8.044 Lecture Notes
Chapter 4: Statistical Mechanics, via the counting of microstates of an isolated system (Microcanonical Ensemble)

Lecturer: McGreevy

Reading: Greytak, Notes on the Microcanonical Ensemble.
Baierlein, Chapters 2 and 4.3.

4.1   Basic principles and definitions

The machinery we are about to develop will take us from simple laws of mechanics (what's the energy of a configuration of atoms) to results of great generality and predictive power.

A good description of what we are about to do is:
combine statistical considerations with the laws of mechanics applicable to each particle in an isolated microscopic system.

The usual name for this is: "The Microcanonical Ensemble"

Ensemble we recognize, at least. This name means: counting states of an isolated system. Isolated means that we hold fixed

$$N, \text{ the number of particles}$$

$$V, \text{ the volume}$$

(walls can't move and do work on unspecified entities outside the room.)

$$E, \text{ the energy of all } N \text{ particles}$$

Previously E this was called U. It's the same thing.

Sometimes other things are held fixed, sometimes implicitly: magnetization, polarization, charge.

For later convenience, we adopt the following lawyerly posture. Assume that the energy lies within a narrow range:

$$E \leq \text{energy} \leq E + \Delta, \qquad \Delta \lll E .$$

Think of Δ as the imprecision in our measurements of the energy – our primitive tools cannot measure the energy of the system to arbitrary accuracy. This is a technicality, in that all of our results will be independent of Δ. We don't need this lawyering for V.

  This data fixes a *macrostate.*
By a *microstate* we'll mean a specification of the state of all the many constituents of the system, *e.g.* the positions and velocities of a classical ideal gas, or, the quantum states of the individual particles in a quantum ideal gas.

  "Accessible microstates" $\equiv$ those (many) microstates consistent with fixed values of N, V, E.

  Consider an ensemble of copies of the system all with the same macrostate, but in different microstates.

  Now we're going to think about what we mean by equilibrium, from a microscopic point of view. [1]

  Suppose that at $t = 0$ the probability of finding any accessible microstate is equal. That is, every accessible microstate occurs with equal probability.
Now let the laws of mechanics evolve each state in the ensemble forward in time.
Claim: At later times, every accessible microstate still occurs in the ensemble with equal probability.

  $\implies$ **Postulate 1: If** an isolated system is found with equal probability in each accessible microstate, **then** it is in equilibrium.

  Starting with this distribution, it stays that way. This is what we mean by thermal equilibrium.

  Individual microstates in the ensemble are continually changing.
But nothing seen by a macroscopic observer changes.

  It's possible to prove this in both classical mechanics and quantum mechanics with some heavy machinery (Liouville's Theorem). [2]

---

[1]Everything I say between here and **The Fundamental Principle of Statistical Mechanics** is intended as motivation for that Principle. Much of it (the second half) is not rigorous – actually it's worse: it is rigorous in a way which is not physically relevant. If you are unhappy about it, you can ignore it and take the Principle on its own terms as an axiom. Understanding the correct basis of Statistical Mechanics (that is, the actual reason that it successfully describes much of the world) in fact remains a frontier research problem in theoretical physics.

[2](Actually: We haven't shown that there aren't other distributions with this property.)

What about the converse?

Suppose that at $t = 0$ the ensemble contains only a subset of accessible states. For example, suppose all the air molecules are found in half the room.
Evolve each member of this ensemble forward in time.
It is highly unlikely that a system in this ensemble stays in the specified subset of states.

An unnecessarily strong claim is: if we wait long enough, each member of the ensemble will pass through *all* the accessible states. This statement is demonstrably true (and goes by names like "ergodic theorem", "Boltzmann H theorem") for some systems, under some assumptions; however, the fact that systems *equilibrate* is more general.[3]

It is not fully understood, but it is a fact of experience: things equilibrate. We take this as **Postulate 2**:

- If an isolated system is *not* found with with equal probability for each accessible microstate, it is *not* in equilibrium.

- It then evolves to equilibrium. Here by equilibrium, we mean that each accessible microstate is equally probable.

The first of the above statements is the contrapositive of the converse of our **Posulate 1**. It is equivalent to:
If an isolated system is in equilibrium, it is found with equal prob in each of its accessible microstates.

---

[3]Remark which can be ignored: The specific point I am complaining about regarding the assumptions of these theorems is that an estimate of the time it takes for a system to explore *all* of the accessible microstates (in the classical case, one asks how long it takes to come within a fixed small distance from any given point in the phase space) is much longer than it takes in practice for systems to be well-described by equilibrium thermodynamics.

Combining these two statements, we get:

Fundamental Postulate of Statistical Mechanics:

An isolated system is in equilibrium iff all accessible microstates are equally probable.

Claim: All the rest of 8.044 follows from this box. We could have begun the class here if we wanted to be deductive (not always the best way to learn). From it we can derive:
• relationships among macroscopic variables *e.g.* $PV = NkT, U = \frac{3}{2}NkT$ 2nd Law, Curie law for paramagnets...

• probability densities for microscopic quantities (*e.g.* velocity distribution for an ideal gas in equilibrium.)

[End of Lecture 10.]

This will be our starting point for deriving thermodynamics and other things. Whether it is an accurate description of, say, a series of snapshots of a real system which we observe to be an an equilibrium macrostate is not something we are prepared to address. Rather, it is a useful starting point for encoding our partial knowledge about such a system.

A more Bayesian perspective on what we are doing is: we want to encode our partial knowledge about the microstate in a probability distribution on the phase space (this is a name for the space of microstates). All we know is the macroscopic variables E, V, N. The distribution which is consistent with this information but introduces the smallest amount of extraneous, *wrong* information is the uniform distribution on the accessible states. This perspective has been forcefully advocated by E. Jaynes; we will come back to it again later.

Now, in equations:

Quantum version (is easier to state than classical)

In equilibrium, the probability that the system is in state k is

$$
p(k) = \begin{cases} \frac{1}{\Omega}, & \text{if } E \leq \underbrace{\langle k|H|k \rangle}_{\text{energy of state } k} \leq E + \Delta \\[2em] 0, & \text{else.} \end{cases}
$$

Note that $|k\rangle$ needn't be an energy eigenstate.[4] Ω determined by normalization: *count* accessible states:

$$
\Omega = \sum_{\text{accessible } k} 1 = \Omega(E, V, N)
$$

Note: Ω is dimensionless.

In classical mechanics, this counting becomes integration; we get a probability density for x_i, p_i.

---

[4]A tricky point which you should ignore on a first pass (and which you should ignore completely if you haven't taken QM):

You might object to this statement that p is a probability, and not a probability density. What about superposition? That is: how do we determine the probability of being in a state $a|k_1\rangle + b|k_2\rangle$? a, b here are continuously variable, so to construct probabilities we would need a probability density.

I haven't told you how to answer this question; it's harder. The formula above is the answer to a more restricted question. We've *assumed* here that we have picked a basis of the Hilbert space labelled by the quantum number k, and we know that the state is one of these basis vectors. The answer to the more general question, which alas we will not discuss in 8.044 requires *density matrices*.

Classical version:

For N particles in three dimensions:

$$p(\{p,q\}) = \begin{cases} \frac{1}{\Omega}, & \text{if } E \leq \underbrace{H(\{p,q\})}_{\text{energy of state with } \{p,q\}} \leq E + \Delta \\ \\ 0, & \text{else.} \end{cases}$$

This p is now a probability density for $6N$ variables. Those variables are $\{p,q\}$, a specific set of N positions (qs) and N momenta (ps) in three dimensions.[5] The space of $\{p,q\}$ is called *phase space*.

The normalization factor for our phase-space probability density is:

$$\Omega = \Omega(E,V,N) \equiv \int d^{3N}p \, d^{3N}q \; 1 = \int_{\text{values with specified } E} dp_{1x} dp_{1y} dp_{1z} dp_{2x} ... dx_1 dy_1 dz_1 dx_2 ... 1$$

$$units: \qquad [\Omega] = \text{length}^{3N} \cdot \text{momentum}^{3N}$$

When this distribution arises as a limit of quantum mechanics, it comes with a factor of h^{3N}
($[h] = \text{length} \cdot \text{momentum} = \text{action}$)
which makes the result dimensionless. This constant factor has no effect on anything (neither thermodynamics nor microscopic probabilities).

Note that the integrand of these integrals is very boring: 1. All of the physics is in the limits of integration.

Yes, we are using p to denote both probability density and momentum. There are only so many letters.

---

[5]Some of you may note a signficant ambiguity here: if I change coordinates on phase space, the probability distribution transforms like a distribution (recall 'functions of a random variable'); if p is uniform with respect to one choice, it will in general not be uniform with respect to others. This problem simply doesn't exist in the quantum version. The resolution is that p is uniform in certain special sets of coordinates, namely when $\{p,q\}$ are canonically conjugate variables. This is the case which arises when classical mechanics arises as a (large-quantum-numbers) limit of quantum mechanics, as indeed it always does.

Squashing

How do we extract information from this formal distribution? First microscopic data, then thermodynamics.

Suppose we are interested in what one of the particles is doing, *e.g.* in its momentum distribution in the x direction:

$$p(p_{1x}) = \int_{\text{all variables except } p_{1x}} p(\{p,q\}) = \frac{1}{\Omega} \int_{\text{accessible values of all variables with } p_{1x}, V, E, N \text{ fixed}} 1$$

The integrand is still 1; all information is in the limits of integration.

More generally: if we pick X to label a special set of states of the system in which some subset of $\{p, q\}$ are fixed to satisfy some condition (*e.g.* all the x coordinates are fixed to be in the northern half of the room), then

$$
\begin{aligned}
p(X) &= \frac{1}{\Omega} \int_{\text{accessible values of all un-fixed coordinates}} 1 \\[2mm]
&= \frac{\Omega'(\text{consistent with } X)}{\Omega} \\[2mm]
&= \frac{\text{volume of accessible phase space consistent with } X}{\text{volume of accessible phase space}}
\end{aligned}
$$

The prime on Ω' is to distinguish this quantity from the total Ω. In cases (QM) where the states are discrete, this is:

$$p(X) = \frac{\text{how many states satisfy the given condition } X}{\text{how many states altogether}}$$

Now you see why I asked to you to count ways to get various poker hands.

Def: A useful quantity is the "cumulative probability in phase space":

$$\Phi(E, V, N) \equiv \int_{\text{energy}(p,q) \leq E} 1$$

Then the cumulative probability and the probability density are related in the usual way, by differentiating. Our lawyering with the energy window Δ makes an appearance here. If we define the density of microstates at energy E to be ω:
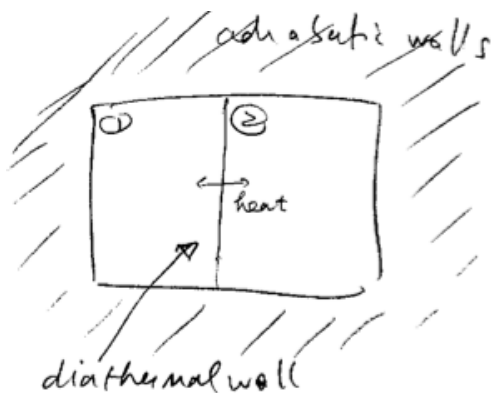
$$\omega(E, V, N) \equiv \partial_E \Phi(E, V, N)$$

Then

$$\Omega(E, V, N) = \omega(E, V, N)\Delta \; .$$

Ω is a probability (that the energy is in the window of width Δ), and ω is a probability density.

4.2   Temperature and entropy

Consider a situation where two systems separated by a diathermal wall are isolated from the rest of the universe:

The total system $\boxed{1} + \boxed{2}$ is isolated.
The total energy E is (therefore) fixed.
The walls don't move: V is fixed, no work is done.

Call the energy of system $\boxed{1}$ E_1. The energy of system $\boxed{2}$ is then $E_2 = E - E_1$.

E_1 varies between elements of the ensemble.

Our job here is to answer the Q:
What is the most probable value of E_1 in thermal equilbrium?
This is a job for the **Fundamental Postulate of Statistical Mechanics**:

$$p(E_1) = \underbrace{\frac{\Omega'\left(\text{system } \boxed{1} \text{ has energy } E_1\right)}{\Omega}}_{\text{equal probability for all accessible microstates}} = \underbrace{\frac{\Omega_1(E_1)\Omega_2(E - E_1)}{\Omega(E)}}_{\text{systems } \boxed{1} \text{ and } \boxed{2} \text{ are independent except for total energy}}$$

In the second step here we enumerated the states of the whole system as a product of the number of states of the two systems.

Maximize $\ln p(E_1)$:

$$0 = \partial_{E_1} \ln p(E_1) = \partial_{E_1}\left(\ln \Omega_1(E_1) + \ln \Omega_2(E - E_1) - \ln \Omega(E)\right)$$

$$\implies \quad 0 = \partial_{E_1} \ln \Omega_1(E_1) + \underbrace{\frac{\partial E_2}{\partial E_1}}_{= -1} \partial_{E_2} \ln \Omega_2(E_2)$$

At the E_1 at which $p(E_1)$ is extremized:

$$\implies \qquad \underbrace{\partial_{E_1} \ln \Omega_1(E_1)}_{\text{a property of system } \boxed{1} \text{ only}} \quad = \quad \underbrace{\partial_{E_2} \ln \Omega_2(E_2)}_{\text{a property of system } \boxed{2} \text{ only}}$$

This is a property of the systems which is equal when $\boxed{1} + \boxed{2}$ are in thermal equilibrium, and which depends only on $\boxed{1}$ or $\boxed{2}$. This must therefore be a function of T.

Consistency with conventions requires the following
Statistical mechanics definition of T:

$$\boxed{\frac{1}{k_B T} = \left(\frac{\partial}{\partial E} \ln \Omega(E) \right)_{\mathchar'26\mkern-10mu dW=0}} \tag{1}$$

$\mathchar'26\mkern-10mu dW = 0$ means (as above) we're not doing work on the system: constant V, $\vec{M}$.

Consistent with 0th Law.

Why $1/T$? Because this definition gives the ideal gas law, that is, it is the same as thermodynamic temperature, defined above. Soon.

Call the equilibrium value of E_1 E_1^{eq}. This is the value of E_1 at which the system has the most microstates available to it.

We've shown that the probability distribution $p(E_1)$ has a critical point. Why should it be a maximum? For most reasonable systems (imagine a gas), $\Omega_1(E_1)$ is a rapidly rising function of E_1: the more energy it has, the more states are accessible to it. On the other hand, $\Omega(E_2 = E - E_1)$ rises rapidly with E_2 and therefore *falls* rapidly in E_1 with E fixed.

This is a high, narrow, gaussian peak. Width is $\sqrt{N}^{-1}$.
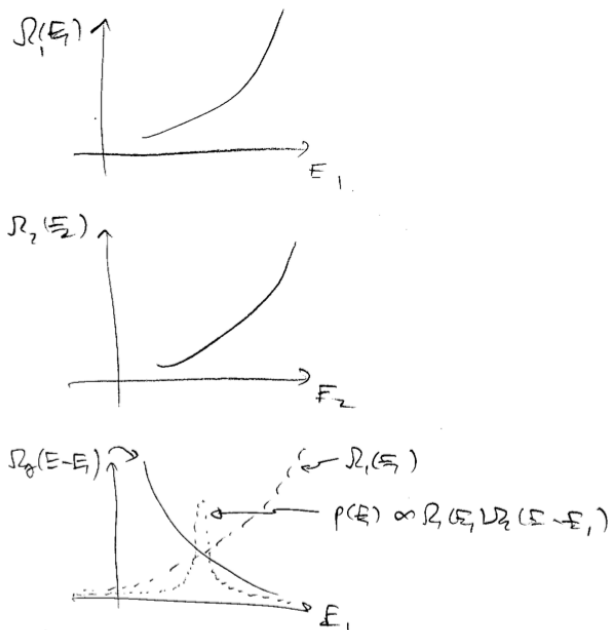
Entropy and the 2nd Law

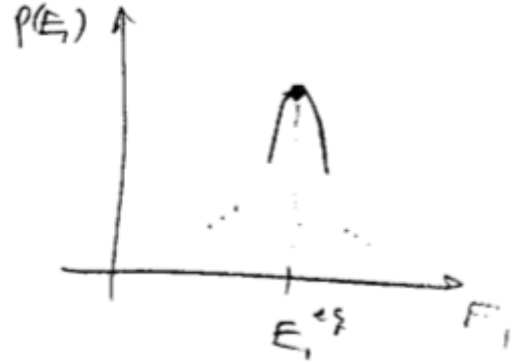Q: What will happen if we prepare the system with some $E_1 \neq E_1^{eq}$?

The postulates above imply that
energy will flow between $\boxed{1}$ and $\boxed{2}$
until $E_1 = E_1^{eq}$. *i.e.* it equilibrates.

E_1^{eq} is a maximum implies that

$$p(E_1) \;\leq\; p(E_1^{eq})$$

$$\frac{\Omega_1(E_1)\Omega_2(E - E_1)}{\Omega(E)} \;\leq\; \frac{\Omega_1(E_1^{eq})\Omega_2(E - E_1^{eq})}{\Omega(E)}$$

rearrange: $\qquad 1 \;\leq\; \dfrac{\Omega_1(E_1^{eq})}{\Omega_1(E_1)} \dfrac{\Omega_2(E - E_1^{eq})}{\Omega_2(E - E_1)}$

Entropy and the 2nd Law, cont'd.

This statement will look more familiar if we take its logarithm, and define (the microcanonical version of) *entropy*:

$$\boxed{S(E, V, N) \equiv k_B \ln \Omega(E, V, N)}$$

With this definition of entropy, we have (log is a monotonic function so the inequalities are preserved):

$$0 = k_B \ln 1 \leq S_1(E_1^{eq}) - S_1(E_1) + S_2(E - E_1^{eq}) - S_2(E - E_1)$$
$$\implies \quad \Delta S_1 + \Delta S_2 \geq 0$$

(where $\Delta S_{1,2}$ are the changes in entropy during the equilibration process)

$$\implies \quad \boxed{\Delta S_{\text{total}} \geq 0} \qquad \text{(2nd Law)}$$

If we start at $E_1 \neq E_1^{eq}$, the total entropy will increase with time.

Now we see that the equilibrium value of E_1 maximizes $S_{\text{total}} = S_1 + S_2$.
At the equilibrium value: • $p(E_1)$ is maximized, • $S(E_1)$ is maximized, • $T_1 = T_2$.

2nd Law: The total entropy of an isolated system always increases as it approaches equilibrium. In equilibrium S is constant in time (as are all thermodynamic variables) and is maximized.
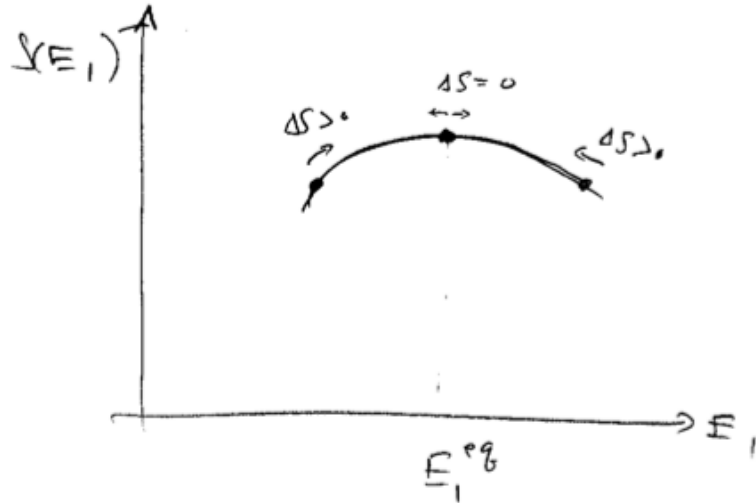
Basic facts about entropy:

• Entropy is a state function. It is manifestly so, given the simple definition: Given a macrostate, count the allowed microstates and take the log.

• Entropy measures (the logarithm of) the number of accessible microstates. Information theory point of view: it's a measure of our ignorance of the microstate given the macrostate.

• Entropy of a composite system is additive:

$$\Omega_{\text{total}} = \Omega_1 \Omega_2 \quad \Longrightarrow \quad S_{\text{total}} = S_1 + S_2$$

Therefore S is extensive. (This is one reason to take the log.)

Claim: $S = S(E, V, N, \vec{M})$. It's defined as a function of all the other *extensive* variables.

• The expression for temperature (1) is now:

$$\boxed{\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{\bar{d}W=0}}$$

Temperature is positive $T > 0$ means that S increases with E.

• 2nd Law: $\Delta S \geq 0$ for spontaneous processes in an isolated system. $\Delta S = 0$ obtained if all processes are reversible.

• <u>Equilibration</u> of an isolated system <u>means</u> evolution to a macrostate <u>maximizing entropy</u>.

Note that when two bodies at different temperatures are brought into contact, our statistical notions show that heat flows from the hotter body to the colder body. This follows since the change in entropy is

$$0 \leq \Delta S = \left(\left(\frac{\partial S_1}{\partial E_1} \right)_{i_1} - \left(\frac{\partial S_2}{\partial E_2} \right)_{i_2} \right) \Delta E_1 = \left(\frac{1}{T_1^i} - \frac{1}{T_2^i} \right) \Delta E_1$$

where $T_{1,2}^i$ are the initial temperatures.

Two lacunae at these point in our discussion:
• We need to show that our definition of temperature agrees with the ideal gas law.
• We need to make a connection with the macroscopic notion of entropy $dS = \frac{\bar{d}Q}{T}$.

4-13

Thermodynamic entropy and statistical entropy are the same.

Consider the limit where $\boxed{1}$ is tiny compared to $\boxed{2}$:

Then $T_2 \equiv T_{\text{reservoir}}$ does not change as heat is exchanged between $\boxed{1}$ and $\boxed{2}$. As before, the whole system $\boxed{1} + \boxed{2}$ is isolated.
And no work is done today – all the walls are fixed.

Now, the definition of T applied to system $\boxed{2}$ is:

$$\left(\frac{\partial S_2}{\partial E_2}\right)_{dW=0} = \frac{1}{\underbrace{T_{\text{reservoir}}}_{\text{this is constant}}}$$

$$dS_2 = \frac{dE_2}{T_{\text{reservoir}}} \qquad \text{Now note: } V \text{ is fixed. No work is done.} \quad dE_2 = dQ_2$$

$$= \frac{dQ_2}{T_{\text{reservoir}}} \qquad \leftarrow dQ_2 \equiv \text{heat entering } \boxed{2}$$

$$= -\frac{dQ_1}{T_{\text{reservoir}}} \qquad dQ_1 = -dQ_2 \quad \text{(since the whole system is isolated)}$$
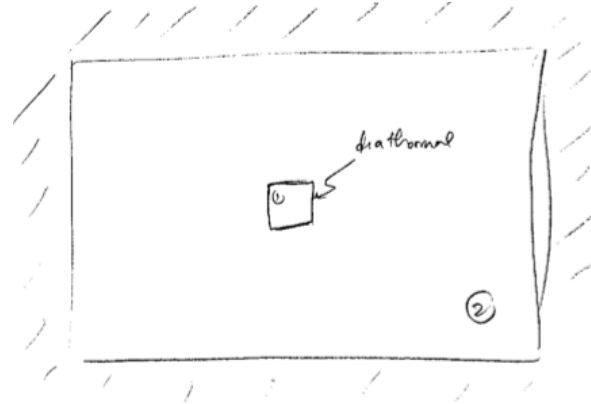
2nd Law: $dS = dS_1 + dS_2 \geq 0$

$$\Longrightarrow \qquad \boxed{dS_1 \geq \frac{dQ_1}{T_{\text{reservoir}}}}$$

This is the form of the 2nd Law valid when exchanging heat with a reservoir.
Equality occurs if the heat is exchanged quasistatically:

$$dS_1 \overset{\text{quasistatic}}{=} \frac{dQ_1}{T}\Big|_{\text{quasistatic}}$$

Note: for quasistatic processes, $T_1 = T_{\text{reservoir}} = T$.

[End of Lecture 11.]

Notice the following thing about doing things quasistatically:
the change in the entropy of the *whole* system ($\boxed{1} + \boxed{2}$) is *zero* in this case. The change in the entropy of system $\boxed{1}$ is determined by the heat exchanged with $\boxed{2}$. That is:

$$\text{quasistatic process of isolated system:} \quad dS = 0$$

$$\text{quasistatic process of system in contact with reservoir:} \quad dS = \frac{dQ}{T} \quad .$$

Now we can rewrite
First Law of Thermodynamics for Quasistatic Processes as:

$$dU \; = \; đQ + đW \qquad \text{\color{green}{this is the general version}}$$

$$= \; TdS - PdV \qquad \text{\color{green}{here we assume quasistatic. else inequality.}}$$

Instead of $-PdV$ here could be any other way of doing work $\sigma dA + f dL + \vec{H}d\vec{M} + ...\mathcal{F}dX$. Here I've added a generic extensive variable X, with its associated conjugate (intensive) force variable $\mathcal{F}$, so $đW = ... + \mathcal{F}dX$.
The TdS term is always there. Solve for dS:

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{H}{T}dM - ... - \frac{\mathcal{F}}{T}dX$$

is an exact differential.

S is a state function, $S = S(E, V, N)$

$$dS = \underbrace{\left(\frac{\partial S}{\partial E}\right)_{V, N}}_{đW=0} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial \vec{M}}\right)_{E,V,N} \cdot d\vec{M} + ...$$

If there are other ways of doing work on the system, there are more terms here.

Now compare this to our stat mech definition of temperature:

$$\frac{\partial S}{\partial E}\Big|_{đW=0} = \frac{1}{T}$$

$$\frac{\partial S}{\partial V}\Big|_{E,N,\vec{M}} = \frac{P}{T}$$

$$\frac{\partial S}{\partial \vec{M}}\Big|_{E,V,N} = -\frac{\vec{H}}{T}$$

For any extensive variable X, with associated conjugate (intensive) force variable $\mathcal{F}$, so $đW = ... + \mathcal{F}dX$:

$$\frac{\partial S}{\partial X}\Big|_{\text{all other extensives fixed}} = -\frac{1}{T}\mathcal{F}$$

Program for the rest of 8.044:

- Pick a system

- Count microstates. This gives $\Omega(E, V, N...)$.

- Take the log:
$$S(E, V, N) = k_B \ln \Omega.$$

- Take derivatives of S to get all intensive variables:
$$T^{-1} = \frac{\partial S}{\partial E}\Big|, \quad P = T\frac{\partial S}{\partial V}\Big|, ...$$

The hard part is the counting of microstates.

Next: for ideal gas, we can actually do the counting.

4.3 Classical monatomic ideal gas

Here we will see how the formal structure introduced above actually works for a classical monatomic ideal gas. This means $N \sim 10^{24} \gg 1$ featureless classical particles (which do not interact with each other) in a box with sides L_x, L_y, L_z. So the volume is $V = L_x L_y L_z$.

$6N$ microscopic degrees of freedom:

$$3N \ q_i's: \quad x_i, y_i, z_i \text{ for } i = 1...N \text{ atoms}$$

$$3N \ p_i's: \quad p_{xi}, p_{yi}, p_{zi} \text{ for } i = 1...N \text{ atoms}$$

$$0 \le x_i \le L_x, \quad 0 \le y_i \le L_y, \quad 0 \le z_i \le L_z.$$

The ps can go from $-\infty$ to ∞.

For a (non-relativistic) *ideal* gas:

$$\text{Energy of microstate with } \{p, q\} = H(\{p, q\}) = \sum_{i=1}^{3N} \frac{\vec{p_i}^2}{2m} \ .$$

Comments:

- Notation: $\vec{p_i}^2 \equiv p_{xi}^2 + p_{yi}^2 + p_{zi}^2$. I may sometimes omit the little arrow. So the sum in H is over all $3N$ momenta.

- The energy doesn't care about where the atoms are – H is independent of the q_i. In a non-ideal gas there would be extra terms depending on $q_i - q_j$ – that is, interactions.

- Here, there is only translational kinetic energy. No contribution from *e.g.* molecular vibrations or rotations. Hence, monatomic gas. Later we will consider diatomic gases which have more microscopic variables.

- This is all the input from mechanics, very simple.

The reasoning is also simple – we just have to count.

$$\Omega = \Omega(E, V, N) \quad = \quad \int_{\underbrace{\text{accessible microstates}}_{\text{all the physics is hidden in here}}} dq_1 dq_2 ... dq_{3N} dp_1 dp_2 ... dp_{3N} \cdot 1$$

(The condition for accessibility doesn't care about the qs.)

$$= \quad \left(\int_0^{L_x} dx \right)^N \left(\int_0^{L_y} dy \right)^N \left(\int_0^{L_y} dz \right)^N \int_{E \leq H(\{p\}) \leq E+\Delta} dp_1 ... dp_{3N} \cdot 1$$

$$= \quad \left(\underbrace{L_x L_y L_z}_{V} \right)^N \int_{E \leq H(\{p\}) \leq E+\Delta} dp_1 ... dp_{3N} \cdot 1$$

It will be more expedient to do the cumulative count:

$$\Phi(E, V, N) = V^N \int_{H(\{p\}) \leq E} dp_1 ... dp_{3N} \cdot 1.$$

Now recall that $H(\{p\}) = \frac{1}{2m} \sum_{i=1}^{3N} p_i^2$, so the condition $H(\{p\}) \leq E$ defines a

$$3N\text{-dimensional } \textit{ball}: \quad \{p | \sum_i p_i^2 \leq 2mE\}$$

(the interior of a $3N - 1$-dimensional sphere) of radius $R = \sqrt{2mE}$.

Math fact: The volume of a $3N$-dimensional ball of radius R is

$$v_{3N} = \frac{\pi^{3N/2}}{\left(\frac{3N}{2}\right)!} R^{3N}$$

What I mean by $x!$ here is defined by the recursion relation $x! = x(x-1)!$ with the initial conditions

$$1! = 1 \text{ and } \frac{1}{2}! = \frac{1}{2}\sqrt{\pi}.$$

[6] Note that Stirling's formula still works for non-integer x. Check that the formula for v_{3N} works for $N = 1$.

---

[6]For x which are not half-integers the right generalization is the Gamma function $x! = \Gamma(x + 1) \equiv \int_0^\infty ds\, s^x e^{-s}$. We won't need this in 8.044.

$$\boxed{\Phi(E, V, N) = \frac{V^N \pi^{3N/2}}{\left(\frac{3N}{2}\right)!} (2mE)^{3N/2}}$$

Now use Stirling:

$$\ln((3N/2)!) \simeq \frac{3N}{2} \ln \frac{3N}{2} - \frac{3N}{2}$$

$$\Longrightarrow \qquad \frac{3N}{2}! \simeq \left(\frac{3N}{2e}\right)^{3N/2}$$

(here $e = 2.7181...$)

$$\Phi = V^N \left(\frac{4\pi emE}{3N}\right)^{3N/2}$$

$$\omega = \frac{d\Phi}{dE} = \frac{3N}{2} \frac{1}{E} V^N \left(\frac{4\pi emE}{3N}\right)^{3N/2}$$

$$\Omega(E, V, N) = \omega\Delta = \frac{3N}{2} \frac{\Delta}{E} V^N \left(\frac{4\pi emE}{3N}\right)^{3N/2}$$

This result is wrong for two reasons, one clerical and one important.

- The clerical reason is the convention about the normalization of Ω. The Ω we wrote above is dimensionful. To make Ω dimensionless, we divide by $h^{3N} = (2\pi\hbar)^{3N}$. This factor had better not matter in any thermodynamic observable. (Notice that our lawyering factor Δ also multiplies the answer for Ω.) This will provide a small check on our answers.

- The important thing we left out is that the atoms are *identical*. The counting we have done above treats the atoms as if they are wearing nametags which we could use to distinguish which atom is which, and then we could say things like "Oh yeah, that's Bob over there. His kinetic energy is 2eV." This is not an accurate account of the microphysics. The 1st atom here and the 2nd atom there is the **same microstate** as the 2nd atom here and the 1st atom there:

4-19
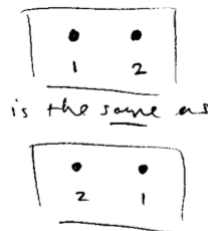
This is actually a result from quantum mechanics. In classical mechanics it is possible to make a gas out of particles with nametags. (Actually we could just label them by their histories.) That's just not the kind of gas we get from atoms and molecules, which are really described by quantum mechanics, in a way which sometimes (conveniently for us) reduces to classical mechanics. Any two atoms of the same kind (same isotope, same state of nucleus, same state of electrons) are identical. This is a deep fact[7].

We've over-counted the configurations of the atoms by a factor of $N!$. If we continue to leave out this factor of $1/N!$ from the correct answer we will run headfirst into a problem with the extensivity of our thermodynamic variables which are supposed to be extensive, which is called Gibbs' Paradox (on the next pset). Gibbs recognized that this factor of $1/N!$ had to be there in the late 19th century, by this reasoning. That is bad-ass. In some ways this can therefore be considered the first QM result.

So our final result for the microcanonical state count is:

$$\Omega(E, V, N) \;=\; \frac{3N\Delta}{2E} V^N \left(\frac{em}{3\pi\hbar^2}\right)^{3N/2} \left(\frac{E}{N}\right)^{3N/2} \underbrace{\left(\frac{e}{N}\right)^N}_{\text{Stirling approx to } 1/N!}$$

$$=\; \underbrace{\left(\frac{V}{N}\right)^N \left(\frac{E}{N}\right)^{3N/2}}_{\text{this is the important bit for thermo}} \frac{3N\Delta}{2E} \left(\frac{em}{3\pi\hbar^2}\right)^{3N/2} e^N \qquad (2)$$

Now that we've done the counting, let's follow the rest of the 8.044 prescription. Find S then T then P. The physical results are about to fall out here fast and easily like few other times your experience of physics.

$$S = k_B \ln \Omega = k_B N \left(\ln\left(\frac{V}{N}\right) + \frac{3}{2}\ln\left(\frac{E}{N}\right) + \underbrace{1 + \frac{3}{2}\ln\left(\frac{em}{3\pi\hbar^2}\right)}_{\text{only differences in entropy matter, classically}} \right) + \underbrace{k_B \ln\left(\frac{3N\Delta}{2E}\right)}_{\ll N. \text{ ignore.}}$$

$$\implies\quad S = k_B N \left(\ln\left(\frac{V}{N}\right) + \frac{3}{2}\ln\left(\frac{E}{N}\right) + \text{const} \right)$$

Note that S is extensive: cut the box in half and see what happens to the entropy. If we divide E, V, N by 2, we can see from the expression above that S is divided by 2.

You can see that if we had not put in the $1/N!$, S would not have been extensive.

---

[7]It is actually a result of quantum field theory: *e.g.* the electrons in the atoms are all identical because they are all quanta of the same *field*.

Internal energy of monatomic ideal gas, from scratch

Now use the stat mech definition to compute T:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_V = k_B N \cdot \frac{3}{2} \cdot \frac{1}{E}$$

Rearrange this to find an expression for the internal energy of a monatomic ideal gas:

$$\boxed{E = \frac{3}{2} N k_B T} \quad \text{for a monatomic ideal gas.}$$

Previously, we stated this without derivation.

$$\implies \quad C_V = \left(\frac{\partial E}{\partial T}\right)_V = \frac{3}{2} N k_B,$$

$$\implies \quad C_P = C_V + N k_B = \frac{5}{2} N k_B, \quad \implies \gamma = \frac{5}{3}.$$

Ideal gas law from scratch

Now let's find P:

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_E = k_B N \cdot \frac{1}{V}$$

Rearrange to find a more familiar expression for the Equation of State of an ideal gas:

$$\boxed{PV = Nk_B T} \quad \text{ideal gas law, from scratch}$$

We have derived this statement, starting from the basic postulate of statistical mechanics. Be impressed.

It can be useful to rewrite S in terms of the other thermodynamic variables:

$$S = k_B N \left(\ln\left(\frac{V}{N}\right) + \frac{3}{2}\ln\left(\frac{E}{N}\right) + c_1\right)$$

$$S = k_B N \left(\ln\left(\frac{V}{N}\right) + \frac{3}{2}\ln\left(T\right) + c_2\right)$$

$$S = k_B N \left(\ln\left(\frac{T}{P}\right) + \frac{3}{2}\ln\left(T\right) + c_3\right)$$

$$S = k_B N \left(-\ln\left(PT^{5/2}\right) + c_3\right)$$

where $c_{1,2,3}$ are constants, independent of the thermodynamic state (these constants drop out when we take derivatives, as they did above). So we see that along a path where the entropy is constant, $PT^{5/2}$ is constant. A constant entropy, reversible path has $\Delta S = \frac{\Delta Q}{T} = 0$, no heat in or out. This determines the shape of an adiabat. It agrees with the answer we got before.

[End of Lecture 12.]

Microscopic information from the stat mech postulate

So far we got thermodynamic relations out, now let's think about microscopic probability densities for ideal gas. We'll do two examples:

1) What is $p(x_i)$? $p(x_i) \equiv$ the probability density for the x-position of the ith particle, in equilibrium.

Q: WHICH ONE IS THE ith ONE? Here I am imagining that I can label one of the particles. The factor of $(N-1)!$ for the remaining particles will drop out of the normalized probabilities.

$$p(x_i) \;=\; \frac{\Omega'(x_i)}{\Omega} = \frac{\text{phase space volume with } x_i \text{ held fixed}}{\text{phase space integral over all } 6N \text{ vars}}$$

$$=\; \frac{L_x^{N-1} L_y^N L_z^N \cdots}{L_x^N L_y^N L_z^N \cdots}$$

$$=\; \frac{1}{L_x} \qquad \text{for } 0 \le x_i \le L_x$$

This was to be expected. Since the positions of the atoms in an ideal gas are statistically independent, we infer that the probability that all the atoms are in the North half of the room $= \frac{1}{2}^N \sim 2^{-10^{24}}$.

2) What is $p(p_{xi})$? $p(p_{xi}) \equiv$ the probability density for the x-component of the momentum of the ith particle.

$$p(p_{xi}) = \frac{\Omega'(p_{xi})}{\Omega} = \frac{1}{\Omega} \int_{E \leq H(\underbrace{\{p,q\}}_{\text{list includes } p_{xi}}) \leq E+\Delta} dq_1...dq_{3N} \underbrace{dp_1...dp_{3N}}_{\text{except } p_{xi}}$$

In this expression, we integrate only over values of the other p_i where the energy $\varepsilon \equiv p_{xi}^2/(2m)$ (the KE$_x$ of atom 1) plus the energy of all the other p_i adds up to E. So we have energy $E - \varepsilon$ to distribute amongst the other $3N - 1$ p_i. The cumulative probability is the volume of a ball in $3N - 1$ dimensions, of radius $2m\sqrt{E - \varepsilon}$. So Ω' is the same as Ω in Eqn. (2) but with $3N \to 3N - 1$ ps and $E \to E - \varepsilon$:

$$\Omega' = \left(\frac{V}{N}\right)^N \left(\frac{E - \varepsilon}{3N - 1}\right)^{\frac{3N-1}{2}} \frac{(3N - 1)\Delta}{2(E - \varepsilon)} \left(\frac{em}{\pi\hbar^2}\right)^{\frac{3N-1}{2}} e^N .$$

Many things cancel in the ratio:

$$p(p_{xi}) = \frac{\Omega'}{\Omega} = \underbrace{\frac{V^N}{V^N}}_{\text{range of accessible positions same.}} \underbrace{\frac{h^{3N}}{h^{3N}}}_{\text{same normalization factor.}} \frac{\left(\frac{E-\varepsilon}{3N-1}\right)^{\frac{3N-1}{2}}}{\left(\frac{E}{3N}\right)^{\frac{3N}{2}}} \frac{(4\pi em)^{\frac{3N-1}{2}}}{(4\pi em)^{\frac{3N}{2}}} \left(\frac{\frac{N-1}{E-\varepsilon}}{\frac{N}{E}}\right)$$

Note that the units work: $[\Omega] = [1], [\Omega'] = \frac{1}{[p_{xi}]}$. The last factor approaches 1 in the thermodynamic limit $N \to \infty, \varepsilon \ll E$. This leaves:

$$p(p_{xi}) = \left(\frac{3}{4\pi em}\right)^{1/2} \underbrace{\frac{(E - \varepsilon)^{\frac{3N-1}{2}}}{E^{3N/2}}}_{\equiv A} \underbrace{\frac{N^{3N/2}}{(N - 1/3)^{\frac{3N-1}{2}}}}_{\equiv B}$$

Note that $\varepsilon \ll E$, or else the distribution is zero.

$$A \simeq \frac{1}{\sqrt{E}} \left(1 - \frac{\varepsilon}{E}\right)^{3N/2}$$

Use $E = \frac{3}{2}Nk_BT$:

$$A \simeq \frac{1}{\sqrt{\frac{3}{2}Nk_BT}} \left(1 - \frac{\left(\frac{\varepsilon}{k_BT}\right)}{\frac{3}{2}N}\right)^{\frac{3}{2}N} \simeq \frac{1}{\sqrt{\frac{3}{2}Nk_BT}} e^{-\frac{\varepsilon}{k_BT}}$$

(For the second step, recall: $\lim_{N \to \infty} \left(1 + \frac{1}{N}\right)^N = e$, the base of the natural log.) Similarly,

$$B \simeq \sqrt{N} \left(\frac{1}{\left(1 - \frac{1}{3N}\right)}\right)^{\frac{3}{2}N} \simeq \sqrt{N}e^{1/2} .$$

Put back together the probability distribution for the x-momentum of one particle:

$$p(p_{xi}) \simeq \left(\frac{3}{4\pi em}\right)^{1/2} \frac{1}{\sqrt{\frac{3}{2}Nk_BT}}\sqrt{N}e^{1/2}e^{-\frac{\varepsilon}{k_BT}} = \frac{1}{\sqrt{2\pi mk_BT}}e^{-p_{xi}^2/(2mk_BT)}$$

Correctly normalized, correct units. To find the velocity distribution from the momentum distribution, recall $v_{xi} = p_{xi}/m$ and our discussion of functions of a random variable:

$$p(v_{xi}) = \left(\frac{m}{2\pi k_BT}\right)^{1/2} e^{\frac{-\frac{1}{2}mv_{xi}^2}{k_BT}} \tag{3}$$

This is the distribution (notice that the units are right and it's correctly normalized) that I presented out of nowhere in Chapter 2. It's called the Maxwell velocity distribution, and we just derived it from the Fundamental Postulate – equal probabilities for accessible microstates.

Note that the probability in (3) is proportional to $e^{-\frac{\mathrm{KE}_x}{k_BT}}$ – the quantity in the exponent is the energy in question divided by the energy scale associated with the temperature. This is a *Boltzmann factor*, with which we will have much truck in the coming weeks.

With this victory in hand we conclude our discussion of the classical, monatomic, ideal gas. Later we will remove and modify these modifiers.

Some words about the bigger picture: this straight-ahead method of just doing the integral defining $\Omega(E)$ doesn't usually work. (One more system next for which it works, a few more on pset 6.)

To do better, we need a more powerful method (the Canonical Ensemble, which is the subject of Chapter 6). Before that, in Chapter 5 we finish up our discussion of Thermodynamics – in particular with this understanding in terms of counting accessible microstates, we are better prepared to discuss entropy as a state variable.

4.4 Two-state systems, revisited

Here we will consider, as in Chapter 1, a collection of classical two-state systems: equivalent to a grid of boxes which can be ON or OFF.

- N independent 2-state systems.

- Each system has two states – it is "a bit" (or a spin).
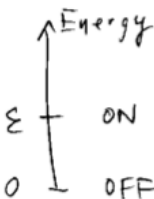
-   OFF has energy 0, ON has energy ε

- (Partial) analog of a collection of QM 2-state systems. Such a thing is a useful model for lots of physical contexts.
  – spins in a magnetic field (a paramagnet)
  – molecules on a surface
  – ions in a crystal which can sit in one of two different sites in each unit cell.
  But: here no superposition.

- We will treat ε as a fixed parameter for now.
  Later, we'll think of it as something we can control externally: the magnetic field to which we subject the paramagnet.

-
$$N = \underbrace{N_0}_{\#\ \text{OFF}} + \underbrace{N_1}_{\#\ \text{ON}} \qquad\qquad E = 0 + \varepsilon N_1$$

- N_1, or equivalently E, specifies the **macro**state.

How many microstates?

It's the same as the number of ways to choose N_1 (indistinguishable-from-each-other) cards from a deck of N.

$$\Omega(E) = \frac{N!}{N_1!(N-N_1)!} \quad \text{with } N_1 = E/\varepsilon$$

$$S \equiv k_B \ln \Omega$$

(Stirling) $\Longrightarrow$

$$= k_B \left(N \ln N - N_1 \ln N_1 - (N-N_1) \ln(N-N_1) \underbrace{-N + N_1 + (N-N_1)}_{=0} \right)$$

$$\boxed{S(E)/k_B = N \ln N - N_1 \ln N_1 - (N-N_1)\ln(N-N_1)} \quad N_1 = \frac{E}{\varepsilon}$$
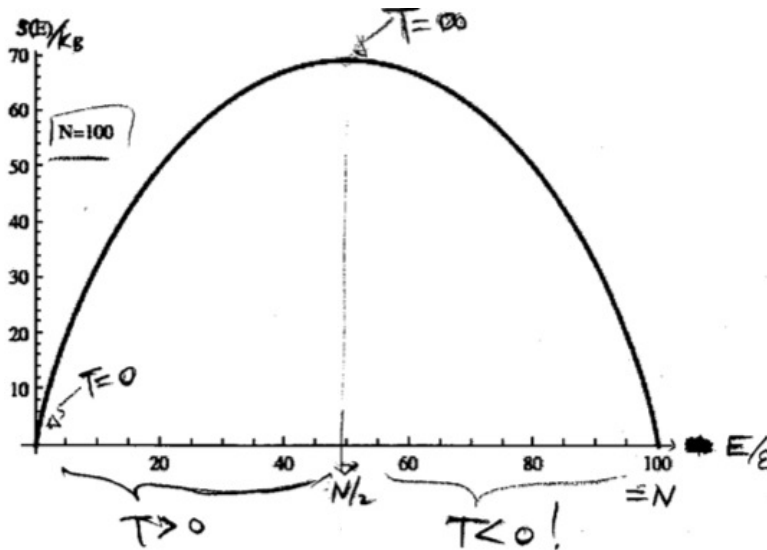
Now we can get all the stuff:

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{1}{\varepsilon}\frac{\partial S}{\partial N_1} = \frac{k_B}{\varepsilon}\left(-\ln N_1 - 1 + \ln(N-N_1) + 1 \right)$$

$$\frac{\varepsilon}{k_B T} = \ln\left(\frac{N-N_1}{N_1}\right) = \ln\left(\frac{N_0}{N_1}\right)$$

$$e^{\frac{\varepsilon}{k_B T}} = \frac{N}{N_1} - 1$$

$$N_1 = \frac{N}{1 + e^{\frac{\varepsilon}{k_B T}}} \quad \Longrightarrow \quad \boxed{E = \frac{\varepsilon N}{1 + e^{\frac{\varepsilon}{k_B T}}}} \tag{4}$$

Plots:

$$\frac{\partial S}{\partial E} = \frac{1}{T}$$

From the plot we see that $N_1 \to 0$ at low T, $N_1 \to N/2$ at high T.

At high T, ON and OFF are equally probable; the fact that ON has higher energy becomes irrelevant for $k_B T \gg \varepsilon$.

Comments on negative temperature: $T < 0$ is

- possible. It has been realized experimentally.

- hot: A system with $T < 0$ increases its entropy by giving up energy. Therefore, a system comprised of a $T < 0$ system and ordinary stuff can't ever be in equilibrium – the $T < 0$ system will keep giving up heat, because by doing so it can *increase* its entropy. Any negative temperature is hotter than any positive temperature.

- not usually encountered in physics, because usually when we study a system at higher energy, we encounter *more* possible states. For example: if we put enough energy into an ordinary system, we can break it apart; the relative positions of its constituents then represent new degrees of freedom which were not previously accessible. (The idea that by studying things at higher energies we can probe their more microscopic constituents is in fact the basis of elementary particle physics.)

- the subject of Baierlein §14.6, if you want to learn more.
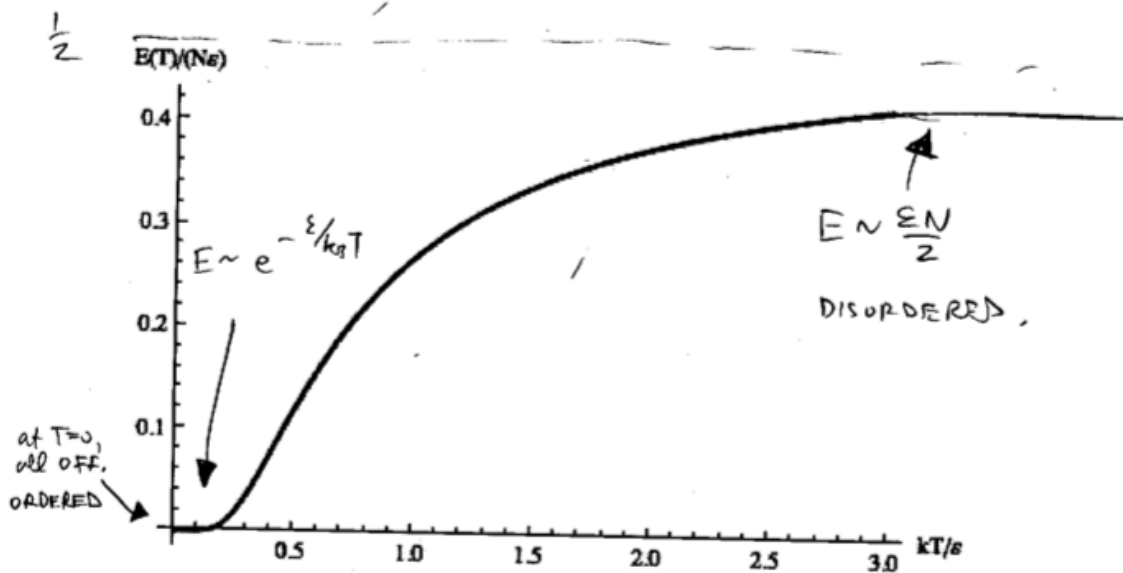
- Exponentially activated behavior at low T, $E \sim e^{-\varepsilon/(k_B T)}$ is due to the presence of a nonzero energy gap ε. This exponential is another Boltzmann factor, and like I said it will play a big role later on in 8.044.

- As $T \to \infty$, $E \to \frac{\varepsilon N}{2}$: This is the maximally disordered state, maximum entropy.

Heat capacity:

$$C \equiv \frac{dE}{dT}\bigg|_\varepsilon = Nk_B \left(\frac{\varepsilon}{k_B T}\right)^2 \frac{e^{\frac{\varepsilon}{k_B T}}}{\left(1 + e^{\frac{\varepsilon}{k_B T}}\right)^2}$$

low T: $\quad C \sim Nk_B \left(\frac{\varepsilon}{k_B T}\right)^2 e^{-\frac{\varepsilon}{k_B T}}$

hight T: $\quad C \sim \left(\frac{\varepsilon}{k_B T}\right)^2 \frac{1}{4}$
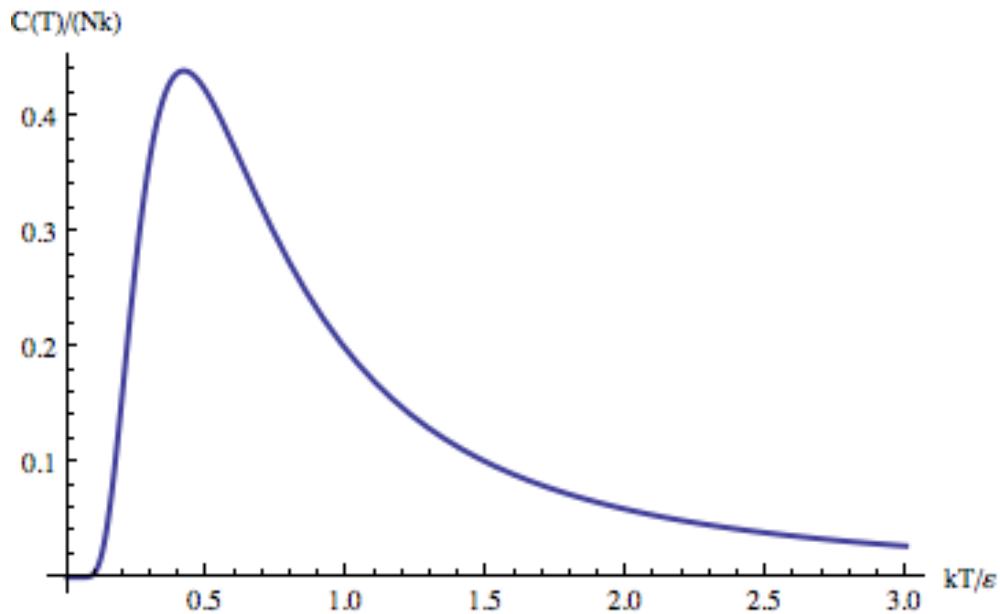
In the formulae above we've assumed that all the 2-state systems have the *same* level spacing. A generalization of the above formula for the energy to the case where the level spacings are differnet (which may be useful for a certain extra credit problem) is

$$E = \sum_{i=1}^{N} \frac{\varepsilon_i}{1 + e^{\frac{\varepsilon_i}{k_B T}}} \ .$$

This formula is actually quite difficult to derive using the microcanonical ensemble, though you can see that it gives back the formula for the energy (4) above. It will be simple using the canonical ensemble, in Chapter 6.
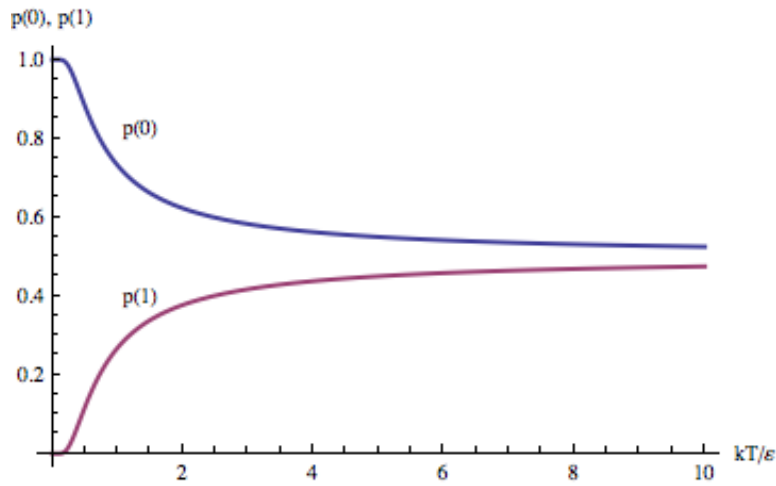
Probability that a bit is ON

Let $n = 0$ mean OFF, $n = 1$ mean ON.

$$p(n) = \frac{\Omega'(N \to N-1, N_1 \to N_1 - n)}{\Omega}$$

$$= \underbrace{\frac{(N-1)!}{N!}}_{=\frac{1}{N}} \underbrace{\frac{N_1!}{(N_1 - n)!}}_{= \begin{cases} 1 & \text{if } n = 0 \\ N_1 & \text{if } n = 1 \end{cases}} \underbrace{\frac{(N - N_1)!}{(N - 1 - N_1 + n)!}}_{= \begin{cases} N - N_1 & \text{if } n = 0 \\ 1 & \text{if } n = 1 \end{cases}}$$

$$p(n) = \begin{cases} 1 - \frac{N_1}{N} & \text{for } n = 0 \\[2mm] \frac{N_1}{N} & \text{for } n = 1 \end{cases}$$

This is a check on the formalism.

$$p(1) = \frac{N_1}{N} = \frac{1}{1 + e^{\frac{\varepsilon}{k_B T}}}$$

Recap of logic

Model the system. What is the configuration space? What is the Hamiltonian?

Count. Find $\Omega(E, V, N)$, the number of accessible microstates.[8]

- On the one hand, this gives thermodynamics:

$$S = k_B \ln \Omega$$

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V,N...}, \qquad \frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{E,N,...}$$

- On the other hand, this gives distributions for microscopic variables:

$$p(\text{specific property } X) = \frac{\Omega'(X)}{\Omega}$$

The counting is the hard part. We can only do this for a few simple cases.

More powerful method is in Chapter 6. Chapter 5 is a better-informed second pass at thermodynamics.

[End of Lecture 13.]

---

[8]Recall that our definition of Ω includes our imprecision in the energy measurement Δ: it is the number of microstates with energy between E and $E+\Delta$. Also, my statement that this is the 'number' of microstates is precise in the quantum mechanical case. In the classical case, it is a volume of phase space, in units of h^{3N}; interpreting this as a number of states requires making the connection with a quantum system to which this classical system is an approximation.