Unit 2-10: The Average Energy vs the Most Probable Energy in the Canonical Ensemble; Stirling's Formula

In the previous section we showed how the relative width of the canonical probability density for the energy $\mathcal{P}(E)$ scaled to zero in the thermodynamic limit $N \to \infty$. Here we look at some other aspects of this distribution. In particular we look at the difference between the average energy and the most probable energy.

In our discussion of energy fluctuations in the canonical ensemble in Notes 2-9, we expanded $E - TS(E)$ to second order about the minimizing energy $\bar{E}$,

$$E - TS(E) \approx A_{\text{micro}} + \frac{\delta E^2}{2TC_V} \tag{2.10.1}$$

This caused the probability density $\mathcal{P}(E) \propto \mathrm{e}^{-[E-TS(E)]/k_BT}$ to be a Gaussian distribution in $\delta E = E - \bar{E}$, with $\bar{E}$ the value of energy that minimizes $E - TS$, or equivalently the value of energy that maximizes the probability $\mathcal{P}(E)$. So $\bar{E}$ is the most probable value of the energy. Since, at this level of approximation, $\mathcal{P}(E)$ is symmetric in δE, one has that $\langle \delta E \rangle = 0$ and so $\bar{E}$ is also the average energy $\langle E \rangle$. To see the difference between the average value $\langle E \rangle$ and the most probable value $\bar{E}$, it is therefore necessary to continue the expansion in Eq. (2.10.1) to third order,

$$E - TS(E) \approx A_{\text{micro}} + \frac{\delta E^2}{2TC_V} - \frac{1}{3!}T\left(\frac{\partial^3 S}{\partial E^3}\right)\bigg|_{E=\bar{E}} \delta E^3 \tag{2.10.2}$$

Note, since $S \sim N$ and $E \sim N$ then $(\partial^3 S/\partial E^3) \sim 1/N^2$. So we can write,

$$E - TS(E) \approx A_{\text{micro}} + \frac{\delta E^2}{2TC_V} - \frac{\gamma}{N^2}\delta E^3 \tag{2.10.3}$$

where γ is some intensive parameter that does not increase as N increases (though γ can vary with T).

Now we compute the difference between the average and the most probable energies

$$\langle \delta E \rangle = \langle E \rangle - \bar{E} \tag{2.10.4}$$

We have

$$\langle \delta E \rangle = \frac{\displaystyle\int \frac{d\delta E}{\Delta E}\, \mathrm{e}^{-[E-TS(E)]/k_BT}\, \delta E}{\displaystyle\int \frac{d\delta E}{\Delta E}\, \mathrm{e}^{-[E-TS(E)]/k_BT}} \tag{2.10.5}$$

$$= \frac{\displaystyle\int d\delta E\, \exp\left[\frac{-\delta E^2}{2k_BT^2C_V} + \frac{\gamma\delta E^3}{N^2k_BT}\right]\delta E}{\displaystyle\int d\delta E\, \exp\left[\frac{-\delta E^2}{2k_BT^2C_V} + \frac{\gamma\delta E^3}{N^2k_BT}\right]} \qquad \text{the } \mathrm{e}^{-A_{\text{micro}}/k_BT} \text{ factors cancel} \tag{2.10.6}$$

$$= \frac{\displaystyle\int d\delta E\, \exp\left[\frac{-\delta E^2}{2k_BT^2C_V}\right]\left(1 + \frac{\gamma\delta E^3}{N^2k_BT}\right)\delta E}{\displaystyle\int d\delta E\, \exp\left[\frac{-\delta E^2}{2k_BT^2C_V}\right]\left(1 + \frac{\gamma\delta E^3}{N^2k_BT}\right)} \qquad \text{expanding the } \delta E^3 \text{ term in the exponentials} \tag{2.10.7}$$

The above has the form,

$$\langle \delta E \rangle = \frac{\displaystyle\int dx\, \mathrm{e}^{-x^2/2\sigma^2}\left(x + ax^4\right)}{\displaystyle\int dx\, \mathrm{e}^{-x^2/2\sigma^2}\left(1 + ax^3\right)} \qquad \text{with } \sigma^2 = k_BT^2C_V \text{ and } a = \frac{\gamma}{N^2k_BT}. \tag{2.10.8}$$

The integrals go from $x = -\bar{E}$ to $+\infty$. Since the width of the Gaussian exponential is $\sigma \sim \sqrt{C_V} \sim \sqrt{N}$, while $\bar{E} \sim N$, we can replace the lower limit by $-\infty$, and then do the Gaussian integrals. Because Gaussian distributions are symmetric, only the even moments are non-vanishing. We therefore get,

$$\langle \delta E\rangle = \frac{\displaystyle\int dx\, \mathrm{e}^{-x^2/2\sigma^2}\ \left(ax^4\right)}{\displaystyle\int dx\, \mathrm{e}^{-x^2/2\sigma^2}} = \frac{a\displaystyle\int dx\, \mathrm{e}^{-x^2/2\sigma^2}\ \left(x^4\right)}{\sqrt{2\pi\sigma^2}} = a\langle x^4\rangle = 3a\sigma^4 \tag{2.10.9}$$

Thus,

$$\langle \delta E\rangle = 3\left(\frac{\gamma}{N^2 k_B T}\right)\left(k_B T^2 C_V\right)^2 \sim \left(\frac{1}{N^2}\right)\left(N^2\right) \sim O(1) \tag{2.10.10}$$

Therefore the relative difference between the average and the most probable energies scales as,

$$\frac{\langle E\rangle - \bar{E}}{\langle E\rangle} = \frac{\langle \delta E\rangle}{\langle E\rangle} \sim \frac{1}{N} \to 0 \qquad \text{as } N \to \infty. \tag{2.10.11}$$

So, as $N \to \infty$, the relative difference between the average and the most probable energies scales to zero faster than the relative width of the distribution, $\sqrt{\langle \delta E^2\rangle}/\langle E\rangle \sim 1/\sqrt{N}$.

Now consider the probability density $\mathcal{P}(E)$ as a whole. We know that in the thermodynamic limit $N \to \infty$, $\langle E\rangle \sim \bar{E} \sim N$, while the width $\sqrt{\langle E^2\rangle - \langle E\rangle^2} \sim \sqrt{N}$. Thus the location of the peak in $\mathcal{P}(E)$ increases faster than the width increases, and so the relative width $\sim 1/\sqrt{N}$ vanishes.

To look at the $N \to \infty$ limiting form of $\mathcal{P}(E)$ we need instead to look at the probability density $\mathcal{P}_u$ for the energy per particle $u = E/N$, since $\langle u\rangle$ approaches a finite constant as $N \to \infty$. We have,

$$\mathcal{P}_u(u)du = \mathcal{P}(E)dE \qquad \Rightarrow \qquad \mathcal{P}_u(u) = \mathcal{P}(E)\frac{dE}{du} = \mathcal{P}(E)\,N \tag{2.10.12}$$

Using $\Omega(E) = \mathrm{e}^{S(E)/k_B}$, we therefore have from Notes 2-8,

$$\mathcal{P}_u(u) = N\,\frac{\Omega(E)\,\mathrm{e}^{-E/k_BT}}{\Delta E\ Q_N} = N\,\frac{\mathrm{e}^{-[E-TS(E)]/k_BT}}{\displaystyle\int dE\,\mathrm{e}^{-[E-TS(E)]/k_BT}} \qquad\qquad E \equiv uN \tag{2.10.13}$$

We can then expand in $\delta E = E - \bar{E}$ about $\bar{E}$, the minimum of $E - TS(E)$,

$$E - TS(E) = \bar{E} - TS(\bar{E}) + \delta E - T\sum_{n=1}^{\infty}\frac{1}{n!}\left(\frac{\partial^n S}{\partial E^n}\right)\bigg|_{E=\bar{E}}\delta E^n \tag{2.10.14}$$

The first two terms give just $\bar{E} - TS(\bar{E}) = A_{\text{micro}}$. As was shown in Notes 2-9, the terms linear in δE cancel. The coefficient of the δE^n term is of order $S/E^n \sim 1/N^{n-1}$. To leading order, therefore, we can keep just the lowest $n = 2$ term. All other terms will give corrections that become negligible compared to the leading term as $N \to \infty$. Thus we can stick with the approximation of Eq. (2.10.1).

So, using Eq. (2.10.1), $E/N = u$, and $\delta u = u - \bar{E}/N = u - \bar{u}$, we then get

$$\mathcal{P}_u(u) = \frac{\mathrm{e}^{-\delta u^2/2\sigma^2}}{\displaystyle\int d\delta u\,\mathrm{e}^{-\delta u^2/2\sigma^2}} = \frac{\mathrm{e}^{-\delta u^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \qquad \text{where } \sigma^2 = k_B T^2 C_V/N^2. \tag{2.10.15}$$

We thus see that, as $N \to \infty$, the probability density $\mathcal{P}_u(u)$ becomes a Gaussian distribution centered at $u = \bar{u}$ (i.e. $\delta u = 0$), with width $\sigma = \sqrt{k_B T^2 C_V}/N \sim 1/\sqrt{N}$.

Thus, in the thermodynamic limit $N \to \infty$, the width of that Gaussian distribution vanishes, and $\mathcal{P}_u(u)$ approaches the delta function $\delta(u - \bar{u})$. The canonical ensemble at temperature T, such that the most probable energy per particle is $\bar{u} = \bar{E}/N$, becomes the same as the microcanonical ensemble at fixed energy per particle $\bar{u}$.

Stirling's Formula

Just for fun, we will now use the saddle point approximation to derive Stirling's formula for $n!$

Consider the integral,

$$I_n = \int_0^\infty dx\, x^n\, \mathrm{e}^{-x} \tag{2.10.16}$$

We can integrate by parts to get,

$$I_n = -\left[x^n\, \mathrm{e}^{-x}\right]_0^\infty + \int_0^\infty dx\, n\, x^{n-1}\, \mathrm{e}^{-x} \tag{2.10.17}$$

The boundary term vanishes at its limits so,

$$I_n = n \int_0^\infty dx\, x^{n-1}\, \mathrm{e}^{-x} = n I_{n-1} \tag{2.10.18}$$

Apply the above recursively to get,

$$I_n = n I_{n-1} = n(n-1) I_{n-2} = n(n-1)(n-2) I_{n-3} = n!\, I_0 \tag{2.10.19}$$

Since $I_0 = \int_0^\infty dx\, \mathrm{e}^{-x} = 1$, we thus have,

$$I_n = n! \tag{2.10.20}$$

Now we will evaluate I_n using the saddle point approximation.

Define $U(x) = -x + n \ln x$. Then,

$$I_n = \int_0^\infty dx\, \mathrm{e}^{U(x)} \tag{2.10.21}$$

The maximum of $U(x)$ is when $dU/dx = -1 + n/x = 0$, or at $\bar{x} = n$. We can therefore expand about this maximum. Using,

$$\begin{aligned}
U(x) &= -x + n \ln x & &\Rightarrow & U(\bar{x}) &= -n + n \ln n \\
U'(x) &= -1 + n/x & &\Rightarrow & U'(\bar{x}) &= 0 \\
U''(x) &= -n/x^2 & &\Rightarrow & U''(\bar{x}) &= -1/n \\
U'''(x) &= 2n/x^3 & &\Rightarrow & U'''(\bar{x}) &= 2/n^2 \\
U''''(x) &= -6n/x^4 & &\Rightarrow & U''''(\bar{x}) &= -6/n^3
\end{aligned} \tag{2.10.22}$$

we expand in $\delta x = x - \bar{x}$ to get

$$U(x) \simeq -n + n \ln n - \frac{1}{2!}\frac{\delta x^2}{n} + \frac{1}{3!}\frac{2\delta x^3}{n^2} - \frac{1}{4!}\frac{6\delta x^4}{n^3} + \cdots \tag{2.10.23}$$

$$= -n + n \ln n - \frac{\delta x^2}{2n} + \frac{\delta x^3}{3n^2} - \frac{\delta x^4}{4n^3} + \cdots \tag{2.10.24}$$

Then,

$$I_n = n! = \int_0^\infty dx\, \mathrm{e}^{U(x)} = \mathrm{e}^{-n + n \ln n} \int_0^\infty dx\, \mathrm{e}^{-\delta x^2/2n}\, \mathrm{e}^{[(\delta x^3/3n^2) - (\delta x^4/4n^3) + \cdots]} \tag{2.10.25}$$

The first exponential term in the integral has the form of a Gaussian with a peak at $x = \bar{x} = n$ (i.e. $\delta x = 0$) and width $\sigma = \sqrt{n}$. As n gets large, the location of the peak in the integrand therefore increases much faster than the width of the integrand, and so we can replace the lower limit of the integration from 0 to $-\infty$.

The argument of the second exponential in the integral consists of terms of order $\delta x^m / n^{m-1}$. Since the width of the integrand is of order $|\delta x| \sim \sigma \sim \sqrt{n}$, these terms are then of order $\sim \sigma^m / n^{m-1} \sim n^{m/2}/n^{m-1} \sim n^{1-m/2}$, where $m \geq 3$. These terms are all therefore small as n gets large, and we can therefore expand the second exponential in the integral.

Using the above, and shifting the integration variable from x to δx, we have,

$$I_n = n! = \mathrm{e}^{-n+n\ln n} \int_{-\infty}^{\infty} d\delta x \, \mathrm{e}^{\delta x^2/2n} \left[1 + \frac{\delta x^3}{3n^2} - \frac{\delta x^4}{4n^3} + O\left(\frac{\delta x^5}{n^4} \right) \right] \tag{2.10.26}$$

$$= \mathrm{e}^{-n+n\ln n} \sqrt{2\pi n} \left[1 + \frac{\langle \delta x^3 \rangle}{3n^3} - \frac{\langle \delta x^4 \rangle}{4n^3} + \cdots \right] \tag{2.10.27}$$

where $\langle \delta x^m \rangle$ represents the average value of δx^m for a Gaussian distribution of mean zero and width $\sigma = \sqrt{n}$. We therefore have $\langle \delta x^3 \rangle = 0$ and $\langle \delta x^4 \rangle \sim \sigma^4 \sim n^2$.

So,

$$I_n = n! = \mathrm{e}^{-n+n\ln n} \sqrt{2\pi n} \left[1 + O\left(\frac{1}{n} \right) \right] \tag{2.10.28}$$

And so,

$$\boxed{\ln n! = n \ln n - n + \frac{1}{2} \ln n + \frac{1}{2} \ln 2\pi + O\left(\frac{1}{n} \right)} \tag{2.10.29}$$

The first two terms on the right give Stirling's formula, the next terms are higher order corrections.