Introduction to Molecular Dynamics

Alexey V. Lyulin

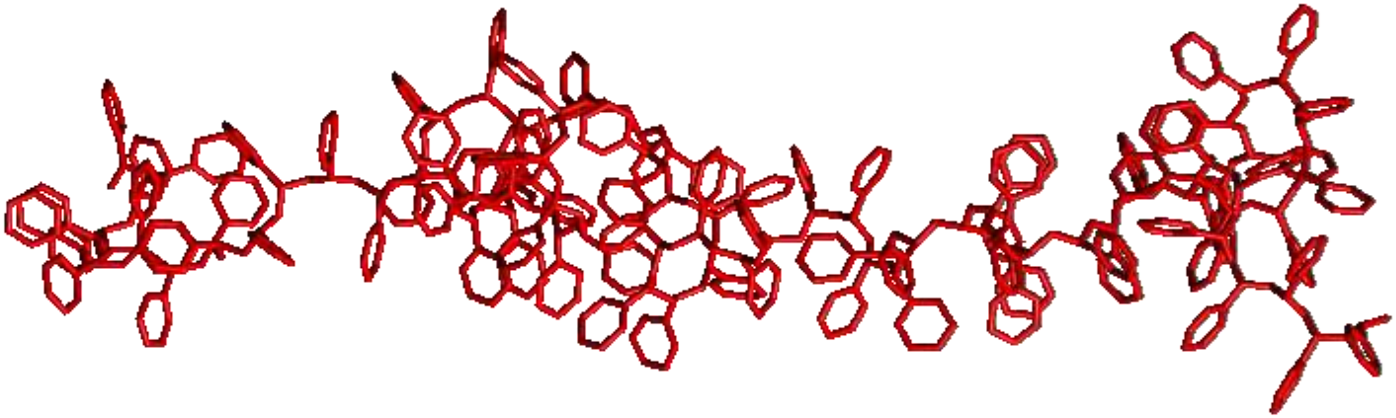
Soft Matter and Biological Physics,
Technische Universiteit Eindhoven, The Netherlands
a.v.lyulin@tue.nl

Han-sur-Lesse winter(??)school, June 13-14 2022

Outline

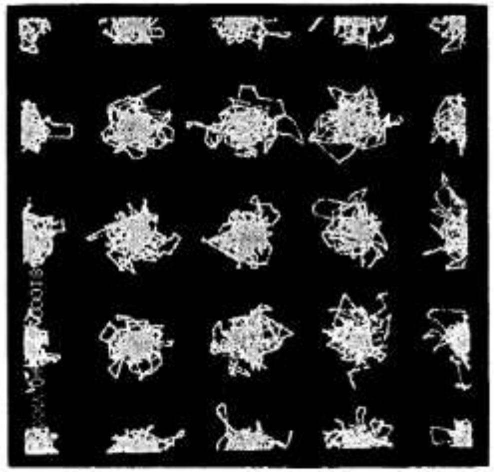
- history
- methods and algorithms
- program organization, program units
- force fields
- examples
- some tricks of the trade

Molecular Dynamics



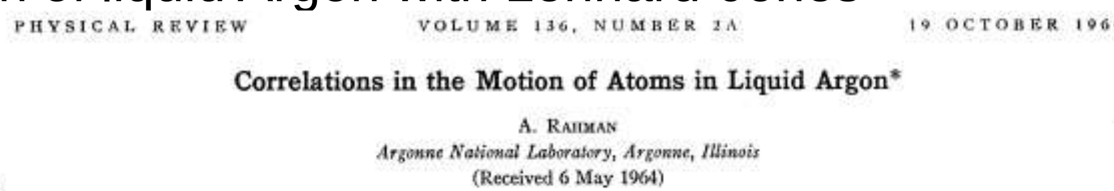
Historical Background

1957: Alder en Wainwright
Molecular dynamics simulation of the 3d system of hard spheres



Solid phase, 32 particles

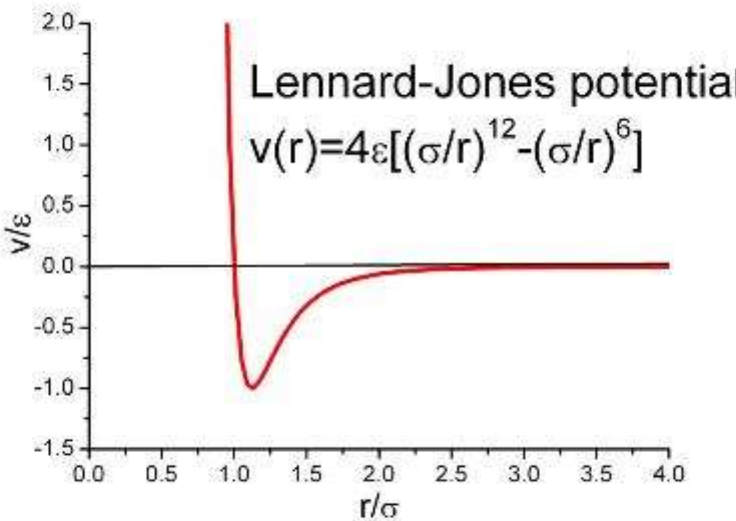
1964: Rahman
Molecular dynamics simulation of liquid Argon with Lennard-Jones potential



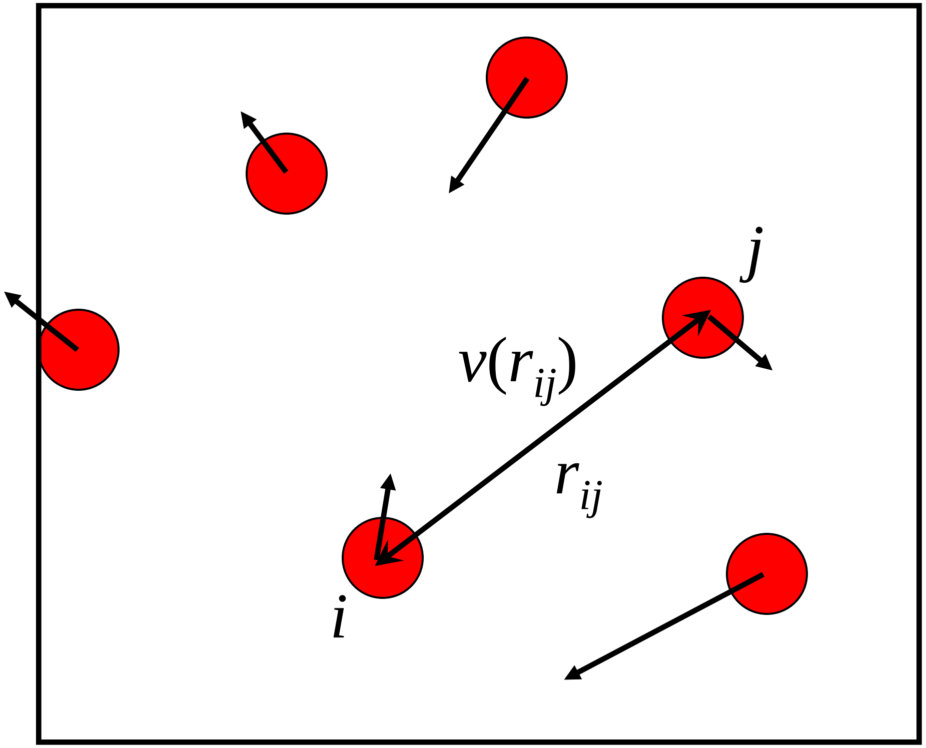
PHYSICAL REVIEW VOLUME 136, NUMBER 2A 19 OCTOBER 1964

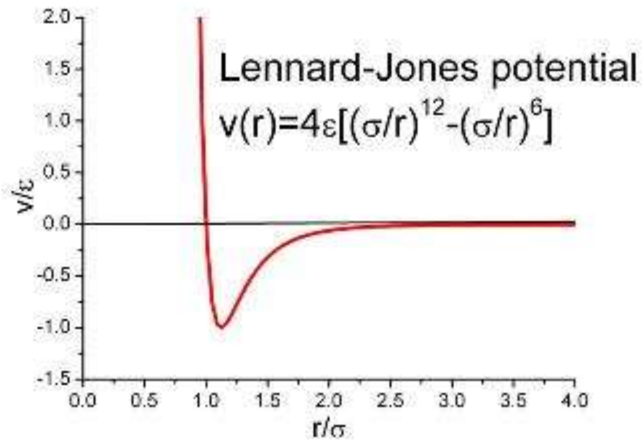
Correlations in the Motion of Atoms in Liquid Argon*

A. RAHMAN
Argonne National Laboratory, Argonne, Illinois
(Received 6 May 1964)



Molecular Dynamics





Very simple in principle: solution of classical Newton equations!

$$m_i \ddot{\vec{\mathbf{r}}}_i = m_i \vec{\mathbf{a}}_i = \vec{\mathbf{F}}_i = \sum_{j \neq i}^{N} \vec{\mathbf{F}}_{ij}$$

Problems, N particles

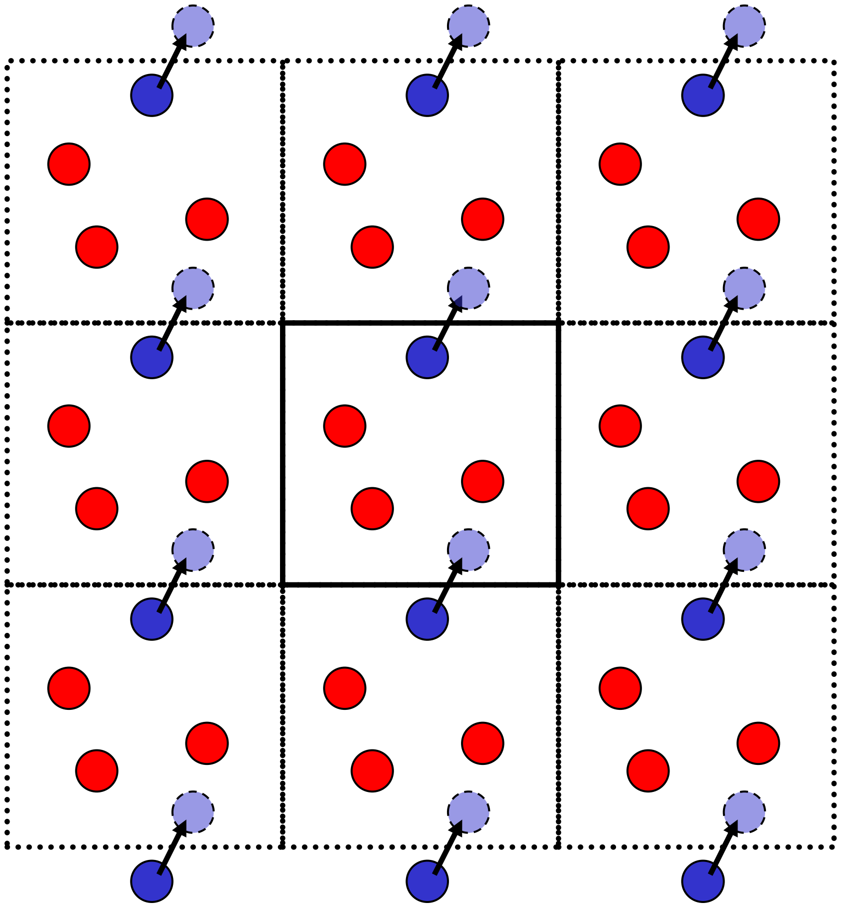
1) Interactions with the walls
   - difficult to describe
   - influence of walls for small systems

2) Infinite range of the interactions
   - problem for large systems: ???? interactions ???

$$N(N-1)/2$$

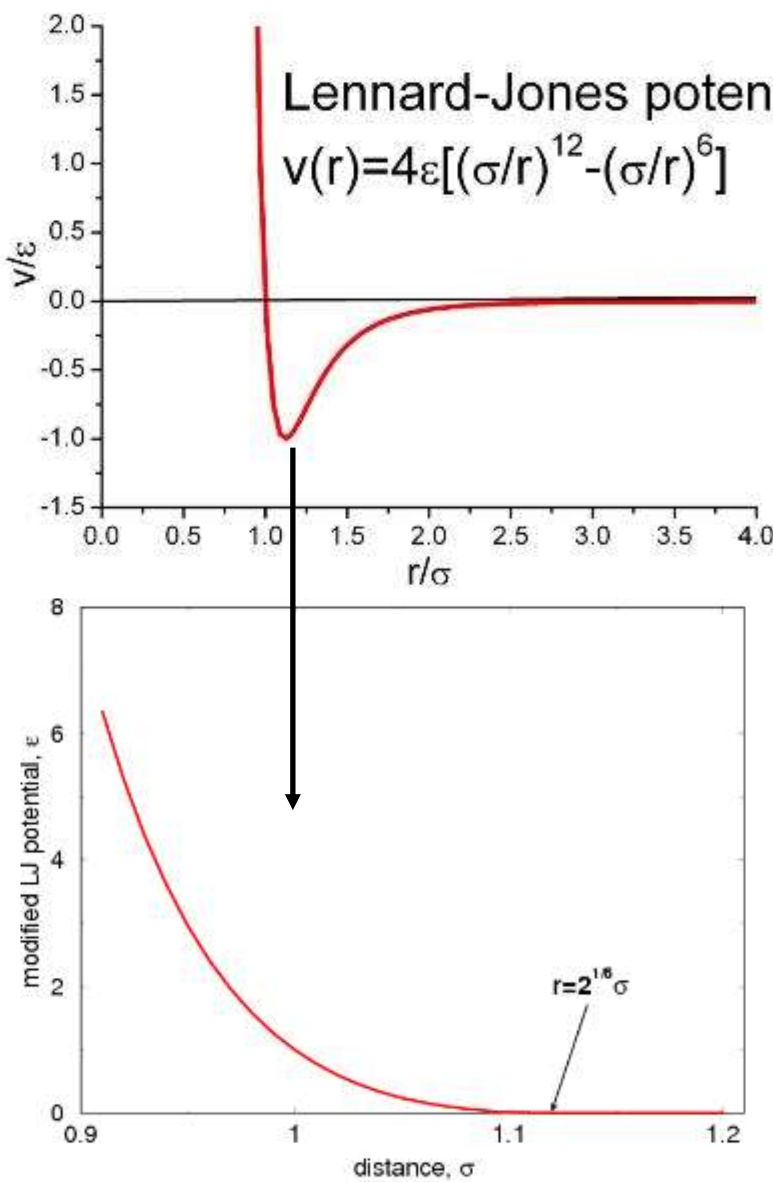
3) Time discretisation of Newton equations
   - conservation laws (energy, momenta)
   - naive discretisation does not work

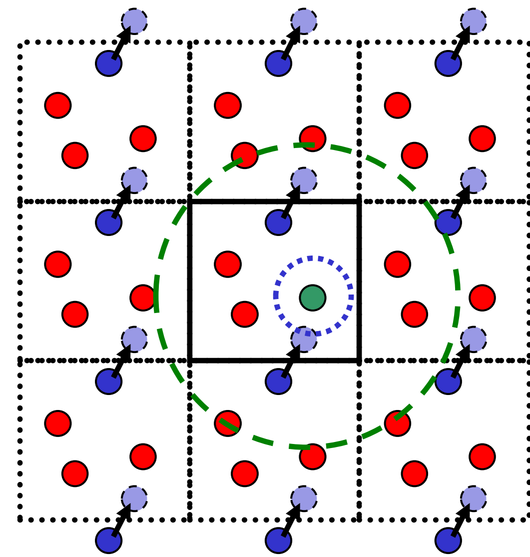
1) Periodic Boundary Conditions



$\Rightarrow$ no more walls, like in the infinite system

2) Potential cut-off





$\Rightarrow$ less CPU time

shifted Lennard-Jones potential

$$v(r_{ij}) = \begin{cases} 4\varepsilon\left(\left(\dfrac{\sigma}{r_{ij}}\right)^{12} - \left(\dfrac{\sigma}{r_{ij}}\right)^{6}\right) - v(r_c) & r_{ij} < r_c = \sqrt[6]{2}\sigma, \\ 0 & r_{ij} \geq r_c. \end{cases}$$

force $\qquad \vec{\mathbf{F}} = -\vec{\nabla} v(r)$

equation of motion $\qquad m_i \ddot{\vec{\mathbf{r}}}_i = m_i \vec{\mathbf{a}}_i = \vec{\mathbf{F}}_i = \sum_{j \neq i}^{N} \vec{\mathbf{F}}_{ij}$

Euler integration

Consider the general first-order o.d.e. $y' = f(x, y)$

We chose equally spaced grid-points $x_n = x_0 + nh$

Approximation of the solution $y_{n+1} = y_n + f(x_n, y_n)h$

The local error $y_{n+1} = y_n + f(x_n, y_n)h + O(h^2)$

to integrate over an interval of order unity requires $O(h^{-1})$ steps

The global error $O(h^{-1}) \cdot O(h^2) = O(h)$

First-order method !

3) Verlet method

$$\cancel{\frac{d\vec{\mathbf{v}}}{dt} = \vec{\mathbf{a}} \Rightarrow \vec{\mathbf{v}}(t+\delta t) = \vec{\mathbf{v}}(t) + \delta t \vec{\mathbf{a}}(t)} \quad \ldots + O(\delta t^2)$$

$$\cancel{\frac{d\vec{\mathbf{r}}}{dt} = \vec{\mathbf{v}} \Rightarrow \vec{\mathbf{r}}(t+\delta t) = \vec{\mathbf{r}}(t) + \delta t \vec{\mathbf{v}}(t)} \quad \ldots + O(\delta t^2)$$

$\Rightarrow$ very bad method!

How to reduce the level of errors introduced into the integration?

$$\vec{x}(t+\Delta t) = \vec{x}(t) + \vec{v}(t)\Delta t + \frac{\vec{a}(t)\Delta t^2}{2} + \frac{\vec{b}(t)\Delta t^3}{6} + O(\Delta t^4)$$

$+$

$$\vec{x}(t-\Delta t) = \vec{x}(t) - \vec{v}(t)\Delta t + \frac{\vec{a}(t)\Delta t^2}{2} - \frac{\vec{b}(t)\Delta t^3}{6} + O(\Delta t^4).$$

$$\vec{x}(t+\Delta t) = 2\vec{x}(t) - \vec{x}(t-\Delta t) + \vec{a}(t)\Delta t^2 + O(\Delta t^4).$$

$$\vec{v}(t) = \frac{\vec{x}(t+\Delta t) - \vec{x}(t-\Delta t)}{2\Delta t} + O(\Delta t^2).$$

velocity is step behind!

velocity Verlet method

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

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

1. Firstly, the new positions are calculated

2. Velocities at the mid-step are computed $\vec{v}(t+\frac{1}{2}\delta t)=\vec{v}(t)+\frac{1}{2}\delta t\vec{a}(t)$

3. The new forces and new accelerations are computed

4. The velocity move is completed $\vec{v}(t+\delta t)=\vec{v}(t+\frac{1}{2}\delta t)+\frac{1}{2}\delta t\vec{a}(t+\delta t)$

Errors

$$\text{error}\left(x(t_0+\Delta t)\right) = O(\Delta t^4)$$

$$x(t_0+2\Delta t) = 2x(t_0+\Delta t) - x(t_0) + \Delta t^2 x''(t_0+\Delta t) + O(\Delta t^4)$$

$$\text{error}\left(x(t_0+2\Delta t)\right) = 2\text{error}\left(x(t_0+\Delta t)\right) + O(\Delta t^4) = 3\,O(\Delta t^4)$$

$$\text{error}\left(x(t_0+3\Delta t)\right) = 6\,O(\Delta t^4)$$

$$\text{error}\left(x(t_0+4\Delta t)\right) = 10\,O(\Delta t^4)$$

$$\text{error}\left(x(t_0+5\Delta t)\right) = 15\,O(\Delta t^4)$$

n=2	3	
n=3	6	
n=4	10	
n=5	15	
n=6	??	21
n	??	n(n+1)/2

$$\text{error}\left(x(t_0+n\Delta t)\right) = \frac{n(n+1)}{2}\,O(\Delta t^4)$$

$$T = n\Delta t \rightarrow n = T/\Delta t$$

$$\text{error}\left(x(t_0+T)\right) = \left(\frac{T^2}{2\Delta t^2} + \frac{T}{2\Delta t}\right) O(\Delta t^4)$$

$$\text{error}\left(x(t_0+T)\right) = O(\Delta t^2)$$

Phase Space

$\vec{\mathbf{r}}_1, \cdots, \vec{\mathbf{r}}_N ; \vec{\mathbf{p}}_1, \cdots, \vec{\mathbf{p}}_N \rightarrow \vec{\mathbf{\Gamma}}$ phase space, dimension: $6N$

$A(\vec{\mathbf{\Gamma}})$ physical property

$H(\vec{\mathbf{\Gamma}}) = T(\vec{\mathbf{p}}_\mathbf{1}, \ldots, \vec{\mathbf{p}}_\mathbf{N}) + U(\vec{\mathbf{r}}_\mathbf{1}, \ldots, \vec{\mathbf{r}}_\mathbf{N})$ Hamiltonian = energy

$$E_K = \frac{1}{N} \sum_{i=1}^{n} \frac{p_i^2}{2m_i} \overset{\text{equipartition}}{=} D \frac{1}{2} k_B T$$

kinetic energy per particle

$$\Rightarrow T(\vec{\mathbf{\Gamma}}) = \frac{1}{N D k_B} \sum_i \frac{p_i^2}{m_i}$$

temperature

$$P(\vec{\mathbf{\Gamma}}) = \frac{1}{V} \left(N k_B T + \frac{1}{D} \sum_{i=1}^{N} \vec{\mathbf{r}} \cdot \vec{\mathbf{F}}_i^{\text{int}} \right)$$

pressure

$$A_{\text{obs}} = \langle A \rangle_{\text{time}} = \left\langle A(\vec{\mathbf{\Gamma}}(t)) \right\rangle_{\text{time}} = \lim_{t_{\text{obs}} \to \infty} \frac{1}{t_{\text{obs}}} \int_0^{t_{\text{obs}}} A(\vec{\mathbf{\Gamma}}(t)) dt$$

macroscopic observable

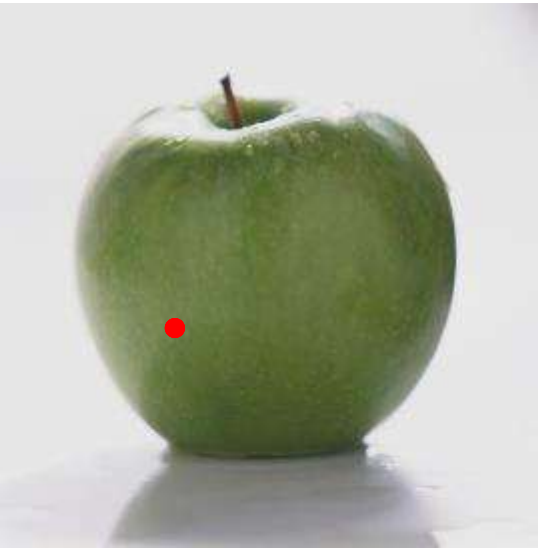
Ensembles

$$\rho_{NVE}(\vec{\boldsymbol{\Gamma}}) \propto \delta(H(\vec{\boldsymbol{\Gamma}}) - E)$$

microcanonical ensemble

$$\Omega = \int_{\vec{\boldsymbol{\Gamma}}} d\vec{\boldsymbol{\Gamma}}\, \delta(H(\vec{\boldsymbol{\Gamma}}) - E)$$

partition function



$$\langle A \rangle_{\text{ensemble}} = \frac{1}{\Omega} \int_{\vec{\boldsymbol{\Gamma}}} d\vec{\boldsymbol{\Gamma}}\, \rho_{NVE}(\vec{\boldsymbol{\Gamma}}) A(\vec{\boldsymbol{\Gamma}})$$

ensemble-averaged

$$\langle A \rangle_{\text{ensemble}} = \langle A \rangle_{\text{time}}$$

"ergodicity"

$$\rho_{NVT}(\vec{\boldsymbol{\Gamma}}) \propto \exp(-H(\vec{\boldsymbol{\Gamma}}) / k_B T)$$

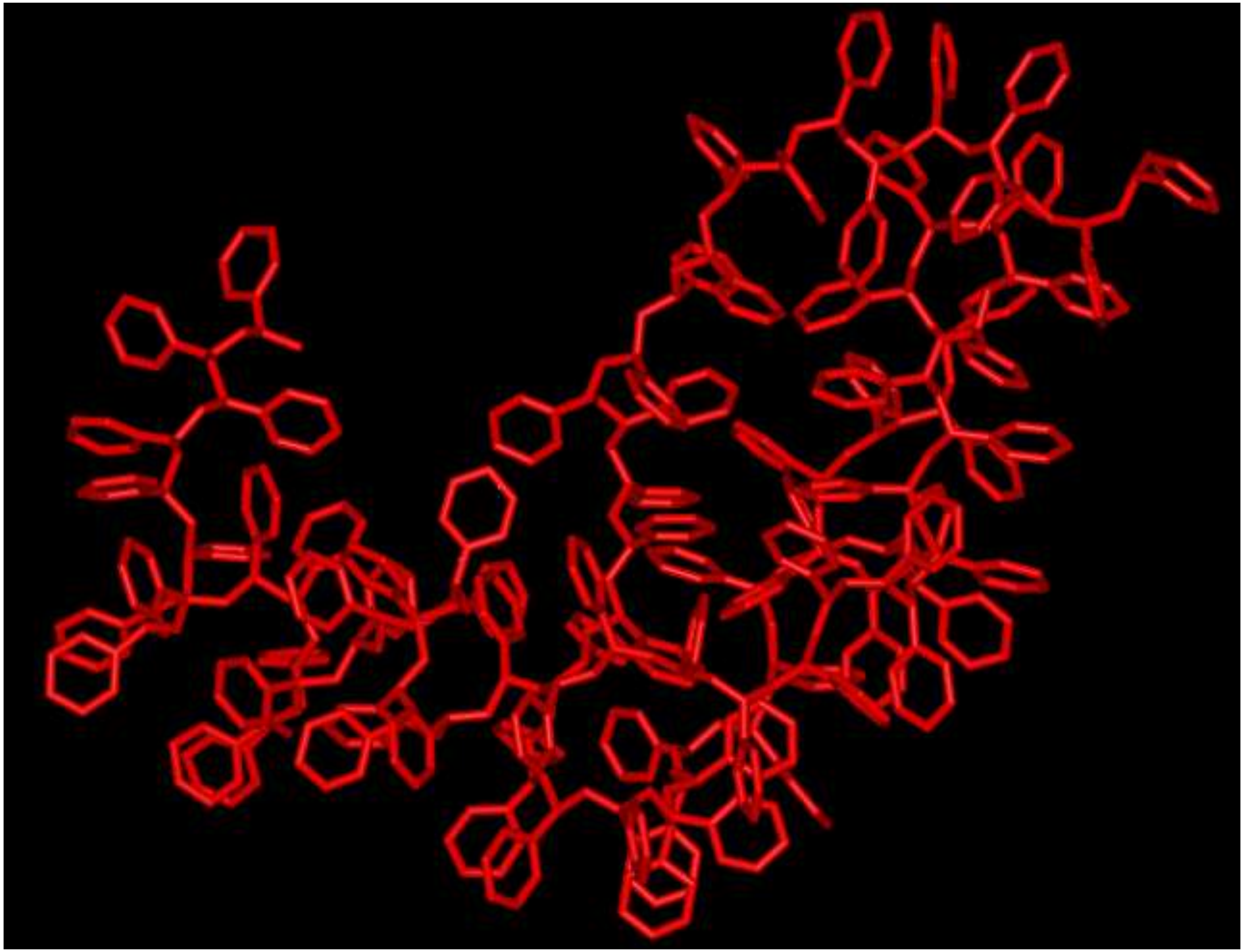
canonical ensemble

$$Q = \int_{\vec{\boldsymbol{\Gamma}}} d\vec{\boldsymbol{\Gamma}}\, \exp(-H(\vec{\boldsymbol{\Gamma}}) / k_{\mathrm{B}} T)$$

partition function

all statististical ensembles equiv. in thermodyn. limit

$$\langle A \rangle_{\text{ensemble, canonic}} = \frac{1}{Q} \int_{\vec{\boldsymbol{\Gamma}}} d\vec{\boldsymbol{\Gamma}}\, \rho_{NVT}(\vec{\boldsymbol{\Gamma}}) A(\vec{\boldsymbol{\Gamma}}) = \langle A \rangle_{\text{ensemble,microcanonic}}$$



N=const
V=const
T=const

Molecular Dynamics

For each atom in every molecule, we need:

- Position (r)
- Momentum (m + v)
- Forces (m + a)

- Thermostat

Typical program organization

• **INPUT**: parameters that specify the conditions of the run
(e.g., initial temperature, number of particles, density, time step).

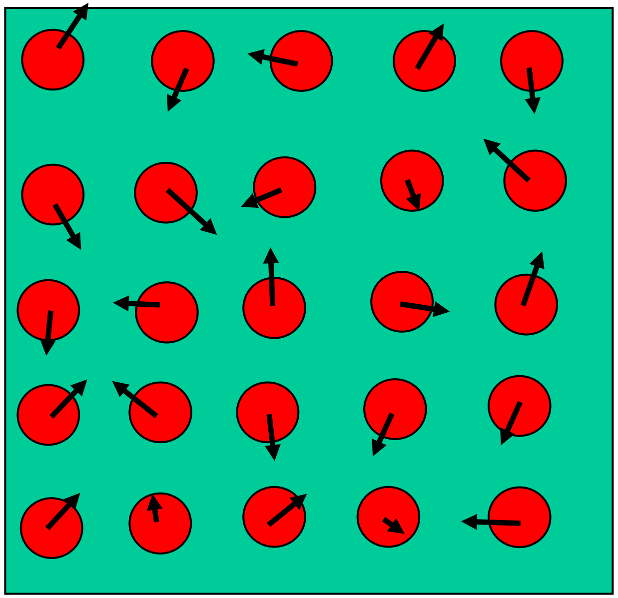
• **INITIALIZATION**: initial positions and velocities.

DO LOOP
• Compute forces on all particles.
• Integrate Newton’s equations of motion.
ENDDO

• **ANALYSIS**: Compute the averages of measured quantities.

Initialization

- place each particle on a cubic lattice;
- give random velocities according to some distribution.
- shift velocities so that the velocity of the center of mass is zero
- scale the resulting velocities to adjust the mean kinetic energy to the desired value.



Initialization

Compute velocity center of mass: $\bar{v}_\alpha = \dfrac{\sum_{j=1}^{N} m_j v_{\alpha,j}}{N \sum_{j=1}^{N} m_j} \quad , \alpha = x, y, z$

Compute kinetic energy: $E_k = \dfrac{1}{2N} \sum_{j=1}^{N} m_j v_j^{\,2}$

$$E_K = \frac{1}{N} \sum_{i=1}^{n} \frac{p_i^2}{2m_i} = D \frac{1}{2} k_B T$$

$$\Rightarrow T = \frac{1}{NDk_B} \sum_i \frac{p_i^2}{m_i}$$

Set velocity center of mass to zero: $v_{\alpha,j} = v_{\alpha,j} - \bar{v}_\alpha \quad , \alpha = x, y, z$

Compute Forces

$$\text{ecut} = 4\varepsilon\left[\left(\frac{\sigma}{r_c}\right)^{12} - \left(\frac{\sigma}{r_c}\right)^{6}\right] \quad - \text{potential at cutoff } r = r_c$$

for each pair of particles *i,j*:

$$\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$$

$$\vec{r}_{ij,\alpha} = \vec{r}_{ij,\alpha} - \text{box round}\left(\frac{r_{ij,\alpha}}{L_{box}}\right)$$

$$\text{if } r^2_{ij} < r^2_c$$

- calculate forces using Lennard-Jones potential
- update forces on molecules *i,j*
- update potential energy, $U_{new} = U_{old} + 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] - \text{ecut}$

Physical units in simulations

- σ length
- m mass
- ε energy
- $\sigma (m/\varepsilon)^{1/2}$ time
- ε/k_B temperature

Initial data for simulations

- lattice size (boxsize)
- number of particles (Npart < 2000)
- integration time interval = tmax-tmin
- time step size (dt) = 0.005

How Long? How Large?

Simulation runs are typically short: T_{run} ~10^3-10^6 MD steps, corresponding to perhaps a few nanoseconds of real time

Consider variable a, $<a> = 0$;

time correlation function $<a(t_0)a(t_0 + t)>$;

assuming that the system is in equilibrium, this function is independent of the choice of time origin and may be written as $<a(0)a(t)>$.

It will decay from an initial value $<a(0)a(0)>$ to a long-time limiting value at t >> $\tau_{\text{relaxation}}$ which one??

$$\lim_{t\to\infty} < a(0)a(t) > = < a(0) > < a(t) > = 0$$

$$\mathbf{T_{run} >> \tau_{relaxation}}$$

Similarly, define a spatial correlation function $\langle a(0)a(r)\rangle$ relating values computed at different points *r* apart.

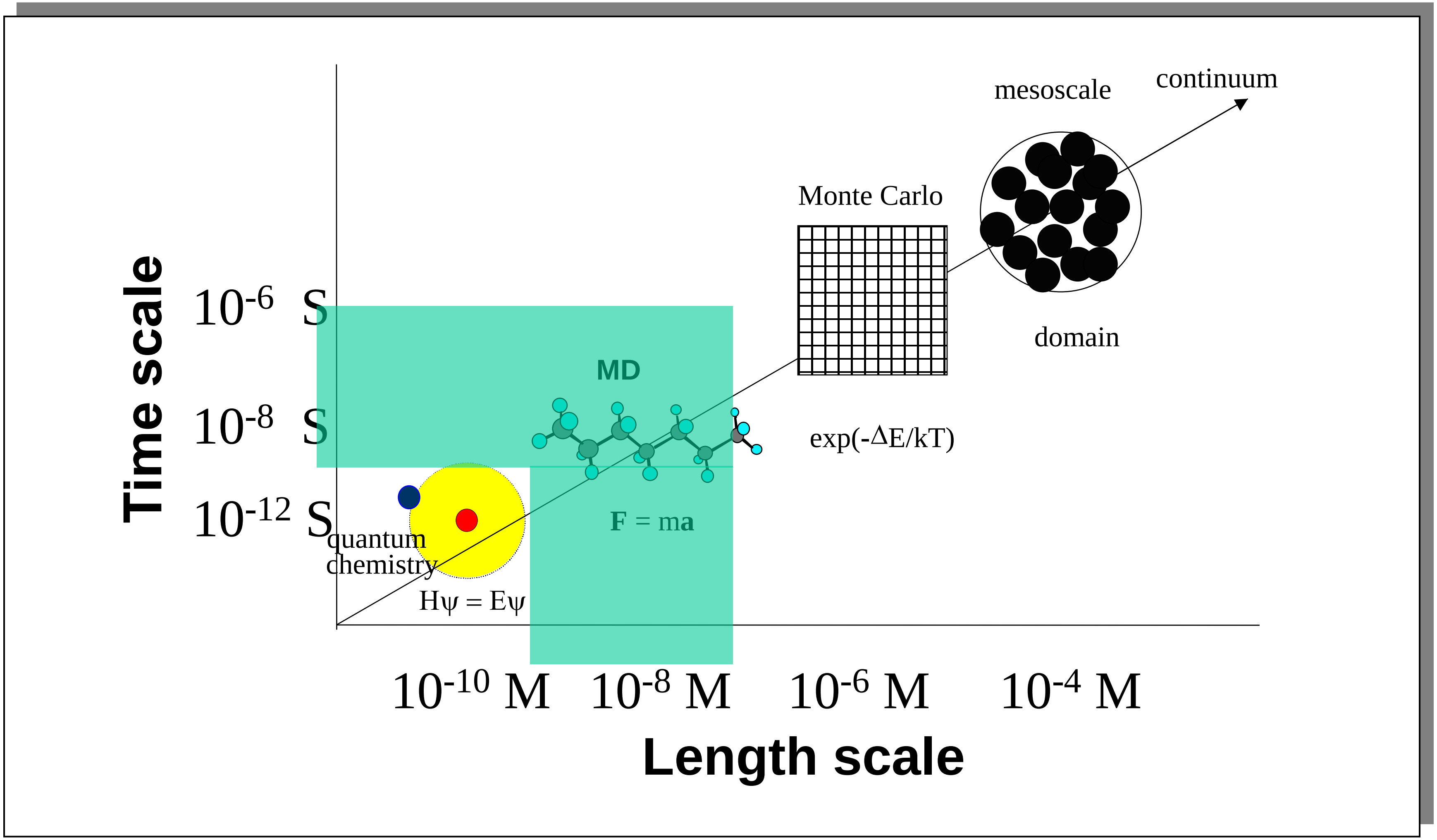
This function decays from a short-range nonzero value to zero over a characteristic distance ξ_a, the correlation length.

It is essential for simulation box sizes L to be large compared with ξ_a. Only then can we guarantee that reliably-sampled statistical properties are obtained.

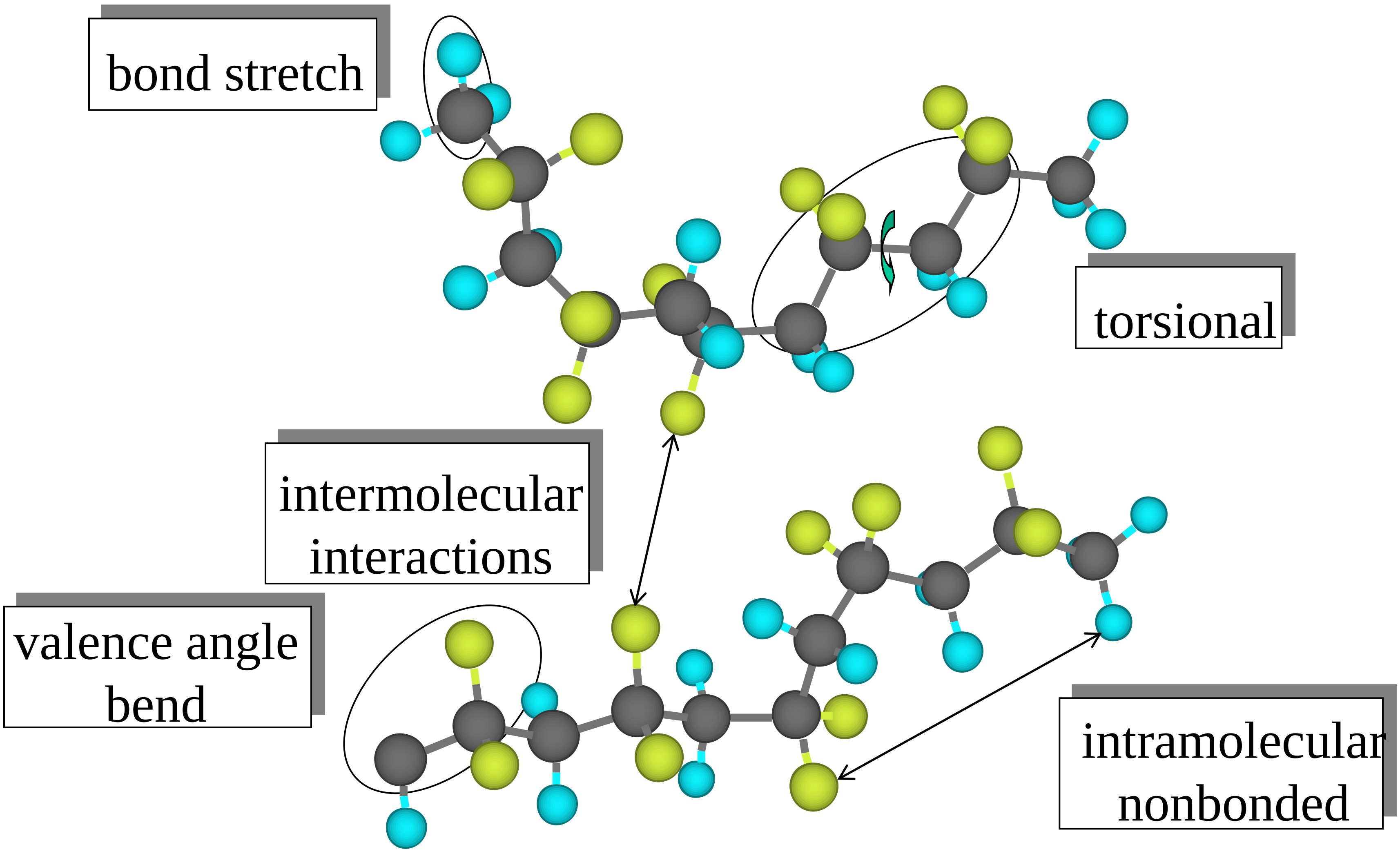
$$\mathbf{L_{box} \gg \xi_a}$$

$$\mathbf{r_{cutoff} < L_{box}/2}$$

What and Where: Scales in Simulations



The Force Field



- Bonded neighbours
- Non-bonded atoms (either other atoms in the same molecule, or atoms from different molecules)

$$V(R) = E_{bonded} + E_{non-bonded}$$

Non-Bonded Atoms

- van der Waals Potential
- Electrostatic Potential

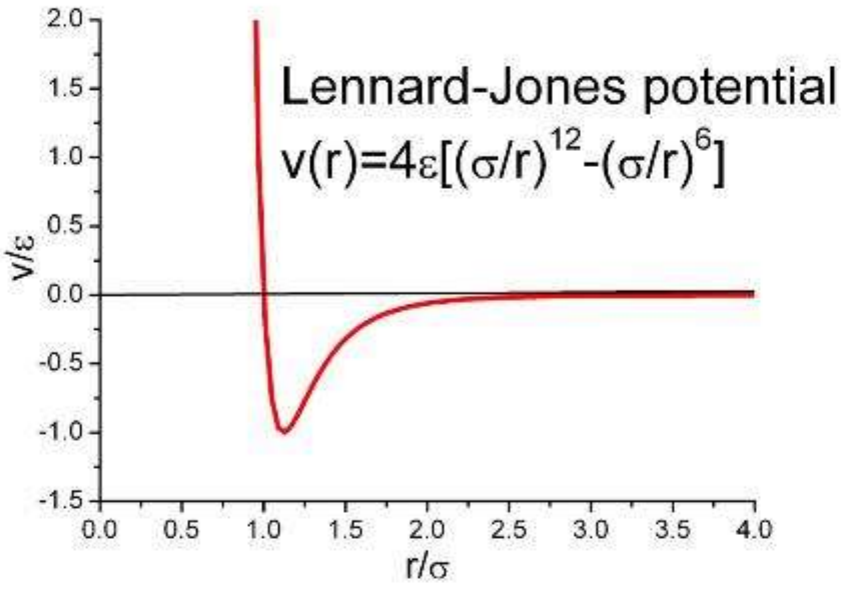
$$E_{non-bonded} = E_{vdW} + E_{Coulomb}$$

van der Waals Potential

- Neutral atoms attract each other at short distances;
- Once the atoms are close enough to have overlapping electron clouds, they will repel each other.

van der Waals Potential

$$E_{vdW} = \sum_{i \neq j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}} \right)$$



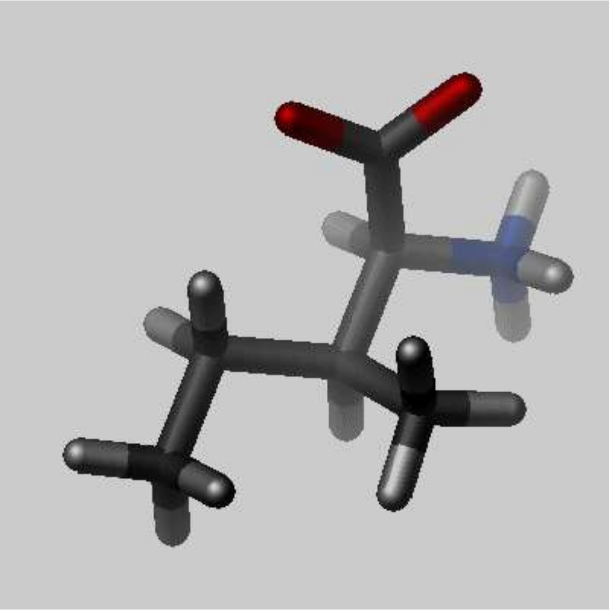
Electrostatic Potential

- Opposite charges attract;
- Like charges repel;
- The force of the attraction is inversely proportional to the square of the distance

$$E_{Coulomb} = \sum_{i \neq j} \frac{q_i q_j}{D r_{ij}}$$

Bonded Atoms

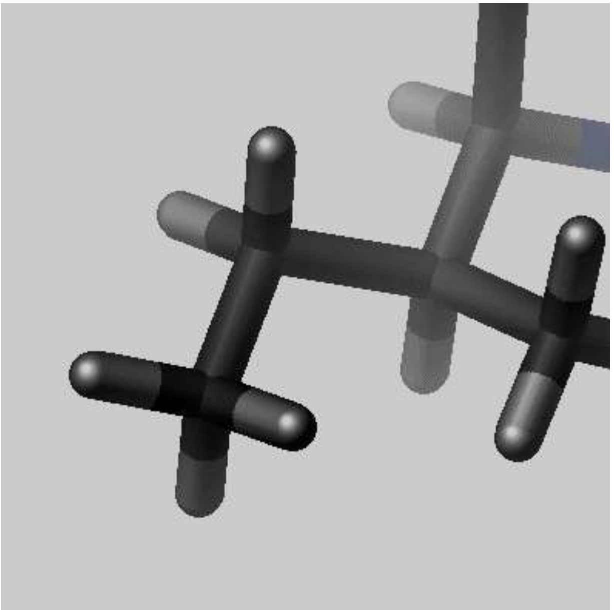
- Stretching the bond
- Bending the angle between bonds
- Rotating around bonds



$$E_{bonded} = E_{bond-stretch} + E_{angle-bend} + E_{rotation}$$

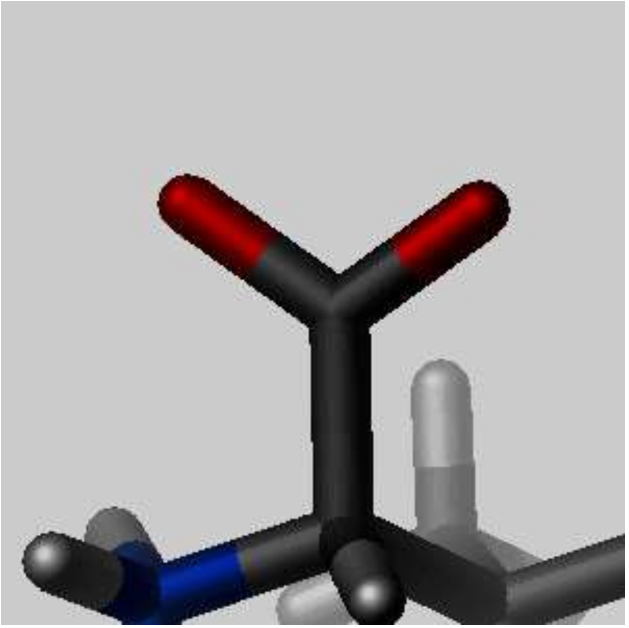
Bond Length (stretching) Potential

spring constant and the equilibrium bond length are dependent on the atoms involved.



$$E_{bond-stretch} = \sum_{1,2\,pairs} \frac{1}{2} K_b (b - b_0)^2$$

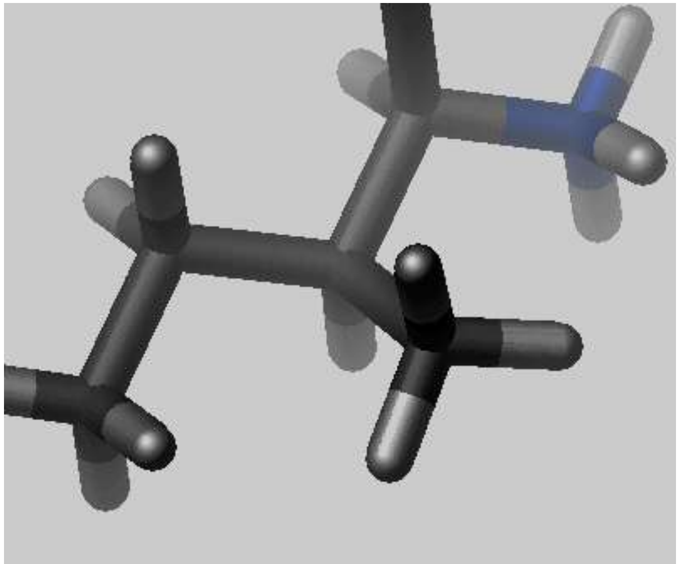
Bond Angle (bending) Potential



$$E_{angle-bend} = \sum_{angles} \frac{1}{2} K_\theta (\theta - \theta_0)^2$$

Torsional Potential

Described by a dihedral angle and coefficient of symmetry ($n=1,2,3$), around the middle bond.



$$E_{rotate-along-bond} = \sum_{1,4\,pairs} \frac{1}{2} K_{\varphi} (1 - \cos(n\varphi))$$

F^{pot}, The Force Field

$$\mathbf{F}^{pot} = -\nabla V(\mathbf{r})$$

$V(r) = V_{stretch} + V_{bend} + V_{tor} + V_{LJ} + V_{Coulomb}$

$$V_{stretch} = \frac{1}{2} k_{ij}(r_{ij} - r^0_{ij})^2$$

$$V_{bend} = \frac{1}{2} k_{ijk}(\Theta_{ijk} - \Theta^0_{ijk})^2$$

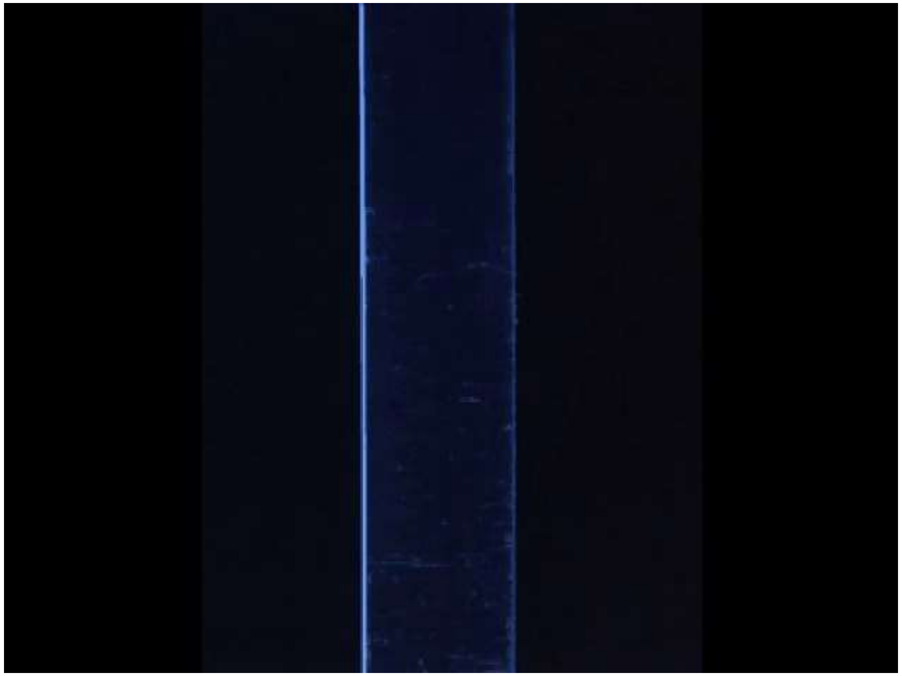
$$V_{tor} = \frac{1}{2} k_{ijkl}(1 - \cos n\phi_{ijkl})$$

$$V_{LJ} = A_{ij} r_{ij}^{-12} - C_{ij} r_{ij}^{-6}$$

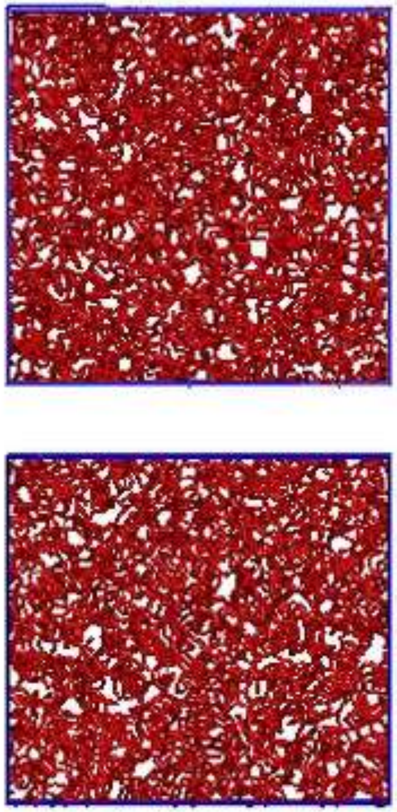
$$V_{Coulomb} = \frac{q_i q_j}{r_{ij}}$$

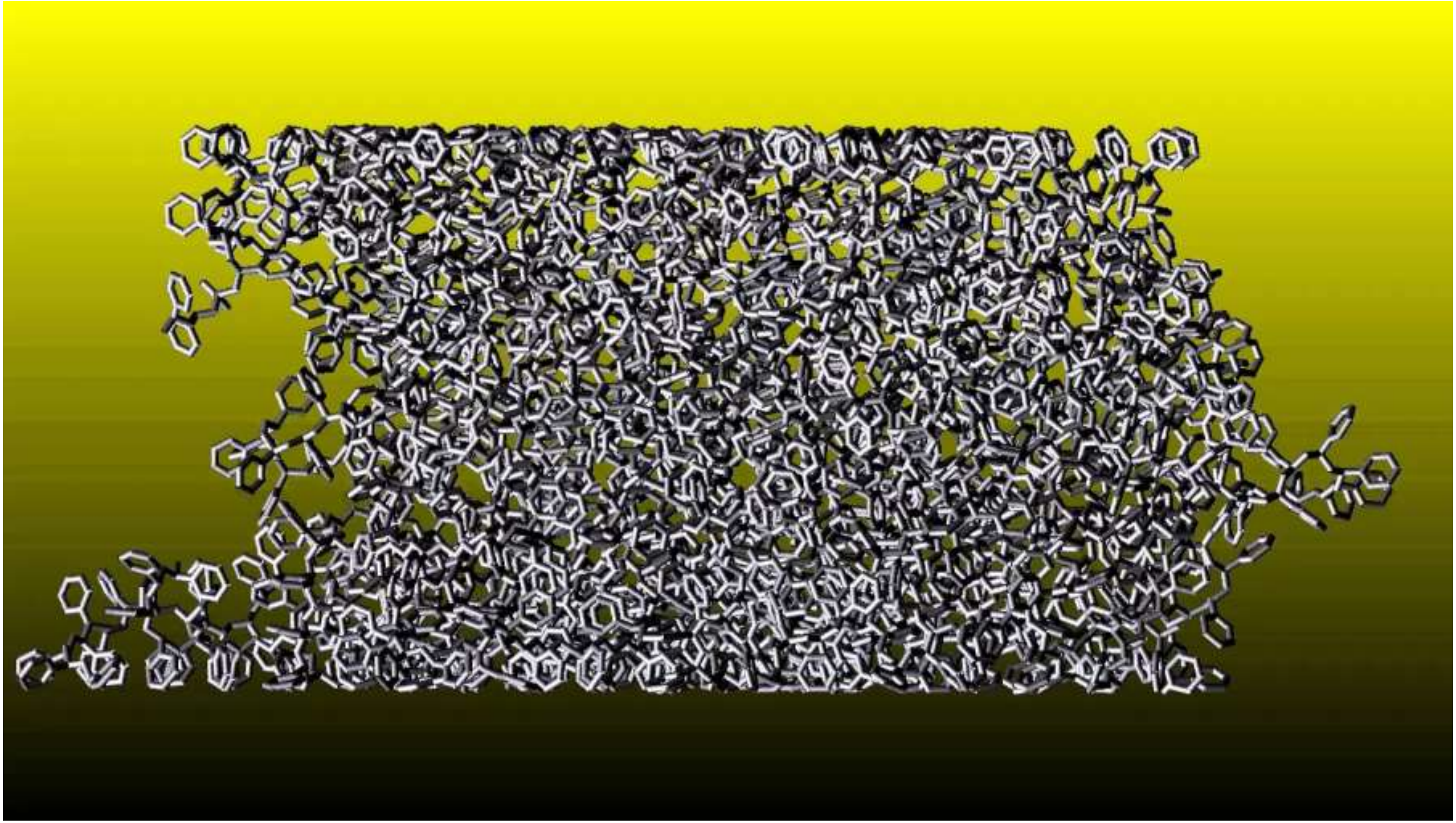
Mechanical rejuvenation

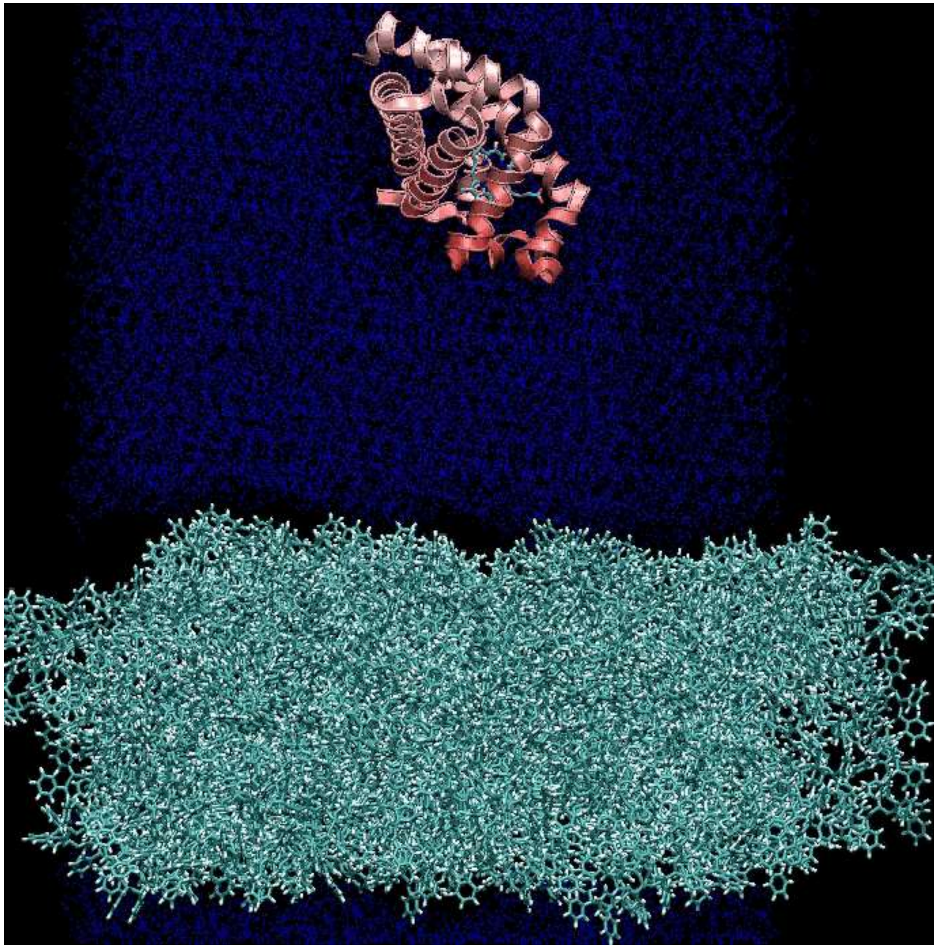
Polystyrene

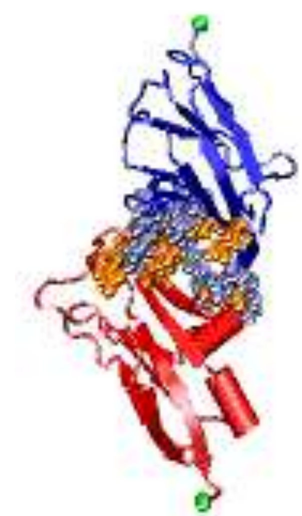


PS: $T << T_g$









Thermostat

(scaling of velocities method)

$$E_{kin}(t) = \sum_{i=1}^{N} \frac{m_i v_i^2(t)}{2}$$

$$\left\langle E_{kin} \right\rangle_t = \frac{3N}{2} k_B T(t) = \frac{3N}{2} k_B T_{fixed}$$

$$\vec{v}_i^{\,new} = \lambda \vec{v}_i$$

$$E_{kin}^{new} = \sum_{i=1}^{N} \frac{m_i \lambda^2 v_i^2}{2}, \quad \left\langle E_{kin}^{new} \right\rangle = \frac{3N}{2} k_B T_{fixed}$$

$$\lambda^2 \frac{3N}{2} k_B T(t) = \frac{3N}{2} k_B T_{fixed} \quad \Rightarrow \quad \lambda = \sqrt{\frac{T_{fixed}}{T(t)}}$$

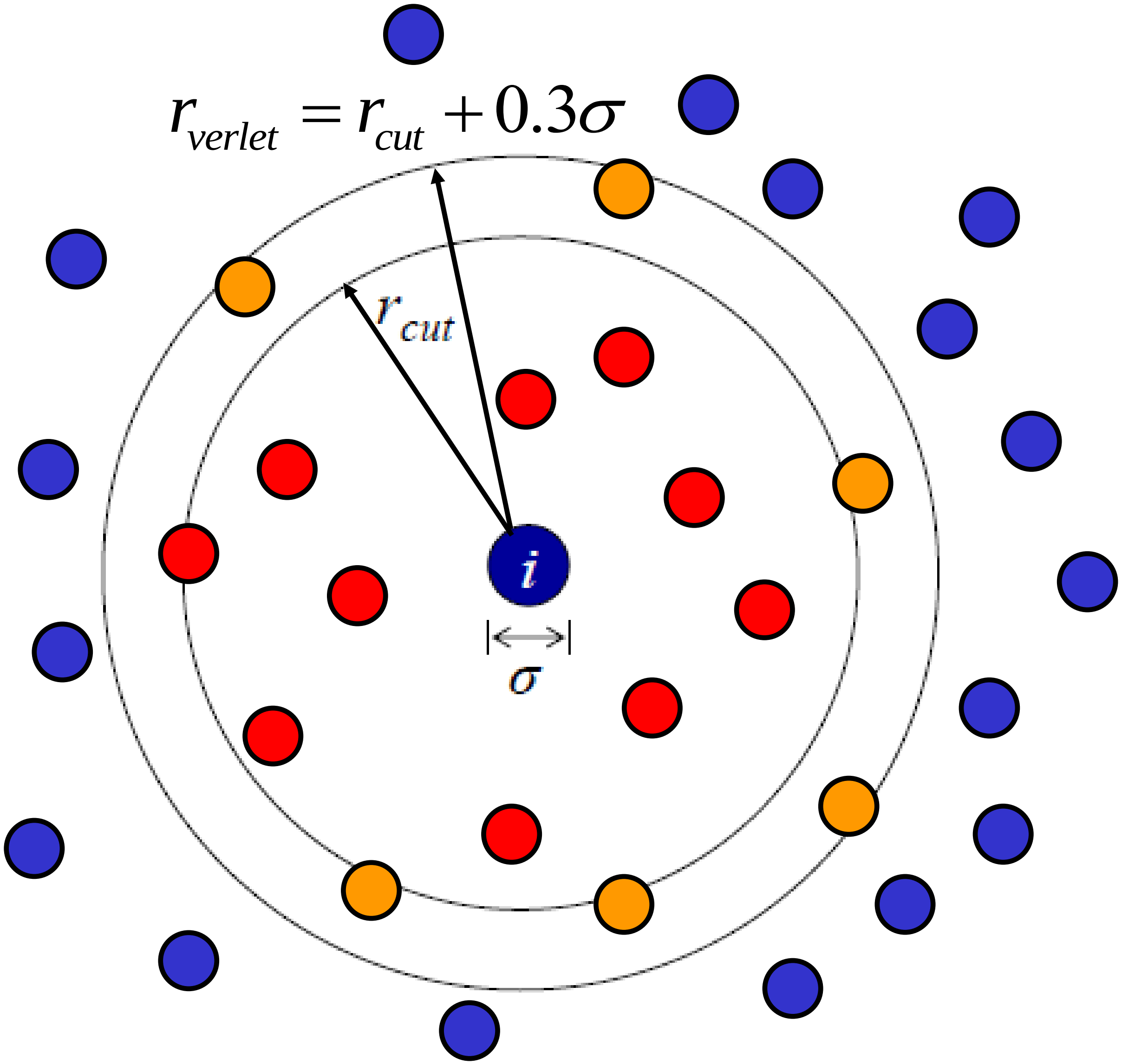
Verlet neighbour list

it is time-consuming!

```
do i=1,N
     do j=1,N
        if(i = = j) cycle
           dx  = x(j)-x(i)
           dy  = y(j)-y(i)
           dz  = z(j)-z(i)
           rsq = dx*dx + dy*dy + dz*dz
           r   = sqrt(rsq)
     enddo
enddo
```

Since atoms move within one time step only < 0.1 – 0.2 Å, the neighbours remain the same for many time steps

Verlet neighbour list



Verlet neighbour list

- Updating neighbour lists
- typical update frequency ~ 10-20 steps

- Correcting for error
- Some fraction of potential energy is always ignored – add correction

$$E_{correction} = 2\pi\rho N \int_{r_{verlet}}^{\infty} r^2 U(r)\,dr$$

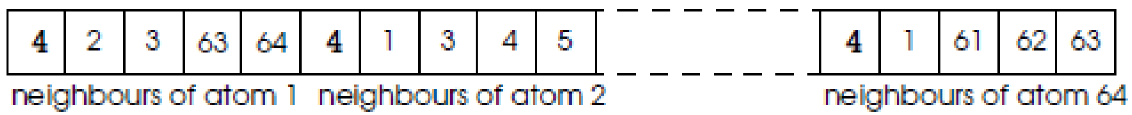
$$E_{correction,LJ} = 8\pi\rho N\varepsilon\left[\frac{\sigma^{12}}{9r_{verlet}^{9}} - \frac{\sigma^{6}}{3r_{verlet}^{3}}\right]$$

Implementation

- In practice the neighbour list can look e.g. like the following:

- Here $\mathtt{NNei}_1$ is the number of neighbours of atom $\mathtt{i}$
- $\mathtt{j}_1$, $\mathtt{j}_2$, ... are the indices of neighbouring atoms

- So, if we would have a 64 atom system, where every atom has 4 neighbours, the neighbour list could look like this
- :



- A practical implementation of creating the list:

```
startofineighbourlist=1
do i=1,N
   nneighboursi=0
   do j=1,N
      if (i==j) cycle
      dx=x(j)-x(i);
      dy=y(j)-y(i);
      dz=z(j)-z(i);
      rsq=dx*dx+dy*dy+dz*dz
      if (rsq <= rskincutsq) then
         nneighboursi=nneighboursi+1
         neighbourlist(startofineighbourlist+nneighboursi)=j
      endif
   enddo
   neighbourlist(startofineighbourlist)=nneighboursi ! Write in number of i's neighbours into list
   startofineighbourlist=startofineighbourlist+nneighboursi+1 ! Set starting position for next atom
enddo
```

Periodic boundaries omitted for brevity.

Autocorrelation functions

time correlation function: $C_{AB}(t) = \langle A(0)B(t) \rangle$

if A=B – autocorrelation function: $C_{AA}(t) = \langle A(0)A(t) \rangle$

$$\langle A(0)B(t) \rangle = \langle A(-t)B(0) \rangle$$

$$C_{AA}(0) = \langle A^2 \rangle$$

normalization: $\tilde{C}_{AA}(t) = \dfrac{\langle A(0)A(t) \rangle - \langle A \rangle^2}{\langle A^2 \rangle - \langle A \rangle^2}$ $\rightarrow$ $\tilde{C}_{AA}(0) = 1$

$\tilde{C}_{AA}(t = \infty) = 0$

Time correlation function may be evaluated as a time average, assuming the system is ergodic –
the phase space average is equal to a time average

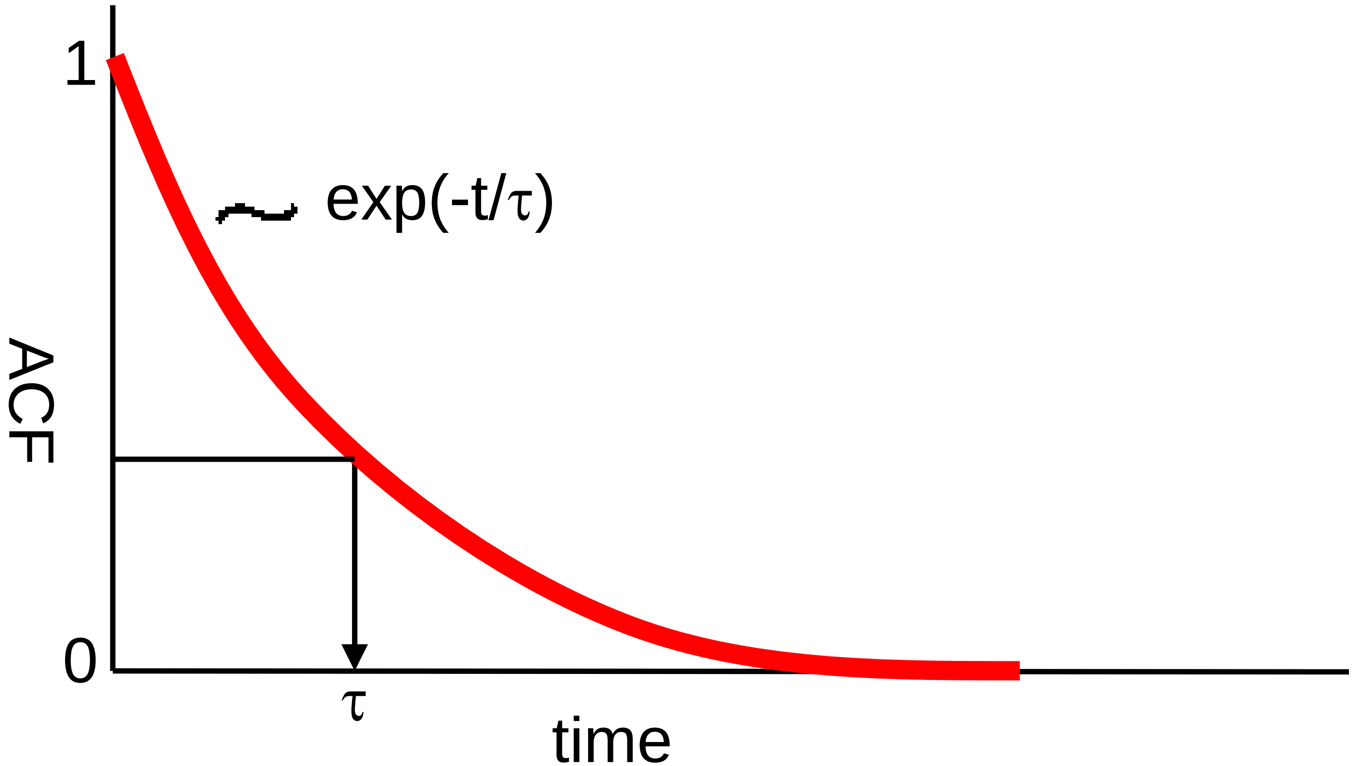
$$C_{AA}(t) = \lim_{T\to\infty} \frac{1}{T-t} \int_0^{T-t} d\tau A(\tau) A(t+\tau)$$

molecular dynamics simulations:

phase space trajectory is determined at discrete time steps, the integral is expressed as a sum

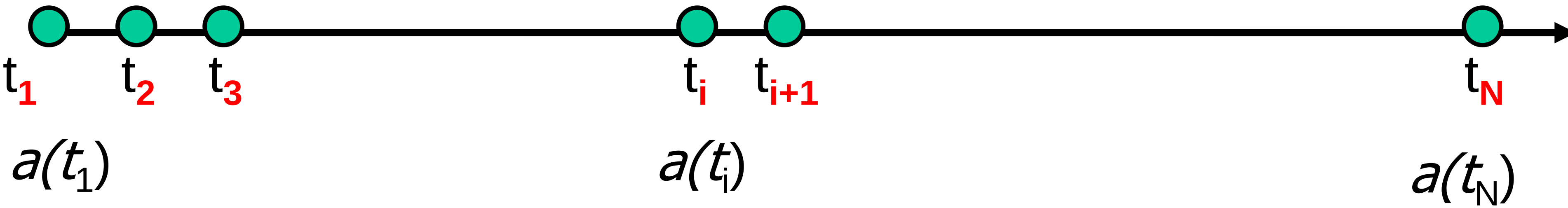
$$C_{AA}(k\Delta t) = \frac{1}{N-k} \sum_{j=1}^{N-k} A(t_k) A(t_{k+j}), \qquad k = 0,1,2...,N_c$$

N – total number of time steps
Δt – time step
$N_c << N$


1
~ exp(-t/τ)
ACF
0
τ
time

Practical realization

equidistant time intervals



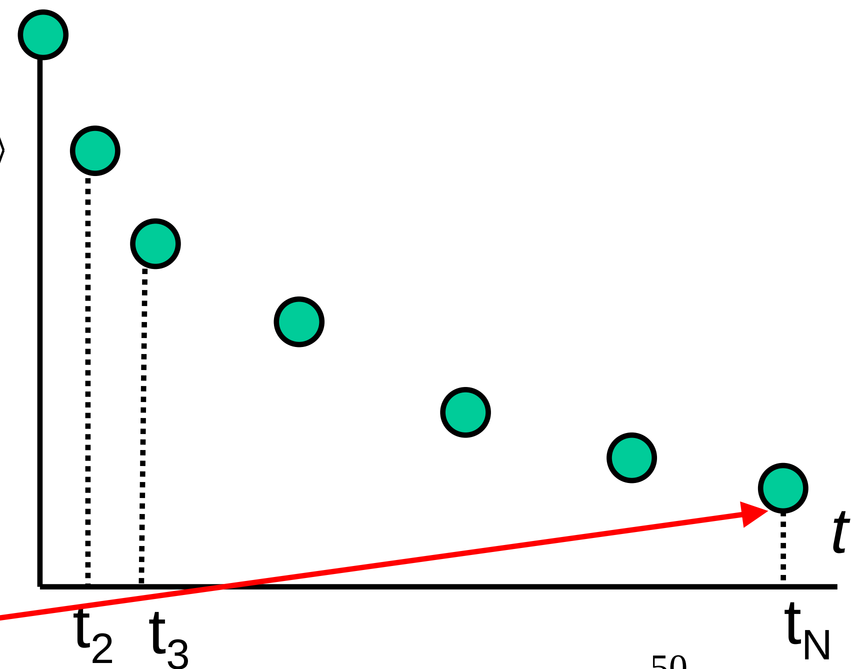
$$\frac{a(t_1)a(t_1)+a(t_2)a(t_2)+\ldots+a(t_N)a(t_N)}{N}=\left\langle a^2\right\rangle$$

$$\frac{a(t_1)a(t_2)+a(t_2)a(t_3)+\ldots+a(t_i)a(t_{i+1})+\ldots}{N-1}=\left\langle a(0)a(\Delta t)\right\rangle$$

$$\frac{a(t_1)a(t_3)+a(t_2)a(t_4)+\ldots+a(t_{N-2})a(t_N)}{N-2}=ACF(t_3)$$

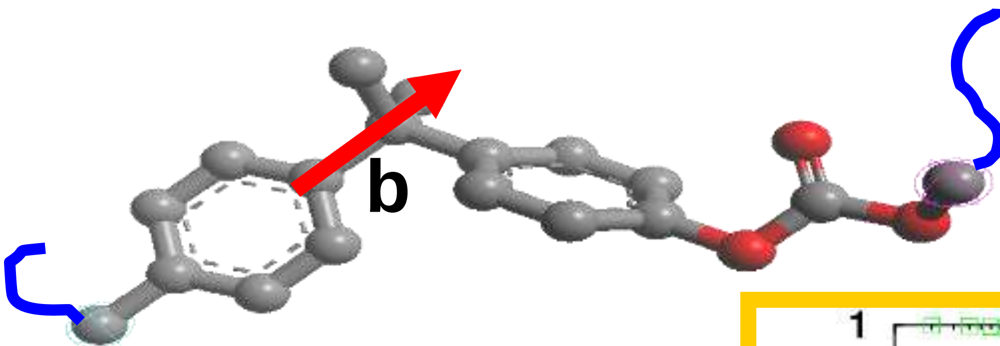
.....

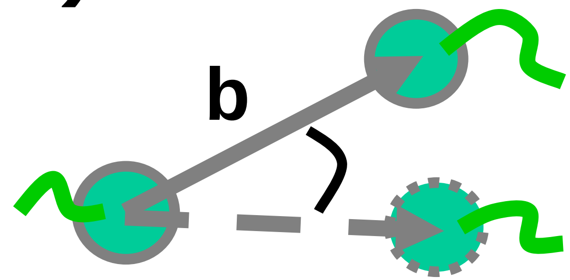
ACF=<a(0)a(t)>



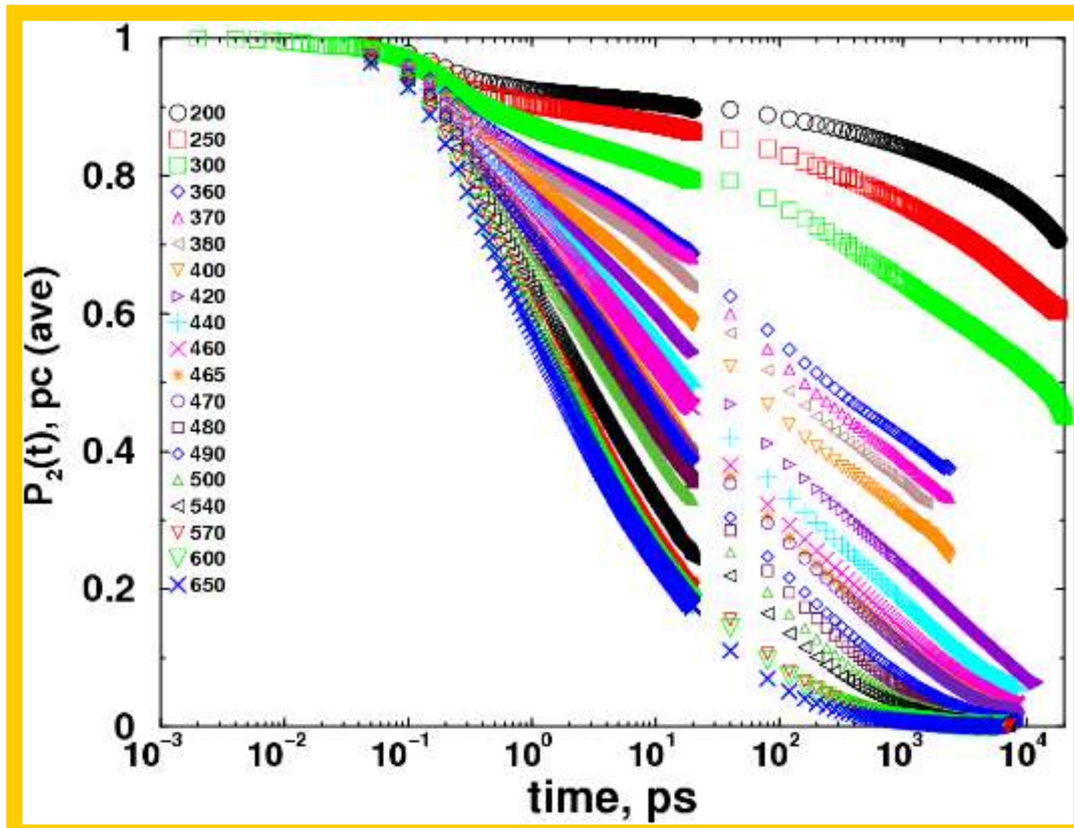
How many terms exist in order to calculate this point?

Example: Orientational mobility
(polycarbonate)





$$P_2(t) = < 3/2(\mathbf{b}(0)\mathbf{b}(t))^2 - 1/2 >$$



MD Simulation exercise

- N particles, in 3 D
- Basic Verlet algorithm
- Force-field: full Lennard-Jones potential
- $\varepsilon = 1$
- $\sigma = 1$