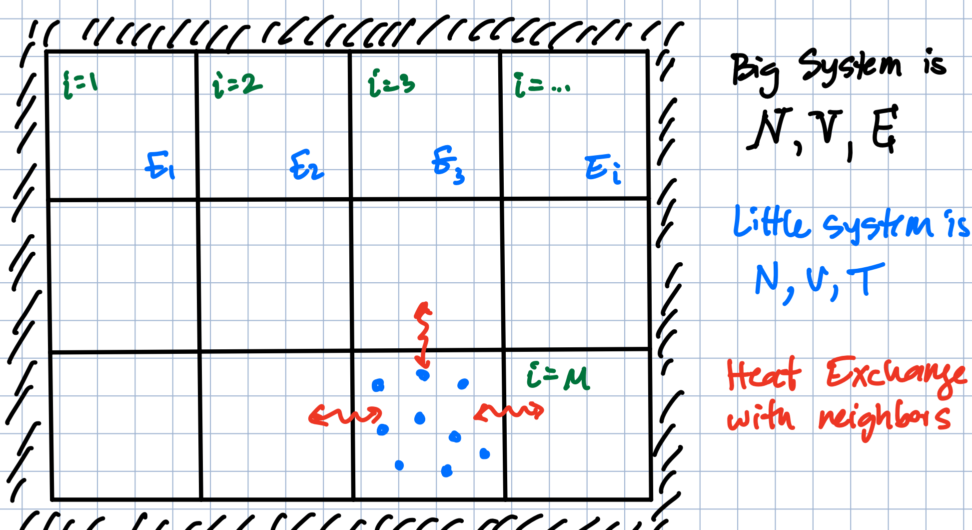
Lecture 17 – The Canonical Ensemble

A. Introduction to the Canonical Ensemble

We saw that we could define an ensemble of systems that had a fixed number of particles N, volume V, and energy E. However, energy and temperature are thermodynamic conjugates. E is not privileged over T. Also, it is more common for us to experimentally encounter a system that is at constant T. So, we would like to construct an ensemble with constant N, V, and T. This type of ensemble is called a canonical ensemble.

We are going to construct a canonical ensemble as follows. We want to use a microcanonical ensemble so we don't need any new postulates or assumptions.



- There are M "little systems" making up the ensemble.

$$\mathcal{N} = MN, \quad \mathcal{V} = MV$$

- The other systems act as a heat bath for the NVT system, keeping it at a constant temperature.

- The energy of the NVT system fluctuates and is a random variable E_i.
- In the M members of the ensemble, there will be m_j that have energy E_j.

$$m_j = \sum_{i=1}^{M} \delta(E_i, E_j)$$

+1 if $E_i = E_j$

m_j: occupation number

J: number of energy levels

$$\delta(E_i, E_j) = \begin{cases} 1 & \text{if } E_i = E_j \\ 0 & \text{otherwise} \end{cases}$$

- The total energy and total number of ensembles are constrained

$$\sum_j m_j E_j = E_t$$

$\underline{m} = \{m_j\}$, vector of all occupation #'s

$$\sum_j m_j = M$$

There are many different ways the energy of the microcanonical ensemble E_t can be distributed in the M NVT systems

Example: $M = 6$ $E_t = 8$ $J = 4$ $E_j = \{0, 1, 2, 3\}$

Distribution 1

i	1	2	3	4	5	6
E_i	2	1	1	2	1	1

$\underline{m} = [0, 4, 2\ 0]$

$$\sum_j m_j = 0 + 4 + 2 + 0 = 6 = M \checkmark$$

$$\sum_j m_j E_j = 0 \cdot 0 + 4 \cdot 1 + 2 \cdot 2 + 0 \cdot 3 = 4 + 4 = 8 = E_t \checkmark$$

Distribution 2

i	1	2	3	4	5	6
E_i	0	3	1	2	1	1

$\underline{m} = [1, 3, 1\ 1]$

$$\sum_j m_j = 1 + 3 + 1 + 1 = 6 = M \checkmark$$

$$\sum_j m_j E_j = 1 \cdot 0 + 3 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 = 3 + 2 + 3 = 8 = E_t \checkmark$$

B. Probability Density for the Canonical Ensemble

For the microcanonical ensemble, we

1. Postulated a probability density: uniform density of configurations with equal energy.
2. Used the probability density to find the density of states $\Omega(E)$.
3. Used the Boltzmann formula to find the entropy: $S = k_B \ln \Omega$

We want to follow an analogous procedure for the canonical ensemble. Let's first find the probability density of finding one of the members of our NVT ensemble at energy E_j

$$P_j = P_{NVT}(E_j) = ?$$

If we could look at the statistics of our ensemble, we would see that

$$P_j = \frac{\langle m_j \rangle}{M} \quad (*)$$

$\leftarrow$ average # of times we observe an NVT ensemble w/ energy E_j in $\mathcal{NVE}$

$\leftarrow$ total # of NVT ensembles

In the microcanonical ensemble, we said that any configuration with energy E was equally probable (principle of equal a priori probability). Here, the big system is an $\mathcal{NVE}$ ensemble. That means that any vector $\underline{m}$ where $\sum_j m_j E_j = \mathcal{E}$ is equally probable. So, we just to count the number of ways we can arrange the energy in the M bins. This will tell us the probability distribution.

$W(\underline{m})$, # of ways of arranging the states for a given $\underline{m}$

$$W(\underline{m}) = \frac{M!}{m_1!\, m_2! \cdots m_J!} = \frac{M!}{\prod_{j=1}^{J} m_j!}$$

Multinomial coefficient

Using $W(\underline{m})$ we can write a probability mass function for a configuration,

$$P_{NVE}(\underline{m}) = \frac{W(\underline{m})}{\sum_{\underline{m}} W(\underline{m})}$$ PMF of the configuration $\underline{m}$ in the NVE ensemble (multidimensional of all $\underline{m}$)

Using this probability mass function we can find the average

$$P_{NVT}(E_j) = \frac{\langle m_j \rangle}{M} = \frac{1}{M}\sum_{\underline{m}} P_{NVE}(\underline{m})\, m_j(\underline{m}) = \frac{1}{M}\,\frac{\sum_{\underline{m}} W(\underline{m})\, m_j(\underline{m})}{\sum_{\underline{m}} W(\underline{m})}$$

Rewriting (A) using $W(\underline{m})$

Example: $M = 6$ $\quad E_t = 8$ $\quad J = 4$ $\quad E_j = \{0, 1, 2, 3\}$

Distribution 1: $m = [0, 4, 2, 0]$

$$W(\underline{m}) = \frac{6!}{0!\,4!\,2!\,0!} = \frac{3\,\cancel{6}\cdot 5\cdot \cancel{4\cdot 3\cdot 2\cdot 1}}{\cancel{4\cdot 3\cdot 2\cdot 1}\cdot \cancel{2\cdot 1}} = 15$$

# of ways to arrange the $4 \times E = 1$ and $2 \times E = 2$ into 6 bins.

Distribution 2: $m = [1, 3, 1, 1]$

$$W(\underline{m}) = \frac{6!}{1!\,3!\,1!\,1!} = \frac{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1} = 6\cdot 5\cdot 4 = 120$$

Many more ways to come up with this distribution of energies

note: $\sum_{\underline{m}} W(\underline{m})$ = Add up all possible configurations (Hard to write as simple sum b/c of constraints.)

$m_j(\underline{m})$: # of ensembles with energy E_j when configuration is $\underline{m}$.

($m_4 = 0$ in Dist 1 and $m_4 = 1$ in Dist 2.)

Show the Python code for M ensembles with $J = 4$ energy levels and fixed energy $E_t = 8$.

Question: Is W the same as the density of states Ω? No, because it is related to the distribution of the microstates. It is related to S in NVT, which has a T-dependence.

Ok, this is all great, but what do I do when $M \to \infty$ (large ensemble) and J is big (large number of energies). Computing the multinomial coefficient will be very hard. Answer: when the $m_j \to \infty$, the multinomial coefficient becomes a very sharply peaked normal distribution around $m_j = m_j^*$. So sharp that it becomes a Dirac delta function $\delta(m_j - m_j^*)$ and only the most probable configuration matters. ✱ This maximizes the entropy of the NVE system.

Because the distribution is so sharp, the average is the same as the most likely state m^*. Therefore, our probability becomes

$$\langle m_j \rangle = m_j^* \Rightarrow P_{NVT}(E_j) = \frac{m_j^*}{M}$$

Example: Show that $\langle m_j \rangle = m_j^*$ when $M \to \infty$ for $J=2$ using Python. Proof is in the Appendix.

C. The Boltzmann Distribution

We want to find the most probable configuration for the multi-dimensional $W(m)$ subject to the constraints

$$\sum_j m_j = M \quad \leftarrow \text{total \# of ensembles}$$

$$\sum_j m_j E_j = E_t \quad \leftarrow \text{total energy}$$

We will use the method of Lagrange multipliers for the constraints. Let

$$f(m) = \ln W(\underline{m}) - \alpha\left(\sum_j m_j - M\right) - \beta\left(\sum_j m_j E_j - E_t\right)$$

Find the extremum:

$$\left.\frac{df}{dm_j}\right|_{m=m^*} = 0$$

maximizing $W(\underline{m})$ is the same as maximizing $P(\underline{m})$ because $P = W(\underline{m})/\text{const.}$ Easier to maximize the log, too.

$$\frac{df}{dm_j} = \underbrace{\frac{d}{dm_j}\left(\ln W\right)}_{1} - \alpha \underbrace{\frac{d}{dm_j}\left(\sum_j m_j - M\right)}_{2} - \beta \underbrace{\frac{d}{dm}\left(\sum_j m_j E_j - E_i\right)}_{3}$$

1: $\frac{d}{dm_j}(\ln W) = \frac{d}{dm_j} \ln\left(\frac{M!}{m_1!\, m_2! \cdots m_j \cdots m_J!}\right)$ ← Don't need to sub for m_J because of constraint

$$= \frac{d}{dm}(\cancel{\ln M!}) - \frac{d}{dm}(\cancel{\ln m_1!}) - \frac{d}{dm}(\cancel{\ln m_2!}) - \dots - \frac{d}{dm}(\ln m_j!)$$
$$- \dots - \frac{d}{dm}(\cancel{\ln m_J!})$$

$$= -\frac{d}{dm_j}(\ln m_j!) = -\frac{d}{dm}\left(m_j \ln m_j - m_j\right) = -\left(\ln m_j + \cancel{\frac{m_j}{m_j}} \cancel{-1}\right)$$

$$= -\ln m_j$$

2: $\frac{d}{dm_j}\left(\sum_j m_j - M\right) = \frac{d}{dm_j}\left(m_1 + m_2 + \dots + m_j + \dots + m_{J-1} + m_J - M\right) = 1$

3: $\frac{d}{dm_j}\left(\sum_j m_j E_j - E_i\right) = \frac{d}{dm_j}\left(m_1 E_1 + m_2 E_2 + \dots + m_j E_j + \dots + m_{J-1} E_{J-1} + m_J E_J - E_i\right)$

$$= E_j$$

Put it together:

$$\left.\frac{df}{dm_j}\right|_{m_j = m_j^*} = -\ln m_j^* - \alpha - \beta E_j = 0$$

$$\boxed{m_j^* = e^{-\alpha} e^{-\beta E_j}}$$

ok, finally, lets get the distribution

$$P(E_j) = \frac{m_j^*}{M} = \frac{e^{-\alpha} e^{-\beta E_j}}{M}$$

Summing over all the systems allows us to find α,

$$\sum_{i=1}^{M} P(E_i) = 1$$

$$\sum_{i=1}^{M} \frac{e^{-\alpha} e^{-\beta E_i}}{M} = 1 \Rightarrow \frac{e^{-\alpha}}{M} \sum_{i=1}^{M} e^{-\beta E_i} = 1 \Rightarrow e^{\alpha} = M \sum_{i=1}^{M} e^{-\beta E_i}$$

Putting it all together gives us the probability density!

$$\boxed{P_{NVT}(E_i) = \frac{e^{-\beta E_i}}{\sum_{i=1}^{M} e^{-\beta E_i}}}$$

We have yet to identify β.
We will see that $\beta = 1/k_B T$.

Boltzmann Distribution!

Summary of derivation

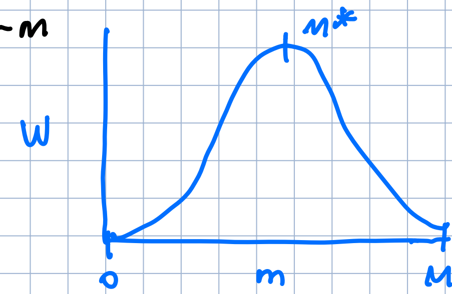
- We create an ensemble of NVT distributions that gives a total $\mathcal{N} \mathcal{V} \mathcal{E}$ system.
- Each NVT system has a random energy E_j.
- Any distribution of the energies, E_j that sums to $\mathcal{E}$ is equally likely to occur, but some of them occur many ways.
- For a large ensemble, only the most likely way to distribute the energies (the one with the most ways) will occur.
- We can solve this by maximizing the system, keeping the # of ensembles and energy constant.
- This gives the Boltzmann distribution.

D. Appendix: Proof that P_{NVE} becomes $\delta(m-m^*)$ at $M \to \infty$

Assume Binomial Coefficient ($J=2$) to make tractable.

$$W(m) = \frac{M!}{m_1!\, m_2!} \qquad m_1 = m \qquad m_2 = M-m$$

$$= \frac{M!}{m!\,(M-m)!} = \binom{M}{m}$$



(i.) Determining the maximum of the distribution

The maximum of $\ln W$ is the same as the maximum of W.

$$\ln W = \ln M! - \ln m! - \ln (M-m)!$$

$$= M\ln M - M - (m\ln m - m) - [(M-m)\ln(M-m) - (M-m)]$$

$$= M\ln M - \cancel{M} - m\ln m + \cancel{m} - M\ln(M-m)$$

$$+ m\ln(M-m) + \cancel{M} - \cancel{m}$$

$$= M\ln\left(\frac{M}{M-m}\right) - m\ln\left(\frac{m}{M-m}\right)$$

Now to find the maximum, we look for the extremum

$$\frac{d}{dm}(\ln W)\Big|_{m^*} = 0$$

$$\frac{d}{dm}\left[M\cancel{\ln}M - M\ln(M-m) - m\ln m + m\ln(M-m)\right] = 0$$

$$-\frac{M}{M-m}\cdot(-1) - \ln m - \underbrace{\frac{m}{m}}_{1} + \underbrace{\ln(M-m)\cdot(-1)}_{\to\ \text{combine}} + \underbrace{\frac{m}{M-m}(-1)}_{\to\ \text{combine}} = 0$$

$$\underbrace{\frac{(M-\cancel{m})}{\cancel{M}-m}}_{+1} + \ln\left(\frac{M-m}{m}\right) - \cancel{1} = 0 \Rightarrow \frac{M-m}{m} = e^0 = 1$$

$$M - m = m \Rightarrow 2m = M \Rightarrow \boxed{m^* = \frac{M}{2}}$$

(ii). Determine the distribution as $M \to \infty$

The distribution (PMF) is

$$P(m) = \frac{\binom{M}{m}}{\sum_m \binom{M}{m}} = \frac{1}{2^M}\binom{M}{m}$$

To find the distribution at large M, we will look at the cumulant generating function. Find the probability generating function first

$$G(z) = E[z^m]$$

$$G(z) = \sum_{m=0}^{M} \frac{1}{2^M} \underbrace{\binom{M}{m}}_{\text{binomial series}} z^m = \frac{1}{2^M}(1+z)^M = \left(\frac{1+z}{2}\right)^M$$

The moment generating function is

$$M(s) = E(e^{sm}) = G(e^s) = \left(\frac{1+e^s}{2}\right)^M$$

The cumulant generating function is

$$K(s) = \ln M(s) = \ln\left[\left(\frac{1+e^s}{2}\right)^M\right] = M\ln(1+e^s) - M\ln 2$$

Now find cumulants

$$k_1 = \mu = \frac{dK}{ds}\bigg|_{s=0} = \frac{M}{1+e^s}e^s\bigg|_{s=0} = \frac{M(1)}{1+1} = M/2$$

the maximum! $m^* = M/2$

$$k_2 = \sigma^2 = \frac{d^2K}{ds^2}\bigg|_{s=0} = \left[Me^s\cdot\left(\frac{1}{1+e^s}\right) + \frac{-Me^s}{(1+e^s)^2}e^s\right]_{s=0} = \left[\frac{M}{2} - \frac{M}{2^2}\right] = \frac{M}{4}$$

(used Mathematica for $k_3, \ldots$)

$k_3 = 0 \qquad k_4 = -M/8 \qquad k_5 = 0 \qquad k_6 = M/4$

Series solution

$$K(s) = \frac{M}{2}s + \frac{M}{4}\frac{s^2}{2!} - \frac{M}{8}\frac{s^4}{4!} + \frac{M}{4}\frac{s^6}{6!} + \ldots$$

Compare to normal (Gaussian):

$$k_{Gauss} = \mu s + \frac{\sigma^2}{2} s^2$$

Do the higher order terms matter as $M \to \infty$? (No) s is of the order of $1/M$ (e^{sm}). Let's rescale.

$$\text{let } \tilde{s} = Ms \to s = \tilde{s}/M$$

$$K(\tilde{s}) = \frac{M}{2}\frac{\tilde{s}}{M} + \frac{M}{4}\frac{\tilde{s}^2}{2M^2} - \frac{M}{8}\cdot\frac{\tilde{s}^4}{24M^4} + \frac{M}{4}\frac{\tilde{s}^6}{720M^6}$$

$$k(\tilde{s}) = \frac{1}{2}\tilde{s} + \frac{1}{8M}\tilde{s}^2 - \frac{1}{192M^3}\tilde{s}^4 + \frac{1}{2880M^5}\tilde{s}^6$$

The width will also go to 0 like $M^{-1/2}$.

These terms will drop out very quickly as $M \to \infty$

Conclusion: By CLT $K(s) \to$ Gaussian with a maximum at $M/2$ and width $\sqrt{M/4}$. It will get sharper and sharper as $M \to \infty$.

(iii). More accurate Stirling's Approximation

For plotting the Stirling's approximation, $O(\ln n)$ corrections make a difference. Let's do this one out real quick

$$W = \frac{M!}{m!(M-m!)} \qquad \ln n! \sim n\ln n - n + \tfrac{1}{2}\ln(2\pi n)$$

$$\begin{aligned}
\ln W &= \ln M! - \ln m! - \ln(M-m)! \\
&= M\ln M - M + \tfrac{1}{2}\ln(2\pi M) - \left[m\ln m - m + \tfrac{1}{2}\ln(2\pi m)\right] \\
&\quad - \left[(M-m)\ln(M-m) - (M-m) + \tfrac{1}{2}\ln(2\pi(M-m))\right] \\
&= M\ln M - \cancel{M} + \tfrac{1}{2}\ln(2\pi M) - m\ln m + \cancel{m} - \tfrac{1}{2}\ln(2\pi m) \\
&\quad - M\ln(M-m) + \cancel{M} - \tfrac{1}{2}\ln(2\pi(M-m)) + m\ln(M-m) - \cancel{m}
\end{aligned}$$

$$= M \ln\left(\frac{M}{M-m}\right) - m \ln\left(\frac{m}{M-m}\right) + \frac{1}{2} \ln\left[\frac{\cancel{2}\pi M}{(\cancel{2}\pi m)(2\pi(M-m))}\right]$$

$$= M \ln\left(\frac{M}{M-m}\right) - m \ln\left(\frac{m}{M-m}\right) + \underbrace{\frac{1}{2} \ln\left(\frac{1}{2\pi}\frac{M}{M-m}\right)}_{\text{This extra term helps the approximation at moderate } m.}$$

The normalization constant for the exact binomial coefficient is 2^M!

In other words

$$\sum_{m=0}^{M} \binom{M}{m} = 2^M$$