Essentials of Ensemble Averages

E. J. Maginn, J. K. Shah

Department of Chemical and Biomolecular Engineering
University of Notre Dame
Notre Dame, IN 46556 USA

Monte Carlo Workshop
Universidade Federal do Rio de Janeiro
Rio de Janeiro Brazil

UNIVERSITY OF NOTRE DAME

Thermodynamic Properties from Ensemble Averages

- Impossible to compute Q directly for non-trivial systems
- Simulations designed to compute *ensemble averages* (or, equivalently, time averages for MD).
- Recall that averages may be computed if probability density known. For example

$$\langle A \rangle = \int \rho_{NVT}(\mathbf{q}^N, \mathbf{p}^N) A(\mathbf{q}^N, \mathbf{p}^N) \tag{1}$$

$$\langle A \rangle = \frac{\int dp^{3N} dq^{3N} \exp\left[-\beta(\mathcal{H}(p^{3N}, q^{3N})\right] A(p^{3N}, q^{3N})}{\int dp^{3N} dq^{3N} \exp\left[-\beta(\mathcal{H}(p^{3N}, q^{3N})\right]} \tag{2}$$

- Still requires enumeration $\mathbf{p}^N$ and $\mathbf{q}^N$
- Even for $n = 100$ particles, this is a HUGE space.

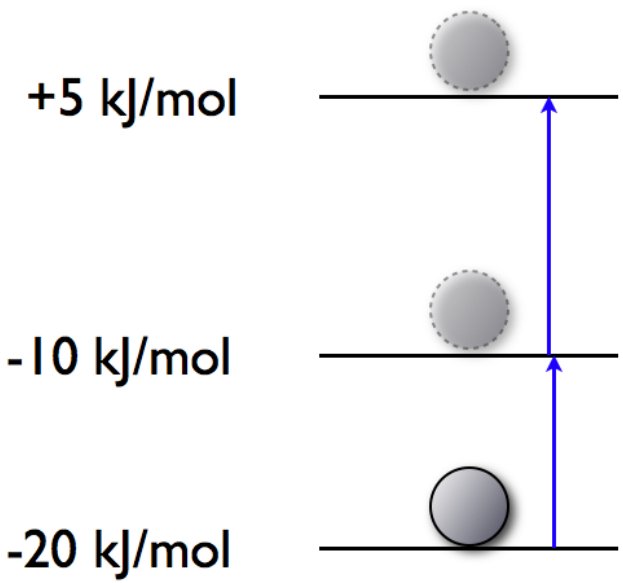
Thermodynamic Properties from Ensemble Averages

- Thermodynamic properties obtained from ensemble averages. Examples:
- with $\rho^{ens}(\mathbf{r}^N) = \rho^{NVT}(\mathbf{r}^N) = \exp[-\beta\mathcal{V}(\mathbf{r}^N)]$
- Excess internal energy: $U^{ex} = \langle\mathcal{V}\rangle_{NVT}$
- Pressure (virial theorem): $PV - Nk_BT = \left\langle \frac{1}{3}\sum_i \mathbf{r}_i\mathbf{F}_i^{int} \right\rangle_{NVT}$
- Excess chemical potential: $\mu^{ex} = \langle\exp[-\beta\mathcal{V}_{test}]\rangle_{NVT}$

Example: Averages in a 3-state system

Consider the following three-state system at constant NVT.

- What is $< E >$ at 200K, 300K, 1000K and 10 000K?
- How does this compare to the arithmetic average energy?
- What percentage of time will the system be in each state?
- Compute this on your own (by “hand” or write a small program / spreadsheet to do it)



Three State System

- Average Energy

$$\frac{E_1 \exp(-E_1/(RT)) + E_2 \exp(-E_2/(RT)) + E_3 \exp(-E_3/(RT))}{\exp(-E_1/(RT)) + \exp(-E_2/(RT)) + \exp(-E_3/(RT))} \quad (3)$$

- 200K: -19.98 kJ/mol
- 300K: -19.82 kJ/mol
- 1000K: -16.86 kJ/mol
- 10 000K: -9.58 kJ/mol
- Arithmetic average: $(5 - 10 - 20)/3 = -8.33$ kJ/mol

UNIVERSITY OF NOTRE DAME

Three State System

Percentage time in each state is the probability of each state

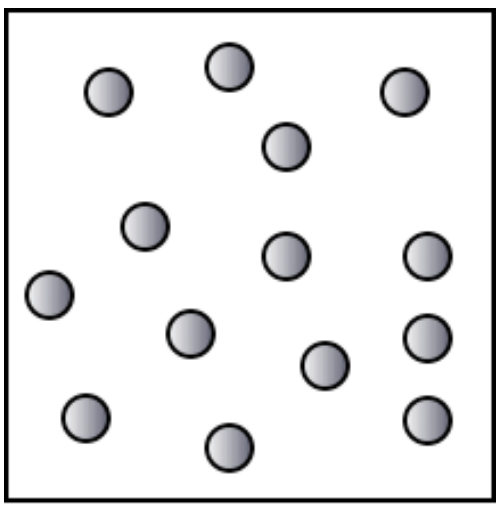
$$P_i = \frac{\exp(-E_i/(RT))}{\exp(-E_1/(RT)) + \exp(-E_2/(RT)) + \exp(-E_3/(RT))} \quad (4)$$

- 200K: $P_1 = 0.000\%$, $P_2 = 0.244\%$, $P_3 = 99.756\%$
- 300K: $P_1 = 0.004\%$, $P_2 = 1.782\%$, $P_3 = 98.214\%$
- 1000K: $P_1 = 3.663\%$, $P_2 = 22.252\%$, $P_3 = 74.085\%$
- 10 000K: $P_1 = 28.181\%$, $P_2 = 33.753\%$, $P_3 = 38.067\%$

Why we can't compute averages from explicit integration for fluids?

The number of "states" available to a fluid are enormous

- 300 molecules
- 300×3 positions
- Number of relative positions is nearly infinite

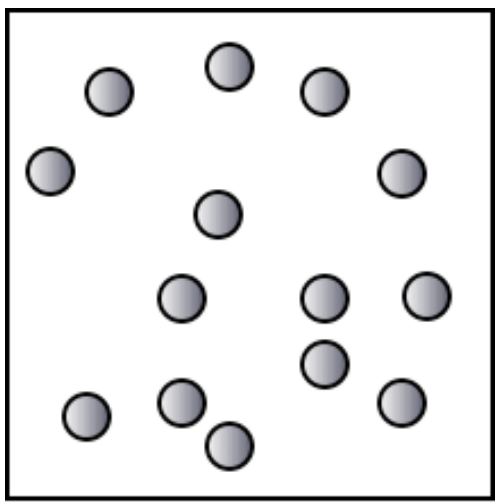


E_1

Why we can't compute averages from explicit integration for fluids?

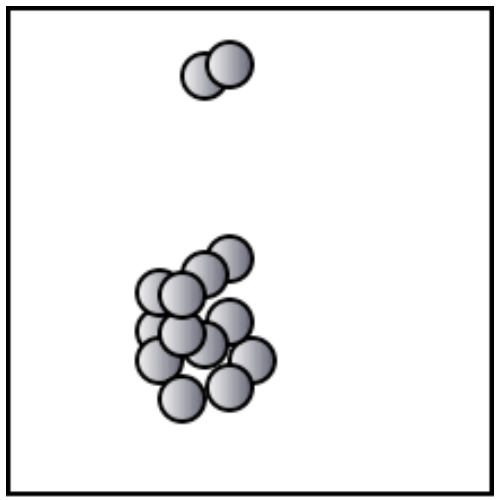
The number of "states" available to a fluid are enormous

- 300 molecules
- 300×3 positions
- Number of relative positions is nearly infinite



E_2

Why we can't compute averages from explicit integration for fluids?



E_3

Even highly improbable configurations must be counted (and actually make up most configurations!)

Thermodynamic Properties from Ensemble Averages

What if we could generate configurations that appeared in a simulation with probability

$$P_{sim}\left(\mathbf{q}^N\right) = \rho_{NVT} \tag{5}$$

Then for N configurations

$$\langle A \rangle_{NVT} = \frac{1}{N}\sum_{i=1}^{N} A(\mathbf{q}_i^N) \tag{6}$$

That would be great!

Generating Configurations Accordings to a Probability Distribution

Molecular dynamics does this

- Equations of motion: $(\mathbf{q}^N, \mathbf{p}^N)$ evolve with time

How would this be done without having to integrate equations of motion?