STATISTICS:

The goal here is to understand the concept of entropy from the perspective of statistical analysis. Rather than trying to be completely general, we will use a specific example that we will refer to as an Einstein Solid.

Imagine a solid consisting of $N$ 3-dimensional oscillators. These are "atoms" that can vibrate along $x$ or $y$ or $z$. For the motion along any given axis, the energy is quantized according to the rule

$$E_n = (n + \frac{1}{2}) \hbar \omega$$

$n = 0, 1, 2, ...$

At temperature $T = 0$, every oscillator will be in its ground state $E = \frac{1}{2} \hbar \omega$ for a total zero-point energy of $3N \frac{1}{2} \hbar \omega$.

If we add some thermal energy to the solid, some of the atoms will be promoted to states with $n > 0$. We will define the "internal energy" $U$ to be the net vibrational energy of the system minus the zero point energy. So then

$$U = \sum_{i=1}^{3N} n_i \hbar \omega$$

where $n_i$ is the quantum number associated with each vibrational degree of freedom. Notice that the sum runs over all atoms and over all three orthogonal directions.

COUNTING MICROSTATES: Suppose we give some finite energy $U$ to the solid. How may the energy be divided? If $N$ is large and $U$ is many times $\hbar \omega$, there are a very large number of ways for the energy to be shared. Each distinct way of dividing the energy is referred to as a "microstate," and to do a statistical analysis we need to count the...
number of microstates.

Because the energy of the system is quantized we can only give energy in multiples of \( \hbar \omega \) \( \Rightarrow U = q \hbar \omega \) where \( q \) must be an integer. Then \( q \) is the number of units of excitation energy that we have given the system.

Here is how to think up the various microstates we can have. Take \( q \) pieces of paper and make a \( \ast \) on each. Each \( \ast \) represents \( \hbar \omega \) of energy. Then take \( 3N-1 \) pieces of paper and make a \( / \) on each. Now put them all in a hat and start drawing out papers. Suppose we get the following pattern.

\[ \text{\( q+3N-1 \) total} \]

2 atoms
share 10 \( \Rightarrow 5 / 5 \)
10 \( \ast / \ast \)

\[ \ast / \ast / \ast / \ast / / / / / \]

This pattern represents a possible microstate of the system. In this particular microstate, oscillator 1 has 2 units of energy, oscillator 2 has 1 unit, oscillator 3 has 4 units etc. \( \Rightarrow \) \( \text{e.g. atom 1, x motion} \)

\( \text{e.g. atom 1, y motion} \)

\( \text{atom 1 \pm motion} \)

How many distinct (different) microstates can we make? On the next try we will probably draw the papers out in a different order. How many different orders are possible? The answer is "the number of permutations of \( q+3N-1 \) objects" \( \Rightarrow (q+3N-1)! \).

But different permutations may give the same pattern \( \Rightarrow \) the same microstate. We get the same pattern if we exchange any \( \ast \) with another \( \ast \) \( \Rightarrow q! \) permutations or any \( / \) with any other \( / \) \( \Rightarrow (3N-1)! \). So the number of distinct microstates is

\[ \Omega(U,N) = \frac{(q+3N-1)!}{q!(3N-1)!} \]
These multiplicities are big numbers. In a "normal" system N might be $10^{23}$ (like a mole) and $(3 \times 10^{23})!$ is gigantic. Even for "small" numbers of atoms — like 1000 — $\Omega$ is a large number.

**EXAMPLES:** $N = 1000$

- $g = 3000, \quad \Omega = 10^4$
- $g = 6000, \quad \Omega = 10^5$
- $g = 10,000, \quad \Omega = 10^7$

There are two points to make here:

1) $\Omega$'s are big.
2) $\Omega$ always increases as we add energy to the system.

Point 2) is easy to demonstrate using Stirling's approximation.

**EINSTEIN SOLIDS IN CONTACT**

Suppose we take two Einstein solids, both with $N$ atoms, and place them into thermal contact with each other. Then heat could flow from one to the other. Of course we assume that the total thermal energy is conserved.

Now there are many different "macrostates" available to the system [here the macrostate depends only on $U_A$], subject only to the condition

$$U_A + U_B = \text{constant}$$

$$\Rightarrow \quad g_A + g_B = \quad g_{\text{TOTAL}} \quad (\# \text{ of units of energy})$$

Suppose the energy is shared in a particular way with $g_A = \text{some value}$ and $g_B = g_{\text{TOTAL}} - g_A$. How likely is that to happen?

We use the assumption that every microstate is equally likely $\Rightarrow \quad \text{prob}(g_A, g_B) \propto \Omega_{\text{TOTAL}}$, where $\Omega_{\text{TOTAL}}$ is the total multiplicity of the combined system. For example, suppose we have 2 atoms that share $g_{\text{TOTAL}} = 8 \Rightarrow \text{the}$
possible macrostates are \((q^A, q^B) = (8, 0), (7, 1), (6, 2)\)...

How many microstates will we have for, e.g., \((6, 2)\).

\[ q = 6 \text{ hw shared between 2 atoms} \Rightarrow 462 \text{ microstates.} \]

\[ q = 2 \text{ hw} \quad \Rightarrow \quad 21 \]

Total \[ 462 \times 21 = 9702. \]

In general \[ \Omega_{\text{tot}} = \Omega_A \cdot \Omega_B = \frac{(q^A + 3N_A - 1)!}{q^A!(3N_A - 1)!} \cdot \frac{(q^B + 3N_B - 1)!}{q^B!(3N_B - 1)!}. \]

Here is what I get for \(\Omega_{\text{tot}}\):

- \((8, 0) (0, 8) \Rightarrow 1287\)
- \((7, 1) (1, 7) \Rightarrow 4752\)
- \((6, 2) (2, 6) \Rightarrow 9702\)
- \((5, 3) (3, 5) \Rightarrow 14112\)

\((4, 4)\) \[15876\] < most probable way to share.

Let's increase the number of atoms (1 energy units) and try again. In this case I will use

\[ N_A = N_B = 10,000 \quad q_{\text{total}} = 50,000 \]

Then

\[ \Omega(25,000; 25,000) = 10^{32910} \]

\[ \Omega(24,000; 26,000) = 10^{32901} \]

Different by \(q\) orders of magnitude.

So

\[ q^A = 25800 \text{ is a billion times more likely than } 24000. \]

As we make \(N\) and \(q\) larger and larger the calculation of \(\Omega_{\text{tot}}\) becomes more and more sharply peaked and we can be "certain" that the energy \(q\)

will be almost equally shared between the two systems.

Let's do a bit of math to see how the available energy is divided between two systems when \(N_A \neq N_B\). As before \(q_{\text{tot}}\) is the total # of energy units available, and we can
calculate the number of microstates available to the system for a given \( q_A, q_B \) where \( q_B = q_{tot} - q_A \) of course.

\[
\Omega_{tot} = \frac{(q_A + 3N_A - 1)!}{q_A!(3N_A - 1)!} \cdot \frac{(q_B + 3N_B - 1)!}{q_B!(3N_B - 1)!}
\]

Instead of working with \( \Omega_{tot} \) and trying to maximize that quantity lets use \( \log \Omega_{tot} \). In addition lets assume that \( q_A \) and \( 3N \) are very large. So we can then drop the "1" relative to \( 3N \). In addition we can use the Stirling Approximation, according to which

\[
\ln n! = n \ln n - n + \frac{1}{2} \ln (2\pi n) + O\left(\frac{1}{n}\right)
\]

Then

\[
\ln \Omega_A = (q_A + 3N_A) \ln (q_A + 3N_A) - (q_A + 3N_A) - [q_A \ln q_A - q_A] - [3N_A \ln 3N_A - 3N_A]
\]

\[
= (q_A + 3N_A) \ln (q_A + 3N_A) - q_A \ln q_A - 3N_A \ln 3N_A
\]

With this and

\[
\ln \Omega_{tot} = \ln (\Omega_A \cdot \Omega_B) = \ln \Omega_A + \ln \Omega_B
\]

we have (setting \( q_B = q_{tot} - q_A \))

\[
\ln \Omega = (q_A + 3N_A) \ln (q_A + 3N_A) - q_A \ln q_A - 3N_A \ln 3N_A + (q_{tot} - q_A + 3N_B) \ln (q_{tot} - q_A + 3N_B) - (q_{tot} - q_A) \ln (q_{tot} - q_A)
\]

\[
- 3N_B \ln 3N_B
\]

What choice for \( q_A \) gives the maximum \( \Omega \) \( \Rightarrow \) \( \max \ln \Omega \)

\[
\frac{1}{q_A} \ln \Omega = 0 = \ln (q_A + 3N_A) + \frac{q_A + 3N_A}{q_A + 3N_A} - \ln q_A + \frac{q_A}{q_A}
\]

\[
- \ln (q_{tot} - q_A + 3N_B) - \frac{q_{tot} - q_A + 3N_B}{q_{tot} - q_A + 3N_B} + \ln (q_{tot} - q_A)
\]

\[
+ \frac{q_{tot} - q_A}{q_{tot} - q_A}
\]
So we have:

\[
\ln \left( \frac{q_A + 3N_A}{q_A} \right) - \ln q_A - \ln \left( \frac{q_{\text{tot}} - q_A + 3N_B}{q_{\text{tot}} - q_A} \right) + \ln(q_{\text{tot}} - q_A) = 0
\]

\[
\ln \left( \frac{q_A + 3N_A}{q_A} \right) = \ln \left( \frac{q_{\text{tot}} - q_A + 2N_B}{q_{\text{tot}} - q_A} \right) = \ln \left( \frac{q_B + 3N_B}{q_B} \right)
\]

\[
\frac{q_A + 3N_A}{q_A} = \frac{q_B + 3N_B}{q_B} \Rightarrow 1 + 3 \frac{N_A}{q_A} = 1 + 3 \frac{N_B}{q_B}
\]

\[
\frac{q_A}{N_A} = \frac{q_B}{N_B}
\]

The "larger" system gets more of the energy \(-\) in equilibrium, the average energy per atom is the same in both systems.

**ENTROPY:** We are now in a position to finally make a definition of entropy. Let's suppose we start with two separated einstein solids, where one has much more energy per atom than the other. As we bring the solids into contact energy passes from A to B as the system evolves towards the macrostate that has the greatest multiplicity (\( \Omega_A \), goes down \( \Omega_B \), goes up \( \Omega_{\text{tot}} \), goes up) and once this has happened the system will never return to its previous, highly improbable state.

\( \Omega_{\text{tot}} \) only increases, just as \( S \) only increases. The definition we want is

\[
S = k_B \ln \Omega
\]

\(-k_B\) Boltzmann constant \( = 1.38 \times 10^{-23} \text{ J/K} \).

Notice then that for the combined A + B system:

\[
S_{\text{tot}} = k_B \ln \Omega_{\text{tot}} = k_B \ln \Omega_A \cdot \Omega_B = k_B \ln \Omega_A + k_B \ln \Omega_B = S_A + S_B
\]

which is what we would like.

**EXAMPLE:** Suppose we allow heat energy \( Q = 1 \text{ Joule} \) to be conducted from a reservoir at \( 350 \text{ K} \) to one at
250K. The entropy change of the entire system is then
\[
\Delta S = \frac{Q}{T_{\text{cold}}} - \frac{Q}{T_{\text{hot}}} = \frac{1 J}{250 K} - \frac{1 J}{350 K} = 1.14 \times 10^{-3} J/K.
\]

How much has \( \Omega \) changed?
\[
\Delta S = k \ln \frac{\Omega_{\text{final}}}{\Omega_{\text{initial}}} = k \ln \frac{\Omega_{\text{final}}}{\Omega_{\text{initial}}}
\]
\[
\ln \frac{\Omega_f}{\Omega_i} = \frac{(1.14 \times 10^{-3} J/K)}{(1.38 \times 10^{-23} J/K)} = 8.3 \times 10^{19}
\]
\[
\Omega_f = \Omega_i e^{8.3 \times 10^{19}}
\]

Even a tiny \( \Delta S \), say \( \Delta S = 10^{-20} J/K \) corresponds to
\[
\ln \frac{\Omega_f}{\Omega_i} = \frac{10^{-20}}{1.38 \times 10^{-23}} = 725
\]
\[
\Omega_f = \Omega_i e^{725} = \Omega_i \cdot 10^{315}
\]

**EQUILIBRIUM CONDITION:**

Let's suppose we bring two objects to gether and allow them to exchange heat energy until equilibrium is reached. Object A has energy \( U_A \) and object B has energy \( U_B = U_{\text{total}} - U_A \). The equilibrium condition is that heat is exchanged (i.e., \( U_A \) decreased while \( U_B \) increases) until \( \Omega_{\text{TOTAL}} \) is maximized.

\[
\frac{d \Omega_{\text{TOTAL}}}{d U_A} = 0 = \frac{d}{d U_A} \ln \Omega_{\text{TOTAL}} = 0
\]

\[
\frac{d}{d U_A} \left[ \ln \Omega_A + \ln \Omega_B \right] = 0 \Rightarrow \frac{d}{d U_A} \left[ k \ln \Omega_A + k \ln \Omega_B \right] = 0
\]

\[
\frac{d}{d U_A} (S_A + S_B) = 0
\]
We should think of $S_A$ as a function of $U_A$ and $N_A$, and $S_B$ as a function of $U_B$ and $N_B$, but $U_B = U_{\text{tot}} - U_A$ so
\[
\frac{d}{dU_A} S_B = \frac{dS_B}{dU_B} \frac{dU_B}{dU_A} = - \frac{dS_B}{dU_B}
\]
the equilibrium condition is
\[
\frac{dS_A}{dU_A} = \frac{dS_B}{dU_B}
\]
If this condition doesn't hold, $S_{\text{tot}}$ and consequently $\Omega$ could be increased by exchange of energy.

Which way does the heat flow when $\frac{dS_A}{dU_A} \neq \frac{dS_B}{dU_B}$? Both quantities are positive. We always increase $\Omega$ by adding energy. We gain entropy by adding heat to the system that has the steeper slope $\Rightarrow$ heat flows spontaneously from the object with lower $\frac{dS}{dU}$ to the one with higher $\frac{dS}{dU}$. So $\frac{dS}{dU}$ is not the temperature.

But the quantity
\[X = \frac{1}{dU} \ln \frac{dS}{dU}\]
has the property that heat flows from the object with higher $X$ to the object with lower $X$. In fact this will be our definition of $T$
\[T = \frac{1}{dS} \ln \frac{dS}{dU}\]

**EXAMPLE:** Let's do a little exercise to convince ourselves that this is a sensible definition of $T$. For an Einstein solid, we showed previously that
\[\ln \Omega = (g+3N) \ln (g+3N) - g \ln g - 3N \ln 3N\]
when $g$ and $N$ are large. Then the entropy is
\[S = k \ln \Omega\]
and to get the temperature we are supposed to take the derivative with respect to $U$. But for an Einstein solid $U$ is just proportional to $g$ $U = ghw$ and so we
write \( \frac{1}{T} = \frac{dl}{du} = \frac{dl}{dg} \cdot \frac{dg}{du} = \frac{1}{kW} \frac{dl}{dg} \)

\[
= \frac{1}{kW} \frac{d}{dg} \left[ k \left[ (q+3N) \ln(q+3N) - q \ln q - 3N \ln 3N \right] \right]
\]

\[
= \frac{k}{kW} \left[ \ln(q+3N) + \frac{q+3N}{q+3N} - \ln q - \frac{q}{q} \right]
\]

\[
= \frac{k}{kW} \left( \ln(q+3N) - \ln q \right) = \frac{1}{kW} \ln \left( \frac{q+3N}{q} \right)
\]

\[
= \frac{k}{kW} \ln \left(1 + \frac{3N}{q} \right).
\]

At this point lets imagine going to the limit of high temperatures, which means that the atoms have lots of energy \( \Rightarrow q \gg 3N \). Then we can make the approximation that for small \( x \)

\[
\ln(1+x) \approx x \Rightarrow \ln(1 + \frac{3N}{q}) \approx \frac{3N}{q}
\]

and

\[
\frac{1}{T} = \frac{k}{kW} \frac{3N}{q} = \frac{k \cdot 3N}{u} \Rightarrow u = 3NkT
\]

Here \( N = \# \) of atoms = \( n \) \( N_A \) where \( n = \# \) of moles. With this and \( kN_A = R \) we get

\[
U = n \cdot 3R T
\]

lying to an old rule \( C_V = 3R \) (per mole).

\[ \text{CONSISTENCY WITH } \Delta S = \Delta Q / T \]

Notice that our old formula \( \Delta S = \Delta Q / T \) looks roughly the same as our definition of \( T \): \( \Rightarrow \frac{1}{T} = \frac{\Delta S}{\Delta Q} \leftrightarrow T = \frac{\Delta S}{\Delta Q} \), but, of course, \( \Delta Q \) is not always the same as \( \Delta U \). Up to now we have mostly thought of Einstein solids which (we assume) do not change in volume and \( \Delta V \), do no work. So we haven’t needed to distinguish \( \Delta U \) from \( \Delta Q \). But in general—say for an ideal gas—\( \Delta W \) is not always zero.
The following arguments are subtle. Suppose we want to move a system from any macrostate \( \mathcal{A} \) to a new macrostate \( \mathcal{B} \) which is only a tiny distance away, for example in the PV plane. Our goal is to find the entropy change \( \Delta S = S_B - S_A \).

Instead of going directly, we could do two steps. 1) Keep volume fixed and add or remove heat as needed. Here \( \Delta W = 0 \) so \( \Delta U_1 = \Delta Q \), and we have \[ \Delta S_1 = \frac{\Delta S}{\Delta U_1} \Delta U = \left( \frac{\partial S}{\partial U} \right)_V \Delta Q \]

2) Follow an adiabatic curve to point \( \mathcal{B} \). Here \( \Delta U \neq 0 \), but the important point is that an adiabatic process is reversible, so \( \Delta S_2 = 0 \).

Finally, remember that \( S \) is a state variable, so \( \Delta S_{A \rightarrow B} \) is independent of the exact path, and so

\[ \Delta S_{\text{tot}} = \Delta S_1 = \frac{\partial S}{\partial U} \frac{\Delta U}{\Delta Q} = \frac{\Delta Q}{T} \]

[As long as the step from \( A \) to \( B \) is infinitesimal, \( \Delta W_{\text{tot}} \) is the same for the new path as for the original one and so \( \Delta Q_{\text{tot}} \) is also path independent.]

**MAXWELL-BOLTZMANN FORMULA:** There is one more very useful piece of information we can extract from the concepts that we have developed so far. Let's imagine a very large system (like a reservoir) and focus on one very small component of the system -- maybe just one atom or one molecule. The tiny system is in thermal contact with the large system. For example, our small system might be one diatomic molecule in a large container of gas. We imagine that the small system can have various energies -- e.g., \( E_1, E_2, E_3, \ldots \).
different rotational states with energies $E = \frac{\ell(2\pi)^2}{2I}$ or different vibrational states $E_n = (n + \frac{1}{2})\hbar\omega$. Our goal is to find the probability that the system has any particular value of $E$, i.e., find $P(E_n)$. As time goes on, $E_n$ can change as heat flows in or out of the big reservoir.

We make the usual assumption that $P(E_n)$ is the number of microstates of the combined system when the small system has energy $E_n$. Then, $P(E_n) \propto \Omega_n^{(R)} n_R$ where $g_n = \#$ of microstates of the small system and $\Omega_n^{(R)}$ is the number of microstates of the reservoir. But

$$S_R = k_B \ln \Omega_R \Rightarrow \Omega_R = e^{S_R/k}$$

Considering 2 states we have

$$\frac{P(E_2)}{P(E_1)} = \frac{g_2}{g_1} \frac{e^{S_R^{(2)}/k}}{e^{S_R^{(1)}/k}} = \frac{g_2}{g_1} e^{\frac{[S_R^{(2)} - S_R^{(1)}]/k}{k}}$$

Suppose $E_2 > E_1$, then to go from state 1 to state 2 we need to extract heat $dQ = E_2 - E_1$ from $R$ and so the entropy of $R$ is reduced by $dQ/T$. In other words

$$S_R^{(2)} - S_R^{(1)} = -\frac{(E_2 - E_1)}{T}$$

$$P(E_2)/P(E_1) = \frac{g_2}{g_1} e^{\frac{-(E_2 - E_1)/kT}{kT}} = \frac{g_2}{g_1} e^{-E_2/kT}$$

Since this applies for any choice of $E_1 + E_2$ we have

$$P(E_n) \propto g_ne^{-E_n/kT}$$

This is the famous MB formula.