

Molecular Simulation

Background

Why Simulation?

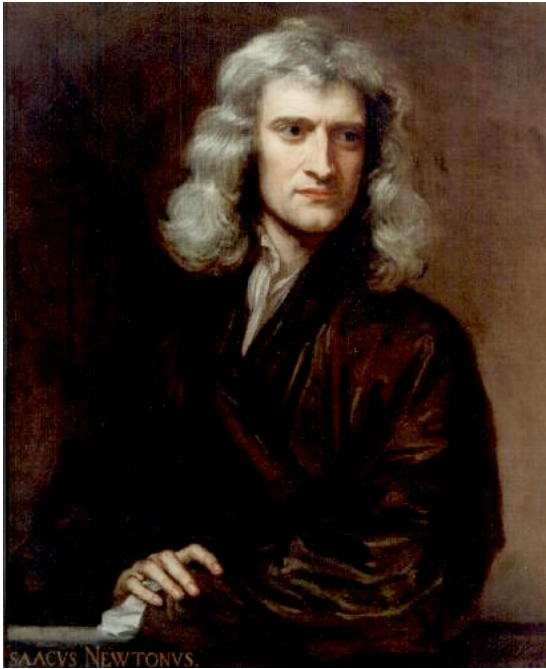
- 1. Predicting** properties of (new) materials
- 2. Understanding** phenomena on a molecular scale.
- ~~**3. Reproducing** known phenomena.~~

Testing simulations by:

Reproducing known phenomena.

THE question:

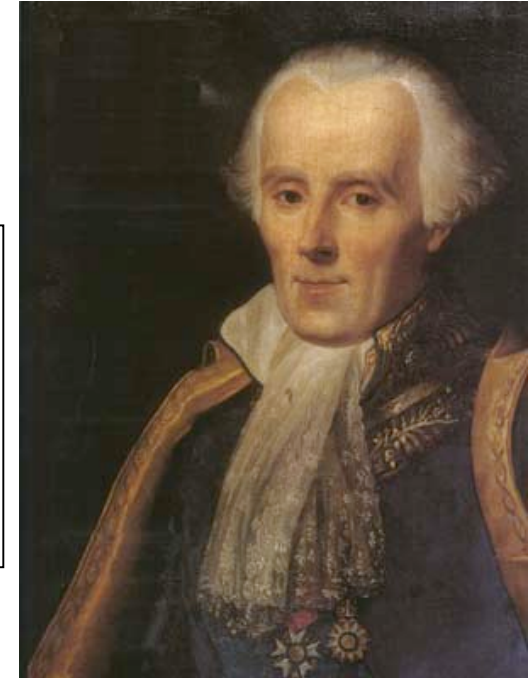
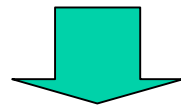
**“Can we predict the
macroscopic properties of
(classical) many-body
systems?”**



NEWTON: $F = m a$

LAPLACE:

Nous devons donc envisager l'état présent de l'univers comme l'effet de son état antérieur et comme la cause de celui qui va suivre. Une intelligence qui, pour un instant donné, connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'Analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux.



“*Translation*” In principle “Yes”.

Provided that we know the position, velocity and interaction of all molecules, then the future behavior is predictable, ... **BUT**

... There are so many molecules.

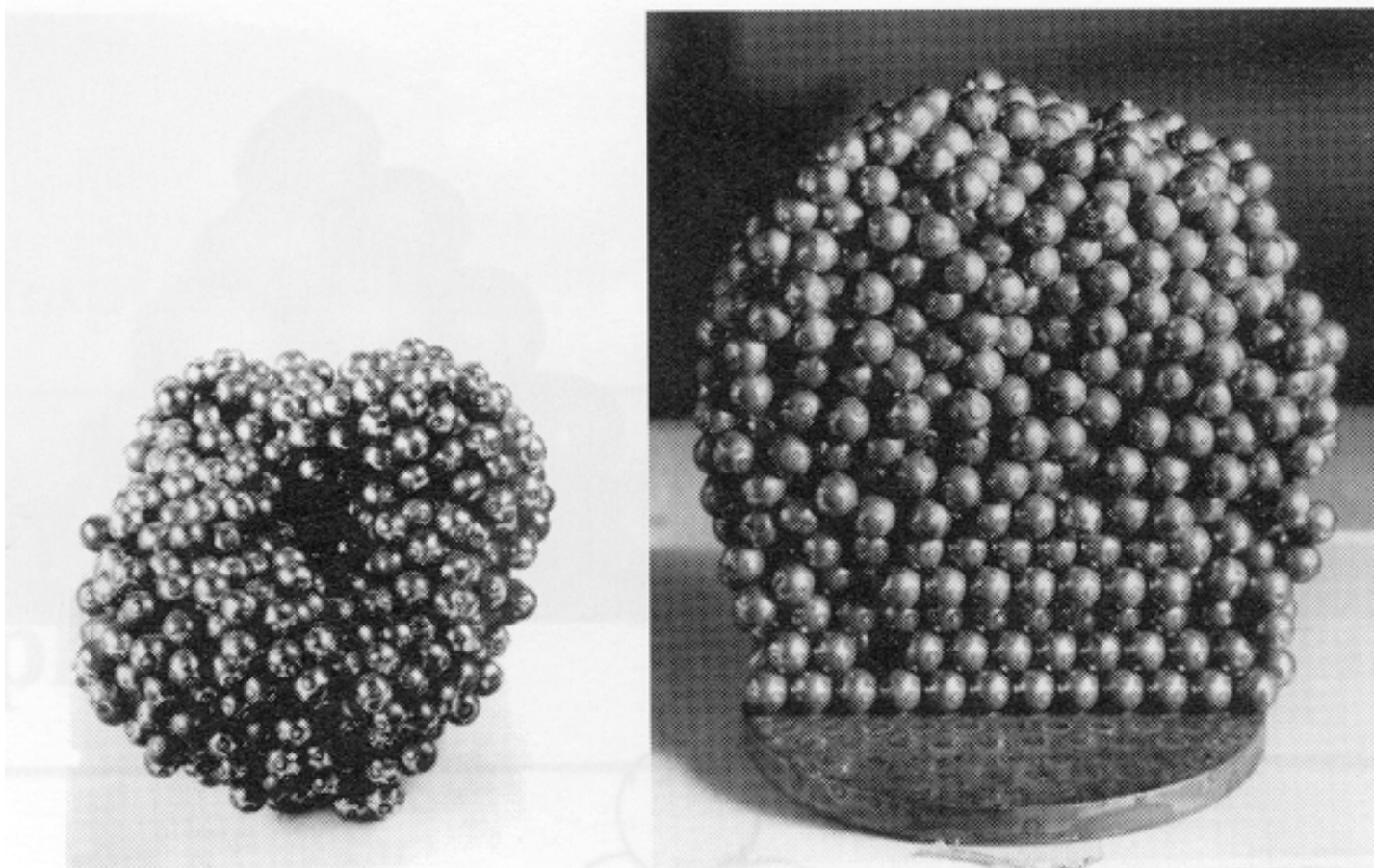
This is why, before the advent of the computer, it was impossible to predict the properties of real materials.

What was the alternative?

1. Smart tricks (“theory”)

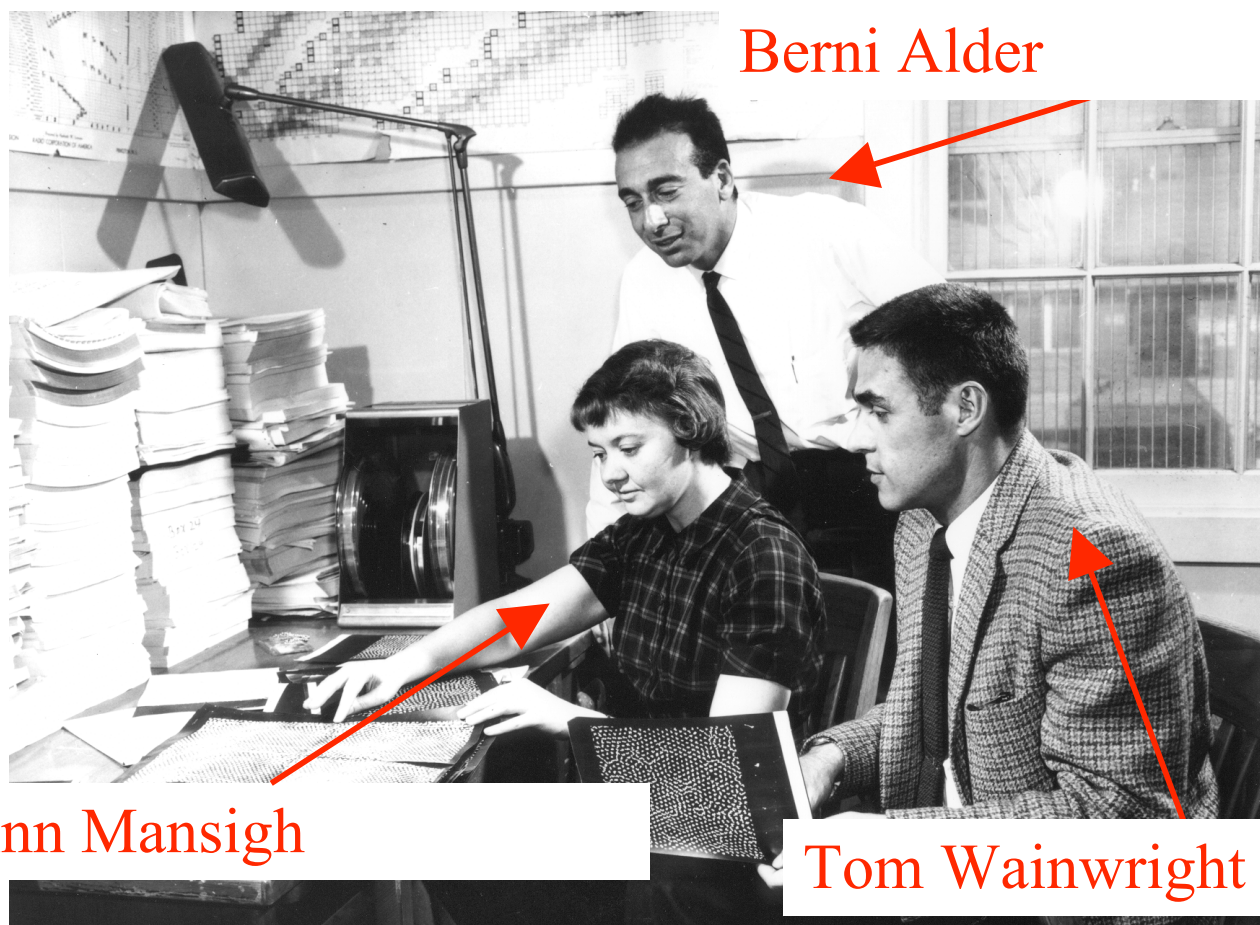
Only works in special cases

2. Constructing model (“molecular lego”)...



J.D. Bernal's "ball-bearing model"
of an atomic liquid...

The computer age (1953...)



With computers we can follow the behavior of **hundreds** to **hundreds of millions** of molecules.

The limits of Simulation

Brute-force simulations can never bridge all the scales between microscopic (nanometers/picoseconds) and macroscopic (cells, humans, planets).

Hence: we need different levels of description (“coarse graining”) - and we need input from experiments at many different levels to validate our models.

The limits of Experiments

Increasingly, experiments generate far more data than humans can digest.

Result: “Experulation”.

Simulations are becoming an integral part of the analysis of experimental data.

Intermezzo:

Essential Statistical Mechanics

A brief summary of:

**Entropy, temperature, Boltzmann distributions
and the Second Law of Thermodynamics**

The basics:

1. Nature is quantum-mechanical

2. Consequence:

Systems have discrete quantum states.

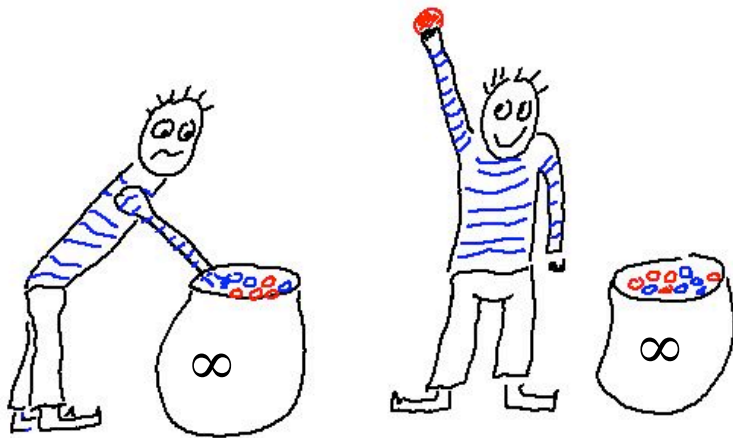
For finite “closed” systems, the number of states is finite (but usually very large)

3. Hypothesis: **In a closed system, every state is equally likely to be observed.**

4. Consequence: **ALL of equilibrium
Statistical Mechanics and Thermodynamics**

First: **Simpler example (standard statistics)**

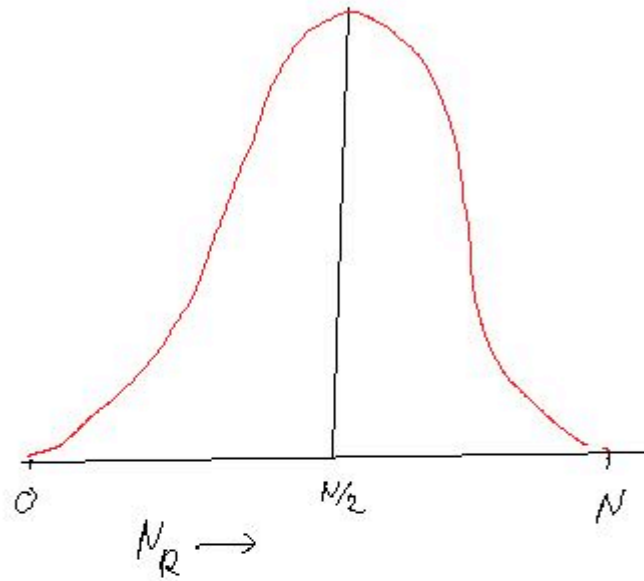
Draw N balls from an infinite vessel that contains an equal number of **red** and **blue** balls



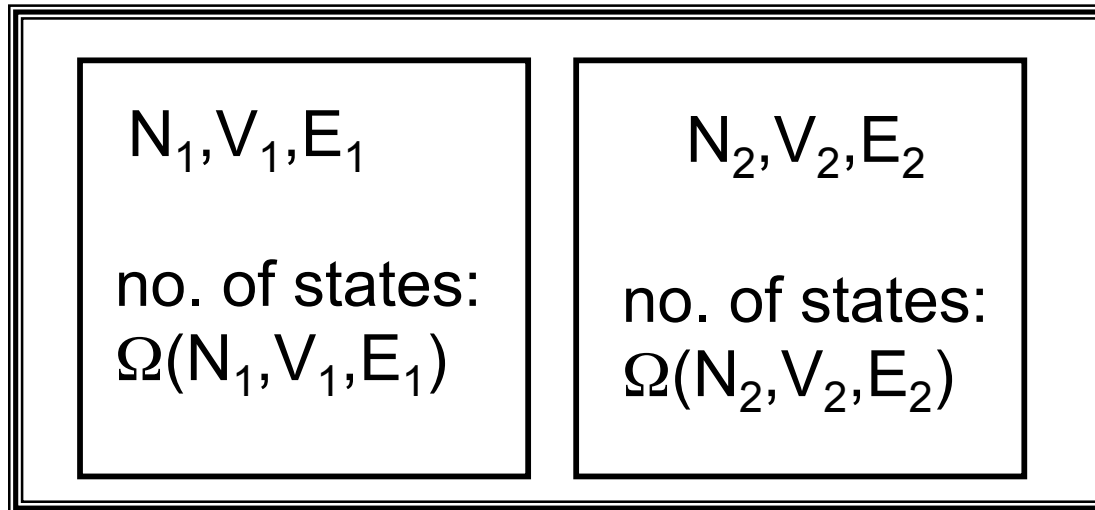
Number of possibilities to draw N_R red balls and N_B blue balls: $\Omega(N_R, N_B) = \frac{N!}{N_R!N_B!}$.

Most likely $N_R = N_B = N/2$.

$$\Omega(N/2, N/2) = \frac{N!}{(N/2)!(N/2)!}$$

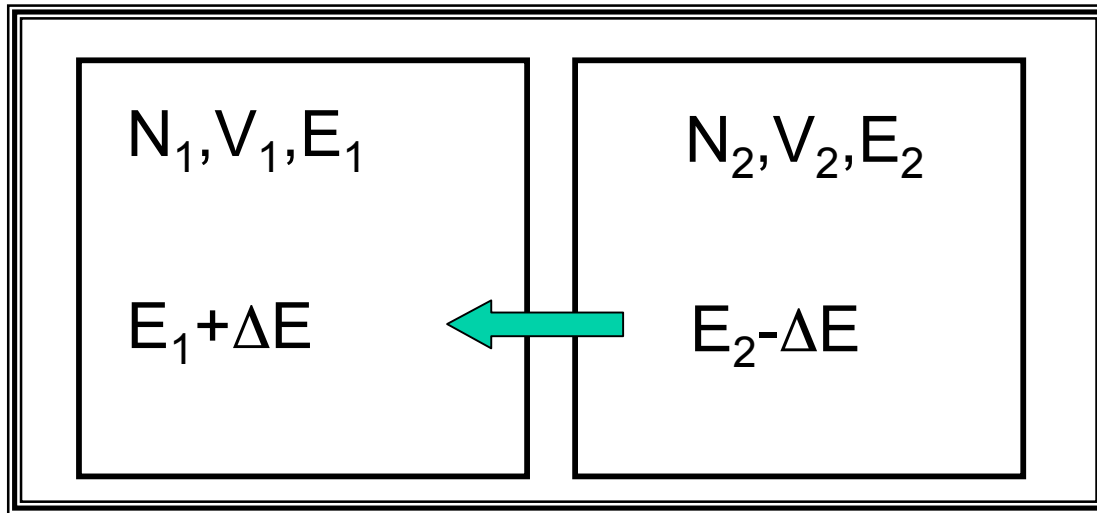


Now consider two systems with total energy E .



$$\Omega(E_1, E_2) = \Omega_1(N_1, V_1, E_1) \times \Omega_2(N_2, V_2, E_2)$$

This function is **very** sharply peaked (for macroscopic systems)



Now, allow energy exchange between **1** and **2**.

The most likely energy distribution is the one that maximizes $\Omega_1 \times \Omega_2$.

It is more convenient (but equivalent) to maximize $\ln(\Omega_1 \times \Omega_2)$.

So:

$$\left(\frac{\partial \ln \Omega(E_1, E - E_1)}{\partial E_1} \right)_{N, V, E} = 0$$

With:

$$\ln \Omega(E_1, E - E_1) = \ln \Omega_1(E_1) + \ln \Omega_2(E - E_1)$$

$$\left(\frac{\partial \ln \Omega_1(E_1)}{\partial E_1}\right)_{N_1, V_1} + \left(\frac{\partial \ln \Omega_2(E_2)}{\partial E_1}\right)_{N_2, V_2} = 0$$

but $dE_2 = -dE_1$, hence:

$$\left(\frac{\partial \ln \Omega_1(E_1)}{\partial E_1}\right)_{N_1, V_1} = \left(\frac{\partial \ln \Omega_2(E_2)}{\partial E_2}\right)_{N_2, V_2}$$

This is the condition for thermal equilibrium (“no spontaneous heat flow between 1 and 2”)

Normally, thermal equilibrium means: equal temperatures...

Let us define:

$$\beta(E, V, N) \equiv \left(\frac{\partial \ln \Omega(E, V, N)}{\partial E} \right)_{N, V}$$

Then, thermal equilibrium is equivalent to:

$$\beta(E_1, V_1, N_1) = \beta(E_2, V_2, N_2)$$

This suggests that β is a function of T.

Relation to classical thermodynamics:

We know that, in equilibrium,
 $\ln \Omega$ is a maximum.

Conjecture: $\ln \Omega = S$

Almost right.

Good features:

- Extensivity
- Third law of thermodynamics comes for free

Bad feature:

- It assumes that entropy is dimensionless but (for unfortunate, historical reasons, it is not...)

We have to live with the past, therefore

$$S = k_B \ln \Omega$$

With $k_B = 1.380662 \cdot 10^{-23}$ J/K

In thermodynamics, the absolute (Kelvin) temperature scale was defined such that

$$\frac{\partial S}{\partial E} = \frac{1}{T} \quad \text{But we found (defined):}$$

$$\beta(E, V, N) \equiv \left(\frac{\partial \ln \Omega(E, V, N)}{\partial E} \right)_{N, V}$$

And this gives the “statistical” definition of temperature:

$$\frac{1}{T} = k_B \left(\frac{\partial \ln \Omega(E, V, N)}{\partial E} \right)_{N, V}$$

In short:

Entropy and temperature are both related to the fact that we can COUNT states.

How large is Ω ?

For macroscopic systems, **super-astronomically** large.

For instance, for a glass of water at room temperature:

$$\Omega \approx 10^{23} 10^{25}$$

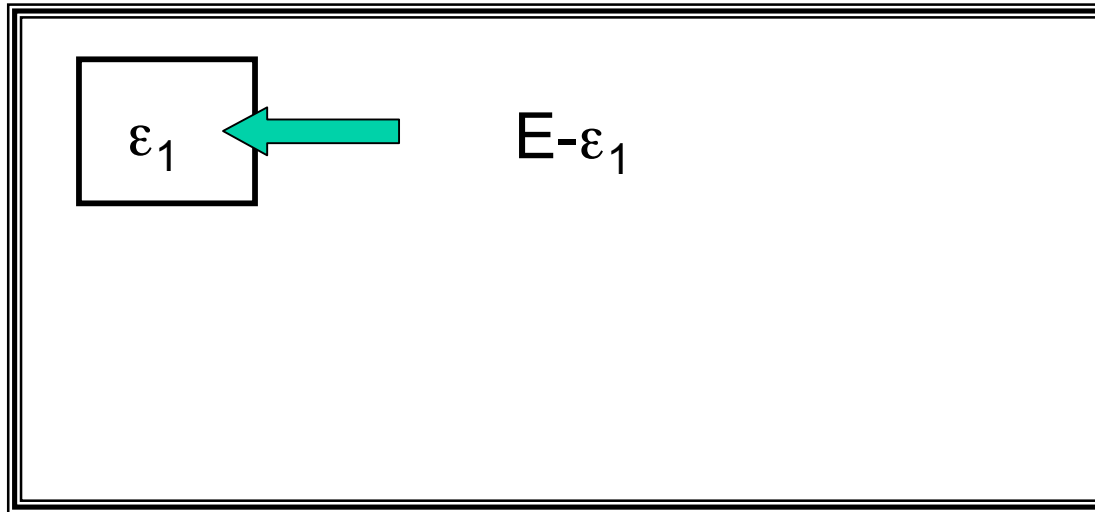
Macroscopic deviations from the second law of thermodynamics are not forbidden, but they are extremely unlikely.

A mathematical relation:

$$10^{-10^{23}} \neq 0$$

A physical relation:

$$10^{-10^{23}} = 0$$



Consider a “small” system (a molecule, a virus, a mountain) in thermal contact with a much larger system (“bath”).

The total energy is fixed. The higher the energy of the small system, the lower the energy of the bath.

What happens to the total number of accessible states?

$$\ln \Omega_B(E - \epsilon_i) = \ln \Omega_B(E) - \epsilon_i \frac{\partial \ln \Omega_B(E)}{\partial E} + \mathcal{O}(1/E)$$

But, as $\beta=1/k_B T$:

$$\ln \Omega_B(E - \epsilon_i) = \ln \Omega_B(E) - \epsilon_i/k_B T + \mathcal{O}(1/E)$$

The probability that the small system is in a **given** (“labeled”) state with energy ϵ_i is

$$P(\epsilon_i) = \frac{\Omega_B(E - \epsilon_i)}{\sum_i \Omega_B(E - \epsilon_i)}$$

using

$$\Omega_B(E - \epsilon_i) = \Omega_B(E) \exp(-\epsilon_i/k_B T)$$

we get

$$P(\epsilon_i) = \frac{\exp(-\epsilon_i/k_B T)}{\sum_i \exp(-\epsilon_i/k_B T)}$$

This is the Boltzmann distribution:

“Low energies are more likely than high energies”

The probability to find the system in state I is:

$$p_i = \frac{\exp(-\beta\epsilon_i)}{Q}$$

$$\beta \equiv 1/k_B T$$

Hence, the average energy is

$$\langle E \rangle = \sum_i p_i \epsilon_i = \sum_i \frac{\epsilon_i \exp(-\beta\epsilon_i)}{Q}$$

Therefore

$$\langle E \rangle = - \frac{\partial \ln Q}{\partial \beta}$$

This can be compared to the thermodynamic relation

$$E = \left(\frac{\partial F/T}{\partial 1/T} \right)$$

This suggests that the partition sum

$$Q = \sum_i \exp(-\beta \epsilon_i)$$

is related to the Helmholtz free energy through

$$F = -k_B T \ln Q$$

How?

We have assumed quantum mechanics (wavefunctions), but often we are interested in the classical limit

$$\sum_i e^{-\beta E_i} \rightarrow \frac{1}{h^{3N} N!} \int \int d\mathbf{p}^N d\mathbf{r}^N e^{-\beta \left[\sum_i \frac{p_i^2}{2m_i} + U(\mathbf{r}^N) \right]}$$

$\frac{1}{h^{3N}}$ \rightarrow Volume of phase space

$\frac{1}{N!}$ \rightarrow Particles are indistinguishable

Integration over the momenta can be carried out for most systems:

$$\int d\mathbf{p}^N \exp \left\{ -\beta \left[\sum_i \frac{p_i^2}{2m_i} \right] \right\} = \left[\int dp \exp \left\{ -\beta \frac{p^2}{2m} \right\} \right]^{3N} = \left(\frac{2\pi m}{\beta} \right)^{\frac{3}{2}N}$$

Steps:

1. We can introduce functions of operators. E.g.

$$\langle i | e^A | i \rangle = \sum_n \langle i | \frac{A^n}{n!} | i \rangle$$

Example. For an energy eigenstate i :

$$\langle i | e^{-\beta H} | i \rangle = \sum_n \langle i | \frac{(-\beta H)^n}{n!} | i \rangle = \sum_n \frac{(-\beta E_i)^n}{n!} = e^{-\beta E_i}$$

2. The “trace” of a quantum mechanical operator does not depend on the basis set

$$\sum_{\text{all } i} \langle i|A|i \rangle = \sum_{\text{all } j} \langle j|A|j \rangle$$

Example:

$$\sum_{\text{eigenstates } i} \langle i|e^{-\beta H}|i \rangle = \sum_{\text{other states } j} \langle j|e^{-\beta H}|j \rangle$$

$$Q = \sum_{\text{eigenstates } i} e^{-\beta E_i} = \sum_{\text{other states } j} \langle j|e^{-\beta H}|j \rangle$$

3. We can write the unit operator I as

$$I = \sum_{\text{complete set of states } j} |j\rangle\langle j|$$

Example:

$$I = \sum_{\text{all wavevectors } q} |q\rangle\langle q|$$

The momentum \mathbf{p} is related to the wavevector \mathbf{q} , via

$$\mathbf{p} = \hbar\mathbf{q} = \hbar\{n_x, n_y, n_z\} \frac{2\pi}{L}$$

Now apply to partition function:

$$Q = \sum_{\text{eigenstates } i} e^{-\beta E_i} = \sum_{\text{other states } j} \langle j | e^{-\beta H} | j \rangle$$

Choose for j the position eigenstates r :

$$Q = \frac{1}{N!} \sum_r \langle r | e^{-\beta H} | r \rangle$$

This we cannot compute.

Now write: $H = K + U$, then

$$Q = \frac{1}{N!} \sum_r \langle r | e^{-\beta(K+U)} | r \rangle$$

This, we can still not compute. What we *can* compute is:

$$\sum_q \langle q | e^{-\beta K} | q \rangle \approx \frac{V^N}{h^{3N}} \int dp^N e^{-\beta K}$$

or (possibly):

$$Z = \frac{1}{N!} \sum_r \langle r | e^{-\beta(U)} | r \rangle = \frac{1}{N!} \int dr^N e^{-\beta U(r^N)}$$

We could use these simplifications if:

$$e^{-\beta(K+U)} \stackrel{?}{=} e^{-\beta K} e^{-\beta U}$$

But for non-commuting operators A and B,

$$e^{A+B} \neq e^A e^B$$

However: in the classical limit, we can write

$$e^{-\beta(K+U)} \approx e^{-\beta K} e^{-\beta U}$$

In the classical limit we get:

$$Q = \frac{1}{N!} \sum_r \langle r | e^{-\beta(K+U)} | r \rangle \approx$$
$$\frac{1}{N!} \sum \langle r | e^{-\beta U} | r \rangle \langle r | q \rangle \langle q | e^{-\beta K} | q' \rangle \langle q' | r \rangle$$
$$\frac{1}{N!} \int dr^N e^{-\beta U(r^N)} \frac{V^N}{h^{3N}} \int dp^N e^{-\beta K} \times \frac{1}{V^N}$$

Or...

$$Q_{\text{classical}} = \frac{1}{h^{3N} N!} \int dr^N dp^N e^{-\beta[U(r^N) + K(p^N)]}$$

Remarks

Define *de Broglie* thermal wave length:

$$\Lambda \equiv \left(\frac{h^2 \beta}{2\pi m} \right)^{1/2}$$

Partition function:

$$Q(N, V, T) = \frac{1}{\Lambda^{3N} N!} \int d\mathbf{r}^N \exp \left[-\beta U \left(\mathbf{r}^N \right) \right]$$

Check: ideal gas

$$Q(N, V, T) = \frac{1}{\Lambda^{3N} N!} \int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)]$$
$$= \frac{1}{\Lambda^{3N} N!} \int d\mathbf{r}^N 1 = \frac{V^N}{\Lambda^{3N} N!}$$

Free energy:

$$\beta F = -\ln\left(\frac{V^N}{\Lambda^{3N} N!}\right) \approx N(\ln \rho \Lambda^3 - 1)$$

Pressure:

$$P = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{N}{\beta V}$$

Energy:

$$E = \left(\frac{\partial \beta F}{\partial \beta}\right) = \frac{3N}{\Lambda} \frac{\partial \Lambda}{\partial \beta} = \frac{3}{2} N k_B T$$

Relating macroscopic observables to
microscopic quantities

Example:

Heat capacity

Pressure

Diffusion coefficient

Fluctuation expression for heat capacity.

Recall:

$$F = -k_B T \ln Q$$

with

$$Q = \sum_i \exp(-\beta \epsilon_i)$$

Then the heat capacity is

$$C_V = \frac{\partial E}{\partial T} = \left(\frac{\partial E}{\partial \beta} \right) \left(\frac{\partial \beta}{\partial T} \right) = - \frac{1}{k_B T^2} \frac{\partial E}{\partial \beta}$$

Using our expression for **E**:

$$C_V = - \frac{1}{k_B T^2} \frac{\partial}{\partial \beta} \sum_i \frac{\epsilon_i \exp(-\beta \epsilon_i)}{Q}$$

Both the numerator and denominator depend on β .

$$C_V = + \frac{1}{k_B T^2} \left(\sum_i \frac{\epsilon_i^2 \exp(-\beta \epsilon_i)}{Q} - \left[\sum_i \frac{\epsilon_i \exp(-\beta \epsilon_i)}{Q} \right]^2 \right)$$

And, finally:

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

COMPUTING THE PRESSURE:

$$F = -kT \ln Q$$

$$Q = \frac{1}{h^{3N} N!} \int d\mathbf{r}^N \exp[-\mathcal{U}(\mathbf{r}^N)/k_B T]$$

$$\frac{\partial F}{\partial V} = -P$$

Introduce “scaled” coordinates:

$$\mathbf{r} \equiv L\mathbf{s}$$

$$Q(N, V, T) = \frac{V^N}{\Lambda^{3N} N!} \int d\mathbf{s}^N \exp[-\beta\mathcal{U}(\mathbf{s}^N; L)]$$

$$\frac{\partial F}{\partial V} = -P$$

$$\frac{\partial F}{\partial V} = k_B T \frac{\partial \ln Q}{\partial V}$$

$$Q(N, V, T) = \frac{V^N}{\Lambda^{3N} N!} \int ds^N \exp[-\beta \mathcal{U}(\mathbf{s}^N; L)]$$

$$P = k_B T \frac{\partial \ln V^N \int ds^N \exp[-\beta \mathcal{U}(\mathbf{s}^N; L)]}{\partial V}$$

$$P = \frac{Nk_B T}{V} + k_B T \frac{\partial \ln \int ds^N \exp[-\beta \mathcal{U}(\mathbf{s}^N; L)]}{\partial V}$$

$$\frac{\partial \mathcal{U}(\mathbf{s}^N; L)}{\partial V} = \sum_{i=1}^N \frac{\partial \mathcal{U}(\mathbf{r}^N)}{\partial \mathbf{r}_i} \frac{\partial \mathbf{r}_i}{\partial V}$$

$$\frac{\partial \mathbf{r}_i}{\partial V} = \frac{1}{3L^2} \frac{\partial L \mathbf{s}_i}{\partial L} = \frac{1}{3L^2} \mathbf{s}_i$$

$$\frac{\partial \mathcal{U}(\mathbf{s}^N; L)}{\partial V} = \sum_{i=1}^N \frac{\partial \mathcal{U}(\mathbf{r}^N)}{\partial \mathbf{r}_i} \cdot \frac{\mathbf{r}_i}{3V}$$

$$P = \frac{Nk_B T}{V} + k_B T \frac{\partial \ln \int d\mathbf{s}^N \exp[-\beta \mathcal{U}(\mathbf{s}^N; L)]}{\partial V}$$

$$P = \frac{Nk_B T}{V}$$

$$\frac{\int d\mathbf{s}^N \sum_{i=1}^N \frac{\partial \mathcal{U}(\mathbf{r}^N)}{\partial \mathbf{r}_i} \cdot \frac{\mathbf{r}_i}{3V} \exp[-\beta \mathcal{U}(\mathbf{s}^N; L)]}{\int d\mathbf{s}^N \exp[-\beta \mathcal{U}(\mathbf{s}^N; L)]}$$

$$P = \frac{Nk_B T}{V} - \left\langle \sum_{i=1}^N \frac{\partial \mathcal{U}(\mathbf{r}^N)}{\partial \mathbf{r}_i} \cdot \frac{\mathbf{r}_i}{3V} \right\rangle$$

$$P = \frac{Nk_B T}{V} + \frac{1}{3V} \left\langle \sum_{i=1}^N \mathbf{f}_i \cdot \mathbf{r}_i \right\rangle$$

$$P = \frac{Nk_B T}{V} + \frac{1}{3V} \left\langle \sum_{i=1}^N \mathbf{f}_i \cdot \mathbf{r}_i \right\rangle$$

For pairwise additive forces:

$$\mathbf{f}_i = \sum_{j \neq i} \mathbf{f}_{ij}$$

Then

$$P = \frac{Nk_B T}{V} + \frac{1}{3V} \left\langle \sum_{i,j=1, i \neq j}^N \mathbf{f}_{ij} \cdot \mathbf{r}_i \right\rangle$$

$$P = \frac{Nk_B T}{V} + \frac{1}{3V} \left\langle \sum_{i,j=1, i \neq j}^N \mathbf{f}_{ij} \cdot \mathbf{r}_i \right\rangle$$

i and j are dummy variable hence:

$$\sum_{i,j=1, i \neq j}^N \mathbf{f}_{ij} \cdot \mathbf{r}_i = \sum_{j,i=1, j \neq i}^N \mathbf{f}_{ji} \cdot \mathbf{r}_j$$

And we can write

$$\sum_{i,j=1, i \neq j}^N \mathbf{f}_{ij} \cdot \mathbf{r}_i = \frac{1}{2} \sum_{j,i=1, j \neq i}^N (\mathbf{f}_{ij} \cdot \mathbf{r}_i + \mathbf{f}_{ji} \cdot \mathbf{r}_j)$$

But as action equals reaction (Newton's 3rd law):

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji}$$

And hence

$$\sum_{j,i=1,j \neq i}^N (\mathbf{f}_{ij} \cdot \mathbf{r}_i + \mathbf{f}_{ji} \cdot \mathbf{r}_j) = \sum_{j,i=1,j \neq i}^N \mathbf{f}_{ij} \cdot (\mathbf{r}_i - \mathbf{r}_j)$$

Inserting this in our expression for the pressure, we get:

$$P = \frac{Nk_B T}{V} + \frac{1}{6V} \left\langle \sum_{i,j=1,i \neq j}^N \mathbf{f}_{ij} \cdot \mathbf{r}_{ij} \right\rangle$$

Where $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$

What to do if you cannot use the virial expression?

$$P = -\frac{\partial F}{\partial V} \approx \frac{F(V - \Delta V) - F(V)}{\Delta V}$$