

Statistics + Physics

Kyle Devereaux, BMC March 13, 2024

0.1 Intro to some basic physics

In most intro physics classes in high school and college, one learns about fundamental laws applied to systems with not that many particles. Classic examples are using Newton's second law to solve for a ball rolling down an incline, using Schrodinger's equation to solve for the position of a quantum particle in a box, and using Maxwell's equation to solve how a charge behaves in an electric field.

An underlying theme however is that in these classes, you only work with scenarios with only a few particles - usually just one or two. They are useful examples, but usually we have more to worry about in the real world. Most objects familiar to us in daily life, have on the order of $\sim 10^{23}$ atoms (Avogadro's number). A cup of water has $\approx 8 \times 10^{24}$ molecules and $\approx 2 \times 10^{25}$ atoms. So what if someone asks you why water freezes at 32F? We don't even know what temperature is.

It's not that we are missing more fundamental laws - classical, quantum, and EM physics are sufficient to solve for such information - but it is *very* hard to do. Classical mechanics tells us that knowing the position \mathbf{x}_i and velocity \mathbf{v}_i of every particle in a system to predict its future dynamics exactly. We just need to solve $F = ma$ for every atom in the water to see when they freeze:

$$\mathbf{F}_i(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{10^{23}}) = m \cdot \mathbf{a}_i = m \frac{d^2 \mathbf{x}_i}{dt^2}$$

But then we have a system of $\sim 10^{23}$ coupled differential equations! No computer could solve it exactly. We need a whole different approach.

0.2 Ideas from statistics

Common properties of materials you know, like pressure, viscosity, temperature, are all just *averages* of the properties of its constituents. If you think of a bottle of water, the most basic description each of the particles has some position \mathbf{x}_i and some velocity \mathbf{v}_i , pressure on the bottle is just the average of all the tiny forces each atom of water exerts on the plastic when it collides with the surface. Same with density, which is just the average number of particles per unit volume

$$\rho = \frac{\langle N \rangle}{\delta r^3}$$

Here we use $\langle N \rangle$ to denote the average. These global properties we call "macroscopic" to distinguish them from the "microscopic" properties of the constituents.

0.3 Microstates and macrostates

The microstate is all the information of the microscopic state of the system, in other words, the set of position and velocity (x_i, v_i) for each for the $N = 10^{23}$ particles.

The macrostate is some total coarse-grained property. i.e. the density, number or particles, energy.

Let X be the set of all possible microstate of the system, then if $x \in X$ is one particular microstate, then for some macrostate M , is usually the case that M is a function of x

$$M = M(x)$$

Our end goal: find relations between different macrostates (without having any microstate buisness). We want to derive famous relations like the ideal gas law $PV = NRT$, note that this is a relation between four macrostates and one constant of nature.

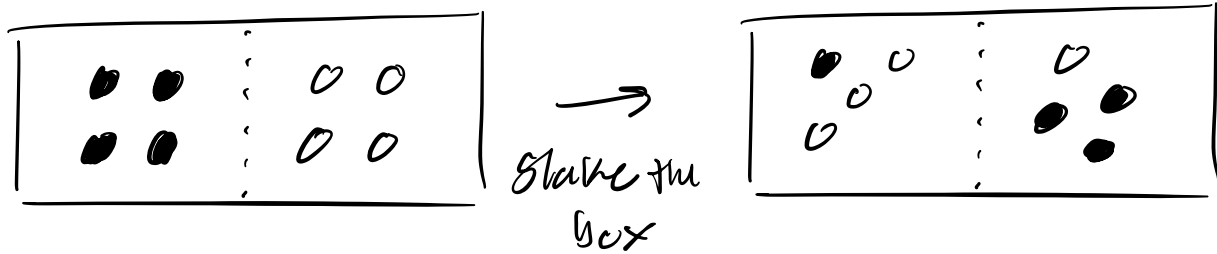
0.4 Microcanonical ensemble

Energy E is an especially important macrostate of "closed" systems - closed meaning the energy of the system doesn't change over time. You can imagine a bunch of particles in a box colliding, with no energy transfer in/out of the box, the total energy E is constant in time, and the microstates of the system are not funtions of time $x = x(t)$. The fundamental statement of the microcanonical ensemble is the probabilities of microstates are given by

$$P(x) = \begin{cases} \frac{1}{g(E)} & H(x) = E \\ 0 & H(x) \neq E \end{cases}$$

Where $g(E)$ is the total number of microstates with energy E and $H(x)$ is the energy of the microstate x . The second case just stays that all microstates of the system, to have non-zero probability, intuitively have to have the same energy as the system. The proof for this is ergodic theory and is a topic for graduate classes.

Example: Box with two halves



Suppose we shake the box and want to measure the microstate,

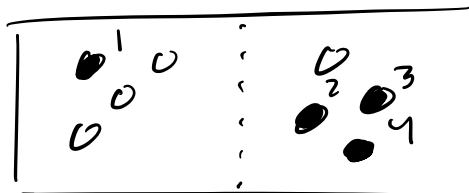
"net change of black balls on left"
$$M = \frac{N_L^B - N_L^W}{2}$$
 correct on average

Before any math, what's your guess? $M=0$?
 If so, we also want to know how often the guess is wrong. aka standard deviation.
 let's find the probability $P(M)$:

Assume shaking totally randomizes everything.
 So each ball 50% chance of ending in left.

label the balls

- 1 L
- 2 R
- 3 R
- 4 R



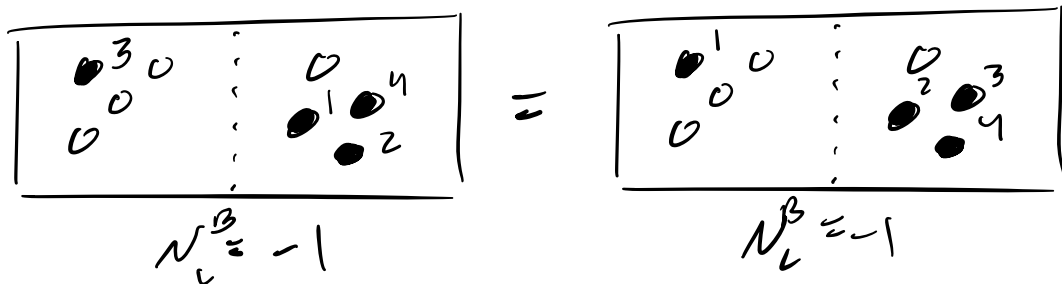
our micro state: $\gamma = \{L, R, R, R\}$

how many configurations? $\rightarrow 2^N$

does each have same probability? Yes $\frac{1}{2^N}$

$$P(x) = \frac{1}{2^N}$$

But we can't double count!



this is because balls are indistinguishable "same up to relabelling"
 $P(N_L^B) : x \mapsto N_L$ is not injective

So for each value of N_L , we need to count the # of microstates:

"multiplicity" $g(N, N_L) =$ number of ways to put N_L black balls on left

what is $g(N, N_L)$?

$$g = \binom{N}{N_L} = \frac{N!}{N_L! (N - N_L)!}$$

$$P(N_L^B) = g(N, N_L) \cdot P(x) = \frac{1}{2^N} \binom{N}{N_L}$$

This is a discrete probability distribution over the variable $0 \leq N_L \leq N$. As a probability distro, it needs to be normalized:

$$\sum_{N_L} P(N_L^B) = 1$$

$$\rightarrow \sum_{N_L=0}^N P(N_L^B) = \sum_{N_L=0}^N \frac{1}{2^N} \binom{N}{N_L^B}$$

$$= \sum_{N_L=0}^N \left(\frac{1}{2}\right)^{N_L} \left(\frac{1}{2}\right)^{N-N_L} \binom{N}{N_L^B}$$

$$\rightarrow = \left(\frac{1}{2} + \frac{1}{2}\right)^N = 1 \quad \text{good}$$

binomial thm: $(x+y)^n = \sum_{k=0}^n x^{n-k} y^k \binom{n}{k}$

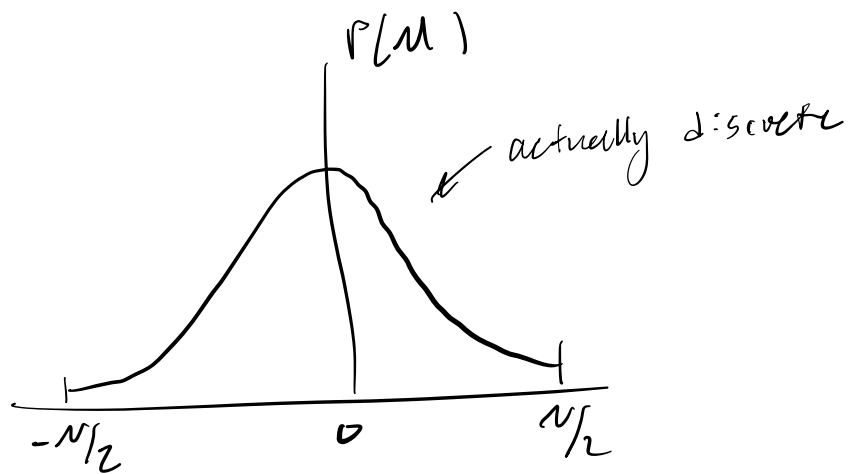
Really we should write $P(N_L) = \frac{1}{2^N} \binom{N}{N_L}$

Now let's write it in terms of macrostate

M :

$$M = \frac{N_L^B - N_L^W}{2} \quad N_L^B = 2M - N_L^W = 2M - \left(M - \frac{N}{2}\right) = M + \frac{N}{2}$$

$$P(M) = P\left(N_L^B = M + \frac{N}{2}\right) = 2^{-N} \frac{N!}{\left(\frac{N}{2} + M\right)! \left(\frac{N}{2} - M\right)!}$$



Now with the probability distro, we just need to use stats to find avg and standard deviation:

$$\text{avg } \langle M \rangle = \sum_m M \cdot P(M)$$

Std is harder. For large N , we can use the Stirling approximation

$$N! \approx \sqrt{2\pi} N^{N+\frac{1}{2}} e^{-N}$$

$$P(M) \approx \frac{1}{\sqrt{\pi N/2}} e^{-2 \frac{M^2}{N}} \quad \leftarrow \text{gaussian}$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{x^2}{\sigma^2}}$$

$$\rightarrow \sigma^2 \approx \frac{N}{4} \quad \sigma = \frac{\sqrt{N}}{2}$$

Relatively, $\frac{\sigma}{N} \propto \frac{1}{\sqrt{N}}$, so for really large N , like $N \sim 10^{23}$, $\langle M \rangle \approx 0$ and $M \approx 0$ very closely.
 "thermo dynamic limit"

Example: spin

Suppose N particles can have one of 2 "spins": one up with energy $-E_0 < 0$, and one down with energy E_0 .

So the total energy for a microstate is

$$H(x) = -E_0 \sum_{i=1}^N \sigma_i$$

where $\sigma_i = +1$ particle i up
 -1 particle i down

microstate: $x = \{\sigma_i\}$

$N_u = \#$ up particles

$N_d = \#$ down particles

$$N = N_u + N_d$$

$$\begin{aligned} \text{Total energy is: } E &= -E_0(N_u - N_d) \\ &= -E_0(2N_u - N) \end{aligned}$$

Multiplicity \rightarrow same as previous example

$$g(N, N_u) = \frac{N!}{N_u!(N - N_u)!}$$

$$N_u = \frac{1}{2} \left(N - \frac{E}{E_0} \right)$$

$$M = \frac{E}{2E_0}$$

$$g(E, N) = \frac{N!}{\left(\frac{N}{2} - \frac{E}{2E_0}\right)! \left(\frac{N}{2} + \frac{E}{2E_0}\right)!} = \frac{N!}{\left(\frac{N}{2} - M\right)! \left(\frac{N}{2} + M\right)!}$$

Taking $N \gg 1$, we used Stirling approximation again:

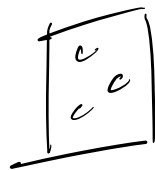
$$g(E, N) \approx \exp[-N(\alpha_- \ln \alpha_- + \alpha_+ \ln \alpha_+)]$$

$$\alpha_{\pm} = \frac{1}{2} \pm \frac{E}{N E_0}$$

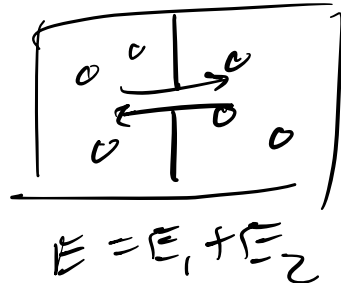
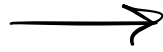
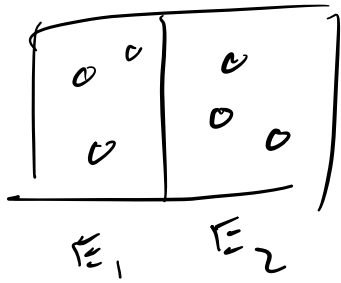
Microcanonical ensemble

$$\rightarrow P(x) = \frac{1}{g(E)} \quad \text{when } H(x) = E$$

Entropy, equilibrium



So far have only considered closed systems. Suppose we have two closed systems with microstates x_1 and x_2 and energies E_1 and E_2 resp. What happens when we connect the systems (i.e. allow energy transfer between the two, remove the wall)? Now only $E = E_1 + E_2$ is conserved.



$$P(x_1, x_2) = P_{E_1}(x_1) P_{E_2}(x_2)$$

$$P_{E_i}(x_i) = \frac{1}{g(E_i)}$$

$$H_i(x_i) = E_i$$

$$P_E(x_1, x_2) = \frac{1}{g(E)}$$

$$H_1(x_1) + H_2(x_2) = E \\ = E_1 + E_2$$

originally the probabilities were independent so $P(x_1, x_2) = P_{E_1}(x_1) P_{E_2}(x_2)$, but now this is not the case, but we still use microcanonical ensemble:

$$g(E) = \sum_{E_1} g_1(E_1) g_2(E - E_1)$$

$$P(E_1) = \frac{g_1(E_1) \cdot g_2(E - E_1)}{g(E)}$$

lets go back to the spin example and now allow the spins to interact with each other.

$$g(E, N) \approx \exp[-N(\alpha_- \ln \alpha_- + \alpha_+ \ln \alpha_+)]$$

Also, we introduce a new quantity to make calculations easier

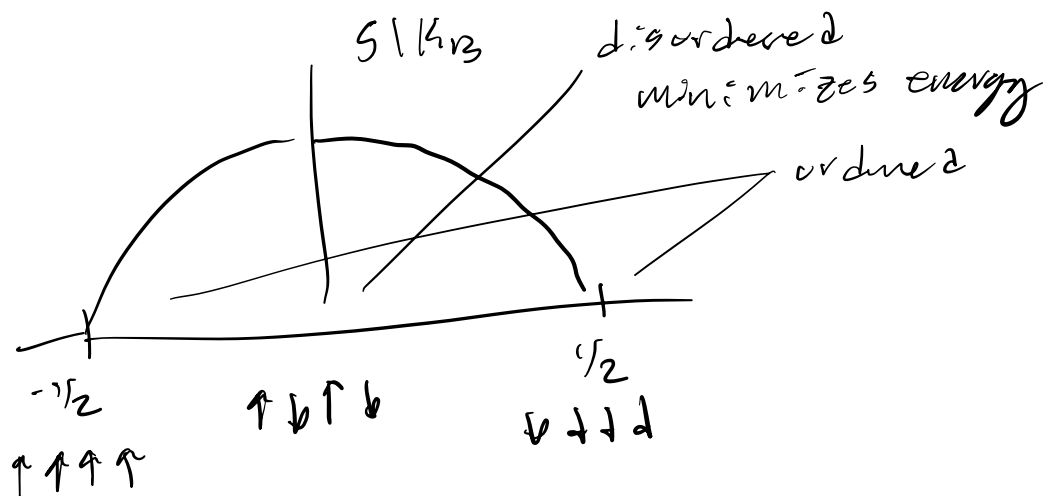
Entropy! $S = k_B \ln(g)$

just a constant, doesn't matter

Removes multiplication in favor of addition.

Physical significance: "disorder" of the system, kinda given by g too.

$$S(E, N) = -k_B N (\alpha_- \ln \alpha_- + \alpha_+ \ln \alpha_+)$$



$$g(E) = \sum_{E_1} g_1(E_1) g_2(E - E_1) = \sum_{E_1} e^{(S_1(E_1) + S_2(E - E_1))/k_B}$$

$$P(E_1) = \frac{g_1(E_1) g_2(E - E_1)}{g(E)} = \frac{\exp(\frac{1}{k_B}(S_1(E_1) + S_2(E - E_1)))}{\sum_{E_1} \exp(\frac{1}{k_B}(S_1(E_1) + S_2(E - E_1)))}$$

what is the most probable energy? \rightarrow

Maximize $P(E_1)$ wrt:

$$\max_{E_1} \frac{\exp\left(\frac{1}{k_B}(S_1 E_1 + S_2 (E - E_1))\right)}{\sum_{E_1} \exp\left(\frac{1}{k_B}(S_1 E_1 + S_2 (E - E_1))\right)}$$

\leftarrow exponent always increases
 \leftarrow constant

$$= \max_{E_1} (S_1 E_1 + S_2 (E - E_1))$$

$$E_1 \frac{dS_1}{dE_1} + (E - E_1) \frac{dS_2}{dE_2} = 0$$

Maximized when $\frac{dS_1}{dE_1} = \frac{dS_2}{dE_2} \rightarrow T_1 = T_2$

This is how temperature is defined!

$$\frac{dS}{dE} = \frac{1}{T} \quad \left. \frac{dS}{dE} \right|_E = \frac{1}{T(E)}$$

So the most likely configuration is when the two temperatures are the same $T_1 = T_2$, — makes sense!

∴ if we have time

$$\frac{1}{T(E)} = \frac{dS}{dE} \quad T(S) = \frac{dE}{dS}$$

$$\underline{dE = T dS} \quad \text{famous}$$

0.5 References

Charles Kittel and Herbert Kroemer, *Thermal Physics (2nd Edition)*, W.H. Freeman 1980