Part II

Statistical Mechanics
Mathematical Preliminaries II

On Counting

The unordered multiplicity of the number of ways of choosing $n$ sticks from a collection of $N$ total sticks is called "$N$ choose $n". It is written $\binom{N}{n}$ and is equal to

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

The ordered (label-distinguished) way of choosing $n$ sticks from a total of $N$ sticks is $N!/n!$.

(Recall that the "factorial" notation $f! = f(f-1)(f-2)(f-3)...(2)(1)$.

If you are throwing an $l$-sided dice, after $N$ throws the total possible number of outcomes (labeling their order in the throws) is $l^N$.

Stirling Formula - When we are counting possible outcomes we will often run into awful-looking expressions involving factorials. However, often the expressions involve factorials of large numbers. There is a very useful formula that well approximates factorials of large numbers; it is called Stirling's formula.

$$N! \sim \sqrt{2\pi N} N^N e^{-N}(1 + O\left(\frac{1}{N}\right))$$

for $N$ large (in practice useful when $N > 6$).

On Sums

The Euler summation formula is very useful;

$$\sum_{n=0}^{\infty} f(n) = \int_{0}^{\infty} f(x) \, dx + \frac{1}{2} f(0) - \frac{1}{12} f'(0) + \frac{1}{720} f''(0) + \ldots$$

On Integrals

Steepest Descent -or- Stationary Phase - often when doing problems in statistical physics and quantum mechanics it will be necessary to compute integrals of exponentials of a positive function $f(x)$, for example

$$I = \int e^{-Nf(x)} \, dx$$

Very often these integrals may be well approximated at large $N$ by use of "steepest descent". The point is that $e^{-Nf(x)}$ is very small at large $N$. The major contribution to the integral then occurs near the minimum of $f(x)$. Suppose the minimum of $f(x)$ occurs at $x = x_0$. Thus since the most of the contribution to the integral comes from around the point $x_0$, we may approximate $f(x)$ there by use of the Taylor expansion.

$$f(x) \sim f(x_0) + f'(x_0)(x-x_0) + f''(x_0)(x-x_0)^2/2 + \ldots$$

where $f'(x_0) = \frac{df(x)}{dx}\bigg|_{x=x_0}$ and $f''(x_0) = \frac{d^2f(x)}{dx^2}\bigg|_{x=x_0}$. Now, since $x_0$ is a minimum, we must have $f'(x_0) = 0$ Subbing this into the integral $I$ we have

$$I \sim e^{-Nf(x_0)} \int e^{-Nf''(x_0)/2(x-x_0)^2} \, dx$$
And, this is simply a Gaussian integral that is easily handled. The answer is given in the first mathematical preliminaries handout. Just for completeness, we here use the result of before and write down the answer for the above integral.

\[
I \sim e^{-Nf(x_0)} \sqrt{\frac{2\pi}{Nf''(x_0)}}
\]

Useful Integrals

\[
\int_0^\infty \frac{x^3}{e^x - 1} \, dx = \frac{\pi^4}{15} \quad (I)
\]

\[
\int_0^\infty \frac{x\,dx}{e^x + 1} = \frac{\pi^2}{12} \quad (II)
\]

\[
\int_0^\infty \frac{x^{1/2}}{e^x - 1} \, dx = 1.306\sqrt{\pi} \quad (III)
\]

\[
\int_0^\infty \frac{x^{\nu-1}}{e^x - 1} \, dx = \Gamma(\nu)\zeta(\nu) \quad (IV)
\]

\[
\int_0^\infty \frac{x^{\nu-1}}{e^x + 1} \, dx = (1 - 2^{1-\nu})\Gamma(\nu)\zeta(\nu) \quad (V)
\]

where the last two rather general looking integrals, \(\Gamma\) is the gamma function, (in integers \(\nu\), \(\Gamma(\nu) = (\nu - 1)!\)) and \(\zeta\) is the Riemann zeta function, which for special values of the argument, namely, when \(\nu = 2m\), where \(m\) is an integer, has the value,

\[
\zeta(2m) = \frac{2^{2m-1}\pi^{2m}|B_{2m}|}{(2m)!} \quad (VI)
\]

where \(B_{2m}\) are the Bernoulli numbers, \(B_0 = 1, B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}, B_8 = -\frac{1}{90} \ldots\)
Problem Set 7

More on Counting

1.) Recall that in class we derived a formula (the binomial coefficients) for counting the number of outcomes, distinguished by the total number of heads \( n \), after \( N \) throws. This formula used the fact that there are only two possibilities \( (l = 2) \) for each coin toss.

Suppose that we have a 3 sided die, call the sides a, b, and c. We throw the die \( N \) times. Write down an expression for the number of total (ordered) distinct outcomes. Now write down a formula that counts the number of outcomes that have \( n_a \) a’s appearing and \( n_b \) b’s appearing.

If I throw my three headed die 9 times, what are my chances of never rolling an ‘a’? What are my chances of rolling at least 3 “a”’s and but only one “b”?

(HINT: one way to do these it to carefully go through it case by case, and logically include all the disjoint posibilities)

2.) Two six sided die are thrown together a large number of times. Person A receives a point each time the sum of the thrown die is an odd number and person B receives a point if the sum is even. Who will accumulate the most points? Suppose that instead of 6 sided die, three sided ones are used. Who will accumulate the most points in this case? If there are only 6 throws, what is the chance that Person A or Person B wins.

3.) Study the example of the computation of the multiplicity function of the harmonic oscillator. For that case,

\[ g(N, n) = \frac{(N + n - 1)!}{n!(N - 1)!} \]

Use Stirlings approximation to show that, for large \( N \) but holding \( n \) at some fixed value, the multiplicity function \( g(N, n) \) grows as

\[ g(N, n) \sim A(n)N^n \]

for fixed \( n \). \( A \) is a function of \( n \) alone, and you can also find it’s approximate functional form. Note how different this is than the spin chain that we spoke about in class.
Problem Set 8

The Micro-canonical Ensemble

1.) Problem 13-3

2.) Ehrenfest’s Dogs - Consider the following problem, originally due to Paul Ehrenfest, also one of the contributors to the foundation of modern statistical mechanics. There are two dogs, dog 1 and dog 2. Each dog has \( N \) slots where fleas can hide. Initially dog 1 has \( N \) fleas and dog 2 has none.

Now, suppose that per unit time the probability that a particular flea will jump to the other dog (thereby leaving the one that it is on) is \( \tau \). Write down a simple differential equation describing the number of fleas on each dog as function of time. (assume \( N \) is large.) Solve this equation for the number of fleas on each dog as a function of time, starting from the time that the dogs are first brought together.

Now, since each dog has \( N \) slots for fleas, write down an expression for the number of configuration of \( n_1 \) fleas on dog 1. Call that function a multiplicity \( g(N, n_1) \). Now write down an expression for the multiplicity function, \( g \), of the configuration of \( n_1 \) fleas on dog one and \( n_2 = N - n_1 \) fleas on dog 2 combined.

Compute the entropy of this system as \( \log(g) \) as a function of time and sketch it.

3.) Consider a system composed of \( N \) spins \( \sigma_i \), \( i = 1, 2, 3, ..., N \), each taking one of three values \( \sigma_i = -1, 0, +1 \).

Let the energy of the system be given as

\[
U = -mB \sum_{i=1}^{N} \sigma_i = -2mBs
\]

where \( s = (N_\uparrow - N_\downarrow)/2 \).

Write a compact expression (sum) for the multiplicity \( g(N, U) \) as a function of energy.

Now specialize to the \( U \sim 0 \) limit. Use Stirling’s approximation and approximate \( g(N, U) \) as a single (the largest) term.

Calculate the average number of zero-spins (\( \sigma_i = 0 \)) in the \( U = 0 \) limit.

Derive an expression for the \( U \) as a function of temperature.
Lecture 23

Introduction to Statistic Mechanics

Having addressed the development of some of the key concepts in classical thermodynamics we now study this century’s development of thermodynamics, so called statistical mechanics. While classical thermodynamics was a general framework for describing the bulk behavior of a suitably generic macro-states, this generality is also classical thermo’s deficiency. For example, the Clausius-Clapeyron equation is adept at describing the depression of a phase transition temperature as one changes the pressure, but it does so in terms of quantities $L$ and $v_g$ for example that are impossible to compute from the point of view of classical thermodynamics. Classical thermodynamics doesn’t distinguish between the boiling of water and the freezing of Napthalene. We are inescapably led to the conclusion that we need a tool that is more *specific* (but just as powerful and universal!) than classical thermodynamics.

The major breakthrough in particularizing thermodynamics came in the early part of the last century with the work of L. Boltzman and finally with the advent of quantum mechanics. It led to a full fruition of the ideas of the early kinetic theorists, such as J. W. Gibbs.

The point to ponder here is: how do the microphysical properties of the subsystems play out in the behavior of the macro-state? Recall that the early success of kinetic theory was based on the fact that the microphysical picture (mass points whizzing about) was so generic, but we know and can easily measure the difference between, for example, the heat capacity of helium gas and nitrogen gas. Somehow differences in the microscopic properties of the subsystems effects the behavior of the equilibria.

Today we cover a few basic definitions and some mathematical preliminaries.

**State** - Recall the definition of macro- and micro-state of before. We now shall think of the macro-state as referring to an equivalence between a large number of distinguishable microstates of a system. Very often we only consider the accessible configurations of the subsystems, meaning those configurations that the system can get in a time small compared with the timescale that we are studying the whole system. Also, as described in one of those first lectures, when we wish to ascertain the value of some quantity such as the energy of a macro-state, we will average over the set of microstates that represents that macrostate.

**Probability** - or likelihood, represents how often a system finds itself in a particular microstate, or set of microstates. Probabilistic methods will figure prominently throughout the subsequent discussion and will take a little getting used to. The good news is that for most of the work you need not be a trained statistician. Just as in classical thermodynamics, crude methods often suffice.

**Partitions and Counting** - Partitions refer to collections of microstates that may be part of a particular realization of the macro-state under study. Often we will be interested in counting the possible partitions and assigning probability to them in such a way that the total probability of the system being in any of the individual partitions will be one.

**Ensemble** - The collection of accessible, distinguishable microstates that are equivalent descriptions of the same macro-state. **Multiplicity** - The multiplicity of a macro-state is simply the number of micro-states in the associated ensemble.
The Fundamental Assumption of Statistical Mechanics

The fundamental assumption of statistical mechanics can be stated in one line;

*A system in a given macro-state is equally likely at any instant to be in any one micro-state of the ensemble describing the macro-state.* Or more tersely, *All microstates are equally likely* Thus if the probability of each micro-state is \( 1/g \), where \( g \) is the multiplicity of the macro-state. Call the ensemble \( \mathcal{Y} \). In notation \( g = |\mathcal{Y}| = \sum_{s \in \mathcal{Y}} 1 \). Suppose that we want to compute the *average value* (also called the *expectation value*) of a quantity \( W \) associated to the whole system. If it is computable for each microstate then by the above fundamental assumption, the average value is

\[
<W> = \sum_{s \in \mathcal{Y}} \text{Probability}(s) \ W(s) = \frac{1}{g} \sum_{s \in \mathcal{Y}} W(s) \quad (23.1)
\]

Example: Binary Spin System.
Example of Spin System

As formulated, statistical mechanic’s basic assumptions seem vague and too simplistic. They are to prove powerful and descriptive to things way beyond what would be expected from such humble beginings. But how do we go about applying these ideas to things like a gas? We’ve spoken about our ensemble as consisting of lists of the positions ans momenta of each particle, and so there are necessarily an infinite number of different lists and so what could assigning them each a probability of $1/\infty$ mean? To illustrate some of the very basic, almost primitive, ideas that form the basis of all our later considerations, I think that it would be usefull to struggle through a particularly simple system, called a spin system. We will generalize these ideas to more complicated things like gasses later.

A spin system is a collection of spins. Think of each spin as a little magnet. Suppose further that it can either point up or down only (though later we will generalize this in several ways.) Immagine that we are wish to consider the gross behavior of a large collection of spins, say $N$ of them ($N$ very large.) Since they only point either up or down there are $2^N$ distinct lists of spin arrangements. That is the multiplicity of our ensemble. Suppose we don’t distinguish between locations but only between various states by the total number of up-spins $N^\uparrow$. Let the number of spins pointing down be $N^\downarrow = N - N^\uparrow$. We wish to find the multiplicity of the partition of the ensemble, all states of which have $N^\uparrow$ up-spins. The book derives it by way of defining a generating function for the partitions with various numbers of up-spins. Suffice to say, the number is simply the binomial coefficient $\binom{N}{N^\uparrow}$, or

$$g(N, N^\uparrow) = \binom{N}{N^\uparrow} = \frac{N!}{(N-N^\uparrow)!N^\uparrow!} = \frac{N!}{N^\uparrow!N^\downarrow!} \quad (24.1)$$

Where $g(N, N^\uparrow)$ is called the multiplicity of the partition with $N^\uparrow$ up spins.

It is often useful to define the spin excess, $s$ as

$$s = \frac{N^\uparrow - N^\downarrow}{2} \quad (24.2)$$

Note that $-N/2 < s < N/2$. You should think of $s$ as a macroscopic variable, a state variable of the system, a bit like the volume is a state variable of a gas. Thus

$$g(N, s) = \frac{N!}{(N/2-s)!(N/2+s)!} \quad (24.3)$$

For large $N$ we can easily see that this function must be very sharply peaked about $s = 0$. For large $N$, we may use Stirling’s formula to rewrite the various factorials that appear in Eq.(24.3). To simplify things even further, it will only be necessary to retain part of Stirlings formula

$$N \text{ large } \rightarrow \quad N! \sim N^N e^{-N} \quad (24.4)$$

Plugging this into Eq.(24.3) and simplifying we find

$$g(N, s) \sim \frac{N^N}{(N/2+s)^{N/2+s}(N/2-s)^{N/2-s}} \quad (24.5)$$
Now, suppose that we are interested in seeing just how peaked this distribution is at \( s = 0 \),
note that for \( s << N \) we may write,
\[
g(N, s) \sim \left( \frac{2^2}{(1 + 2s/N)^{1+2s/N}(1 - 2s/N)^{1-2s/N}} \right)^{N/2} \sim g(N, 0) e^{-4s^2/N} \tag{24.6}
\]

where \( g(N, 0) \sim 2^N \). Note two keys things about this. First, because \( s = 0 \) is a maximum of the multiplicity function, we know that at \( s = 0 \) the multiplicity function must be flat. Therefore an approxiamtely gaussian shape is expected. Second, we find that the width, \( \delta s = 2/\sqrt{N} \), of this peak is small in proportion to the full range of \( s \), that is, \( \frac{\delta s}{\sqrt{N}} \sim 1/\sqrt{N} \).
This tells us that as \( N \) gets very large, virtually all the states are near \( s = 0 \). We will see other properties like this "central limit" arise in many systems.

Finally, suppose that we placed these spins in a magnetic field \( B \), pointing up say. Each individual spin will "feel" the presence of the field, and to each one we may assign an energy \( -mB\sigma_i \), where \( m \) is the magnetic moment and \( \sigma_i = \pm 1 \) is a variable representing the \( i \)'th spin in the spin system. The total energy of the system would then be
\[
U = -mB \sum_i \sigma_i = -2mBs \tag{24.7}
\]

And so we may speak about the multiplicity function \( g(N, U) \), representing the number of configurations of our spin system of a particular energy \( U \).
The fact that the multiplicity function is so strongly peaked leads one of the following consideration. Suppose that we had two systems, one with energy \( U_1 \), and another with energy \( U_2 \). Suppose that we bring them into intimate contact where they can share heat, etc, and generally interact strongly with one another.

Suppose that the multiplicity function of each system is a function of its energy, as was the case above. Denote the total energy by \( U = U_1 + U_2 \). The multiplicity of the combined state of the two baths is
\[
g = \sum_{U_1} g_1(U_1)g_2(U - U_1) \tag{24.8}
\]

We know that if the systems are each very large, that the individual multiplicity functions \( g_1 \) and \( g_2 \) will be very strongly peaked. We may therefore reasonably hope that in limit that the systems are very large that the sum in Eq.(24.8) will be dominated by a single term (!). Call the value of \( U_1 \) for that special term \( U^* \). Thus,
\[
g \sim g_1(U^*)g_2(U - U^*) \tag{24.9}
\]

and the statement that it is a maximum indicates that \( \frac{\partial g}{\partial U^*} = 0 \). Applied to Eq.(24.9) we find
\[
\frac{\partial \log g_1(U^*)}{\partial U^*} = \frac{\partial \log g_2(U - U^*)}{\partial (U - U^*)} \tag{24.10}
\]

And that indeed, for any number of systems joined together this derivative of the logarithm of the multiplicity function with respect to the energy in the (sub-)system is the same. What could it represent? The only thing that we know about bringing such general systems into thermal equilibrium is that the temperatures must equilibrate. Thus we conclude that
\[
\frac{\partial \log g(U)}{\partial U} = \frac{1}{kBT} \tag{24.11}
\]
for some constant $k_B$. This constant is conventionally called Boltzmann’s constant (!) and thus Eq.(24.11) suggests that we identify $k_B \log g(U)$ with the entropy of the system

$$
S = k_B \log(g(U))
$$

(24.12)

A very interesting conclusion indeed! So to compute the entropy of a system we need only specify the number of configurations $g$ of the system at that energy and use Eq.(24.12)!

This relation, Eq(24.12) called Boltzmann’s relation, fundamentally connects entropy with our inability to determine which (micro-) state in particular a system is in...it is a measure of our ignorance of the actual microphysical state. This is basically the bridge between micro- and macro- thermo; this derivation and Boltzmann’s ideas were apparently quite controversial in the day...finally, this equation and the name “Boltzmann” are -eerily- the only adornments on Ludwig Boltzmann’s tombstone.

Finally, we return to our example of the spin system and compute the entropy if it’s energy functional is as given $U = -2mBs$ (Eq.(24.7)). Then $g(N, s)$ from Eq.(24.3) we write using the Stirling approximation (see form in Eq.(24.5)) that the entropy

$$
S = k_B \log(g) = k_B N \log(N) - k_B \left(\frac{N}{2} + s\right) \log\left(\frac{N}{2} + s\right) - k_B \left(\frac{N}{2} - s\right) \log\left(\frac{N}{2} - s\right)
$$

(24.13)

so that using the definition of temperature as $\frac{1}{T} = \frac{\partial S}{\partial U} |_v$ (see Eq.(24.11)), we have

$$
\frac{1}{T} = k_B \left[-\log\left(\frac{N}{2} - s\right) + \log\left(\frac{N}{2} - s\right)\right] \frac{\partial s}{\partial U} = -\frac{k_B}{2mB} \log\left(\frac{N}{2} - s\right) \log\left(\frac{N}{2} + s\right)
$$

(24.14)

and so inverting this for the spin excess $s$ we have

$$
\frac{N/2 - s}{N/2 + s} = e^{2mB/k_BT}
$$

(24.15)

which we can use to derive a formula for the internal energy of the system as a function of the temperature and the applied field,

$$
U = -2mBs = -mBN \frac{e^{2mB/k_BT} - 1}{e^{2mB/k_BT} + 1}
$$

(24.16)

This formula finally relates the microphysical description of the degrees of freedom and their individual energies to the energy of the macroscopic state. Note that as $T \to 0$ the $U \to -mBN$ and the system freezes in a state that is completely magnetized. Alternatively in the $T \to \infty$ limit, the $U \to 0$, meaning that in the high temperature limit the system completely demagnetizes, as expected.

**Nernst’s Law:** It is also instructive to study the formula for the entropy Eq.(24.13) in these limits. Note that in the $T \to 0$ limit, the spin excess $s \to N/2$ which means that the entropy $S \to 0$. This is generally expected for systems that do not have a (massively degenerate) ground state; that the entropy must vanish in that limit. Think of the following simple case; the system has a unique ground state and there is an energy gap to get to the next level. Thus, at temperatures low compared with that energy $g(U) = 1$ since there is only one level available, indicating that $S = 0$ via Eq.(24.12). The fact that the entropy should vanish at zero temperature was empirically first noted experimentally by Nersnt during a
Ph.D. defense of one of his students who had measured the entropy functions of various gasses and their mixtures, but it is sometimes referred to as “Nernst’s law”.

The high temperature limit is also informative: as $T \to \infty$ the spin excess $s \to 0$ so the entropy in that limit goes to $S \to N k_B \log(2) = k_B \log(2^N)$, as expected since the total number of accessible states of the system of $N$ binary spins is $2^N$. The increase in the entropy has occurred with a decrease in our knowledge of precisely which state the system is in...informationally, the entropy here is seen as a measure of our lack of information on the instantaneous state of the system.
Lecture 25

The Ising Model

In this lecture we will use some of the techniques that we have developed to the last lectures (on the so-called micro-canonical ensemble) to solve for the thermodynamic behavior of a spin chain, called the one-dimensional Ising model. It is a particularly illustrative example that will lead to useful generalizations.

The one-dimensional Ising model is a spin system in which there are interactions between the spins themselves. The simplest case is that in which \( N \) identical spins, each pointing either up or down, are laid out (imagine they are evenly spaced) on a line. Each spin interacts with its nearest neighbors. Generally we may represent this interaction in terms of an energy, and also may easily represent it in terms of a single parameter \( J \) the spin-spin coupling

\[
U = - \sum_{i=1}^{N} J \sigma_i \sigma_{i+1} \tag{25.1}
\]

where \( \sigma_i = \pm 1 \) represents the value of the spin (pointing up (1) or down (-1)) at the \( i \)’th site. In general we could also add a term like Eq.(24.7) to the definition of the total energy in Eq.(25.1) were we interested in representing the Ising model in a fixed magnetic field. Before we add that complexity, let’s just study the thermodynamics of the system in Eq.(25.1). In order to compute the entropy (and thereby find the equation of state, etc) we will need to compute the multiplicity function \( g \) as a function of the energy \( U \). To this end note that we may rewrite the sum in Eq.(25.1) as

\[
U = J \sum_{l=1}^{N-1} P(l) = -JN + 2Jn \tag{25.2}
\]

where \( P(l) = 1 \) if the spins \( \sigma_l \) and \( \sigma_{l+1} \) are anti-aligned, and \( P(l) = 0 \) if aligned. Thus, the energy \( U \) is simply related to the number \( n \) of ”discontinuities” in the chain (that is, it counts the number of magnetic 'domains’). But for each pair of spins there are only two possibilities; either alligned or anti-aligned. Thus, thought of in terms of discontinuities, the interacting spins system has itself been reduced to a type of non-interacting spin system in which the \( P(l) \) are the independent spins! These \( P(l) \) are called the psuedo-spins. Therefore, counting the number of microstates with a given \( U \) (i.e. fixed \( n \)) is as simple as the non-interacting spin system of lecture 24 (assuming \( N, n \) both large and periodic boundary conditions):

\[
w(U, N) = \binom{N}{n} = \frac{N!}{n!(N-n)!} \tag{25.3}
\]

Now we can use Stirling’s approxiamation to estimate the entropy of our sample.

\[
S = k_B \log(w(U, N)) \sim -(N-n) \log(N-n) - n \log(n) + \text{Const.} \tag{25.4}
\]

and so note that using Eq.(24.11) yields an expression for the temperature

\[
\frac{1}{T} = \left. \frac{\partial S}{\partial U} \right|_{V} = \left. \frac{dn}{dU} \frac{\partial S}{\partial n} \right|_{V} \tag{25.5}
\]
and so
\[
\frac{1}{T} = -\frac{k_B}{2J} \log\left(\frac{N - n}{n}\right)
\] (25.6)
which solving for \(n\) leads to
\[
\frac{1}{2}(U/J + N) = n = \frac{N}{1 + e^{\frac{2J}{k_B T}}}
\] (25.7)
resulting in
\[
U = -NJ \tanh\left(\frac{J}{k_B T}\right)
\] (25.8)
which indeed has the correct limiting behavior we expect from the temperature dependence of interacting spins. Indeed, as \(T \to \infty\) the spins are randomized and so one finds \(U \to 0\) there. On the other hand, at low temperatures, we have \(U \to -JN\) and so we see that the system magnetizes and all the spins align with each other. They can do so all pointing up or all pointing down.
Note also the symmetry of the final \(U\) with respect to \(J \to -J\). This indicates that in 1-d in the thermodynamic limit there is no thermal distinction between a ferromagnet and an antiferromagnet. An antiferromagnetic ground state is one in which adjacent spins are antialigned. Clearly this symmetry will not be true in higher dimension OR in a modified Ising model, for example, including interactions between spins that is beyond nearest neighbor.
Lecture 26

The Drude Model

As a starting point for connecting statistical mechanics with properties, we now introduce the Drude model for electrical conduction. The Drude model (Paul Drude, 1900) is a kinetic theory "ideal gas"-like model for the electrons in a metal, and, although clearly a naive, classical picture is surprisingly successful at capturing at least many qualitative features of electrical conduction.

The Drude model assumes that the electrons inside a metal are basically free, only cohesively held by the many positive charges (ion 'cores') in the metal which they experience as an overall, basically spatially homogeneous attractive potential $-V_0$. "Basically" free here means that we are also ignoring the long range (repulsive) interactions between the electrons, which, again, we think of contributing to the background overall potential. In many systems with mobile charges and polarizable cores the electric influence of each electron may be thought of as being effectively 'screened' by the subtle motions of these other charges anyways, so this assumption that the electrons don't interact at long range appears to be reasonable.

With those assumptions we have picture of the electrons forming a nearly ideal gas inside a box whose boundary is the metal. Let $N$ be the number of electrons per unit volume, each of charge $e$. The electrons are going every which way, but have an average speed $v$. With no externally applied field, there are as many electrons moving to the left as to the right, and so there is no net current in any direction.

Take a sample of this metal having cross-sectional area $A$ and length $L$. Suppose that we place a voltage $V$ across the metal. Thus we expect the electrons feel a force due to the a local internal electric field $F = eE = eV/L$. In response to this force they accelerate all in the same direction $a = F/m = eV/(mL)$. This means that the whole cloud of electrons can acquire a velocity in the direction of the applied voltage difference, that is, the can constitute a current.

However, were the cloud accelerating with acceleration $a$ the velocity would keep increasing, i.e. the current would keep increasing, an unphysical conclusion. Thus, the third major assumption in Drude’s model is that electrons individually are only ballistic for short periods of time, on average a time $\tau$, so it is only during time $\tau$ that they can accumulate some coherent net "drift" velocity $v_d = a\tau$. Think of after each $\tau$ seconds on average, the electron collides with something in the metal (a core, another electron, a defect, etc) and has its velocity randomly reset to a new direction, entirely forgetting the coherent 'drift' motion it had acquired during the previous $\tau$ seconds.

So, then, the average areal current density of this drifting cloud is $j = eNv_d = \frac{e^2NV\tau}{Lm}$, so that the total current is

$$I = Aj = \frac{Ae^2NV\tau}{Lm} = \frac{V}{\rho L/A} = V/R$$

(26.1)

We thus recover Ohm’s law where $R = \rho L/A$ is the resistance (Ohms) of the metal segment and $\rho = \frac{mN}{e^2}$ is called the resistivity of the metal, an intrinsic property of that material. Charts of resistivity are easy to find in books and on the web. Note some interesting consequences of this model;
1) The sign of the mobile charges doesn’t matter. Some materials have both + and - mobile charge carriers (semiconductors, electrolytic solutions) and all that matters is the their overall number density and their mean time between scattering events $\tau$.

2) Increasing the (charge) carrier density $N$ decreases the resistivity. This has important consequences for understanding how electrical properties vary with temperature, not just in metals, but in insulators and semiconductors and solutions...

3) The speed at which the electrons may be moving is inconsequential. All that matters is the time between collisions $\tau$. Of course, if a contribution to $\tau$ is a fixed defect density in the material, then the $\tau$ does indirectly depend on the velocity in that if sped up the electrons they would have less time between scattering events off these fixed defects. For example, suppose that the static defect density is $n$, and the cross-section for electrons to scatter off these defects is $\sigma$, then the mean-free path between scattering for the electrons is $d \sim \sqrt{\frac{1}{2n\sigma}}$ and so the $\tau = \frac{d}{v} \sim \frac{1}{\sqrt{2n\sigma}}$ and so would decrease with increased mean electron speed $v$.

Plasma Frequency - another related/useful concept for thinking about the behavior of metals in terms of a free ‘gas’ of electrons is the natural frequency at which the cloud of electrons would ‘shake’, also called the plasma frequency, $f_p$. It is straightforward to show this frequency is

$$f_p = \frac{1}{2\pi} \sqrt{\frac{Ne^2}{m\epsilon}}$$

where $\epsilon$ is the permittivity of the environment that the electrons find themselves in. Let’s work some examples of these ideas to get a better ‘feel’ for the Drude model:

**Example 0:** Compute the Drude parameters for Sodium metal. Given data is: $\rho = 47.7 \text{n}\Omega\cdot\text{m.}$ Density is 0.97 g/cc. Each sodium atom contributes $1e^{-}$ to conduction.

**Example 1:** From P. B. Johnson and R. W. Christy, *Phys. Rev. B* 6, 4370 (1972), a Drude dispersion model has for the AC permittivity of a metal,

$$\epsilon(\omega) = \epsilon_\infty - \frac{\omega_p^2}{\omega(\omega + i\tau)}$$

with $\omega_p = 2\pi f_p$ (26.3)

For the bulk silver the parameters are $\epsilon_\infty = 3.7$, $\omega_p = 1.4 \times 10^{16} \text{ rad/s}$, and $\tau = 4x10^{14} \text{ s}$. However, use a much smaller relaxation time for conduction, $\tau = 0.45 \times 10^{14} \text{ s}$.

**Example 2:** From J. R. Devore, *J. Opt. Soc. Am.* 41, 416 (1951). For TiO$_2$, a transparent oxide (in the visible light spectrum), its AC permittivity $\epsilon_{\text{TiO}_2}$ is nearly constant over that frequency range and can be well described by the formula

$$\epsilon_{\text{TiO}_2} = 5.913 + \frac{0.2441}{\frac{\lambda^2}{\lambda^2 - 0.0803}}$$

(26.4)

where $\lambda$ is the wavelength of the EM waves in microns

**Example 3:** from *Phys. Rev. B* 90, 045313 (2014), "Strong coupling in hyperbolic metamaterials,” by Prashant Shekhar and Zubin Jacob
In which they have $\frac{1}{\tau}$, the electron scattering rate, set at $1 \times 10^{13}$ s$^{-1}$, and the plasma frequency in the doped semiconductor InGaAs as $\omega_p = 9.43 \times 10^{14}$ rad/s (which they claim corresponds to a doping density of $N = 2.5 \times 10^{19}$ cm$^3$. Pretty heavy doping.

For $\epsilon_0$ use the vacuum permittivity constant, but the wrinkle here is that the charge carriers are not 'bare' electrons, but electrons that are 'dressed' by their interactions with the lattice and other electrons in such a way that they behave as if they had a mass that was $m = 0.0665m_0$, just a fraction of the bare (vacuum) electron mass $m_0$. Use these data to connect/estimate the other parameters in the Drude model for InGaAs.
Problem Set 9

The Canonical Ensemble

1) For the Ising model that we’ve discussed in class (1-d, nearest neighbor interaction J only) derive an expression for the energy fluctuation \( \Delta U^2 = \langle U^2 \rangle - \langle U \rangle^2 \). Give your answer as a function of \( N, T \) and \( J \).

2) Suppose a model quantum dot has non-degenerate states of energy \( n\epsilon \), where \( n \) is an integer greater than or equal to zero. Now consider a microsystem consisting of 4 of these quantum dots. Suppose that the total average energy of the system is \( U = 6\epsilon \). Identify the possible configurations and the average occupation number for the various energy levels if
a) The dots are gaseous bosons
b) The dots are gaseous fermions

3) A system with two non-degenerate energy levels \( \epsilon_0 \) and \( \epsilon_1 \) (with \( \epsilon_1 > \epsilon_0 \)) is populated by \( N \) distinguishable particles and is held at temperature \( T \).
   a) Show that the average energy per particle is given by
      \[
      u = \frac{U}{N} = \frac{\epsilon_0 + \epsilon_1 e^{-\beta \Delta \epsilon}}{1 + e^{-\beta \Delta \epsilon}} \quad \text{with} \quad \Delta \epsilon = \epsilon_1 - \epsilon_0 \quad \beta = \frac{1}{k_B T}
      \]
   b) Show that as \( T \to 0 \), \( u = \epsilon_0 + \Delta \epsilon e^{-\beta \Delta \epsilon} \) and that \( T \to \infty \) leads to \( u \to \frac{\epsilon_0 + \epsilon_1}{2} - \frac{\beta}{4} (\Delta \epsilon)^2 + \ldots \)
   c) Show that the specific heat at constant volume is
      \[
      c_V = k_B \left( \frac{\Delta \epsilon}{k_B T} \right)^2 \frac{e^{-\beta \Delta \epsilon}}{(1 + e^{-\beta \Delta \epsilon})^2}
      \]
   d) compute now \( c_V \) in the limits if \( T \to 0 \) and \( T \to \infty \), and sketch \( c_V \) versus \( \Delta \epsilon / k_B T \).

4) Consider a gas of \( N \) particles in volume \( V \). Around each atom there are two zones, as pictures. The inner zone (of volume \( b \)) is a hard-core repulsion into which no other atom can tread (infinite positive potential energy for another atom there). The region \( A \) has volume \( a \) and is associated with an attractive potential energy of \( -E_1 \) if another atom is in that region. Work in the dilute gas limit, thus compute the partition function of the gas as \( Z_N = \frac{1}{N!} Z_1^N \) where \( Z_1 \) is the partition function of a single gas atom in the environment of a chamber of volume \( V \) filled with \( N \) other atoms.
   a) Write down the partition function \( Z_N \) in terms of \( V, a, b, E_1 \) etc.
   b) Use (a) to determine the pressure as a function of volume and temperature... that is, derive the associated equation of state.
   c) Show that in the small \( a \) and \( b/V \) limit the the result of (b) is indistinguishable to the Van derWaals equation of state (!!)
Lecture 27

Introduction to the Canonical Ensemble

In the preceding we have used the fundamental assumption of statistical mechanics, and very general notions of energy, energy conservation and equilibrium to reveal a "link" between the microphysical picture and the description of a macroscopic state. The "link" between these two descriptions is most succinctly summarized by Boltzmann’s relation

\[ S = k_B \log(g(U)) \]  \hspace{1cm} (27.1)

And from this quantity we may ascertain the equation of state, and from there, a complete description of the equilibrium state follows.

But Eq.(27.1) is too difficult to use explicitly in all except the most simple situations. The problem is \( g(U) \). How, for example, are we supposed to compute it, even approximately, for an ideal gas? Counting states looks difficult ... and also ascertaining how to divide the fixed total energy \( U \) among the gas molecules. Indeed, it can be done in some cases, but we would like to reformulate Eq.(27.1) to make it simpler to use when we don’t know the energy \( U \) of the system precisely.

Imagine connecting the system of interest up to a large reservoir (for example think of a large reservoir of spins) at some temperature \( T \). Now, the energy of our system will not in general be constant as it may freely donate and receive energy from the bath of spins. By the fundamental assumption of statistical mechanics we take the probability of each microstate describing the same macrostate to be the same. Let the total energy of our system and the (spin-) heat bath be \( U + U_1 \) where \( U_1 \) denotes the energy in the system that we want to focus on. Now, suppose that the energy of our small subsystem goes up somewhat, to \( U_1 + E \). Then, assuming \( E \) to be small, the system and bath are still nearly at their equilibrium macroscopic states. Thus, with a small error the relative probability of such an excursion \( P(E)/P(0) \) is simply the reciprocal of the total number of states accessible to the system+bath system

\[ \frac{P(E)}{P(0)} = \frac{g_1(U_1)}{g_1(U_1 + E)} = \frac{g_b(U - E)}{g_b(U)} \]  \hspace{1cm} (27.2)

where the last equality follows from the large bath limit, and the fact that if you take a small amount of energy out of a very large heat bath, \( g_1(U_1)g_b(U) \sim g_1(U_1 + E)g_b(U - E) \). After all, the equilibrium partitioning of the energy into \( U_1 \) and \( U \) is determined by the flatness of the product of the multiplicity functions of the bath and the system.

Now, we have from Eq.(27.1) that \( g(u) = e^{S(U)/k_B} \) and so Eq.(27.2) may be written

\[ \frac{P(E)}{P(0)} = e^{\frac{S_b(U-E) - S_b(U)}{k_B}} \]  \hspace{1cm} (27.3)

in which \( S_b(U) \) refers to the entropy of the very large bath. Since \( E \) is small w.r.t. \( U \) we may expand the exponent out to linear order to find

\[ \frac{P(E)}{P(0)} = e^{-\frac{E}{k_B}} = e^{-\frac{E}{k_B}T} \]  \hspace{1cm} (27.4)

Again, this is the relative probability of a state in equilibrium at temperature $T$ experiencing a gain in energy $E$.

We now see that we may reformulate our description of probability of a particular microstate. Instead of describing it as $1/g(U)$, where $U$ is the total (fixed) energy of the state, we may talk about the probability at temperature $T$ that the system has absorbed (variable) energy $E$ as $\sim e^{-E/k_B T}$. This factor is called Boltzmann’s factor.

Before we spoke about the (micro-canonical) ensemble $\mathcal{E}$ as that collection of microscopic states that would be regarded as the same macroscopic state. Different macroscopic states can be distinguished between on the basis of their total energy. Therefore if we now allow our system to absorb energy from the reservoir, we would no longer be at a microstate in our original ensemble! We must henceforth enlarge our notion of ensemble to include all configurations of the system regardless of their energy. This ensemble of configurations is called the canonical ensemble and we will denote it by $\mathcal{E}'$.

Since it gives us the relative probability of a state $f \in \mathcal{E}'$ absorbing a certain amount of energy from a bath at a fixed temperature, we can write an absolute (normalized) probability simply by requiring that the sum of the probabilities of all the configurations should equal one. That is, to convert from relative to absolute probability we need a normalize our probabilities with a constant $Z$

$$P(E) = \frac{1}{Z} e^{-\frac{E}{k_B T}}$$

so that requiring the sum over all (absolute) probabilities the configurations $f$ in the ensemble $\mathcal{E}'$ for our system to equal one

$$\sum_{f \in \mathcal{E}'} P(E(f)) = 1$$

thus we see immediately that this implies that

$$Z = \sum_{f \in \mathcal{E}'} e^{-\frac{E}{k_B T}} = \sum_{f \in \mathcal{E}'} e^{-\beta E} \quad \beta = \frac{1}{k_B T}$$

This function $Z$ is called the partition function. It will play a central role in the further development of the theory.

We will see that the partition function also has complete thermodynamic information. As a step towards demonstrating that, we can readily demonstrate how one can use the partition function to compute the internal energy $U$. The internal energy for a system that is connected to a heat bath is just its average energy $U = \langle E \rangle$. This is just the energy of each microstate weighted by the probability of being in that state, given by Eq.(27.5) in terms of the states energy $E$ only. Thus

$$U = \langle E \rangle = \sum_{f \in \mathcal{E}'} E(f) P(E(f)) = \frac{1}{Z} \sum_{f \in \mathcal{E}'} E e^{-\beta E}$$

which we realize as the negative of the derivative of the partition function w.r.t. $\beta$,

$$U = -\frac{1}{Z} \partial_\beta Z = k_B T^2 \partial_T \log(Z)$$

These few lines will be very useful in the following lectures and captures the economy of the statistical mechanical formulation in terms of the partition function.
Application: Binary Spin System Let’s revisit the non-interacting spin system of Lecture 24 and compute its average energy in terms of the partition function.

We have \( Z = \sum_{\mathbf{E}} e^{-\beta \mathbf{E}} \) with \( \mathbf{E} = -m B \sum_{i=1}^{N} \sigma_i \) where \( \sigma_i = \pm 1 \) and the set of states \( \mathbf{E} \) being all possible spin configurations of the \( N \) spins. Now, note that \( e^{-\beta \mathbf{E}} = \Pi_i e^{-\beta m B \sigma_i} \) and so the sum over independent spin arrangements breaks into a product over the partition functions of the individual spins! This simplification is generic for all systems in which there are uncoupled subsystems...more on that later. So for now,

\[
Z = \sum_{\mathbf{E}} e^{-\beta \mathbf{E}} = \Pi_i \sum_{\sigma} e^{-\beta m B \sigma} = (e^{\beta m B} + e^{-\beta m B})^N \quad (27.11)
\]

And we can now compute the average energy \( U = -\partial_{\beta} \log(Z) \) which you can verify reproduces the earlier result of Lecture 24...with much less work!
Lecture 28

The Ideal Gas

As another example of how one may use the canonical ensemble and the partition function to describe a system, let's apply these ideas to a familiar system: the ideal gas.

Consider the lesson from last time: for independent (and non-interacting) distinguishable particles, we know that the partition sum for $N$ particles factorizes into the product of the partition sum for the individual particles.

$$Z_{distinguishable} = Z_1^N$$

(28.1)

This was the case for the "discontinuities" in the spin chain from last time. For an ideal gas, since each of the particles are identical (undistinguishable), were we to switch any two particles (both their positions and momenta) we would have the same configuration. Thus, for a gas of undistinguishable particles, Eq.(28.1) over counts the number of distinct configurations contributing to the partition sum. It overcounts precisely by a factor of $N!$, since there are $N!$ ways of switching a "labeling" on the $N$ particles and since the particles are indistinguishable, all labelings must be regarded as the same configuration. Thus, for a gas of indistinguishable non-interacting particles,

$$Z_N = \frac{1}{N!} Z_1^N$$

(28.2)

If you have difficulty understanding the above argument, accept Eq.(28.2) as a magic formula for now. I hope that what comes next will clarify things a bit.

Thus, to compute the partition function of a gas of $N$ particles, we need only compute the partition sum of a single molecule of that gas and then use Eq.(28.2). Now, recall that the partition sum is

$$Z = \sum_{\text{configurations}} e^{-E/k_B T}$$

(28.3)

where the sum is over distinct microphysical configurations. Now, for a single gas molecule in a container of volume $V$, what constitutes distinguishable configurations? Well, certainly the position $\vec{x}$ and the momentum $\vec{p}$ at that position uniquely determine the "state" or "configuration" of that molecule! Thus, we really need to perform the sum

$$Z_1 = \sum_{\vec{x},\vec{p}} e^{-E(\vec{x},\vec{p})/k_B T}$$

(28.4)

But how do we perform the sum, as there is a continuum of positions $\vec{x}$ and $\vec{p}$? One approach is to replace the sum with an integral. This replacement can be made semi-rigorous several different ways, and we will spare the reader details (that may otherwise be presented in the lecture). Thus, with this understanding of how to perform the sum,

$$Z_1 = G \int d^3\vec{x} d^3\vec{p} e^{-E(\vec{x},\vec{p})/k_B T}$$

(28.5)

Where $G$ is some overall normalization constant that we use to make this replacement of the sums by integrals (note below that $G$ will never enter in to any physical quantity).
Now, for a simple ideal gas, the energy $E(\vec{x}, \vec{p})$ of a particle does not depend on the position of the gas molecule BUT only upon its momentum (we will generalize this all later.)

$$E(\vec{x}, \vec{p}) = \frac{\vec{p}^2}{2m}$$  \hspace{1cm} (28.6)

where $\vec{p} = (p_1, p_2, p_3)$ and $m$ is the mass of the molecule. Since $E$ is independent of the position of the molecule, we may easily perform the $d^3 \vec{x}$ integral to find Eq.(28.5) reduces to

$$Z_1 = GV \int \int \int dp_1 dp_2 dp_3 e^{-\left(p_1^2 + p_2^2 + p_3^2\right)/k_B T}$$  \hspace{1cm} (28.7)

Where recall, the volume $V = \int d^3 \vec{x}$ and $d^3 \vec{p} = dp_1 dp_2 dp_3$. We can now use our formula sheet passed out at the beginning of the semester to compute this gaussian integral. Looking back to that sheet we find

$$Z_1 = GV \left(2\pi mk_B T\right)^{\frac{3}{2}}$$  \hspace{1cm} (28.8)

and so, using Eq.(28.2), we can write down the partition function of a gas of $N$ molecules as

$$Z = \frac{G^N V^N}{N!} \left(2\pi mk_B T\right)^{\frac{3N}{2}}$$  \hspace{1cm} (28.9)

Next, let’s use Eq.(28.9) and the relation from last lecture to compute the average energy of the gas. We know that from Eq.(27.10) we have

$$U = \langle E \rangle = \frac{1}{Z} \frac{\partial Z}{\partial T} = \frac{3}{2} Nk_B T$$  \hspace{1cm} (28.10)

which agrees with our earlier intuition from kinetic theory!

We now go on to show that the partition function $Z$ really has information about other thermodynamic functions.

**Fluctuations**: We can easily construct other observables using the partition function. For example, we can find the average value of the energy squared, $\langle E^2 \rangle$. Indeed,

$$\langle E^2 \rangle = \sum_{s \in \text{configs}} P(E_s) E_s^2 = \frac{1}{Z} \sum_{s \in \text{configs}} E_s^2 e^{-\beta E_s} = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}$$  \hspace{1cm} (28.11)

where we have used the trick that each time we differentiate with respect to $\beta$ we “pull down” a factor of the energy. Note that $\langle E^2 \rangle$ is not $\langle E \rangle^2$! In general the average of a power $\langle \mathcal{O}^n \rangle$ of an observable is not just the power of its expectation value $\langle \mathcal{O} \rangle^n$, the difference being that fluctuations about the expected average value contribute to $\langle \mathcal{O}^n \rangle$ but not to $\langle \mathcal{O} \rangle$ itself.

One can compute virtually every desired thermodynamic property from the partition function.

**Application 1**: Spin system computation of $\Delta U = \langle E^2 \rangle - \langle E \rangle^2$...pretty straightforward since we have $Z$ and have already computed $U = \langle E \rangle$ in the previous lecture. Do it and connect it with...

**Connection with Random Walk and Diffusion**: Note the connection between the random walk and non-interacting spin system. $\sigma = 1$ is a jump to the right and $\sigma = -1$ is a jump the same distance BUT to the left. Then the total net distance you have gone
is \( d = \sum_i \sigma_i \), which is proportional to what we had called the internal energy in the non-interacting spin system. SO, look at the < \( d^2 \) > in the limit of large \( N \).

**Connection with extensivity:** In the microcanonical ensemble approach, uncorrelated systems '1' and '2' have a joint multiplicity function \( g(U) = g_1(u_1)g_2(U - u_1) \) and since the entropy is a logarithm of the multiplicity function, the entropy of the whole system is the sum of the entropies of the uncorrelated systems.

Similar story with the canonical partition function. In that case, for uncorrelated distinguishable systems, the total partition function \( Z = Z_1Z_2 \) is just the product. They are connected and so assumed to be in thermal equilibrium (that is, there is only one \( \beta \) in the product) so that - again since the log of a product is the sum of the logs - the total energy as given by Eq.(28.10) for example, is the sum of the energy of the subsystems.

SO basically extensitivity in the classical thermo context here is a consequence of (1) the systems becoming uncorrelated and (2) the logarithmic connection between probability and thermodynamic observables.

**The Gibbs Paradox:** Gibbs noticed that this scaling leads to a distinctly different entropy function depending on whether the molecules making up a gas were identical OR distinguishable. Consider two chambers of the same volume filled with the same gas at the same temperature and pressure. If the entropy of each chamber is \( S \) then the total entropy of the system is \( 2S \). But according to Boltzmann, the entropy is also the logarithm of the multiplicity function \( g \). The systems are uncorrelated SO the total \( g \) of the combined system is \( g = g_1^2 \) and then indeed the entropy of the total system is double the entropy of the parts.

Note the volume dependence of the multiplicity function. To put \( N \) molecules in a volume \( V \), first decimate the volume into \( M \) cells, each of volume \( V/M \). Ignoring for now the momentum factor of the multiplicity function (which won’t depend on the volume), we find that the \( g \sim M^N \) so that in the limit of large \( M \), we see that the entropy goes as \( S = k_B \log(g) = N \log(V) + \ldots \) where \( \ldots \) are volume independent terms.

Now look at combining these chambers, say, by removing a plenum between them. Nothing really changes. But each gas gets to expand into the new volume (of the other gas). Since we have shown that the entropy of a gas scales with the log of the volume, we see that now each gas of \( N \) atoms now has entropy \( S = Nk_B \log(2V) \) so that the entropy has somehow increased by removing the plenum between the chambers, going from \( S = 2Nk_B \log(V) \) before removal to \( S = 2Nk_B \log(2V) \) after removal, indicating an entropy change of \( S = 2Nk_B \log(2) \) even though nothing has really changed!

The resolution of the paradox is to realize that since these particles of gas are identical, we have overcounted the states; really the placement of the atoms into the \( M \) boxes should be counted up to an overall permutation of the \( N \) atoms with each other. There are \( N! \) ways of permuting \( N \) objects... and \( (2N)! \) ways of permuting \( 2N \) atoms. Thus, the \( g \)'s are really \( g \sim M^N/N! \) so that the entropy is really \( S = Nk_B \log(V/N) \) in the large \( M \) limit.

With this understanding now, before you remove the plenum, you have \( S_{\text{tot}} = S_1 + S_2 = 2Nk_B \log(V/N) \). After you remove the plenum you have \( S_{\text{tot}} = k_B \log((2V)^{2N}/(2N)!) \) which limits to the same expression. Once identical gasses mix, the counting of distinct microstate reduces the logarithmic increase due to the volume change.

Of course, this is not the case if the gas molecules are different...then there is a real change in the entropy, also called the entropy of mixing.
Maxwell Distribution

We can use the result of last time to derive a formula for the distribution of velocities of atoms in a gas. We know that the probability of each atom having an energy $E$ is

$$P(E) = \frac{1}{Z} e^{-\frac{E}{k_B T}} \quad (29.1)$$

but, assuming that this energy is almost entirely kinetic energy, (which it is for most gases at room) $E = \frac{1}{2} m \vec{v}^2$ we find

$$P(\vec{v}) = \frac{1}{Z} e^{-\frac{m \vec{v}^2}{2 k_B T}} \quad (29.2)$$

Where, using the fact that

$$\sum_{\vec{v}\epsilon v} P(\vec{v}) = 1 \quad (29.3)$$

and that

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} \, dx = \sqrt{\pi} \alpha \quad (29.4)$$

we can find the value of $Z$, the normalization constant of the probability function. From Eq.(25.3) it is

$$Z = \int d\vec{v} e^{-\frac{m \vec{v}^2}{2 k_B T}} = \int dv_x dv_y dv_z e^{-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2 k_B T}} = \left(\frac{2 \pi k_B T}{m}\right)^\frac{3}{2} \quad (29.5)$$

Thus

$$P(\vec{v}) = \left(\frac{m}{2 \pi k_B T}\right)^\frac{3}{2} e^{-\frac{m \vec{v}^2}{2 k_B T}} \quad (29.6)$$

which is called the Maxwell-Boltzmann distribution.

Another Definition:

$$v_{rms} = \sqrt{\langle \vec{v}^2 \rangle} \quad (29.7)$$

Discuss applications of this formula: Average Kinetic Energy
Connection to the Free Energy, etc.

**Goal:** To identify the log of the partition function as the free energy divided by $k_B T$.

We now show that the canonical ensemble can be used to find all the thermodynamic quantities of interest. Recall that the Helmholtz free energy is $F = U - TS$ and that

$$dF = -SdT - PdV$$

and so $S = -\frac{\partial F}{\partial T}|_V$. We may bring this term to the other side of the above equation to write,

$$U = F - T \frac{\partial F}{\partial T}|_V = \frac{\partial \beta F}{\partial \beta}|_V$$

(29.8)

Now, from the previous lecture we know that $U = -\frac{\partial \log(Z)}{\partial \beta}$ where the partition function $Z$ has been computed at fixed volume. Thus we are led to the identification,

$$\beta F = -\log(Z) + \text{const.}$$

(29.9)

where the constant is independent of the temperature and can be absorbed into the definition of $Z$, thus,

$$F = -k_B T \log(Z)$$

(29.10)

Since the natural variables of $F$ are $V$ and $T$, it is clear that if we compute the partition function at constant $V$, we can use Eq.(29.10) to find every thermodynamic function.

**Application 1a:** Ideal gas equation of state. Write down the expression for $F$ and show that it leads (via the defining relations for $P$ for example) to (1) the ideal gas equation of state and (2) the ideal gas internal energy formula.

**Application 1b:** We can do almost the same calculation as in 1a but with a gas that has hard core repulsion. Assume around each gas molecule there is a region of volume $b$ from which other molecules are excluded. Recompute the $Z$ and then the $F$ using Eq.(29.10) and from there the new equation of state. Now look back at the equation of state for the Van der Waals gas. Then derive an expression for the total internal energy of this gas. Does the formula look familiar? Why?
**Application 2:** Quantum harmonic oscillator (see the next lecture).

**Further Techniques:** A very general and useful method for computing the partition function of an interacting spin system is the transfer matrix method. Although we have already done the thermodynamics of the 1-d Ising model, one may ask how the picture might change were we to include, for example, an external magnetic field. In that case the straightforward map of products of neighboring spins into sums of independent pseudospins is not possible. The problem, and others like it, seem intractable analytically UNTIL one learns the transfer matrix technique.

Consider a spin system (1-d for simplicity, but this can be generalized) with nearest neighbor interactions and an external field we call $H$. The Hamiltonian (energy function) for the system is thus

$$E(\sigma) = \sum_{i=1}^{N} -J\sigma_i\sigma_{i+1} + H\sigma_i$$  \hfill (29.11)$$

where the spins can be $\sigma = \pm 1$. We wish to compute the partition function $Z = \sum_{\text{config}} e^{-\beta E(\sigma)}$ where the sum ranges over all possible spin configurations. Note that in this sum, the individual spin contributions become products over the possible spins at that site, schematically as,

$$Z = \Pi_{i=1}^{N} e^{\beta J\sigma_i\sigma_{i+1}} e^{-\beta H\sigma_i}$$  \hfill (29.12)$$

where we must include in each product the contribution from the spin from the next site. One very useful way to do this is to think of each factor in the above as a matrix quadratic form that is a function of the $\sigma_i$ thought of as vectors, that is, think of the $\sigma_i$ as being (1,0) if spin up and (0,1) if spin down. Then in terms of a 2x2 matrix, we can write

$$M_i = e^{\beta J\sigma_i\sigma_{i+1}} e^{-\beta H\sigma_i} = \begin{bmatrix} e^{\beta(J-H)} & e^{\beta(-J-H)} \\ e^{\beta(-J+H)} & e^{\beta(J+H)} \end{bmatrix}$$  \hfill (29.13)$$

where we think of this matrix as being between the possible states $\sigma_i$ (on the RHS) and $\sigma_{i+1}$ (on the LHS of the matrix). Multiplying matrices $\Pi_{i=1}^{N} M_i$ automatically keeps track of all the terms in the space of all configurations of spins! Finally, thinking of the $N$ as large and for convenience tying the ends of the string together into a big loop (that is, letting the $N$'th spin be next to the $i = 1$ spin) we can compute the partition function Eq.(29.12) via

$$Z = \Pi_i M_i = M^N$$  \hfill (29.14)$$

where the last thing follows from the fact that we assumed $J$ and $H$ were not varying with location, thus all the $M_i$ were the same. Now, since $N$ is large, the Eq.(29.14) can be approximated at large $N$ by the largest eigenvalue of the $M$ matrix raised to the $N$'th power. Which, when you think about it, means that you have essentially reduced the system to a non-interacting spin system again, but retained the dependence on the parameters $J$ and $H$. Very cool....

We can be more explicit about this ....
Mathematical aside: To help you recall what you learned about eigenvectors and eigenvalues....

Lecture Question: Use Eq.(29.10) to derive an expression for the free energy of an ideal monoatomic gas.
Further Applications

**Equipartition Theorem:** The equipartition theorem indicates that for a classical system bound to a heat bath of temperature $T$ the average energy in a quadratic degree of freedom is $k_B T / 2$. By quadratic degree of freedom we mean any independent motion/momentum that appears quadratically and isolated in the microphysical formula for the total energy. Although this sounds very specialized, it has a wide approximate use and is very useful for complicated coupled and non-linear systems.

A system's (classical) degrees of freedom are the co-ordinates you use to classically specify the state of the system. For a gas this could be a list of the positions $(x, y, z)$ and momenta $(p_x, p_y, p_z)$ of each atom (Question: why don’t you need the co-ordinate locations momenta of each electron and nucleon instead? What temperature regime are we mostly working in and physically why does that help you answer this question?) To summarize, for a gas of $N$ atoms, the gas has $6N$ degrees of freedom since each atom needs 6 numbers to completely specify its position and motion. Note that for a diatomic molecule we would need to specify not only the location (of the center of mass of each molecule) and the (center of mass) momentum of the whole molecule BUT we also need to specify the molecular orientation and how fast it is spinning about its center of mass.

(SPECIAL QUANTUM MECHANICAL NOTE ABOUT ACCOUNTING FOR ANGULAR MOMENTUM:) Isolated atoms can have a net angular momentum, think spinning tops. However this fixed net angular momentum cannot change without typically al lot of energy. So at low temperatures this internal angular motion degree of freedom makes no contribution to the thermodynamics. For molecules a similar situation exists except that since it is so much larger, the angular motions can be excited at much lower temperatures. Think that larger wavelengths mean lower energy. But there is a wrinkle here...we haven’t said anything about the motion wrt what axis. Here’s the deal; for a diatomic molecule, at relatively low temperatures it is easy to get it rotating about an axis perpendicular to the bond direction. There are two such directions. However, to get it rotating around the axis along the bond is about as hard as rotating the isolated atoms by themselves (think about what you would see if you looked down that axis, afterall). Thus, for reasonable (whatever that is!) temperatures we only have angular motion in a diatomic molecule about two axes. For a non-symmetric general non-linear triatomic and higher molecule (for example, water) however there are three independent angular motions about the three principle axes of the moment of inertia tensor.

SO to summarize, its 0, 2, 3 for atom, diatomic molecule and larger molecule respectively for the angular degrees of freedom of the molecule. Thus, to form the partition function of the system, with only $N$ quadratic degrees of freedom (though there may be other DOF on which the energy depends in a non-quadratic way), we need to sum (but, as for the ideal gas where we had classical DOF, we need to integrate) $e^{-\beta E}$ over all the configurations of the system. Since the $E$ only depends quadratically on these DOF, the integration is just gaussian integration, and so,

$$Z = \int dx_1 \ldots dx_R dp_1 \ldots dp_R e^{-\beta (\sum_i (\alpha_i x_i^2 + \gamma_i p_i^2))} = \# \beta^{R/2}$$

where $R$ is for example $6N$ or otherwise the total count of the quadratic degrees of freedom (and $\#$ is some overall factor of no thermodynamic consequence). Then, by Eq.(28.10), we
have that $U = \frac{R}{2} k_B T$. Thus, by way of review, for the monoatomic ideal gas of $N$ atoms, the energy for each particle is $E = \frac{(p_x^2 + p_y^2 + p_z^2)}{2m}$ we have $R = 3N$ total degrees of freedom, so the Eq.(28.10) indicates that the total internal energy is $U = \frac{3}{2} N k_B T$. Let us delineate some other useful cases:

1) **Diatomic ideal gas**: For the diatomic gas of $N$ molecules, the energy per particle goes as $E = \frac{(p_x^2 + p_y^2 + p_z^2)}{2m} + \frac{1}{2I_x} L_x^2 + \frac{1}{2I_y} L_y^2$ where $L_x$ and $L_y$ are the total (classical) angular momentum about the $\hat{x}$ and $\hat{y}$ axis, here assumed perpendicular to the molecular axis $\hat{z}$. Thus, $R = 5N$ and the internal energy is $U = \frac{5}{2} N k_B T$ BUT the equation of state is the usual ideal gas law.

2) **More complicated molecular gas**: Take a more complicated molecular gas, for example, ammonia, NH$_3$. In toy pictures of ammonia it is viewed as a pyramidal structure. (They way physicists think about ammonia in its ground state is more like a platonic ball with a core of nitrogen and a surface of hydrogens all spread out smoothly and evenly on the ball...but that is another story for another time...) In this case, a gas of $N$ molecules would have a $R = 6N$ DOF since there are three translation and three rotation axis, and the energy per molecule goes something like $E = \frac{(p_x^2 + p_y^2 + p_z^2)}{2m} + \frac{1}{2I_x} L_x^2 + \frac{1}{2I_y} L_y^2 + \frac{1}{2I_z} L_z^2$. Thus, $U = 3N k_B T$ and, again the equation of state is the same ideal gas law.

3) **Molecular gas with vibration and rotation**: Take for example the case of a simple diatomic gas like HCl. At a high enough temperature (but well before molecular dissociation) the molecule not only translates and spins about its CM, BUT ALSO the hydrogen can bounce in- and away- from the Cl atom. The idea is that there may be a complicated interatomic potential energy -vs- separation graph, but the molecule in its ground state 'finds' the minimum of this graph. Thinking classically, there is a huge repulsion if the H and the CL are right on top of one another and yet it costs energy to dissociate the molecule SO the H doesn’t like to be too far away from the CL. The 'just right' distance is the putative ground state bond length of the molecule, or thought of as a physicist, the radius of the platonic ball with Cl mostly in the center and a smooth uniform ocean of hydrogen on the surface. Near that minimum the interatomic potential looks quadratic! That is, we can associate an internal quadratic degree of freedom to the relative motion of the Cl and H along the bond direction. This means that for sufficiently high temperatures at which we can excite that motion (Note that due to quantum mechanics, there is a minimum temperature below which you almost never excite this mode...it is “frozen out” ) we would need to add a new quadratic energy term to the energy per molecule, that is, $E = \frac{(p_x^2 + p_y^2 + p_z^2)}{2(m_1 + m_2)} + \frac{1}{2\mu} L_x^2 + \frac{1}{2I_y} L_y^2 + \frac{1}{2I_z} L_z^2 + \frac{k}{2} \hat{z}^2 + \frac{1}{2\mu} p_z^2$ where $\hat{z}$ is the direction along the Cl - H axis and $z$ is the measure of the stretch of that bond. In this energy, the $P$ are the total (center of mass) momentum and the $p$ is the relative momentum of the Cl and the H (with $\mu = m_1 m_2/(m_1 + m_2)$ the reduced mass) Thus, for this case, in a gas of $N$ molecules there are $R = 7N$ dof meaning that the total internal energy in this case is $U = \frac{7}{2} N k_B T$. 

(final thing to do: graph crudely the transition of the $C_v$ as you move from below to above the temperature at which you can excite this spring motion).

4) **Einstein Solid:** Debye’s model for a solid is that of mass points all connected to their nearest neighboring mass points with springs. In a crystal the mass points would occupy each location on a lattice, whereas in an amorphous solid they are more randomly connected to one another. Whatever the case may be, and however complicated the model would be with springs attaching masses, etc, at the end of the day for small excursions from their equilibrium positions, one will find that each mass point has three orthogonal directions along which there is a net spring constant-type force acting on the masspoint. To better understand the connection between these melange of springs and this simplifying fact (just three orthogonal directions for each mass) takes a we bit of linear algebra which we are not going to describe here.

But the conclusion is that for elastic vibrations in a solid, we can think of there being $N$ mass points each having three spring directions. Debye was the first to solve for the spectrum of the elastic modes in a crystalline solid of this type, but Einstein, noting that the equipartition theorem just counts quadratic modes in the energy but independent of the coefficients $\frac{1}{m}$ or $k$, realized that for a solid the energy of each atom (i.e. mass point) is $E = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) + \frac{1}{2}(k_z z^2 + k_y y^2 + k_x x^2)$ and so, by the equipartition result, expected that the total internal energy of a substance should be $U = 3Nk_B T$. This used microphysics to explain the **Dulong-Pettit Observation**, that most elements have a universal specific heat always close to $3Nk_B$ where $N$ is the total number of atoms. This “law of Dulong and Pettit” is very useful also for salts (ionic solids) and certain simple molecular solids.

They noted furthermore that the measured specific heats approach this universal limit as one goes to higher temperature (but still remain solid). Later in this book we will revisit Debye’s model for a solid in more detail to understand this fact.

5) **Virial Theorem:** Here we review the Virial Theorem you learned in classical mechanics, eventually relating it to the Equipartition theorem.

From classical mechanics class, you know that if there is a system of $N$ particles interacting through a homogeneous potential of scaling dimension $\alpha$ (by a homogeneous potential, we mean the interparticle potential $V(x) = Bx^\alpha$ where $x$ is a spatial co-ordinate and $B$ some overall constant), then as long as the system is stable (i.e. all the orbits are bounded) it must be that the (time-) average potential energy and (time-) averaged kinetic energy are related to one another

$<P.E.> = \frac{2}{\alpha} <K.E.>$ \hspace{1cm} (30.5)

Note that $\alpha = 2$ for a harmonic (spring-like) potential and $\alpha = -1$ for the coloumb potential. Of course, the total energy is $U = <P.E.> + <K.E.> = (1 + \frac{2}{\alpha}) <K.E.>$ and since the kinetic energy is always quadratic in the co-ordinate (and thus its thermal (phase-)average is $\frac{2}{\alpha} k_B T$ as given by the equipartition theorem) we expect for the thermal average of the system that

$U = \frac{3(\alpha + 2)}{2\alpha} Nk_B T$ \hspace{1cm} (30.6)

Here are some simple conclusions from this result. First, as one can check by actually doing the integrations, the virial result above is consistent with the actual computation of the partition function. That is, we can generalize it to the following statement; Any term in the
energy functional that depends on a co-ordinate dof (say \(x\)) that has scaling dimension \(\alpha\) contributes a term \(\frac{1}{\alpha} k_B T\) to the average total energy of the system at temperature \(T\). This should strike one as very curious...we have apparently shown in the case of homogeneous potentials that the temporal average is the same as the phase average...i.e. that these systems seem to be ergodic.

Second, we note the result that gravitational systems must be thermodynamically unstable. Its not so much that the total internal energy being negative (for gravity \(\alpha = -1\)) that is disturbing, as the energy is always defined up to an overall constant anyways, BUT that the specific heat of the system would have to be negative. We will in what follows show that this is true in particular situations including stars and blackholes.

6) Ultra-violet “catastrophe” - Take a look at the waves on a string. There are independent modes of standing waves associated with the number of nodes that a wave has. But a there are \(Z^+\) of them...that is an infinite number of such modes. If you were to write down the energy function for each mode, it would be \(E = \frac{p_y^2}{2m} + \frac{ky^2}{2}\) where \(p_y\) is a momentum of the string pieces in the vertical direction and \(m\) is related to the mass of the string whereas \(y\) is the vertical amplitude of the string wave and \(k\) is basically the string tension divided by the length. Forget these details...what is important is that for each vibration mode there are two quadratic modes. This means that by the equipartition theorem at equilibrium at temperature \(T\) the string is not just sitting there but has \(k_B T\) in each oscillatory mode.

While that is not much energy for a large string, there are an infinite number of these modes; so the total internal energy of the string should be infinite! That naive conclusion is called the ultraviolet catastrophe. Return to the string; clearly the string cannot support all the modes. Indeed even for a nearly perfect thin string there is no way that the string can support standing wave modes of a wavelength smaller than the interatomic spacing...SO the maximum node number for a string of atoms is about the number of atoms...returning our accounting the energy to the law of Dulong and Pettit. So, for sound vibrations we are ‘safe’ and the story makes sense.

BUT the real problem with this thinking is that the same considerations apply to the radiation field...i.e. light. Suppose that we have a long thin cavity with mirrors on both ends. If the mirrors come into equilibrium with a heat bath at temperature \(T\) they will radiate back and forth into the cavity until the radiation field comes into equilibrium with the mirrors. Each mode has the same energy formula as the string...it still has two quadratic dof. But now there is no limit to how small the wavelength of the waves can be...always an infinite number of waves fit the cavity, again leading to the conclusion that the cavity at any temperature should have an infinite energy density. The ultraviolet catastrophe for light in a cavity is harder thing to argue away. The ultimate resolution of this ‘paradox’ is that modes cannot be occupied at arbitrarily low amplitude... there are discrete amounts of energy that can excite the string...of course this is a quantum mechanical resolution.
Quantum Mechanical Harmonic Oscillator: We now compute the partition function of a single harmonic oscillator... think of this as a single mode of radiation that fits in the cavity. In contrast to the classical harmonic oscillator (done in the previous lecture on the equipartition theorem), a quantum harmonic oscillator energy levels are discrete and spaced by $\hbar \omega$ where $\hbar = 1.0545714810^{-34}$ m$^2$-kg / s AND $\omega$ is the natural resonant frequency of the wave/system (in rad/s...recall $\omega = 2\pi f$ where $f$ is the frequency). For example, if talking about a spring and a mass the $\omega = \sqrt{k/m}$. The energy levels are

$$E_n = (n + 1/2)\hbar \omega$$ (31.1)

where $n \in \mathbb{Z}^+$ and $n = 0$ is associated with the ground state. Note that the energy of the harmonic oscillators’ ground state is not zero but is $\hbar \omega/2$. This curious fact, that at the lowest energy state the system still has some energy and motion in it, will be elaborated in your quantum class.

Since there is one state for each $n \geq 0$, we can form the partition sum for a quantum harmonic oscillator:

$$Z = \sum_{\text{configs}} e^{-\beta E_{\text{config}}} = \sum_{n=0}^{\infty} e^{-\beta E_n} = e^{-\beta \hbar \omega/2} \sum_{n=0}^{\infty} e^{-n\beta \hbar \omega} = \frac{1}{2 \sinh(\frac{\beta \hbar \omega}{2})}$$ (31.2)

And so using Eq.(28.10), we compute the average total energy at temperature $T$:

$$U = \langle E \rangle = \frac{\hbar \omega}{2} \coth(\frac{\beta \hbar \omega}{2})$$ (31.3)

Let’s investigate the limits once again to get a deeper feel for this result. At low temperature, as $\beta \to \infty$, the $\coth(\infty) \sim 1$, $\frac{\hbar \omega}{2} \coth(\frac{\beta \hbar \omega}{2})$ meaning the average energy just goes over to the ground state energy $\hbar \omega/2$. This makes sense in that as you go to very low temperature the system cannot absorb enough energy from the bath, on average, to even get into the first excited state, and so stays stuck in the ground state.

In the opposite limit, the high temperature limit, $\beta \to 0$, the $\coth(\beta x) \sim 1/\sinh(\beta x) \sim 1/(\beta x)$. Thus in the high temperature limit we recover the equipartition result (i.e. the classical harmonic oscillator limit) $U = k_B T$.

A particularly useful way of writing the result Eq.(31.3) is in terms of a average expectation value,

$$U = \langle E \rangle = E_{<n>} = \hbar \omega(< n > + 1/2) = \frac{\hbar \omega}{2} \coth(\frac{\beta \hbar \omega}{2})$$

so that

$$< n > = \frac{1}{e^{\beta \hbar \omega} - 1}$$ (31.4)

This mean occupation number of a quantum harmonic oscillator at temperature $T$ is called the Planck distribution or Bose-Einstein distribution function. We’ll use it more fully later.
when we talk about so-called black body radiation (that is, a field of light in equilibrium with a heat bath).

For now, however we wish to point out a useful correspondence. In the quantum mechanical harmonic oscillator, all the energy levels are equally spaced. Another way of thinking about the harmonic oscillator levels are as a system with just two levels, one at energy \( \bar{h}\omega \) and the other at an energy an \( \hbar \omega \) higher. Now rather than think of the quantum harmonic oscillator excited to level \( n \) alternatively as \( n \) “particles” in the on excited state of the two level system the energy differences between the both pictures and their respective ground states being the same (but the total energies as being different). We call putting \( n \) particles in the same state a ‘degeneracy’ or ‘multiplicity’ or “occupation” of the state. In this way of thinking, then, the Eq.(31.4) would be for the mean occupation number (number of particles) that exist at that temperature. Particles that can multiply inhabit the same quantum state like this are called **Bosons**. Some example of real world bosons are photons, phonons, gauge particles like gluons (strong nuclear force carriers) and \( W^\pm \) and \( Z \) (electroweak force carriers), but also importantly at low energies atoms with a total even number of fermions such as \(^4\text{He}, \ 7\text{Li}, \ 87\text{Rb}.\)

**Fermi-Dirac particles** All this “multiplicity” of identical excitations is quite the opposite of our lowly non-interacting spin system. Recall there the energy of a single spin was \( E = -mB\sigma \) where the spin \( \sigma = \pm \). Thus, the partition sum for this single sum is

\[
Z = e^{-\beta mB} + e^{\beta mB} = 2cosh(\beta mB) \tag{31.5}
\]

and if we identify the energy difference as \( 2mB = \hbar \omega \), then then we have for the partition function of the spin system,

\[
Z = 2cosh(\beta mB) \rightarrow U = -\frac{\hbar \omega}{2}\text{tanh}\left(\frac{\beta \hbar \omega}{2}\right) \tag{31.6}
\]

Once again we study the high and low temperature limits. In the high temperature limit \( \beta \to 0 \), we have \( U \to 0 \). Thus, each level is equally occupied. In the low temperature limit, \( \beta \to \infty \), and in that limit the \( U \to -\frac{\hbar \omega}{2} \), which is what one would expect since the system has frozen out into the lower energy level.

Now, as we did in the case of the quantum harmonic oscillator, we can write down an expression for the mean occupation number \( < n > \) of the excited state, that is, \( U = \hbar \omega < n > - \frac{\hbar \omega}{2} \) (so that \( n = 0 \) would be the ground state and \( n = 1 \) would be the excited state...there are only two possible states). Plugging this into Eq.(31.6) we find,

\[
<n> = \frac{1}{e^{\beta \hbar \omega} + 1} \tag{31.7}
\]

This is called the Fermi-Dirac distribution function. Compare Eq.(31.7) with that of Eq.(31.4). The interpretation of the distributions is as follows.

In the Bose-Einstein case (Eq.(31.4)) there can be an arbitrarily large average occupation number; as the temperature increases, the boson particles just “pile on”. The Fermi-Dirac distribution is quite different in that as one increases the temperature the \( < n > \to 1/2 \) meaning that the particle spends 1/2 its time in each of the ground state or the excited state. As the temperature decreases the occupation number of a state with positive \( \omega \) (the excited state) goes to 0 (and a state with negative \( \omega \) would become completely occupied, \( < n > \to 1 \)). In any limit the average occupation number is limited to being between 0
and 1. These excitations, called fermi particles, resist multiply occupying a single state. Examples of Fermi particles are electrons, protons, neutrons, quarks, muons, neutrinos, and at low temperatures atoms that have a total ODD number of fermions, such as $^{22}Na$, $^{86}Rb$, $^3He$ and $^6Li$.

Although the distribution doesn’t specifically exclude the possibility that multiple fermions can occupy the same state (just in such a way that the time average or phase average of the occupation is always less than one) we will learn that in individual fermionic systems not connected to a heat bath that indeed no two identical fermions can ever occupy the same quantum state at the same time.
Lecture 31.5

The Transmission Line

Deeply understanding a linear transmission line is a rich source of relevant electrical and physics problems, including the particle in a box. We can think of a transmission line as a series of lumped electrical elements (see Fig. 1), the basic unit of which is the linear four terminal network (see figure 2), so we first review that. Let \((I_{in}, V_{in})\) be the voltage across and the current into the left hand side, and \((I_{out}, V_{out})\) be that on the right hand side. The vector notation is, of course, deliberate in that it is useful when modelling linear system like a passive, linear transmission line. Let \(Z_1\) and \(Z_2\) be the complex impedances associated with the as shown. For any, even complicated, lumped linear circuit, the overall transfer function (the function that relates the input fields \((I_{in}, V_{in})\) to the output fields \((I_{out}, V_{out})\) ) can be reduced to a \(Z_1\) and \(Z_2\).

For linear response, it is very useful to study the transfer function of the 4-terminal network at a single frequency (since any signal in the time domain can be represented as a sum over different frequencies...the Fourier decomposition), \(\omega\). Thus we take

\[
(I_{in}, V_{in})(\omega(t)) = (I_{in}(\omega), V_{in}(\omega))e^{-i\omega t} + c.c.
\]

via current conservation and Ohm’s law, respectively as

\[
I_{out} = I_{in} - V_{out}/Z_2 \quad V_{out} = V_{in} - Z_1 I_{in}
\]

Which we can more succinctly fashion as a matrix relation between \(\vec{v}_{in}\) and \(\vec{v}_{out}\),

\[
\begin{align*}
\vec{v}_{in} &= (I_{in}(\omega), V_{in}(\omega))^t \\
\vec{v}_{out} &= M\vec{v}_{in}
\end{align*}
\]

\[
M = \begin{bmatrix}
1 + Z_1/Z_2 & -1/Z_2 \\
-Z_1 & 1
\end{bmatrix}
\]

Here are a few important things to note about the properties of \(M\) and how they relate to experience. Note that the determinant \(det(M) = 1\), even if there is dissipation in the network. This is actually tied to a deep physical principle, called the canonical commutation relations, that is, \(V\) and \(I\) are naturally canonically conjugate as dynamical variables, just as \(x\) and \(p_x\) are for a mechanical system. There are in principle many ways of writing the energy of a system in terms of ‘position’ and ‘momentum’ variables, but only distinct pairs of choices leads to a way of writing the dynamics (i.e. ‘Newton’s equations’...that is, how the variables depend on time) as the gradients of that energy.

For example, for a 1-d harmonic oscillator the most natural thing to write for the energy is

\[
H = \frac{p_x^2}{2m} + \frac{k}{2}x^2,
\]

as this leads to the equations of motion, \(\dot{p} = -\frac{\partial H}{\partial x}\) and \(\dot{x} = \frac{\partial H}{\partial p_x}\). For an LC circuit the analogous thing would be \(H = \frac{C}{2}V^2 + \frac{L}{2}I^2\), and the time development can be written in the same way as partial derivatives of the energy, in these simple \(V\) (voltage across the elements) and \(I\) (current coursing through them) co-ordinates. Turns out \(det(M) = 1\) preserves that structure so that if \((I_{in}, V_{in})\) are canonically conjugate then so is \((I_{out}, V_{out})\). This canonically pairing persists into a quantum mechanical description.
of the electrical system...with important consequences for the electrical behavior of squids (superconducting quantum interferences devices).

Now imagine these 4 terminal networks as very small spatially, so that we can use Eq.(31.5.2) to write down a differential version of the relation between $\vec{v}_{in}$ and $\vec{v}_{out}$, as

$$\frac{d\vec{v}}{dx} = \begin{bmatrix} 0 & -1/z_2 \\ -z_1 & 0 \end{bmatrix} \vec{v}$$ (31.5.3)

where $z_1$ and $1/z_2$ are the $Z_1$ and $Z_2$ as measured per unit length along the transmission line.

It is useful to think of a concrete example of the $Z$ and the $z$. So, for a simple co-axial or pair-wire transmission line, the $Z_1$ is mostly wire resistance and inductance, so $z_1$ are those again, but now per unit length. The $Z_2$ is typically due to the capacitance of the cable, so that $z_2$ is the capacitance per unit length. In the familiar notation of Steinmetz, we can for this useful example then reduce $z_1 = i\omega L$ and $z_2 = 1/(i\omega C)$ where $L$ and $C$ are the inductance and capacitance per unit length along the cable.

In the differential form of Eq.(31.5.3) we can solve the eigenvalue equation $\frac{d\vec{v}}{dx} = Mvecv = ik\vec{v}$ to find the eigenvalues satisfy the characteristic equation $k^2 - z_1/z_2 = 0$, so that $k = 2\pi/\lambda = \pm \omega \sqrt{LC}$ in our example. This means that the signal $\vec{v}$ is a wave travelling at (phase-)velocity $\omega/k = 1/\sqrt{LC}$.

Note finally, that for a transmission line in which the $z$’s are constant, one can use these eigenvalues to also write down the eigenvectors associated with the motion of the signal $k, \omega$ wave. That is, we find the general solution of Eq.(31.5.3) is $\vec{v}(x, t) = A\vec{v}_+e^{(kx - \omega t)} + B\vec{v}_-e^{-i(kx + \omega t)}$ where $A$ and $B$ are the amplitudes of the rightmoving and leftmoving waves respectively. One can check that the eigenvectors are $\vec{v}_\pm = \begin{pmatrix} -ik \\ \pm z_1 \end{pmatrix}$.

Note to quantify the actual power flowing in the transmission line, one can use

$$P(x, t) = \vec{v}^\dagger \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \vec{v}$$ (31.5.4)

and this can be used to , for example, normalize the propagation eigenvectors $\vec{v}_\pm$ so that the power flow can be written as a simple quadratic form in the amplitudes $A$ and $B$ (and their complex conjugates).
Lecture 32

Particle in a box review

For an ideal gas, the archetype was that of a chamber filled with $N$ non-interacting, indistinguishable particles moving through the chamber bouncing elastically off the wall. Quantum mechanically, the above picture is only useful for a weakly non-interacting gas at sufficiently high temperature and low density (physically the idea of this limit is that the interactions are rare and only associated with energies far below the energy scale $k_B T$). Even for a strictly non-interacting gas at low temperatures and/or at higher density, the quantum mechanical description of a gas starts in terms of states as “waves that fit the box” and their associated occupation numbers. That said, much of the discussion in this warm up is to remind you about what you learned about the particle in the box during your modern physics class.

Consider a cubical box, each side of length $L$. Consider further that deBroglie tells us to consider each particle with momentum $p$ as being associated with a wave of wavelength $\lambda = \frac{2\pi}{\bar{\hbar}/p}$. Define the wavenumber $k = \frac{2\pi}{\lambda} = \frac{p}{\bar{\hbar}}$, then an associated wavefunction can be written as a sum over a set of travelling waves,

$$\Psi = \sum_{\vec{k}} e^{i(\vec{k} \cdot \vec{x} - \omega t)} \tag{32.1}$$

where $\vec{k} = \vec{v}c\bar{\hbar}$ and $E = \hbar\omega$ is the energy of the state associated to the wave. This energy is itself a function of $\vec{k}$. For example, for a massive particle moving slower than the speed of light, it is the usual $E = \vec{p} \cdot \vec{p}/2m$. For light (or something moving at the speed of light) it would be $E = c|\vec{p}|$. Thus the state in either case is a sum of travelling waves of the same energy -meaning- the sum runs over all $\vec{k}$ that have a fixed length.

**Boundary Conditions** Now, in order for the waves to 'fit' in the box, the waves making up the state must constructively interfere in the interior of the box and satisfy physical boundary conditions at its edges. There are only a few choices to make here. **Choice #1**: One could -as we will now- require the wave to satisfy periodic boundary conditions. Using periodic boundary conditions means that we require each solution Eq.(32.1) to satisfy $\Psi(\vec{x}) = \Psi(\vec{x} + \vec{l})$ where $\vec{l}$ is any lattice vector in $\vec{l} = (n_1, n_2, n_3)L$ where $n_i$ are any integer values. **Choice #2** is called Neumann boundary conditions, and it is simply to require the wave function Eq.(32.1) to vanish on all sides ('faces') of the box. Finally, **Choice #3** is the oft-employed Dirichlet boundary conditions in which we require that the gradient perpendicular to each 'face' of the box is zero.

Today's take away lesson is that although these three choices are microphysically quite different, they all lead to the same thermodynamic limit for the bulk properties so it doesn't matter which one we choose! They may indeed lead to different accounting of the surface
energy associated with the ensemble of the waves, but the associated surface energy—while interesting in its own right—will clearly not scale with the volume of the system and thus not contribute to the bulk thermodynamic limit. Computing the connection between surface energies and boundary conditions is interesting and left to the earnest reader, but all should take away the lesson that some microphysics (in this case, details of the boundary conditions) details are lost in the thermodynamic (volume-)“scaling” limit.

Taking now periodic boundary conditions, we learn that \( \vec{k} \cdot \vec{l} = 2\pi n \) for all \( \vec{l} \) in the lattice and \( n \) any integer. Effectively this means that the \( \vec{k} \) must lie on the “dual lattice” and that the momenta take only discrete values (i.e. are “quantized”) to

\[
\vec{p} = \frac{h}{L} (n_1, n_2, n_3)
\]

(32.2)

where \( h = 2\pi \hbar \) is so small that for a \( L = 1 \text{cm} \) box, the \( \Delta P \) between two adjacent solutions of the form Eq.(32.2) is about \( 10^{-32} \text{ m-Kg/s} \), an incredibly small value for a momentum difference, so small that even were it for hydrogen atoms it would mean a difference of velocity of only about \( 10^{-5} \text{ m/s} \).

Since it is so small for a typical box size, we will wantonly turn sums over states into integrals etc without any regard for the little details of the edges of the integration not conforming to the lattice of allowed states of Eq.(32.2). Volume scaling (i.e. extensivity) will only depend on volume integrations/sums in the lattice of Eq.(32.2).
Einstein $A$ and $B$ coefficients

A particularly simple and powerful way of understanding the *connection between statistic and equilibrium distribution function* was introduced in the early years of quantum mechanics by none other than Albert Einstein, even before the new ’quantum mechanics” was really fully formulated.

Einstein thought of atoms as mechanical systems (they are, after a fashion) that, when placed in a oven (heated box) come to equilibrium with the walls of the box through radiative transitions. Let's for simplicity consider the case of an atom with a single excitation. Each atom is individually either in a ground state '0' or an excited state '1'. Let $N$ be the total number of atoms, $N = N_0 + N_1$ where $N_0$ and $N_1$ are the numbers of atoms in either the ground state or the excited state respectively. Denote the energy difference between the ground state and the excited state by $E$.

Now, we need to talk about the oven. It can be thought of as a perfect cube, with perfectly (well, almost) absorbing walls and being heated to temperature $T$, the walls fill the oven with photons at all sorts of different energies and moving in all possible different directions. Of course, for the photons to exist in the cavity, they have to 'fit' and so there are really discrete modes that each photon can occupy. Let $n_{\text{bos}}$ be the number of photons that are in a particular mode (think of mode as a combination of both the direction of flight and the polarization state of the photon...there are two distinct polarization states, all others being linear combinations of the two distinct ones.).

Now, motivated by the photoelectric effect, Einstein thought of an atom absorbing a photon from a particular mode as starting in the ground state and then ending up in the excited state. Denote the rate for this process to occur in the presence of a single photon as $B$.

Further denote by $A$ the rate for a single excited atom in isolation to emit a photon back into that particular mode of the oven ('cavity’). This $A$ is simply related to what we call the decay rate of the excited state, also called the spontaneous emission rate.

Einstein’s key point was to realize that like Newtonian mechanics, quantum mechanics should always be reversible, that is, more precisely, time-reversal symmetric. What that means is that any process that runs forward in time can also be viewed as running backwards in time and has the same rate either way. Thus, if the total rate for absorption scales with the number of photons as $n_{\text{bos}}B$ (as we might expect since each photon is acting independently, and the atom could grab any one...as they are indistuishable from one another), there should be a corresponding stimulated emission process that is the time reverse of this, that is the rate for the emission from an excited state atom of an additional photon in the presence of the $n_{\text{bos}} - 1$ photons should be $(n_{\text{bos}} - 1)B$. The initial state of the first process is the final state of the second process, and they are occuring at the same rate $B$, implying time reversal symmetry.

We are now ready to talk about the consequence of the atoms reaching equilibrium in the oven. In equilibrium, the total current of atoms getting excited is $N_0 B n_{\text{bos}}$ (note how similar this to the expected law of mass action in chemistry) whereas the total current of atoms returning to the ground state is $N_1 (A + B n_{\text{bos}})$, where the $N_1 A$ term is included here since the spontaneous emission process occurs in parallel. Clearly it needs to be included in the equation here, else the decay would shut down as $n_{\text{bos}} \rightarrow 0$, which does not agree with our experience of spontaneous decay of an isolated atom.
After the atoms and the oven come into thermal equilibrium the current of atoms decaying back into the ground state and the current of atoms being excited must be equal, that is

\[ N_0 B n_{\text{bos}} = N_1 (A + B n_{\text{bos}}) \]  

(33.1)

Einstein then surmised that the quantum mechanics must lead to \( A = B \) since the \( n_{\text{bos}} = 1 \) process with \( B \) (decay) is the time reverse of the spontaneous decay process given in terms of \( A \). Finally, Einstein expected that the numbers of atoms in each of the states must be proportional to finding any single atom in either state (this is the ergodic hypothesis, that the phase averages are the same as the temporal average of a single subsystem), and that in equilibrium we expect the probability of a single atom to be in a particular state is simply proportional to the Boltzmann factor \( e^{-\beta E} \), where \( \beta = \frac{1}{k_B T} \). Thus,

\[ \frac{N_1}{N_0} = e^{-\beta E} \]  

(33.2)

so that we can solve the foregoing with \( A = B \) to find that \( n_{\text{bos}} \) must satisfy

\[ n_{\text{bos}} = \frac{1}{e^{\beta E} - 1} \]  

(33.3)

and this is the Bose-Einstein distribution function for bosonic particles. Note that this is the number of photons in a particular mode in the cavity. The low temperature limit \( \beta \to \infty \) then corresponds to the usual Boltzmann statistics \( n_{\text{bos}} \sim e^{-\beta E} \), whereas the number of photons in the high temperature limit \( \beta \to 0 \) diverges. We’ll have more to say about this so-called “infrared red” divergence in later chapters.

### Fermi-Dirac Statistics

The framework of the argument above is so simple and clear that it can be extended in several ways; that flexibility and simplicity is the hallmark of a great theoretical argument. First let us redo this argument for a cavity filled with fermions instead of bosons (like light). The ‘atom’ in this case is something that can absorb and emit these fermions into that particular state (like a trapping site in a crystal, for example). Fermions satisfy the Pauli exclusion principle, that no two can be in exactly the same mode (i.e. same cavity-fitting wave and spin).

This means that the amplitude for a stimulated emission process must be identically zero. Thus, the current from the ground state to the excited state is \( N_0 A n_{\text{ferm}} \), whereas the current from the excited state back down to the ground state is \( N_1 A (1 - n_{\text{ferm}}) \), where we have used microphysical reversibility in assigning an intrinsic rate \( A \) to each of these processes and the Pauli exclusion principle is manifest in the second process, as the current shuts off if \( n_{\text{ferm}} = 1 \) (if the state is already occupied then, by Pauli exclusion, the rate of decay into that state must vanish). Thus, in equilibrium, since there the current into the excited state equals the current into the initial state, we have

\[ N_0 A n_{\text{ferm}} = N_1 A (1 - n_{\text{ferm}}) \]  

(33.4)

and, again using the expected Boltzmann equilibrium ratio, Eq.(33.2), we find

\[ n_{\text{ferm}} = \frac{1}{e^{\beta E} + 1} \]  

(33.5)
Note that, once again, the low temperature limit $\beta \to \infty$ goes over smoothly into the Boltzmann ratio for positive energies (goes to ‘1’, or full occupation for negative energies $E < 0$) and that the high temperature limit $\beta \to 0$.
The distribution functions $n$ for the bosons and fermions are sometimes denoted $n_b$ and $n_b$ respectively, sometimes even $n_+$ and $n_-$ respectively.
There are many many other ways to extend the arguments presented here to other situations.
Lecture 34

Particle in a box: Black body radiation

Now that we know that the momenta of particles (waves) in a box must be discrete (quantized) values, we are able to discuss the thermodynamics of a gas of these particles.

As an introduction to this, consider the simplest case: a gas of photons in a cubical \((l, l, l)\) cavity. The walls of the cavity are assumed to be coupled to the photons in the box (that is, they absorb and emit them) at temperature \(T\). Photons, being massless, are always relativistic, and satisfy

\[ E = c|p| \quad (34.1) \]

But recall that the momenta are (quantized) discrete, \(p = \frac{\hbar}{l}(n_1, n_2, n_3)\) with \(n_i \in \mathbb{Z}\). A particular standing wave mode of electromagnetic field in the cavity is given by the choice of integers \(n_1, n_2, n_3\) (which correspond to the number of ”humps” of the norm of the electric field along the \(\hat{x}, \hat{y}, \hat{z}\) directions) and by a choice (+ or -) of the polarization of the wave. Since the walls can emit and absorb photons, each mode of the cavity gets ”populated” by a number of photons on the average equal to the occupation number for a harmonic oscillator of that energy. That is, a mode of energy \(E\), is filled with \(n(E)\)

\[ n(E) = \frac{1}{e^{E/k_BT} - 1} \quad (34.2) \]

photons. (Note that for photons the chemical potential \(\mu\) must be zero! Their number is not conserved...).

Thus, the average total energy of the gas is easy to compute

\[ < U > = \sum_{\text{modes}} E n(E) = 2 \sum_{n_1,n_2,n_3} \frac{E(\vec{n})}{e^{E/k_BT} - 1} \quad (34.3) \]

where the factor of two is from the polarizations. The \(n_i\) can be positive or negative, and, indeed summing over the nodes is simply summing over a three-dimensional lattice (see figure 1). Now,

\[ E(\vec{n}) = \frac{ch}{l} \sqrt{n_1^2 + n_2^2 + n_3^2} \quad (18.4) \]

and so let \(R = \sqrt{n_1^2 + n_2^2 + n_3^2}\), the distance from the origin is \(n\) space to the point \((n_1, n_2, n_3)\) (Note all \(n_i \in \mathbb{Z}\) the energy \(E(\vec{n})\) is really only a function of \(R\). This means that all the states on a ball of radius \(R\) in \(n\)-space have the same energy, \(E = \frac{cR}{l}\). But how many states are there at fixed \(R\)? Well, it is rather straightforward to convince oneself that there are approximately \(4\pi R^2\) states in \(n\) space ”on the ball” of radius \(R\). (This is just the surface area of that ball.) Finally then, since photons are bosons, using Eq.(34.3) and rewriting the sum as an integral Eq.(34.3) may be written as

\[ < U > = 2 \int d^3n \frac{ch \sqrt{\vec{n}^2}}{e^{\frac{cR}{l}k_BT} - 1} \quad (34.5) \]
but, as described, let \( R = \sqrt{n^2} \) thus \( \int d^3n g(R) = \int R^2 \sin \theta d\theta d\phi g(R) = 4\pi \int_0^{+\infty} R^2 dR g(R) \) or for Eq(34.5) We thus have,

\[
\langle U \rangle = \frac{8\pi c h}{l} \int_0^{+\infty} \frac{R^3 dR}{e^{\frac{nk_BT}{ch}}-1} = \frac{8\pi c h}{l} \left( \frac{nk_BT}{ch} \right)^4 \int_0^{+\infty} \frac{x^3 dx}{e^x - 1}
\]

\[
\langle U \rangle = \frac{8\pi V (k_BT)^4}{15c^3h^3}
\]

Note that it is proportional to the fourth power of the temperature (and is proportional to the volume).

### Emissivity and the Stefan-Boltzmann Law

If we now look at the wall of our cavity filled with (thermal) photons, we may ask what is the radiant energy flux off the wall, per unit area. Obviously, we need only consider the flux of energy carried by the half of the photons leaving the wall (see figure 2)

\[
\text{Power} = P = \frac{cU}{2} \cdot \text{flux}
\]

(34.7)

here \( P \) represents the power loss and "flux" is a factor that normalizes the energy transport through a surface perpendicular to the wall. For photons emerging at an angle \( \theta \), their energy flux factor is \( \cos \theta \), since \( c \cos \theta \) represents how fast energy is being transported perpendicular to the wall by photons. Thus since all directions are equally likely, the total flux factor is

\[
\text{flux} = \frac{\int \cos \theta d\Omega}{\int d\Omega}
\]

(34.8)

where \( d\Omega \) represents the solid angle (area of unit sphere) \( d\Omega = \sin \theta d\theta d\phi \) and all the integrals are from \( \theta = 0 \ldots \pi/2 \) and \( \phi = 0 \ldots 2\pi \) (i.e. just over the half sphere of radiation propagating to the right.) Thus

\[
\text{flux} = \frac{2\pi}{\pi} \int_0^{\pi/2} \cos \theta \sin \theta d\theta = \frac{1}{2}
\]

(34.9)

and so

\[
\text{Power} = P = \frac{cU}{4} = \frac{\pi^2 k_B^4}{60c^2h^3} T^4 = \sigma T^4
\]

(34.10)

where \( \sigma = \frac{\pi^2 k_B^4}{60c^2h^3} = 5.6x10^{-8} \) is called Stefan-Boltzman constant.

Finally, what is the pressure due to a gas of photons hitting and reflecting off the walls of the chamber? ....
Lecture 36

The Grand Canonical Ensemble

We now discuss equilibrium in which both energy and particle number is exchanged. In general we know that the multiplicity function depends on both the total energy and total number of particles that make up the system. Imagine connecting the system up to a large bath with which it can exchange particles and heat. The configuration accessible to the system under such conditions are as a set called the grand canonical ensemble. Let us require that both the total energy and particle number of the bath plus the system is fixed

\[
N = n_s + n_b \quad U = u_s + u_b
\]  

(36.1)

Now, we know that the total multiplicity function \(g(U, N)\) of the combined system is

\[
g(U, N) = \sum_{u_s} \sum_{n_s} g_s(u_s, n_s)g_b(U - u_s, N - n_s) \sim g_s(u_s^*, n_s^*)g_b(U - u_s^*, N - n_b^*)
\]

(36.2)

where again we have approximated each of the sums by its single largest term and \(u_s^*\) and \(n_s^*\) are the energies and numbers of particles in the system at this most likely combination. Now, if we are to faithfully represent the sum in Eq.(36.2) by a single largest term, then we must have that

\[
\frac{\partial (g_sg_b)}{\partial u_s}|_{n_s} = 0 = \frac{\partial (g_sg_b)}{\partial n_s}|_{u_s}
\]

(36.3)

and from this and Eq.(36.1) we have that (following as we did the discussion of the canonical ensemble)

\[
\frac{\partial g_s}{g_s \partial u_s}|_{n_s} = -\frac{\partial g_b}{g_b \partial u_s}|_{n_s} = \frac{\partial g_b}{g_b \partial n_b}|_{u_s}
\]

(36.4)

and

\[
\frac{\partial g_s}{g_s \partial n_s}|_{u_s} = -\frac{\partial g_b}{g_b \partial n_s}|_{u_s} = \frac{\partial g_b}{g_b \partial n_b}|_{u_b}
\]

(36.5)

Now, the first of these relations we have already identified... namely, the derivatives in Eq.(36.4) represent a universal function of the temperature. We were able to show that universal function to be the reciprocal of the temperature in energy units (as must be the case on dimensional grounds). which is common to the two systems as they finally reach equilibrium.

The second relation is a like the first in many ways. The only function we know that is common to two systems that exchange particle number is the chemical potential \(\mu\). Testing Eq.(36.5) on systems that we know (such as the ideal gas) we know that this universal function is simply the negative of the ratio of the chemical potential to the temperature!

\[
\frac{\partial g_s}{g_s \partial n_s}|_{u_s} = -\frac{\mu}{k_b T} = \frac{\partial g_b}{g_b \partial n_b}|_{u_b}
\]

(36.6)

So, finally, using Eq.(33.4) and Eq.(33.6) we have that

\[
g(u, n) \sim e^{\frac{u - \mu n}{k_b T}}
\]

(36.7)
Since the probability of the system being in a particular microstate is assumed to be the reciprocal of the multiplicity function (recall the fundamental assumption of statistical mechanics!) we have that the ratio of the probabilities of two states in this grand canonical ensemble is

\[
\frac{P(u_1, n_1)}{P(u_2, n_2)} = e^{\frac{-\Delta u + \mu \Delta n}{k_B T}}
\]  

(36.8)

where \(\Delta u = u_1 - u_2\) and \(\Delta n = n_1 - n_2\), and so we may again construct a grand sum over the probabilities of all the states accessible to the system in this new ensemble. The new sum is called the Gibbs sum

\[
\zeta = \sum_{\mathcal{G}} e^{\frac{u + \mu n}{k_B T}}
\]  

(36.9)

where \(\mathcal{G}\) represents the collection of states making up the grand canonical ensemble, \(i.e.\) states in which the number of particles and the total energy of the state is variable.
Problem Set 10

More on The Canonical Ensemble

1) Compute the internal energy $U$ and pressure $P$ of a non-interacting gas of $N$ distinguishable ultra-relativistic particles. The energy of an individual particle is $E = c|\vec{P}|$.

2) Consider a gas where the molecules have a fixed electric dipole moment $|\vec{d}|$. They are indistinguishable and do not interact with one another BUT they do interact with an externally applied electric field $\vec{E}$ with interaction energy $-\vec{d} \cdot \vec{E}$. Compute the partition function for this gas. Compute $U$ as a function of the applied electric field and temperature. Also, compute the bulk electric polarizability, $<\vec{D}> = \frac{1}{N} <\sum_i \vec{d}_i>$ for the gas.

3) The energy levels of a diatomic molecule have energy given in terms of the total angular momentum $J$ as $E = J(J+1)\tau_0$ where $J$ can take on any integer value and $\tau_0$ is a constant. For each value of $J$ there are $2J+1$ distinguishable microstates.
a) Write down the partition function for a single such molecule at rest.
b) Now let them move as well as point. Find an approximate form for the specific heat at low temperature of a gas of $N$ of these indistinguishable molecules.
c) Finally approximate the sums as full integral over $J$ to find the approximate form of the specific heat at high temperature.

4) a) Compute the density of states of a two-dimensional ultra-relativisitic gas.
b) Use your result to compute the energy density and equation of state of a two-dimensional gas of photons. Assume that there is only one photon polarization possible in 2-dimensions.

5) a) Use the Virial Theorem to compute at high-temperatures the total energy of a system of distinguishable particles all interacting with each other through the power law relation $V(\vec{x}_1 - \vec{x}_2) = \Lambda |\vec{x}_1 - \vec{x}_2|^n$. (Here $\Lambda$ is some constant).
b) Critique what this implies about the average interparticle spacing as a function of temperature in the high temperature limit. Describe what happens in the two cases of $n < 0$ and $n > 0$. 