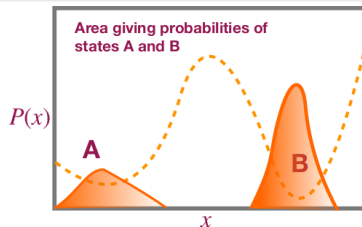
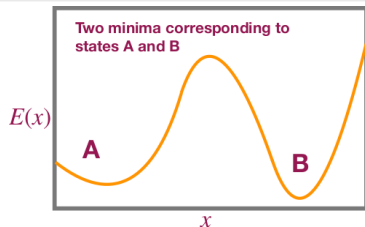


Molecular Modelling and Simulations

Shampa Raghunathan

Boltzmann distribution



- Probability of system being at position x is

$$P(x) = \frac{1}{Q} e^{-\frac{E(x)}{k_B T}}$$

$E(x)$ is the energy at position x ; k_B = Boltzmann constant

- Find Q , the “partition function” function

$$Q = \sum e^{-\frac{E(x)}{k_B T}}$$

such that, the total probability is 1:

$$\sum P(x) = 1$$

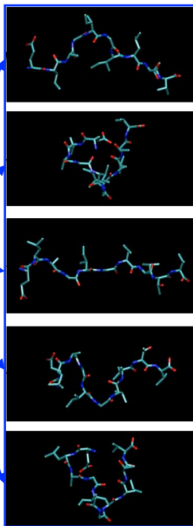
Boltzmann distribution

Macroscopic

Average

$$\langle A \rangle = \frac{1}{Q} \sum_{\alpha} A_{\alpha} e^{-\beta E_{\alpha}}$$

Where, $Q = \sum_{\alpha} e^{-\beta E_{\alpha}}$



Microscopic

$$E_1, P_1 \sim e^{-\beta E_1}$$

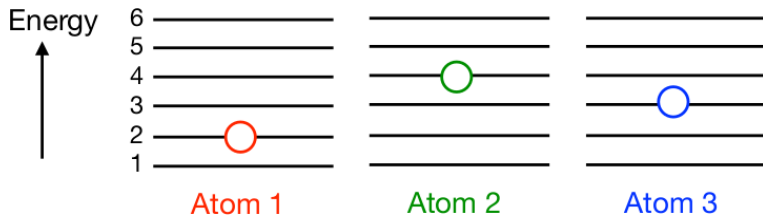
$$E_2, P_2 \sim e^{-\beta E_2}$$

$$E_3, P_3 \sim e^{-\beta E_3}$$

$$E_4, P_4 \sim e^{-\beta E_4}$$

$$E_5, P_5 \sim e^{-\beta E_5}$$

Canonical and microcanonical ensembles compared



Microcanonical ensemble:

Equally probable microstates are permissible, for example, if, total $E = 9$, then $\{2,4,3\}$ and $\{3,3,3\}$ would be allowable microstates (with equal probability) though $\{3,3,4\}$ not allowed

Canonical ensemble:

All microstates would be possible, each one with value $\{n_1^j, n_2^j, n_3^j\}$, where $j = 1, 6$; however, probability of each microstate would not be identical.

Canonical ensemble

Systems with constraints (N, V, T) are referred to as being in the **canonical ensemble**, which is sometimes referred to as the NVT ensemble.

$$Q_{NVT} = \sum_{\alpha} e^{-\beta E_{\alpha}} \quad (1)$$

As, $\beta = 1/k_B T$ now, the averages in the canonical ensemble takes the form,

$$\langle E \rangle = \frac{1}{Q_{NVT}} \sum_{\alpha} E_{\alpha} e^{-\frac{E_{\alpha}}{k_B T}} \quad (2)$$

here, as an example average energy is shown.

Similarly, average kinetic energy is related to temperature

$$\langle K \rangle = \frac{3}{2} N k_B T$$

Maxwell-Boltzmann distribution

From the canonical partition function—the **velocity** (v) distributions in an ideal gas in thermal equilibrium:

$$P(v_x, v_y, v_z) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} e^{-\frac{1}{2} \frac{mv^2}{k_B T}} \quad (3)$$

Derivation

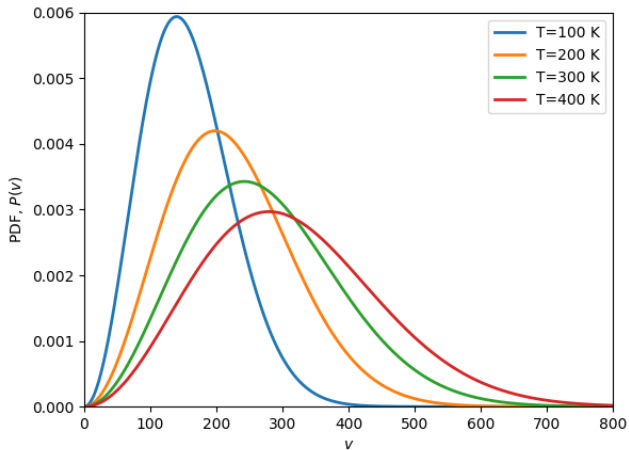
Integrating each component of the velocity of an atom, partition function becomes

$$Q_{NVT} = \int e^{-\frac{1}{2} \frac{mv^2}{k_B T}} dv_x = \sqrt{\frac{2\pi k_B T}{m}} \quad (4)$$

Solving as Gaussian integral of the form $\int e^{-ax^2} dx = \sqrt{\pi/a}$. For each atom we have three identical integrals multiplying each other, one for v_x , one for v_y , and one for v_z . Thus the contribution from each atom

becomes $\left(\sqrt{\frac{2\pi k_B T}{m}} \right)^3$.

Maxwell-Boltzmann distribution



Outline

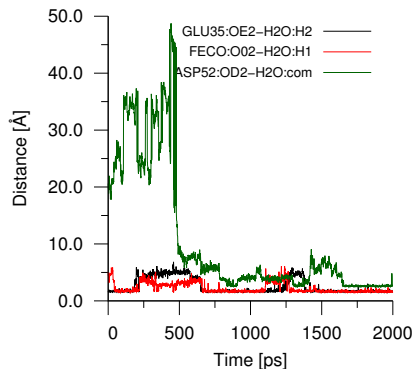
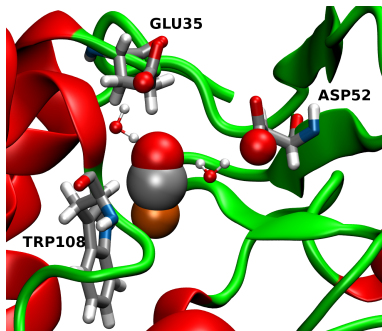
- 1 **Computing properties**
 - Fluctuations
 - Correlation functions

Properties from MD simulations

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

Analyzing simulation results

- ▶ Directly visualization of results using molecular graphics. The results can (of course) be analyzed by visual inspection, however this is not often trivial!



VMD - Visual Molecular Dynamics
(<https://www.ks.uiuc.edu/Research/vmd/>)

Averages

The **average** value of E

$$\langle E \rangle = \frac{1}{\tau_{max}} \sum_{\tau=1}^{\tau_{max}} E(\tau)$$

The **variance** of E is defined as,

$$\sigma_E^2 = \langle E^2 \rangle - \langle E \rangle^2$$

The variance of E gives a measure of the size of the fluctuations in its values around its average value.

Averages

If the data are divided into n_b bins of equal length τ_b then average quantities are given by

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

where, the bin averages are

$$\langle E_i \rangle = \frac{1}{\tau_b} \sum_{\tau=\tau_0}^{\tau_0+\tau_b} E(\tau)$$

Fluctuations

In equilibrium all thermodynamic quantities fluctuate around an average value

$$\sigma_F^2 = \langle (F - \langle F \rangle)^2 \rangle = \langle F^2 \rangle - \langle F \rangle^2 = \langle (\delta F)^2 \rangle$$

From thermodynamics, the derivative of E with respect to T at constant volume is defined **heat capacity** in the canonical ensemble

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{N,V}$$

Further writing in terms of fluctuations of energy

$$C_V = \frac{1}{k_B T^2} (\langle E^2 \rangle - \langle E \rangle^2)$$

Correlation functions

Correlations between two quantities A and B can be given by

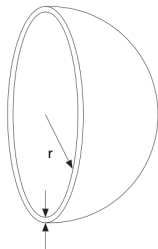
$$C_{AB} = \frac{\langle (A - \langle A \rangle) (B - \langle B \rangle) \rangle}{\sigma_A \sigma_B}$$

where, the standard deviations are $\sigma_A = \sqrt{\langle (A - \langle A \rangle)^2 \rangle}$.

The correlation function C_{AB} is normalized, such that, if the values of A and B are perfectly correlated, then $C_{AB} = 1$. If A and B are independent and not correlated at all, then $C_{AB} = 0$.

Spatial correlations: Structural properties

The **radial distribution function (RDF)** 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.



$$N(r) = \int_0^r \rho g(r) 4\pi r^2 dr$$

$g(r)$ is the pair correlation function.

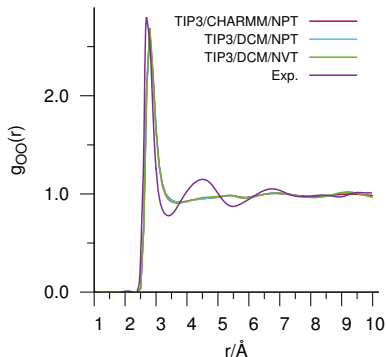
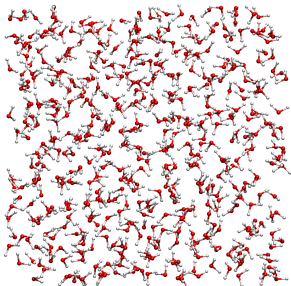
- The integral of $g(r)$ allows to determine the number of neighbors around a central atom.

Radial distribution function, RDF

$$g_{AB}(r) = \frac{N_{AB}(r, \Delta r)}{\rho_B V_S(r, \Delta r)}$$

$N_{AB}(r, \Delta r)$ is the average number of B sites found in a shell, Δr thick, at distance r from the A sites, V_S is the volume of the shell, and ρ_B is the average number density of B sites in the system.

~ (25 Å)³ box with 512 H₂O molecules



Time correlations

Correlations between one time t and another time t' , let's say between $A(t)$ and $B(t')$ is called a **time correlation function**

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

separated by the time interval t , and averaged over the whole trajectory.

An **autocorrelation function** describes a correlation between the value of a quantity at one time and its value at another, *i.e.*,

$$C_{AA}(t) = \frac{\langle (A(t) - \langle A \rangle) \rangle \langle (A(0) - \langle A \rangle) \rangle}{\sigma_A^2}$$

Time correlations: Dynamical properties

The **velocity autocorrelation function** (the correlation of the velocity with itself), which is defined as

$$C_{vv}(t) = \frac{\langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle}{\langle v^2 \rangle}$$

The above equation is averaged over all atoms; replacing $\langle v^2 \rangle = 3k_B T / m$

$$C_{vv}(t) = \frac{m}{3k_B T} \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle$$

- ▶ An important property of the time autocorrelation function $C_{AA}(t)$ is that by taking its Fourier Transform, $F\{C_{AA}(t)\}_\omega$, one gets a spectral decomposition of all the frequencies that contribute to the motion.

Time correlations: Dynamical properties

Diffusion constant D is related to the velocity autocorrelation function by

$$D = \lim_{t \rightarrow \infty} \frac{k_B T}{m} \int_0^t C_{vv}(t') dt' = \lim_{t \rightarrow \infty} \frac{1}{3} \int_0^t \langle \mathbf{v}(t') \cdot \mathbf{v}(0) \rangle dt'$$

D can also be computed from the mean square displacement (MSD) of all atoms using the Einstein relation

$$D = \lim_{t \rightarrow \infty} \frac{1}{6t} \langle |r(t) - r(0)|^2 \rangle, \quad (5)$$

where $r(t)$ is the position of the particle at time t , and averaged over all N particles