

### 5.2.1 Energy

The internal energy is easily obtained from a simulation as the ensemble average of the energies of the states that are examined during the course of the simulation:

$$U = \langle E \rangle = \frac{1}{M} \sum_{i=1}^M E_i$$

### 5.2.4 Temperature

In a canonical ensemble the total temperature is constant. In the microcanonical ensemble, however, the temperature will fluctuate. The temperature is directly related to the kinetic energy of the system as follows:

$$\mathcal{K} = \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m_i} = \frac{k_B T}{2} (3N - N_c) \quad (5.14)$$

In this equation,  $\mathbf{p}_i$  is the total momentum of particle  $i$  and  $m_i$  is its mass. According to the theorem of the equipartition of energy each degree of freedom contributes  $k_B T/2$ . If there are  $N$  particles, each with three degrees of freedom, then the kinetic energy should equal  $3Nk_B T/2$ .  $N_c$  in equation (5.14) is the number of constraints on the system. In a molecular dynamics simulation the total linear momentum of the system is often constrained to a value of zero, which has the effect of removing three degrees of freedom from the system and so  $N_c$  would be equal to 3. Other types of constraint are also possible discuss in section 6.5.

### 5.2.3 Pressure

The pressure is usually calculated in a computer simulation via the virial theorem of Clausius. The *virial* is defined as the expectation value of the sum of the products of the coordinates of the particles and the forces acting on them. This is usually written  $W = \sum x_i \dot{p}_x$ , where  $x_i$  is a coordinate (e.g. the  $x$  or  $y$  coordinate of an atom) and  $\dot{p}_x$  is the first derivative of the momentum along that coordinate ( $\dot{p}_x$  is the force, by Newton's second law). The virial theorem states that the virial is equal to  $-3Nk_B T$ .

In an ideal gas, the only forces are those due to interactions between the gas and the container and it can be shown that the virial in this case equals  $-3PV$ . This result can also be obtained directly from  $PV = Nk_B T$ .

Forces between the particles in a real gas or liquid affect the virial, and thence the pressure. The total virial for a real system equals the sum of an ideal gas part ( $-3PV$ ) and a contribution due to interactions between the particles. The result obtained is:

$$W = -3PV + \sum_{i=1}^N \sum_{j=i+1}^N r_{ij} \frac{d\psi(r_{ij})}{dr_{ij}} = -3Nk_B T \quad (5.12)$$

The real gas part is derived in Appendix 5.3. If  $d\psi(r_{ij})/dr_{ij}$  is written as  $f_{ij}$ , the force acting between atoms  $i$  and  $j$ , then we have the following expression for the pressure:

$$P = \frac{1}{V} \left[ Nk_B T - \frac{1}{3k_B T} \sum_{i=1}^N \sum_{j=i+1}^N r_{ij} f_{ij} \right] \quad (5.13)$$

The forces are calculated as part of a molecular dynamics simulation, and so little additional effort is required to calculate the virial and thus the pressure. The forces are not routinely calculated during a Monte Carlo simulation, and so some additional effort is required to determine the pressure by this route. When calculating the pressure it is also important to check that the components of the pressure in all three directions are equal.

### 5.2.2 Heat capacity

At a phase transition the heat capacity will often show a characteristic dependence upon the temperature (a first-order phase transition is characterised by an infinite heat capacity at the transition but in a second-order phase transition the heat capacity changes discontinuously). Monitoring the heat capacity as a function of temperature may therefore enable phase transitions to be detected. Calculations of the heat capacity can also be compared with experimental results and so be used to check the energy model or the simulation protocol.

The heat capacity is formally defined as the partial derivative of the internal energy with respect to temperature:

$$C_V = \left( \frac{\partial U}{\partial T} \right)_V \quad (5.8)$$

The heat capacity can therefore be calculated by performing a series of simulations at different temperatures, and then differentiating the energy with respect to the temperature. The differentiation can be done numerically or by fitting a polynomial to the data and then analytically differentiating the fitted function. The heat capacity may also be calculated from a single simulation by considering the instantaneous fluctuations in the energy as follows:

$$C_V = \{ \langle E^2 \rangle - \langle E \rangle^2 \} / k_B T^2 \quad (5.9)$$

An alternative way to write this expression uses the relationship

$$\langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 \quad (5.10)$$

giving

$$C_V = \langle (E - \langle E \rangle)^2 \rangle / k_B T^2 \quad (5.11)$$

A derivation of this result is provided in Appendix 5.2.

The heat capacity can therefore be obtained by keeping a running count of  $E^2$  and  $E$  during the simulation, from which their expectation values  $\langle E^2 \rangle$  and  $\langle E \rangle$  can be calculated at the end of the calculation. Alternatively, if the energies are stored during the simulation then the value of  $\langle (E - \langle E \rangle)^2 \rangle$  can be calculated once the simulation has finished. This second approach may be more accurate due to round-off errors;  $\langle E^2 \rangle$  and  $\langle E \rangle^2$  are usually both large numbers and so there may be a large uncertainty in their difference.

# STATISTICAL MECHANICS :

## MICROCANONICAL ENSEMBLE

$$\Omega(N, V, E)$$

$$S = k \ln \Omega$$

## CANONICAL ENSEMBLE

$$Q(N, V, T)$$

$$A = -kT \ln Q$$

## ISOTHERMAL-ISOBARIC ENSEMBLE

$$\Delta(N, T, P)$$

$$G = -kT \ln \Delta$$

$$\mu = \left\{ 1 + \frac{\Delta t}{\tau} B [P(t) - P_0] \right\}^{1/3}$$

$$\alpha = \left\{ 1 + \frac{\Delta t}{\tau} \left[ \frac{T_0}{T(t)} - 1 \right] \right\}^{1/2}$$

## ACTIVATED DYNAMICS:

$$k = \langle \frac{1}{2} \rangle K \langle |E'| \rangle \left[ \frac{\rho(E^*)}{\int \rho(E) dE} \right]$$

ADD UMBRELLA TERM,

$$U(E) = (K/2) (E - E^*)^2$$

START USUAL MD FROM SEVERAL  
POINTS NEAR TRANSITION STATE

## BROWNIAN (STOCHASTIC) DYNAMICS:

LANGEVIN EQUATIONS OF MOTION,

$$m \frac{d^2x}{dt^2} = F(x) - f \frac{dx}{dt} + R(t)$$

FRICION                  RANDOM FORCE

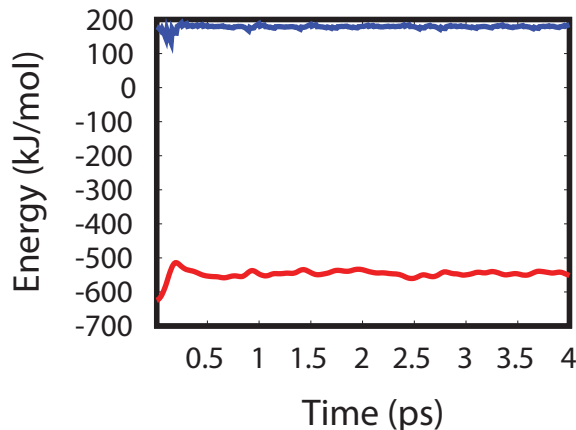
MODELS BEHAVIOR OF AN ENSEMBLE  
OF DIFFUSING MOLECULES

# Running an MD Simulation

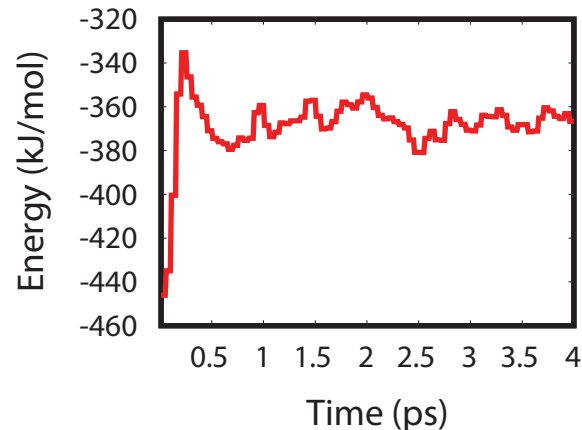
## ▶ Equilibration Phase

- ▶ Before data-collection and results can be analyzed the system must be prepared via equilibration
  - ▶ Minimize energy
  - ▶ Velocity/Pressure scaling (move T/P to desired value)
  - ▶ Heat cycles
  - ▶ Tempering of potential parameters
- ▶ During equilibration monitor thermodynamic properties and structure
- ▶ After achieving stability, perform “production” run
- ▶ Equilibration of the 2D argon system

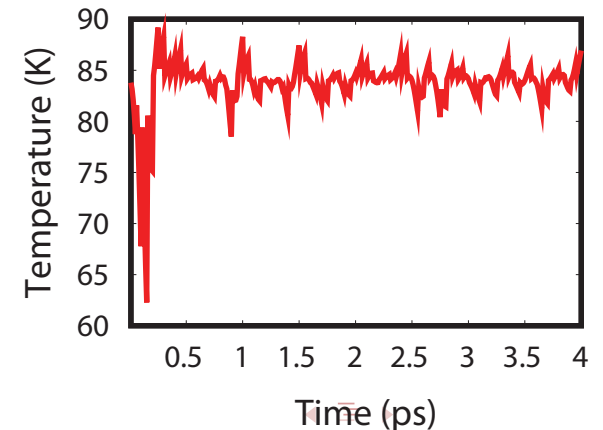
Potential and Kinetic Energy



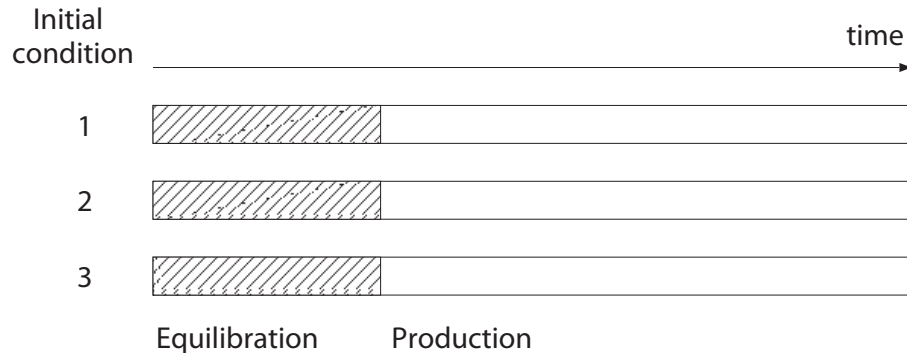
Total Energy



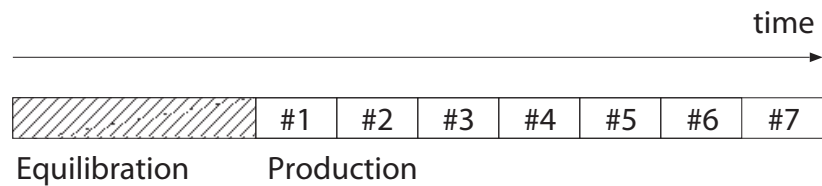
Temperature



- ▶ The properties (such as time averages) should not depend on the initial conditions !
- ▶ Compute averages from several simulations :

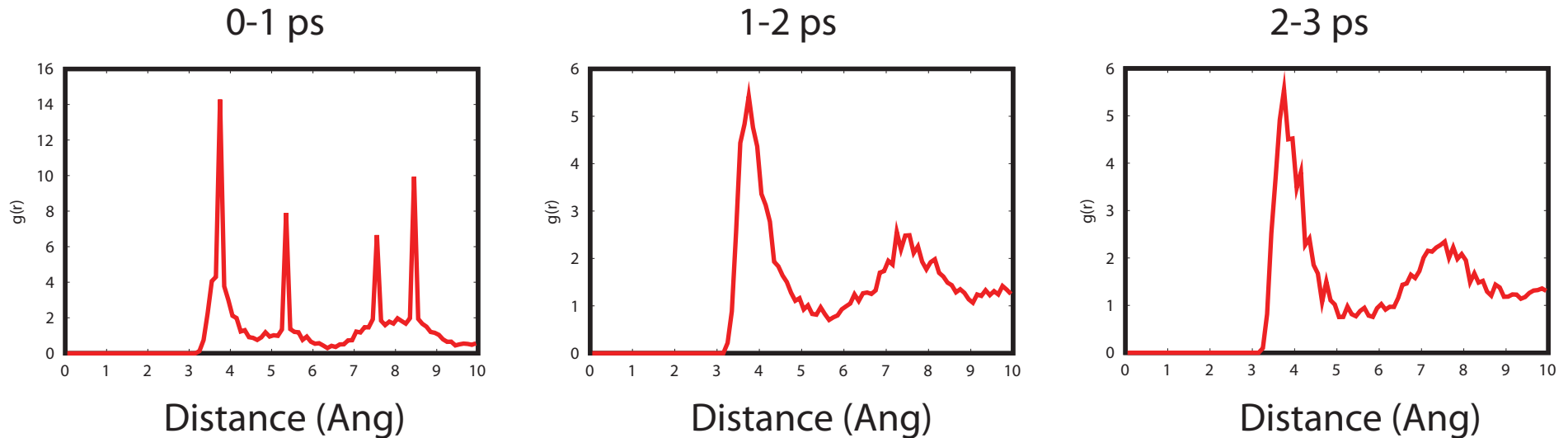


- ▶ Compute block-averages :



- ▶ Difficulty: Block-averages might be the same, because the “equilibration” is very slow
- ▶ Sometimes several simulations are performed with different system sizes to check equilibration

# Evolution of the radial distribution function of the 2D argon system



## ► Production phase

- It is important to continue to check the properties monitored during the equilibration phase during the production phase. They may still not be stable, in which case the beginning or the whole of the simulation might have to be discarded.

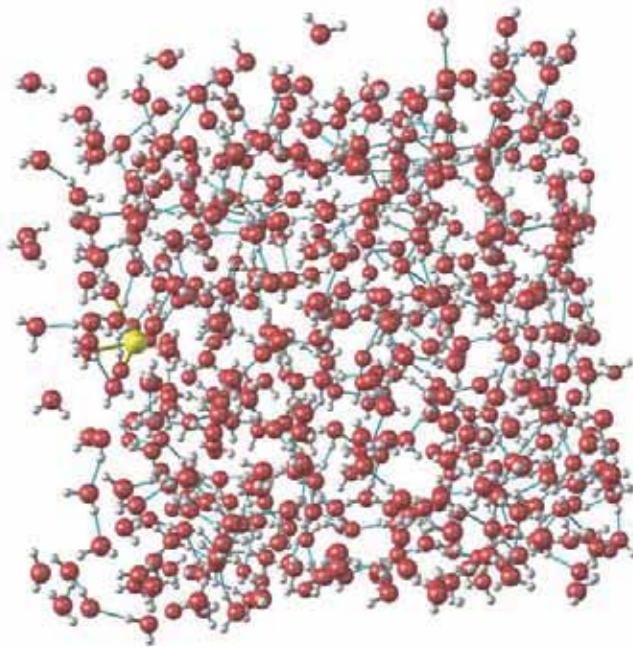


# Properties from MD Simulations

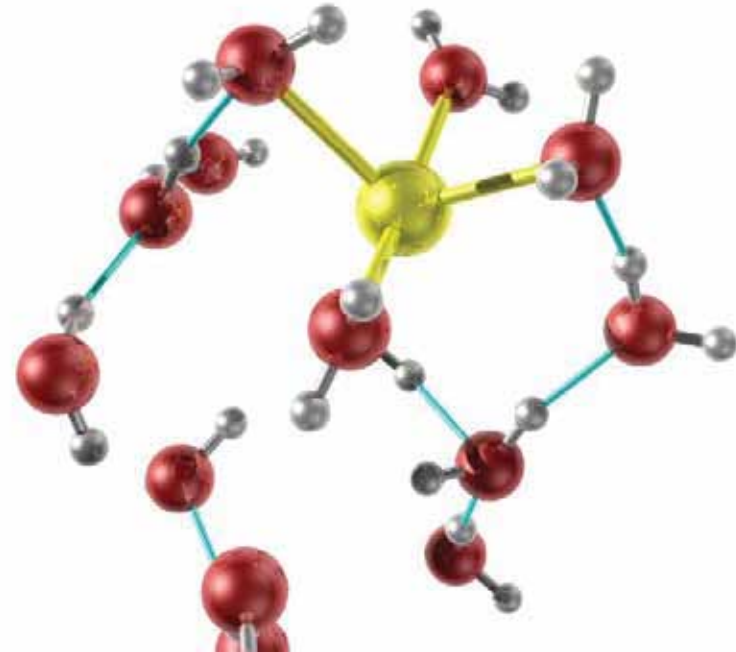
- ▶ Time averages
  - ▶ Thermodynamic properties (energies, pressure ...)
  - ▶ Structural properties
  - ▶ ...
- ▶ Dynamic quantities
  - ▶ Time correlation functions (and their FT's, related to spectroscopic properties.)
  - ▶ Transport properties (diffusion ...)

# Analyzing Simulation Results

- ▶ Directly visualize the results using molecular graphics. The results can (of course) be analyzed by inspection, although this is not as trivial as it may sound !



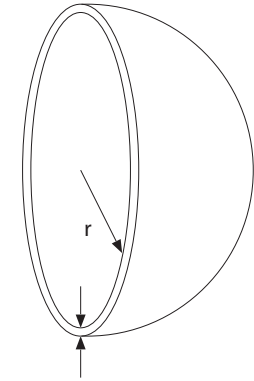
Snapshot from a simulation containing 512 water molecules and one Na<sup>+</sup> ion



Local environment of Na<sup>+</sup> (aq)

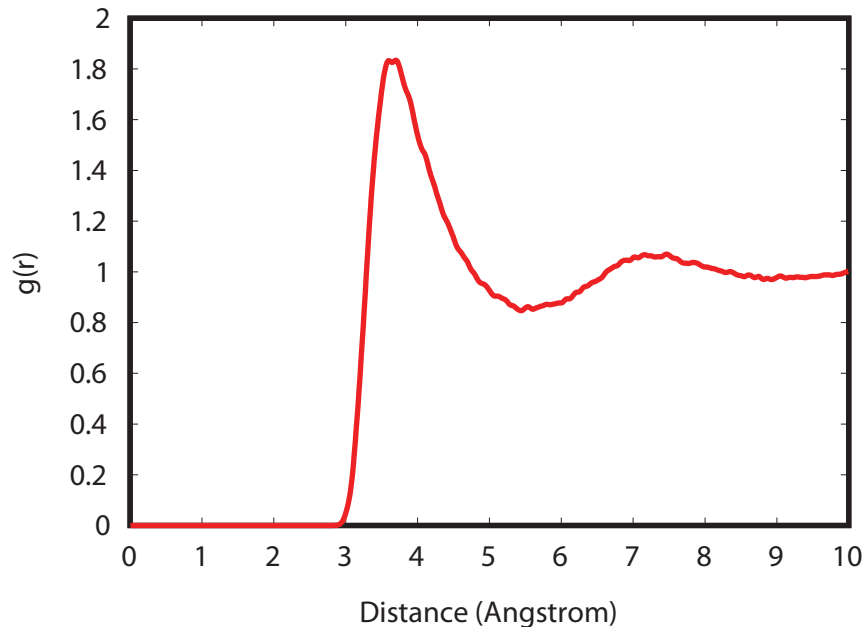
# Structural Properties

The radial distribution function gives a measure of the local structure. It corresponds to the local concentration of particles in a (thin) spherical shell at the distance  $r$  around a central particle, relative to a uniform distribution of particles.

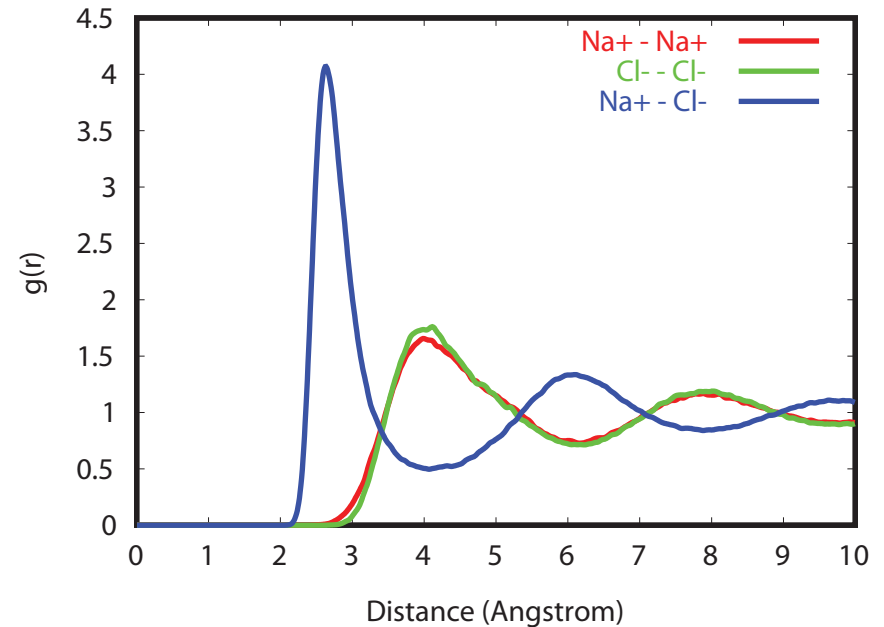


## ► Examples

Liquid Argon



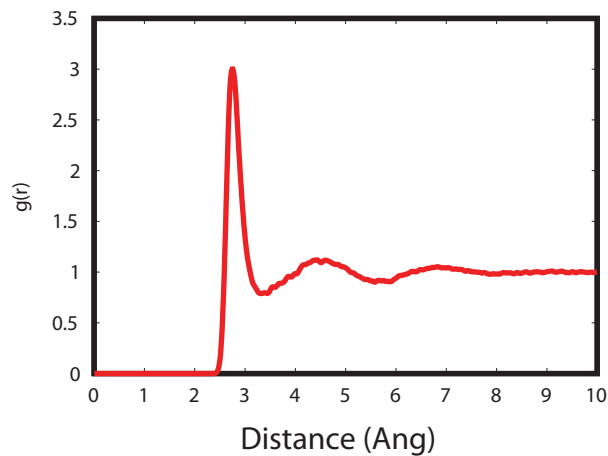
NaCl Melt



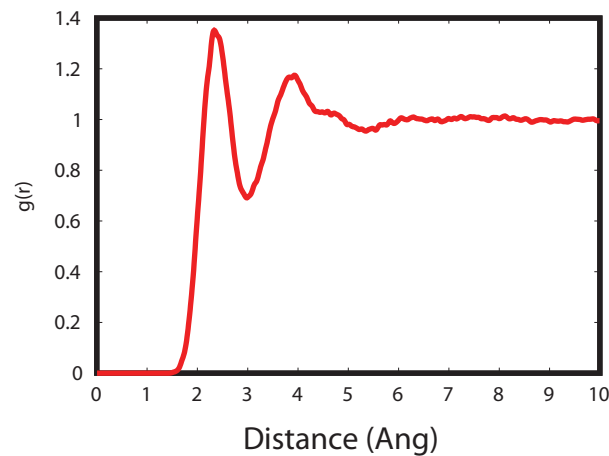
# Radial Distribution Functions

- Useful for molecular liquids. Site-site radial distribution functions for water :

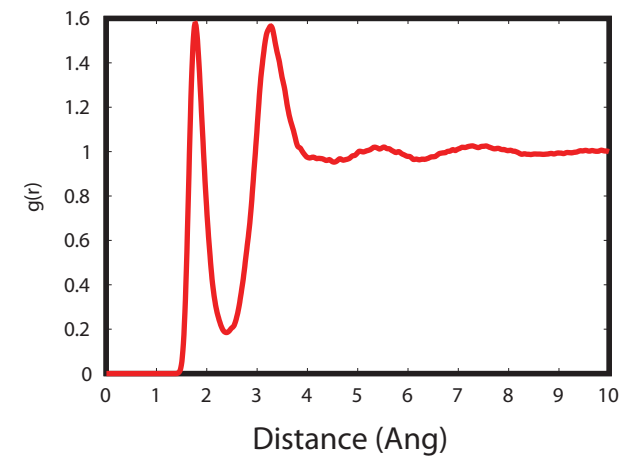
Oxygen-Oxygen



Hydrogen-Hydrogen



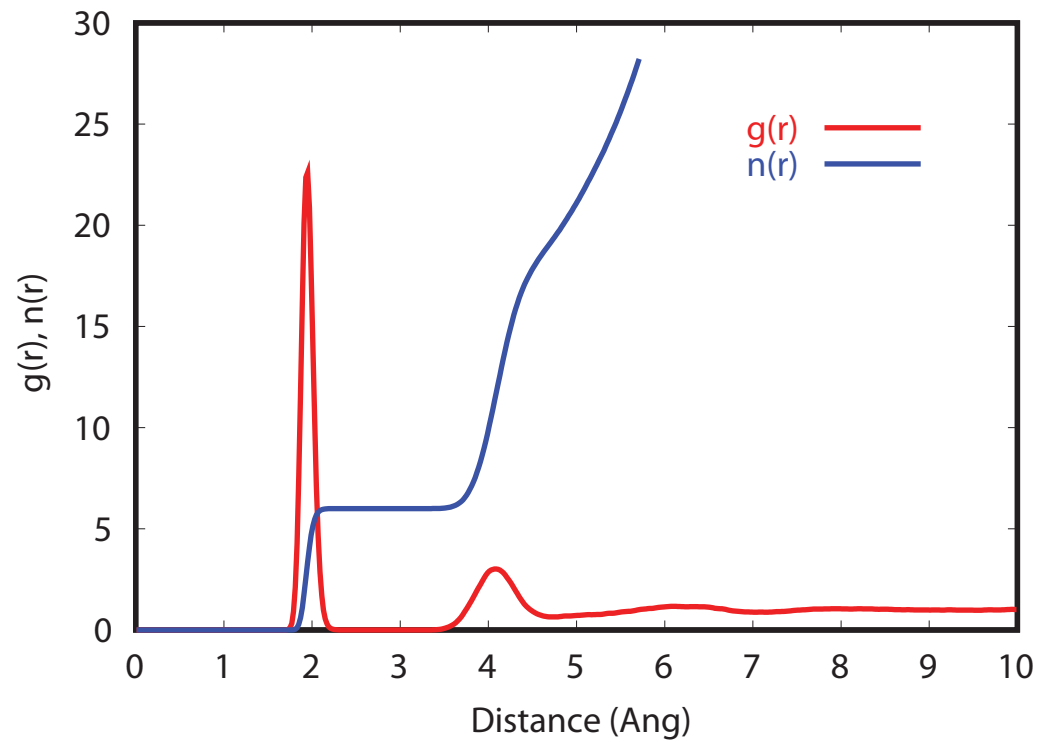
Oxygen-Hydrogen



For large molecules the number of site-site distribution functions obviously becomes very large and only a subset is usually computed

- For molecules it is also possible to compute various angular dependent functions, and/or compute so-called spatial distribution functions

- ▶ Integrating the radial distribution function gives the number of particles surrounding the central particle



Structure of water around an  $\text{Al}^{3+}$  ion:  $g_{\text{Al}-\text{O}}(r)$  and  $n_{\text{Al}-\text{O}}(r)$

# Dynamical Properties

## ▶ Time Correlation Functions

- ▶ Correlations between two different quantities A and B are measured using a time correlation function :

$$C_{AB}(t) = \langle A(t) \cdot B(0) \rangle$$

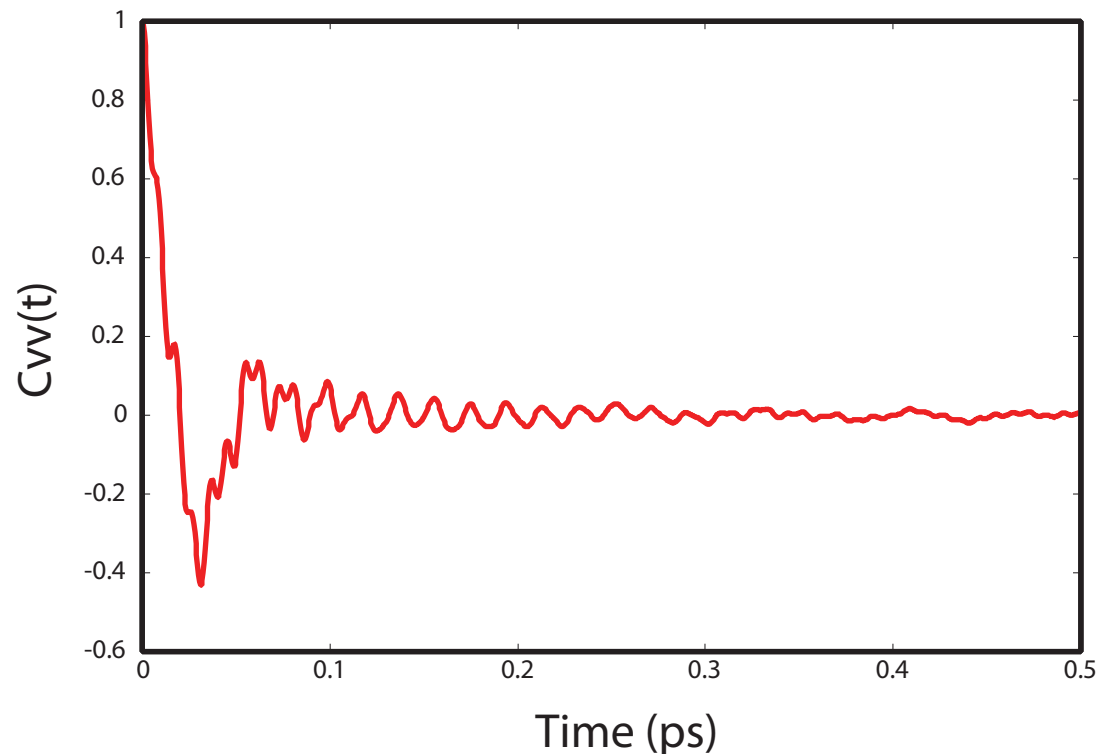
- ▶ Such time correlation functions are interesting since :
  - ▶ They give a picture of the dynamics in the system
  - ▶ Their time integrals are often related to various transport properties
  - ▶ Their Fourier transforms are often related to experimental spectra
- ▶ If A and B are different properties, C is called a cross correlation function
- ▶ If A and B are the same property, C is called an auto correlation function. The auto correlation function is a measure of the “memory” of the system for some property

## ▶ Dynamical Properties: Time Correlation Functions (cont)

- ▶ If  $A(t)$  is a property of many particles the correlation function is collective
- ▶ If  $A(t)$  is a property of a single particle the function is a single particle correlation function
- ▶ The single particle velocity auto-correlation (VAC) function :

$$C_{vv}(t) = \langle v(t) \cdot v(0) \rangle$$

- ▶ Example : Hydrogen in liquid water

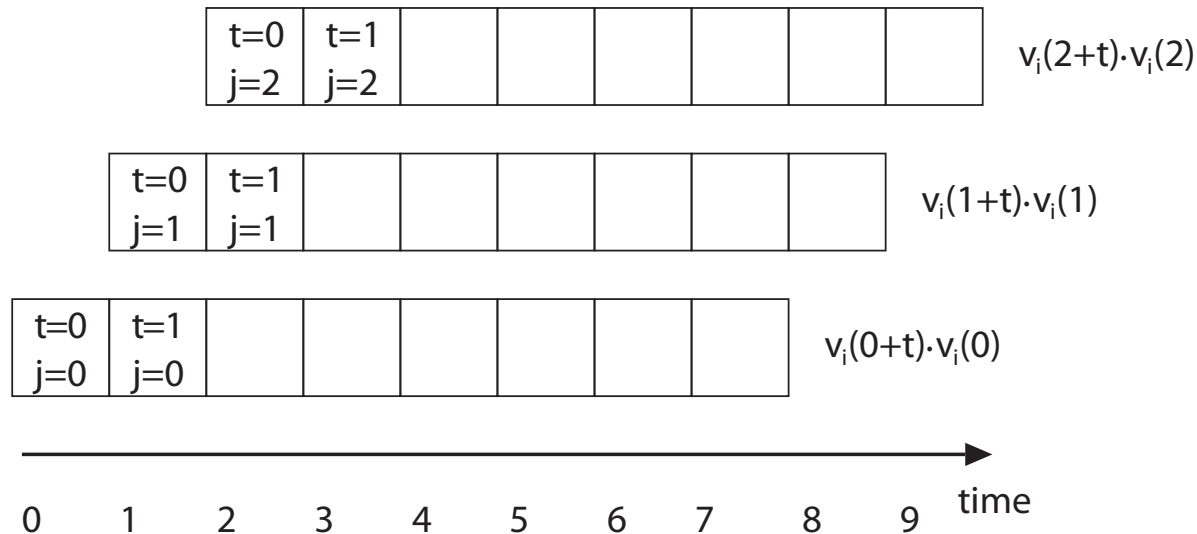


► Dynamical Properties: Time Correlation Functions (cont)

- The average in the velocity auto-correlation function is typically taken over all particles in the system and for a number of different time origins

$$\langle v(t) \cdot v(0) \rangle = \frac{1}{N} \sum_i^N \langle v_i(t) \cdot v_i(0) \rangle$$

$$\langle v_i(t) \cdot v_i(0) \rangle = \frac{1}{M} \sum_j^M v_i(t_j + t) \cdot v_i(t_j)$$





▶ Dynamical Properties: Time Correlation Functions (cont)

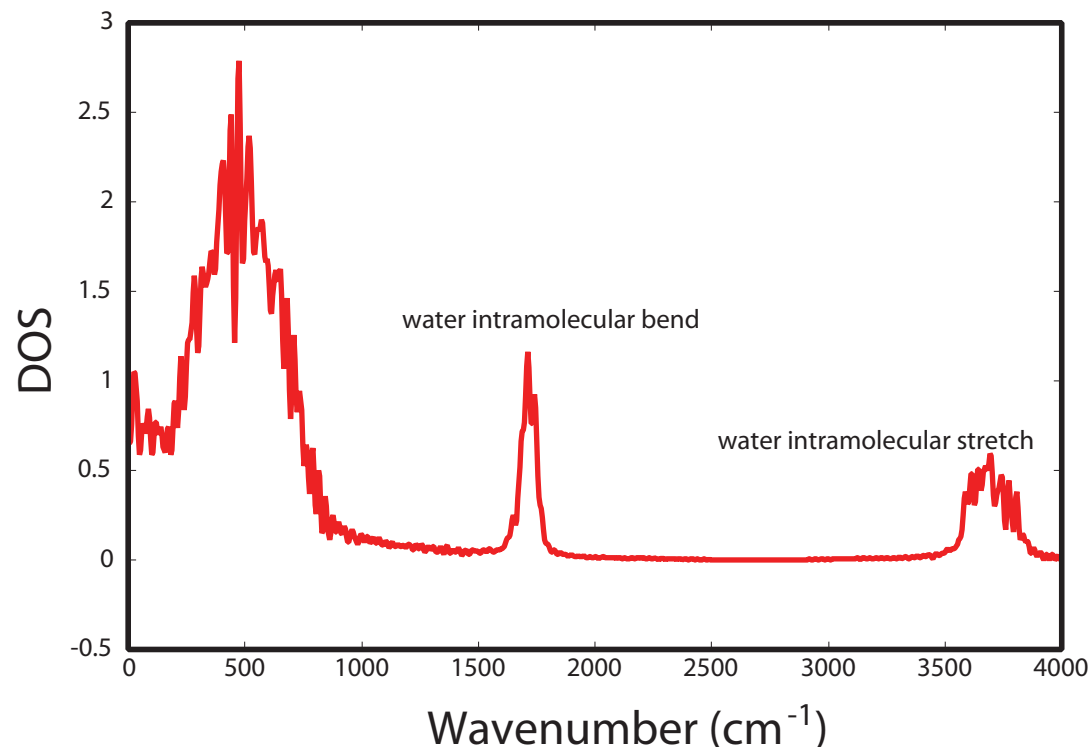
- ▶ The normalized time correlation function is

$$C_{AB}(t) = \frac{\langle A(t) \cdot B(0) \rangle}{\langle A(0) \cdot B(0) \rangle}$$

- ▶ Fourier transforming the correlation function

$$\hat{C}_{AB}(\omega) = \int_{-\infty}^{\infty} C_{AB}(t) e^{-i2\pi\omega t} dt$$

- ▶ Fourier transform of the hydrogen in liquid water VAC



# Transport Properties

- ▶ Integrating the velocity auto-correlation function gives the diffusion coefficient

$$D = \frac{1}{3} \int_0^{\infty} \langle v(t) \cdot v(0) \rangle dt$$

- ▶ This is an expression of the general type

$$\gamma = \int_0^{\infty} \langle \dot{A}(t) \cdot \dot{A}(0) \rangle dt$$

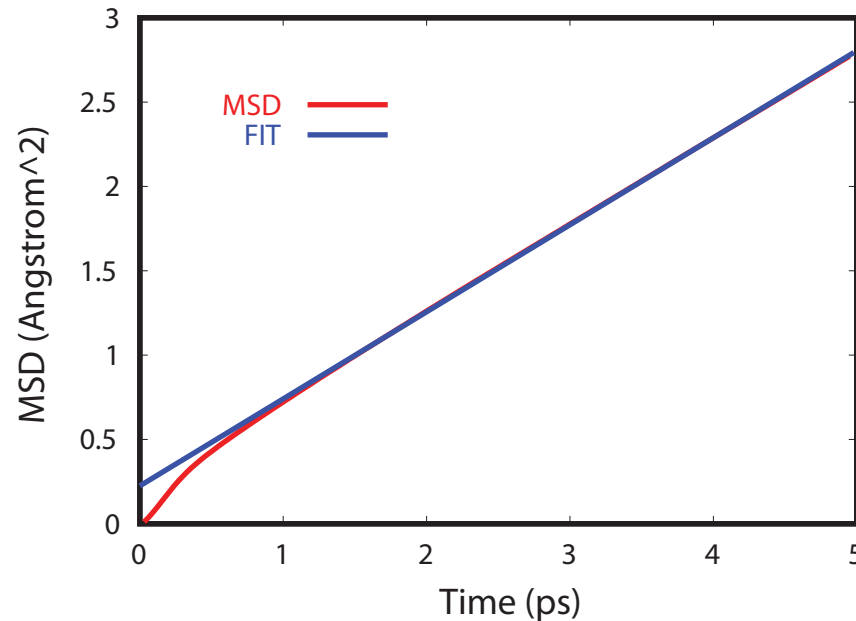
- ▶ The corresponding “Einstein relation” is

$$2t\gamma = \langle (A(t) - A(0))^2 \rangle$$

- ▶ An alternate way to compute the diffusion coefficient

$$2tD = \frac{\langle |r(t) - r(0)|^2 \rangle}{3}$$

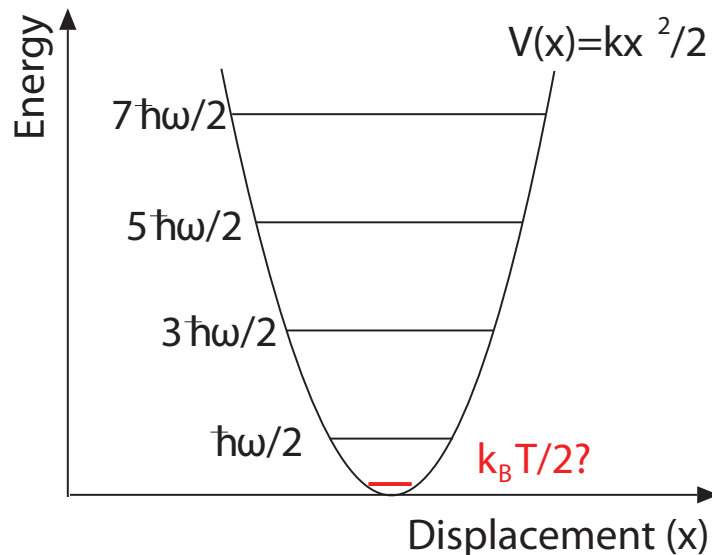
- ▶ The Einstein relation holds at long times !



- ▶ Examples of other dynamical properties that can be studied using time correlation functions and/or Einstein relations :
  - ▶ Total dipole moment auto-correlation function :  
Related to (infrared) absorption spectrum
  - ▶ Auto-correlation function of elements of the pressure tensor :  
Related to the viscosity
  - ▶ Orientational correlation functions :  
Related to various spectroscopic techniques (NMR, IR, Raman ...)

# Handling Fast Vibrational Motion

- ▶ Vibrational motion with high frequencies ( $\hbar \omega \approx k_B T$ ) are really quantized



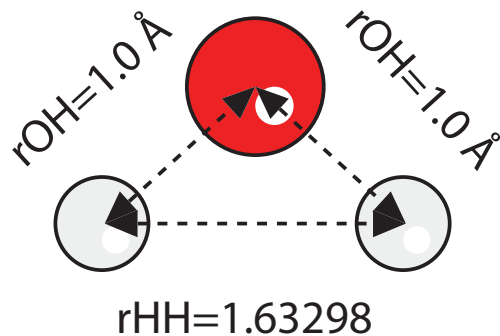
- ▶ Also, since the frequencies are very high, short timesteps are required
- ▶ Flow of energy might be slow, due to poor coupling between the degrees of freedom. This can lead to problems with equilibration.

# Treatment of Rigid Molecules

- ▶ One “solution” is to make molecules / bonds rigid !
- ▶ Rigid molecules
  - ▶ Separate motion into translation and rotation ; Separate equations of motion for the center of mass, and some representation of the rotation of the molecule (use Euler angles or quaternions)
- ▶ Rigid bonds
  - ▶ Constraint dynamics (for “holonomic constraints”)
  - ▶ Appropriate for molecules that are partially flexible, such as a polymer
  - ▶ Rigid small molecules can also be handled by introducing fixed “bonds”, three per atom (or, actually,  $3N - 6$  bonds per molecule)

# Constraint Dynamics

- ▶ Relatively simple algorithms exist: SHAKE and RATTLE
- ▶ SHAKE enforces the (for instance) distance between two atoms to be constant



- ▶ Holonomic constraints :

$$f(q_1, q_2, \dots, t) = 0$$

$$r_{ij}^2 - d_{ij}^2 = 0$$

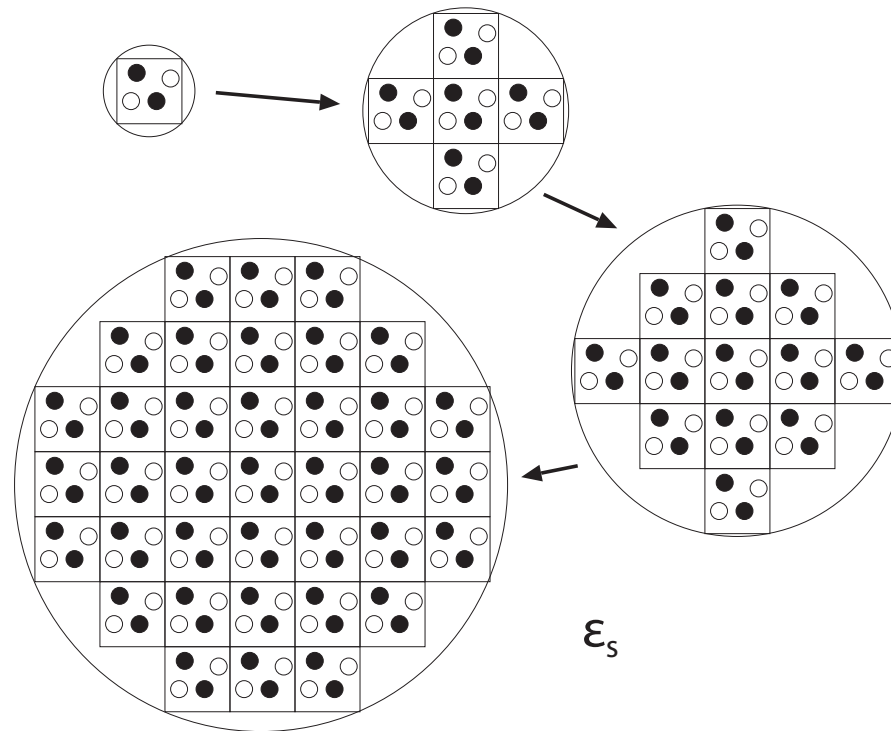
- ▶ The SHAKE method is tightly connected to the integrator used, the variant for the velocity Verlet integrator is termed RATTLE
- ▶ The method introduces an extra force directed along the bond between two atoms at time zero (i.e. before the integration)
- ▶ First the integration step is completed as if there were no constraint force
- ▶ Then all constraint forces are solved for, one by one, iteratively

# Long Range Interactions

- ▶ A long range interaction decays no faster than  $r^{-d}$ , where “d” is the dimensionality of the system
- ▶ The problem : The interaction decay is so slow that we cannot just truncate it at a reasonably short distance
- ▶ Even worse : Conditionally convergent !
- ▶ Important members of this class :
  - ▶ charge-charge ( $r^{-1}$ )
  - ▶ charge-dipole ( $r^{-2}$ )
  - ▶ dipole-dipole ( $r^{-3}$ )
  - ▶ charge-quadrupole ( $r^{-3}$ )
- ▶ The Coulomb interaction :

$$V = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^N \sum_{j=i+1}^N \frac{q_i q_j}{r_{ij}}$$

- ▶ Different methods to treat this kind of interaction have been devised (Ewald, reaction field, various multipole methods)
- ▶ Here only the ordinary Ewald method is (briefly) considered
- ▶ Sum over periodic images built up in spherical layers :



- ▶ The very large sphere is surrounded by a medium with relative permittivity  $\epsilon_s$

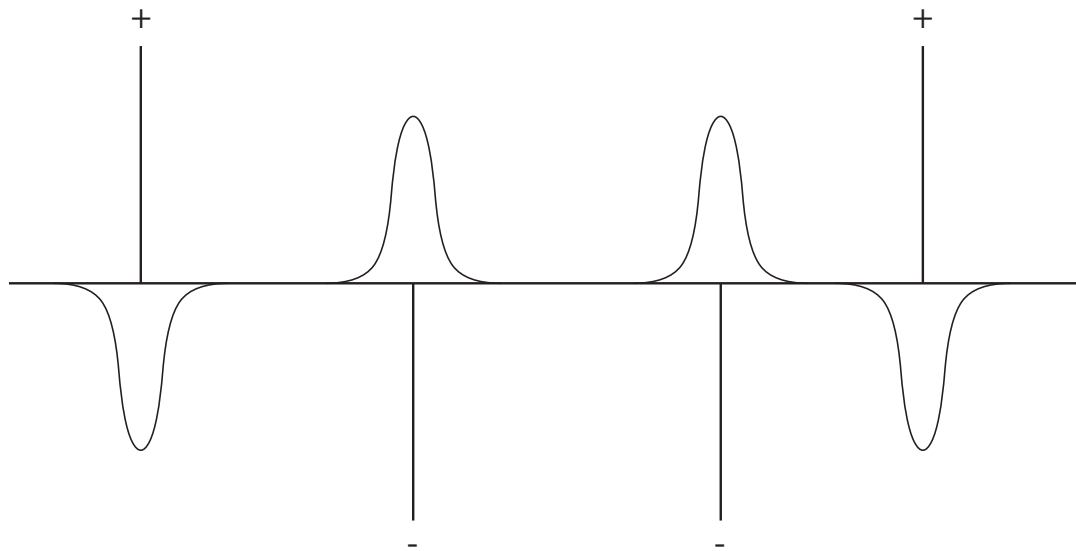


- ▶ The potential energy can be written as :

$$V = \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_n' \left[ \sum_i^N \sum_j^N \frac{q_i q_j}{|r_{ij} + n|} \right]$$

where  $n = (n_x L, n_y L, n_z L)$ .  $n_x, n_y, n_z$  are integers and  $L$  the size of the central image. The ' in the sum :  $i \neq j$  for  $|n| = 0$

- ▶ The Ewald method : Add screening charge distribution with opposite charge and equal magnitude ( $\alpha^3 \pi^{-3/2} e^{-\alpha^2 r^2}$ )

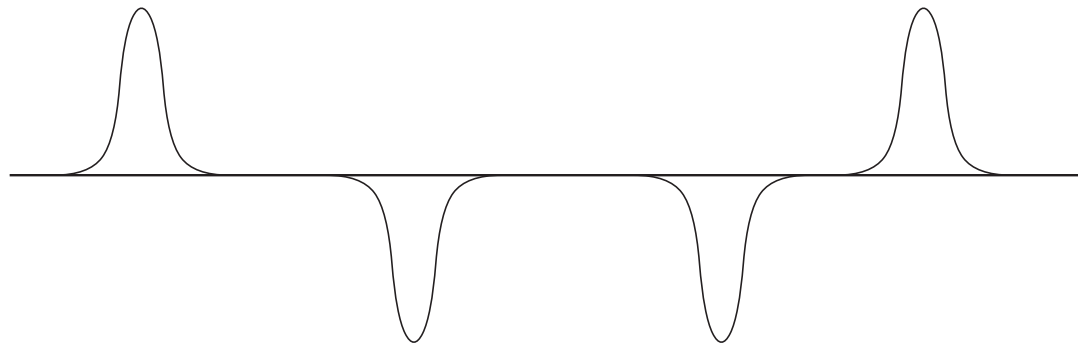


- ▶ The interaction between the charges is now short ranged :

$$V_{\text{real}} = \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_n' \left[ \sum_i^N \sum_j^N q_i q_j \frac{\text{erfc}(\alpha|r_{ij} + n|)}{|r_{ij} + n|} \right]$$

$$\text{erfc}(x) = 2\pi^{-1/2} \int_x^\infty e^{-t^2} dt$$

- ▶ For suitable values of the  $\alpha$  parameter,  $|n|$  can be truncated to 0
- ▶ The original potential is restored by adding a canceling charge distribution :



- ▶ The canceling distribution is summed in reciprocal (Fourier) space :

$$V_{\text{reciprocal}} = \frac{1}{\epsilon_0 L^3} \frac{1}{2} \sum_{k \neq 0} \frac{1}{|k|^2} e^{-\frac{|k|^2}{4\alpha^2}} \left[ \sum_i^N \sum_j^N q_i q_j \cos(r_{ij} \cdot k) \right]$$

- ▶ The sum goes over reciprocal vectors ,  $k = 2\pi n/L$
- ▶ A correction term needs to be subtracted as the sum in reciprocal space includes the interaction of the canceling distribution at  $r_i$  with itself :

$$V_{\text{self}} = -\frac{\alpha}{4\pi^{3/2} \epsilon_0} \sum_i^N q_i^2$$

- ▶ The expression  $V = V_{\text{real}} + V_{\text{reciprocal}} + V_{\text{self}}$ , corresponds to the potential energy for the large sphere surrounded by a medium with  $\epsilon_s = \infty$