Unit 2-9: Equivalence of the Microcanonical and the Canonical Ensembles in the Thermodynamic Limit

When we introduced the Helmholtz free energy $A(T,V,N)$ in Notes 1-5, it was defined mathematically as the Legendre transform of $E(S,V,N)$, or equivalently, $-A/T$ was the transform of $S(E,V,N)$. In Notes 2-4 we saw that, in ensemble theory, $S(E,V,N)$ is determined from the number of states $\Omega(E,V,N)$ in the *microcanonical* ensemble at fixed energy E. We then called $\Omega(E,V,N)$ the microcanonical partition function. We can therefore denote the Helmholtz free energy that we get by taking the Legendre transform of $S(E,V,N)=k_B\ln\Omega(E,V,N)$ as the *microcanonical* Helmholtz free energy, $A_{\text{micro}}(T,V,N)$.

In our discussion of the *canonical* ensemble in Notes 2-8, we obtained the Helmholtz free energy from the canonical partition function, $A(T,V,N)=-k_bT\ln Q_N(T,V)$. We can denote this as the *canonical* Helmholtz free energy. How do we know that the microcanonical Helmholtz free energy A_{micro} and the canonical Helmholtz free energty A are the same?

In Notes 1-6 we motivated the physical meaning of the Helmholtz free energy by reference to a system in contact with a thermal reservoir, in a discussion similar to that which motivated the canonical ensemble. Here we will show explicitly that $A_{\text{micro}}=A$ in the thermodynamic limit $N\to\infty$.

Energy Fluctuations

Before we show that $A_{\text{micro}}=A$, we first consider the behavior of the energy in the canonical ensemble. In the microcanonical ensemble the system energy E is fixed. In the canonical ensemble the system energy fluctuates, with an average $\langle E\rangle$ that is fixed by the temperature T. In Notes 2-4 we argued, for the case of an ideal gas divided in half by a thermally conducting wall, that such fluctuations become negligible as the size of the system gets large. Here we will explicitly demonstrate this for a system in the canonical ensemble.

Consider the variance of the energy in the canonical ensemble. Using the canonical probability density for the system to have energy E, $\mathcal{P}(E)=\dfrac{\Omega(E)}{\Delta E\;Q_N}\,\mathrm{e}^{-\beta E}$, we have,

$$\mathrm{Var}[E]\equiv\left\langle\left(E-\langle E\rangle\right)^2\right\rangle=\langle E^2\rangle-\langle E\rangle^2=\int dE\,\mathcal{P}(E)\,E^2-\left(\int dE\,\mathcal{P}(E)\,E\right)^2 \tag{2.9.1}$$

$$=\int\frac{dE}{\Delta E}\,\frac{\Omega(E)}{Q_N}\,\mathrm{e}^{-\beta E}\,E^2\;-\;\left(\int\frac{dE}{\Delta E}\,\frac{\Omega(E)}{Q_N}\,\mathrm{e}^{-\beta E}\,E\right)^2 \tag{2.9.2}$$

Consider now the derivative,

$$\frac{\partial\langle E\rangle}{\partial\beta}=\frac{\partial}{\partial\beta}\left(\int\frac{dE}{\Delta E}\,\frac{\Omega(E)}{Q_N}\,\mathrm{e}^{-\beta E}\,E\right)=\int\frac{dE}{\Delta E}\,\frac{\Omega(E)}{Q_N}\,\mathrm{e}^{-\beta E}\,(-E^2)-\left(\int\frac{dE}{\Delta E}\,\frac{\Omega(E)}{Q_N}\,\mathrm{e}^{-\beta E}\,E\right)\frac{1}{Q_N}\frac{\partial Q_N}{\partial\beta} \tag{2.9.3}$$

where the second term follows because $Q_N(T,V)$ depends on T and hence on $\beta=1/k_BT$, and so $\partial(1/Q_N)/\partial\beta=-(1/Q_N^2)(\partial Q_N/\partial\beta)$.

Now from Eq. (2.8.19) of Notes 2-9, we have,

$$\frac{1}{Q_N}\frac{\partial Q_N}{\partial\beta}=\frac{\partial}{\partial\beta}\ln Q_N=-\langle E\rangle \tag{2.9.4}$$

So Eq. (2.9.3) then becomes,

$$\frac{\partial\langle E\rangle}{\partial\beta}=-\langle E^2\rangle+\langle E\rangle^2=-\mathrm{Var}[E] \tag{2.9.5}$$

Therefore,

$$\mathrm{Var}[E] = \langle E^2\rangle - \langle E\rangle^2 = -\frac{\partial\langle E\rangle}{\partial\beta} = -\frac{\partial\langle E\rangle}{\partial(1/k_BT)} = k_BT^2\,\frac{\partial\langle E\rangle}{\partial T} = k_BT^2\,C_V \tag{2.9.6}$$

where $C_V = (\partial E/\partial T)_{V,N}$ is the specific heat at constant volume. This gives a sometime useful formula for the specific heat at constant volume in the canonical ensemble,

$$\boxed{C_V = \frac{1}{k_BT^2}\left[\langle E^2\rangle - \langle E\rangle^2\right]} \tag{2.9.7}$$

The main point is that $\mathrm{Var}[E]$ scales as an *extensive* variable, since E is extensive and $\beta = 1/k_BT$ is intensive. So $\mathrm{Var}[E] \sim N$, where N is the number of particles and is a measure of the size of the system.

Now the width of the probability density $\mathcal{P}(E)$ is given by its standard deviation σ_E,

$$\sigma_E \equiv \sqrt{\mathrm{Var}(E)} = \sqrt{\langle E^2\rangle - \langle E\rangle^2} \sim \sqrt{N} \tag{2.9.8}$$

and the energy is extensive so $\langle E\rangle \sim N$. So the relative width of the probability density, which is a measure of the significance of the energy fluctuations, scales with the system size as,

$$\frac{\sigma_E}{\langle E\rangle} \sim \frac{\sqrt{N}}{N} \sim \frac{1}{\sqrt{N}} \to 0 \qquad \text{as } N \to \infty. \tag{2.9.9}$$

So in the thermodynamic limit $N \to \infty$, the relative fluctuations of the energy in the canonical ensemble vanish. This is one indication that the canonical and the microcanonical ensembles should give equivalent results in the thermodynamic limit.

Equivalence of A and A_{micro}

We now investigate the effect that energy fluctuations have on the canonical Helmholtz free energy A, as compared to the microcanonical Helmholtz free energy A_{micro}.

Microcanonical A_{micro}

To compute A_{micro} we:
1) Compute $S(E,V,N) = k_B \ln\Omega(E,V,N)$ from the microcanonical partition function $\Omega(E,V,N)$.
2) Take the Legendre transform of S with respect to E to get $(-A_{\mathrm{micro}}/T) = S - (E/T)$

We can also write the Legendre transform as follows:

$$\frac{-A_{\mathrm{micro}}(T,V,N)}{T} = \max_E\left[S(E,V,N) - \frac{E}{T}\right] \qquad \Rightarrow \qquad A_{\mathrm{micro}}(T,V,N) = \min_E\left[E - TS(E,V,N)\right] \tag{2.9.10}$$

If $\bar{E}$ is this minimizing value of E, then we have,

$$\boxed{A_{\mathrm{micro}}(T,V,N) = \bar{E} - TS(\bar{E},V,N)} \tag{2.9.11}$$

Canonical A

To compute A we first compute the canonical partition function $Q_N(T,V)$ and then take $A(T,V,N) = -k_BT\ln Q_N(T,V)$.

Consider now the computation of $Q_N = \mathrm{e}^{-A/k_BT}$.

$$Q_N = \mathrm{e}^{-A/k_BT} = \int \frac{dE}{\Delta E}\,\Omega(E,V,N)\,\mathrm{e}^{-E/k_BT} \qquad \text{use } S = k_B \ln \Omega \tag{2.9.12}$$

$$= \int \frac{dE}{\Delta E}\,\mathrm{e}^{S(E,V,N)/k_B}\,\mathrm{e}^{-E/k_BT} \tag{2.9.13}$$

$$= \int \frac{dE}{\Delta E}\,\mathrm{e}^{-[E-TS(E,V,N)]/k_BT} \tag{2.9.14}$$

Consider the exponent $E - TS(E,V,N)$ and expand to second order about its minimum $\bar{E}$. With $E = \bar{E} + \delta E$, and recalling that $(\partial S/\partial E)_{E=\bar{E}} = 1/T$ since $\bar{E}$ minimizes $E - TS$, we have,

$$E - TS(E,V,N) = \bar{E} - TS(\bar{E},V,N) + \delta E - T\left(\frac{\partial S}{\partial E}\right)\bigg|_{E=\bar{E}}\delta E - \frac{1}{2}T\left(\frac{\partial^2 S}{\partial E^2}\right)\bigg|_{E=\bar{E}}\delta E^2 \tag{2.9.15}$$

$$= A_{\mathrm{micro}} + \delta E - T\left(\frac{1}{T}\right)\delta E - \frac{1}{2}T\left(\frac{\partial(1/T)}{\partial E}\right)\bigg|_{E=\bar{E}}\delta E^2 \tag{2.9.16}$$

$$= A_{\mathrm{micro}} + \frac{1}{2T}\left(\frac{\partial T}{\partial E}\right)_{V,N}\delta E^2 = A_{\mathrm{micro}} + \frac{1}{2T}\frac{1}{(\partial E/\partial T)_{V,N}}\delta E^2 \tag{2.9.17}$$

$$= A_{\mathrm{micro}} + \frac{\delta E^2}{2TC_V} \tag{2.9.18}$$

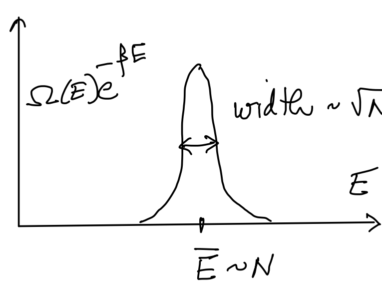
Using Eq. (2.9.18) in (2.9.14), we can now compute,

$$Q_N(T,V) = \mathrm{e}^{-A/k_BT} = \int \frac{d\delta E}{\Delta E}\,\mathrm{e}^{-A_{\mathrm{micro}}/k_BT}\,\mathrm{e}^{-\delta E^2/2k_BT^2C_V} \tag{2.9.19}$$

$$= \mathrm{e}^{-A_{\mathrm{micro}}/k_BT}\int \frac{d\delta E}{\Delta E}\,\mathrm{e}^{-\delta E^2/2k_BT^2C_V} \tag{2.9.20}$$

The integrand in the integral over δE has the form of a Gaussian which is sharply peaked at $\delta E = 0$ (i.e. $E = \bar{E}$) with a width $\sqrt{\langle \delta E^2 \rangle} = \sqrt{k_BT^2C_V}$. Since $C_V \sim N$ and $\bar{E} \sim N$ are both extensive, the relative width of this Gaussian is,

$$\frac{\sqrt{\langle \delta E^2 \rangle}}{\bar{E}} \sim \frac{\sqrt{N}}{N} \sim \frac{1}{\sqrt{N}} \to 0 \qquad \text{as } N \to \infty. \tag{2.9.21}$$



Since the contribution to the Gaussian integral comes almost entirely from the range of δE within a few widths about the peak, and the width is a very small fraction of $\bar{E}$, it is therefore an excellent approximation to take the limits of integration on δE to be $\pm\infty$ (rather than $-\bar{E}$ and $+\infty$) and explicitly do the Gaussian integral.

Recalling the normalization of a Gaussian, $\int_{-\infty}^{\infty} dx\,\mathrm{e}^{-x^2/2\sigma^2} = \sqrt{2\pi\sigma^2}$, we get,

$$Q_N(T,V) = \mathrm{e}^{-A/k_BT} = \mathrm{e}^{-A_{\mathrm{micro}}/k_BT}\,\frac{\sqrt{2\pi k_BT^2C_V}}{\Delta E} \tag{2.9.22}$$

Taking the natural log of both sides then gives,

$$A = A_{\mathrm{micro}} - k_BT\ln\left(\frac{\sqrt{2\pi k_BT^2C_V}}{\Delta E}\right) = A_{\mathrm{micro}} - \frac{k_BT}{2}\ln\left(\frac{2\pi k_BT^2C_V}{(\Delta E)^2}\right) \tag{2.9.23}$$

So

$$\boxed{A - A_{\rm micro} = -\frac{k_BT}{2}\ln\left(\frac{2\pi k_BT^2C_V}{(\Delta E)^2}\right)} \tag{2.9.24}$$

Now since A and $A_{\rm micro}$ are both extensive $\sim N$, and $C_V \sim N$ is also extensive, and $\Delta E \sim E/N$ is intensive, then we have the relative difference in the canonical and the microcanonical Helmholtz free energies,

$$\frac{A - A_{\rm micro}}{A} \sim \frac{\ln N}{N} \to 0 \qquad \text{as } N \to \infty. \tag{2.9.25}$$

We conclude that $A = A_{\rm micro}$, and hence the canonical and the microcanoncial ensembles will give the same thermodynamic results, provided we are in the thermodynamic limit $N \to \infty$.

Because of this equivalence, although we developed the canonical ensemble as a description of a system with fluctuating energy in contact with a thermal reservoir, we could also just as well use it to describe a system at fixed energy E in thermal isolation from its surroundings. One just has to choose the temperature T of the canonical ensemble so that it gives $\langle E\rangle$ as the fixed energy E. If we are in the thermodynamic limit, the fluctuations of E away from $\langle E\rangle$ will be negligible and will not effect the thermodynamic properties.

Note, the approximation we made to evaluate the integral in Eq. (2.9.14), where we expanded the argument of the exponential to second order about its minimum and then did the resulting Gaussian integration, is known as the *saddle point approximation*. We will soon see that Stirling's formula for $\ln N!$ is a result of a saddle point approximation.