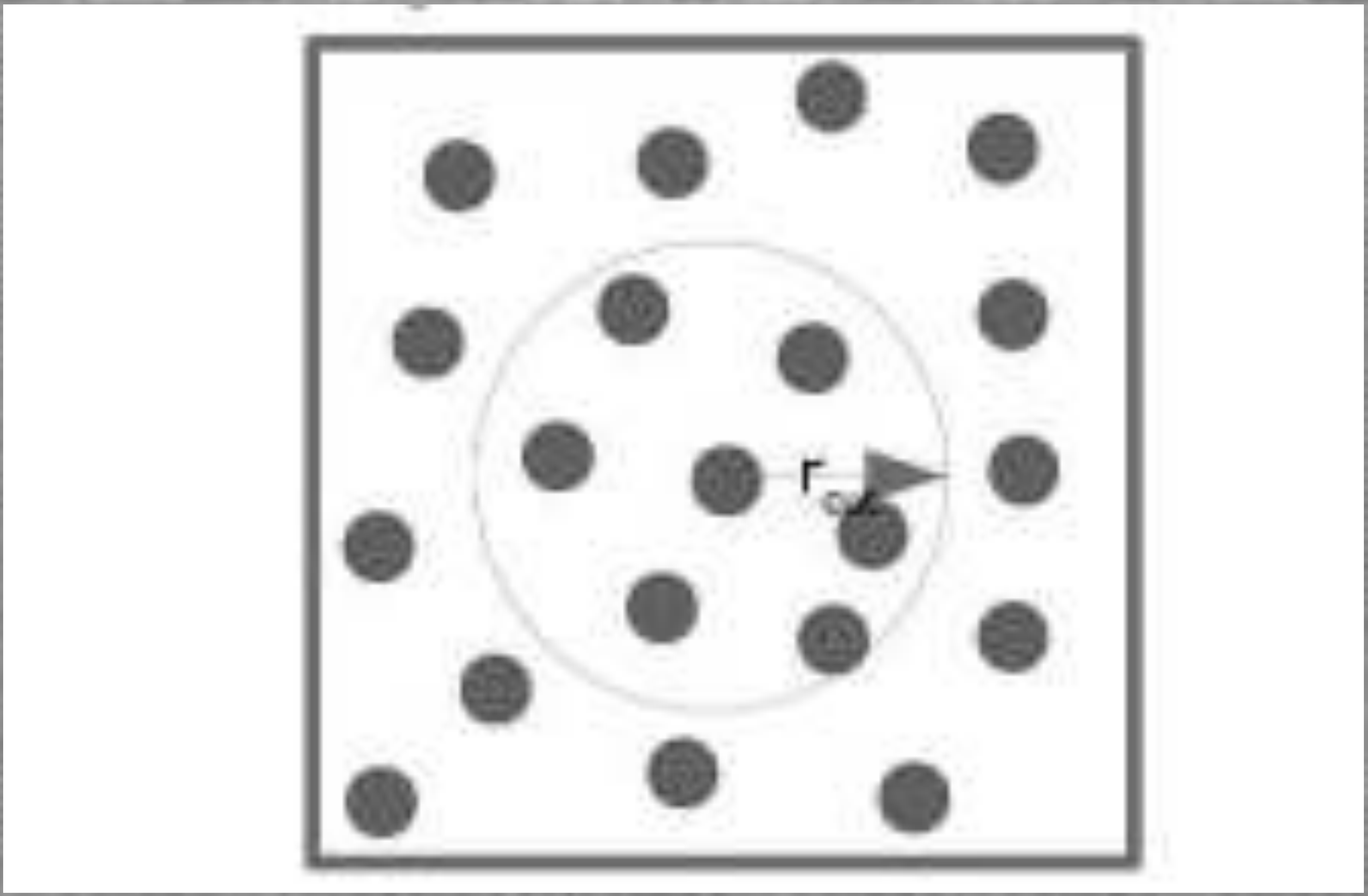
An Introduction to Molecular Dynamics Simulations

Macroscopic properties are often determined by molecule-level behavior.

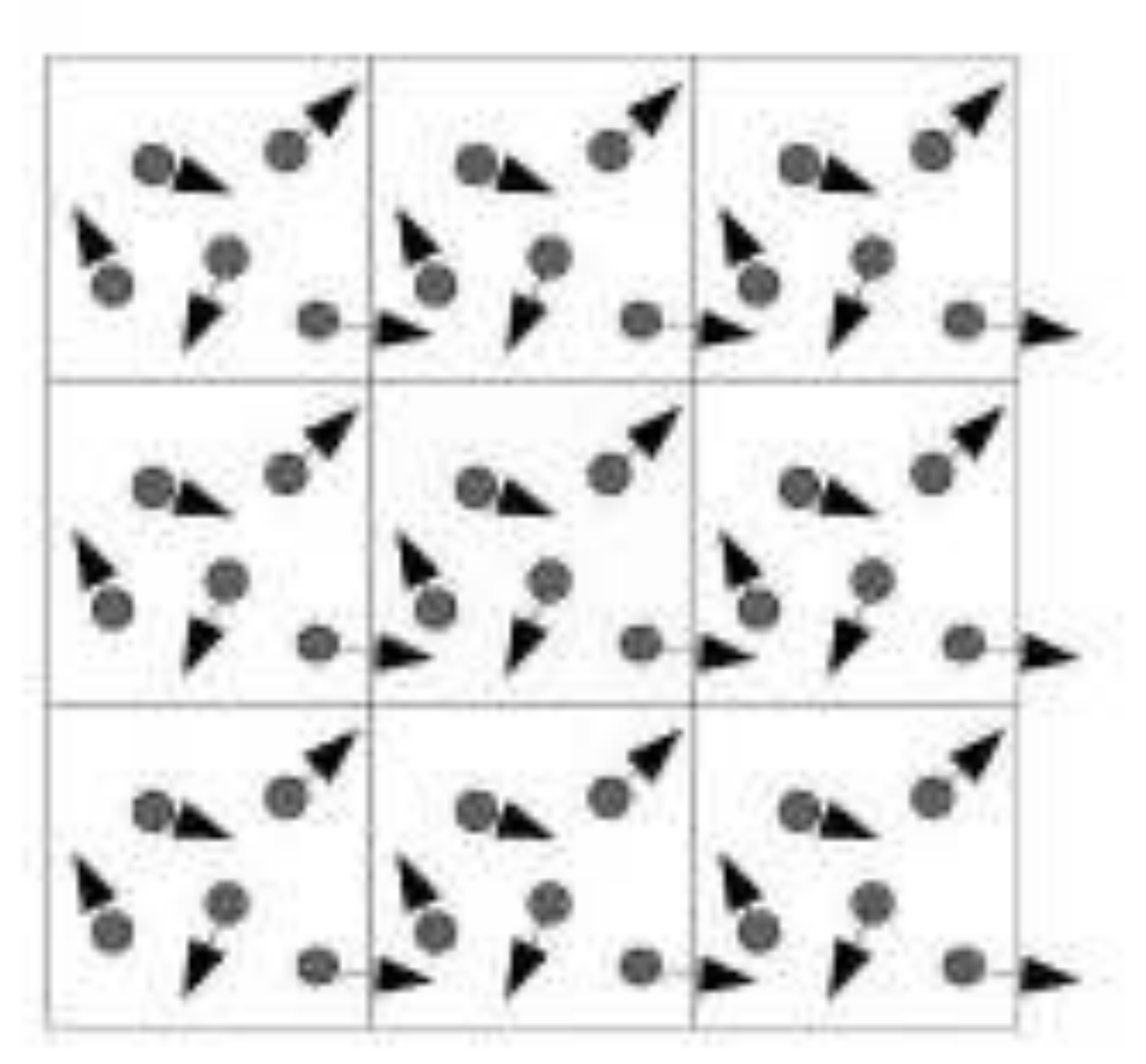
Quantitative and/or qualitative information about macroscopic behavior of macromolecules can be obtained from simulation of a system at atomistic level.

Molecular dynamics simulations calculate the motion of the atoms in a molecular assembly using Newtonian dynamics to determine the net force and acceleration experienced by each atom. Each atom $\boldsymbol{i}$ at position $\boldsymbol{r_i}$, is treated as a point with a mass $\boldsymbol{m_i}$ and a fixed charge $\boldsymbol{q_i}$.

The model system

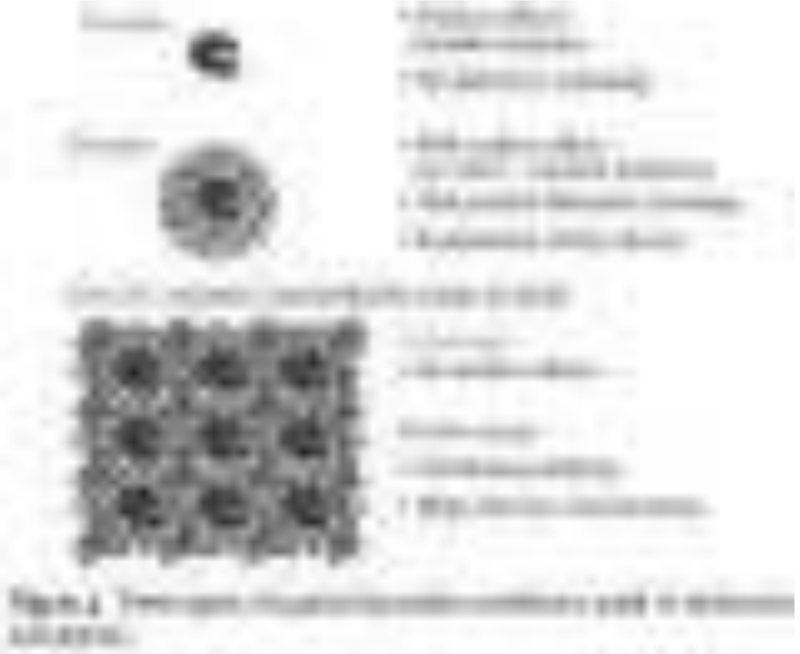


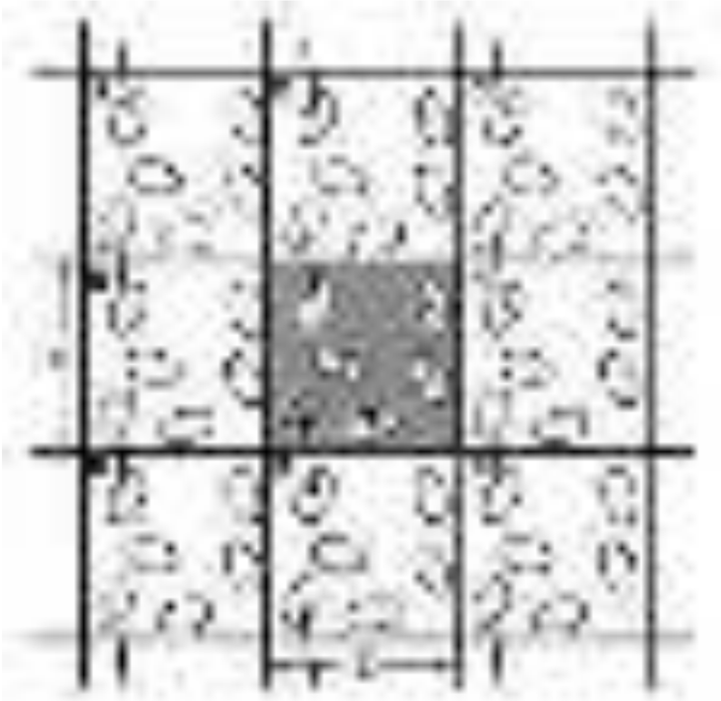
Periodic Boundary conditions



Molecular Dynamics Simulations

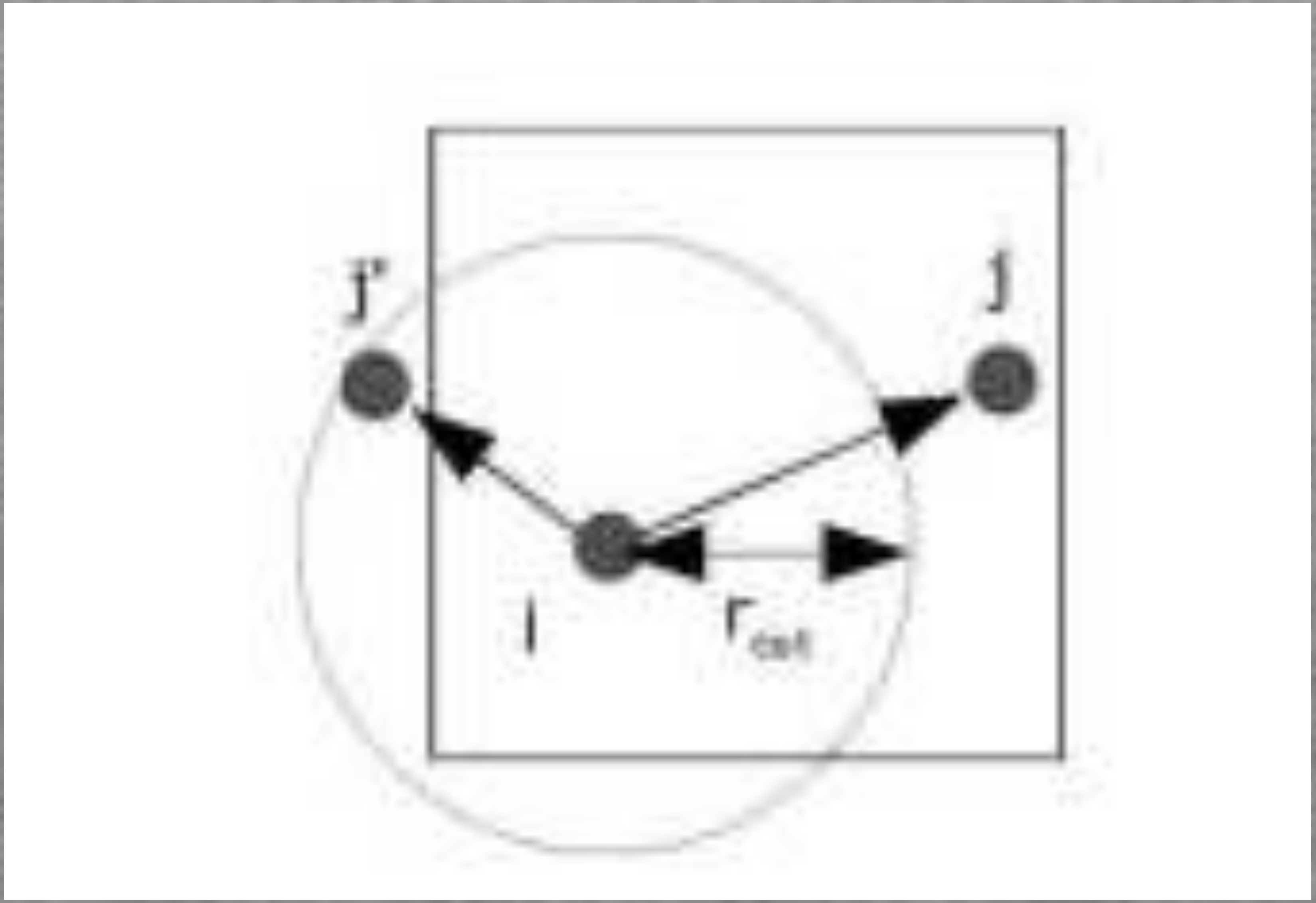
- **Periodic boundary conditions**
  - Our systems are very small 10^5 or 10^6 << Avogadro number
  - Boundaries will have a large effect on the calculated properties
  - Use periodic boundary conditions. The system is surrounded by translated copies of itself





van Gunsteren, W.F., et al., Biomolecular modeling: Goals, problems, perspectives. Angew. Chem. Int. Edit., 2006. 45(25): p. 4064-4092.

Cutt-off and minimun image convection



Steps in Molecular Dynamics Simulations

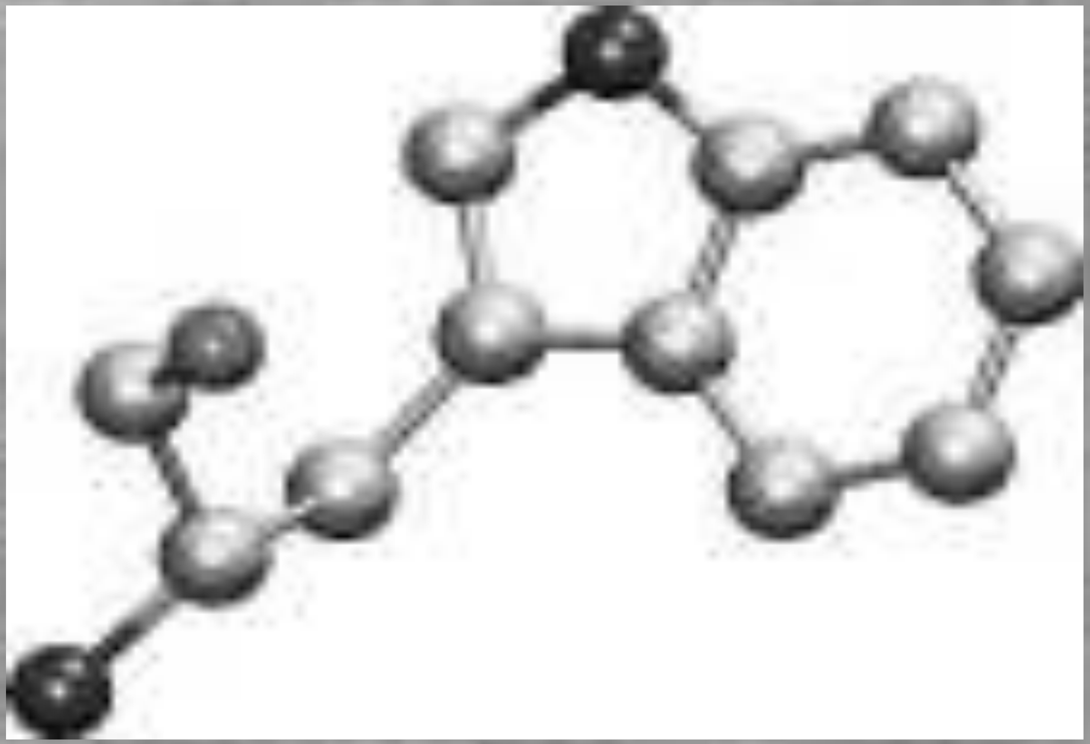
1) Build realistic atomistic model of the system

2) Simulate the behavior of your system over time using specific conditions (temperature, pressure, volume, etc)

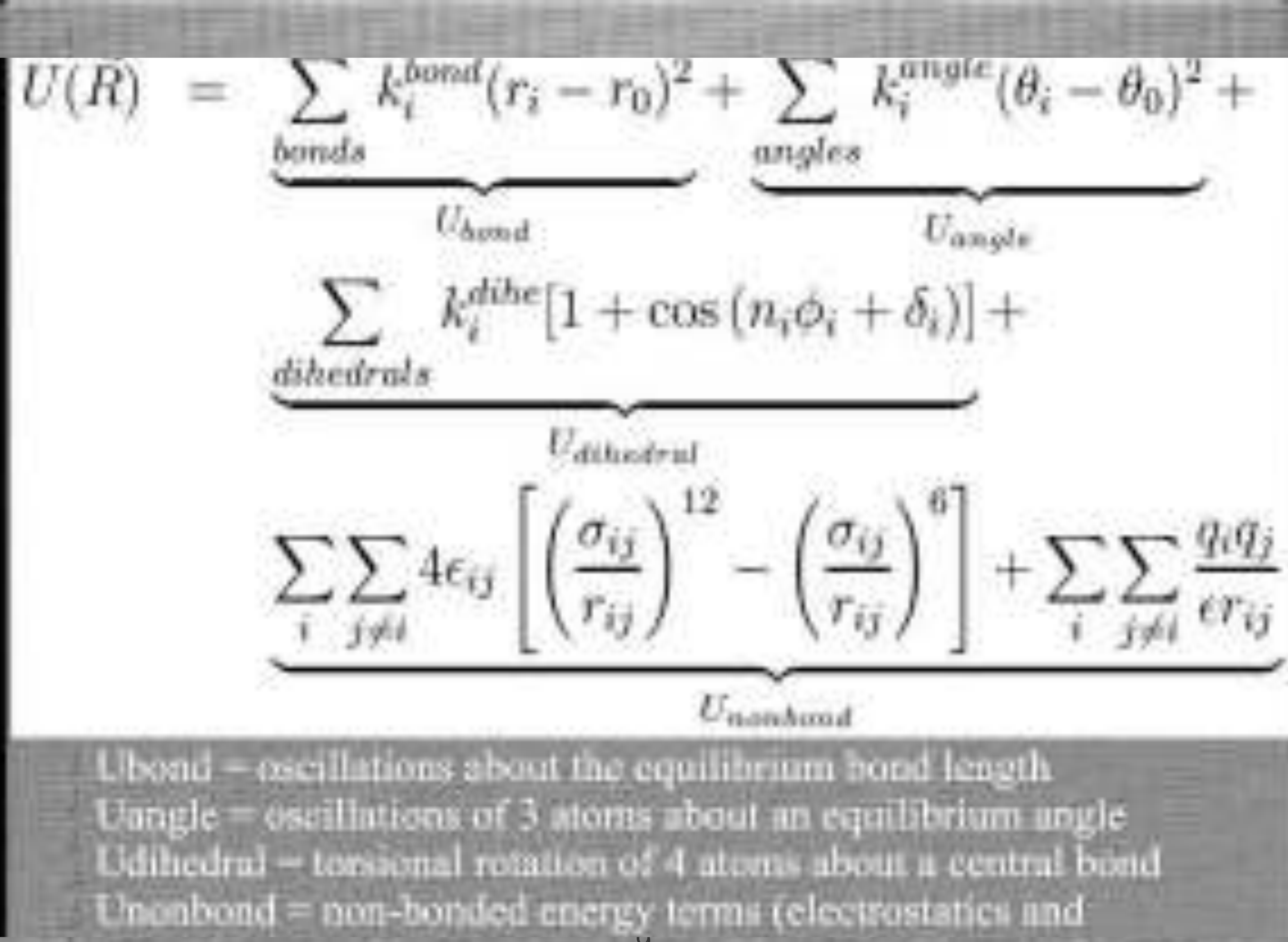
3) Analyze the results obtained from MD and relate to macroscopic level properties

What is a Forcefield?

In molecular dynamics a molecule is described as a series of charged points (atoms) linked by springs (bonds).



To describe the time evolution of bond lengths, bond angles and torsions, also the nonbond van der Waals and elecrostatic interactions between atoms, one uses a forcefield.
The forcefield is a collection of equations and associated constants designed to reproduce molecular geometry and selected properties of tested structures.



$$U(R) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} +$$

$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} +$$

$$\underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

Ubond = oscillations about the equilibrium bond length
Uangle = oscillations of 3 atoms about an equilibrium angle
Udihedral = torsional rotation of 4 atoms about a central bond
Unonbond = non-bonded energy terms (electrostatics and

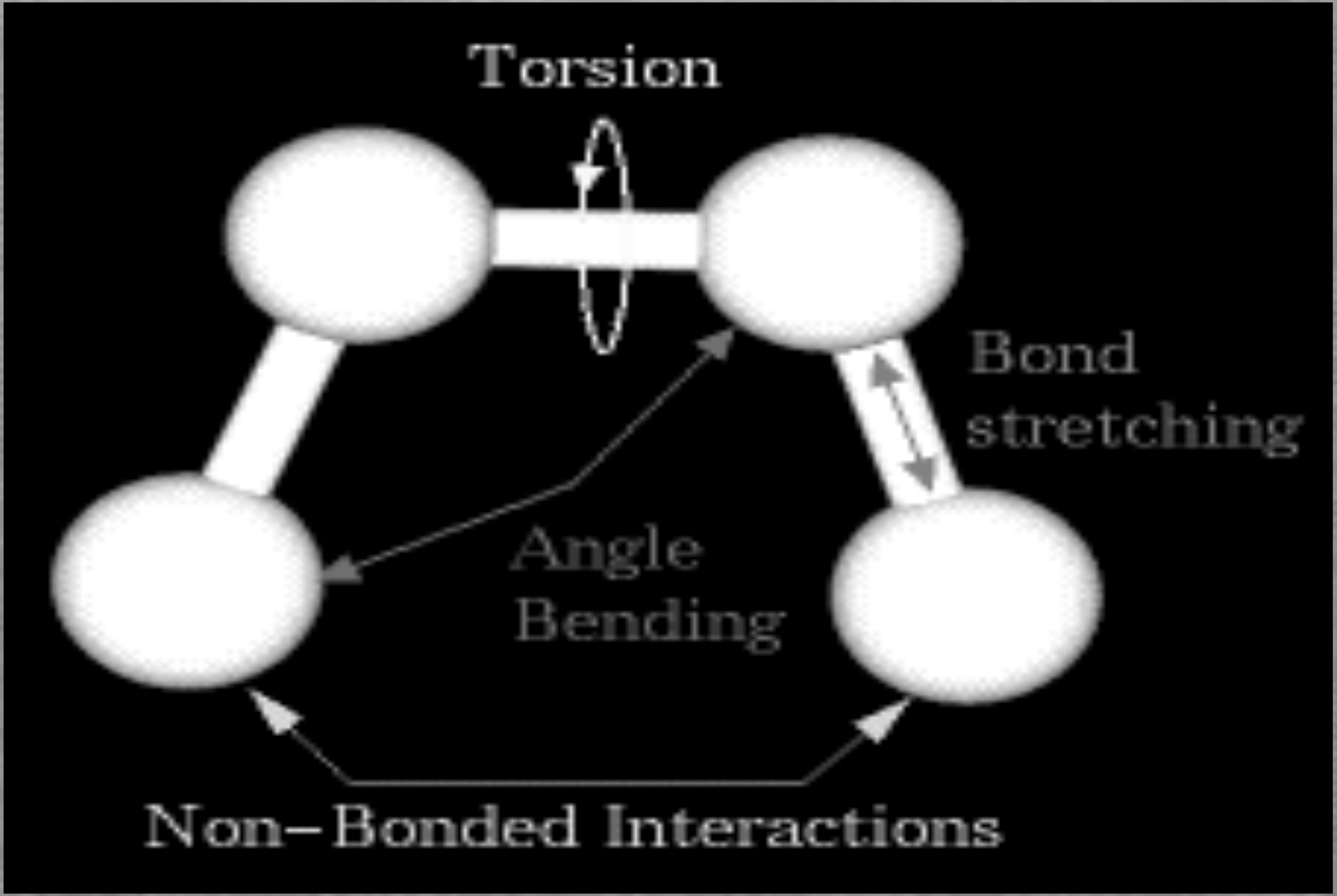
Energy Terms Described in the ………………… forcefield

Bond

Angle

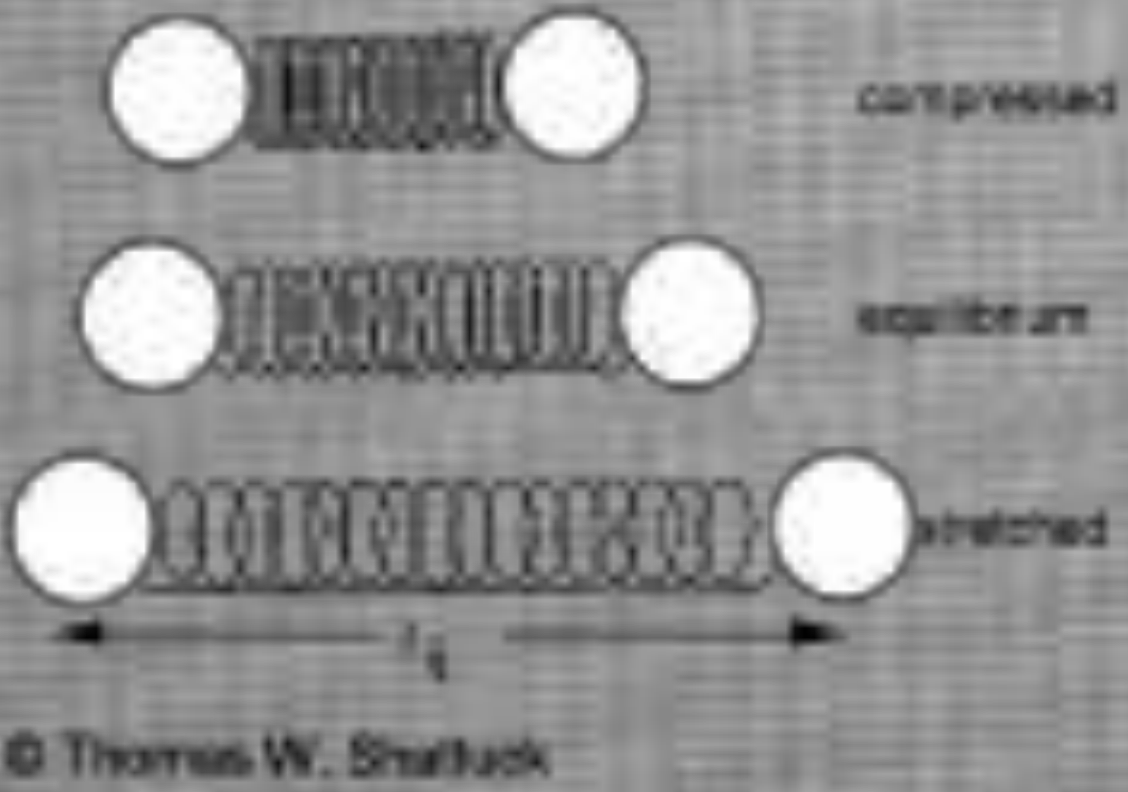
Dihedral

Improper

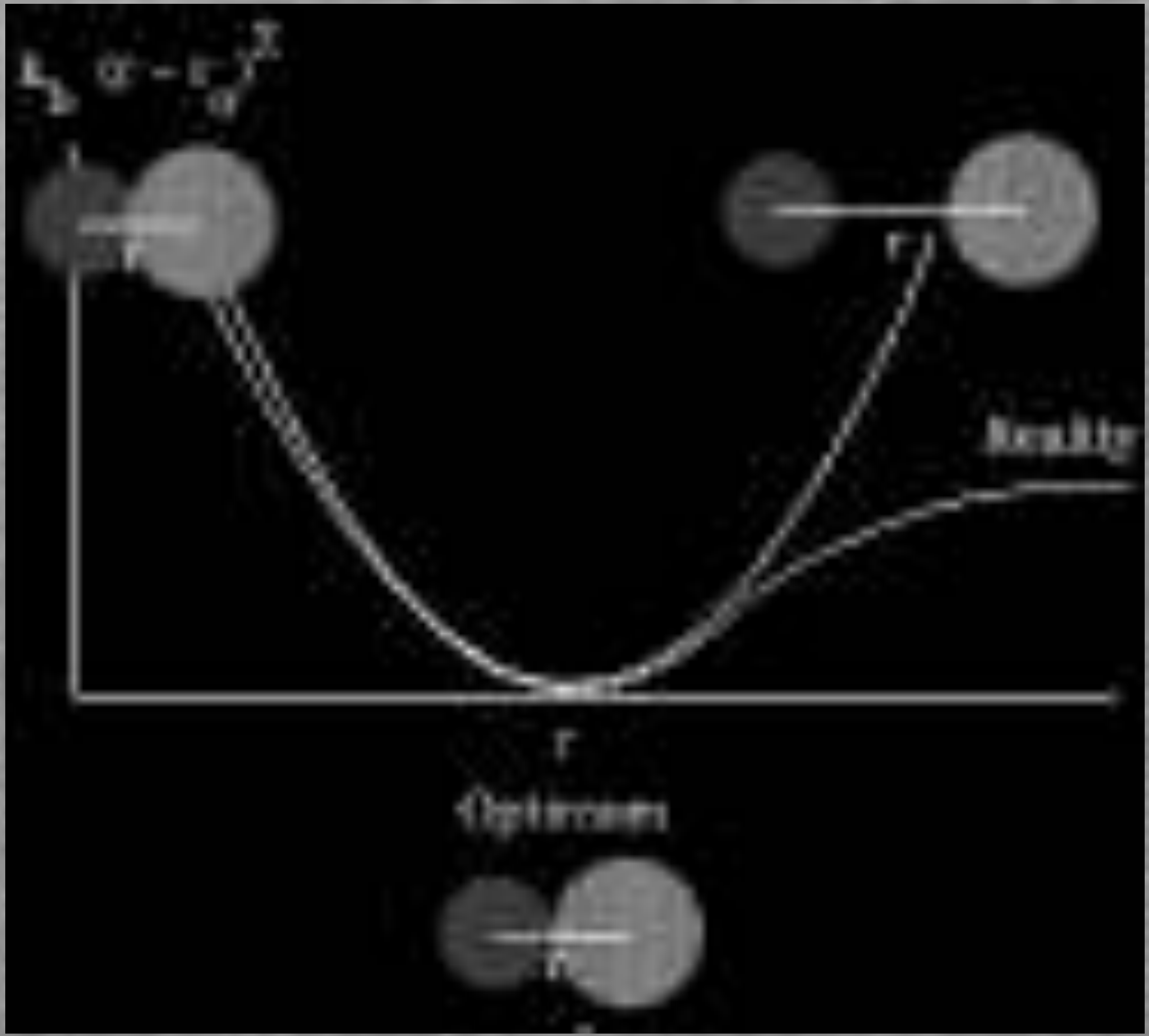

Torsion
Bond stretching
Angle Bending
Non–Bonded Interactions

Bonds:
Harmonic bond Stretching

$V(r_{ij})$ represents the energy required to stretch or compress a covalent bond:



© Thomas W. Shattuck

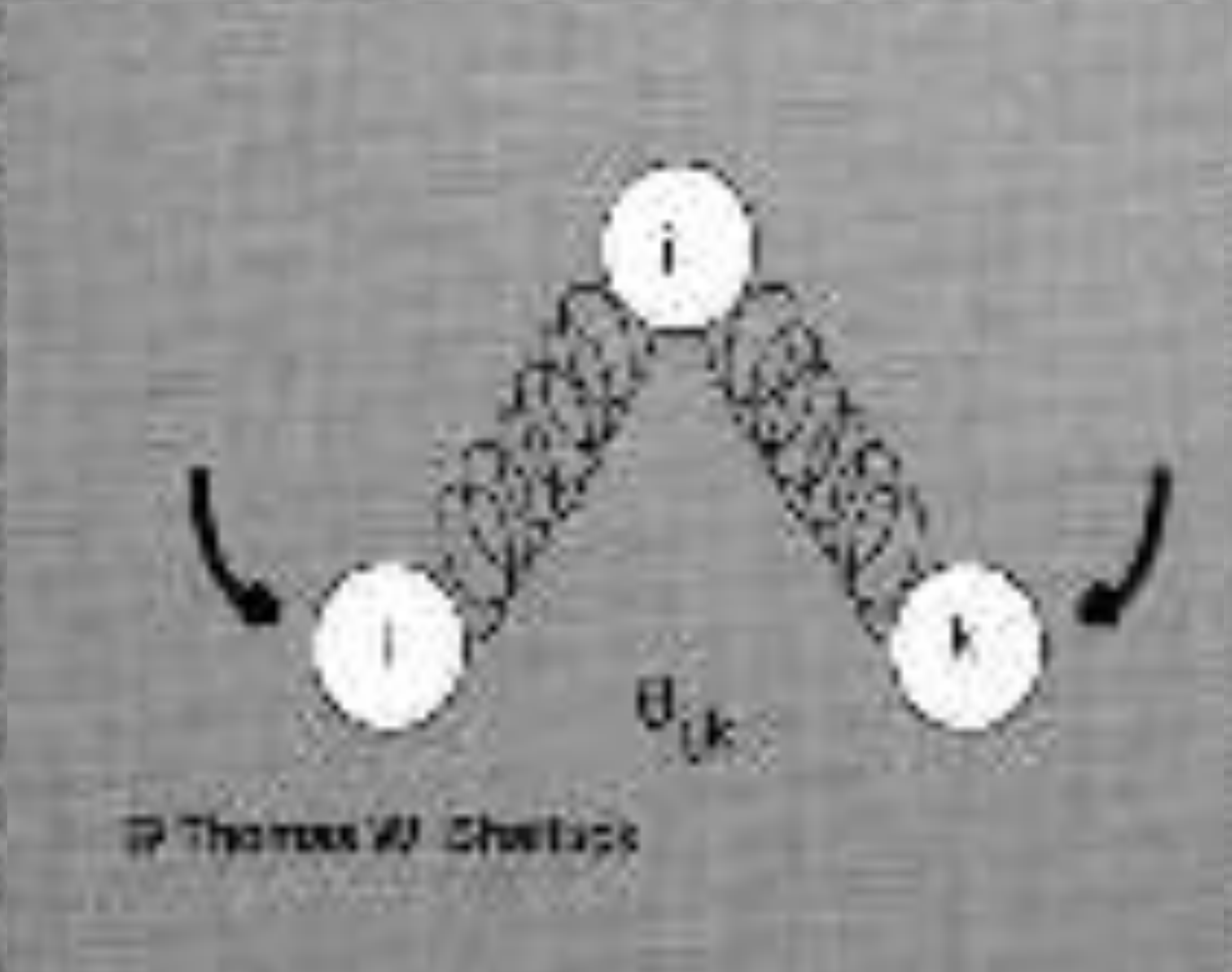


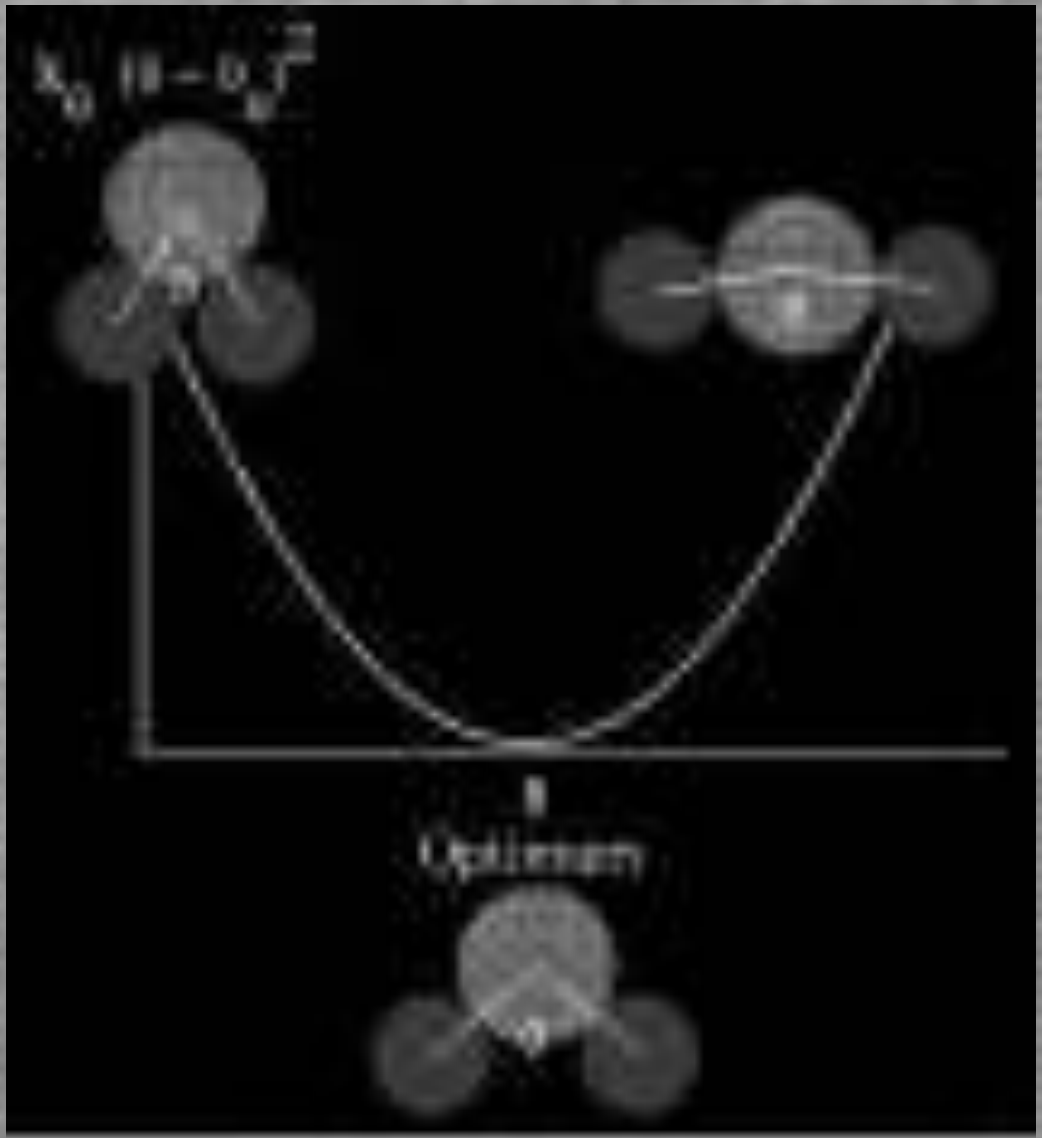
A bond can be thought of as a spring having its own equilibrium length, r_o, and the energy required to stretch or compress it can be approximated by the Hookean potential for an ideal spring:

$$E = \sum_{\text{bonds}} k_b (r - r_o)^2$$

Angles:
Harmonic angle bending

E_{bend} is the energy required to bend a bond from its equilibrium angle, θ_o:



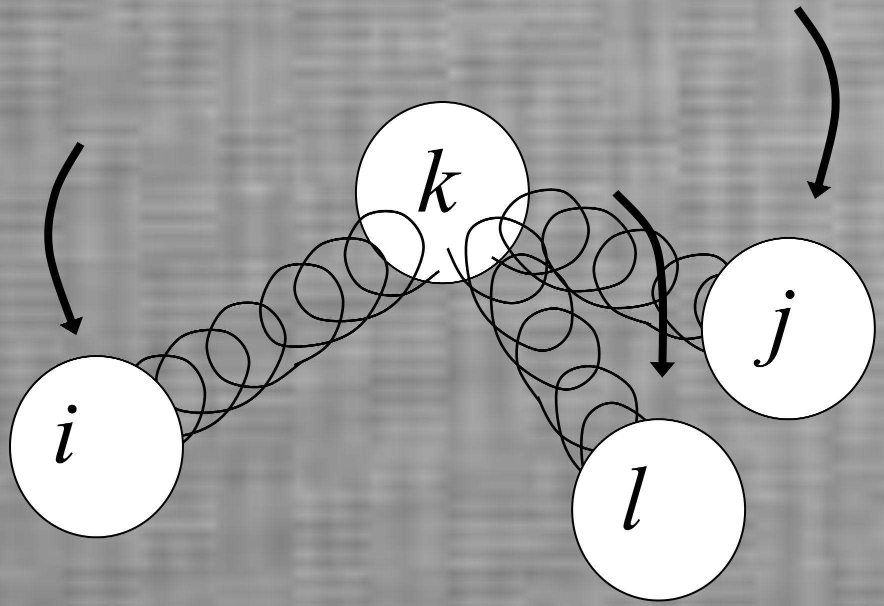


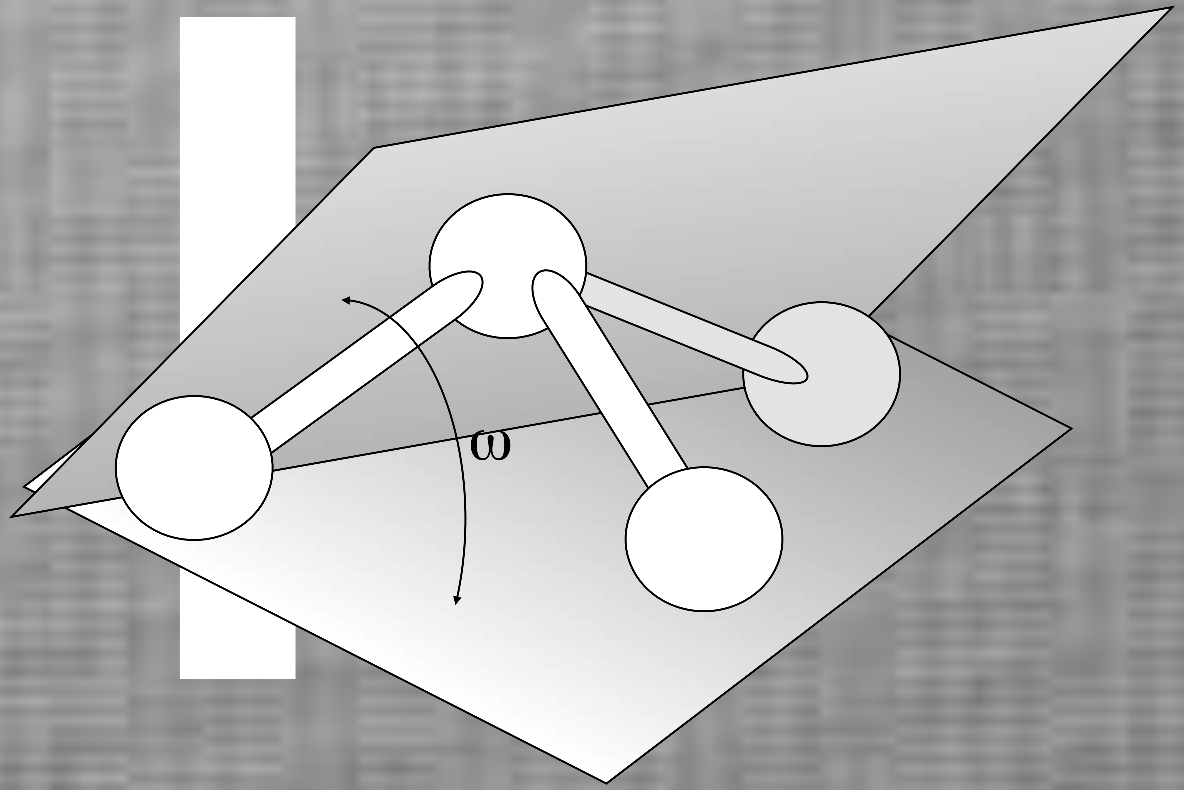
Again this system can be modeled by a spring, and the energy is given by the Hookean potential with respect to angle:

$$E = \sum_{angles} k_\theta (\theta - \theta_o)^2$$

Improper dihedrals: planarity and harmonic

$E_{improper}$ is the energy required to deform a planar group of atoms from its equilibrium angle, ω_o, usually equal to zero:





© Thomas W. Shattuck

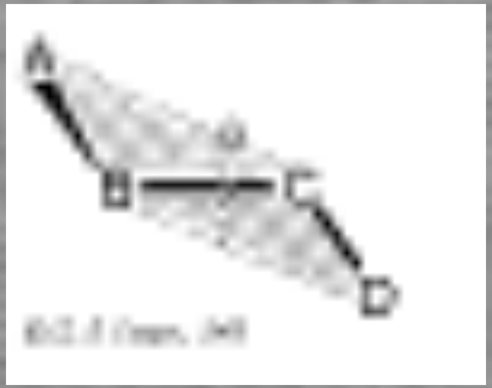
Again this system can be modeled by a spring, and the energy is given by the Hookean potential with respect to planar angle:

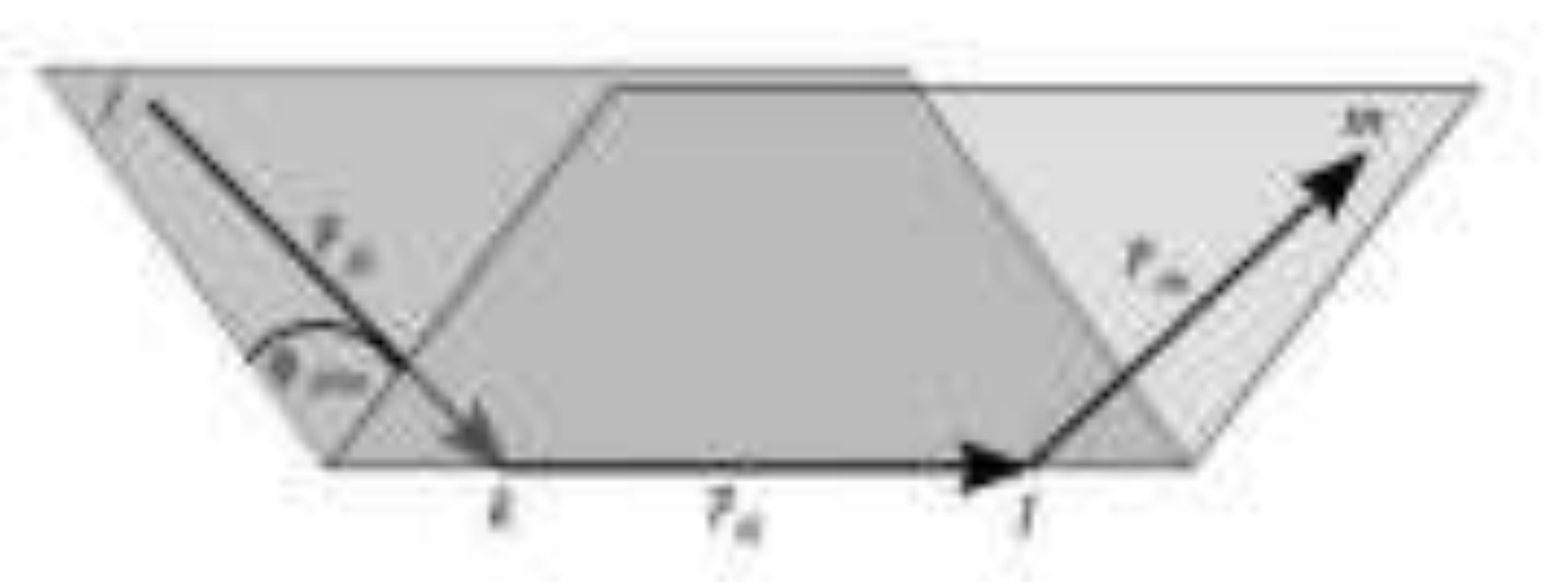
$$E_{improper} = \frac{1}{2} k_{o,ijkl} (\omega_{ijkl} - \omega_o)^2$$

Dihedral Angle potentials

The dihedral angle potentials describe the interaction arising from torsional forces in molecules. (They are sometimes referred to as torsion potentials.)

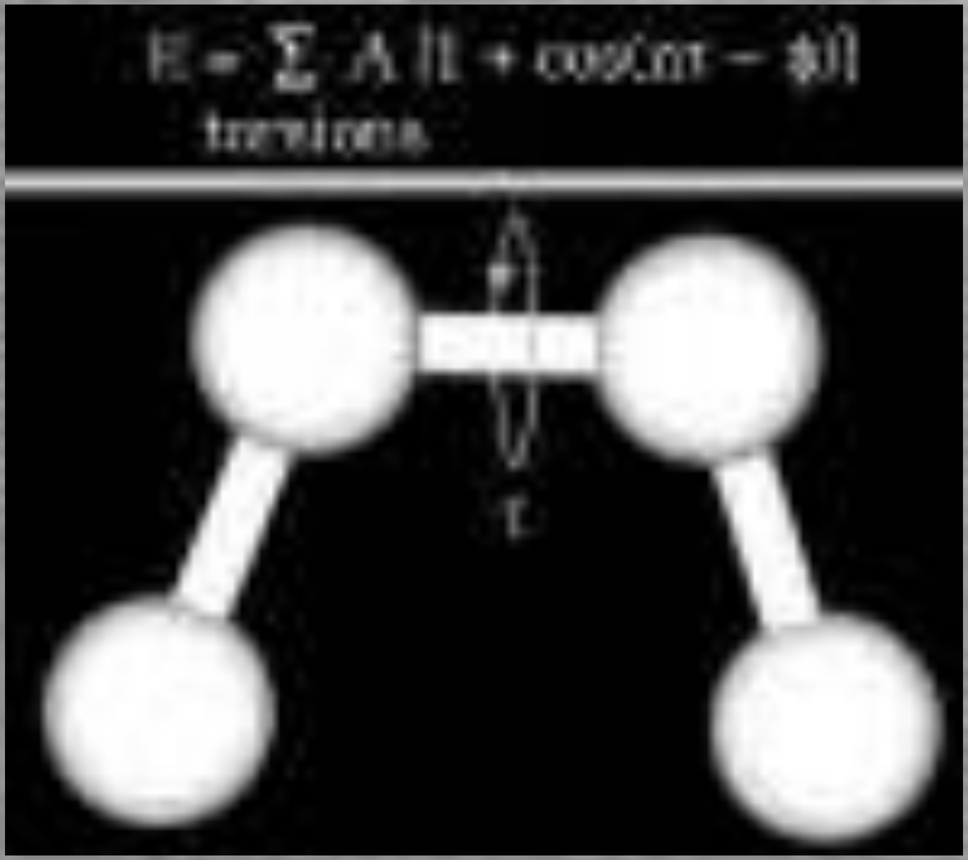
Formally the dihedral angle (also known as a torsion angle) between four atoms A-B-C-D is defined as the angle between the the planes ABC (marked by red lines) and BCD (marked in purple):

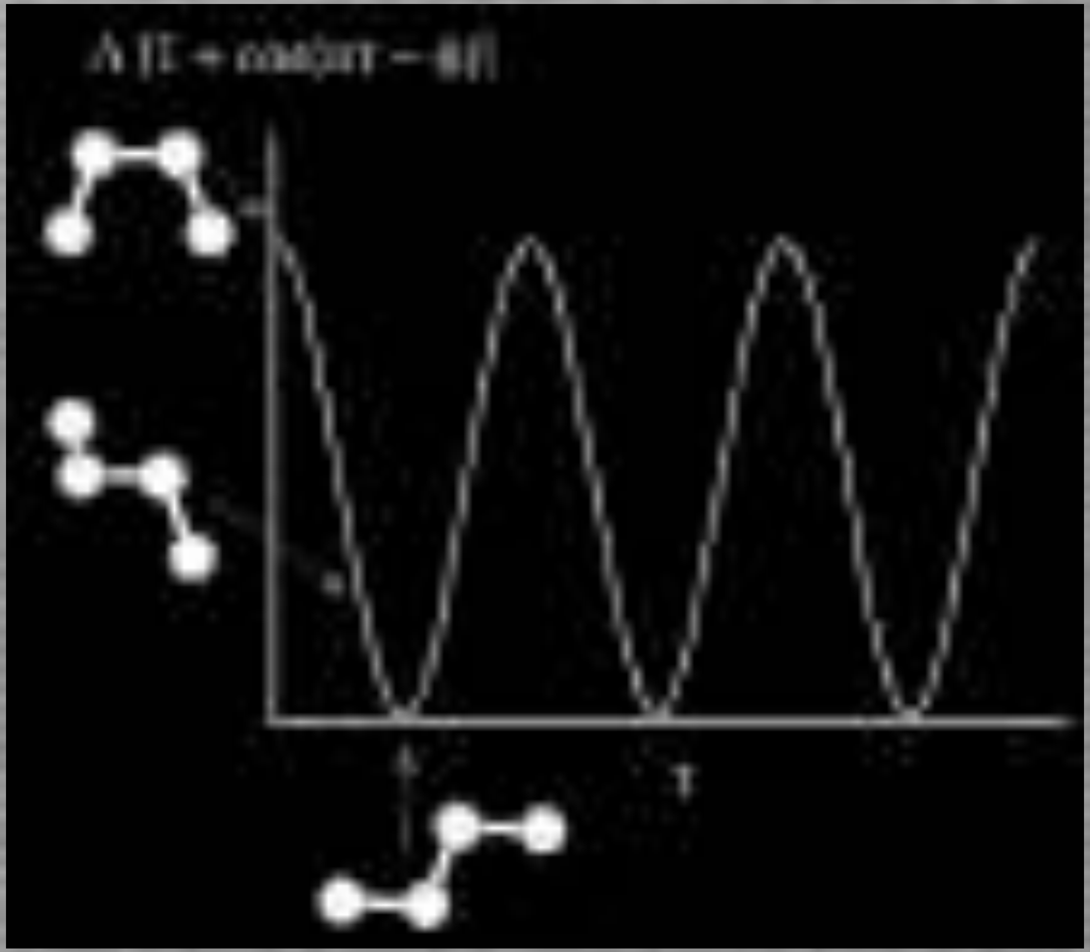




single cosine

E_{tor} is the energy of torsion needed to rotate about bonds:

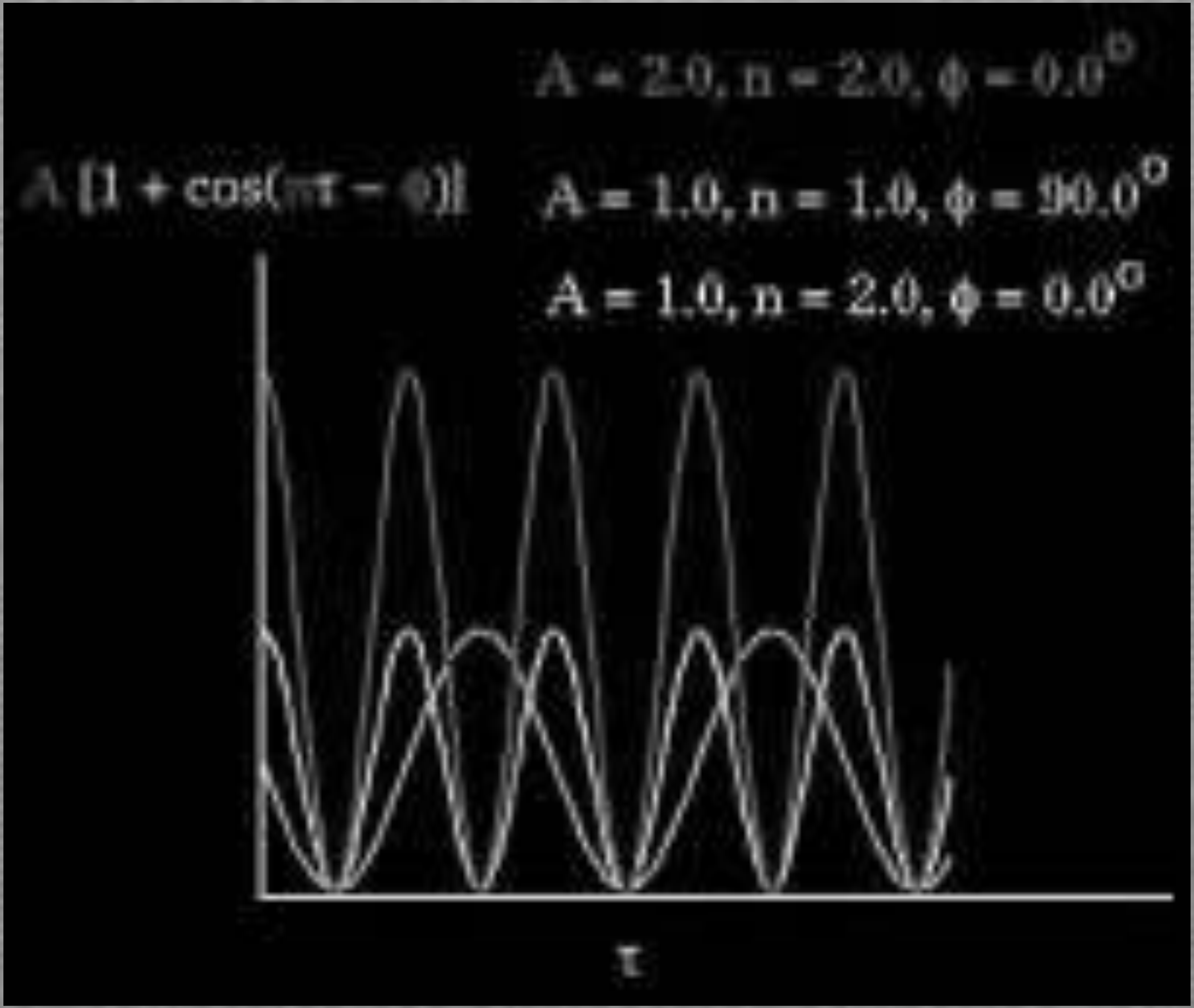




Torsional interactions are modeled by the potential:

$$E_{tor} = \tfrac{1}{2} k_{tor,1} (1 - \cos \phi) + \tfrac{1}{2} k_{tor,2} (1 - \cos 2\phi) + \tfrac{1}{2} k_{tor,3} (1 - \cos 3\phi)$$

asymmetry (butane) 2-fold groups e.g. COO- standard tetrahedral torsions


A = 2.0, n = 2.0, φ = 0.0°
A [1 + cos(nτ − φ)]
A = 1.0, n = 1.0, φ = 90.0°
A = 1.0, n = 2.0, φ = 0.0°
τ

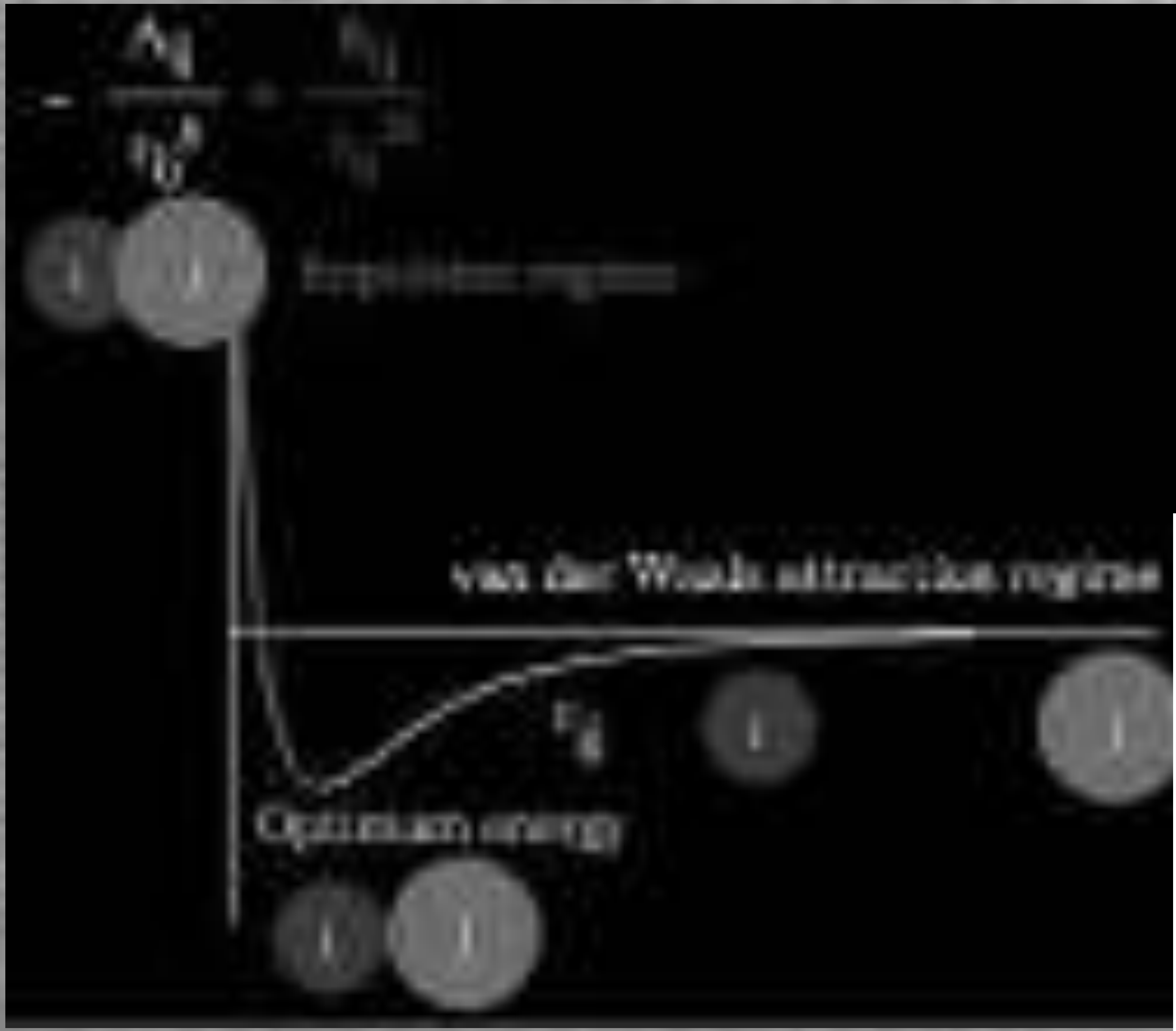
The intermolecular potential function (NON- bonded) short ranged interactions:

- Short ranged (**van der Waals)** potentials
- The three body potentials
- The Tersoff Covalent potentials
- Four Body potentials
- Metal potentials
- External fields

Short Ranged (van der Waals) Potentials

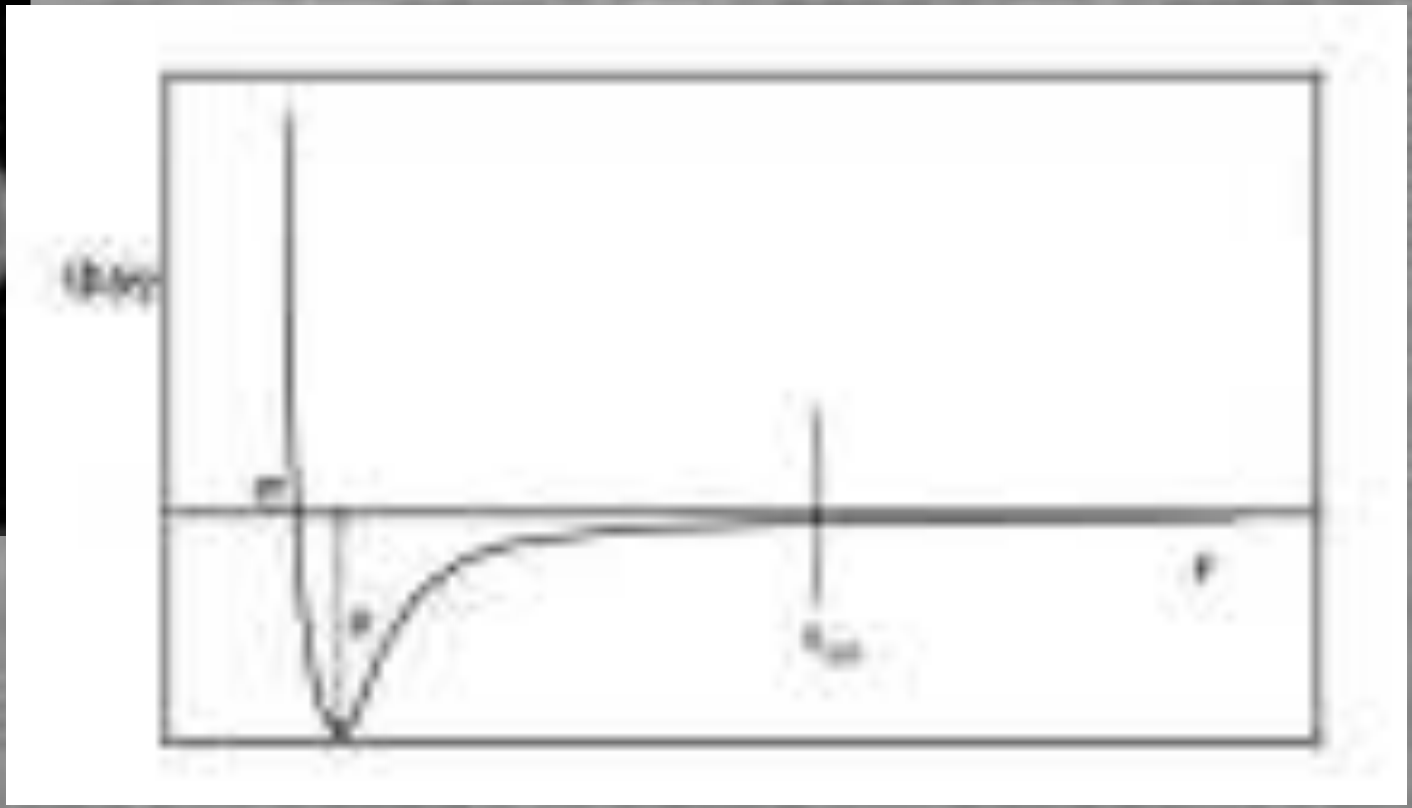
12-6 potential

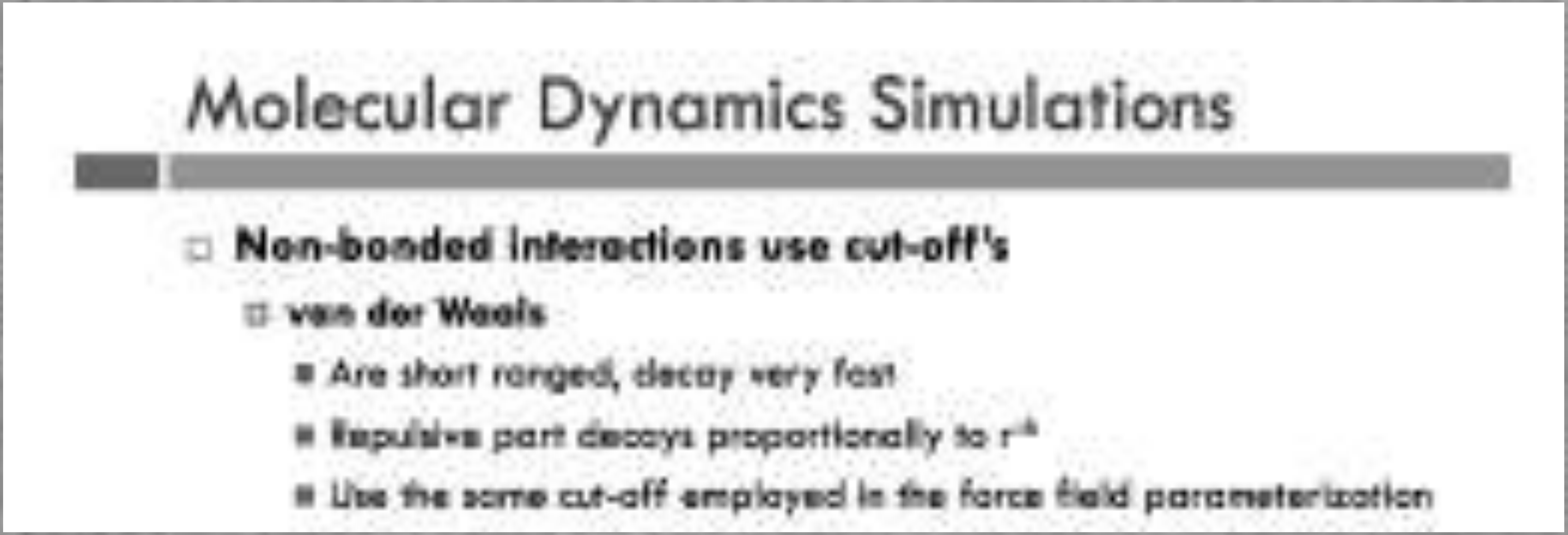
E_{vdW} is the steric exclusion and long-range attraction energy (QM origins):



$$E = \sum_i \sum_j \left(\frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} \right) + \sum_i \sum_j \frac{q_i q_j}{r_{ij}}$$

van der Waals term   Electrostatic term

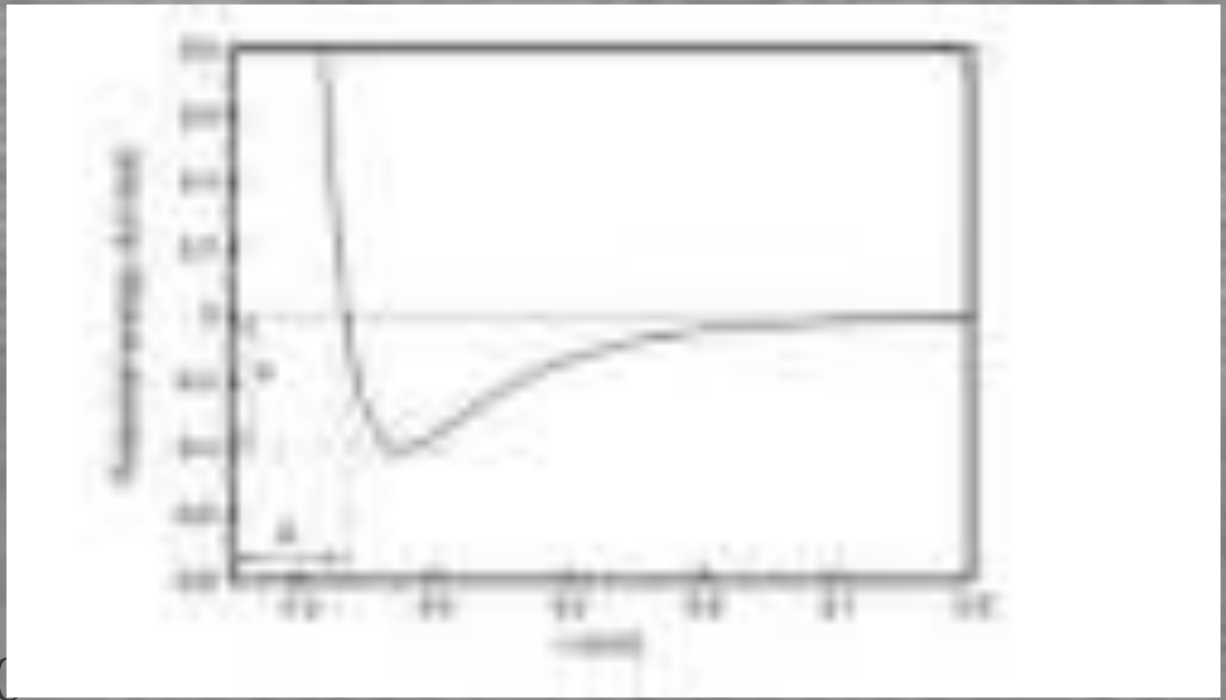




Molecular Dynamics Simulations

- Non-bonded interactions use cut-off's
  - van der Waals
    - Are short ranged, decay very fast
    - Repulsive part decays proportionally to r^{-6}
    - Use the same cut-off employed in the force field parameterization

DL_POLY uses two cut-offs (8 Å and 1.4 Å)



ILJ *vs.* LJ

$$V_{ILJ} = \varepsilon\left[\frac{m}{n(r)-m}\left(\frac{r_0}{r}\right)^{n(r)} - \frac{n(r)}{n(r)-m}\left(\frac{r_0}{r}\right)^{m}\right]$$

$$V_{LJ} = \varepsilon\left[\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^{6}\right]$$

$$n(r) = \beta + 4.0\left(\frac{r}{r_0}\right)^{2}$$

$$C_6 = -\frac{n(r)}{n(r)-m}\varepsilon r_o^6$$

less attractive long range

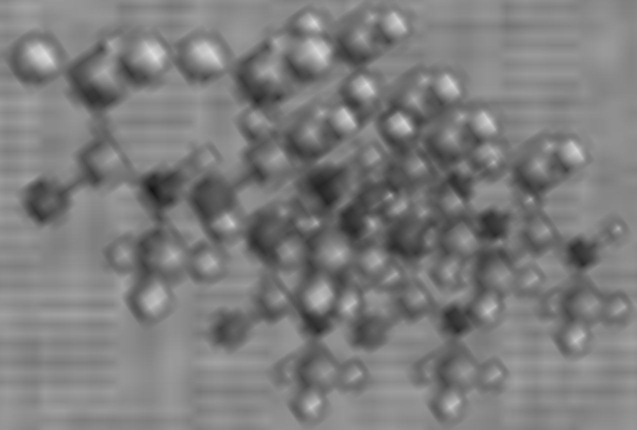
$$C_6 = -2\varepsilon r_o^6$$

$$n(r) = \beta + 4.0\left(\frac{r}{r_0}\right)^{2}$$

less repulsive short range

- + part of the formulation: represents the size repulsion contribution
- - part of the formulation: represents the effective dispersion attractive to the same pair
- β is an adjustable parameter (like ambient characteristic)
- use the same value of ε, r_0 for different molecules => structural meaning

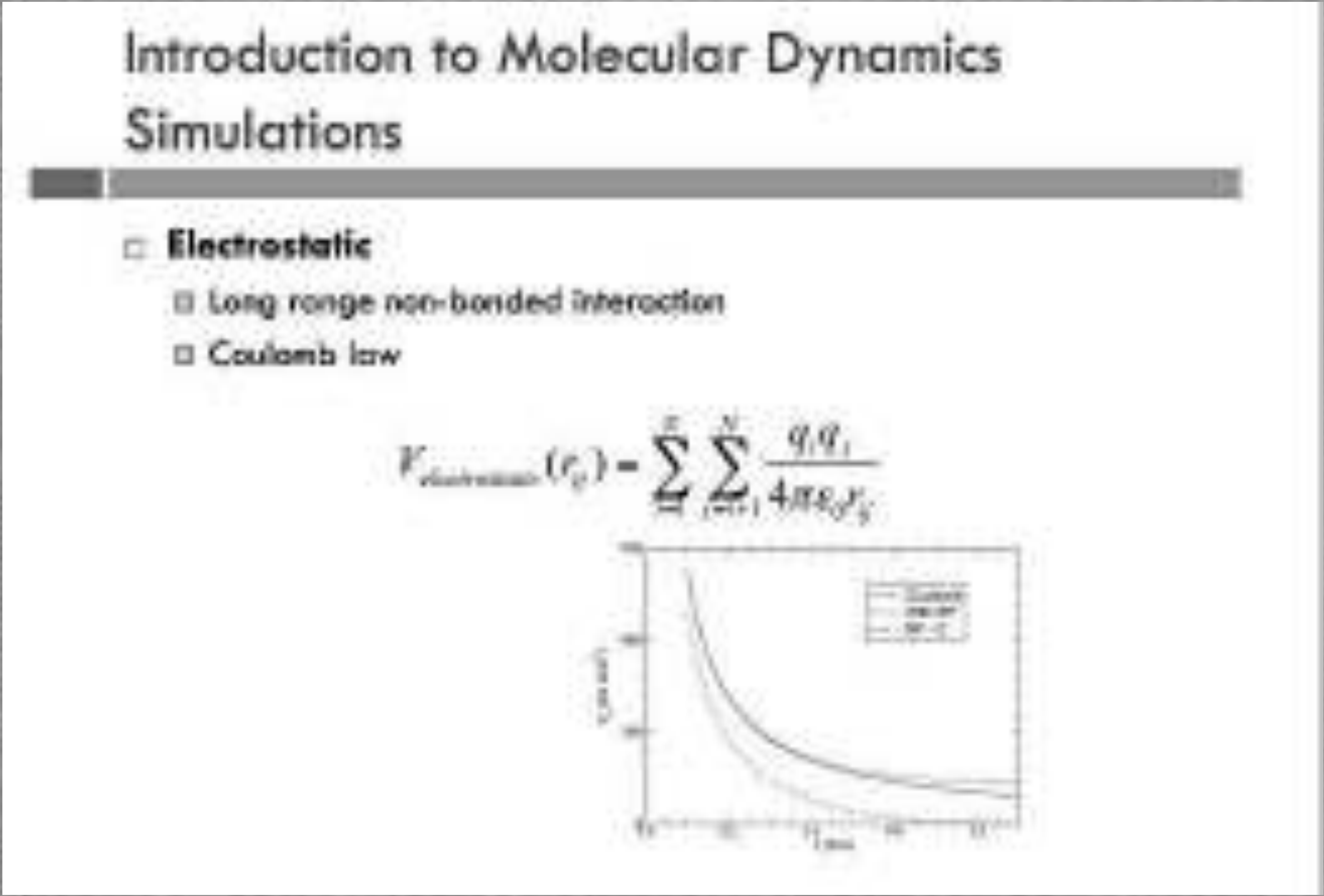
- β is sensible to the charge distribution and to the choice of the interaction of the reference centers.
- The initial value of β can be estimated from the cubic root of the polarizability of the involved partners.[1]
- due to the (i.e H bonding) for the charge transfer effects the values of β need to be varied within a well defined interval for optimization.



[1] Capittelli et al., *Chem. Phys.*, 2007, **338**, 62.

Long ranged Electrostatic (Coulombic)interactions

- **Direct Coulomb Sum**
- Truncated and Shifted Coulomb sum
- **Ewald Sum**
- Smoothed Particle Mesh Ewald (SPME)
- Hautman Klein Ewald (HKE)for systems with 2D periodicity
- Reaction Field
- Dynamics shell model
- Relaxed shell model



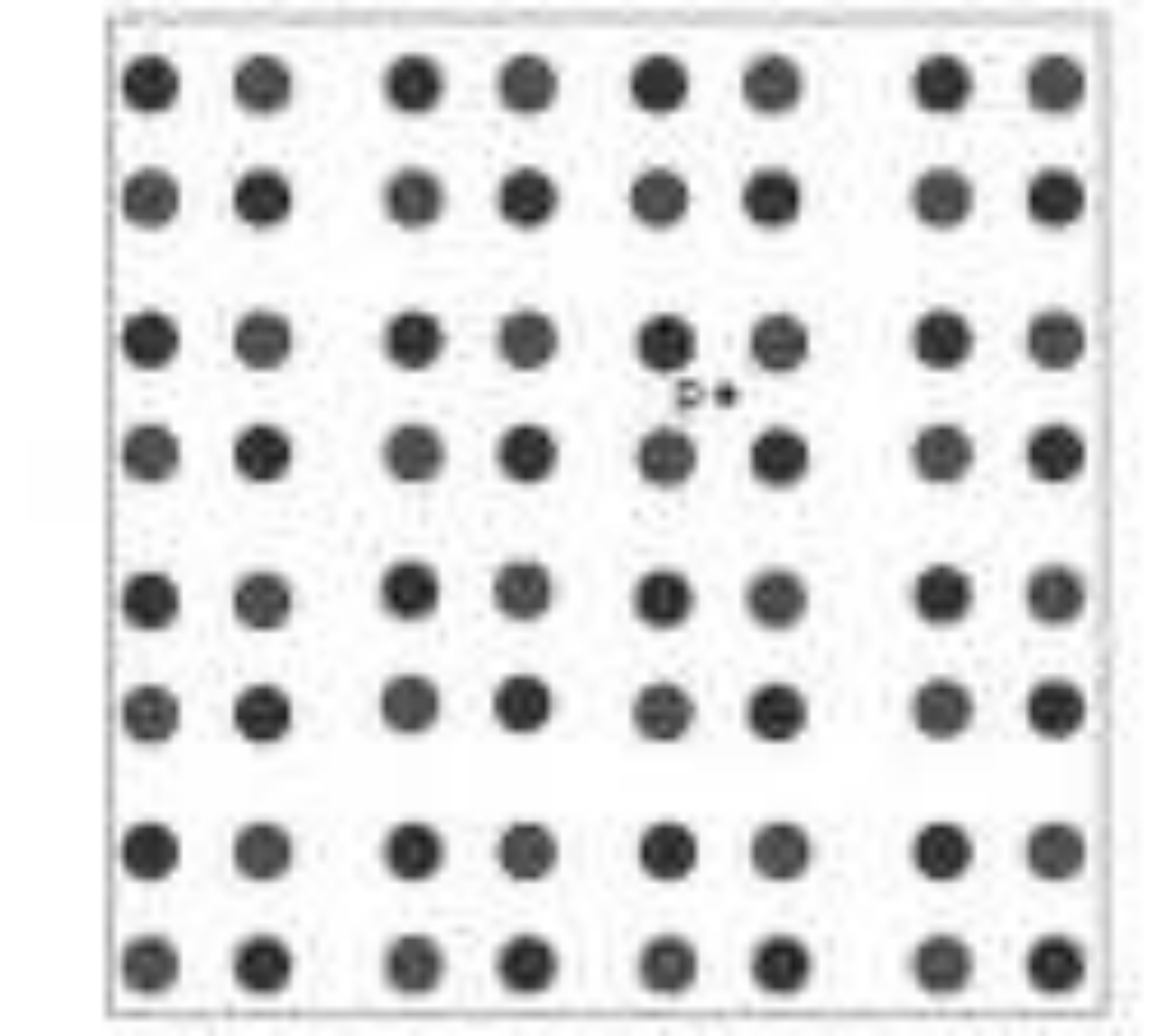
Introduction to Molecular Dynamics Simulations

- **Electrostatic**
  - Long range non-bonded interaction
  - Coulomb law

$$V_{electrostatic}(r_{ij}) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

Ewald sum

$$\Phi^C(\{\vec{r}_i\}) = \frac{1}{4\pi\epsilon_0} \sum_{\vec{L}=\vec{0}}^{\infty} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{q_i q_j}{|\vec{L}+\vec{r}_j-\vec{r}_i|},$$

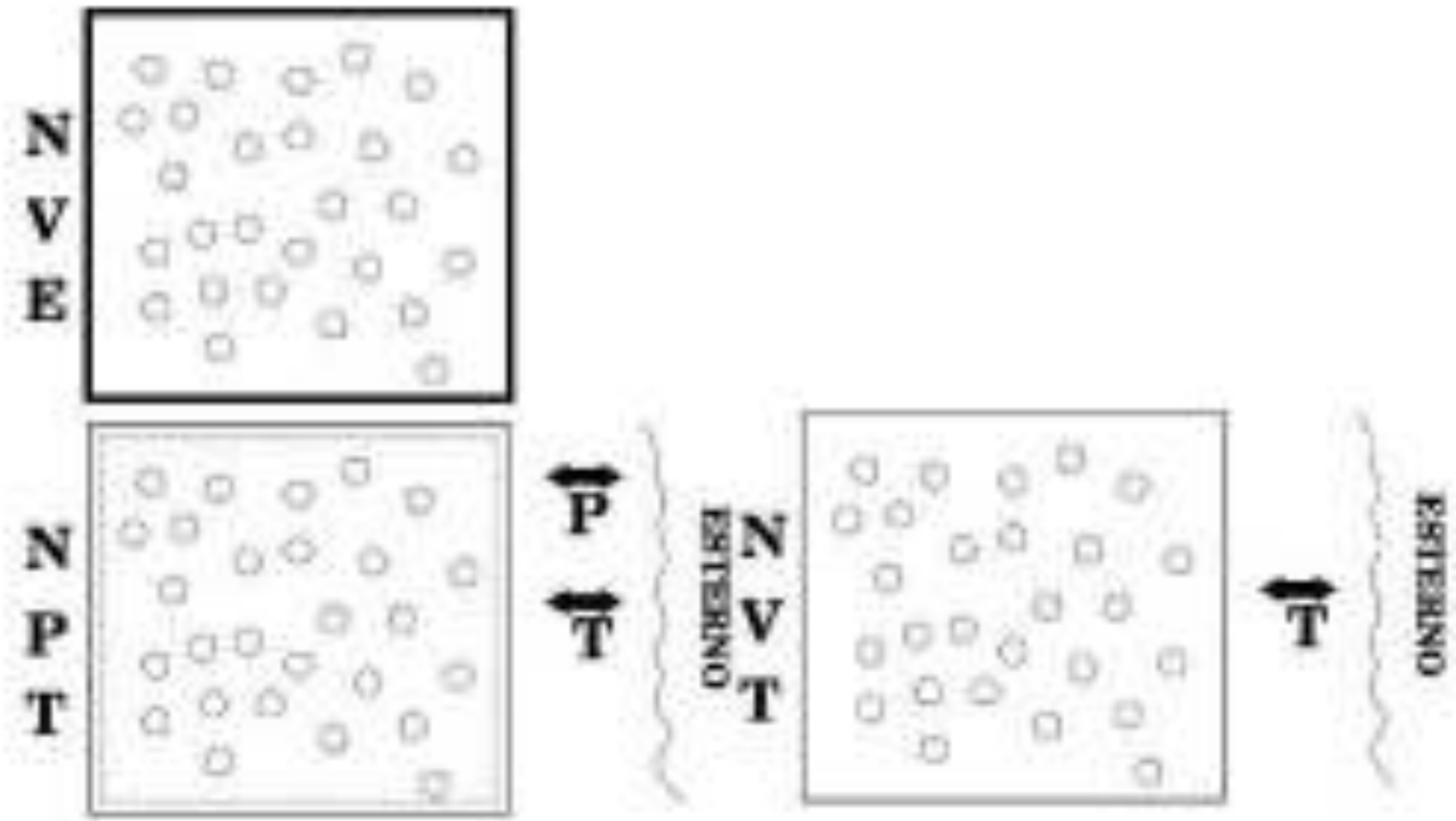


A “Probe” Unit Positive Charge in a Periodic Lattice of Ions.

Type of data	Type of system	Phase	Type of properties	Force field parameters
structural data (expt)	small molecules	crystalline solid phase	molecular geometry: bond lengths, bond angles	b_0, θ_0, ξ_0
spectroscopic data (expt)	small molecules	gas phase	molecular vibrations: force constants	K_b, K_θ, K_ξ
thermodynamic data (expt)	small molecules, mixtures, solutions	condensed phase	heat of vaporization, density, partition coefficient, free energy of solvation	van der Waals: $C_6(i,j)$, $C_{12}(i,j)$, q(foo
dielectric data (expt)	small molecules	condensed phase	dielectric permittivity, relaxation	charges q_i
transport data (expt)	small molecules	condensed phase	diffusion and viscosity coefficients	$C_6(i,j)$, $C_{12}(i,j)$, q_i
electron densities (theor.)	small molecules	gas phase	quantum-chemical calculation of atom charges	charges q_i(initial)

van Gunsteren, W.F., et al., Biomolecular modeling: Goals, problems, perspectives. Angew. Chem. Int. Edit., 2006. 45(25): p. 4064-4092.

Microcanonical emsembles


NVE
NPT
P
T
ESTERNO
NVT
T
ESTERNO

Molecular Dynamics Simulations

- **Ensembles**
  - **Isoenergetic**
    - System in a closed, rigid and insulated container (NVE)
    - The system has constant energy (E)
    - All states are equally likely

$$P(s) = \frac{1}{\Omega(E)}$$

where Ω(E) are the number of states with energy E

Baptista, AM. Molecular Simulation Group,Instituto de Tecnologia Química e Biológica. www.itqb.unl.pt/simulation. Personal communication

Molecular Dynamics Simulations

- **Ensembles**
  - **Isothermal**
    - Closed system at constant temperature and volume (NVT)
    - Energy can be exchanged with the bath at temperature T
    - A state with low energy is more likely than a state with high energy

$$P(s) = \frac{e^{-E_s/RT}}{\sum_s e^{-E_s/RT}}$$

Baptista, AM. Molecular Simulation Group,Instituto de Tecnologia Química e Biológica. www.itqb.unl.pt/simulation. Personal communication

Molecular Dynamics Simulations

- **Ensembles**
  - **Isothermal-isobaric**
    - Closed system connected to a piston and at constant temperature (NPT)
    - Energy can be exchanged with the bath at temperature T
    - Volume (V) can be "exchanged" with the bath at pressure P
    - A state with low E+pV is more likely than a state with high E+pV

$$P(s) = \frac{e^{-\left(E_s + pV_s\right)/RT}}{\sum_s e^{-\left(E_s + pV_s\right)/RT}}$$

Baptista, AM. Molecular Simulation Group,Instituto de Tecnologia Química e Biológica. www.itqb.unl.pt/simulation. Personal communication

Molecular Dynamics Simulations

- **Ensembles**
  - **Grand canonical ensemble**
    - System in a rigid container able to exchange particles with a reservoir and, in contact with a heat bath (N μ T)
    - Energy can be exchanged with the bath at temperature T
    - The number of molecules (N) can vary by exchange with the reservoir at chemical potential μ

$$P(s) = \frac{e^{-\left(E_s - \mu N_s\right)/RT}}{\sum_s e^{-\left(E_s - \mu N_s\right)/RT}}$$

Baptista, AM. Molecular Simulation Group,Instituto de Tecnologia Química e Biológica. www.itqb.unl.pt/simulation. Personal communication

Molecular Dynamics Simulations

- **Thermostats**
  - **Berendsen**
    - The system is coupled to an external heat bath with fixed T_{bath}
    - The heat bath supplies/removes energy to/from the system
    - Velocities are scaled proportionally to $T_{bath} - T(t)$

$$\frac{dT(t)}{dt} = \frac{1}{\tau}\left(T_{bath} - T(t)\right)$$

τ determines the how tightly the bath and the system are coupled together (ps). Usually $\tau = 0.4$ ps.

- Scaling factor for the velocities

$$\lambda^2 = 1 + \frac{\delta t}{\tau}\left(\frac{T_{bath}}{T(t)} - 1\right)$$

Leach, A.R., Molecular modelling: principles and applications. 2nd ed. 2001: Prentice Hall.

Molecular Dynamics Simulations

- **Thermostats**
  - **Nosé-Hoover**
    - Equations of motion are changed by introducing a friction force
    - The friction force is proportional to the particle velocity and friction parameter ξ

$$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i} - \xi \frac{dr_i}{dt}$$

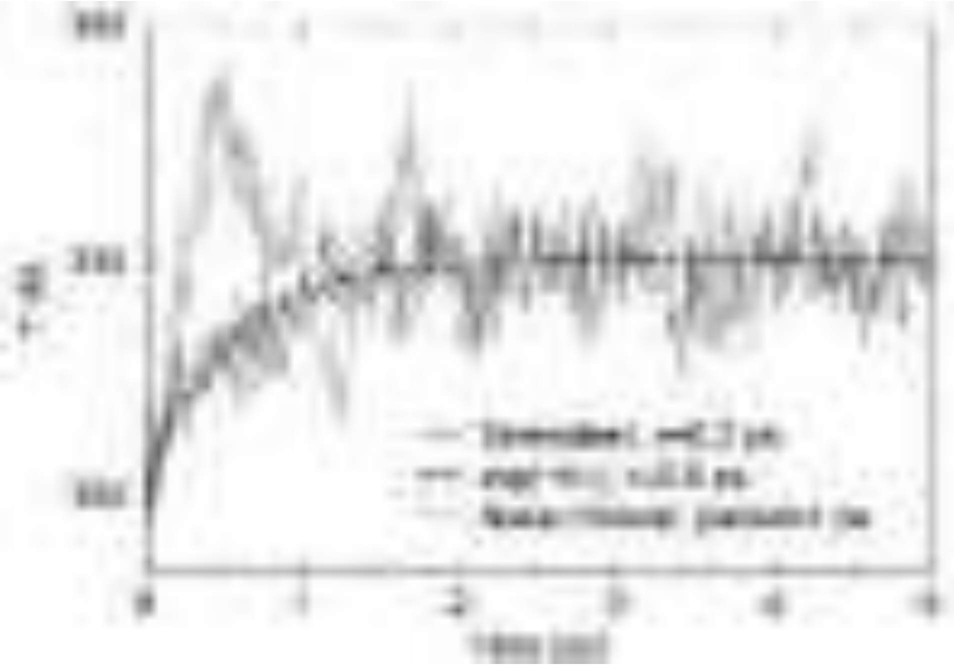
where the friction parameter ξ is given by

$$\frac{d\xi}{dt} = \frac{1}{Q}\left(T - T_0\right)$$

Leach, A.R., Molecular modelling: principles and applications. 2nd ed. 2001: Prentice Hall.

Molecular Dynamics Simulations

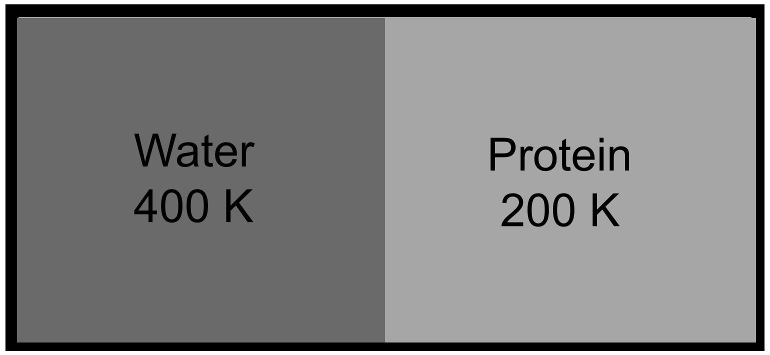
- **Difference between Berendsen and Nosé thermostats**
  - Berendsen gives exponential decay of the system towards de desired temperature. Does not generate a proper ensemble
  - Nosé can induce oscillations on the system. Gives a correct ensemble
  - Example: Temperature increase of a water box (300 K to 330 K)



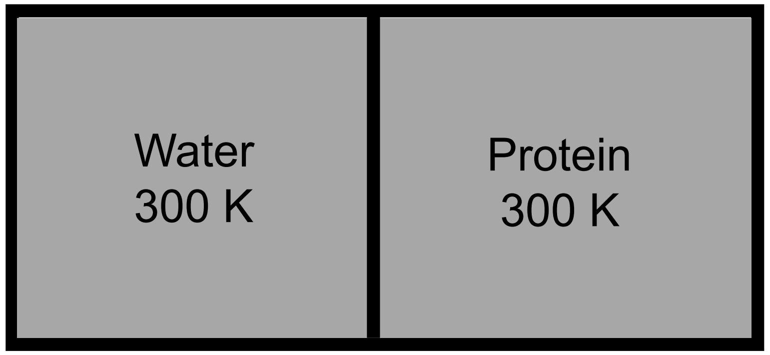
Hess, B. Gromacs Workshop at CSC 2007. http://www.csc.fi/english/research/sciences/chemistry/courses/gmx2007. Personal communication

Molecular Dynamics Simulations

- **Thermostats**
  - **Heat groups**
    - Energy exchange between different components is not perfect
    - Different components can have different temperatures but the global temperature system is “correct”
    - Heat group concept



1 heat group



2 heat groups

Molecular Dynamics Simulations

- **Barostats**
  - **Berendsen**
    - Analogous to the Berendsen thermostat
    - The system maintains constant pressure by changing its volume
    - The rate of change of the pressure is given by:

$$\frac{dP(t)}{dt} = \frac{1}{\tau_P}\left(P_{bath} - P(t)\right)$$

τ_P determines the how tightly the bath and the system are coupled together (ps). Usually $\tau_P = 1.0$ ps.

    - Scaling factor for the box. Atomic coordinates are scaled by $\lambda^{1/3}$

$$\mu^2 = 1 - \kappa \frac{\delta t}{\tau_P}\left(P(t) - P_{bath}\right)$$

κ is the experimental isothermal compressibility.

Leach, A.R., Molecular modelling: principles and applications. 2nd ed. 2001: Prentice Hall.

Molecular Dynamics Simulations

- **Integration of the equations of motion**
  - Integration is broken in many small fixed time δt steps
  - The algorithms must:
    - Conserve energy and momentum
    - Be time-reversible
    - Permit a long time step δt
  - Algorithms used:
    - Verlet algorithm
    - Leap-frog algorithm

Leach, A.R., Molecular modelling: principles and applications. 2nd ed. 2001: Prentice Hall.

Molecular Dynamics Simulations

- **Verlet algorithm**
  - Uses the positions and accelerations at time t, and the positions from the previous step, **r**(t-δt), to calculate the new positions at t +δt, **r**(t+δt).

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

  - Velocities are obtained by:

$$\boldsymbol{v}(t) = \left[\boldsymbol{r}(t+\delta t) - \boldsymbol{r}(t-\delta t)\right] / \, 2\delta t$$

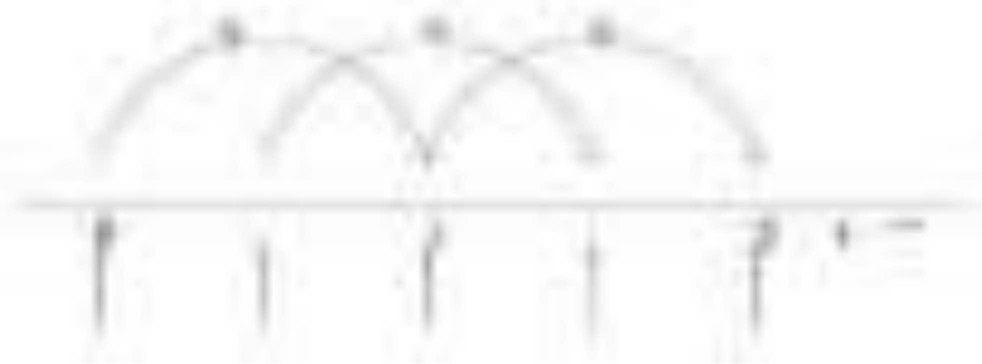
Leach, A.R., Molecular modelling: principles and applications. 2nd ed. 2001: Prentice Hall.

Molecular Dynamics Simulations

- **Verlet algorithm**
  - Positions and velocities are not synchronized
  - Numerically more stable

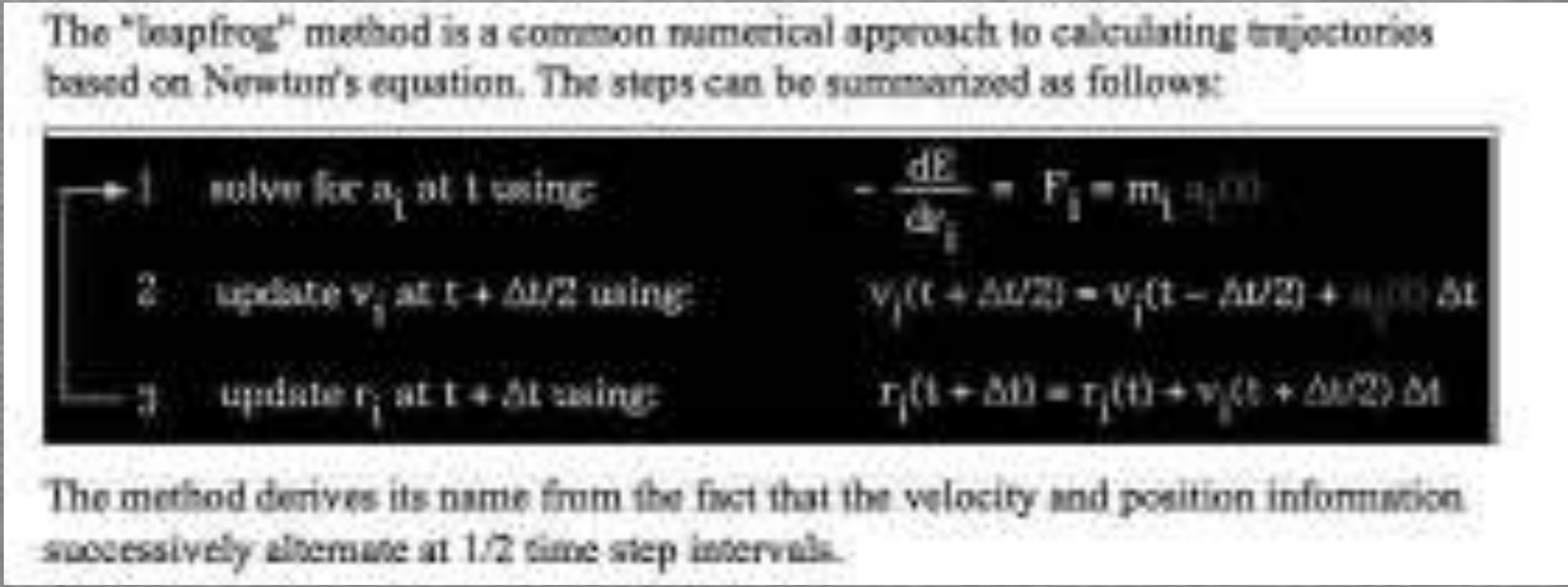
$$\boldsymbol{r}(t+\delta t)=\boldsymbol{r}(t)+\delta t\boldsymbol{v}\left(t+\frac{1}{2}\delta t\right)$$

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



Leach, A.R., Molecular modelling: principles and applications. 2nd ed. 2001: Prentice Hall.

The difference between the Verlet algorithm and the **leapfrog algorithm** is that the latter evaluates the velocities at half-integer time steps, which are artificial points in time, exactly half way between two time steps. The algorithm uses these half-integer time steps velocities to compute the new positions. To derive the leapfrog algorithm from the Verlet algorithm let us first define the velocity at half-integer time steps



The "leapfrog" method is a common numerical approach to calculating trajectories based on Newton's equation. The steps can be summarized as follows:

1 solve for a_i at t using: $-\frac{dE}{dr_i} = F_i = m_i a_i(t)$

2 update v_i at $t + \Delta t/2$ using: $v_i(t + \Delta t/2) = v_i(t - \Delta t/2) + a_i(t)\,\Delta t$

3 update r_i at $t + \Delta t$ using: $r_i(t + \Delta t) = r_i(t) + v_i(t + \Delta t/2)\,\Delta t$

The method derives its name from the fact that the velocity and position information successively alternate at 1/2 time step intervals.

The steps to take in solving the equations of motions using the leapfrog algorithm are summarized in the list below. It is assumed that the initial positions, **r**(t), and velocities, **v**(t−Δt/2), are known.

1. Calculate **F**(t).
2. Calculate new velocities **v**(t+Δt/2) using **v**(t−Δt/2) and **F**(t).
3. Calculate new positions **r**(t−Δt/2) using **r**(t) and **v**(t−Δt/2).
4. Repeat these steps until a certain stop criterium.

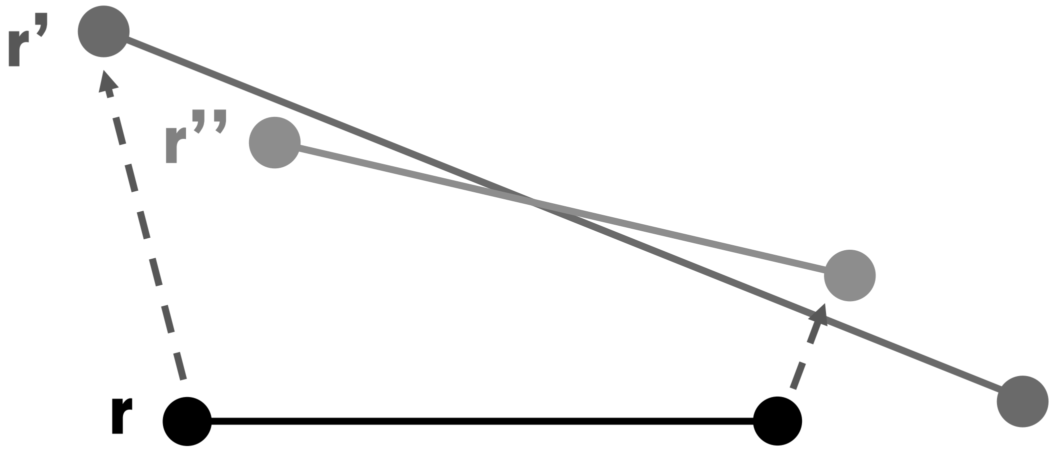
Molecular Dynamics Simulations

- **Constrains**
  - High frequency motions misbehave in classical simulation because they are of quantum nature
  - Most of the times, high frequency motions are not essential to the properties that we want to study
  - Bond constraints and, in some cases, angle constrains should be used
  - Removing the bond vibrations with constraints allows one to increase the time step from 0.5 or 1 fs to 2 or 4 fs

Leach, A.R., Molecular modelling: principles and applications. 2nd ed. 2001: Prentice Hall.

Molecular Dynamics Simulations

- **Constraint algorithms**
  - SHAKE
    - The SHAKE algorithm changes a set of unconstrained coordinates **r**’ to a set of coordinates **r**” that fulfil a list of distance constraints, using a set **r** as reference.
    - Iterative method



van der Spoel, D., et al., Gromacs user manual. Version 4.0. 2008.

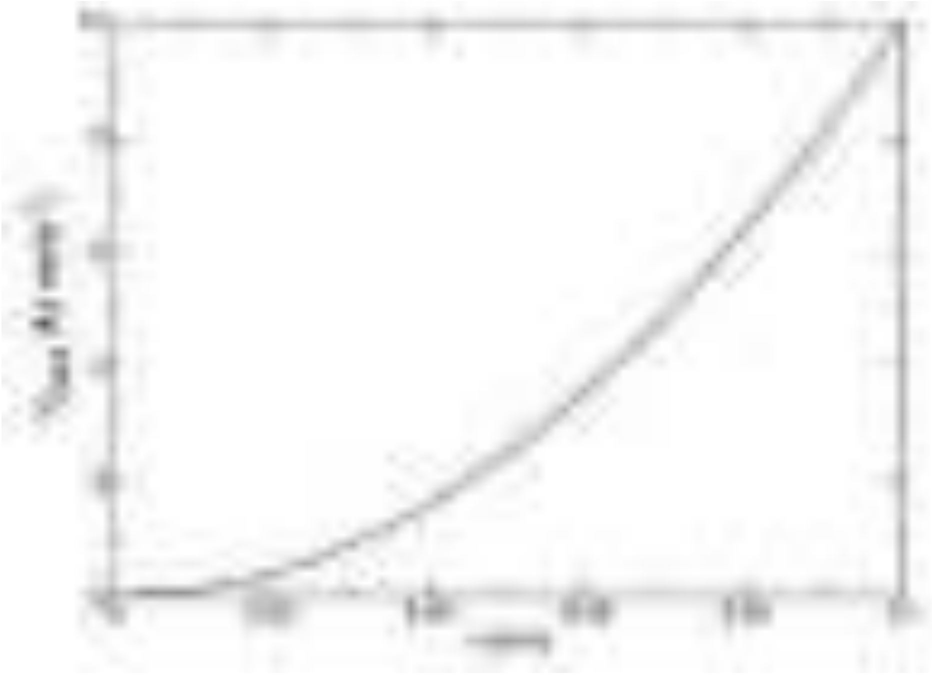
Molecular Dynamics Simulations

- **Constraint algorithms**
  - LINCS
    - Non-iterative method (faster than SHAKE)
    - it can only be used with bond constraints and isolated angle constraints
  - SETTLE
    - Analytical solution of SHAKE specifically for water molecules

van der Spoel, D., et al., Gromacs user manual. Version 4.0. 2008.

Molecular Dynamics Simulations

- **Restraints**
  - Sometimes are confused with constrains
  - Particles remain close to a reference position
  - Positions, distances, angles, dihedral
  - Avoid disastrous deviations during equilibrations
  - Include knowledge from experimental data (ex: NMR)



$$V_{pr}(r_i) = \frac{1}{2} k_{pr} \left| r_i - R_i \right|^2$$

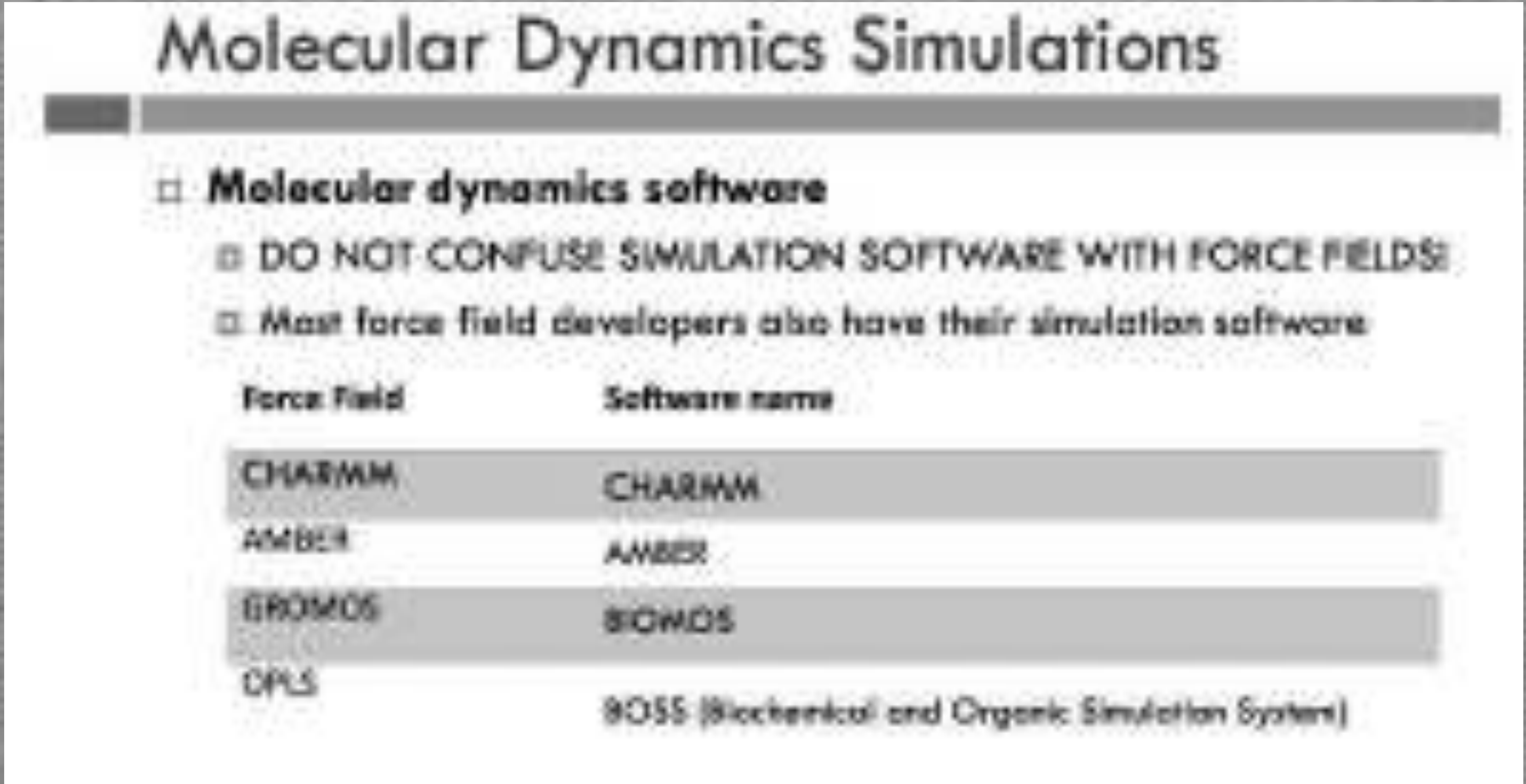
van der Spoel, D., et al., Gromacs user manual. Version 4.0. 2008.

Molecular Dynamics Simulations

- **Basic MD algorithm**

```
Input: Position of all atoms of the system
START
    1. Generated velocities by a Maxwell random distribution
    2. Compute forces on each particle
    3. Compute the scaling factors $\lambda$ and $\mu$
    4. Update and scale velocities (Temperature)
    5. Update new unconstrained coordinates (Move the particles)
    6. Apply constraint algorithm to coordinates (SHAKE)
    7. Correct velocities for constraints (due to the constraints used)
    8. Scale coordinates and box (Pressure)
    9. Write output
Repeat steps 2-9 N times
END
```

van der Spoel, D., et al., Gromacs user manual. Version 4.0. 2008.



Molecular Dynamics Simulations

- **Molecular dynamics software**
  - DO NOT CONFUSE SIMULATION SOFTWARE WITH FORCE FIELDS!
  - Most force field developers also have their simulation software

Force Field	Software name
CHARMM	CHARMM
AMBER	AMBER
GROMOS	BIOMOS
OPLS	BOSS (Biochemical and Organic Simulation System)

DL_POLY