PHYS 521: Statistical Mechanics

Homework #2

Prakash Gautam

April 26, 2018

1. Show that for a given N_r with $\sum_i^N p_i = 1$, the uncertainty function $S(\{p_i\})$, takes its maximum value when $p_i = \frac{1}{N}$ for all i, that is $S(\{p_i\}) = A(N)$
Solution:
The uncertainty function is $S(\{p_i\}) = -C\sum_i p_i \ln p_i$. We want to maximize this function subject to the constraint $\sum_i p_i = 1$. Using Lagrange's multiplier method to find the extremum of function, we can define a new function $S - \lambda\left(\sum_i p_i - 1\right)$

$$\begin{aligned}
\frac{\partial S'}{\partial p_j} &= \frac{\partial}{\partial p_j}\left[-C\sum_i p_i \ln p_i - \lambda\left(\sum_i p_i - 1\right)\right] \\
&= -C\sum_i\left(\delta_{ij}\ln p_i + \frac{1}{p_j}p_i\delta_{ij}\right) - \lambda\left(\sum_i \delta_{ij}\right) \\
&= -C\left(\ln p_j + 1\right) - \lambda \qquad (1)
\end{aligned}$$

But for extremum condition of this function the partial derivative with respect to every p_j should vanish. Thus we get

$$\ln p_j = -\frac{\lambda}{C} - 1 \qquad \Rightarrow \qquad p_j = \exp\left[-\frac{\lambda}{C} - 1\right]$$

The RHS of above expression is a constant, lets call that constant M so $p_i = M$ for some constant M but since probability has to add 1 we get

$$\sum_j p_j = 1; \qquad \Rightarrow \sum_j M = 1 \qquad \Rightarrow MN = 1 \qquad \Rightarrow M = \frac{1}{N}$$

Substuting this back we get

$$p_j = \frac{1}{M}$$

Thus the uncertainty function takes it maximum value when $p_i = 1/N$ for all p_i $\square$

2. Consider a urn problem discussed in class: An urn is filled with balls, each numbered $n = 0, 1, 2\ldots$. The average value of n is $\langle n\rangle = 2/7$. Calculate the probabilities p_0, p_1

and p_2 which yield the maximum uncertainty. Find the expectation value, based on these probabilities $\langle n^3 \rangle - 2\langle n \rangle$.
Solution:
The expectation value of n is given by

$$\langle n \rangle = p_0 \cdot 0 + p_1 \cdot 1 + p_2 \cdot 2 \quad \Rightarrow \quad p_1 + 2p_2 = 2/7$$

This is one of the constraints for maximizing the uncertainty function, the other constraint equation is $p_0 + p_1 + p_2 = 1$. Using these as we calculated in (1) we have

$$S' = \frac{S}{C} - \alpha(p_1 + 2p_2 - 2/7) - \beta(p_0 + p_1 + p_2 - 1)$$

Taking derivative with respect to α and β and equating to zero gives

$$\begin{aligned} \ln p_0 + 1 - \beta &= 0 \\ \ln p_1 + 1 - \alpha - \beta &= 0 \\ \ln p_2 + 1 - 2\alpha - \beta &= 0 \end{aligned}$$

These three equations along with two constraint equation form five equation in five unkonwn $p_0, p_1, p_2, \alpha, \beta$. We can solve this equation to get the numeric value of the parameters. Solving for the parameters we get

$$p_0 = \frac{15}{21} \qquad p_1 = \frac{4}{21} \qquad p_2 = \frac{1}{21}$$

Now the requried function is

$$\begin{aligned} \langle n^3 \rangle - 2\langle n \rangle &= p_0 \cdot 0 + p_1 \cdot 1^3 + p_2 \cdot 2^3 - \langle n \rangle \\ &= p_1 + 8p_2 - 2\frac{2}{7} \\ &= \frac{4}{21} + 8\frac{1}{21} - \frac{4}{7} \\ &= 0 \end{aligned}$$

The requried value is 0

□

3. Assuming the entropy, S and the number of microstates, Ω of a physical system are related through an arbitrary functional form $S = f(\Omega)$, show that the additive character of S (extensive parameter) and the multiplicative parameter Ω meaning $\Omega = \Omega_1, \Omega_2, \ldots$, is the number of microscopic states for a subsystem necessarily require that the function $F(\omega)$ is of the form

$$S = k \ln(\Omega)$$

where k is a (universal) constant. The form was first written down by Max Plank.
Solution:

Given the multiplicative parameter $\Omega = \Omega_1 \cdot \Omega_2 \ldots \Omega_r$. The extensive parameter as a function of this parameter which is a additive function be S. Thus we have

$$S\left(\Omega_1 \cdot \Omega_2 \ldots \Omega_r\right) = S(\Omega_1) + S(\Omega_1) + \ldots + S(\Omega_r)$$
$$S(\Omega) = \sum_j^r S(\Omega_j)$$

Differentiating with respect to Ω_i on both sides

$$\frac{\mathrm{d}}{\mathrm{d}\Omega_i} S\left(\Omega\right) = \frac{\mathrm{d}}{\mathrm{d}\Omega_i} \sum_j^r S(\Omega_j)$$
$$\frac{\mathrm{d}S(\Omega)}{\mathrm{d}\Omega} \frac{\mathrm{d}\Omega}{\mathrm{d}\Omega_i} = \sum_j^r \frac{\mathrm{d}S(\Omega_j)}{\mathrm{d}\Omega_i} \delta_{ij}$$

But since the derivative of product $\Omega = \Pi_j \Omega_j$ with respect to Ω_i is just the product without that parameter $\frac{\mathrm{d}\Omega}{\mathrm{d}\Omega_i} = \Pi_{j \neq i} \Omega_j$. Multiplying both sides by Ω_i we get

$$\Omega_i \left(\Pi_{j \neq i} \Omega_j\right) \frac{\mathrm{d}S(\Omega)}{\mathrm{d}\Omega} = \Omega_i \frac{\mathrm{d}S(\Omega_i)}{\mathrm{d}\Omega_i} \qquad \Rightarrow \qquad \Omega \frac{\mathrm{d}S(\Omega)}{\mathrm{d}\Omega} = \Omega_i \frac{\mathrm{d}S(\Omega_i)}{\mathrm{d}\Omega_i}$$

But the expression $\frac{1}{x}dx \equiv d(\ln(x))$ recognizing similar expression in both sides of the equality above we get

$$\frac{\mathrm{d}S(\Omega)}{\mathrm{d}(\ln \Omega)} = \frac{\mathrm{d}S(\Omega_i)}{\mathrm{d}(\ln \Omega_i)}$$

The expression in RHS is independent of expression on right. Since the product of the parameters can be varied while still keeping one of the parameters Ω_i constant. So the expression can only be equal to each other if they are equal to a constant.

$$\frac{\mathrm{d}S(\Omega)}{\mathrm{d}(\ln \Omega)} = k \qquad \Rightarrow \qquad dS(\Omega) = kd(\ln \Omega)$$

Integrating this expression we get

$$S(\Omega) = k \ln \Omega$$

Which is the required expression. □

4. Show that in $\ln x \leq x - 1$, if for all real positive x. The equality holds for $x = 1$.
Solution:
Rearranging the equation $\ln x - x \leq -1$. Let us define a function $g(x) = \ln x - x$. Differentiating this function with respect to x we get

$$g'(x) = \frac{1}{x} - 1 = \frac{1 - x}{x} = -\frac{x - 1}{x}$$

Since for all positive values of x i.e., $\forall x > 0$ we have

$$x - 1 < x \Rightarrow \frac{x-1}{x} < 1 \qquad \Rightarrow g'(x) = -\frac{x-1}{x} < -1$$

let $f(x) = \ln(1+x) - x$ so that $f(0) = 0$.

Clearly

$$f'(x) = -\frac{x}{1+x}$$

and hence $g'(x) > 0$ if $-1 < x < 0$ and $f'(x) < 0$ if $x > 0$. It follows that that $f(x)$ in increasing in $(-1, 0]$ and decreasing in $[0, \infty)$. Thus we have $f(x) < f(0)$ if $-1 < x < 0$ and $f(x) < f(0)$ if $x > 0$. It thus follows that $f(x) \leq f(0) = 0$ for all $x > -1$ and there is equality only when $x = 0$. So we can write

$$\ln(1+x) \leq x \qquad \forall x \geq -1$$

Since x is just a dummy variable we can transform $x \to x - 1$ to get

$$\ln(x) \leq x - 1 \qquad \forall x \geq 0$$

This completes the proof. □

5. Prove that $\log_2 X = \frac{\log X}{\log 2}$. Interpret the meaning of

$$S = -\sum_i p_i \log_2(p_i)$$

.

Solution:

Let $y = \log_2 X$. Raising both sides to 2 a gives us

$$2^y = 2^{\log_2 x} \qquad \Rightarrow \qquad 2^y = x$$

Taking logarithm on both side with respect to base 110 we get

$$\log X = \log\left(2^y\right) \qquad \Rightarrow \log X = y \log 2 \qquad \Rightarrow \qquad y = \frac{\log X}{\log 2}$$

But by our assumption $y = \log_2 X$ thus we have

$$\log_2 X = \frac{\log X}{\log 2}$$

In digital electronics and in information theory where they represent the signal information in binary, the logarithm of a number with respect to 2 gives the total number of bits required to represent the number. Multiplying the number of bits $\log_2 N$ by the probability of the number gives the total average number of bits required.

So then the entropy function $S = -\sum_i p_i \log_2(p_i)$ represents the infromation content of the signal. □