# Brief Review of Statistical Mechanics

## Introduction

Statistical mechanics: "a branch of physics which studies macroscopic systems from a microscopic or molecular point of view" (McQuarrie,1976)
Also see (Hill,1986; Chandler, 1987)

Stat mech will inform us about

- how to set up and run a simulation algorithm
- how to estimate macroscopic properties of
   interest from simulations

Two important postulates

- in an isolated system with constant E, V, and N,
   all *microstates* are equally likely.
- time averages are equivalent to *ensemble*
   averages (principle of ergodicity).

The *ensemble* formalism allows extension of these postulates to more useful physical situations.

**Ensembles**

Ensemble: "the (virtual) assembly of all possible microstates (that are) consistent with the constraints with which we characterize the system macroscopically" (Chandler, 1987)

Microstate: the complete specification of a system at the most detailed level.

In the most rigorous sense, the microstate of a system is its quantum state, which is obtained by solving the Schrodinger equation. Quantum states are discrete and have discrete probabilities.

However, the microstate of a *classical* system is completely specified by the positions ($\mathbf{r}$) and momenta ($\mathbf{p}$) of all particles. Such microstates are part of a continuum and must be described with probability density functions.

Note: We will sometimes use quantum notation for compactness, but the focus of this part of the course is on classical systems.

Macroscopic constraints: generally, these are thermodynamic properties (energy, temperature, pressure, chemical potential, ...)

## Role of simulation

For a given model system, the tasks of a molecular simulation are to:

(1) Sample microstates within an ensemble, with the appropriate statistical weights

(2) During the sampling, calculate and collect molecular-level information that aids in understanding the physical behavior of the system

(3) Employ a large enough sample size to ensure that the collected information is meaningful
(Note: even with today's powerful computers, it is generally impossible to sample all of the microstates of a model system.)

## *Microcanonical ensemble.*

Constant E, V, N.



Walls are insulating
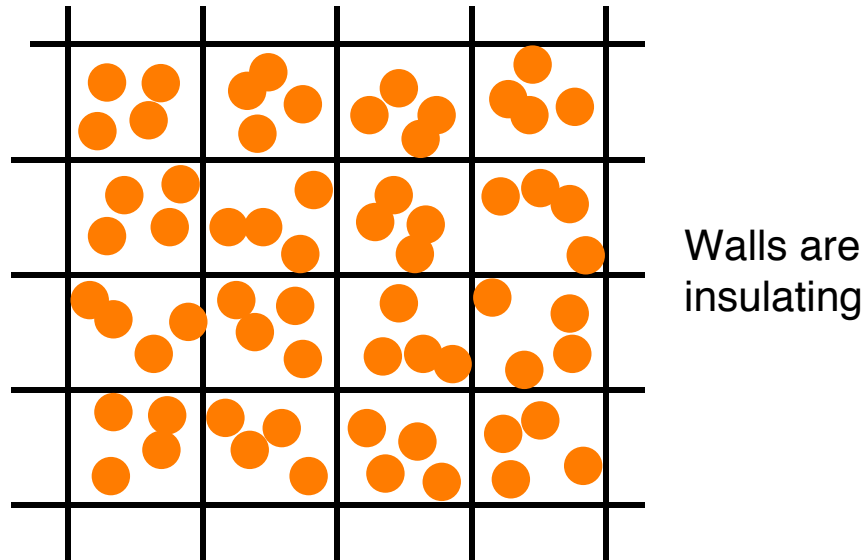
According to the first postulate, all of these states are equally likely.

$$\pi_i = \frac{1}{\Omega}$$

Properties that can be measured over a simulation:
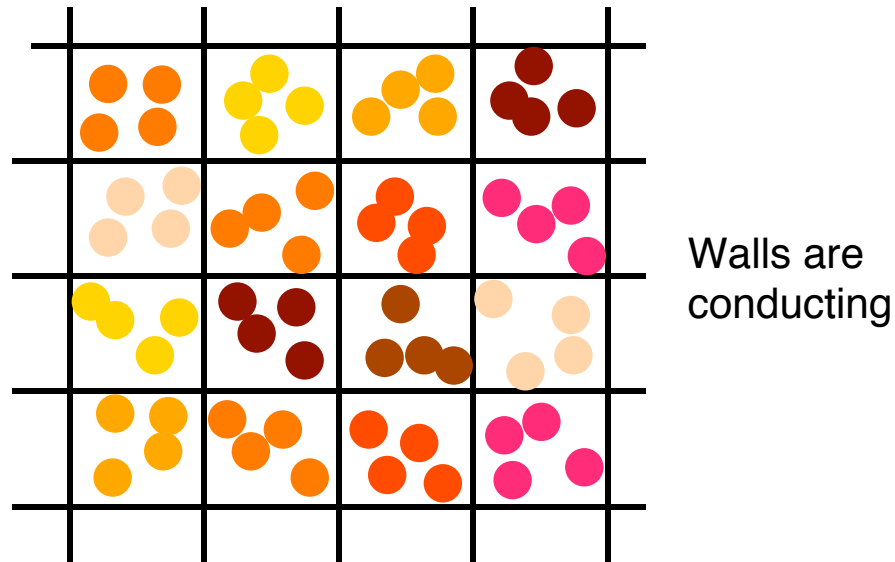  T, p.  And, with more effort, $\mu$.

Note: The conjugate variables are the ones that can be "measured" as outcomes.

$$dE = TdS - pdV + \mu dN$$

$$dS = \left(\frac{1}{T}\right)dE + \left(\frac{p}{T}\right)dV - \left(\frac{\mu}{T}\right)dN$$

## *Canonical ensemble.*

Constant T, V, N.



Walls are conducting

Now the boxes may have different energies!

Can show (Hill, 1986; McQuarrie, 1976) that the probability of a state is

$$\pi_i(E_i) = \frac{e^{-\beta E_i}}{Q}, \quad \text{where} \quad \beta = \frac{1}{k_B T}$$
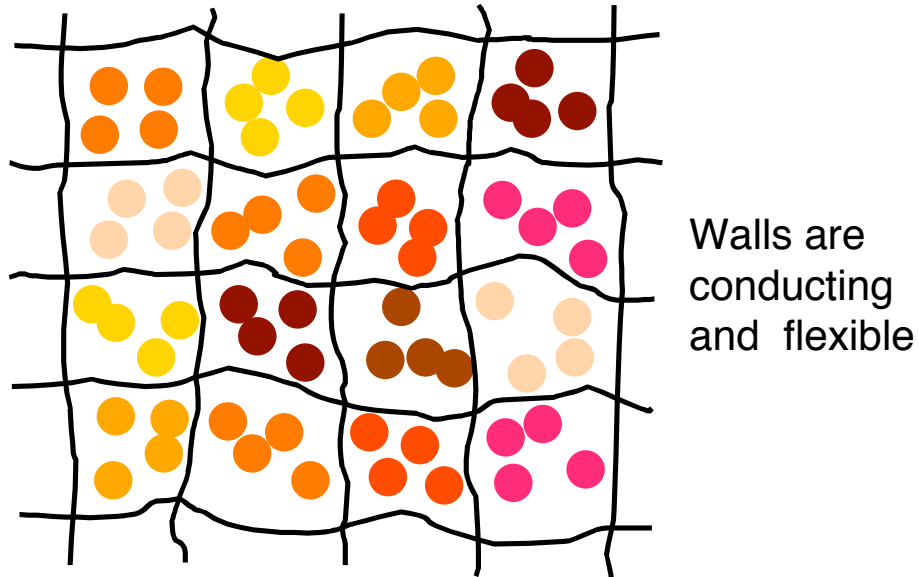
Properties that can be measured:
  E, p. And, with more effort, $\mu$.

The governing equation is

$$d\left(S - \frac{E}{T}\right) = -E\,d\left(\frac{1}{T}\right) + \left(\frac{p}{T}\right)dV - \left(\frac{\mu}{T}\right)dN$$

### *Isothermal-isobaric ensemble.*

Constant T, p, N.



Walls are
conducting
and  flexible

Now the boxes may have different energies and
volumes!

Can show (Hill,1986; McQuarrie, 1976) that the
probability of a state is

$$\pi_i(E_i, V_i) = \frac{e^{-\beta E_i}e^{-\beta p V_i}}{\Delta}$$

Properties that can be measured:
    E, V.  And, with more effort, $\mu$.

The governing equation is

$$d\left(S - \frac{E}{T} - \frac{pV}{T}\right) = -E\,d\left(\frac{1}{T}\right) - V\,d\left(\frac{p}{T}\right) - \left(\frac{\mu}{T}\right)dN$$

## *Grand canonical ensemble.*

Constant T, V, $\mu$.



Walls are conducting and permeable

Now the boxes may have different energies and numbers of molecules!

Can show (Hill,1986; McQuarrie, 1976) that the probability of a state is

$$\pi_i(E_i, N_i) = \frac{e^{-\beta E_i} e^{\beta \mu N_i}}{\Xi}$$
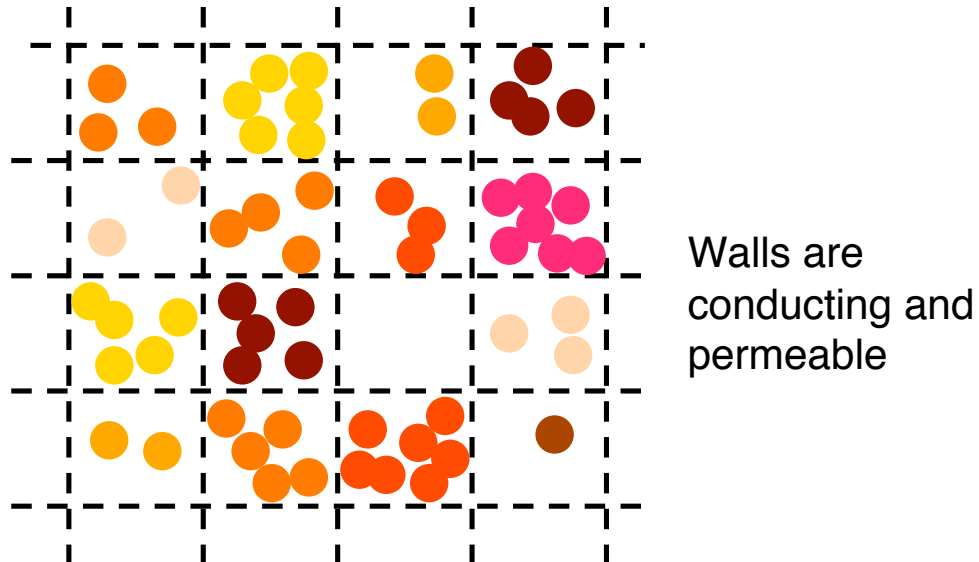
Properties that can be measured:
    E, p, and N.

The governing equation is

$$d\left(S - \frac{E}{T} + \frac{N\mu}{T}\right) = -Ed\left(\frac{1}{T}\right) + \left(\frac{p}{T}\right)dV + Nd\left(\frac{\mu}{T}\right)$$

# Bridges to macroscopic thermodynamics

*Partition functions.*

The normalization constants are referred to as partition functions, and they have (macroscopic) physical significance.

Summary, in "quantum" notation:

| Ensemble | Partition Function | Bridge Equation |
|---|---|---|
| Micro-canonical | $\Omega(E,V,N) = \sum_i 1$ | $\dfrac{S}{k_B} = \ln\Omega$ |
| Canonical | $Q(T,V,N) = \sum_i e^{-\beta E_i}$ | $-\beta A = \ln Q$ |
| Isothermal-Isobaric | $\Delta(T,p,N) = \sum_V e^{-\beta pV} \sum_i e^{-\beta E_i(V)}$ $= \sum_V e^{-\beta pV}\, Q(T,V,N)$ | $-\beta G = \ln\Delta$ |
| Grand canonical | $\Xi(T,V,\mu) = \sum_N e^{\beta\mu N} \sum_i e^{-\beta E_i(N)}$ $= \sum_N e^{\beta\mu N}\, Q(T,V,N)$ | $\beta pV = \ln\Xi$ |

We can translate these quantum formulas into quasi-classical expressions that are appropriate for our force-field-based molecular modeling approach.

Recall that the microstate of a classical system is not defined by a quantum state, but rather by the position and momentum of each particle. To get a partition function, we must sum (actually integrate, for the classical case) over all possible microstates.

Example: canonical partition function

$$Q = \frac{1}{N!}\frac{1}{h^{3N}}\int\ldots\int d\mathbf{p}_1\ldots d\mathbf{p}_N \; d\mathbf{r}_1\ldots d\mathbf{r}_N \; e^{-\beta H}$$

where

$$H\left(\mathbf{p}^{\{N\}},\mathbf{r}^{\{N\}}\right) = K\left(\mathbf{p}^{\{N\}}\right) + U\left(\mathbf{r}^{\{N\}}\right)$$

If we could evaluate this partition function for a model system at various temperatures and volumes, we could obtain thermodynamic properties via

$$-\beta A(\beta, V, N) = \ln Q$$

$$E = -\left.\frac{\partial(-\beta A)}{\partial \beta}\right|_{V,N}$$

$$\beta p = \left.\frac{\partial(-\beta A)}{\partial V}\right|_{T,N}$$

However, it is *not even close* to practical to calculate the partition function for a typical system.

Example:
Say we have a system with N=100 atoms, and we wish to evaluate the configurational (**r**) part of the canonical partition function. We have 3N=300 coordinates over which to integrate.

Try a Simpson's rule integration, where 10 function evaluations per coordinate are employed (a fairly coarse grid).

$\#$ of function evaluations $= (10)(10)L\ (10) = 10^{300}$

$$\left(10^{300}\,\text{operations}\right)\left(\frac{\text{sec}}{5\times 10^{12}\,\text{ops}}\right) \approx 10^{288}\,\text{sec}$$

$$\approx 10^{280}\,\text{yrs}$$

Not likely in the forseeable future!

Instead, we focus on obtaining averages of properties of interest.

### *Formalism for simple averages.*

One can easily define the average of any property that has a simple dependence on the microstate variables.

For example, in quantum notation

$$\langle B \rangle = \sum_i B_i \, \pi_i$$

In the canonical ensemble, for example, we have

$$\langle B \rangle = \frac{\sum_i B_i \, e^{-\beta E_i}}{Q}$$

In classical notation, a similar expression can be written for any quantity that has a simple, direct functional dependence on molecular positions and momenta.

$$\langle B \rangle = \frac{\dfrac{1}{N!}\dfrac{1}{h^{3N}} \int \cdots \int d\mathbf{p}_1 \ldots d\mathbf{p}_N \, d\mathbf{r}_1 \ldots d\mathbf{r}_N \, B\!\left(\mathbf{p}^{\{N\}}, \mathbf{r}^{\{N\}}\right) e^{-\beta H}}{Q}$$

This seems no better than before – we can't calculate the numerator or denominator!

But…

### *Obtaining simple averages from simulation.*

Recall that it is the task of a simulation algorithm to generate microstates in an ensemble, in proportion to their statistical weights (probabilities).

This can be done
- probabilistically (Monte Carlo)
- deterministically (molecular dynamics)[*]

[*]actually a time average – we'll revisit this issue

Either way, given a set of $n_{obs}$ representative microstates, we can write the average of B as

$$\langle B \rangle = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} B\left(\mathbf{p}_i^{\{N\}}, \mathbf{r}_i^{\{N\}}\right)$$

*as long as those microstates were generated according to the appropriate weighting function $\pi$ for that ensemble!*

The quality of the result that you obtain depends on $n_{obs}$.

## *Separation of the energy.*

In the classical limit, the kinetic energy (K) and potential energy (U) can be considered as separable in the Hamiltonian.

This allows Q to be factored

$$Q = \frac{1}{N!}\frac{1}{h^{3N}}\left(\int...\int d\mathbf{p}_1...d\mathbf{p}_N \ e^{-\beta K}\right)\left(\int...\int d\mathbf{r}_1...d\mathbf{r}_N \ e^{-\beta U}\right)$$

and the momentum part will drop out of averages that depend only on positions.

$$\langle B \rangle = \frac{\dfrac{1}{N!}\dfrac{1}{h^{3N}}\left(\int...\int d\mathbf{p}_1...d\mathbf{p}_N \ e^{-\beta K}\right)\left(\int...\int d\mathbf{r}_1...d\mathbf{r}_N \ B\left(\mathbf{r}^{\{N\}}\right)e^{-\beta U}\right)}{\dfrac{1}{N!}\dfrac{1}{h^{3N}}\left(\int...\int d\mathbf{p}_1...d\mathbf{p}_N \ e^{-\beta K}\right)\left(\int...\int d\mathbf{r}_1...d\mathbf{r}_N \ e^{-\beta U}\right)}$$

$$= \frac{\int...\int d\mathbf{r}_1...d\mathbf{r}_N \ B\left(\mathbf{r}^{\{N\}}\right)e^{-\beta U}}{\int...\int d\mathbf{r}_1...d\mathbf{r}_N \ e^{-\beta U}}$$

This is why Monte Carlo, where there is no time, can still be useful in obtaining thermodynamic properties.

### *Examples of simple averages.*

Kinetic energy.

The classical definition is

$$K = \sum_{i=1}^{N} \frac{1}{2m_i} |\mathbf{p}_i|^2$$

So we have

$$\langle K \rangle = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} K\left(\mathbf{p}_i^{\{N\}}\right) = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} \left( \sum_{i=1}^{N} \frac{1}{2m_i} |\mathbf{p}_i|^2 \right)$$

Temperature.

Start with the rigorous definition of $\langle K \rangle$

$$\langle K \rangle = \frac{\dfrac{1}{N!}\dfrac{1}{h^{3N}}\left(\int \cdots \int d\mathbf{p}_1 \ldots d\mathbf{p}_N \, K(\mathbf{p}^{\{N\}})e^{-\beta K}\right)\left(\int \cdots \int d\mathbf{r}_1 \ldots d\mathbf{r}_N \, e^{-\beta U}\right)}{\dfrac{1}{N!}\dfrac{1}{h^{3N}}\left(\int \cdots \int d\mathbf{p}_1 \ldots d\mathbf{p}_N \, e^{-\beta K}\right)\left(\int \cdots \int d\mathbf{r}_1 \ldots d\mathbf{r}_N \, e^{-\beta U}\right)}$$

$$= \frac{\int \cdots \int d\mathbf{p}_1 \ldots d\mathbf{p}_N \, K(\mathbf{p}^{\{N\}})e^{-\beta K}}{\int \cdots \int d\mathbf{p}_1 \ldots d\mathbf{p}_N \, e^{-\beta K}}$$

Can show that for $K = \displaystyle\sum_{i=1}^{N}\frac{1}{2m_i}|\mathbf{p}_i|^2$ , this reduces to

$$\langle K \rangle = \frac{3}{2}Nk_BT$$

*The equipartition principle*: any contribution to the Hamiltonian that is additive and quadratic will contribute ($k_BT/2$) to the energy. (McQuarrie,1976 pp. 121-123)

So the average temperature is

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

Potential energy.

This is simply the interaction potential within and between molecules (intramolecular and intermolecular, respectively). It depends only on the atomic coordinates, not the momenta.

So we may write

$$\langle U \rangle = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} U\left(\mathbf{r}_i^{\{N\}}\right)$$

We may write a similar relationship for any particular component of the potential energy. For example, for a polymeric system we may be interested in the torsional component of the potential:

$$\langle U_{tors} \rangle = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} U_{tors}\left(\mathbf{r}_i^{\{N\}}\right)$$

Pressure.

Pressure can be obtained from the *virial formula* (Allen and Tildesley,1987)

$$PV = Nk_BT + \langle W \rangle$$

where W is the *virial*

$$W = \frac{1}{3} \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{f}_i$$

Usually this is reduced into a form that is independent of origin before application. For example, for pairwise forces, the virial takes the form

$$W = \frac{1}{3} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \equiv \frac{1}{3} \sum_{i} \sum_{j>i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij}$$

Three-body and higher forces are more complicated, but similar formulas can be derived.

Chemical potential.

Chemical potential is not a "mechanical variable" – it does not have a "simple, direct functional dependence on molecular positions and momenta."

Calculating the chemical potential of a species from a molecular simulation can be quite challenging; we will discuss this in detail later.

For now, the following definition (in the canonical ensemble) provides a clue on how it might be calculated

$$\beta\mu = -\frac{\partial(-\beta A)}{\partial N}\bigg|_{\beta,V}$$

**Ergodicity**


***Ergodic hypothesis.***

"…a large number of observations made on a single system at $n$ arbitrary instants of time have the same statistical properties as observing $n$ arbitrarily chosen systems at the same time from an ensemble…" (McQuarrie,1976)


time average ⇔ ensemble average


We believe that it "holds for all many-body systems in nature." (Chandler,1987)


***Implications for molecular simulation.***
(Allen and Tildesley,1987; Frenkel and Smit,1996)

molecular dynamics
    simulation in time
    average over snapshots in time

Monte Carlo
    probabilistic walk through configurations
    average over (properly weighted) configurations

When we simulate a given model system under a given set of macroscopic constraints, we should get statistically identical results from MD and MC.

# *Ergodic problems in molecular simulation.*

It is possible for your simulation to become effectively "stuck" in certain regions of phase space ($\mathbf{r}^{\{N\}}, \mathbf{p}^{\{N\}}$) and not explore other regions of importance.
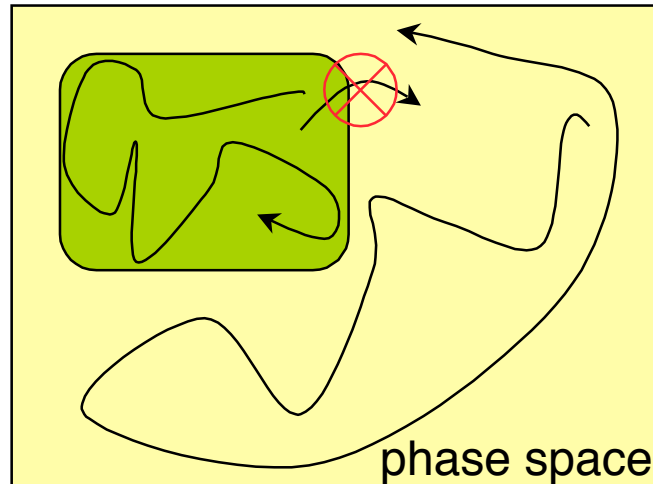


phase space

(Figure by D.A. Kofke)

This can happen because of
- effective (kinetic) barriers
- true disconnects in phase space of model

When it happens
- the simulated system is effectively (or truly) *non-ergodic*
- you may have *phase space sampling problems*
    - fail to achieve a representative sampling of microstates
    - calculated averages are not true averages of your model system.

Phase-space sampling problems can occur with both MD and MC algorithms.  There are several ways to look for, and avoid, such problems.

Molecular dynamics
      - increase simulation time
      - vary starting conditions
      - use annealing

Monte Carlo
      - same remedies as MD, plus…
      - use special (bolder) algorithms

## Role of simulation, re-visited

For a given model system, the tasks of a molecular simulation are to:

(1) Sample microstates within an ensemble, with the appropriate statistical weights

(2) During the sampling, calculate and collect molecular-level information that aids in understanding the physical behavior of the system

(3) Employ a large enough sample size to ensure that the collected information is meaningful
(Note: even with today's powerful computers, it is generally impossible to sample all of the microstates of a model system.)