1. **An ideal gas on a lattice.** In class we introduced Boltzmann’s microscopic definition of entropy:

\[ S_B = k_B \ln \Omega. \]

A more general expression for entropy due to Gibbs is:

\[ S_G = -k_B \sum_\nu P(\nu) \ln P(\nu), \]

where \( \nu \) is a microstate and \( P(\nu) \) is the probability of that microstate.

(i) For an isolated system (with fixed energy \( E \), number of molecules \( N \), and volume \( V \)) show that \( S_B \) and \( S_G \) are identical. Recall that in such a system, all allowed microstates are equally likely.

Let’s start with the Gibbs form using the uniform distribution over microstates \( P(\nu) = 1/\Omega \).

\[ S_G = -k_B \sum_\nu \frac{1}{\Omega} \ln \frac{1}{\Omega} = k_B \sum_\nu \frac{1}{\Omega} \ln \Omega = k_B \ln \Omega = S_B \]

(ii) For the rest of this problem, consider a collection of \( N \) indistinguishable particles arranged on a lattice of \( M \) cells. Each cell can be occupied by at most one particle, and particles in different cells do not interact. Calculate the total number \( \Omega(M, N) \) of possible configurations for this system.

We need to place \( N \) particles in \( M \) cells. If multiple particles could fit into a single cell we would have \( M^N \) possible ways to arrange the particles, but we said that each cell can only hold a single particle. Therefore we need to choose \( N \) of the \( M \) cells to hold a particle:

\[ \Omega(M, N) = \frac{M!}{N!(M-N)!}. \]

(iii) Assuming \( M, N, \) and \( M-N \) are all very large, use Stirling’s approximation to write the Boltzmann entropy per cell, \( S_B/M \), as a function of \( f \equiv N/M \) alone.

Using Stirling’s approximation for \( M!, \), \( N!, \) and \( (M-N)! \), we get

\[
\frac{S}{k_B} = \ln \Omega \\
= \ln M! - \ln N! - \ln(M-N)! \\
\approx M \ln M - M - N \ln N + N - (M-N) \ln(M-N) + (M-N) \\
= M \ln M - N \ln N - (M-N) \ln(M-N)
\]
We divide through by $M$ and use $N = fM$ to replace all instances of $N$ yielding

\[
\frac{S}{k_B M} = \ln M - f \ln fM - \ln M(1 - f) + f \ln M(1 - f)
\]

\[
= \ln \frac{1}{1 - f} - f \ln \frac{f}{1 - f}
\]

\[
= -(1 - f) \ln(1 - f) - f \ln f
\]

(iv) The occupation state of one cell in this lattice system is not affected by that of any other cell. As a result, the total entropy can be written as $S = MS_{\text{cell}}$, where $s_{\text{cell}}$ is the entropy of a single cell. Using Gibbs’ definition of entropy, calculate $s_{\text{cell}}$ in terms of $p_1$ and $p_0$, the probabilities of finding a particular cell occupied or unoccupied, respectively.

Each cell can be occupied with probability $p_1$ or unoccupied with probability $p_0$, so the Gibbs entropy per cell is particularly simple to compute:

\[
s_{\text{cell}}/k_B = -p_0 \ln p_0 - p_1 \ln p_1.
\]

(v) By expressing $p_0$ and $p_1$ in terms of $f$, the fraction of occupied cells, demonstrate that the Boltzmann and Gibbs lattice gas entropies are the same.

This is simple enough once we note that $p_0 = f$ and $p_1 = 1 - f$ (each site must be either occupied or unoccupied). Inserting into (iv) gives

\[
\frac{S_G}{k_B M} = \frac{s_{\text{cell}}}{k_B} = -f \ln f - (1 - f) \ln(1 - f),
\]

in agreement with the Boltzmann expression in (iii). (You might ponder why the Boltzmann expression relied on the Stirling approximation while the Gibbs expression did not.)

(vi) Show that for the low density lattice gas ($f \ll 1$),

\[
S \approx -k_B V [\rho \ln(\rho v) - \rho],
\]

where $v$ is the volume of a single lattice cell, and $\rho = N/V = N/(Mv)$ is the density.

We first want to Taylor expand our previous results for small $f$. Then we re-express terms using $\rho = N/(Mv)$ and $\nu \rho = N/M = f$.

\[
\frac{S}{k_B M} = -f \ln f - (1 - f) \ln(1 - f)
\]

\[
= f(1 - \ln f) + O(f^2)
\]

\[
= -(\ln \nu \rho - 1) \nu \rho.
\]

Multiplying through by $M$ gives

\[
S = -nk_B [\ln \nu \rho - 1] = -V k_B [\rho \ln \nu \rho - \rho].
\]
(vii) From macroscopic thermodynamics we know that

\[ dS = \frac{1}{T} dE + \frac{p}{T} dV - \frac{\mu}{T} dN, \]

so

\[ \frac{p}{T} = \left( \frac{\partial S}{\partial V} \right)_{E,N}. \]

Use the expression you just obtained for the entropy of a low density lattice gas to explicitly compute the partial derivative. Express your result as the familiar ideal gas law: \( pV = Nk_B T \). (You may be more familiar with \( pV = nRT \). Make sure you think through the distinction between the two forms if it’s not immediately obvious to you.)

You must be a little careful in differentiating \( S \) with respect to \( V \). At first glance it may look like \( S \) is a constant times \( V \), but volume terms are also hidden in the \( \rho \) terms:

\[ S = -k_B V \left[ \frac{N}{V} \ln \frac{N}{M} - \frac{N}{V} \right] = k_B \left[ N \ln \frac{N}{M} - N \right] = -k_B N \left[ \ln \frac{N_v}{V} - 1 \right]. \]

Now we can take the partial derivative.

\[ \frac{p}{T} = \left( \frac{\partial S}{\partial V} \right)_{E,N} = k_B N \frac{V}{N_v} \frac{N_v}{V^2} = k_B N \frac{V}{V}. \]

The difference between this and \( pV = nRT \) is whether you measure the number of particles in moles. That is, \( R = k_B N_A \) with \( N_A \) being Avogadro’s number. You may have seen a different (more mechanical) derivation of an ideal gas law that considers the force imparted by particles hitting a wall, which requires you to consider the momenta of the particles. The power of the current presentation is that it emphasizes how general the ideal gas law really is. It’s not limited to gas molecules flying through vacuum until they collide with a wall. Rather, the “equation of state” \( pV = nRT \) derives from noninteracting degrees of freedom which have statistically independent fluctuation. Anytime you have that decorrelation, you should anticipate something with an ideal gas law structure (though we may call the pressure the osmotic pressure, for example).

(viii) An ideal gas is one with particles which do not interact. Is the lattice gas model whose entropy you found in parts (iii) through (v) an ideal gas? In other words, are there any interactions between the particles? Explain how your answer does or does not agree with your derivation of an ideal gas law in (vii).

No, the gas in the first parts is not ideal. Even though the particles don’t interact at a distance, they do affect each other because only one particle can be in every site. A truly non-interacting situation would have allowed there to be multiple particles at a site. Nevertheless, in the low-density limit, our interacting gas model acts like an ideal gas because the particles are so dilute that there is no real difference between explicitly forbidding them from being on top of each other and allowing that possibility (but knowing that it will essentially never happen when there are so few particles).
2. The climb up.

“[The Boltzmann distribution] is the summit of statistical mechanics, and the entire subject is either the slide-down from this summit, as the principle is applied to various cases, or the climb-up to where the fundamental law is derived.” —Richard Feynman

In class we presented a reasonable but less than rigorous derivation of the Boltzmann distribution,

\[ P(\nu) \propto e^{-\beta E(\nu)}, \]

for microstates \( \nu \) of a system that can exchange energy with a very large bath. This result was obtained from the fundamental relationship

\[ P(\nu) \propto \Omega_B(E_T - E(\nu)) \]

through a Taylor expansion of \( \ln \Omega_B \), where \( E_T \) denotes the total energy shared by system and bath, and \( \Omega_B(E_B) \) is the number of bath microstates with energy \( E_B \). Here you will develop this argument a little more carefully (and perhaps a little more convincingly—you may have wondered, for example, why we were Taylor expanding \( \ln \Omega_B \) and not \( \Omega_B \)).

We will assume that \( \Omega_B \) has a large deviation form,

\[ \Omega_B = [\omega_B(\epsilon_B)]^{N_B}, \]

where \( N_B \) is the number of molecules in the bath, \( \epsilon_B \equiv E_B/N_B \) is the corresponding energy per molecule, and \( \omega_B(\epsilon_B) \) is a smooth function that does not depend on the size of the bath. As in lecture the dependence of \( \Omega_B \) on \( E_B \) will be used to define a property \( \beta \) of the bath:

\[ \beta \equiv \left( \frac{\partial \ln \Omega_B}{\partial E_B} \right)_{N_B,V_B} = \frac{1}{\omega_B(\epsilon_B)} \frac{\partial \epsilon_B}{\partial E_B} = \frac{1}{\omega_B} \frac{\partial \epsilon_B}{\partial E_B} \]

(i) Show that \( \beta \) is insensitive to the extent of the bath. In particular, relate \( \beta \) to \( \omega_B \) and derivatives of \( \omega_B \) with respect to \( \epsilon_B \). Explain why this relationship indicates independence of the bath’s size. (As we will see in future lectures, this \( \beta \) is an inverse temperature, so it should be pleasing \( \beta \) doesn’t depend on bath size. When I say to put a test tube in a 25°C water bath, I shouldn’t have to specify if the bath is 1 liter or 2 liters in volume. Assuming the volume of the bath is very big compared to the system, I should be able to specify \( \beta \) alone.)

\[ \beta = \left( \frac{\partial \ln \Omega_B}{\partial E_B} \right)_{N_B,V_B} = N_B \frac{\partial}{\partial E_B} \ln \omega_B(\epsilon_B) = \frac{N_B}{N_B} \frac{\partial \omega_B}{\partial \epsilon_B} \frac{\partial \epsilon_B}{\partial E_B} = 1 \frac{\partial \omega_B}{\partial \epsilon_B} \]

Notice that all factors of \( N_B \) have canceled. Since \( \omega_B \) is a smooth function that does not depend on the size of the bath, it cannot depend on \( N_B \). Increasing the system size thus yields the exact same value of \( \beta \).

(ii) Show that

\[ \left( \frac{\partial \Omega_B}{\partial E_B} \right)_{N_B,V_B} = \beta \Omega_B \]
By definition,
\[ \beta = \left( \frac{\partial \ln \Omega_B}{\partial E_B} \right)_{N_B,V_B} = \frac{1}{\Omega_B} \left( \frac{\partial \Omega_B}{\partial E_B} \right)_{N_B,V_B}, \]
and multiplying by \( \Omega_B \) on both sides of the equation directly gives the desired result.

(iii) Show that
\[ \left( \frac{\partial^2 \Omega_B}{\partial E_B^2} \right)_{N_B,V_B} = \beta^2 \Omega_B + c. \]
Identify the quantity \( c \) and explain why it can be neglected in the limit \( N_B \to \infty \).

\[
\left( \frac{\partial^2 \Omega_B}{\partial E_B^2} \right)_{N_B,V_B} = \partial \left( \frac{\partial \Omega_B}{\partial E_B} \right)_{N_B,V_B} + \Omega_B \left( \frac{\partial \beta}{\partial E_B} \right)_{N_B,V_B} = \beta^2 \Omega_B + c,
\]
where
\[ c = \Omega_B \left( \frac{\partial^2 \ln \Omega_B}{\partial E_B^2} \right)_{N_B,V_B} = \Omega_B \left( \frac{\partial \beta}{\partial E_B} \right)_{N_B,V_B} = \Omega_B \frac{\partial \beta}{\partial \epsilon_B} \frac{\partial \epsilon_B}{\partial E_B} = \frac{\Omega_B \partial \beta}{N_B \partial \epsilon_B}. \]

Because neither \( \beta \) nor \( \epsilon_B \) depend on the system size \( N_B \), the derivative on the right hand side also does not depend on \( N_B \). The ratio of \( c/(\beta^2 \Omega_B) \) decays as \( 1/N_B \), and hence \( c \) can be neglected in the \( N_B \to \infty \) limit. **Note that \( c \) itself does not vanish in this limit; it simply grows less rapidly than does the first term.**

(iv) Calculate
\[ \left( \frac{\partial^n \Omega_B}{\partial E_B^n} \right)_{N_B,V_B} \]
for arbitrary (integer) \( n \). Do not include any terms that are negligible in the limit \( N_B \to \infty \).

You can continue using the chain rule as in (iii) and observe that all terms involving a derivative of \( \beta \) with respect to \( E_B \) will have contributions which are negligible in the \( N_B \to \infty \) limit. That leaves only
\[ \lim_{N_B \to \infty} \left( \frac{\partial^n \Omega_B}{\partial E_B^n} \right)_{N_B,V_B} = \beta^n \Omega_B. \]
A complete answer to this problem would prove this result by induction.
(v) Consider the Taylor expansion

\[ \Omega_B(E_T - E) = \Omega_B(E_T) - E \left( \frac{\partial \Omega_B}{\partial E_B} \right)_{N_B,V_B} + \frac{1}{2} E^2 \left( \frac{\partial^2 \Omega_B}{\partial E_B^2} \right)_{N_B,V_B} + \ldots \]

\[ = \sum_{n=0}^{\infty} \frac{1}{n!} (-E)^n \left( \frac{\partial^n \Omega_B}{\partial E_B^n} \right)_{N_B,V_B}, \]

where all partial derivatives are implicitly evaluated at \( E = 0 \).

Combining your results together with this expansion, show that the relationship

\[ \Omega_B(E_T - E(\nu)) \propto e^{-\beta E(\nu)} \]

is exact in the limit \( N_B \to \infty \) of an infinitely large bath.

3. The slide down. Remarkably, the decomposition between system and bath (as well as our counting tools with partition functions) work equally well in a quantum setting as a classical setting. Consider a one-dimensional quantum harmonic oscillator, which has equally spaced discrete energy levels: \( \frac{1}{2} \hbar \omega, \frac{3}{2} \hbar \omega, \frac{5}{2} \hbar \omega \ldots \) (Warning: this \( \omega \) has nothing to do with what we were calling \( \omega \) in Problem 2, but the notation is reasonably standard in both cases. I’m sticking with it despite the possibility for confusion.) We assume the harmonic oscillator is in contact with a large thermal bath at temperature \( T \).

(i) We have seen that the probability of occupying each state is given by the Boltzmann distribution:

\[ P(\nu) = \frac{e^{-E(\nu)/k_B T}}{Q}, \]

where \( E(\nu) \) is the energy of microstate \( \nu \) and \( Q \) is a normalization factor, also called the canonical partition function. For the quantum harmonic oscillator we have a relatively simple expression for \( E(\nu) \). Use the equal spacing of energy levels to express the normalization factor \( Q \) as a function of \( T \). [Hint: a geometric series can be summed exactly.]

The canonical partition function is given by

\[ Q = e^{-\beta \hbar \omega /2} \sum_{n=0}^{\infty} \left( e^{-\beta \hbar \omega} \right)^n = \frac{e^{-\beta \hbar \omega /2}}{1 - e^{-\beta \hbar \omega}}. \]
I asked for this as a function of \( T \), not \( \beta \), which is simple enough:

\[
Q(T) = \frac{\exp\left(-\frac{\hbar \omega}{2k_B T}\right)}{1 - \exp\left(-\frac{\hbar \omega}{k_B T}\right)} = \frac{1}{2} \csc h\left(\frac{\hbar \omega}{2k_B T}\right).
\]

You might not have recognized that the expression could be rewritten compactly in terms of the hyperbolic cosecant, which is fine. I’ve chosen to write it this way since it makes the differentiation that much simpler. But you could get to all the right answers if you leave everything in terms of the exponentials.

(ii) By differentiating \( Q \) appropriately, determine the average energy \( \langle E \rangle \), which should also be a function of \( T \).

Notice that \( \ln Q \) is a cumulant generating function, so

\[
\langle E \rangle = -\frac{\partial}{\partial \beta} \ln Q = -\frac{\hbar \omega}{2} \coth\left(\frac{\hbar \omega}{2k_B T}\right).
\]

Or in terms of the exponentials, this would be:

\[
\langle E \rangle = \frac{\hbar \omega}{2} \frac{\exp\left(\frac{\hbar \omega}{2k_B T}\right) + \exp\left(-\frac{\hbar \omega}{2k_B T}\right)}{\exp\left(\frac{\hbar \omega}{2k_B T}\right) - \exp\left(-\frac{\hbar \omega}{2k_B T}\right)}
\]

It’s always a good idea to check that the limiting behavior agrees with what you’d expect. When \( T \) approaches 0, you get \( \langle E \rangle \rightarrow \hbar \omega/2 \), the ground state energy. That’s good! And when \( T \) approaches \( \infty \), the average energy also diverges.

(iii) A common experiment is to measure how much energy must be put into a system to increase the temperature, the so-called heat capacity. In other words, the heat capacity measures the rate of energy increase with increasing temperature:

\[
C = \frac{\partial \langle E \rangle}{\partial T}.
\]

By writing \( \langle E \rangle \) in terms of a derivative of \( \ln Q \), demonstrate that \( C \) is related to a second derivative of \( \ln Q \). Use this observation to compute both the typical size of the energy fluctuations \( \langle \delta E^2 \rangle \) and the heat capacity \( C \). Both of these quantities should be functions of the temperature.

The relationship between heat capacity and energy fluctuations is given by:

\[
C = \frac{\partial \langle E \rangle}{\partial T} = \frac{\partial \langle E \rangle}{\partial \beta} \frac{\partial \beta}{\partial T} = -\frac{\partial^2 \ln Q}{\partial \beta^2} \left(-\frac{1}{k_B T^2}\right) = \frac{\langle \delta E^2 \rangle}{k_B T^2}
\]

Since we have the partition function \( Q \), we can take the two derivatives with respect to \( \beta \) to explicitly compute the energy fluctuations.

\[
\langle \delta E^2 \rangle = \frac{\partial^2}{\partial \beta^2} \ln \left(\frac{1}{2} \csc h\left(\frac{\hbar \omega}{2k_B T}\right)\right) = \frac{(\hbar \omega)^2}{4} \csc h^2\left(\frac{\hbar \omega}{2k_B T}\right).
\]
You may have done this in terms of all of the exponentials, in which case you’d have arrived at

\[ \langle \delta E^2 \rangle = \left( \frac{\hbar \omega}{\exp \left( \frac{\hbar \omega}{2k_B T} \right) - \exp \left( -\frac{\hbar \omega}{2k_B T} \right)} \right)^2. \]

(iv) Now imagine that the quantum harmonic oscillator is actually a three-dimensional harmonic oscillator but that the oscillations in the \( x \), \( y \), and \( z \) dimensions are independent. Then

\[ Q_{3D} = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} \sum_{z=0}^{\infty} e^{-(E_x+E_y+E_z)/k_B T}. \]

By separating out the three sums, relate \( Q_{3D} \) to \( Q_{1D} \) that you computed in part (a). What will be the new values of \( \langle E \rangle \), \( \langle \delta E^2 \rangle \), and \( C \) for the three-dimensional oscillator?

The new cumulant generating function for energy will simply be \( \ln Q_{3D} = 3 \ln Q_{1D} \), with one factor of \( \ln Q_{1D} \) coming from each degree of freedom. As a result, the first two cumulants, \( \langle E \rangle \) and \( \langle \delta E^2 \rangle \), both increase by a factor of 3. Since \( C \) is proportional to the second cumulant, it will also increase by a factor of 3.

\[
\langle E \rangle = -\frac{3\hbar \omega}{2} \coth \left( \frac{\hbar \omega}{2k_B T} \right)
\]

\[
\langle \delta E^2 \rangle = \frac{3(\hbar \omega)^2}{4} \operatorname{csch}^2 \left( \frac{\hbar \omega}{2k_B T} \right)
\]

\[
C = \frac{3(\hbar \omega)^2}{4k_B T^2} \operatorname{csch}^2 \left( \frac{\hbar \omega}{2k_B T} \right)
\]

(v) Repeat your logic from (iv) to determine \( \langle E \rangle \), \( \langle \delta E^2 \rangle \), and \( C \) for \( N \) three-dimensional harmonic oscillators. Einstein considered this as a model for vibrations of the positions of \( N \) atoms around their equilibrium positions in a crystal. At that time, measurements of heat capacity in macroscopic materials could be handled in the lab even if the single-atom vibrations could not be directly measured. By detecting how \( C \) varied with temperature \( T \), Einstein argued that one could infer microscopic information about the nature of the vibrations. Pretty cool!

The same logic applies, but now 3 becomes \( 3N \).

\[
\langle E \rangle = -\frac{3N\hbar \omega}{2} \coth \left( \frac{\hbar \omega}{2k_B T} \right)
\]

\[
\langle \delta E^2 \rangle = \frac{3N(\hbar \omega)^2}{4} \operatorname{csch}^2 \left( \frac{\hbar \omega}{2k_B T} \right)
\]

\[
C = \frac{3N(\hbar \omega)^2}{4k_B T^2} \operatorname{csch}^2 \left( \frac{\hbar \omega}{2k_B T} \right)
\]