Reminders

① Midterm exam Monday Oct 18 (STEM classroom)

10:10 – ≈ 11:20

- Material from TD & Fundamentals of SM (HW's 1–4; lectures thru today & monday)

- Will provide any non-trivial formulas

② HW4 due Friday Oct 15

Recap of L13

* Internal excitations (e.g. vibrational, rotational, intrinsic spin, ...)

$$\epsilon(\vec{p}) \rightarrow \epsilon_\alpha(\vec{p}) = \epsilon(\vec{p}) + \epsilon_\alpha$$

$$\therefore \quad Z_1 = Z_1^{trans.} \times Z_1^{int} \Rightarrow Z_N = \frac{1}{N!}\left(Z_1^{trans}\right)^N \times \left(Z_1^{int}\right)^N$$

$\Rightarrow$ corrections to F, S, E, C_V etc.

* Quantized nature of ϵ_α (QM)

$\Rightarrow$ modes "frozen out" at low T; deviates from EP thm.

Ensemble theory revisited

* I'd like to revisit the derivation of P_i of the canonical + grand canonical ensembles from a different perspective due to Gibbs, since it comes in use later on when we discuss bose/fermi distributions (+ also because it connects with ideas from other fields like information theory if we decide to explore that later on.)

* The key concept is a variational principle where the entropy is maximized under different constraints, depending on the ensemble.

→ N_s mental copies of our system
(i.e., # of copies in the ensemble)

→ each system can be in any of M different micro states

→ # of arrangements for n_i systems in microstate i $\equiv W = \dfrac{N_s!}{n_1!\, n_2! \cdots n_M!}$

where $\sum_{i=1}^{M} n_i = N_s$

* Sometimes W is called "the ignorance" since larger values mean there are more possible configs. of the ensemble, so we know less about its state.

* Following Gibbs, define entropy-like qty $(N_s S)$ for ensemble

$$N_s S = \log W$$

$$= \log(N_s!) - \sum_i \log(n_i!)$$

* Take the freedom to take N_s (+ hence n_i) large $\Rightarrow$ Stirling Approx:

$$N_s S = N_s \log N_s - \cancel{N_s} - \sum_i n_i \log n_i + \overset{N_s}{\cancel{\sum_i n_i}}$$

$$= \Big(\sum_i n_i\Big) \log N_s - \sum_i n_i \log n_i$$

$$= \sum_i n_i \left(\log N_s - \log n_i\right)$$

$$= -\sum_i n_i \log \frac{n_i}{N_s}$$

Since $\frac{n_i}{N_s}$ = fraction of systems in state i

$$\Rightarrow \boxed{S = -\sum_i \frac{n_i}{N_s} \log \frac{n_i}{N_s} = -\sum_i P_i \log P_i}$$

* Now we re-derive p_i for the different ensembles by maximizing S (entropy) under different constraints
(treat p_i as variational parameters)

1. Microcanonical

$$\frac{\partial S}{\partial p_i} = 0 \quad \text{with constraint} \quad \sum_i p_i = 1$$

$$\text{or} \quad \left(\sum_i p_i - 1\right) = 0$$

Reminder: If you want to maximize $f(\vec{x})$ subject to some constraint $g(\vec{x}) = 0$, then you maximize without any constraint the modified function $(f(\vec{x}) - \lambda g(\vec{x}))$ wrt $\vec{x}$ & λ (say $\vec{x} \in \mathbb{R}^N$)

i.e.,

$$\left.\begin{aligned} &\frac{\partial}{\partial x_i}\left(f(\vec{x}) - \lambda g(\vec{x})\right) = 0 \\ &\frac{\partial}{\partial \lambda}\left(f(\vec{x}) - \lambda g(\vec{x})\right) = 0 \end{aligned}\right\} \text{solve } N+1 \text{ equations}$$

* if multiple constraints $g_\alpha(\vec{x}) = 0 \quad \alpha = 1 \ldots p$

$$\left.\begin{aligned} &\frac{\partial}{\partial x_i}\left(f(\vec{x}) - \sum_\alpha \lambda_\alpha g_\alpha(\vec{x})\right) = 0 \\ &\frac{\partial}{\partial \lambda_\beta}\left(f(\vec{x}) - \sum_\alpha \lambda_\alpha g_\alpha(\vec{x})\right) = 0 \end{aligned}\right\}$$

Review of Lagrange Multipliers

* Say you want to extremize $f(\vec{x})$ subject to the constraint $g(\vec{x}) = 0$ (let $\vec{x} \in \mathbb{R}^N$)

* The constraint $g(\vec{x}) = 0$ specifies a hypersurface of $(N-1)$ dimensions

* Say you're at a point $\vec{x}_0$ on the hypersurface. If you move to $\vec{x}_0 \to \vec{x}_0 + \delta\vec{x}$

$$dg(\vec{x}_0) = g(\vec{x}_0 + \delta\vec{x}) - g(\vec{x}_0)$$
$$= \delta\vec{x} \cdot \vec{\nabla} g \Big|_{\vec{x}_0}$$

$\therefore$ We stay on the hypersurface $g(\vec{x}) = 0$ if $\delta\vec{x} \perp \vec{\nabla} g$

* Geometrically, finding extremum of $f(\vec{x})$ subject to $g(\vec{x}) = 0$ means finding $\vec{x}_0$ where $df = \delta\vec{x} \cdot \vec{\nabla} f \Big|_{\vec{x}_0} = 0$ in all directions on the $(N-1)$ dimension hypersurface

$$\therefore \vec{\nabla} f \Big|_{\vec{x}_0} \parallel \vec{\nabla} g \Big|_{\vec{x}_0}$$

* Therefore, a necessary condition for finding the extremum of $f(\vec{x})$ subject to $g(\vec{x}) = 0$ is

$$\vec{\nabla} f(\vec{x})\Big|_{\vec{x}=\vec{x}_0} = \lambda\, g(\vec{x})\Big|_{\vec{x}=\vec{x}_0}$$

arbitrary scalar
"Lagrange Multiplier"

or

$$\boxed{\vec{\nabla}\left[f(\vec{x}) - \lambda g(\vec{x})\right]\Big|_{\vec{x}_0} = 0} \qquad \otimes$$

* Eq. $\otimes$ alone is not a sufficient condition, because all it guarantees is that we are finding an extremum of $f(\vec{x})$ on some contour $g(\vec{x}) = C$ (not necessarily $g(\vec{x}) = 0$). Different $\lambda \Leftrightarrow$ different C's.

$\therefore$ We supplement $\otimes$ w/ the condition $g(\vec{x}) = 0$, which can be implemented by finding the unconstrained extremum of $\mathcal{L}(\vec{x},\lambda) = f(\vec{x}) - \lambda g(\vec{x})$ in the $(N+1)$-dim space $(\vec{x},\lambda)$.

$$\boxed{\begin{aligned} \frac{\partial}{\partial x_i}\left[f(\vec{x}) - \lambda g(\vec{x})\right] &= 0 \\ \frac{\partial}{\partial \lambda}\left[f(\vec{x}) - \lambda g(\vec{x})\right] &= 0 \end{aligned}}$$

ex: Find extremum of $f(x,y) = 3x^2 - 4xy + y^2$
w/ constraint $g(x,y) = 3x + y = 0$

① $\frac{\partial}{\partial x}\left[(3x^2 - 4xy + y^2) - \lambda(3x+y) \right] = 0$

② $\frac{\partial}{\partial y}\left[(3x^2 - 4xy + y^2) - \lambda(3x+y) \right] = 0$

③ $3x + y = 0$

① $\Rightarrow$ $6x - 4y - 3\lambda = 0$

② $\Rightarrow$ $-4x + 2y - \lambda = 0$

③ $\Rightarrow$ $3x + y = 0$

$$\left.\begin{aligned} ①+②&\Rightarrow\ 18x - 10y = 0 \\ ③&\Rightarrow\ 3x + y = 0 \end{aligned}\right\}$$ solving for x, y

$x = 0$
$y = 0$

Note: If there are multiple constraints $g_i(\vec{x}) = 0 \quad i = 1 \ldots p$, then solve the $(N+p)$ equations

$$\vec{\nabla}\left[f(\vec{x}) - \sum_{i=1}^{p} \lambda_i g_i(\vec{x}) \right] = 0$$

$$g_1(\vec{x}) = 0$$

$$\vdots$$

$$g_p(\vec{x}) = 0$$

... Back to our problem to derive different ensemble P_i's via maximizing S' wrt P_i with diff. constraint

$$S' = -\sum_i \frac{n_i}{N_s} \log \frac{n_i}{N_s} = -\sum_i P_i \log P_i$$

(1) Microcanonical

$$\frac{\partial S'}{\partial P_i} = 0$$ with constraint $\sum_i P_i = 1$

or $\left(\sum_i P_i - 1\right) = 0$

$$\therefore \quad \frac{\partial}{\partial p_j}\left(-\sum_i p_i \log p_i - \lambda\left(\sum_i p_i - 1\right)\right) = 0$$

$$\Rightarrow \quad -\log p_j - 1 - \lambda = 0$$

$$= \log p_j = -\lambda - 1$$

$$\Rightarrow p_j = e^{-1-\lambda} \equiv p \text{ (constant)}$$

since $\sum_{i=1}^{M} p_i = 1 = pM$

$\Rightarrow p = \frac{1}{M}$ as before ($M \to \Omega$)

(2.) Canonical

$\frac{\partial S}{\partial p_i} = 0$ w/ constraint

(1) $\sum_i p_i = 1$

(2) $\langle E \rangle = \sum_i p_i E_i$

$$\Rightarrow \frac{\partial}{\partial p_j}\left[-\sum_i p_i \log p_i - \lambda\left(\sum_i p_i - 1\right) - \beta \sum_i p_i E_i\right] = 0$$

$$= -\log p_j - 1 - \lambda - \beta E_j$$

$\Rightarrow \log p_j = \log C - \beta E_j \qquad (\log C \equiv -1 - \lambda)$

$p_j = Ce^{-\beta E_j}$; $\sum_j p_j = 1 \Rightarrow C = \frac{1}{\sum_j e^{-\beta E_j}}$

* Similarly, you can find $p(i)$ for GCE by

$$\frac{\partial S}{\partial p_i} = 0 \quad \text{w/ constraints} \quad \begin{array}{l} (1)\ \sum_i p_i = 1 \\ (2)\ \sum_i p_i N_i = \langle N \rangle \\ (3)\ \sum_i p_i E_i = \langle E \rangle \end{array}$$

where now $\beta\mu$ appears as the Lagrange multiplier of the $\sum_i p_i N_i = \langle N \rangle$ constraint.

Comments

* You might feel this is very different than, e.g., the earlier derivation we did of the CE where we considered our system connected to a large heat bath it can exchange energy with.

* But nowhere does the derivation rely on the detailed nature of the heat bath.

* Gibbs's approach is effectively taking the heat bath as comprised of $(N_s - 1)$ copies of the system!

* This is really the same approach as before, (i.e., treating combined S+R in microcanonical to infer the p_i's for the sys.)

albeit slightly in disguise

* Therefore, the problem is to maximize the multiplicity

$$\Omega_{TOT}(\mathcal{E}_{TOT}) = W = \frac{N_s!}{n_1!\, n_2! \cdots n_m!} \quad \text{wrt. } n_i\text{'s}$$

with constraints

① $$\sum_{i=1}^{m} n_i E_i = \mathcal{E}_{TOT}$$

② $$\sum_{i=1}^{m} n_i = N_s$$

* But dividing ① + ② by N_s, we see this means

$$① \Rightarrow \sum_{i=1}^{m} \left(\frac{n_i}{N_s}\right) E_i = \sum_{i=1}^{m} p_i E_i = \frac{\mathcal{E}_{TOT}}{N_s} = \langle E \rangle_{system}$$

$$② \Rightarrow \sum_{i=1}^{m} p_i = 1$$

i.e., it's maximizing W (which is the same as maximizing $\log W$ which is the same as maximizing $S = \frac{1}{N_s} \log W$)

Subject to 1.) $\langle E \rangle = \sum_i p_i E_i$

2.) $\sum_i p_i = 1$

* This Gibbs approach is powerful + generalizeable

e.g.: Systems w/ net macroscopic ∡ momentum (e.g., neutron star) at constant T, V, N

$$\Rightarrow \frac{\partial S}{\partial p_i} = 0$$ w/ constraints

1.) $\sum_i p_i = 1 \Rightarrow \left(\sum_i p_i - 1\right) = 0$

2.) $\left(\langle E\rangle - \sum_i p_i E_i\right) = 0$

3.) $\left(\langle L^z\rangle - \sum p_i L_i^z\right) = 0$

$$\therefore \frac{\partial}{\partial p_j}\left[-\sum_i p_i \log p_i - \lambda_1\left(\sum_i p_i - 1\right) - \lambda_2\left(\sum_i p_i E_i - \langle E\rangle\right) - \lambda_3\left(\sum_i p_i L_i^z - \langle L^z\rangle\right)\right] = 0$$

$$\Downarrow$$

Ultimately find

$$p_j = \frac{1}{Z} e^{-\beta(E_j - \omega L_j^z)}$$

(λ_1 related to Z

$\lambda_2 = \beta$

$\lambda_3 = \omega\beta$)

∡ velocity

(Know from mechanics that

$E_j - \omega L^z$ = energy of the rotating body in the body-fixed frame.)