

Statistical Physics

33-765

Nathaniel D. Hoffman

Spring Semester, 2020

Contents

LECTURE 1: OVERVIEW OF PROBABILITY

Monday, January 13, 2020

0.1 Probability in a Nutshell

We will begin with an object called a random variable (“RV”) A . It can take many different values each time the experiment is run, and the value given follows some distribution: $A \in \{a_1, \dots, a_n\} \mapsto \Pr_A(a)$, the probability that the RV A takes the value a .

Additionally, we want the probabilities to be normalized: $\sum_a \Pr_A(a) = 1$.

Now suppose we have two random variables, A and B .

Definition 0.1.1 (Joint Probability). $\Pr_{A,B}(a,b)$ is the probability that A takes the value a and B takes the value b .

Definition 0.1.2 (Marginal Probability). $\Pr_A(a) = \sum_b \Pr_{A,B}(a,b)$

Definition 0.1.3 (Conditional Probability). $\Pr(a | b)$ is the probability of a given b , since it is conceivable that the random variable B is informative about the outcome of A or vice versa.

Example. Suppose we roll a die, and $A \in \{1, 2, 3, 4, 5, 6\}$ while $B \in \{\text{even, odd}\}$.

$$\Pr(\text{roll a } 6) = \frac{1}{6}$$

so

$$\Pr(\text{roll a } 6 | \text{even}) = \frac{1}{3}$$

and

$$\Pr(\text{even} | \text{roll a } 6) = 1$$

◇

It's intuitive that

$$\Pr_{A,B}(a,b) = \Pr(a | b) \Pr_B(b) = \Pr(b | a) \Pr_A(a)$$

Therefore

$$\Pr(a | b) = \frac{\Pr(b | a) \Pr_A(a)}{\Pr_B(b)} \quad (\text{Bayes' Theorem})$$

0.1.1 Statistical Independence

We call a pair of random variables statistically independent if the joint probabilities factor into the marginals:

$$\Pr_{A,B}(a,b) = \Pr_A(a)\Pr_B(b) \iff A \text{ and } B \text{ are statistically independent}$$

In some ways, statistical independence is a good thing, since it makes calculations technically easier. Unfortunately, it's typically boring, since this means A is not informative about B and vice versa, so knowing things about one variable will tell you nothing about the other.

If two RVs are statistically independent,

$$\Pr(a | b) = \frac{\Pr_{A,B}(a,b)}{\Pr_B(b)} = \frac{\Pr_A(a)\Pr_B(b)}{\Pr_B(b)} = \Pr_A(a)$$

Now suppose we have more than two RVs, $\{A_1, A_2, \dots, A_N\}$.

Definition 0.1.4 (Pairwise Independence).

$$\Pr_{A_i, A_j}(a_i, a_j) = \Pr_{A_i}(a_i)\Pr_{A_j}(a_j) \quad i \neq j$$

Definition 0.1.5 (Mutual Independence).

$$\Pr_{A_1, A_2, \dots, A_N}(a_1, a_2, \dots, a_N) = \Pr_{A_1}(a_1)\Pr_{A_2}(a_2)\cdots\Pr_{A_N}(a_N)$$

Note

Mutual independence obviously implies pairwise independence. However, pairwise independence **does not** imply mutual independence!

0.2 Functions of Random Variables

Suppose we have a random variable A with probability $\Pr_A(a)$. We can use a function to generate a new random variable, $F(A)$. What is $\Pr_F(f)$, the probability that our new random variable takes a value f ?

$$\Pr_F(f) = \sum_a \delta_{f,F(a)} \Pr_A(a) \quad (\text{Transformation Theorem for Probabilities})$$

To understand this, notice that, for example, if $A \in \{-3, -2, -1, 0, 1, 2, 3\}$ and $F(x) = x^2$, there are now two ways to get $f = 4$ (namely, $a = -2$ or $a = 2$). Therefore, we have to look at both of these a values (the δ -function) and of course we must also consider the probability to get each a .

0.3 Expectation Values

Definition 0.3.1 (Expectation Value).

$$\langle A \rangle_{\Pr_A} = \sum_a \Pr_A(a)$$

is called the expectation value of A , and is simply a weighted average over all the possible values A can take.

$\langle \cdots \rangle$ acts like a linear operator:

$$\langle \alpha A + \beta B \rangle = \alpha \langle A \rangle + \beta \langle B \rangle$$

LECTURE 2:
Monday, January 13, 2020

0.4 Expectation Values of Functions of Random Variables

If we have a function $F(A)$ on a random variable,

$$\begin{aligned}\langle F \rangle &= \sum_f f \Pr_F(f) \\ &= \sum_f f \sum_a \delta_{f,F(a)} \Pr_A(a) \\ &= \sum_a \Pr_A(a) \underbrace{\sum_f f \delta_{f,F(a)}}_{F(a)} \\ &= \sum_a F(a) \Pr_A(a)\end{aligned}$$

Therefore, we can calculate the expectation value of F using either the probability distribution of the function itself or by using the probability distribution of the random variable.

0.4.1 Moments

The n -th moment is defined as

$$\langle F^n \rangle = \sum_f f^n \Pr_F(f) = \sum_a [F(a)]^n \Pr_A(a)$$

We can then define “centered” moments as

$$\langle (F - \langle F \rangle)^n \rangle = \dots$$

The most important of these is the variance, the second centered moment:

$$\sigma_F^2 = \langle (F - \langle F \rangle)^2 \rangle = \langle F^2 - 2F \langle F \rangle + \langle F \rangle^2 \rangle$$

Notice that $\langle F \rangle^2$ is just a number, so we already know its expectation value. Again, recall that expectation values are linear so we can also remove the $2 \langle F \rangle$ factor:

$$\sigma_F^2 = \langle F^2 \rangle - \langle F \rangle^2$$

From the original definition, it is obvious that the variance is strictly non-negative. If the variance vanishes, the random variable never deviates from its expectation value, so it is not random. We can define the square root of the variance as the standard deviation of F : σ_F . It's slightly nicer than the variance because it has the same units as the random variable, but often the square root makes it difficult to work with.

Definition 0.4.1 (Covariance).

$$\text{Cov}(A, B) = \langle AB \rangle - \langle A \rangle \langle B \rangle$$

For a single variable, $\sigma_A^2 = \text{Cov}(A, A)$.

Definition 0.4.2 (Correlation Coefficient).

$$c_{A,B} = \frac{\text{Cov}(A, B)}{\sigma_A \sigma_B}$$

This coefficient is dimensionless and invariant under linear transformations of the random variables:

$$A \mapsto a_0 + a_1 A \quad \text{and} \quad B \mapsto b_0 + b_1 B \implies c_{A,B} \mapsto c_{A,B}$$

Additionally,

$$-1 \leq c_{A,B} \leq +1$$

Definition 0.4.3 (Correlation). If the covariance vanishes, we say that the variables are uncorrelated.

Independence implies non-correlation, but non-correlation does not imply independence.

The proof of the forward direction is a result of the joint probabilities factoring to allow the covariance to cancel to zero. The disproof of the opposite direction can be made by clever counterexample. Correlation is a measure of linear dependence, but two random variables can be non-linearly dependent.

0.4.2 Sums of Random Variables

$$S = \sum_{j=1}^N F_j$$

$$\langle S \rangle = \sum_{j=1}^N \langle F_j \rangle$$

$$\begin{aligned} \sigma_S^2 &= \langle S^2 \rangle - \langle S \rangle^2 \\ &= \left\langle \sum_{j=1}^N F_j \sum_{k=1}^N F_k \right\rangle - \left\langle \sum_{j=1}^N F_j \right\rangle \left\langle \sum_{k=1}^N F_k \right\rangle = \sum_{j,k} (\langle F_j F_k \rangle - \langle F_j \rangle \langle F_k \rangle) = \sum_{j,k} \text{Cov}(F_j, F_k) \end{aligned}$$

If all $\{F_j\}$ are uncorrelated, then $\text{Cov}(F_j, F_k) = \delta_{jk} \sigma_{F_j}^2$. In this special case,

$$\sigma_S^2 = \sum_{j,k} \delta_{jk} \sigma_{F_j}^2 = \sum_j \sigma_{F_j}^2$$

Let's now look at the special case where all of the random variables is independent and each has an identical probability distribution. For example, if we roll a bunch of identical dice, the results are all statistically independent, but the probabilities are the same.

$$\langle F_j \rangle = \mu$$

Define

$$\sigma_{F_j}^2 = \sigma^2$$

$$\langle S \rangle = N\mu$$

so

$$\sigma_S^2 = N\sigma^2 \implies \sigma_S = \sqrt{N}\sigma$$

This also implies that

$$\frac{\sigma_S}{\langle S \rangle} = \frac{1}{\sqrt{N}} \frac{\sigma}{\mu}$$

Let's look at an even more special case, where each random variable has a binomial distribution. In this case, each F_j corresponds to N individual success/fail experiments where p is the probability of success. In other words, there are two possible values with probabilities p and $1 - p$. Call the success outcome 1 and failure 0 such that $\Pr(1) = p$ and $\Pr(0) = 1 - p$.

$$S = \sum_{j=1}^N = F_j$$

so

$$\langle S \rangle = N \langle F_j \rangle = N(1 \cdot p + 0 \cdot (1 - p)) = Np$$

$$\sigma_S^2 = N\sigma_{F_j}^2 = N \cdot (\langle F_j^2 \rangle - \langle F_j \rangle^2) = Np(1 - p)$$

so

$$\frac{\sigma_S}{\langle S \rangle} = \frac{1}{\sqrt{N}} \cdot \sqrt{\frac{1-p}{p}}$$

As N gets bigger, the sum gets less random!

We can work out the exact probability distribution:

$$\Pr(S, N; p) = \binom{N}{S} p^S (1-p)^{N-S} \quad (\text{Binomial Distribution})$$

where

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

Let's make sure this distribution is normalized. We can do this using the binomial theory:

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

so

$$\sum_{S=0}^N \Pr(S, N; p) = \sum_{S=0}^N \binom{N}{S} p^S (1-p)^{N-S} = [p + (1-p)]^n = 1$$

Definition 0.4.4 (Stirling's Approximation for Factorials).

$$\begin{aligned}\ln N! &= \ln(1 \cdot 2 \cdot 3 \cdots N) \\ &= \ln 1 + \ln 2 + \dots + \ln N \\ &= \sum_{n=1}^N \ln n \\ &\approx \int_1^N dn \ln n = N \ln N - N + 1 \approx N \ln N - N\end{aligned}$$

if N is large.

LECTURE 3: PROBABILITY DENSITIES

Wednesday, January 15, 2020

A useful consequence of Stirling's approximation is

$$\ln \binom{n}{k} = -n [\varphi \ln \varphi + (1 - \varphi) \ln 1 - \varphi] \quad \varphi = \frac{k}{n}$$

0.5 Continuous Random Variables

Instead of asking for the probability to get a particular point, we instead ask about the probability of covering regions in event space.

Example. Let $x \in \mathbb{R}$ be uniformly picked between 0 and 10. What is the probability that $x \in [7, 9]$?

$$\Pr(7 \leq x \leq 9) = \frac{2}{10}$$

$\Pr(x = 8)$ is meaningless in general

We can also define probability densities:

$$\Pr(x) = \begin{cases} \frac{1}{10} & x \in [0, 10] \\ 0 & x \notin [0, 10] \end{cases}$$

In this case,

$$\Pr(7 \leq x \leq 9) = \int_7^9 dx \Pr(x) = 0.2$$

This does not mean $\Pr(8) = \frac{1}{10}$, since we need to integrate over this density. \diamond

In general, integrals of probability densities are probabilities. For $x \in \Omega$, the probability density is always non-negative ($\Pr(x) \geq 0$). Normalization gives us $\int_{\Omega} dx \Pr(x) = 1$.

- We also require $dx \Pr(x)$ to be dimensionless, so $[\Pr(x)] = \frac{1}{[dx]}$. Note that the notation for probability density is identical to the notation for a probability, so we must be cautious when deciding which is meant.
- Since probability densities generally have units, their numerical values need not be smaller than 1.
- Probability densities can also diverge—their values can become infinite, as long as the integral is well-defined.

Example.

$$\Pr(x) = \begin{cases} \frac{1}{2\sqrt{x}} & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Notice that

$$\int_0^1 dx \Pr(x) = \sqrt{x} \Big|_0^1 = 1$$

so there is no problem with the divergence at 0. \diamond

- All of our knowledge of discrete random variables carries over with $\sum \rightarrow \int$ and $\delta_{ij} \rightarrow \delta(x - y)$.

Example. If we have a random variable $x \in \mathbb{D}$ with a probability density $\Pr_X(x)$ and a function $F(x)$, what is the probability density $\Pr_F(f)$? We can extend transformation theorem to show that

$$\Pr_F(f) = \int_{\mathbb{D}} dx \delta(f - F(x)) \Pr_X(x) \quad (\text{Transformation Theorem for Probability Densities})$$

Note that this integral does not simplify in general because the δ -function contains a function of x instead of just x . One option to simplify this is to make a variable substitution for $F(x)$, but this only works in regions where $F(x)$ is invertible. \diamond

0.6 Ideal Gasses

By “ideal” we mean that gas atoms or molecules do not interact:

$$H = \sum_{i=1}^N \frac{p_i^2}{2m}$$

It is reasonable to assume in this case that the positions are independent of the momenta (a very classical assumption).

$$\Pr(\{q_1, q_2, \dots, q_N, p_1, p_2, \dots, p_N\}) = \Pr_q(\{q_1, \dots, q_N\}) \Pr_p(\{p_1, \dots, p_N\})$$

In shorthand, we will write $\Pr(q, p) = \Pr_q(q) \Pr_p(p)$. We will treat these terms separately, and for now, we will concern ourselves with the position.

For an ideal gas, we may further assume that all positions are independent of one another.

$$\Pr_q(q) = \prod_{i=1}^N \Pr_{q_i}(q_i) = \prod_{i=1}^N \Pr_{q_1}(q_1)$$

since we also assume the particles are identically distributed. Notice that the second term requires an Avogadro’s number of probability densities, while the final term only requires one.

Let’s now imagine an isolated box divided into two compartments with a hole in the division. We have gas particles in both sides, and because of the hole, they can move from the left side to the right side and back. The left side has N_A particles, volume V_A , and energy E_A , while the right hand side has N_B , V_B , and E_B , where

$$\begin{aligned} N_A + N_B &= N \\ V_A + V_B &= V \\ E_A + E_B &= E \end{aligned}$$

For any given particle, what is the probability of a particular particle being in compartment A ?

$$\Pr = \frac{V_A}{V}$$

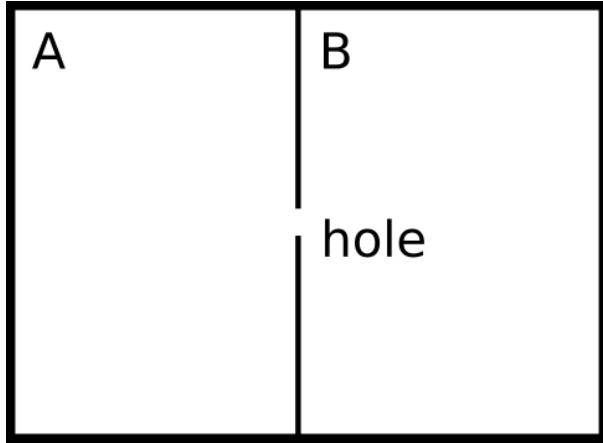


Figure 0.6.1: Box with hole

Why is this true? Physicists call this the equal a priori principle. We are assuming these particles are evenly distributed in this box. If we make this assumption, we can later derive predictions, and as it turns out, an incredible amount of nontrivial things can be predicted and experimentally verified.

LECTURE 4: IDEAL GASSES
Friday, January 17, 2020

From last lecture, we had a box of gas divided into two volumes (A and B). We decided that the probability of finding a given particle in volume A was

$$p = \frac{V_A}{V}$$

How do we know this? We should be able to somehow calculate this, since we know the Hamiltonian of the system and the equations of motion. Recall that last time we called this the “equal a priori” assumption. Proving these basic ideas from first principles is incredibly difficult, but they follow experimental results very well.

After assuming the equal a priori probability for all particles (after assuming a microstate), we can now calculate the probability of finding N_A particles in compartment A :

$$\Pr(N_A | N) = \binom{N}{N_A} p^{N_A} (1-p)^{N-N_A} = \binom{N}{N_A} \left(\frac{V_A}{V}\right)^{N_A} \left(1 - \frac{V_A}{V}\right)^{N-N_A} = \binom{N}{N_A} \left(\frac{V_A}{V}\right)^{N_A} \left(\frac{V_B}{V}\right)^{N_B}$$

We have N trials, and for each trial we have a probability p that they will make it over to A .

$$\langle N_A \rangle = N \frac{V_A}{V} \propto N$$

$$\sigma_{N_A} = \sqrt{N \frac{V_A}{V} \frac{V_B}{V}} \propto \sqrt{N}$$

since the distribution is binomial. Note that $\frac{\sigma_{N_A}}{\langle N_A \rangle} \propto \frac{1}{\sqrt{N}}$.

Note that with an Avogadro number of particles, the fluctuations are on the order of trillions. However, relatively speaking, they order on trillionths ($\frac{1}{\sqrt{N}}$). That insight led Boltzmann to suggest that the observed *macrostate* is simply the most probable one. If you were to ask how many particles are observed on the left-hand side, Boltzmann would say you just need to maximize $\Pr(N_A | N)$ (we would find, unsurprisingly, the expectation value that we just calculated).

Define an object $\Omega_q(N, V) = \frac{V^N}{N!}$. We can now rewrite our probability as

$$\Pr(N_A | N) = \frac{\Omega_q(N_A, V_A) \Omega_q(N_B, V_B)}{\Omega_q(N, V)}$$

Maximizing this probability is the same as maximizing the logarithm of the probability (since the logarithm is monotonic):

$$\ln \Pr(N_A | N) = \ln \Omega_q(N_A, V_A) + \ln \Omega_q(N_B, V_B) - \ln \Omega_q(N, V)$$

Let's redefine this logarithm:

$$S_q(N, V) = k[\ln \Omega_q(N, V) + XN]$$

k and X are just there to account for things we will encounter later, the maximizations still stay the same. Maximizing $\ln \Pr$ subject to $N_A + N_B = N$ and our other constraints is equivalent to maximizing

$$S_{q,\text{tot}}(N_A, N_B, \dots) = k \ln \Pr(N_A, N) + S_q(N, V) = S_q(N_A, V_A) + S_q(N_B, V_B)$$

This is just a rewriting of the logarithms of the Ω_q 's above. $S_q(N, V) \propto N$, and similarly with $S_q(N_A, V_A)$ and $S_q(N_B, V_B)$. How big is $k \log \Pr(N_A, N)$? The probability distribution has to be normalized—the area under our binomial distribution must be 1. We know its width is \sqrt{N} , so the height of our distribution has to be $\propto \frac{1}{\sqrt{N}}$. Therefore, the logarithm is proportional to $\ln N$, which is incredibly smaller than N . In other words, all the physics happens in the S_q 's. From our maximizing argument, we need to maximize $S_q(N_A, V_A) + S_q(N_B, V_B) \approx S_q(N, V)$.

$$S_q(N, V) = k \left[\ln \frac{V^N}{N!} + XN \right] = k [N \ln V - N \ln N + N + XN] = kN \left[\ln \frac{V}{N} + 1 + X \right]$$

LECTURE 5: Wednesday, January 22, 2020

Last lecture, we looked at the case of particles in a box partitioned into two compartments. We wrote down a probabilistic treatment using the binomial distribution, but we can push this forward to gain insight into a function which we can calculate and maximize to find the state of these particles. However we are only halfway there, since that analysis only concerned the positions of the particles and not their momentums.

Today, consider a box separated by a wall which has no hole. However, this wall is diathermal, meaning energy can pass through the wall.

We consider the energy of each compartment and of the total to be

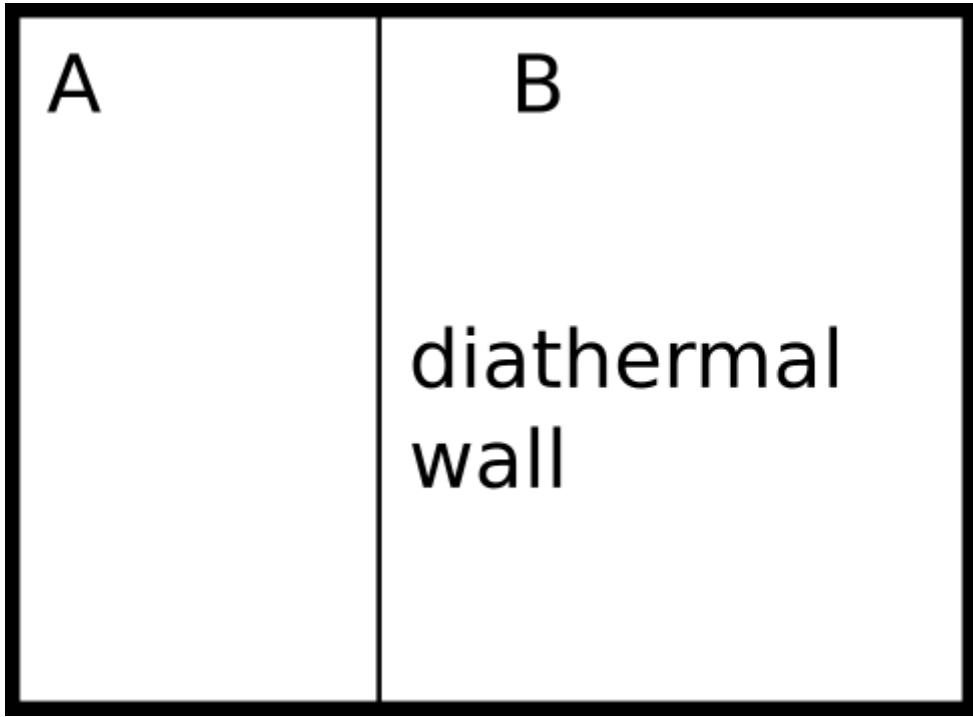
$$E_\alpha = \sum_{i \in \alpha} \frac{\vec{p}_{\alpha,i}^2}{2m}$$

There are many microstates which we can define (a single microstate is a set of momenta $\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N\}$), but only certain ones will satisfy the constraint $E = \sum_i \frac{\vec{p}_i^2}{2m}$.

These microstates live on a $3N - 1$ -dimensional hyperspace. Consider that all microstates are equally likely. This means

$$\Pr(\{\vec{p}\}) = \frac{\delta(E - \sum_i \frac{\vec{p}_i^2}{2m})}{\int d^{3N}p \delta(E - \sum_i \frac{\vec{p}_i^2}{2m})}$$

This is the original sin of stat mech: we will not derive this equation, but take it for granted. It does seem reasonable, since it picks out allowed microstates and is obviously normalized.

**Figure 0.6.2:** Box with Diathermal Wall

What is the probability of getting a value E_A ? We can use the transformation theorem:

$$\begin{aligned} \Pr_A(E_A) &= \int d^{3N}p \Pr(\{\vec{p}\}) \delta\left(E_A - \sum_{i \in A} \frac{\vec{p}_i^2}{2m}\right) \\ &= \frac{\int d^{3N}p \delta\left(E - \sum_j \frac{\vec{p}_j^2}{2m}\right) \delta\left(E_A - \sum_{i \in A} \frac{\vec{p}_i^2}{2m}\right)}{\int d^{3N}p \delta\left(E - \sum_j \frac{\vec{p}_j^2}{2m}\right)} \\ &= \frac{\int dp_A \delta\left(E_A - \sum_{i \in A} \frac{\vec{p}_i^2}{2m}\right) \int dp_B \delta\left(\overbrace{E - \sum_{j \in A} \frac{\vec{p}_j^2}{2m}}^{E_B} - \sum_{j \in B} \frac{\vec{p}_j^2}{2m}\right)}{\int dp \delta\left(E - \sum_j \frac{\vec{p}_j^2}{2m}\right)} \\ &= \frac{\int dp_A \delta\left(E_A - \sum_{i \in A} \frac{\vec{p}_i^2}{2m}\right) \int dp_B \delta\left(E_B - \sum_{i \in B} \frac{\vec{p}_i^2}{2m}\right)}{\int dp \delta\left(E - \sum_i \frac{\vec{p}_i^2}{2m}\right)} \end{aligned}$$

This final equation is Equation 6.3 in the book, and is just given as obvious (while it clearly isn't).

These integrals all have the same basic form. Let's call them functions:

$$\Pr_A(E_A) = \frac{\Omega_E(E_A, N_A) \cdot \Omega_E(E_B, N_B)}{\Omega_E(E, N)}$$

where

$$\Omega_E(E, N) = \int dp \delta\left(E - \sum_i \frac{\vec{p}_i^2}{2m}\right)$$

The structure is now precisely the same as our previous work with position, although this integral is not exactly as simple as the summation we had before. Let's expand this integral so we can understand what's going on:

$$\Omega_E(E, N) = \int d^3 p_1 d^3 p_2 \cdots d^3 p_N \delta \left(E - \frac{1}{2m} \sum_i \vec{p}_i^2 \right)$$

This is an integral over a $3N$ -dimensional space. Fortunately, the function we are integrating is not terrible, since the variable we are integrating over enters in the form

$$\sum_i \vec{p}_i^2$$

which is just the squared distance from the origin of the momentum vector. If we were in three dimensions, this would be a spherically symmetric problem, since $x^2 + y^2 + z^2 = r^2$, and we would be tempted to go into spherical coordinates and just integrate over r . In three dimensions, we take $d^3 p \rightarrow dp 4\pi r^2$. We have to extend this to N -dimensions. For now, let's write this differential as

$$d^3 p_1 d^3 p_2 \cdots d^3 p_N \rightarrow dp p^{3N-1} S_{3N}$$

where S_{3N} is the surface area of a unit sphere in \mathbb{R}^{3N} (such that $S_3 = 4\pi$).

Now our Ω_E function looks like:

$$\begin{aligned} \Omega_E(E, N) &= \int_0^\infty dp p^{3N-1} S_{3N} \delta(E - \frac{1}{2m} p^2) \\ &= \int_0^\infty dp p p^{3N-2} S_{3N} \delta \left(E - \frac{p^2}{2m} \right) \end{aligned}$$

Let's now make the substitution

$$x = \frac{p^2}{2m} \rightarrow p = \sqrt{2mx} \quad dx = \frac{p dp}{m} \rightarrow m dx = p dp$$

so

$$\Omega_E(E, N) = \int_0^\infty dx m(2mx)^{\frac{3N-2}{2}} S_{3N} \delta(E - x) = m S_{3N} (2mE)^{\frac{3N-2}{2}}$$

Notice we already have the energy dependence—the S_{3N} term is just an N -dependent scale factor. However, it would be nice to know what it actually is before we compute the final probability.

What is the surface area of a sphere in M -dimensional space?

$$\int dx e^{-x^2} = \pi^{\frac{1}{2}}$$

Now we will take this integral and raise it to the power of M :

$$\begin{aligned} \pi^{\frac{M}{2}} &= \left[\int dx e^{-x^2} \right]^M \\ &= \int dx_1 dx_2 \cdots dx_M e^{-(x_1^2 + x_2^2 + \cdots + x_M^2)} \\ \text{Use spherical coordinates: } &= \int_0^\infty dx x^{M-1} S_M e^{-x^2} \end{aligned}$$

Now use the substitution

$$t = x^2 \rightarrow x = \sqrt{t} \quad dx = \frac{1}{2\sqrt{t}} dt$$

$$\begin{aligned} \pi^{\frac{M}{2}} &= \int_0^\infty \frac{dt}{2\sqrt{t}} t^{\frac{M-1}{2}} S_M e^{-t} \\ &= \frac{1}{2} S_M \int_0^\infty dt t^{\frac{M}{2}-1} e^{-t} \\ &= \frac{1}{2} S_M \Gamma \left(\frac{M}{2} \right) \end{aligned}$$

The gamma function has the property that

$$\Gamma(N+1) = N! \quad \text{for } N \in \mathbb{N}$$

We can finally write our equation for S_M :

$$S_M = \frac{2\pi^{\frac{M}{2}}}{\Gamma(\frac{M}{2})} = \frac{2\pi^{\frac{M}{2}}}{(\frac{M}{2}-1)!} = \frac{M\pi^{\frac{M}{2}}}{(\frac{M}{2})!}$$

so we can finally write down

$$\Omega_E(E, N) = \frac{3N\pi^{\frac{3N}{2}}}{(\frac{3N}{2})!} m(2mE)^{\frac{3N-2}{2}}$$

LECTURE 6:

Friday, January 24, 2020

From last lecture, we showed

$$\Pr_A(E_A) = \frac{\Omega_E(E_A, N_A)\Omega_E(E_B, N_B)}{\Omega_N(E, N)}$$

where

$$\Omega_E(E, N) = \frac{3N\pi^{3N/2}}{(\frac{3N}{2})!} m(2mE)^{\frac{3N-2}{2}}$$

Since $E_A + E_B = E$, we can rewrite this as

$$\Pr_A(E_A) = C E_A^{\frac{3N_A-2}{2}} (E - E_A)^{\frac{3N_B-2}{2}}$$

Excursion

Take a Gaussian distribution:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

The natural logarithm of this distribution is

$$\ln p(x) = \ln \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(x-\mu)^2}{2\sigma^2}$$

Notice now that

$$\frac{\partial \ln p(x)}{\partial x} = -\frac{x-\mu}{\sigma^2}$$

and

$$\frac{\partial^2 \ln p(x)}{\partial x^2} = -\frac{1}{\sigma^2}$$

So for things that look close to Gaussian, we can find the mean and standard deviation by taking derivatives of the natural logarithm of the distribution.

Our probability distribution is very close to a Gaussian thanks to the Central Limit Theorem, which we demonstrated on the homework.

$$\ln \Pr_A(E_A) = C + \frac{3N_A-2}{2} \ln(E_A) + \frac{3N_B-2}{2} \ln(E - E_A)$$

and

$$\frac{d \ln \Pr_A(E_A)}{d E_A} = \frac{3N_A-2}{2} \frac{1}{E_A} - \frac{3N_B-2}{2} \frac{1}{E - E_A} = 0$$

implies

$$E_{A,\max} = E \frac{3N_A - 2}{3N - 4} \approx E \frac{N_A}{N}$$

or

$$\frac{E_{A,\max}}{N_A} = \frac{E}{N} = \frac{E_{B,\max}}{N_B}$$

Taking the second derivative

$$\frac{\partial^2 \ln \Pr_A(E_A)}{\partial E_A^2} = \frac{3N_A - 2}{2} \frac{-1}{E_A^2} - \frac{3N_B - 2}{2} \frac{1}{(E - E_A)^2}$$

so

$$\sigma_{E_A}^2 = \frac{2}{3} \frac{E_{A,\max}^2}{N} \frac{N_B}{N_A} = \frac{2}{3} \left(\frac{E}{N}\right)^2 \frac{N_B}{N_A} N$$

so $\sigma_{E_A} \propto \sqrt{N}$.

Let's now define

$$S_{E,\alpha} = k \ln(\Omega_E(E_\alpha, N_\alpha))$$

such that

$$S_{E,\text{total}}(E_A, N_A) = k \underbrace{\ln \Pr_A(E_A)}_{\sim \ln N} + \underbrace{S_E(E, N)}_{\sim E} = \underbrace{S_E(E_A, N_A)}_{\sim E_A} + \underbrace{S_E(E_B, N_B)}_{\sim E_B}$$

since the distribution $\Pr_A(E_A)$ has a standard deviation $\sim \sqrt{N}$, the height of the distribution must be $\sim \frac{1}{\sqrt{N}}$. So we can say that

$$S_E(E, N) = kN \left[\frac{3}{2} \ln \frac{E}{N} + \text{const} \right]$$

by the Stirling approximation.

Recall that, for ideal gasses,

$$\Pr(\{q, p\}) = \Pr_q(\{q\}) \Pr_p(\{p\})$$

we just showed that to maximize this, we should maximize $\ln \Pr$:

$$k \ln \Pr(E_A, V_A, N_A) = k \ln \Pr(N_A, V_A) + k \ln \Pr(E_A, N_A) = S_q(N_A, V_A) + S_E(E_A, N_A)$$

so

$$S(E, V, N) = k \left[\ln \frac{V^N}{N!} + \ln \frac{E^{\frac{3N}{2}}}{(\frac{3N}{2})!} + x \right] \xrightarrow{\text{Stirling}} kN \left[\ln \frac{V}{N} + \frac{3}{2} \ln \frac{E}{N} + x \right]$$

where we will find that this constant x (labeled as just const in the derivation above) is in fact

$$x = \frac{3}{2} \ln \frac{4\pi m}{3h^2} + \frac{5}{2}$$

where h is Planck's constant! How does that get in there? We have been doing all these calculations classically so far, but we will see later how quantum mechanics comes into this calculation.

0.7 Equilibrium Conditions

We've been saying that the most likely thing to happen in a system is what does happen, per Boltzmann. If we start out in a system that is not in the most likely state, this statement means that it will eventually trend towards the most likely state, which can be calculated by maximizing the entropy which we have just defined. Let's start with the energy. We start with a system and we want to know where it is going to end up:

$$0 = \frac{\partial}{\partial E_A} [S_A(E_A, V_A, N_A) + S_B(E_B, V_B, N_B)]$$

since technically $E_B = E - E_A$ so we can differentiate with respect to E_A since E_B is implicitly dependent on it.

$$0 = \frac{d}{dE_A} S_A(E_A, V_A, N_A) + \frac{d}{dE_B} S_B(E_B, V_B, N_B) \frac{\partial}{\partial E_A} E_B$$

so

$$\frac{\partial S_A}{\partial E_A} = \frac{\partial S_B}{\partial E_B} \quad (\text{Energy Equilibrium Condition})$$

This equation must hold *only if* the system can move energy from one chamber to the other.

A similar condition must hold for particles:

$$\frac{\partial S_A}{\partial N_A} = \frac{\partial S_B}{\partial N_B} \quad (\text{Particle Equilibrium Condition})$$

Again, this only has to hold if the particles can move from one chamber to the other.

There is a third condition, which concerns the volume. We haven't discussed this yet, but imagine the wall separating the chambers can slide back and forth. If it can, it will do so until

$$\frac{\partial S_A}{\partial V_A} = \frac{\partial S_B}{\partial V_B} \quad (\text{Volume Equilibrium Condition})$$

Often, more than one of these transmissions can occur. In general, when particles can move back and forth, they will carry energy, so the particle and energy equilibrium conditions must both be satisfied.

0.8 Distinguishable Particles

Nothing in the classical laws of physics prevent the particles from being distinguishable. When we list the microstate, we are artificially distinguishing the particles by calling all of the momenta $p_1, p_2 \dots$ and the positions $q_1, q_2 \dots$. We have to then divide out all of the permutations, since exchanging identical particles will keep the energy the same, so we don't want to double-count all of the microstates under exchanges (dividing by some fancy factorials). We could also choose to not distinguish the particles from the start and do the derivation again, and we should get the same equations for everything. Some books don't get this "correctly" and the difference is known as the Gibbs paradox.

Now imagine we have particles that have the exact same interactions, but some of them are actually different, like some particles are hydrogen and others are neon. If we separate these particles into chambers of just hydrogen and just neon and poke a hole between the chambers, someone looking at the difference between particles would say that the entropy has increased, since the particle species will mix. However, someone who isn't looking at the difference would say that the entropy remains the same. The entropy depends on how you choose to distinguish your particles.

LECTURE 7: THE MAXWELL BOLTZMANN DISTRIBUTION

Wednesday, January 29, 2020

0.9 The Maxwell Boltzmann Distribution

Start with our microstate:

$$\Pr(\{\vec{q}, \vec{p}\}) = \frac{\delta(E - H(q, p))}{N! h^{3N} \underbrace{\Omega(E, V, N)}_{\int \frac{dq dp}{h^{3N} N!} \delta(E - H(q, p))}}$$

Next, marginalize over all particles except for the first one:

$$\begin{aligned}\Pr(\vec{\mathbf{q}}_1, \vec{\mathbf{p}}_1) &= \int d^3 q_2 d^3 q_3 \cdots d^3 q_N d^3 p_2 d^3 p_3 \cdots d^3 p_N \Pr(\{\vec{\mathbf{q}}, \vec{\mathbf{p}}\}) \\ &= \frac{1}{h^{3N} N! \Omega(E, V, N)} \int d^3 q_2 d^3 q_N d^3 p_2 \cdots d^3 p_N \delta \left(E - \sum_{j=1}^N \frac{\vec{\mathbf{p}}_j^2}{2m} \right)\end{aligned}$$

Now notice that

$$\Omega \left(E - \frac{\vec{\mathbf{p}}_1^2}{2m}, V, N-1 \right) = \frac{1}{h^{3(N-1)} (N-1)!} \int d^3 q_2 \cdots d^3 q_N d^3 p_2 \cdots d^3 p_N \delta \left(E - \frac{\vec{\mathbf{p}}_1^2}{2m} - \underbrace{\sum_{j=2}^N \frac{\vec{\mathbf{p}}_j^2}{2m}}_{H_{N-1}} \right)$$

Therefore,

$$\Pr(\vec{\mathbf{q}}_1, \vec{\mathbf{p}}_1) = \frac{\Omega \left(E - \frac{\vec{\mathbf{p}}_1^2}{2m}, V, N-1 \right)}{h^3 N \Omega(E, V, N)}$$

We don't actually care about $\vec{\mathbf{q}}_1$, since nothing in this probability distribution depends on it.

$$\Pr(\vec{\mathbf{p}}_1) = \frac{V}{h^3 N} \frac{\Omega \left(E - \frac{\vec{\mathbf{p}}_1^2}{2m}, V, N-1 \right)}{\Omega(E, V, N)}$$

Again, we are going to want to maximize this probability to find the most likely microstate, and as usual, it will be easier to maximize the logarithm:

$$\ln \Pr(\vec{\mathbf{p}}_1) = \underbrace{\ln \left(\Omega \left(E - \frac{\vec{\mathbf{p}}_1^2}{2m}, V, N-1 \right) \right)}_{\sim \ln \Omega(E, V, N-1) - \frac{\vec{\mathbf{p}}_1^2}{2m} \frac{\partial \ln(\Omega(E, V, N-1))}{\partial E} + \dots} + \text{const.}$$

Here we are using a Taylor expansion. Let's define

$$\beta = \frac{\partial}{\partial E} [\ln \{\Omega(E, V, N-1)\}]$$

so

$$\ln \Pr(\vec{\mathbf{p}}_1) = \text{const.} - \beta \frac{\vec{\mathbf{p}}_1^2}{2m} + \text{higher order terms}$$

Therefore, to good approximation,

$$\Pr(\vec{\mathbf{p}}_1) = \text{const.} \times e^{-\beta \vec{\mathbf{p}}_1^2 / 2m} = \left(\frac{\beta}{2\pi m} \right)^{3/2} e^{-\beta \frac{\vec{\mathbf{p}}_1^2}{2m}}$$

since the probability distribution must be normalized. Component-wise, we can write

$$\Pr(p_{1x}) = \sqrt{\frac{\beta}{2\pi m}} e^{-\beta \frac{p_{1x}^2}{2m}}$$

What does this actually tell us? What is this β factor that we created? Recall that we define the change in momentum as

$$\Delta p = F \Delta t$$

in simple Newtonian mechanics. If we put particles in a box with a wall of area A , over a time scale Δt , how many particles will hit the wall, and how much momentum will they impart onto the wall? For any given velocity, there is a maximum distance $v_x \Delta t$ for which all particles within this distance will hit the

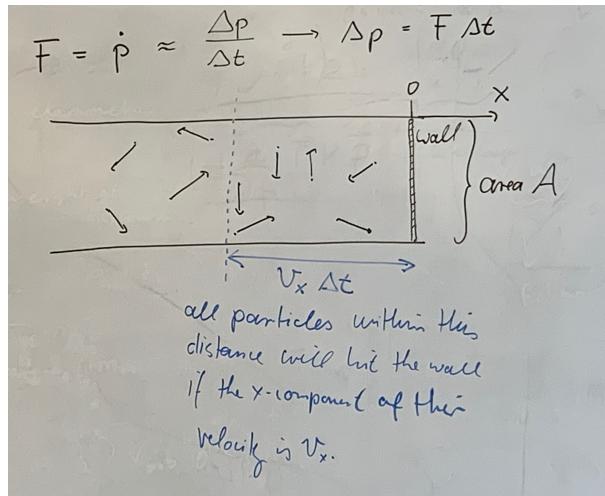


Figure 0.9.1: Pressure on wall with area A

wall if the x -component of their velocity is v_x . Particles at that distance which are moving slower will never hit the wall in Δt , and particles moving faster will have a greater chance to hit.

We can write this change as

$$F \Delta t = \Delta p_x = \int_0^\infty dp_x \left\{ \underbrace{\sqrt{\frac{\beta}{2\pi m}} e^{\beta \frac{p_x^2}{2m}}}_{\text{fraction of particles with momentum within } p_x \text{ and } p_x + dp_x} \times \underbrace{2p_x}_{\text{momentum exchange}} \times \underbrace{\frac{N}{V}}_{\text{total number of particles in hitting volume}} \times A \times \Delta t \times \underbrace{\frac{p_x}{m}}_{v_x} \underbrace{\Delta t v_x = L_x}_{\text{volume of the hitting particles}} \right\}$$

Evaluating this integral, we find

$$\Delta p_x = \sqrt{\frac{\beta}{2\pi m}} 2 \frac{NA}{V} \Delta t \frac{1}{m} \underbrace{\int_0^\infty dp_x p_x^2 e^{-\beta p_x^2/2m}}_{\frac{m}{2\beta} \left(\frac{2\pi m}{\beta} \right)^{\frac{1}{2}}} = \frac{N}{V} A \Delta t \frac{1}{\beta}$$

The pressure is defined as $P = \frac{F}{A}$ so

$$P = \frac{N}{V \beta}$$

Now let's compare. Our calculation tells us that

$$PV = \frac{N}{\beta}$$

The ideal gas law says that

$$PV = Nk_B T$$

Therefore

$$\beta = \frac{1}{k_B T}$$

Let's then go back to entropy:

$$\frac{\partial S}{\partial E} = \frac{\partial}{\partial E} [k \ln(\Omega(E, V, N))] = k \frac{\partial}{\partial E} \left(\ln \left[\Omega(E, V, \overbrace{N}^{\approx N-1}) \right] \right) = \frac{k}{k_B T} = \frac{1}{T}$$

if we allow that k in the entropy to be the Boltzmann constant. Any other choice in k just scales the relative temperatures.

Let's look at some other derivatives that we care about. We would expect the derivative of entropy with respect to volume to be the pressure, because if we have a moving wall, the point at which the entropy maximizes is where the pressures equilibrate, so the wall should stop moving and the volumes should also stop changing.

$$\frac{\partial S}{\partial V} = \frac{\partial}{\partial V} \left\{ k_B N \left[\frac{3}{2} \ln \frac{E}{N} + \ln \frac{V}{N} + \text{const.} \right] \right\} = \frac{k_B N}{V} = \frac{P}{T}$$

using the ideal gas law $PV = Nk_B T$.

Conversely,

$$\frac{P}{T} = \frac{\partial S}{\partial V}$$

gives the ideal gas law, so

$$\frac{1}{T} = \frac{\partial}{\partial E} \left\{ k_B N \left[\frac{3}{2} \ln \frac{E}{N} + \ln \frac{V}{N} + \text{const.} \right] \right\} = \frac{3}{2} k_B N \frac{1}{E}$$

so

$$E = \frac{3}{2} N k_B T$$

This is the “caloric” equation of state for a monoatomic ideal gas (as compared to the thermal equation of state—the ideal gas law).

LECTURE 8: DERIVATIVES OF ENTROPY

Friday, January 31, 2020

Last time we looked into the meaning of the partial derivative of the entropy with respect to things like the energy and the volume and discovered that these partials have to be the same in cases with two connected chambers. We have another partial derivative that we didn't go over yet:

$$\frac{\partial S}{\partial N} = ?$$

We know this for an ideal gas:

$$\begin{aligned} \frac{\partial S_{\text{ideal}}}{\partial N} &= \frac{\partial}{\partial N} \left\{ N k_B \left[\frac{3}{2} \ln \frac{E}{N} + \ln \frac{V}{N} + X \right] \right\} \\ &= k_B \left[\frac{3}{2} \ln \frac{E}{N} + \ln \frac{V}{N} + X \right] + k_B N \left[-\frac{3}{2} \frac{1}{N} - \frac{1}{N} \right] \\ &= k_B \left[\frac{3}{2} \ln \frac{E}{N} + \ln \frac{V}{N} - \frac{5}{2} + X \right] \end{aligned}$$

So far we haven't had to worry about the X , but here it doesn't go away. We don't know what it is (yet). People simply define

$$\frac{\partial S(E, V, N)}{\partial N} = -\frac{\mu}{T}$$

where we define μ as the “chemical potential”—probably one of the worst misnomers ever made, since it sounds very related to the field of chemistry. This is the equilibrium condition when particles are allowed to move back and forth. There is energy passed between them, so T must be the same. Of course, it would be nice to understand what it is and what it does, but when we look at the expression from the ideal gas, it doesn't look nice. We will later find different language in which to discuss the chemical potential which will clarify what all of these terms mean.

Recap

Knowing $S(E, V, N)$ gives us all the thermodynamic information we want.

$$\frac{\partial S}{\partial E} = \frac{1}{T} \quad \frac{\partial S}{\partial V} = \frac{P}{T} \quad \frac{\partial S}{\partial N} = -\frac{\mu}{T}$$

Hence, this relation is called a “fundamental relation”. In other words, the entropy as a function of energy, volume, and number of particles is a thermodynamic potential. The word “potential” suggests that “good things happen” when you differentiate.

- However, not every thermodynamic observable is a thermodynamic potential. For instance, $E(T, V, N)$ does not tell us everything about the equilibrium conditions of a system. However, we will call $E(T, V, N)$ and equation of state.
- Additionally, if we have the entropy expressed in different variables, say $S(T, V, N)$, then it is not a thermodynamic potential. The entropy itself is not a thermodynamic potential— $S(E, V, N)$ is.

It seems, in our case, that we have one thermodynamic potential and three equations of state (the three partial derivatives of the thermodynamic potential). For *extensive systems*, it turns out that these three equations are *not* independent. What do we mean by extensive? If you were to scale up a system (make the number of particles bigger, the volume bigger, and the energy bigger). If you scale these by the same factor and the entropy stays the same, the system is extensive.

0.9.1 Differential of Entropy

Let us look at the “differential” of the entropy. If we moved E , V , and/or N a bit, what happens to S ?

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN$$

where the subscripts imply taking the derivative while keeping the other variables constant. We actually know what all of these partial derivatives are, so we can write

$$\begin{aligned} dS &= \frac{1}{T} dE + \frac{P}{T} dV - \frac{\mu}{T} dN \\ T dS &= dE + P dV - \mu dN \end{aligned}$$

This can be written as

$$dE = T dS - P dV + \mu dN$$

We can see that $E(S, V, N)$ is a thermodynamic potential, since its derivatives will tell us everything we want to know about the system. It would be very useful to memorize this equation. This is one of the most important equations we will be using early in this course.

0.10 Microstates and Macrostates

- Microstates

Describes the microscopic configuration of a system

Typically probabilistic

$\text{Pr}(\{q, p\})$

Realm of statistical physics

- Macrostate

Describes the system by specifying a small number of constraints

$$\{E, V, N\}$$

Realm of thermodynamics

If the state is “pure” (meaning the probability distribution is a δ function), then the microstate is a property of the system. However, when we go to the macrostate, we find that we have an incomplete description of the system. This loss of entropy is the origin of entropy. In physics, we are trying to go from the stuff we know in the microstate to develop the macrostate.

LECTURE 9: THERMODYNAMICS
Monday, February 03, 2020

0.11 Equilibrium Macrostates

- State of balance
- Macroscopic description is constant in time
- No macroscopic flux of either energy or particles

Definition 0.11.1 (State Functions). Once you have chosen the macroscopic variables that define the constraints of your system, any function of them is a function of state. Its value is fixed once you know the state. As trivial as this sounds, it will have interesting implications in the future.

0.11.1 Postulates of Thermodynamics

1. Equilibrium states exist
2. The values assumed by the extensive parameters of an isolated composite system in the absence of an internal constraint are those that maximize the entropy over the set of all constrained macroscopic states.

This is a statement about what equilibrium looks like. It talks about extensive parameters, which for now we will think of as “amount of stuff” parameters. It could be the number of particles, their momentum, or energy, but not pressure. Extensive things are things you add up to get the total, so pressure and temperature are not extensive. We are also talking about an isolated composite system, meaning that it is divided into at least two parts. The absence of constraints means you are allowing some variables to change. For instance, a compartment divided by a wall has a constraint on the number of particles in a given chamber, but removing the wall allows particles to move from one side to the other.

⇒ Entropy exists, and it characterizes equilibrium

3. Additivity: The entropy of a composite system is the sum of the entropies of the subsystems.

Entropy is not a conserved quantity. If we start with two isolated systems with different entropies and bring them together, the equilibrium entropy is typically higher than the sum of the isolated entropies.

4. Monotonicity:

$$\frac{\partial S}{\partial E} > 0 \quad \text{hence, } T > 0$$

$$\frac{\partial S}{\partial V} > 0 \quad \text{hence, } P > 0$$

5. Analyticity: $S(E, V, N)$ is an analytic (meromorphic) function, possibly with countably many exceptions (phase transitions).
6. Extensivity: $S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N)$. In other words, the entropy is an extensive function.

Consequences: Choose $\lambda = \frac{1}{N}$ such that $S\left(\frac{E}{N}, \frac{V}{N}, 1\right) = \frac{1}{N}S(E, V, N)$ or $S(E, V, N) = NS\left(\frac{E}{N}, \frac{V}{N}, 1\right) = NS(e, v, 1)$ where e is the specific energy and v is the specific volume. This function only depends on two arguments, and we will call it the specific entropy s :

$$S(E, V, N) = Ns(e, v)$$

As a side note, there is a counterexample to trivial extensivity in a self-gravitating system. If we put a bunch of particles together to make a massive object, they will have some gravitational energy. If we take more particles and make a bigger planet (suppose we use twice the amount of particles and it has twice the volume), the gravitational energy will be $U = -G \frac{M^2}{R} \sim -G \frac{(\rho R^3)^2}{R} = -G\rho^2 R^5 \sim -V^{5/3}$. The energy then does not scale at the same magnitude, and this is because gravitation is not a contact force, so the energy has some dependence on scale.

0.11.2 Laws of Thermodynamics

- 0: Equilibrium is transitive
- 1: Heat is a form of energy, and energy is conserved
- 2: After the release of a constraint in a closed system, the entropy can at most go up
- 3: As $T \rightarrow 0$, the entropy goes towards a constant

The reason for the last law is quantum mechanics. If we took the entropy of an ideal gas as we've defined it now, we would find it violates this third law.

Many of the principles of thermodynamics are written in differential form, since we want to see what happens if we tweak certain parameters in a system. Our first example is

$$\underbrace{\mathrm{d}U}_{\text{System energy}} = \underbrace{\mathrm{d}Q}_{\text{Thermal energy added to system}} + \underbrace{\mathrm{d}W}_{\text{Work done onto system}} \quad (\text{First Law of Thermodynamics})$$

Different sign conventions on these right-hand terms exist, so beware! Also, the bars on the differentials are important and we will define their meaning soon.

LECTURE 10:

Wednesday, February 05, 2020

Excursion – Exact and Inexact Differentials

Consider $f(x) dx$. This can always be written as the differential of some function $F(x)$:

$$F(x) = \int^x dt f(t)$$

$$\mathrm{d}F(x) = \frac{\partial F(x)}{\partial x} dx = f(x) dx$$

However, this is no longer generally true in more than one dimension. Consider $f_x(x, y) dx + f_y(x, y) dy$. Does a function exist such that

$$\mathrm{d}F = f_x dx + f_y dy$$

If so, this differential is exact. If this is true,

$$\mathrm{d}F = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$$

As a consequence,

$$\begin{aligned}\frac{\partial f_x}{\partial y} &= \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial x} \right) \\ &= \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y} \right) \\ \frac{\partial f_x}{\partial y} &= \frac{\partial f_y}{\partial x}\end{aligned}$$

If this doesn't hold, our differential will never be exact. This condition is called an "integrability condition". If it holds, the differential is called "closed". Closed differentials are exact if the domain is simply connected.

The "bar" notation denotes differentials that are not exact. There is no F such that the equations are satisfied. In the equation $dU = dQ + dW$

How would you integrate an inexact differential?

$$dF = f_x(x, y) dx + f_y(x, y) dy$$

Define

$$\vec{f}(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} \quad d\vec{r} = \begin{pmatrix} dx \\ dy \end{pmatrix}$$

so that

$$\int_C dF = \int_C \vec{f} \cdot d\vec{r} = \int_a^b \vec{f}(\vec{r}(s)) \cdot \frac{\partial \vec{r}}{\partial s} ds$$

If the differential was exact, any path integral would only depend on the endpoints and not on the path taken. Therefore, any closed loop integral would be zero.

Differentials that are not closed still have integrating factors which close them!

$$dF(x, y) \rightarrow a(x, y)dF(x, y) = dG(x, y)$$

For example:

$$dF = y dx + 1 dy$$

Obviously, $\frac{\partial y}{\partial y} = 1 \neq \frac{\partial 1}{\partial x} = 0$, so this differential is not exact, but look at

$$e^x dF = e^x y dx + e^x dy$$

which is exact.

Can we find the G that gives this $e^x dF$? Yes, we can pick a convenient starting point $(0, 0)$ and integrate to some fixed point (x, y) . If the differential is exact, then the answer does not depend on the path.

$$\begin{aligned}G(x, y) &= \int_S dF(x', y') e^{x'} \\ &= \int_S^{(x,y)} \left[y' e^{x'} dx' + e^{x'} dy' \right] \\ &= \int_0^x dx' 0 e^{x'} + \int_0^y dy e^{x'} = ye^x\end{aligned}$$

Here, we take a straight line path along the x-axis till $x' = x$, then move straight up/down to y . Let's test if this is an exact differential.

$$dG = \frac{\partial G}{\partial x} dx + \frac{\partial G}{\partial y} dy = ye^y dx + e^x dy$$

Friday, February 07, 2020

Recall the First Law in differential form:

$$dU = dQ + dW$$

This equation implies that there is no state function for heat or work. To understand this equation, let's revisit the conservation of energy. Imagine a box with cross-sectional area A and a piston. We can apply some sort of force F to that piston, moving it an amount dx , and there's some system inside the box. The work done on the system is (assuming the direction of the force is in the opposite direction of the x axis):

$$dW = F(-dx) = \frac{F}{A}(-dx A) = -P dV$$

We could also write this as

$$dV = -\frac{1}{P}dW$$

Recall that we can use an integrating factor (here it's $-\frac{1}{P}$) to get an exact differential from an inexact differential. The work that we do on the system only depends on the path taken and not on the endpoints, whereas the volume is clearly a well-defined number at each point and doesn't depend on what "path" was taken to get to that volume.

We can also look at a small change in heat:

$$\begin{aligned} dS &= S(U + dQ, V, N) - S(U, V, N) \\ &= S(U, V, N) + \frac{\partial S}{\partial U}dQ + \dots - S(U, V, N) \\ &= \frac{\partial S}{\partial U}dQ = \frac{1}{T}dQ \end{aligned}$$

so $\frac{1}{T}$ is an integrating factor for dQ .

Using these integrating factors, we can see that

$$dU = T dS - P dV (+\mu dN)$$

which we derived using statistical mechanics. Again, memorize this.

0.12 Heat Engines

0.12.1 Reversible and Irreversible Processes

Imagine we have some system, we release some constraint, and as a consequence, the system spontaneously goes to a new equilibrium state and the entropy increases. The reverse does not happen spontaneously. This is called an irreversible process. The counterpart to this is a reversible process. Imagine we move a system through a sequence of equilibrium states. If you move it back to the starting point (not necessarily along the same path), no change in entropy has occurred. This doesn't mean that the entropy was constant along the path, but once you return to the start, the entropy returns to its initial value. This is a reversible process. In practice, reversible processes are "quasistatic"—they require "infinitely slow" motion. However, there do exist quasistatic processes which are not reversible. As a counterexample, imagine a box with a thick wall between two compartments. One side of this box is hot and the other is cold, and suppose the wall conducts heat but poorly. Depending on how poorly the wall conducts heat, we can make the equilibration of the compartments as slow as we want, although this is certainly not a reversible process. The change of state is very slow but still irreversible.

Definition 0.12.1. Heat Engine A heat engine is a cyclically operating device that takes "heat" and turns it into "work".

From the First Law, we already have a constraint on the amount of work we can do:

$$dW \leq dQ$$

where dW is the work **done by** the engine and dQ is the heat going into the system. Notice that this is a different sign convention than when we first defined it. One might think we could also take some of the dU to exceed this limit, but this is not allowed because the heat engine is cyclic, so it must return to the same state on each cycle. Therefore, $dU = 0$ over any cycle.

Next, the Second Law tells us that, in fact,

$$dW < dQ$$

even under ideal reversible conditions! We can imagine a system where no energy is ever lost to friction or other irreversible processes, and one might expect such a machine would be perfectly efficient. Next time, we will show why this isn't in fact the case.

LECTURE 12: HEAT ENGINES, CONTINUED

Monday, February 10, 2020

0.13 Heat Engine Efficiency

In the last lecture, we learned that

$$dW \leq dQ$$

where W is the work done by a heat engine and Q is the heat put into the system. The inequality is not surprising, you shouldn't be able to create energy from nothing. However, the fact that these are never actually equal is not trivial. Suppose we put heat into the engine (reversibly). The entropy will increase by

$$dS = \frac{dQ}{T}$$

When the engine runs cyclically, the entropy must also return to its original state. There's nothing wrong with entropy decreasing, it just will not happen spontaneously. The only way to do this is to remove heat from the engine (so that dQ is negative). We have to subtract the exact same amount of heat from the engine that we put in to make the entropy return to its starting point, so there is no heat left over to do work. However, there is another factor which we can change (the temperature) which will allow us to remove a different amount of heat. If we get rid of heat at a lower temperature, there will still be heat left over to do work.

We put heat into the engine at temperature T_H and remove heat at temperature T_C . In the best case scenario,

$$dW = dQ_H + dQ_C$$

Again, be cautious with the sign conventions. dQ_C is negative, since it is heat being removed from the system.

Over one cycle,

$$0 = dS = \frac{dQ_H}{T_H} + \frac{dQ_C}{T_C}$$

The first equality holds because we run the system cyclically. The second inequality holds because we run the system reversibly. Therefore,

$$dQ_C = -\frac{T_C}{T_H} dQ_H$$

so

$$dW = dQ_H - \frac{T_C}{T_H} dQ_H = dQ_H \left(1 - \frac{T_C}{T_H}\right)$$

Therefore, the efficiency of the engine (bang for our buck in very scientific terms) is

$$\eta := \frac{dW}{dQ_H} = 1 - \frac{T_C}{T_H} \quad (\text{Carnot Efficiency})$$

This is the maximal efficiency of an engine that runs cyclically and reversibly. Notice we made no assumptions about the process or the mechanics of the engine. We only used the facts that the entropy cannot be different at the end of a cycle and the process must be reversible. Typically, these efficiencies tend to be on the order of 30–40%. Since we run the engine reversibly, we can actually imagine reversing it. This would mean we input work and use it to take heat out of the cold reservoir and put it into the hot reservoir. This is not something heat will do spontaneously. If you care about the cold reservoir, you can make a refrigerator. If you care about the hot reservoir, you can make a heat pump. In both cases, we can now define something called an efficiency ratio. For the refrigerator,

$$\varepsilon_R = \frac{dQ_C}{-dW} = \frac{T_C}{T_H - T_C}$$

This ratio *can be* bigger than 1! For a heat pump,

$$\varepsilon_{HP} = \frac{-dQ_C}{-dW} = \frac{1}{\eta} > 1$$

Heat pumps are more than 100% efficient!

0.14 Thermodynamic Potentials

We know that $S(U, V, N)$ contains “all” the thermodynamic information we care about. However, its natural variables are not always convenient. Can we express the information content of $S(U, V, N)$ via a different function which depends on more convenient variables?

LECTURE 13: THE LEGENDRE TRANSFORMATION
Wednesday, February 12, 2020

0.15 Legendre Transformations

Suppose we are given a function $f(x)$ along with its derivative $p = p(x) = f'(x)$. Can we construct a function $g(p)$ which contains “the same information” as $f(x)$? In other words, can we construct a $g(p)$ which can be transformed in a similar way back into $f(x)$? Naively, we could try solving for $x(p)$ by inverting the function $p = f'(x)$ or $x = f'^{-1}(p)$. Unfortunately, this is only possible if f' is monotonic which implies f is either convex or concave. Let’s assume the universe is nice and we have such a function. Now we insert $x(p)$ into $f(x)$ so that our new function is

$$g(p) = f(x(p))$$

This does not work, despite how nice it looks, because $g(p)$ does not contain enough information to reconstruct $f(x)$.

Example. Suppose $f(x) = \frac{1}{2}(x - x_0)^2$. This is a parabola, so it is convex. $p = f'(x) = x - x_0$ so that $x(p) = p + x_0$. Therefore

$$g(p) = f(x(p)) = \frac{1}{2}(p + x_0 - x_0)^2 = \frac{1}{2}p^2$$

We have lost all information about where the parabola is centered. The answer no longer contains x_0 . ◇

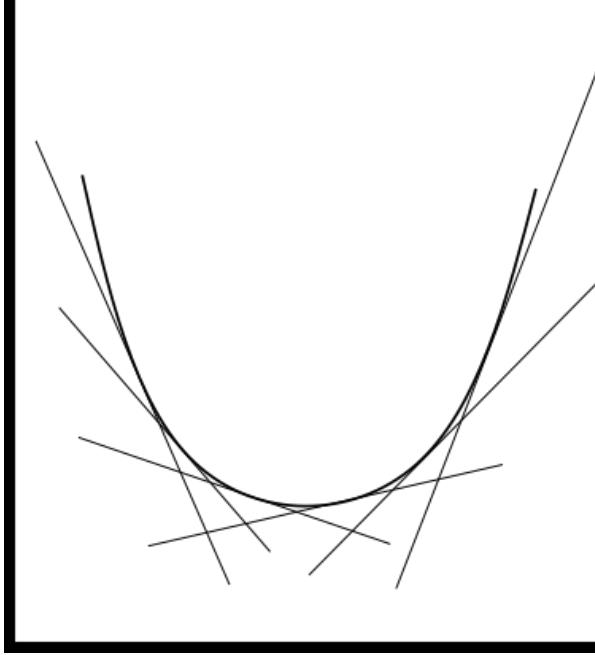


Figure 0.15.1: A convex function is bounded by its tangents and each tangent has a unique y -intercept

We need to think about a way to specify a function via its slopes. In ?? we see that any function can be described by its entire family of tangents. In fact, if the function is strictly convex or concave, it is bounded by its tangents and each tangent line will have a unique y -intercept. We can surely use this to reconstruct the original function.

For a given point on the function, the tangent line can be described as

$$t_{x_0}(x) = f(x_0) + f'(x_0)(x - x_0)$$

The y -intercept of this tangent is

$$t_{x_0}(0) = f(x_0) - x_0 f'(x_0)$$

If f is convex, then $f'(x)$ is monotonic, and there is a unique relation between the intercept and slope. Therefore, $p = f'(x_0)$ can be uniquely solved for $x_0(p)$. Hence, either x_0 or p can be used to specify a given tangent. Let's use the slope to characterize the intercept:

$$t_p(0) = f(x_0(p)) - x_0(p)p$$

For any choice of p , we can calculate the intercept, so we know the entire family of tangents. Therefore, for a convex or concave function, knowing the derivative gives us an exact definition of all the tangent lines.

$$g(p) = t_p(0) = f(x_0(p)) - x_0(p)p \quad (\text{Legendre Transformation})$$

We can formalize this as

$$g(p) = \min_x \{f(x) - xp\} \quad \text{convex} \quad f$$

For concave function, just take the maximum. Let's find the minimum:

$$0 = \frac{\partial}{\partial x} \{f(x) - xp\} = f'(x) - p$$

or $p = f'(x)$. Since f is convex, we can solve this for x :

$$x = f'^{-1}(p) = x(p)$$

Therefore

$$g(p) = f(x(p)) - x(p)p$$

How do we construct the reverse transformation? For any point x , the point that intercepts our original line is the largest value any tangent line attains at that value of x :

$$\begin{aligned} \max_p \{t_p(0) + px\} &= \max_p \{f(x_0(p)) - x_0(p)p + px\} \\ 0 = \frac{\partial}{\partial p} \{f(x_0(p)) - x_0(p)p + px\} &= \underbrace{\frac{\partial f}{\partial x_0}}_p \frac{\partial x_0}{\partial p} - \frac{\partial x_0}{\partial p}p - x_0(p) + x \end{aligned}$$

so the first two terms cancel. We are left with $x = x_0$. Now inserting this back into our original equation, we find

$$\max_p \{t_p(0) + px\} = f(x) - xp + px = f(x) : W$$

Putting this all together, we can write this transform as a proper pair of transforms:

$$g(p) = \min_x \{f(x) - xp\}$$

$$f(x) = \max_p \{g(p) + px\}$$

As it turns out, there is an equivalent transformation if we swap the $+$ and $-$ signs in both transforms.

$$g'(p) = \frac{\partial g}{\partial p} = \frac{\partial}{\partial p} \{f(x(p)) - px(p)\} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial p} - p \frac{\partial x}{\partial p} - x(p) = -x$$

Notice that this is the same way we defined $p = f'(x)$ with a different sign:

$$df = p dx \quad dg = -x dp$$

LECTURE 14: LEGENDRE TRANSFORMATIONS

Friday, February 14, 2020

Recall from last lecture that we defined the Legendre transformation:

$$\begin{aligned} g(p) &= \min_x \{f(x) - xp\} && \text{(Legendre Transformation)} \\ f(x) &= \max_p \{g(p) + px\} && \text{(Inverse Legendre Transformation)} \end{aligned}$$

where

$$\begin{aligned} f'(x) &= p & \text{or} & \quad df = p dx \\ g'(p) &= -x & \text{or} & \quad dg = -x dp \end{aligned}$$

The plus and minus in the transform and its inverse can be switched with no effect to the validity of the transformation.

0.16 Helmholtz Free Energy

We start with the entropy $S(U, V, N)$ which we know is monotonically increasing in U since $\frac{\partial S}{\partial U} = \frac{1}{T}$. We know it's also concave in U . We can solve this for $U(S, V, N)$ (we don't care about V and N right

now, they're just coming along for the ride). The same information is contained in both equations, so this is still a thermodynamic potential.

$$dU = \underbrace{T}_{\left(\frac{\partial U}{\partial S}\right)_{V,N}} dS - \underbrace{P}_{\left(\frac{\partial U}{\partial V}\right)_{S,N}} dV + \underbrace{\mu}_{\left(\frac{\partial U}{\partial N}\right)_{S,V}} dN$$

Because $S(U)$ is concave, $U(S)$ is convex (and still monotonically increasing). S is an awkward variable. Let's change it to $T = \left(\frac{\partial U}{\partial S}\right)_{V,N}$. We can do this with a Legendre transform:

$$F(T, V, N) = \min_S \{U(S, V, N) - TS\} \quad (\text{Helmholtz Free Energy})$$

This function F is the Helmholtz free energy. We could choose either a minus or plus sign, but minus is conventionally used. In order to turn $T dS$ into $-S dT$, we need a minus sign. Say we would like to calculate dF :

$$\begin{aligned} dF &= d(U - TS) \\ &= dU - (T dS + S dT) \\ &= T dS - P dV + \mu dN - T dS - S dT \\ &= -S dT - P dV + \mu dN \end{aligned}$$

If we had chosen a + here, the $T dS$ terms would not have canceled nicely. This will turn out differently for other situations. With this function, we now have some additional definitions for the entropy, pressure, and chemical potential:

$$-S = \left(\frac{\partial F}{\partial T}\right)_{V,N} \quad -P = \left(\frac{\partial F}{\partial V}\right)_{T,N} \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}$$

Notice that we now have a function $S(T, V, N)$, which is *not* a thermic potential.

Consider a system at constant T and N such that $dT = dN = 0$:

$$dF = -P dV = dW$$

The maximum amount of work that can be extracted from a system at fixed T and N is equal to the free energy difference between the initial and final states. Note the distinction between the energy and free energy. If you let a system do work and want to see how much work was done, you can't just use the difference in energies ΔU , you have to use the difference in the Helmholtz free energies ΔF . We can do this by replacing other variables other than S with T .

0.17 Enthalpy

Instead of replacing S with T , replace V with P . The result is called the enthalpy:

$$H(S, P, N) = \min_V \{U(S, V, N) + PV\} \quad (\text{Enthalpy})$$

We need the transformation to have a different sign because dU has a $-P dV$ term.

$$\begin{aligned} dH &= T dS + V dP + \mu dN \\ T &= \left(\frac{\partial H}{\partial S}\right)_{P,N} \quad V = \left(\frac{\partial H}{\partial P}\right)_{S,N} \quad \mu = \left(\frac{\partial H}{\partial N}\right)_{S,P} \end{aligned}$$

Let's look at a system with constant P and N (imagine a liquid in a test tube, if the tube is open, the pressure is just 1atm and the number of particles is not changing):

$$dH = T dS = dQ$$

The change in heat under these conditions is the change in enthalpy.

0.18 Gibbs Free Energy (Free Enthalpy)

What if we exchanged both S for T and V for P ?

$$G(T, P, N) = \min_{S, V} \{U(S, V, N) - TS + PV\} \quad (\text{Gibbs Free Energy})$$

$$dG = -S dT + V dP + \mu dN$$

$$-S = \left(\frac{\partial G}{\partial T}\right)_{P, N} \quad V = \left(\frac{\partial G}{\partial P}\right)_{T, V} \quad \mu = \left(\frac{\partial G}{\partial N}\right)_{T, P}$$

For a system with constant T and P ,

$$dG = \mu dN$$

The change in the Gibbs free energy when adding one particle is the chemical potential.

0.19 Grand Potential

Now let's exchange S for T and N for μ :

$$\Omega(T, V, mu) = \min_{S, N} \{U(S, V, N) - TS - \mu N\} \quad (\text{Grand Potential})$$

$$d\Omega = -S dT - P dV - N d\mu$$

$$-S = \left(\frac{\partial \Omega}{\partial T}\right)_{V, \mu} \quad -P = \left(\frac{\partial \Omega}{\partial V}\right)_{T, \mu} \quad -N = \left(\frac{\partial \Omega}{\partial \mu}\right)_{T, V}$$

This potential is useful when the number of particles is not fixed, and turns out to be very helpful in quantum statistics.

One Final Trick

The expressions

$$\min_{S, \dots} \{U(S, V, N) - TS \pm \dots\}$$

and

$$\min_{U, \dots} \{U - TS(U, V, N) \pm \dots\}$$

are equivalent! We are just running over the same set of U , V , and N , but labeling U and S differently.

LECTURE 15: EXTENSIVITY

Monday, February 17, 2020

Variables in thermodynamics are typically intensive or extensive. Intensive variables are the ones which do not scale as we change the size of the system, whereas extensive variables do. For example, entropy is extensive. Note that

$$dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN$$

If we take a system defined by variables U , V , and N , scaling the system to λV , λN and λU will demonstrate degree-one extensivity:

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N)$$

Knowing this, we can use one of Euler's theorems:

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \lambda^m f(x_1, \dots, x_n)$$

then

$$\frac{d}{d\lambda} f(\lambda sxu_1, \dots, \lambda x_n) = m\lambda^{m-1} f(x_1, \dots, x_n)$$

This is equivalent to

$$\sum_{i=1}^n \frac{\partial f}{\partial x^i} \Big|_{\lambda x^i} \underbrace{\frac{d\lambda x_i}{dx^i}}_{x^i} = m\lambda^{m-1} f(x_1, \dots, x_n)$$

In the limit $\lambda \rightarrow 1$,

$$\lim_{\lambda \rightarrow 1} \sum_{i=1}^n \frac{\partial f}{\partial x^i} \Big|_{\lambda x^i} x^i = m f(x_1, \dots, x_n) = \sum_{i=1}^n \frac{\partial f}{\partial x^i} x^i$$

Using this result, we can write

$$S = \left. \frac{\partial S}{\partial U} \right|_{V,N} U + \left. \frac{\partial S}{\partial V} \right|_{U,V} V + \left. \frac{\partial S}{\partial N} \right|_{V,U} N$$

Plugging in our known derivatives of S , we have

$$S = \frac{1}{T} U + \frac{P}{T} V - \frac{\mu}{T} N$$

or

$$U = TS - PV + \mu N$$

We had this before, but as a differential. This only works because of extensivity. However, this only holds well with microsystems with short-range interactions. We previously showed that gravitational systems are not consistent with this theorem.

We know that

$$dU = T dS - P dV + \mu dN$$

but the relationship that we just found means that, due to extensivity,

$$dU = T dS + S dT - P dV - V dP + N d\mu + \mu dN$$

which implies

$$0 = S dT - V dP + N d\mu \quad (\text{Gibbs-Duhem Relation})$$

This tells us that T , P , and μ can not independently vary. This tells us that there is no thermodynamic potential that can be defined using only intensive variables. You need to have an extensive variable.

Let's rewrite this as

$$d\mu = -\frac{S}{N} dT + \frac{V}{N} dP$$

Because of the way we are writing the differential, we have to put it in terms of the differentials dT and dP : $\frac{S(N,T,P)}{N} = s(T, P)$ and $\frac{V}{N} = v(T, P)$ is the specific volume.

$$d\mu = -s(T, P) dT + v(T, P) dP$$

We can also write the extensivity equation in terms of the entropy:

$$S = \frac{U}{T} + \frac{P}{T} V + \frac{\mu}{T} N$$

such that

$$dS = dU \frac{1}{T} + U d\left(\frac{1}{T}\right) + dV \frac{P}{T} + V d\left(\frac{P}{T}\right) - d\left(\frac{\mu}{T}\right) N - \frac{\mu}{T} dN$$

Again, from our original definition of dS , this implies

$$0 = U d\left(\frac{1}{T}\right) + V d\left(\frac{P}{T}\right) - N d\left(\frac{\mu}{T}\right)$$

We can define new variables $\tilde{P} = \frac{P}{T}$ and $\tilde{T} = \frac{1}{T}$ such that

$$d\left(\frac{\mu}{T}\right) = \frac{U}{N} d\tilde{T} + \frac{V}{N} d\tilde{P}$$

or

$$d\left(\frac{\mu}{T}\right) = \frac{U(\tilde{T}, \tilde{P}, N)}{N} d\tilde{T} + \frac{V(\tilde{T}, \tilde{P}, N)}{N} d\tilde{P}$$

From this, we can get relations like

$$\frac{\partial}{\partial \tilde{T}} (\mu \tilde{T}) \Big|_{\tilde{P}} = \frac{U(\tilde{T}, \tilde{P}, N)}{N} = u(\tilde{T}, \tilde{P})$$

and

$$\frac{\partial}{\partial \tilde{P}} (\mu \tilde{T}) \Big|_{\tilde{T}} = \frac{V(\tilde{T}, \tilde{P}, N)}{N} = v(\tilde{T}, \tilde{P})$$

These relations might be useful for a particular experiment. Our goal in doing this is to write things we know in many different ways. Let's now use Legendre transforms to write thermodynamic potentials using these new relations, particularly the Euler relation $U = TS - PV - \mu N$. Using extensivity, we can write

$$\begin{array}{lll} H = U + PV & = TS & + \mu N \\ F = U - TS & = & -PV + \mu N \\ G = U - TS + PV & = & + \mu N \\ \Omega = U - TS - \mu N & = & -PV \end{array}$$

0.20 Various Partial Derivatives

The majority of Chapter 14 in the textbook concerns

$$\frac{\partial(\text{this})}{\partial(\text{that})} \Big|_{\text{something}}$$

or partial derivatives where something is held constant. Let's look at an ideal gas with $U = \frac{3}{2}Nk_B T$:

$$\frac{\partial U}{\partial V} \Big|_{S,N} = -P < 0$$

$$\frac{\partial U}{\partial V} \Big|_{V,N} = \frac{\partial}{\partial V} \left(\frac{3}{2}PV \right) = \frac{3}{2}P > 0$$

$$\frac{\partial U}{\partial V} \Big|_{T,N} = 0$$

As we can see here, general derivatives like this are meaningless unless we specify what is being kept constant. We have some additional partial derivatives defined by $dU = T dS - P dV + \mu dN$, but what if we want more from this? We can take second derivatives, and these should give us more information.

	$\frac{\partial}{\partial S}$	$\frac{\partial}{\partial V}$	$\frac{\partial}{\partial N}$
$\frac{\partial}{\partial S}$	$\frac{\partial^2}{\partial S^2}$	$\frac{\partial}{\partial S} \frac{\partial}{\partial V}$	$\frac{\partial}{\partial S} \frac{\partial}{\partial N}$
$\frac{\partial}{\partial V}$		$\frac{\partial^2}{\partial V^2}$	$\frac{\partial}{\partial V} \frac{\partial}{\partial N}$
$\frac{\partial}{\partial N}$			$\frac{\partial^2}{\partial N^2}$

LECTURE 16: SECOND DERIVATIVES OF THERMODYNAMIC POTENTIALS

Wednesday, February 19, 2020

Let's examine the derivatives $\frac{\partial^2 U}{\partial S^2}$, $\frac{\partial^2 U}{\partial S \partial V}$, and $\frac{\partial^2 U}{\partial V^2}$. First, let's write out our usual equation and transform it with a product rule:

$$\begin{aligned} dU &= T dS - P dV + \mu dN \\ d(U + PV) &= T dS + V dP + \mu dN \end{aligned}$$

We can now define some derivatives of this equation:

$$\left. \frac{1}{V} \frac{\partial V}{\partial T} \right|_{P,N} = \alpha$$

Here, α is the thermal expansion at a constant pressure (usually written in units related to the volume).

$$K_T = -\left. \frac{1}{V} \frac{\partial V}{\partial P} \right|_{T,N}$$

K_T is the isothermal compressibility.

$$C_V = \left. \frac{1}{N} \frac{dQ}{dT} \right|_{V,N} = \left. \frac{1}{N} \frac{T dS}{dT} \right|_{V,N} = \left. \frac{T}{N} \frac{\partial S}{\partial T} \right|_{V,N}$$

is the heat capacity at constant volume and

$$C_P = \left. \frac{T}{N} \frac{\partial S}{\partial T} \right|_{P,N}$$

is the heat capacity at constant pressure.

We need to relate partial derivatives of various quantities to other quantities which are usually easier to measure. One method is by Maxwell relations:

$$\begin{aligned} U \longrightarrow dU &= T dS - P dV + \mu dN \implies \left. \frac{\partial T}{\partial V} \right|_{S,N} = -\left. \frac{\partial P}{\partial S} \right|_{V,N} \\ F \longrightarrow dF &= -S dT - P dV + \mu dN \implies \left. \frac{\partial \mu}{\partial T} \right|_{N,V} = -\left. \frac{\partial S}{\partial N} \right|_{T,V} \\ &\vdots \end{aligned}$$

Suppose we wanted to find a relation for

$$\left. \frac{\partial T}{\partial P} \right|_{S,\mu} = \left. \frac{\partial ?}{\partial ?} \right|_?$$

We need the differentials of P , S , and μ , so we want one of the potentials of the form

$$T dS + V dP - N d\mu$$

Knowing that we can switch the order of second derivatives (take one derivative first),

$$\frac{\partial^2 H}{\partial S \partial P} = \left. \frac{\partial T}{\partial P} \right|_{S,\mu} = \left. \frac{\partial V}{\partial S} \right|_{P,\mu}$$

Additionally, we can often write

$$\left. \frac{\partial A}{\partial B} \right|_{C,D} = \frac{1}{\left. \frac{\partial B}{\partial A} \right|_{C,D}}$$

The Maxwell relations can usually be written in this form. The other method for deriving these relations is using Jacobians. In general,

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \left. \frac{\partial v}{\partial x} \right|_y & \left. \frac{\partial v}{\partial y} \right|_x \\ \left. \frac{\partial u}{\partial x} \right|_y & \left. \frac{\partial u}{\partial y} \right|_x \end{vmatrix}$$

Using relations with products of determinants, we can prove that

$$\frac{\partial(u, v)}{\partial(x, y)} = \frac{\partial(u, v)}{\partial(A, B)} \frac{\partial(A, B)}{\partial(x, y)}$$

Another interesting property is that exchanging rows or columns in the Jacobian introduces minus signs:

$$\frac{\partial(A, B, C)}{\partial(X, Y, Z)} = -\frac{\partial(C, B, A)}{\partial(X, Y, Z)} = -\frac{\partial(A, B, C)}{\partial(Y, X, Z)}$$

Another interesting property is that

$$\frac{\partial(u, y)}{\partial(x, y)} = \begin{vmatrix} \left. \frac{\partial u}{\partial x} \right|_y & \left. \frac{\partial u}{\partial y} \right|_x \\ \underbrace{\left. \frac{\partial y}{\partial x} \right|_y}_0 & \underbrace{\left. \frac{\partial y}{\partial y} \right|_x}_1 \end{vmatrix} = \left. \frac{\partial u}{\partial x} \right|_y$$

Because dU is an exact derivative,

$$d(dU) = 0 = dT dS - dP dV + d\mu dN$$

Suppose we fix N , then

$$dT dS = dP dV$$

or

$$\frac{\partial(T, S)}{\partial(P, V)} = 1$$

Additionally, we can use the properties we found above to write

$$\left. \frac{\partial P}{\partial T} \right|_{V, N} = \frac{\partial(P, V)}{\partial(T, V)} = \frac{\partial(P, V)}{\partial(P, T)} \frac{\partial(P, T)}{\partial(T, V)} = \left. \frac{\partial V}{\partial T} \right|_P \frac{1}{\frac{\partial(T, V)}{\partial(P, T)}} = \frac{\left. \frac{\partial V}{\partial T} \right|_P}{-\left. \frac{\partial V}{\partial P} \right|_T} = \frac{\alpha}{K_T}$$

We have shown that

$$\left. \frac{\partial P}{\partial T} \right|_{V, N} = \frac{\alpha}{K_T}$$

the right-hand side of which is something that can be measured.

Another relation that can be derived is

$$C_P - C_V = T \frac{V}{N} \frac{\alpha^2}{K_T} > 0$$

We will prove this in a future lecture.

LECTURE 17: MAXWELL RELATIONS

Friday, February 21, 2020

From last lecture, we decided we don't really care about the number of particles, and we can write partial derivatives in terms of Jacobians:

$$\left. \frac{\partial P}{\partial T} \right|_{V, N} \equiv \left. \frac{\partial P}{\partial T} \right|_V = \frac{\partial(P, V)}{\partial(T, V)} = \frac{\partial(P, V)}{\partial(\textcolor{red}{P}, \textcolor{red}{T})} \frac{\partial(\textcolor{red}{P}, \textcolor{red}{T})}{\partial(T, V)}$$

Typically, we want to introduce a Jacobian made of the variables used in the original derivative (in this case, P and T). It usually makes for a simple transformation, but it is obviously not the only possibility. For the magic to work out, we need things to be in the same “slot”, and switching ordering means introducing minus signs:

$$-\frac{\partial(V, P)}{\partial(T, P)} \frac{\partial(P, T)}{\partial(\textcolor{red}{V}, \textcolor{red}{T})} = -\left. \frac{\partial V}{\partial T} \right|_P \left. \frac{\partial P}{\partial V} \right|_T$$

We can then use inverses:

$$-\left. \frac{\partial V}{\partial T} \right|_P \left. \frac{\partial P}{\partial V} \right|_T = \frac{\frac{1}{V} \left. \frac{\partial V}{\partial T} \right|_P}{-\frac{1}{V} \left. \frac{\partial V}{\partial P} \right|_T} = \frac{\alpha}{K_T}$$

0.21 The Joule-Thomson Effect

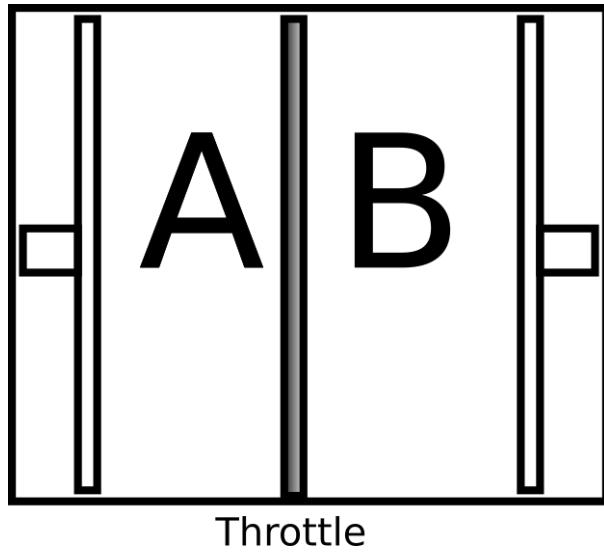


Figure 0.21.1: Demonstration of the Joule-Thomson Effect

Fun fact, Thomson is William Thomson, aka Kelvin. Let's imagine a container with two pistons (see ??). The gas in compartment A is at pressure P_A and the gas in B is at P_B . Suppose we start with $P_A > P_B$ and we maintain this while moving the pistons such that the gas moves from compartment A to compartment B . The energy in B can be written

$$U_B = U_A + P_A V_A - P_B V_B$$

Note that this implies the enthalpy stays constant (isenthalpic process):

$$H_A = U_A + P_A V_A = U_B + P_B V_B = H_B$$

How does the temperature of the gas change? For a small pressure difference at constant N and H , we have

$$dT = \left. \frac{\partial T}{\partial P} \right|_{H, N} dP = \mu_{JT}$$

where μ_{JT} is the “Joule-Thomson coefficient”. If $\mu_{JT} > 0$, the gas cools in the process. Let's use Maxwell relations to transform this into something we can measure:

$$\mu_{JT} = \left. \frac{\partial T}{\partial P} \right|_{H, N} = \frac{\partial(T, H)}{\partial P, H} = -\frac{\partial(T, H)}{\partial(T, P)} \frac{\partial(T, P)}{\partial(H, P)} = -\frac{\left. \frac{\partial H}{\partial P} \right|_T}{\left. \frac{\partial H}{\partial T} \right|_P}$$

Next, let's work out what these derivatives are. Recall the enthalpy $H(S, P, N)$ has a differential form:

$$dH = T dS + V dP + \mu dN$$

We agreed at doing this at constant N , so we really only need $dH = T dS + V dP$.

$$\frac{\partial H}{\partial P} \Big|_T = T \frac{\partial S}{\partial P} \Big|_T + V \frac{\partial P}{\partial T} \Big|_T$$

We can justify this with the chain rule:

$$\begin{aligned} \frac{\partial H(S, P, N)}{\partial P} \Big|_{T,N} &= \frac{\partial H}{\partial S} \Big|_{P,N} \frac{\partial S}{\partial P} \Big|_{T,N} + \frac{\partial H}{\partial P} \Big|_{S,N} \frac{\partial P}{\partial P} \Big|_{T,N} + \frac{\partial H}{\partial N} \Big|_{S,P} \frac{\partial N}{\partial P} \Big|_{T,N} \\ &= T \frac{\partial S}{\partial P} \Big|_{T,N} + V \frac{\partial P}{\partial P} \Big|_{T,N} + \mu \frac{\partial N}{\partial P} \Big|_{T,N} \end{aligned}$$

$$\frac{\partial P}{\partial P} = 1$$

so we really only need to find

$$\frac{\partial S}{\partial P} \Big|_T = - \frac{\partial V}{\partial T} \Big|_P$$

using Maxwell's relations.

How does the volume change while keeping the pressure fixed? This is typically thought of as the thermal expansion, but we need that $\frac{1}{V}$ factor to make it into the defined material property:

$$\frac{\partial H}{\partial P} \Big|_T = -TV \frac{1}{V} \frac{\partial V}{\partial T} \Big|_P + V = V(1 - T\alpha)$$

Now for the denominator:

$$\frac{\partial H}{\partial T} \Big|_P = T \frac{\partial S}{\partial T} \Big|_P + V \frac{\partial P}{\partial T} \Big|_P + (\text{d}N \text{ stuff}) = N \frac{T}{N} \frac{\partial S}{\partial T} \Big|_P = NC_p$$

Therefore,

$$\mu_{JT} = - \frac{V(1 - T\alpha)}{NC_p} = - \frac{V}{NC_p}(T\alpha - 1)$$

What happens when we do this with an ideal gas?

$$\alpha = \frac{1}{V} \frac{\partial V}{\partial T} \Big|_P = \frac{1}{V} \frac{\partial}{\partial T} \left(\frac{Nk_B T}{P} \right) = \frac{1}{V} \frac{Nk_B}{P} = \frac{1}{T}$$

Therefore, $\mu_{JT, \text{ideal}} = 0$! We can't use ideal gasses to enact the Joule-Thomson effect to cool or heat a gas.

0.22 Heat Capacity

Heat capacities answer the question of how much heat goes into a system when you heat it at constant pressure or volume. Let's start with constant volume:

$$C_v = \frac{T}{N} \frac{\partial S}{\partial T} \Big|_{V,N} = \frac{T}{N} \frac{\partial(S,V)}{\partial(T,V)} = \frac{T}{N} \frac{\partial(S,V)}{\partial(\mathbf{T},\mathbf{P})} \frac{\partial(\mathbf{T},\mathbf{P})}{\partial(T,V)}$$

Why would we do this? The first Jacobian has no like terms, so we could expand it again and end up with more partial derivatives. The second option is that we did an expansion that we shouldn't do. The third option is to literally write down what the Jacobian is. It's just the determinant of a matrix of a bunch of partial derivatives, so we can write it down. Let's first deal with the easy second term:

$$\frac{\partial(T,P)}{\partial(T,V)} = \frac{\partial P}{\partial V} \Big|_T = \frac{1}{\frac{\partial V}{\partial P} \Big|_T} = \frac{1}{-v \frac{1}{-v} \frac{\partial V}{\partial P} \Big|_T} = \frac{1}{-V\kappa_T}$$

where κ_T is the compressibility. Now let's look at the unusual first term:

$$\frac{\partial(S,V)}{\partial(T,P)} = \begin{vmatrix} \frac{\partial S}{\partial T} \Big|_P & \frac{\partial S}{\partial P} \Big|_T \\ \frac{\partial V}{\partial T} \Big|_P & \frac{\partial V}{\partial P} \Big|_T \end{vmatrix} = \underbrace{\frac{\partial S}{\partial T} \Big|_P}_{*} \underbrace{\frac{\partial V}{\partial P} \Big|_T}_{**} - \underbrace{\frac{\partial V}{\partial T} \Big|_P}_{†} \underbrace{\frac{\partial S}{\partial P} \Big|_T}_{††}$$

Let's now evaluate each of these partial derivatives separately:

$$* \frac{\partial S}{\partial T} \Big|_P = \frac{N}{T} \frac{T}{N} \frac{\partial S}{\partial T} \Big|_P = \frac{N}{T} C_p$$

$$** \frac{\partial V}{\partial P} \Big|_T = -V\kappa_T$$

$$\dagger \frac{\partial V}{\partial T} \Big|_P = V\alpha$$

$$\dagger\dagger \frac{\partial S}{\partial P} \Big|_T \implies -S dT + V dP \implies \frac{\partial S}{\partial P} \Big|_T = -\frac{\partial V}{\partial T} \Big|_P = -V\alpha$$

Therefore,

$$\frac{\partial(S,V)}{\partial(T,P)} = \frac{N}{T} C_p (-V\kappa_T) - V\alpha (-V\alpha)$$

Putting everything together, we find that

$$C_V = \frac{T}{N} \frac{-\frac{NV}{T} C_p \kappa_T + (V\alpha)^2}{-V\kappa T} = C_p - \frac{TV\alpha^2}{N\kappa_T}$$

Typically, this equation is written

$$C_P - C_V = \frac{TV\alpha^2}{N\kappa_T}$$

While α need not be positive, α^2 must be, and in conventional cases, $\kappa_T > 0$ (materials don't expand when you compress them). Therefore, $C_P > C_V$.

0.23 Extremum Principles

Let's recall the reason we extremize entropy. Take a box with a movable wall that for the moment is fixed. If you unbolt the wall, it will move to an equilibrium position and is very unlikely to move back to its original position. No work was done on this box, and no heat went in or out of the box, but the entropy went up. After we released some extensive constraint (volume, in this case), the entropy increased irreversibly. Now that the system has achieved a new equilibrium state, the entropy has achieved its maximum possible value under its new conditions. If we suppose the constraint is parameterized by the position of the wall x ,

$$\left. \frac{\partial S}{\partial x} \right|_U = 0$$

and

$$\left. \frac{\partial^2 S}{\partial X^2} \right|_U < 0$$

so that this point is a maximum.

Now we are going to repeat this but with an interesting change. In this situation, we know that unbolting the wall will irreversibly move us towards the equilibrium. But what if we wanted to do this reversibly? How do we do this? We have to make the wall move slowly, so imagine there is a stick attached to the wall which extends outside the box and someone is holding that stick, preventing it from moving too fast (adiabatically). Because we did it reversibly and no heat went in or out of the system, the entropy remains constant. However, work is being done here, unlike in the previous system. This system "talked" to the outside world, not by heat exchange, but by work exchange, so it is unlikely that the energy of that system stayed the same. Again, this system has kept the entropy the same, but it lowered its energy (in fact, we will prove in the next lecture that the energy has found a new minimum through this process).

LECTURE 19: EXTREMUM PRINCIPLES, CONTINUED

Wednesday, February 26, 2020

In the top image of ??, the process which moves the wall to a new equilibrium state is irreversible. The energy stays constant and the entropy increases, and at the new equilibrium, $\left. \frac{\partial S}{\partial x} \right|_U = 0$ and $\left. \frac{\partial^2 S}{\partial x^2} \right|_U < 0$. In the bottom image, we perform the same process adiabatically such that the entropy stays the same while the energy changes. We will show that the energy has reached a new minimum. In the top image, we can take a total derivative of the entropy to first order:

$$dS = \underbrace{\left. \frac{\partial S}{\partial U} \right|_x}_{\frac{1}{T}} dU + \underbrace{\left. \frac{\partial S}{\partial X} \right|_U}_{0} dX$$

Let's write the next-order expansion in x . Typically we don't do this, but since all the interesting stuff vanished at first-order, it might be important:

$$dS = \underbrace{\left. \frac{\partial S}{\partial U} \right|_x}_{\frac{1}{T}} dU + \underbrace{\left. \frac{\partial S}{\partial X} \right|_U}_{0} dX + \frac{1}{2} \underbrace{\left. \frac{\partial^2 S}{\partial x^2} \right|_U}_{<0} (dx)^2$$

What happens in the second case? Rewriting the previous equation in terms of dU , we see that

$$\begin{aligned} dU &= T dS + 0 \cdot dx - \frac{1}{2} T \left. \frac{\partial^2 S}{\partial x^2} \right|_U (dx)^2 \\ &= \left. \frac{\partial U}{\partial S} \right|_x dS + \left. \frac{\partial U}{\partial x} \right|_S dx + \frac{1}{2} \left. \frac{\partial^2 U}{\partial x^2} \right|_S (dx)^2 \end{aligned}$$

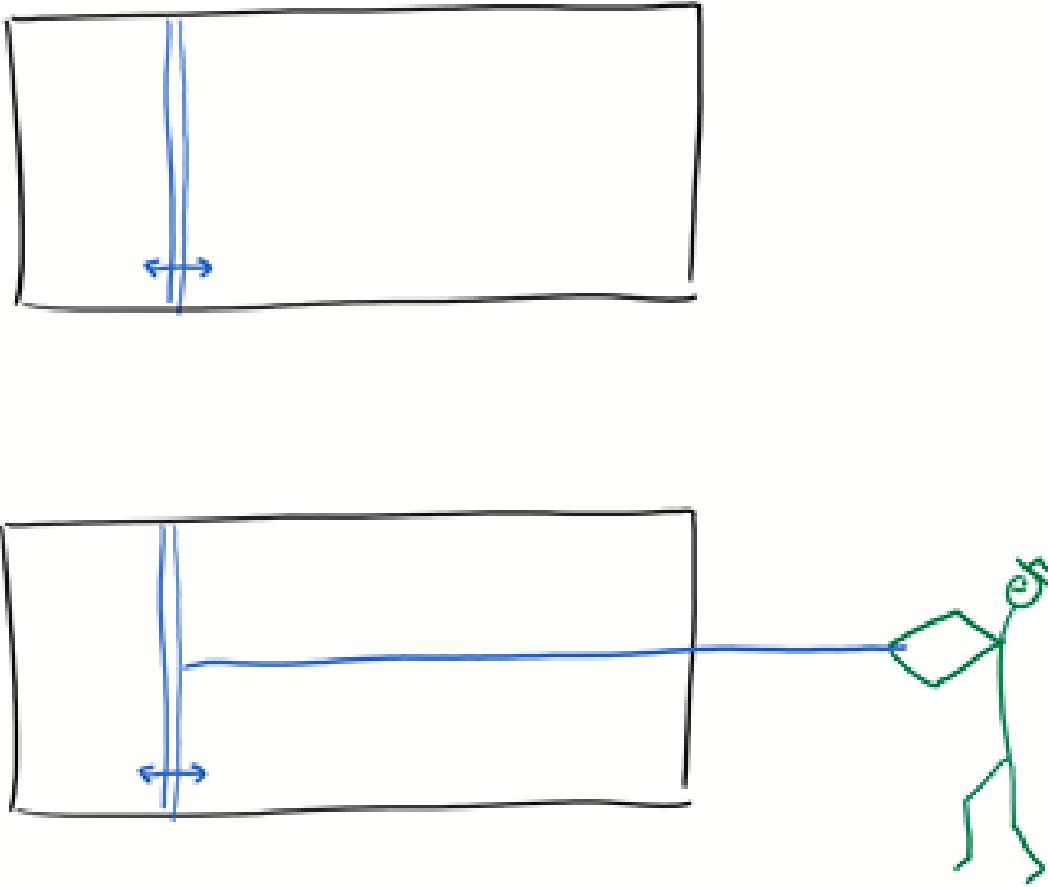


Figure 0.23.1: Top: A constraint being released, changing the entropy.
 Bottom: A constraint being released adiabatically, leaving the entropy constant.
 Credit: Logan Carpenter

where the second line is again the total derivative of U expanded to second order in x . We can identify the terms (the differentials) to show that

$$\left. \frac{\partial U}{\partial x} \right|_S = 0 \quad \left. \frac{\partial^2 U}{\partial x^2} \right|_S > 0$$

for the second scenario. *At constant entropy, the total energy is minimized in thermal equilibrium.*

What if the system is at constant T ? Imagine we put the adiabatic system in contact with a reservoir at constant temperature T_R ? This entire system, reservoir and original system combined, is equivalent to the second system. Energy must be minimized, so the principles of the second case apply to the total combined energy of both subsystems (calling the reservoir energy U_R):

$$\begin{aligned} \text{In equilibrium } & \begin{cases} \left. \frac{\partial}{\partial x} (U + U_R) \right|_S = 0 & (1) \\ \left. \frac{\partial^2}{\partial x^2} (U + U_R) \right|_S > 0 & (2) \end{cases} \\ & S + S_R = \text{const. } \forall x \quad (3) \end{aligned}$$

Since the reservoir interacts with the system only via heat exchange,

$$\begin{aligned} dU_R &= \underbrace{dW_R}_0 + dQ_R = T_R dS_R \\ \Rightarrow \frac{\partial U_R}{\partial x} &= T_R \frac{\partial S_R}{\partial x} \quad (4) \end{aligned}$$

Since our system is at constant $\{T, V, N\}$, let's consider the change in Helmholtz free energy in equilibrium

(the free energy of the system alone, not the system and reservoir combined):

$$\begin{aligned}\frac{\partial F}{\partial x} &= \frac{\partial}{\partial x}(U - TS) \\ &= \frac{\partial U}{\partial x} - T \frac{\partial S}{\partial x} \\ &= \frac{\partial U}{\partial x} - T_R \frac{\partial S}{\partial x}\end{aligned}$$

since the system is in thermal equilibrium with the reservoir (so $T = T_R$).

$$\begin{aligned}\frac{\partial F}{\partial x} &= \frac{\partial U}{\partial x} + T_R \frac{\partial S_R}{\partial x} \quad \text{by (3)} \\ &= \frac{\partial U}{\partial x} + \frac{\partial U_R}{\partial x} \quad \text{by (4)} \\ &= \frac{\partial}{\partial x}(U + U_R) = 0 \quad \text{by (1)}\end{aligned}$$

Therefore, we have reached an extremum in F at equilibrium. Is it a maximum or a minimum? We can go through all of these same steps again at a second derivative level and use equation (2) to show that

$$\frac{\partial^2 F}{\partial x^2} > 0$$

Therefore, at constant $\{T, V, N\}$, the Helmholtz free energy F is at a minimum at equilibrium. Notice the pattern here. $F(T, V, N)$ is minimized when T , V , and N are fixed. When S , V , and N are constant, $U(S, V, N)$ is minimized. The only exception is that $S(U, V, N)$ is maximized when U , V , and N are fixed. Similar extremum principles can be derived for the other thermodynamic potentials.

Recall that

$$dF = -S dT - P dV + \mu dN$$

At constant T and N ,

$$dF = -P dV = dW$$

The maximum work that can be extracted at constant T , V , and N is the change in free energy between initial and final state. Similarly, if the entropy is constant, the energy U determines the maximum work that can be extracted. Note that when we say V is constant, we are talking about the volume of the system as a whole, so $P dV$ is not zero, as this refers to the rearrangement of volumes inside the system.

For a final example, consider a system at constant T , P , and $\{N_i\}$ (multiple species of particles, although this is not important, but it's good to know how to deal with this).

$$dG = -S dT + V dP + \sum_i \mu_i dN_i$$

If the pressure is fixed, what happens? There is now work being done on the system to keep the pressure constant (in addition to heat being exchanged to keep the temperature constant). We can realize this by the system plus a T -reservoir plus a " P -reservoir" at constant energy:

$$\frac{\partial}{\partial x}(U + U_R) = 0 \quad \text{at equilibrium} \quad (1)$$

$$\frac{\partial^2}{\partial x^2}(U + U_R) > 0 \quad \text{at equilibrium} \quad (2)$$

$$S + S_R = \text{const.} \quad \forall x \quad (3)$$

$$V + V_R = \text{const.} \quad \forall x \quad (4)$$

$$dU_R = T_R dS_R - P_R dV_R \quad (5)$$

$$\begin{aligned}
\frac{dG}{dx} &= \frac{\partial}{\partial x}(U - TS + PV) \\
&= \frac{\partial}{\partial x}U - T_R \frac{\partial S}{\partial x} + P_R \frac{\partial V}{\partial x} \\
&= \frac{\partial}{\partial x}U \underbrace{T_R \frac{\partial S_R}{\partial x}}_{(3)} - \underbrace{P_R \frac{\partial V_R}{\partial x}}_{(4)} \\
&= \frac{\partial U}{\partial x} \underbrace{\frac{\partial U_R}{\partial x}}_{(5)} = \frac{\partial}{\partial x}(U + U_R) \underbrace{= 0}_{(1)}
\end{aligned}$$

Additionally, we can do the same thing to show that in this situation, $\frac{\partial^2 G}{\partial x^2} > 0$ using equation (2). We can do similar things with the grand potential and the enthalpy, which all have dimension of energy. The only odd-one-out is the entropy, which does not have energy units.

LECTURE 20: STABILITY CONDITIONS
Friday, February 28, 2020

0.24 Stability Conditions

In our previous lectures, we found properties of systems called extremum conditions. We used the maximization of entropy to show that energy is minimized if entropy is kept constant. By adding temperature and pressure reservoirs, we can translate these conditions to the other thermodynamic potentials to show that they are also minimized when those reservoirs are kept constant.

Now we want to find out what makes a system stable. In our previous example with a piston being released, we know that if the piston is in equilibrium, the total energy of the system is at a minimum. If we wiggle the piston away from that minimum, the energy must increase. Let's say that the left compartment has variables $\{U, S, V, N\}$ and the right compartment has $\{\tilde{U}, \tilde{S}, \tilde{V}, \tilde{N}\}$. We can write this variation of the wall as the following:

$$\Delta U_{\text{tot}} [U(S, V + \Delta V, N) - U(S, V, N)] + [\tilde{U}(\tilde{S}, \tilde{V} - \Delta V, \tilde{N}) - \tilde{U}(\tilde{S}, \tilde{V}, \tilde{N})]$$

Now let's take the opposite perspective. Consider a system with some $U(S, V, N)$. "Clone" it to get an identical copy. Now we bring these two systems together, and we permit an exchange of V . Since A system should be in equilibrium with itself, such a change in volume of either subsystem must increase the energy. We can write this as

$$\Delta U_{\text{tot}} [U(S, V + \Delta V, N) - U(S, V, N)] + \Delta U_{\text{tot}} [U(S, V - \Delta V, N) - U(S, V, N)] \geq 0$$

or

$$0 \leq \frac{U_{\text{tot}}}{(\Delta V)^2} = \frac{U(S, V + \Delta V, N) - 2U(S, V, N) + U(S, V - \Delta V, N)}{(\Delta V)^2} \xrightarrow[\Delta V \rightarrow 0]{} \left. \frac{\partial^2 U}{\partial V^2} \right|_{S, N}$$

The energy of a stable system must be a convex function of volume, for fixed S and N . We know that

$$\left. \frac{\partial U}{\partial V} \right|_{S, N} = -P$$

so

$$-\left. \frac{\partial P}{\partial V} \right|_{S, N} = \left. \frac{\partial^2 U}{\partial V^2} \right|_{S, N} \geq 0$$

so

$$\kappa_S = -\frac{1}{V} \left. \frac{\partial V}{\partial P} \right|_{S, N} \geq 0$$

Therefore, thermodynamic stability requires the adiabatic compressibility to be positive.

Notice, we could do the same thing with any *extensive* variable for any thermodynamic potential and show that the second derivative will probably be positive if the system is stable. Of course, we couldn't do this with something like temperature (we can't just move temperature from one compartment to the other), but let's look at a few more examples. Let's now exchange S instead of V . This means that

$$\frac{\partial^2 U}{\partial S^2} \Big|_{V,N} \geq 0$$

What does this imply? Recall that

$$T = \frac{\partial U}{\partial S} \Big|_{V,N} \implies \frac{\partial T}{\partial S} \Big|_{V,N} = \frac{\partial^2 U}{\partial S^2} \Big|_{V,N} \geq 0 \text{ therefore } C_V \geq 0$$

How about $F(T, V, N)$ with an exchange of V ? By the same argument, we know that $\frac{\partial^2 F}{\partial V^2} \Big|_{T,N} \geq 0$.

$$-P = \frac{\partial F}{\partial V} \Big|_{T,N} \implies -\frac{\partial P}{\partial V} \Big|_{T,N} = \frac{\partial^2 F}{\partial V^2} \Big|_{T,N} \geq 0 \implies \kappa_T \geq 0$$

A final example: Take the enthalpy $H(S, P, N)$ and exchange S . This implies $\frac{\partial^2 H}{\partial S^2} \Big|_{P,N} \geq 0$.

$$T = \frac{\partial H}{\partial S} \Big|_{P,N} \implies \frac{\partial T}{\partial S} \Big|_{P,N} = \frac{\partial^2 H}{\partial S^2} \Big|_{P,N} \geq 0 \implies C_P \geq 0$$

What about our intensive variables, $\{T, P, \mu\}$? These always show up after exchanging an extensive variable with its intensive partner via a Legendre transform.

$$U(S, V, N) \xrightarrow[\text{L.T.}]{\curvearrowright} F(T, V, N)$$

For the first function, we know a convexity condition with respect to S . $F(T)$ must therefore be concave! This works with every single intensive variable. The second derivative with respect to $\{T, P, \mu\}$ must be ≤ 0 .

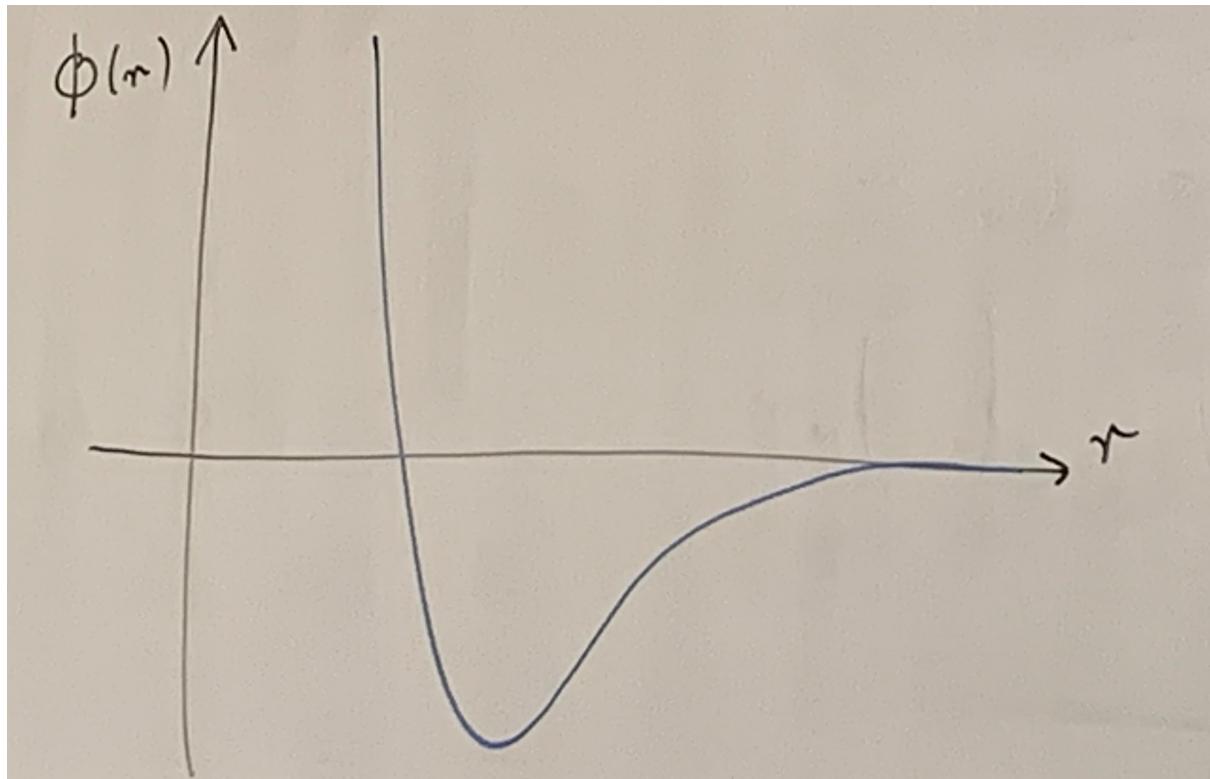
Conclusion

Since all thermodynamic potentials with dimension energy (everything but entropy) come from U , and U is convex in all its extensive variables, we have the following general result: All thermodynamic potentials with dimension energy are convex in their extensive variables and concave in their intensive variables.

In the next chapter, we will be discussing non-ideal gasses. Why might a gas be non-ideal? The answer is that the particles interact with each other. The most common case is that you have a Hamiltonian like

$$H = \underbrace{\sum_{i=1}^N \frac{\vec{p}_i^2}{2m}}_{\text{ideal}} + \sum_{i < j} \Phi(|\vec{r}_i - \vec{r}_j|)$$

Of course, in real life, there are more than just pairwise interactions. You can have triplet interactions and so on if you have more than two particles. A typical potential might look like the following figure (??):

**Figure 0.24.1:** Pairwise Interaction Potential

$$H = \sum_{i=1}^N \frac{\bar{p}_i^2}{2m} + \sum_{i < j} \Phi(|\vec{r}_i - \vec{r}_j|)$$

Typically, this potential goes to infinity as $r \rightarrow 0$, but tends to act attractively as $-\frac{1}{r^6}$ farther away. Between this attraction and repulsion, there's a minimum where particles can bind and interact. The extra repulsion should make the pressure bigger, while the extra attraction should make the pressure smaller, although the winner will probably be determined by the temperature.

We can write the entropy as

$$S(U, V, N) = k_B \ln \int \frac{d^{3N}p d^{3N}q}{h^{3N} N!} \delta(U - H(\{q, p\}))$$

Why don't we just figure out what this is? We can't actually do this integral in any nice way. However, we have an empirical way out. We can make an educated guess for how the free energy changes (in comparison to an ideal gas). For an ideal gas, we can calculate the free energy as

$$F_{\text{ideal}}(T, V, N) = -Nk_B T \left[\frac{3}{2} \ln(k_B T) + \ln\left(\frac{V}{N}\right) + X \right]$$

Van der Waals' inspired modification to this:

$$F(T, V, N) = -Nk_B T \left[\frac{3}{2} \ln(k_B T) + \ln\left(\frac{V - Nb}{N} + X\right) \right] - a \frac{N^2}{V}$$

where a and b are free parameters. Van der Waals' did this for his PhD thesis (no pressure), but why?

- When he decided to replace $V \rightarrow V - Nb$, he figured he should reduce the total accessible volume, since volume is taken up by the gas particles themselves.

- The first N in $-a\frac{N^2}{V} = N[-a\frac{N}{V}]$ keeps the equation extensive. The minus sign refers to a reduction in free energy due to an attraction between nearby particles, and that attraction is proportional to the density $\frac{N}{V}$.

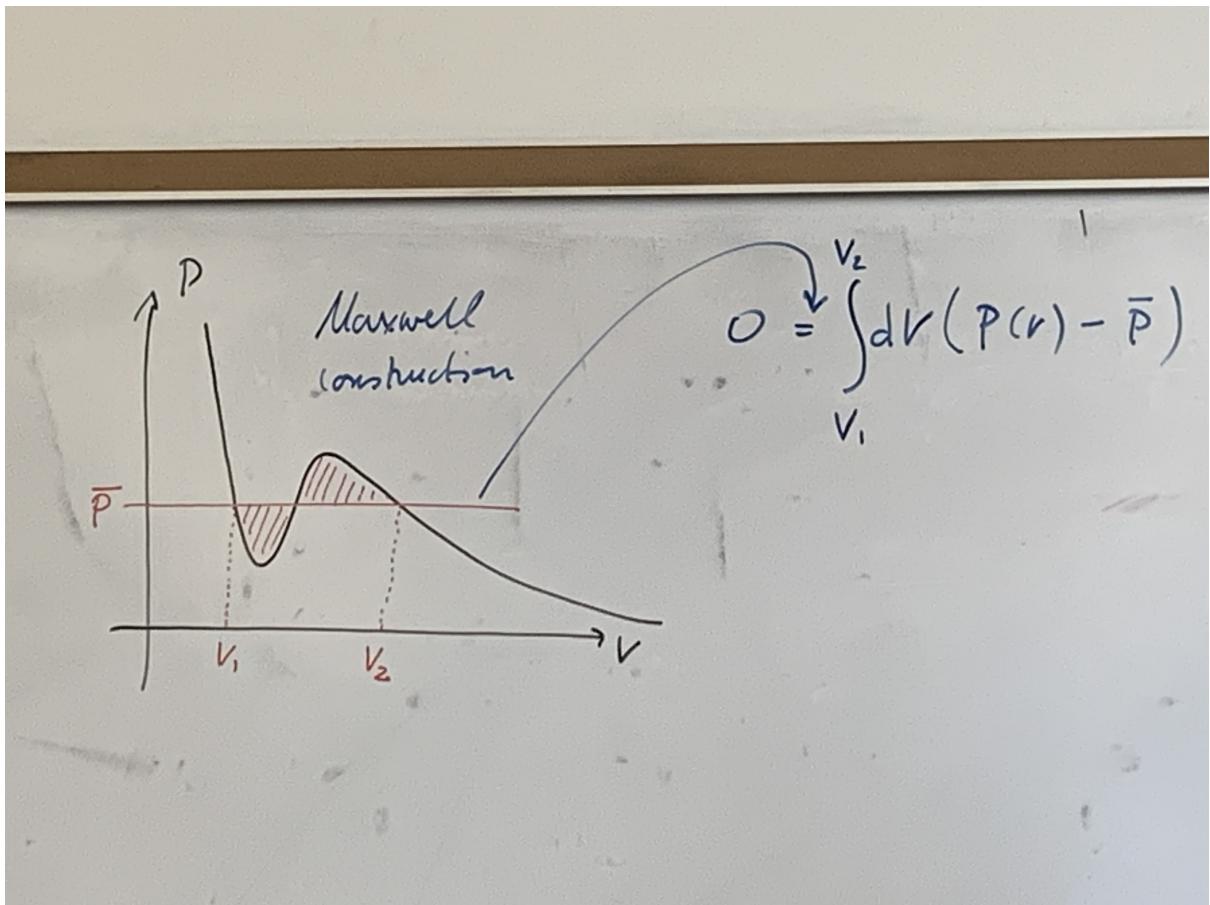


Figure 0.24.2: Diagram of the Maxwell Construction

This free energy has some interesting consequences. If we plot pressure-volume curves, we see that there is a region of temperature for which the curve is not convex (as it should be). In particular, there is a pressure \bar{P} for which the area under the curve is equal to the area above the curve (see ??). This is called a Maxwell construction. We can write this as

$$0 = \int_{V_1}^{V_2} dV (P(V) - \bar{P}) = \int_{V_1}^{V_2} dV \left(-\frac{\partial F}{\partial V} - \bar{P} \right) = -(F(V_2) - F(V_1)) - \bar{P}(V_2 - V_1)$$

so

$$\bar{P} = -\frac{F(V_2) - F(V_1)}{V_2 - V_1} = -F'(V_2) = -F'(V_1)$$

This term is the slope of the line connecting $(V_1, F(V_1))$ with $(V_2, F(V_2))$. The fact that this is equal to the slopes of the tangents of F at V_1 and V_2 is interesting. This line is called a double tangent to F (see ??).

In the next lecture, we will discuss why this is the correct construction.

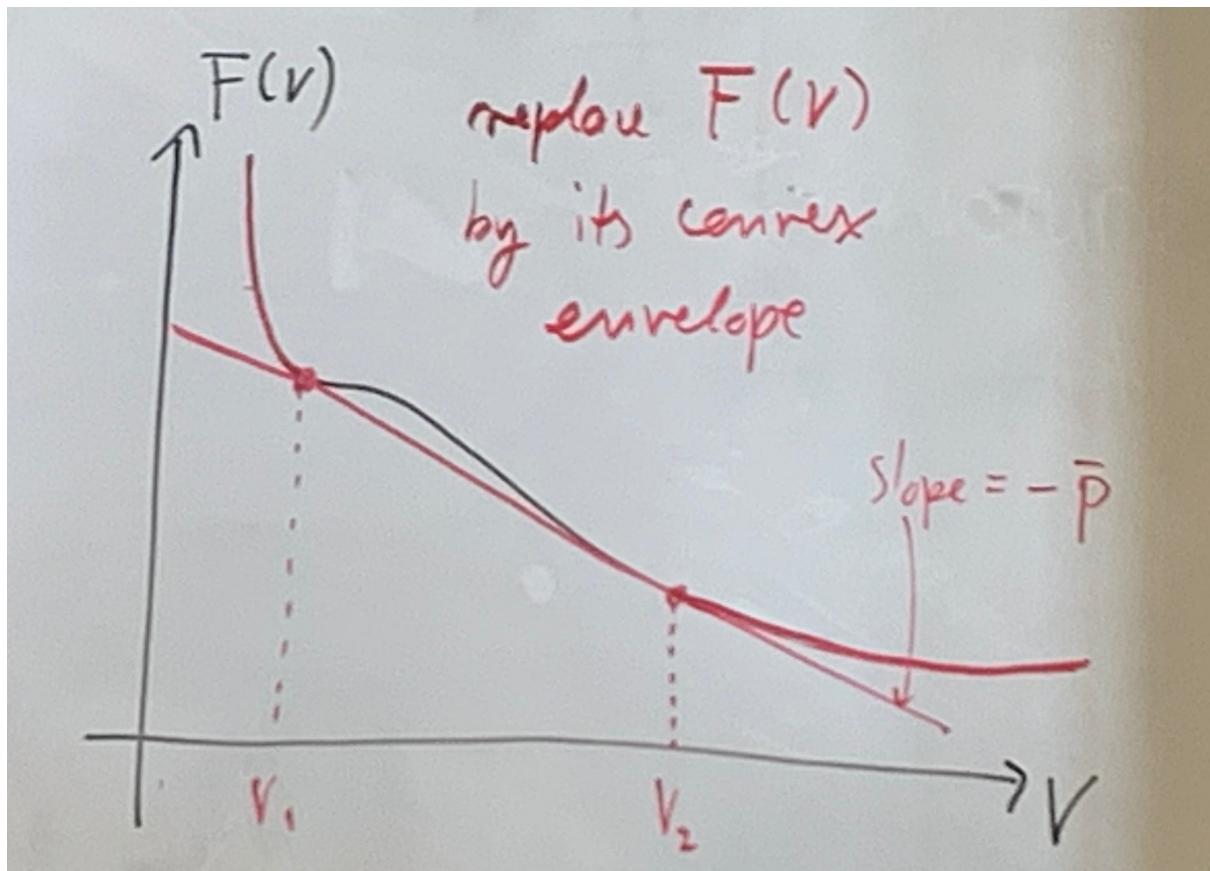


Figure 0.24.3: The double tangent line forms a convex envelope of the free energy

0.25 Phase Transitions

Instead of writing down the entropy from first principles, we made a guess as to how non-ideal gasses might behave to modify the free energy. We found that below a “critical” temperature, the free energy was not necessarily convex, making the pressure-volume curve no longer monotonic. We found a way to cut this non-monotonic part into sections of equal area, and it turns out that moving along this line corresponds to moving along the convex envelope of the free energy. In the Gibbs free energy, this critical region corresponds to a “bow-tie” like the one we saw in a homework problem, and the Maxwell construction cuts the tie. Now we want to know why the convex envelope is correct in the first place.

Recall in if we are in free energy variables and we pick T , V , and N , we can find the corresponding free energy. If the volume is in the critical region, there is a way to further minimize the energy by splitting into two volumes, one in each dip. Once you do that, the total free energy lies on the convex envelope line. The system found an opportunity to lower its free energy by phase separating.

For any volume in that critical range $V \in [V_1, V_2]$, the system can lower its free energy by splitting into a fraction α at the state point V_2 and a fraction $1 - \alpha$ at V_1 . If it does so, the free energy after splitting is

$$F_{\text{split}} = \alpha F(V_2) + (1 - \alpha)F(V_1)$$

We must also conserve the original volume, so $V = \alpha V_2 + (1 - \alpha)V_1$. We can solve this for alpha:

$$\alpha = \frac{V - V_1}{V_2 - V_1}$$

Therefore,

$$F_{\text{split}} = \frac{V - V_1}{V_2 - V_1} F(V_2) + \frac{V_2 - V}{V_2 - V_1} F(V_1)$$

At the endpoints, one of the terms vanishes, and between the end points, this is a linear combination of the endpoint values. Therefore, this phase separation energy is the line connecting the critical points.

This is how phase transitions happen. In the critical region, weird things happen. The compressibility becomes negative, and this is a violation of the stability criteria we established a few lectures ago. This is what allows the system to no longer stay homogeneous.

This construction now allows us to construct phase diagrams (??):

0.25.1 Latent Heat

There is a finite amount of heat which goes into or comes out of the system when it experiences a 1st order phase transition:

$$\Delta Q = T \Delta S$$

Considering constant T and N ,

$$dS = \left. \frac{\partial S}{\partial V} \right|_{T,N} dV = \left. \frac{\partial P}{\partial T} \right|_{V,N} dV$$

Therefore

$$\Delta S = \int_{V_1}^{V_2} dV \left. \frac{\partial S}{\partial V} \right|_{T,N} = \int_{V_1}^{V_2} dV \left. \frac{\partial P}{\partial T} \right|_{V,N} = \left. \frac{\partial P}{\partial T} \right|_{V,N} \int_{V_1}^{V_2} dV = \left. \frac{\partial P}{\partial T} \right|_{V,N} \Delta V$$

We can therefore write

$$\frac{\Delta S}{\Delta V} = \left. \frac{\partial P}{\partial T} \right|_{\text{phase boundary}} \quad (\text{Clausius Clapeyron Equation})$$

Example. Let's look at an example using the phase diagram of water (see ??). On the phase boundary between gas and liquid, the slope is positive, so $\frac{\Delta S}{\Delta V} > 0$. This implies that evaporation increases the entropy. However, if we look at the liquid-ice boundary, the slope is negative, which tells us that freezing water lowers the entropy (ice is an ordered crystal). This also implies that when ice freezes, it expands.

◇

LECTURE 23: THE GIBBS PHASE RULE

Wednesday, March 18, 2020

In the last lecture, we talked about phase separation as a direct result of the Van der Waal's equation. If we make plots of the extensive or intensive variables (pressure, temperature, and volume for example), we see that the Maxwell construction gives us a "cliff" which represents a first-order phase transition. We then used the Clausius Clapeyron Equation to find relations between the entropy and temperature depending on certain slopes.

0.26 Gibbs Phase Rule

What if you have a system with multiple phases coexisting (or multiple chemical species coexisting, like ice cubes in alcohol)? Consider a system with K components and φ phases. By components, we mean things like water, alcohol, etc. In any phase j , the number fraction of component k is

$$x_k^{(j)} = \frac{N_k^{(j)}}{\sum_{k=1}^K N_k^{(j)}}$$

Obviously,

$$\sum_{k=1}^K x_k^{(j)} = 1$$

This means that in every phase, there are $K - 1$ independent concentrations (not K , because they have to add up to 1). In φ phases, this creates $\varphi(K - 1)$ tunable concentrations (the number of concentrations you can freely pick). However, this is not all that is going on. These phases are in contact, so of course, they can exchange molecules. Because of this, the chemical equilibrium requires that the φ coexisting phases must have the same chemical potential:

$$\mu_k^{(1)} = \mu_k^{(2)} = \cdots = \mu_k^{(\varphi)}$$

There are $\varphi - 1$ equalities in the equation above. For every component, there are $\varphi - 1$ conditions that must hold, and for all K components, this creates $K(\varphi - 1)$ constraints. We also have $\varphi(K - 1)$ choices of concentrations, so what are the total number of degrees of freedom?

$$f = \underbrace{\frac{2}{T} \text{ and } P}_{\text{tunable concentrations}} + \underbrace{\varphi(K - 1)}_{\text{constraints}} - \underbrace{K(\varphi - 1)}_{\text{constraints}} = 2 + K - \varphi \quad (\text{Gibbs Phase Rule})$$

For example, pure water ($K = 1$), fluid-ice coexistence would have two phases ($\varphi = 2$), so the number of degrees of freedom is $f = 2 + 1 - 2 = 1$. Therefore, the phase boundary is a 1D curve in the $T - P$ diagram. In the case of fluid-ice-vapor coexistence, we have $\varphi = 3$, so $f = 2 + 1 - 3 = 0$. There are no degrees of freedom left. This corresponds to the “triple point”, and explains why the triple point is uniquely defined.

0.27 The Third Law of Thermodynamics

Nernst: “For $T \rightarrow 0$, the entropy of the system must go to a constant which does not depend on the pressure.”

Planck: “For $T \rightarrow 0$, the entropy goes to 0.”

The classical ideal gas violates the third law. This is because the origin of the third law is in quantum mechanics. We will discuss this in more detail in a later lecture. For now, what are the consequences of the third law?

1. Heat capacities go to zero as $T \rightarrow 0$.

Proof: $c_x(T) = \frac{T}{N} \frac{\partial S}{\partial T} \Big|_{x,N}$ implies $dS(T, x, N) = \frac{Nc_x(T)}{T} dT$, so $\Delta S = N \int_{T_1}^{T_2} dT \frac{c_x(T)}{T}$. If this is the case, assume we pick two temperatures $0 < T_1 < T_2$ which are “close” to 0 in the sense that if $c_x(T)$ were to go to a constant $c_x(0) = c_{x,0} > 0$ for $T \rightarrow 0$, then for $0 < T < T_2$, we can safely approximate $c_x(T)$ by $c_{x,0}$. Basically we’re saying that we can expand $c_x(T)$ around T_2 and take, to good approximation, the first value of the expansion. If we do this,

$$\Delta S \approx N \int_{T_1}^{T_2} dT \frac{c_{x,0}}{T} = Nc_{x,0} \ln\left(\frac{T_2}{T_1}\right)$$

However, this goes to ∞ as $T_1 \rightarrow 0$. Hence, $S(T_1)$ would go to $-\infty$ as $T_1 \rightarrow 0$, in violation of the third law. Hence, $c_x(T)$ must go to zero as $T \rightarrow 0$.

2. $\alpha(T) = \frac{1}{V} \frac{\partial V}{\partial T} \Big|_{P,N} = -\frac{1}{V} \frac{\partial S}{\partial P} \Big|_{T,N}$. If $S \rightarrow$ a constant that does not depend on pressure as $T \rightarrow 0$, then $\alpha(T) \rightarrow 0$ as $T \rightarrow 0$, since as $T \rightarrow 0$, S no longer depends on P .
3. Recall we once derived the relation $c_p - c_v = \frac{\alpha^2 TV}{N\kappa_T}$. Let’s look at

$$\frac{c_p - c_v}{T} = \frac{\alpha^2 V}{N\kappa_T} = -\underbrace{\frac{1}{N} \frac{\partial V}{\partial T}}_{\rightarrow 0} \Big|_P \underbrace{\frac{\partial P}{\partial V}}_{\rightarrow 0} \Big|_T = -\underbrace{\frac{1}{N} \frac{\partial S}{\partial P}}_{\rightarrow 0} \Big|_T \underbrace{\frac{\partial S}{\partial V}}_{\rightarrow 0} \Big|_T \rightarrow 0 \text{ as } T \rightarrow 0$$

Both the numerator and denominator go to zero here, so not only do c_V and c_P both individually go to zero, their difference vanishes *more strongly* than linearly in T .

 LECTURE 24: CLASSICAL STATISTICAL PHYSICS
 Friday, March 20, 2020

Recall that knowing $S(E, V, N)$ tells us basically everything about the system. Once you know (any) thermodynamic potential, “you win”. Notice we are back to using E for energy. Remember that we had a way to compute this:

$$S(E, V, N) = k_B \ln \left[\int \frac{d^{3N}p d^{3N}q}{h^{3N} N!} \delta(E - H(p, q)) \right]$$

This tells us everything about the system, but in general this integral is impossible (we were only able to do it for a very simple case, the ideal gas). Even if we wanted to approximate it using a Taylor expansion, we would get very frustrated with the δ -function.

Let’s go back to a canonical distribution of energy in a system connected to a reservoir separated by a diathermal wall. We can see that the total energy of the “universe” is just the sum of the system and the reservoir,

$$E_T = E + E_R$$

For such a system, we found that

$$P(E) = \frac{\Omega(E)\Omega_R(E_R)}{\Omega_T(E_T)}$$

Since we go by our maxim that the system goes to the state of maximal probability, we can just maximize the logarithm:

$$\ln P(E) = \ln \Omega(E) + \ln \Omega_R(E_T - E) - \ln \Omega_T(E_T)$$

Since the reservoir is large, we have $E \ll E_T$, so we can Taylor expand around it:

$$\ln P(E) = \ln \Omega(E) + \left[\ln \Omega_R(E_T) - \underbrace{\frac{\partial \ln \Omega_R(E_T)}{\partial E_T}}_{\frac{1}{k_B T_R} = \beta_R = \beta = \frac{1}{k_B T}} E + \mathcal{O}(E^2) \right] - \ln \Omega_T(E_T)$$

Notice that $\ln \Omega_R(E_T)$ and $\ln \Omega_T(E_T)$ don’t depend on E . We’ll put them into a new constant, $\ln Z$:

$$\ln P(E) = \ln \Omega(E) - \beta E - \ln Z$$

Now we can re-exponentiate to find

$$P(E) = \frac{1}{Z} \Omega(E) e^{-\beta E}$$

The one thing we don’t know in there is Z , but notice that $P(E)$ is a probability density, so it must be normalized. Therefore,

$$Z(T, V, N) = \int dE \Omega(E, V, N) e^{-\beta E}$$

If we want to give the proper terminology, $Z(T)$ is the “Laplace transform” of $\Omega(E)$.

Z is the “partition function”. It’s called Z after the German *Zustandssumme*, meaning “sum of states”.

Let’s do this derivation again, but in phase space. The probability of being in one particular microstate is

$$P(p, q) = \frac{\Omega_R(E_T - H(p, q))}{\Omega_T(E_T)}$$

$$\ln P(p, q) = \ln \Omega_R(E_T - H(p, q)) - \ln \Omega_T(E_T)$$

Again, if we recognize $H \ll E_T$, we can Taylor expand:

$$\ln P(p, q) = \ln \Omega_R(E_T) - \beta H(p, q) - \ln \Omega_T(E_T) + \dots = -\beta H(p, q) - \ln \tilde{Z}$$

Therefore,

$$P(p, q) = \frac{1}{\tilde{Z}} e^{-\beta H(p, q)}$$

and

$$\tilde{Z}(T, V, N) = \int d^{3N}p d^{3N}q e^{-\beta H(p, q)}$$

Notice this looks very similar to what we did before, but without any Ω functions. Notice also that there is a very obvious sum of states in \tilde{Z} in the form of an integral over all p and q . Asking the probability of finding a state with energy E is a different question than asking the probability of finding a particular state, since states can be degenerate.

Let's now link these two expressions.

$$\begin{aligned} Z(T, V, N) &= \int dE \underbrace{\Omega(E, V, N)}_{\int \frac{d^{3N}p d^{3N}q}{h^{3N} N!} \delta(E - H(p, q))} e^{-\beta E} \\ &= \int \frac{d^{3N}p d^{3N}q}{h^{3N} N!} \int dE \delta(E - H(p, q)) e^{-\beta E} \\ &= \frac{1}{h^{3N} N!} \int d^{3N}p d^{3N}q e^{-\beta H(p, q)} \\ &= \frac{1}{h^{3N} N!} \tilde{Z}(T, V, N) \end{aligned}$$

Therefore, we could write

$$P(p, q) = \frac{1}{Z h^{3N} N!} e^{-\beta H(p, q)}$$

and

$$Z(T, V, N) = \int \frac{d^{3N}p d^{3N}q}{h^{3N} N!} e^{-\beta H(p, q)}$$

This probability density on phase space is called the “canonical state”. These are probably the most important equations in classical statistical physics.

LECTURE 25: THE CANONICAL STATE

Monday, March 23, 2020

0.28 The Canonical State

Is the canonical distribution stable under the time-evolution of Hamiltonian dynamics? We can answer yes because of Liouville’s Theorem.

$$\begin{aligned} \frac{\partial}{\partial t} P(p, q, t) &= - \sum_{j=1}^{3N} \left(\underbrace{\frac{\partial P}{\partial q_j} \dot{q}_j}_{\frac{\partial H}{\partial p_j}} + \underbrace{\frac{\partial P}{\partial p_j} \dot{p}_j}_{-\frac{\partial H}{\partial q_j}} \right) \\ &= - \sum_{j=1}^{3N} \left(\frac{\partial P}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial P}{\partial p_j} \frac{\partial H}{\partial q_j} \right) \\ &= -\{P, H\} \end{aligned}$$

This is a nontrivial statement which follows from Liouville's theorem. We will not prove it. An important property of the Poisson bracket in the last line is that

$$\{A, f(A)\} = 0$$

and the canonical state is just a function of H , so

$$\frac{\partial P_{\text{can}}}{\partial t} = -\{P_{\text{can}}, H\} \propto -\{e^{-\beta H}, H\} = 0$$

Therefore, the canonical state does not change under Hamiltonian dynamics.

0.28.1 Energy Fluctuations

Recall $P(E) = \frac{1}{Z}\Omega(E)e^{-\beta E}$ which increases strongly with E . Typically $\Omega(E) \sim E^f$ with $f \sim N$. Where does $P(E)$ have its maximum?

$$0 = \frac{\partial}{\partial E} \ln P(E) = \frac{\partial}{\partial E} [-\ln Z + \ln \Omega(E) - \beta E] = \frac{f}{E} - \beta$$

so

$$E_{\max} = \frac{f}{\beta} = fk_B T$$

Recall for the ideal gas, $f = \frac{3}{2}N$.

Next, how wide is the peak?

$$-\frac{1}{\sigma_E^2} = \left. \frac{\partial^2 \ln P}{\partial E^2} \right|_{E \rightarrow E_{\max}} = -\frac{f}{E_{\max}^2} = -\frac{1}{f(k_B T)^2}$$

so

$$\sigma_E = \sqrt{fk_B T}$$

so the coefficient of variation is

$$\frac{\sigma_E}{E_{\max}} = \frac{\sqrt{fk_B T}}{fk_B T} = \frac{1}{\sqrt{f}} \sim \frac{1}{\sqrt{N}}$$

so again, the energy fluctuations scale such that at large N , they are small compared to the overall energy of the state.

Finally, let's link this to the Helmholtz free energy.

$$\begin{aligned} \ln P(E) &= -\beta E + \ln \Omega(E) - \ln Z \\ \ln Z &= -\beta E + \ln \Omega(E) - \ln P(E) \\ &= -\beta \underbrace{\left(E - TS \right)}_{\text{Scales with } N} - \overbrace{\ln \underbrace{P(E)}_{\substack{\text{scales with } \ln E \sim \ln N \\ \text{width } \sim \sqrt{E}, \text{ height } \sim \frac{1}{\sqrt{E}}}}}^{\text{width } \sim \sqrt{E}, \text{ height } \sim \frac{1}{\sqrt{E}}} \end{aligned}$$

Therefore, for large N ,

$$-k_B T \ln Z(T, V, N) = E - TS$$

The right-hand side would be the Helmholtz free energy if it was expressed in T , V , and N , and these are exactly the variables of $Z(T, V, N)$, so

$$F(T, V, N) = -k_B T \ln Z(T, V, N)$$

This is another super important equation in statistical mechanics. Alternatively we could write

$$e^{-\beta F} = Z = \int dE \Omega(E) e^{-\beta E}$$

We can actually do even better than this, but to do it, we need a small excursion:

Excuse: Saddle Point Evaluation of Integrals

Here's a fun trick to approximate integrals. Suppose we have a function $f(x)$ that has a single maximum, and perhaps around that maximum we can Taylor expand into a parabola at x_m . Now suppose we want to integrate

$$I_N := \int dx e^{Nf(x)}$$

approximately for large N . As long as $f(x)$ has a single peak, we can Taylor expand $f(x)$ around the maximum and replace $f(x)$ in the integral by that Taylor expansion. Naively, the function can diverge from that parabola arbitrarily far away from the point of expansion, but it turns out this doesn't matter:

$$\begin{aligned} I_N &\approx \int dx e^{N \left[\underbrace{f(x_m)}_{\text{constant}} - \frac{1}{2} |f''(x_m)| (x - x_m)^2 + \dots \right]} \\ &= e^{Nf(x_m)} \sqrt{\frac{2\pi}{N|f''(x_m)|}} \cdot \dots \\ \ln I_n &= Nf(x_m) + \frac{1}{2} \ln \frac{2\pi}{N|f''(x_m)|} + \dots \end{aligned}$$

and

$$\lim_{N \rightarrow \infty} \left(\frac{1}{N} \ln I_N \right) = f(x_m) = \max_x f(x)$$

We don't actually have to do any integral at all, we just need to find the maximum!

Now back to the main problem,

$$\begin{aligned} e^{-\beta F} &= Z = \int dE \Omega(E) e^{-\beta E} \\ &= \int dE e^{S(E)/k_B} e^{-\beta E} \\ &= \int dE e^{-\beta(E - TS(E))} \\ &= N \int de e^{N(-\beta(e - Ts(e)))} \end{aligned}$$

What is

$$-\beta f = \lim_{N \rightarrow \infty} \left[\frac{1}{N} \ln e^{-\beta F} \right]?$$

where f is the specific free energy F/N .

From our saddle point evaluation, we can see that we just need the to find

$$-\beta f = \max_e \{-\beta(e - Ts(e))\} = -\beta \min_e \{e - Ts(e)\}$$

or

$$f = \min_e \{e - Ts(e)\} = \min_s \{e(s) - Ts\}$$

The Legendre transform between the thermodynamic potentials $s(e)$ and $f(T)$ arises naturally as the saddle point approximation linking the partition functions $\Omega(E)$ and $Z(T)$! However, this is only valid for infinitely large systems, and this simple connection will not in general be true for finite systems, particularly computer simulations.

$$\begin{aligned}
 \frac{\partial}{\partial \beta} \ln Z &= \frac{1}{Z} \frac{\partial Z}{\partial \beta} = \frac{1}{Z} \frac{\partial}{\partial \beta} \int \frac{d^{3N} p d^{3N} q}{h^{3N} N!} e^{-\beta H(p,q)} \\
 &= \int d^{3N} p d^{3N} q (-H(p,q)) \underbrace{\frac{e^{-\beta H(p,q)}}{Z h^{3N} N!}}_{\text{canonical state}} \\
 &= -\langle H(p,q) \rangle_{\text{CS}}
 \end{aligned}$$

so

$$E \equiv \langle H \rangle = -\frac{\partial}{\partial \beta} \ln Z = \frac{\partial}{\partial \beta} (\beta F)$$

Can we learn anything new aside from simply reproducing the thermodynamics we already know? Yes, but we need to go to higher derivatives. Recall that

$$c_V = \frac{T}{N} \left. \frac{\partial S}{\partial T} \right|_{V,N} = \frac{1}{N} \left. \frac{\partial E}{\partial T} \right|_{V,N}$$

Notice that $T dS = E$ in this scenario, so

$$\begin{aligned}
 c_V &= \frac{1}{N} \frac{\partial}{\partial T} \langle H \rangle = \frac{1}{N} \frac{\partial \beta}{\partial T} \frac{\partial}{\partial \beta} \langle H \rangle = -\frac{\beta}{NT} \frac{\partial}{\partial \beta} \langle H \rangle \\
 &= -\frac{\beta}{TN} \frac{\partial}{\partial \beta} \frac{\int \frac{d^{3N} p d^{3N} q}{h^{3N} N!} H(p,q) e^{-\beta H(p,q)}}{\int \frac{d^{3N} p d^{3N} q}{h^{3N} N!} e^{-\beta H(p,q)}}
 \end{aligned}$$

To simplify this notation, let us introduce the following:

$$\text{Tr}(\cdot) := \int \frac{d^{3N} p d^{3N} q}{h^{3N} N!} (\cdot)$$

Using this notation,

$$\begin{aligned}
 c_V &= -\frac{\beta}{TN} \frac{\partial}{\partial \beta} \frac{\text{Tr}(He^{-\beta H})}{\text{Tr}(e^{-\beta H})} \\
 &= -\frac{\beta}{TN} \frac{\text{Tr}(-H^2 e^{-\beta H}) \text{Tr}(e^{-\beta H}) - \text{Tr}(He^{-\beta H}) \text{Tr}(-He^{-\beta H})}{[\text{Tr}(e^{-\beta H})]^2} \\
 &= -\frac{\beta}{TN} \left[\frac{\text{Tr}(-H^2 e^{-\beta H})}{\text{Tr}(e^{-\beta H})} + \left(\frac{\text{Tr}(He^{-\beta H})}{\text{Tr}(e^{-\beta H})} \right)^2 \right] \\
 &= \frac{\beta}{TN} [\langle H^2 \rangle - \langle H \rangle^2] \\
 &= \frac{\beta}{TN} \sigma_H^2
 \end{aligned}$$

It turns out that this heat capacity can be written as a result of fluctuations of the Hamiltonian!

$$\underbrace{\sigma_H^2}_{\text{energy fluctuations}} = \underbrace{N}_{\text{extensive}} \underbrace{(k_B T)}_{\text{characteristic energy}} \underbrace{\frac{c_V}{k_B}}_{\text{response function}}$$

This is an example of a “fluctuation-response theorem”. It is generally beyond thermodynamics because it talks about fluctuations—something we talk about in statistical physics.

Notice that $\sigma_H^2 \sim N$, as expected from our last lecture. Can response functions be infinite? It certainly seems mathematically possible. Imagine an isotherm through a critical point of a phase transition. At

the critical point, $\frac{\partial P}{\partial V}|_T = 0$, so $\kappa_T = \infty$ (see ??). In fact, most response functions (compressibility, specific heat, etc.) tend to diverge at critical points. However, this would seem to suggest that $\sigma_H^2 = \infty$ at these points, so how is this even possible? We happen to know that probability densities with infinite variance exist (Cauchy-Lorentz) and they have some curious properties. At the critical point, the system shows fluctuations that are not Gaussian (because the Central Limit Theorem no longer holds)! This is the beginning of some awesome physics which unfortunately are beyond the scope of this class.

LECTURE 27:
Friday, March 27, 2020

Take a simple case where

$$H(p, q) = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} + U(\{q\})$$

We can therefore write

$$\begin{aligned} Z &= \int \frac{d^{3N}p d^{3N}q}{h^{3N} N!} e^{\beta \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} + U(\{q\})} \\ &= \frac{1}{h^{3N} N!} \int d^{3N}p e^{-\beta \sum_{i=1}^N \frac{\vec{p}_i^2}{2m}} \int d^{3N}q e^{-\beta U(\{q\})} \\ &= \frac{1}{h^{3N} N!} \left[\underbrace{\int d^3p_1 e^{-\beta \frac{\vec{p}_1^2}{2m}}}_{\left(\frac{2\pi m}{\beta}\right)^{3/2}} \right]^N \underbrace{\int d^{3N}q e^{-\beta U(\{q\})}}_{=:Q} \\ &= \frac{(2\pi m k_B T)^{3N/2}}{h^{3N} N!} Q \\ &= \left(\underbrace{\frac{\sqrt{2\pi m k_B T}}{h}}_{\frac{1}{\lambda_{\text{th}}}} \right)^{3N} \frac{1}{N!} Q \\ &= \frac{Q}{\lambda_{\text{th}}^{3N} N!} \end{aligned}$$

where Q is often called the configurational partition function (since it deals with the positions of the particles) and λ_{th} is the thermal de Broglie wavelength. The difficult part now is calculating Q . Let's do this for the ideal gas. In this case, $U(\{q\}) = 0$ so $e^{-\beta U(\{q\})} = 1$. Therefore,

$$Q = \int d^{3N}q e^{-\beta U(\{q\})} = \int d^{3N}q = V^N$$

Therefore, for the ideal gas,

$$Z_{\text{IG}} = \left(\frac{V}{\lambda_{\text{th}}^3} \right)^N \frac{1}{N!}$$

Do not lose track of the fact that the temperature dependence is in λ_{th} . We can then calculate the free energy of the ideal gas:

$$\begin{aligned}
F &= -k_B T \ln(Z) = -k_B T \ln \left[\left(\frac{V}{\lambda_{\text{th}}^3} \right)^N \frac{1}{N!} \right] \\
&= -k_B T \left[N \ln \frac{V}{\lambda_{\text{th}}^3} - \ln(N!) \right] \\
&\approx -k_B T \left[N \ln \frac{V}{\lambda_{\text{th}}^3} - N \ln(N) + N \right] \\
&\approx -N k_B T \left[\ln \frac{V}{N \lambda_{\text{th}}^3} + 1 \right] \\
&\approx N k_B T [\ln(n \lambda_{\text{th}}^3) - 1]
\end{aligned}$$

where $n = \frac{N}{V}$ is the particle number density. Notice that the argument of the logarithm is finally dimensionless.

0.28.2 Factorization

Let's consider another Hamiltonian:

$$H(p, q) = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} + \sum_{i=1}^N \tilde{\phi}_i(\vec{q}_i)$$

Notice that the potential and kinetic energy decompose additively into sums of terms, each of which features only one particle. Therefore

$$\begin{aligned}
Z &= \frac{1}{\lambda_{\text{th}}^{3N} N!} \int d^{3N} q \prod_{i=1}^N e^{-\beta \tilde{\phi}_i(\vec{q}_i)} \\
&= \frac{1}{\lambda_{\text{th}}^{3N} N!} \prod_{i=1}^N \int d^3 q_i e^{-\beta \tilde{\phi}_i(\vec{q}_i)} \\
&= \frac{\prod_{i=1}^N Q_i}{\lambda_{\text{th}}^{3N} N!} \\
&= \left(\frac{Q_i}{\lambda_{\text{th}}^3} \right)^N \frac{1}{N!}
\end{aligned}$$

Example. Take for example a single harmonic oscillator in one dimension:

$$H_1 = \frac{p^2}{2m} + \frac{1}{2} k x^2$$

$$\begin{aligned}
Z &= \int \frac{dp dx}{h^4 1!} e^{-\beta \left(\frac{p^2}{2m} + \frac{1}{2} k x^2 \right)} = \frac{1}{h} \sqrt{2\pi m k_B T} \sqrt{\frac{2\pi k_B T}{k}} \\
&= \frac{2\pi k_B T}{h} \sqrt{\frac{m}{k}} = \frac{k_B T}{\hbar \omega}
\end{aligned}$$

Now for N harmonic oscillators,

$$H_N = \sum_{i=1}^N \left(\frac{p_i^2}{2m} + \frac{1}{2} k_i x_i^2 \right)$$

$$Z = \prod_{i=1}^N Z_i = \prod_{i=1}^N \frac{k_B T}{\hbar \omega_i}$$

If all $\omega_i = \omega$,

$$Z = \left(\frac{k_B T}{\hbar \omega} \right)^N$$

What happened to $N!$? Often, when particles are tethered to springs, they become distinguishable via the spring to which they belong. In this case, they cannot be confused, so we lose the degeneracy which is accounted for by $N!$. \diamond

LECTURE 28: THE GRAND CANONICAL ENSEMBLE

Monday, March 30, 2020

Take a system at constant $\{T, V, \mu\}$. Recall that

$$\Omega(T, V, \mu) = U - TS - \mu N$$

Note that this is equal to $-PV$ if the system is extensive. Our microstate still contains N particles and six coordinates, so we still have a $6N$ -dimensional phase space, but now N can vary. How do we deal with this? We can separate phase space into a collection of simple phase spaces at fixed N (??).

Our probability distribution can live anywhere on any of these planes, possibly in multiple places at the same time. To maintain our constant terms, we can imagine the system is in contact with a reservoir.

$$P(E, N) = \frac{\Omega(E, V, N)\Omega_R(E_T - E, V_R, N_T - N)}{\Omega_T(E_T, V_T, N_T)}$$

Here, let's say $E \ll E_T$ and $N \ll N_T$ and take the logarithm as we usually do:

$$\begin{aligned} \ln(P(E, N)) &= \ln(\Omega) + \ln(\Omega_R) - \ln(\Omega_T) \\ &= \ln(\Omega) + \left[\ln(\Omega_R(E_T, V_R, N_T)) - E \underbrace{\frac{\partial \ln(\Omega_R)}{\partial E_T}}_{\beta} - N \underbrace{\frac{\partial \ln(\Omega_R)}{\partial N_T}}_{-\beta\mu} + \dots \right] - \ln(\Omega_T) \\ &= \ln(\Omega) - \beta(E - \mu N) - \ln(\mathcal{Z}) \end{aligned}$$

so we can write

$$P(E, N) = \frac{1}{\mathcal{Z}} \Omega(E, V, N) e^{-\beta(E - \mu N)}$$

where

$$\begin{aligned} \mathcal{Z}(T, V, \mu) &= \sum_{N=0}^{\infty} \int dE \Omega(E, V, N) e^{-\beta(E - \mu N)} \\ &= \sum_{N=0}^{\infty} Z(T, V, N) e^{\beta \mu N} \end{aligned}$$

$$\ln(\mathcal{Z}) = \underbrace{\frac{S}{k_B}}_{\sim N} - \beta(E - \mu N) - \underbrace{\ln(P(E, N))}_{\ln(N)}$$

We will neglect the part that scales like $\ln(N)$, so

$$-k_B T \ln(\mathcal{Z}(T, V, \mu)) = E - TS - \mu N = \Omega(T, V, \mu)$$

where $\Omega(T, V, \mu)$ is the grand potential (from back when we did transforms into things like enthalpy and the Gibbs free energy).

Example. Let's do an example with the ideal gas. Recall that we previously found

$$Z = \left(\frac{V}{\lambda_{\text{th}}^3} \right)^N \frac{1}{N!}$$

and

$$\begin{aligned} \mathcal{Z} &= \sum_{N=0}^{\infty} Z_N e^{\beta \mu N} = \sum_{N=0}^{\infty} \frac{1}{N!} \left(\frac{V}{\lambda_{\text{th}}^3} \right)^N e^{\beta \mu N} \\ &= \exp \left\{ \frac{Ve^{\beta \mu}}{\lambda_{\text{th}}^3} \right\} \\ -PV &= \Omega(T, V, \mu) = -k_B T \ln \left(\exp \left\{ \frac{Ve^{\beta \mu}}{\lambda_{\text{th}}^3} \right\} \right) = -k_B T \frac{V}{\lambda_{\text{th}}^3} e^{\beta \mu} \end{aligned}$$

so

$$P_{\text{IG}}(T, \mu) = k_B T \frac{e^{\beta \mu}}{\lambda_{\text{th}}^3}$$

or

$$\mu(T, P) = k_B T \ln \left(\frac{P \lambda_{\text{th}}^3}{k_B T} \right)$$

◇

0.29 Quantum Statistical Physics

Instead of looking at chapter 22, we will instead make some remarks on quantum mechanics within the context of statistical physics. We would like to first compare classical and quantum mechanics. For simplicity, let's only look at one Cartesian degree of freedom in a non-relativistic system with no spin.

Concept	CM	QM
Arena	Phase Space Γ ($\cong \mathbb{R}^2$)	Hilbert space \mathcal{H} ($\cong L^2(\mathbb{R})$)
Canonical Variables	Real coordinates on Γ , p and q	Self-adjoint operators on \mathcal{H} , P and Q
Elements of Liouville Space (Algebra)	Complex functions of Γ $a(p, q)$	Operators on Hilbert space $A(P, Q)$

LECTURE 29:
Friday, April 03, 2020

Theorem 0.29.1 (Spectral Theorem for a General Quantum State). *Suppose we have a quantum state W with $W k t \psi_n = p_n |\psi_n\rangle$ and $\langle \psi_n | \psi_m \rangle = \delta_{nm}$ and $\sum_n |\psi_n\rangle \langle \psi_n| = 1$. This means that $\{|\psi_n\rangle\}$ is an eigenbasis of the Hilbert space to the state W . This means we can write*

$$W = \sum_n p_n |\psi_n\rangle \langle \psi_n|$$

where $0 \leq p_n \leq 1$ and $\sum_n p_n = 1$, which follows from $\text{Tr}(W) = 1$.

In this framework, what do we mean by expectation value?

$$\begin{aligned}
 \langle A \rangle &= \text{Tr}(WA) = \text{Tr}\left(\sum_n p_n |\psi_n\rangle \langle\psi_n| A\right) \\
 &= \sum_n p_n \text{Tr}(|\psi_n\rangle \langle\psi_n| A) \\
 &= \sum_n p_n \sum_m \langle\psi_m|\psi_n\rangle \langle\psi_n| A |\psi_m\rangle \\
 &= \sum_n p_n \underbrace{\langle\psi_n|A|\psi_n\rangle}_{\text{objective indeterminacy}}
 \end{aligned}$$

The term objective indeterminacy is something that is new in quantum mechanics. The value of this matrix element causes the result to not necessarily be deterministic. The p_n give subjective ignorance, a sense that we have not measured the state carefully enough. This is because the original state that we worked with is not pure.

The entropy of a state is a measure of our subjective ignorance plus, in quantum mechanics, the objective indeterminacy. Classically, the entropy is bounded by $-\infty \leq S \leq k_B \ln\left(\int_{\Gamma} \frac{dp dq}{h}\right) = +\infty$ if $\Gamma = \mathbb{R}^2$. In quantum mechanics, however, $0 \leq S \leq k_B \ln(\dim \mathcal{H}) = +\infty$ if $\mathcal{H} = L^2(\mathbb{R})$. The interesting thing is that the entropy cannot get arbitrarily small, it's bounded below by 0. This looks a lot like it has something to do with the third law of thermodynamics. For quantum mechanics, we can prove this:

$$W \ln(W) = \sum_n p_n \ln(p_n) |\psi_n\rangle \langle\psi_n|$$

This implies

$$S = -k_B \text{Tr}(W \ln(W)) = -k_B \sum_n p_n \ln(p_n)$$

If we have maximal knowledge about the system, then we have minimal entropy. Maximal knowledge means that $p_n = \delta_{n,n_0}$ (we know the system is exactly in one pure state). This makes $S = 0$ since $1 \ln(1) = 0 \ln(0) = 0$. Conversely, if we have minimal knowledge, we have maximal entropy. In that case, $p_n = \frac{1}{N}$ where $N = \dim \mathcal{H}$. If so, then

$$S = -k_B \sum_{n=1}^N \frac{1}{N} \ln\left(\frac{1}{N}\right) = -k_B \ln\left(\frac{1}{N}\right) = k_B \ln(N) = k_B \ln(\dim \mathcal{H})$$

which is infinite if the Hilbert space square-integrable over the reals.

0.30 The Quantum Canonical Ensemble

By analogy with the classical case, we can reasonably guess that the quantum canonical state is given by

$$W = \frac{1}{Z} e^{-\beta H}$$

Note that this is an operator, since H is an operator. The partition function must be

$$Z = \text{Tr}(e^{-\beta H})$$

Since H is self-adjoint (and hopefully compact) its eigenvalues are real and its eigenvectors form an orthonormal basis of Hilbert space:

$$H |n\rangle = E_n |n\rangle$$

such that $|n\rangle$ is an energy eigenvector. This corresponds to a pure energy eigenstate $|n\rangle \langle n|$ with the corresponding eigenvalue E_n . The fact that this basis is orthonormal means $\langle n|m\rangle = \delta_{nm}$ and $H = \sum_n E_n |n\rangle \langle n|$. This is the spectral theorem for (compact) operators.

Since the canonical state is a function of H , it is diagonal in the same basis as H . Therefore,

$$\begin{aligned} Z &= \text{Tr}(e^{-\beta H}) = \sum_n \langle n | e^{-\beta H} | n \rangle \\ &= \sum_n \langle n | e^{-\beta E_n} | n \rangle \\ &= \sum_n e^{-\beta E_n} \end{aligned}$$

It could happen that the Hamiltonian is degenerate. If we have degenerate eigenstates of H , we can (if we want) sum over the energy levels, but then we need to explicitly multiply back the degeneracy:

$$Z = \sum_n e^{-\beta E_n} = \sum_l \Omega(l) e^{-\beta E_l}$$

where l is an energy level and $\Omega(l)$ is the degeneracy of the level.

0.30.1 Thermal Averages

$$\langle A \rangle = \langle A, W \rangle = \text{Tr}(AW) = \sum_n \langle n | AW | n \rangle$$

where $|n\rangle$ is the energy eigenbasis and W is the canonical state. Then we get that

$$\begin{aligned} \langle A \rangle &= \sum_n \langle n | Ae^{-\beta E_n} \frac{1}{Z} | n \rangle \\ &= \sum_n p_n \langle n | A | n \rangle \end{aligned}$$

The matrix element is again a measure of objective indeterminacy while the p_n factors are a measure of subjective ignorance.

LECTURE 31:
Wednesday, April 08, 2020

Let's look at a couple examples of quantum systems so we can calculate some quantum partition functions.

Example. Consider a two-level system with $H = \epsilon n$ with $n \in \{0, 1\}$ such that the energy levels are $\{0, \epsilon\}$. In this case, the partition function is

$$Z = \sum_n = e^{-\beta \epsilon n} = 1 + e^{-\beta \epsilon}$$

so

$$F = -k_B T \ln(Z) = k_B T \ln(1 - e^{-\beta \epsilon})$$

$$U = \frac{\partial \beta F}{\partial \beta} = \frac{\epsilon e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} = \frac{\epsilon}{e^{\beta \epsilon} + 1}$$

$$\langle n \rangle = \sum_{n=0}^1 n \frac{e^{-\beta \epsilon n}}{Z} = \frac{0e^0 + 1e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} = \frac{1}{e^{\beta \epsilon} + 1}$$

so $U = \epsilon \langle n \rangle$. ◊

Example. Consider a symple harmonic oscillator with $H = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{\mathbf{Q}}^2 = \hbar\omega(a^\dagger a + \frac{1}{2}) = \hbar\omega(\hat{N} + \frac{1}{2})$, so

$$\begin{aligned} Z &= \text{Tr}(e^{-\beta H}) = \sum_{n=0}^{\infty} \langle n | e^{-\beta\hbar\omega(\hat{N}+\frac{1}{2})} | n \rangle \\ &= \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+\frac{1}{2})} \\ &= e^{-\frac{1}{2}\beta\hbar\omega} \sum_{n=0}^{\infty} (e^{-\beta\hbar\omega})^n \end{aligned}$$

This is a geometric series which converges to

$$Z = \frac{e^{-\frac{1}{2}\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} = \frac{1}{2 \sinh\left(\frac{\beta\hbar\omega}{2}\right)}$$

Observe that in the limit $T \rightarrow \infty$ or $\hbar \rightarrow 0$, $Z \rightarrow \frac{1}{\beta\hbar\omega} = \frac{k_B T}{\hbar\omega}$, which is the classical result. This explains why we insisted on an \hbar in the classical partition function.

$$F = k_B T \ln\left(2 \sinh\left(\frac{\beta\hbar\omega}{2}\right)\right)$$

$$U = \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1}$$

Also, since $U = \langle H \rangle = \hbar\omega(\langle n \rangle + \frac{1}{2})$,

$$\langle n \rangle = \frac{1}{e^{\beta\hbar\omega} - 1}$$

◊

Example. Finally, consider the Einstein model for a crystal. We have N particles, 3 dimensions, everything (nearest particles in a 3D lattice) connected by springs, so we have $3N$ eigenmodes. In the simplest possible approximation, assume they all have frequency ω .

$$Z = \left(\frac{1}{2 \sinh\left(\frac{\beta\hbar\omega}{2}\right)} \right)^{3N}$$

so

$$U = \frac{3}{2}Nk_B T + \frac{3N\hbar\omega}{e^{\beta\hbar\omega} - 1}$$

One very interesting thing about this system is the heat capacity,

$$c_V(T) = \frac{1}{N} \left. \frac{\partial U}{\partial T} \right|_{V,N} = 3k_B \left(\frac{\hbar\omega}{k_B T} \right)^2 \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2}$$

The interesting thing is the limits. At high temperature, $\beta\hbar\omega \rightarrow 0$, so $c_V \rightarrow 3k_B$. This is the classical result, and is known as the “law of Dulong-Petit”. This doesn’t actually happen in reality when we cool crystals down, and for a while nobody knew why. However, in the limit $\beta\hbar\omega \rightarrow \infty$, $c_V \rightarrow 3k_B(\beta\hbar\omega)e^{-\beta\hbar\omega}$. The important part is the exponential term, which says that in the low temperature limit, the heat capacity becomes exponentially small. Using this, Einstein explained why the heat capacity of crystals goes to 0 as they are cooled.

0.31 Blackbody Radiation

Let's remind ourselves of Maxwell's equations in vacuum:

$$\begin{aligned}\vec{\nabla} \cdot \vec{E} &= 0 \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{E} &= -\dot{\vec{B}} \\ \vec{\nabla} \times \vec{B} &= \frac{1}{c^2} \dot{\vec{E}} \\ \frac{1}{c^2} \ddot{\vec{E}} &= \vec{\nabla} \times \dot{\vec{B}} = -\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = -\vec{\nabla}(\vec{\nabla} \cdot \vec{E}) + \nabla^2 \vec{E} = \nabla^2 \vec{E}\end{aligned}$$

and we get the same with \vec{B} . This means the electric and magnetic fields satisfy the wave equation,

$$\left(\nabla^2 - \frac{1}{c^2} \partial_t^2 \right) \{ \vec{E}, \vec{B} \} = 0$$

Let's look at eigenmodes of a hollow metal cube of length L . We begin with the ansatz

$$E_x(\vec{r}, t) = E_{x0} \sin(\omega t) \cos(k_x x) \sin(k_y y) \sin(k_z z) E_y(\vec{r}, t) = E_{y0} \sin(\omega t) \sin(k_x x) \cos(k_y y) \sin(k_z z) E_z(\vec{r}, t) = E_{z0} \sin(\omega t) \sin(k_x x) \sin(k_y y) \cos(k_z z)$$

This ensures proper boundary conditions. The electric field is perpendicular to the boundary whenever $\{x, y, z\} = 0$. Also, inserting the solution into the wave equation, it solves it. However, there is also a wall at $\{x, y, z\} = L$. We need $k_i L = m_i \pi$ where $m_i \in \mathbb{N}$. Additionally, $k^2 = k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2}$ or $\omega = ck$. Therefore

$$\omega^2 = c^2 k^2 = \left(\frac{\pi c}{L} \right)^2 (m_x^2 + m_y^2 + m_z^2) = \left(\frac{\pi c}{L} \right)^2 \vec{m}^2$$

$$\text{where } \vec{m} = \begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} \in \mathbb{N}^3 - \{ \vec{0} \}.$$

States (“modes”) are enumerated by \vec{m} and the polarization $\sigma \in \{0, 1\}$. The state space is an “octant” in \mathbb{N}_1^3 or $\mathbb{N}_0^3 - \{ \vec{0} \}$. The density of states $D_{\vec{m}}(\vec{m}) = \sum_{\vec{n} \in \mathbb{N}_1^3} \delta(\vec{m} - \vec{n}) \approx 1$. The approximation comes from, rather than integrating over all of phase space, integrating over a box around one of these δ -functions, of which there are $\frac{L}{\pi c}$ in number.

◇

LECTURE 33: THE HARMONIC SOLID

Monday, April 13, 2020

Last time we finished up talking about blackbody radiation. Today we will do something that at first looks rather different but is in fact very similar. Imagine a sequence of masses m connected by springs in one-dimension. The particles lie at a resting displacement of a , and we will describe the displacement from that resting position as x , so the position of the j th particle is $\vec{r}_j = R_j + x_j = ja + x_j$. We can write down the energy of this system:

$$E = \sum_{j=1}^N \left\{ \frac{1}{2} m \dot{x}_j^2 + \frac{1}{2} (x_j - x_{j+1})^2 \right\}$$

However, this final term needs some additional definition at the boundaries ($j = 1, N$). In this case, we will use periodic boundary conditions such that $x_{N+1} = x_1$, as if the particles were actually connected in a circle. Unfortunately, the coupling between particles in the last term doesn't allow this to easily

factor in the partition function, but we can use normal modes to decouple the system. It turns out we can find the normal modes via a Fourier transform:

$$\tilde{x}_k = \frac{1}{\sqrt{N}} \sum_{j=1}^N x_j e^{-ikR_j} = \frac{1}{\sqrt{N}} \sum_{j=1}^N x_j e^{-ika j}$$

We can define the inverse transform as

$$x_j = \frac{1}{\sqrt{N}} \sum_k \tilde{x}_k e^{+ika j}$$

Let's check that this is actually the inverse transformation:

$$\begin{aligned} x_j &= \frac{1}{\sqrt{N}} \sum_k \frac{1}{\sqrt{N}} \sum_{j'} x_{j'} e^{-ikR_{j'}} e^{+ikR_j} \\ &= \sum_{j'} x_{j'} \frac{1}{N} \sum_k e^{\underbrace{ik(R_j - R_{j'})}_{\frac{2\pi}{L} n a(j-j')}} \\ &= \sum_{j'} x_{j'} \frac{1}{N} \begin{cases} N & j = j' \\ 0 & j \neq j' \end{cases} \end{aligned}$$

because the phases average out if $j \neq j'$

$$x_j = \sum_{j'} x_{j'} \delta_{jj'} = x_j$$

Now let's transform the Hamiltonian. We'll first look at the kinetic energy:

$$\begin{aligned} \frac{1}{2} m \sum_j \dot{x}_j^2 &= \frac{1}{2} m \sum_j \left(\frac{1}{\sqrt{N}} \sum_k \dot{\tilde{x}}_k e^{ikR_j} \right) \left(\frac{1}{\sqrt{N}} \sum_{k'} \dot{\tilde{x}}_{k'} e^{ik'R_j} \right) \\ &= \frac{1}{2} m \sum_{k,k'} \dot{\tilde{x}}_k \dot{\tilde{x}}_{k'} \underbrace{\frac{1}{N} \sum_j e^{i(k+k')R_j}}_{\delta_{k,-k'}} \\ &= \frac{1}{2} m \sum_k \dot{\tilde{x}}_k \dot{\tilde{x}}_{-k} \\ &= \frac{1}{2} m \sum_k |\dot{\tilde{x}}_k|^2 \end{aligned}$$

This last step can be justified by looking at how the Fourier transform is defined. $-k$ will give the complex conjugate of the point.

Next, the potential energy:

$$\frac{1}{2} K \sum_j (x_j - x_{j+1})^2 = \frac{1}{2} K \sum_j \left(\frac{1}{\sqrt{N}} \sum_k \tilde{x}_k e^{ikR_j} - e^{ikR_{j+1}} \right) \left(\frac{1}{\sqrt{N}} \sum_{k'} \tilde{x}_{k'} e^{ik'R_j} - e^{ik'R_{j+1}} \right)$$

Recall that $R_{j+1} = R_j + a$

$$\begin{aligned} &= \frac{1}{2} K \sum_{k,k'} \tilde{x}_k (1 - e^{ika}) \tilde{x}_{k'} (1 - e^{ik'a}) \frac{1}{N} \sum_j e^{i(k+k')R_j} \\ &= \frac{1}{2} K \sum_k |\tilde{x}_k|^2 \underbrace{(1 - e^{ika})(1 - e^{-ika})}_{1 - e^{-ika} - e^{ika} + 1 = 2 - 2 \cos(ka) = 4 \sin^2(\frac{ka}{2})} \\ &= \sum_k \frac{1}{2} K(k) |\tilde{x}_k|^2 \end{aligned}$$

where $K(k) = 4K \sin^2(\frac{ka}{2})$. The Hamiltonian is now diagonal, so we can factorize it! Every mode oscillates with frequency

$$\omega(k) = \sqrt{\frac{K(k)}{m}} = \sqrt{\frac{K}{m}} 2 \left| \sin\left(\frac{ka}{2}\right) \right| = 2\tilde{\omega} \left| \sin\left(\frac{ka}{2}\right) \right|$$

This is called the “dispersion relation”. The region from $-\frac{\pi}{a}$ to $\frac{\pi}{a}$ is known as the “first Brillouin zone”. For small ka , $\omega(k) \approx \tilde{\omega}|ka|$ and the phase velocity is $v = \frac{\omega(k)}{k} \approx_{k \rightarrow 0} \tilde{\omega}a$.

The classical energy can be easily calculated by equipartition theorem. There are N quadratic degrees of freedom for both position and momentum, so

$$U = \langle H \rangle = N \cdot 2 \cdot \frac{1}{2} k_B T = N k_B T$$

which tells us that

$$c_V = k_B$$

which follows from the “Dulong-Petit” law.

In quantum, the partition function factorizes, so

$$Z_{QM} = \prod_K Z_{QM,k} = \prod_k \frac{1}{2 \sinh\left(\frac{\beta \hbar \omega(k)}{2}\right)}$$

so

$$F_{QM} = -k_B T \ln(Z_{QM}) = k_B T \sum_k \ln\left(2 \sinh\left(\frac{\beta \hbar \omega(k)}{2}\right)\right)$$

and

$$U_{QM} = \frac{\partial \beta F_{QM}}{\partial \beta} = \sum_k \left\{ \frac{\hbar \omega(k)}{2} + \frac{\hbar \omega(k)}{e^{\beta \hbar \omega(k)} - 1} \right\}$$

This is better than the Einstein assumption that all the modes have the same frequency. What now does this predict about the heat capacity of a 3D solid? This was first done by Debye using the following approximation:

1 Restrict to low temperatures, so we only need low ω or $\hbar \omega(k) \approx \hbar \tilde{\omega}a|k| = \hbar v|k| = \hbar v \frac{\pi}{L} n = \hbar \omega(n)$.

2+3 Make the Brillouin zone a sphere which contains $3N$ modes:

$$3N = \frac{3}{8} \int_0^{n_D} dn 4\pi n^2 = \frac{1}{2} \pi n_D^3 \implies n_D = \left(\frac{6N}{\pi}\right)^{1/3}$$

LECTURE 34: THE DEBYE MODEL

Wednesday, April 15, 2020

We now need to calculate the energy from the Debye model:

$$U_{Debye} = 3 \frac{1}{8} \int_0^{n_D} dn 4\pi n^2 \frac{\hbar \omega(n)}{e^{\beta \hbar \omega(n)} - 1}$$

The prefactor of 3 counts the number of polarizations (two transverse and one longitudinal), the $1/8$ is because we are only integrating over an octant, and we have left out the ground state energy from the equation for U above.

$$U_{Debye} = \frac{3\pi}{2} \int_0^{n_D} dn n^2 \frac{\hbar v \frac{\pi}{L} n}{e^{\beta \hbar v \frac{\pi}{L} n} - 1}$$

Let's define $x \equiv \frac{\beta \hbar v \pi n}{L}$ and $dx = \frac{\beta \hbar v \pi}{L} dn$:

$$U_{Debye} = \frac{3\pi}{2} \int_0^{\frac{\beta \hbar v \pi n_D}{L}} dx x^2 \left(\frac{L}{\beta \hbar v \pi}\right)^3 \frac{1}{\beta} \frac{x}{e^x - 1}$$

We will next define a critical “Debye temperature” $\Theta_D \equiv \frac{\hbar v \pi n_D}{L k_B} = \frac{\hbar v}{k_B} \left(\frac{6\pi^2 N}{L^3} \right)^{1/3}$:

$$U_{\text{Debye}} = \frac{3\pi}{2} \frac{L^3 (k_B T)^4}{(\hbar v \pi)^3} \int_0^{\Theta_D/T} dx \frac{x^3}{e^x - 1}$$

If the upper boundary was infinite, the solution would be easy, but it isn’t. Instead, let’s look in two limits. First, look at large T . This seems inconsistent, since the approximations we made were meant to work for small T , but it turns out there are some subtle cancellations of errors. In this limit, $\frac{\Theta_D}{T} \rightarrow 0$, so we can Taylor expand the integrand about x :

$$\frac{x^3}{e^x - 1} = \frac{x^3}{1 + x + \frac{x^2}{2} + \dots - 1} \sim x^2 + \mathcal{O}(x^3)$$

In that limit,

$$U_{\text{Debye}}(T \rightarrow \infty) \approx \frac{3\pi}{2} \frac{L^3 (k_B T)^4}{(\hbar v \pi)^3} \int_0^{\Theta_D/T} dx x^2 = 3Nk_B T$$

Now remember the dispersion relation is certainly not linear for large k (large T). However, for large temperature, the energy of a harmonic oscillator will go to $\frac{1}{2}k_B T$ regardless. This is the boring limit, since we expect it to work out.

Now let’s look at the (asymptotic) limit as $T \rightarrow 0$:

$$U_{\text{Debye}}(T \rightarrow 0) \approx \frac{3\pi}{2} \frac{L^3 (k_B T)^4}{(\hbar v \pi)^3} \cdot \frac{\pi^4}{15}$$

so

$$\frac{U_{\text{Debye}}}{V} = u = \left(\frac{3\pi^2}{30\hbar^3 v^3} \right) (k_B T)^4$$

This is similar to the blackbody case, although we have $3/30$ rather than $2/30$ due to the polarizations. Additionally, we use the phonon speed v rather than the speed of light.

$$c_V \sim T^3$$

for small T .

0.32 Ideal Quantum Gasses

Let’s suppose we have a quantum state of N particles $\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$. For fermions/bosons we need to completely antisymmetrize/symmetrize this wave function. Note that fermions/bosons are particles with half-integer/integer spin. Why is this so? This is called the Spin-Statistics Theorem. The proof involves a deeper dive into quantum field theory and is not at all obvious.

$$\psi_{\{F,B\}} = X_{\{F,B\}} \sum_P \{\sigma(P), 1\} \psi(\vec{r}_{P(1)}, \vec{r}_{P(2)}, \dots, \vec{r}_{P(N)})$$

where P are the permutations of $\{1, \dots, N\}$ and σ is $+1$ for an even permutation and -1 for an odd permutation and X is some normalization factor. If the original wave function can be written as

$$\psi(\vec{r}_1, \dots, \vec{r}_N) = \prod_{i=1}^N \psi_{\alpha_i}(\vec{r}_i)$$

(a product of single-particle quantum states), then

$$\psi_{\{F,B\}} = X_{\{F,B\}} \sum_P \{\sigma(P), 1\} \prod_{i=1}^N \psi_{\alpha_i}(\vec{r}_{P(i)})$$

In the Fermion case, this is called the “Slater determinant”.

Suppose the Hamiltonian is of the form

$$H_N = \sum_{j=1}^N H_j,$$

a sum of individual Hamiltonians for each particle, and each H_j are identical. If this is true,

$$H_n \psi_{\{F,B\}} = E \psi_{\{F,B\}} = \sum_{j=1}^N \epsilon_{\alpha_j} \psi_{\{F,B\}}$$

so

$$Z = \sum'_{\{\alpha\}} e^{-\beta \sum_{j=1}^N \epsilon_{\alpha_j}} = \sum'_{\{\alpha\}} \prod_{j=1}^N e^{-\beta \epsilon_{\alpha_j}}$$

The sums which have a prime by them are special. These are sums over all quantum states subject to the (anti)symmetrization constraint. This is hard, because it doesn’t allow us to look at an individual particle without considering all the others to maintain symmetry. However, there is a better way of dealing with this, but it requires a clever change of notation:

LECTURE 35: IDEAL QUANTUM GASSES

Friday, April 17, 2020

Now that we have a new way of labeling states, let’s look at the partition function in occupation number representation.

$$Z = \sum_{\{n_\alpha\}}^{\sum_\alpha n_\alpha = N} \prod_\alpha e^{-\beta \epsilon_\alpha n_\alpha}$$

An additional constraint must be made on this sum. For Fermions, we need to consider the Pauli exclusion principle, so the sum will have $n_\alpha \in \{0, 1\}$, whereas Bosons can have $n_\alpha \in \mathbb{N}$.

The fixed N constraint is annoying and makes this hard to calculate, but we can remove this constraint by moving to the grand canonical ensemble:

$$\begin{aligned} Z &= \sum_{N=0}^{\infty} Z_N e^{\beta \mu N} \\ &= \sum_{N=0}^{\infty} \sum_{\{n_\alpha\}}^{\sum_\alpha n_\alpha = N} \prod_\alpha e^{-\beta(\epsilon_\alpha - \mu)n_\alpha} \end{aligned}$$

The interesting thing here is that the first sum makes the constraint $\sum_\alpha n_\alpha = N$ superfluous! Since we sum all N up anyway, there is no reason to separate these into groups of fixed N .

$$\begin{aligned} Z &= \sum_{\{n_\alpha\}} \prod_\alpha e^{-\beta(\epsilon_\alpha - \mu)n_\alpha} \\ &= \sum_{n_1} \sum_{n_2} \dots \prod_\alpha e^{-\beta(\epsilon_\alpha - \mu)n_\alpha} \\ &= \sum_{n_1} e^{-\beta(\epsilon_1 - \mu)n_1} \sum_{n_2} e^{-\beta(\epsilon_2 - \mu)n_2} \dots \\ &= \prod_\alpha \sum_{n_\alpha} e^{-\beta(\epsilon_\alpha - \mu)n_\alpha} \end{aligned}$$

Therefore

$$\begin{aligned}\Omega &= -k_B T \ln(\mathcal{Z}) = \begin{cases} -k_B T \sum_{\alpha} \ln \left(\sum_{n_{\alpha}=0}^1 e^{-\beta(\epsilon_{\alpha}-\mu)n_{\alpha}} \right) \\ -k_B T \sum_{\alpha} \ln \left(\sum_{n_{\alpha}=0}^{\infty} (e^{-\beta(\epsilon_{\alpha}-\mu)})^{n_{\alpha}} \right) \end{cases} \\ &= \begin{cases} -k_B T \sum_{\alpha} \ln(1 + e^{-\beta(\epsilon_{\alpha}-\mu)}) & \text{Fermions} \\ -k_B T \sum_{\alpha} \ln \left(\frac{1}{1-e^{-\beta(\epsilon_{\alpha}-\mu)}} \right) & \text{Bosons} \end{cases} \\ &= \begin{cases} -k_B T \sum_{\alpha} \ln(1 + e^{-\beta(\epsilon_{\alpha}-\mu)}) & \text{Fermions} \\ +k_B T \sum_{\alpha} \ln(1 - e^{-\beta(\epsilon_{\alpha}-\mu)}) & \text{Bosons} \end{cases} = \mp k_B T \sum_{\alpha} \ln(1 \pm e^{-\beta(\epsilon_{\alpha}-\mu)})\end{aligned}$$

When we use $\{\mp, \pm\}$ notation, we will always use it in the order in which the Fermion is in the top of the notation and the Boson is in the bottom.

How many particles do we have in a particular state α ?

$$\begin{aligned}\langle n_{\alpha} \rangle &= \frac{1}{\mathcal{Z}} \left(\prod_{\alpha'} \sum_{n_{\alpha'}} n_{\alpha'} e^{-\beta(\epsilon_{\alpha'}-\mu)n_{\alpha'}} \right) \\ &= \frac{1}{\mathcal{Z}} \left(\prod_{\alpha'} \sum_{n_{\alpha'}} \left(-\frac{1}{\beta} \frac{\partial}{\partial \epsilon_{\alpha'}} \right) e^{-\beta(\epsilon_{\alpha'}-\mu)n_{\alpha'}} \right) \\ &= -\frac{1}{\beta} \frac{1}{\mathcal{Z}} \frac{\partial}{\partial \epsilon_{\alpha}} \mathcal{Z} \\ &= -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_{\alpha}} \ln(\mathcal{Z}) = \frac{\partial \Omega}{\partial \epsilon_{\alpha}} \\ \langle n_{\alpha} \rangle &= \frac{\partial}{\partial \epsilon_{\alpha}} \left[\mp k_B T \sum_{\alpha'} \ln(1 \pm e^{-\beta(\epsilon_{\alpha'}-\mu)}) \right] \\ &= \mp k_B T \frac{\pm e^{-\beta(\epsilon_{\alpha}-\mu)}(-\beta)}{1 \pm e^{-\beta(\epsilon_{\alpha}-\mu)}} \\ &= \frac{e^{-\beta(\epsilon_{\alpha}-\mu)}}{1 \pm e^{-\beta(\epsilon_{\alpha}-\mu)}} \\ &= \frac{1}{e^{\beta(\epsilon_{\alpha}-\mu)} \pm 1} \begin{cases} \text{Fermi} \\ \text{Bose} \end{cases}\end{aligned}$$

This expression appears a lot, so we will write it as

$$\langle n_{\alpha} \rangle = \mathcal{F}_{\pm}(\epsilon_{\alpha} - \mu)$$

For large ϵ , the function $\mathcal{F}(\epsilon)$ is just the Boltzmann factor $e^{-\beta\epsilon}$, regardless of Fermion or Boson statistics. However, the expectation values are very different for small ϵ .

We can write down the density of states for a single particle as

$$D(\epsilon) = \sum_{\alpha} \delta(\epsilon - \epsilon_{\alpha})$$

and the cumulative density of states as

$$W(\epsilon) = \int_{-\infty}^{\epsilon} d\epsilon' D(\epsilon')$$

We can therefore write

$$\Omega = \mp k_B T \int_0^{\infty} d\epsilon D(\epsilon) \ln(1 \pm e^{-\beta(\epsilon-\mu)})$$

Let's do integration by parts, integrating $D(\epsilon)$ and differentiating the logarithm.

$$\Omega = - \int d\epsilon W(\epsilon) \mathcal{F}_{\pm}(\epsilon - \mu)$$

Therefore,

$$\langle N \rangle = \sum_{\alpha} \langle n_{\alpha} \rangle = \int d\epsilon D(\epsilon) \mathcal{F}_{\pm}(\epsilon - \mu) = -\frac{\partial \Omega}{\partial \mu}$$

LECTURE 36:
Monday, April 20, 2020

In the last lecture, we worked out that

$$N(\mu) = \int d\epsilon \frac{D(\epsilon)}{e^{\beta(\epsilon-\mu)} \pm 1}$$

for Fermions and Bosons. This equation is often used to find μ as a function of N . This is useful if we wish to re-express things as a function of N , since we lost this control when we moved to the grand canonical ensemble.

What is the average energy? To do this, note that

$$\begin{aligned} \frac{\partial}{\partial \beta} \ln(1 \pm e^{-\beta(\epsilon-\mu)}) &= \frac{\pm e^{-\beta(\epsilon-\mu)}}{1 \pm e^{-\beta(\epsilon-\mu)}} (-(\epsilon - \mu)) \\ &= \mp f_{\pm}(\epsilon - \mu)(\epsilon - \mu) \\ \implies \epsilon f_{\pm}(\epsilon - \mu) &= \mp \frac{\partial}{\partial \beta} \ln(1 \pm e^{-\beta(\epsilon-\mu)}) + \mu f_{\pm}(\epsilon - \mu) \end{aligned}$$

$$\begin{aligned} E &= \langle H \rangle = \sum_{\alpha} \epsilon_{\alpha} \langle n_{\alpha} \rangle \\ &= \int d\epsilon D(\epsilon) \epsilon f_{\pm}(\epsilon - \mu) \\ &= \int d\epsilon D(\epsilon) \left[\mp \frac{\partial}{\partial \beta} \ln(1 \pm e^{-\beta(\epsilon-\mu)}) + \mu f_{\pm}(\epsilon - \mu) \right] \\ &= \underbrace{\frac{\partial \beta \Omega}{\partial \beta}}_{E - \mu N} - \mu \underbrace{\frac{\partial \Omega}{\partial \mu}}_{-N} \end{aligned}$$

0.33 Density of States for Free Particles in a Cubic Box

In general, we can do this for a box with volume $V = L^d$ for any dimension d . We know that the energy levels should be

$$\frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} \left(\frac{\pi}{L} \vec{n} \right)^2 \quad \vec{n} \in \mathbb{N}_0^d - \{ \vec{0} \}$$

Therefore,

$$\begin{aligned} D(\epsilon) &= \int_{\mathbb{R}_+^d} d^d n \delta \left(\epsilon - \frac{\hbar^2 \pi^2}{2m L^2} \vec{n}^2 \right) \\ &= \frac{1}{2^d} \int_{\mathbb{R}^d} \delta \left(\epsilon - \frac{\hbar^2 \pi^2}{2m L^2} \vec{n}^2 \right) \\ y^2 &= \frac{\hbar^2 \pi^2}{2m L^2} n^2 \quad dy = \frac{\pi \hbar}{\sqrt{2m L^2}} dn \\ D(\epsilon) &= \frac{1}{2^d} \int d^d y \left(\frac{\sqrt{2m L}}{\pi \hbar} \right)^d \delta(\epsilon - y^2) \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2^d} \left(\frac{\sqrt{2m}}{\pi\hbar} \right)^d V \int_0^\infty dy A_d y^{d-1} \delta(\epsilon - y^2) \\
 x = y^2 \quad dy = \frac{1}{2\sqrt{x}} dx \\
 D(\epsilon) &= \frac{1}{2^d} \left(\frac{\sqrt{2m}}{\pi\hbar} \right)^d V \int_0^\infty \frac{dx}{2\sqrt{x}} A_d x^{\frac{d-1}{2}} \delta(\epsilon - x) \\
 &= \frac{1}{2} \frac{(2m)^{d/2}}{h^d} V A_d \epsilon^{d/2-1}
 \end{aligned}$$

Recall that the surface area of a d -sphere is

$$A_d = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}$$

so

$$D(\epsilon) = \left(\frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \epsilon^{\frac{d}{2}-1}$$

This expression does not include any mention of spin, which may be important later.

0.33.1 The Equation of Clapeyron

Recall that $D(\epsilon) = c_d \epsilon^{d/2-1}$. Because of this, we can write

$$D(\epsilon) = \frac{2}{d} \left[\frac{d}{d\epsilon} (\epsilon D(\epsilon)) \right]$$

$$PV = -\Omega = \pm k_B T \int_0^\infty d\epsilon D(\epsilon) \ln(1 \pm e^{-\beta(\epsilon-\mu)})$$

Now we are going to insert this weird rewriting of $D(\epsilon)$:

$$\begin{aligned}
 PV &= \pm k_B T \frac{2}{d} \int_0^\infty d\epsilon \frac{d}{d\epsilon} (\epsilon D(\epsilon)) \ln(1 \pm e^{-\beta(\epsilon-\mu)}) \\
 &= \mp k_B T \frac{2}{d} \int_0^\infty d\epsilon \epsilon D(\epsilon) \frac{\mp \beta e^{-\beta(\epsilon-\mu)}}{1 \pm e^{-\beta(\epsilon-\mu)}} \\
 &= \frac{2}{d} \int d\epsilon \epsilon D(\epsilon) f_\pm(\epsilon - \mu) \\
 &= \frac{2}{d} E
 \end{aligned}$$

so

$$E = \frac{d}{2} PV$$

This is incredible, since this is exactly the classical result, but we derived it using quantum statistics.

0.33.2 Grand Potential of a Free Ideal Quantum Gas

$$\Omega(T, V, \mu) = \mp k_B T \int_0^\infty d\epsilon D(\epsilon) \ln(1 \pm e^{-\beta(\epsilon-\mu)})$$

Define

$$z \equiv e^{\beta\mu}$$

as the “fugacity” and

$$D(\epsilon) = (2s+1) \left(\frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \epsilon^{d/2-1}$$

including the spin degeneracy.

$$\Omega = \mp k_B T (2s+1) \left(\frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \int d\epsilon \epsilon^{\frac{d}{2}-1} \ln(1 \pm ze^{-\beta\epsilon})$$

$$(t = \beta\epsilon \quad dt = \beta d\epsilon)$$

$$\Omega = \mp k_B T (2s+1) \left(\frac{\sqrt{2\pi m k_B T}}{h} \right)^d V \underbrace{\frac{1}{\Gamma(\frac{d}{2})} \int_0^\infty dt t^{\frac{d}{2}-1} \ln(1 \pm ze^{-t})}_{-L_{\frac{d}{2}-1}(\mp z)}$$

where $L_\nu(z)$ is a polylogarithm.

$$\Omega(T, V, \mu) = \pm k_B T (2s+1) \frac{V}{\lambda_{\text{th}}^d} L_{\frac{d}{2}+1}(\mp Z)$$

Now we can use some of the properties of the polylog:

$$\frac{PV}{k_B T} = -\beta\Omega = \mp(2s+1) \frac{V}{\lambda_{\text{th}}^d} L_{\frac{d}{2}+1}(\mp z)$$

and

$$N = -\frac{\partial\Omega}{\partial\mu} = z \frac{\partial}{\partial z}(-\beta\Omega) = \mp(2s+1) \frac{V}{\lambda_{\text{th}}^d} L_{\frac{d}{2}}(\mp z)$$

These two equations can be seen as a parametric representation of the thermal equation of state (with z being the parameter). If we take the ratio of these equations, we find that

$$\frac{PV}{Nk_B T} = \frac{L_{\frac{d}{2}+1}(\mp z)}{L_{\frac{d}{2}}(\mp z)} = \begin{cases} \geq 1 & \text{Fermions} \\ = 1 & \text{Boltzmann (Classical)} \\ \leq 1 & \text{Bosons} \end{cases}$$

In the case of Fermi/Bose statistics, we find an additional repulsions/attraction between particles which is not present in the classical case.

LECTURE 37: BOSE-EINSTEIN CONDENSATION

Wednesday, April 22, 2020

0.34 Bose-Einstein Condensation

For a Boson gas, at $T = 0$, every particle is in the ground state. This is actually very unremarkable and not what we mean by BE condensation. The interesting thing is that for a given density $n = \frac{N}{V}$, there exists a temperature $T_E > 0$ such that for $0 < T < T_E$, the ground state is “macroscopically occupied”.

$$n_0 \equiv \lim_{V \rightarrow \infty} \frac{\langle n_{\alpha_0} \rangle}{V} > 0$$

Observe that

$$\langle n_{\alpha_0} \rangle = \frac{1}{e^{\beta(\underbrace{\epsilon_0}_0 - \mu)} - 1} = \frac{z}{1-z} \sim L^d$$

since the fugacity is

$$z \sim \frac{L^d}{1+L^d} \sim 1 - cL^{-d}$$

The fugacity is extremely close to 1 so μ is extremely close to 0 = ϵ_0 . What about the first excited state? Observe that $\epsilon_k \sim k^2 \sim \frac{1}{L^2}$ so $\epsilon_0 \sim \frac{\gamma}{L^2}$.

$$\begin{aligned}\langle n_{\alpha_1} \rangle &= \frac{z}{e^{\beta\epsilon_1} - z} = \frac{z}{e^{\beta\gamma/L^2} - z} = \frac{1 - cL^{-d}}{1 + \frac{\beta\gamma}{L^2} - 1 + cL^{-d}} \\ &= \frac{L^d - c}{\beta\gamma L^{d-2} + c} \\ &\sim \begin{cases} L^2 & d = 3 \implies \frac{\langle n_{\alpha_1} \rangle}{\langle n_{\alpha_0} \rangle} = \frac{1}{L} \rightarrow 0 \\ L^2 & d = 2 \implies \frac{\langle n_{\alpha_1} \rangle}{\langle n_{\alpha_0} \rangle} = 1 \\ L & d = 1 \implies \frac{\langle n_{\alpha_1} \rangle}{\langle n_{\alpha_0} \rangle} = 1 \end{cases}\end{aligned}$$

In 1 and 2 dimensions, the first excited state occupancy is like the ground state (boring). Condensation only happens in 3 dimensions. It may happen that the total number of particles in all excited states cannot go beyond a certain number:

$$N = (2s+1) \frac{V}{\lambda_{\text{th}}^d} L_{d/2}(z) \leq (2s+1) \frac{V}{\lambda_{\text{th}}^d} L_{d/2}(1)$$

Now recall that $L_{d/2}(z)$ diverges for $z \rightarrow 1$ for $d = 1$ and $d = 2$. However, $L_{3/2}(1) = \zeta(3/2) = 2.612 \dots < \infty$. This formula claims that the total number of particles in a system is finite. This cannot be true. In reality, we are writing down the total number of particles in excited states. The ground state particles are not counted in the integral because the density of states $D(\epsilon = 0) = 0!$ Therefore, once we reach the limit of particles given here, every particle added will condense into the ground state.

$$\begin{aligned}N_{\text{excited, max}} &= (2s+1) \frac{V}{\lambda_{\text{th}}^3} \gamma\left(\frac{3}{2}\right) \\ N_0 &= N - N_{\text{excited, max}} = N - (2s+1) \frac{V}{\lambda_{\text{th}}^3} \gamma\left(\frac{3}{2}\right)\end{aligned}$$

so

$$\frac{N_0}{N} = 1 - (2s+1) \frac{V}{N \lambda_{\text{th}}^3} \gamma\left(\frac{3}{2}\right)$$

Recall that $\lambda_{\text{th}} \sim \frac{1}{\sqrt{T}}$, so

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_E}\right)^{3/2}$$

such that

$$T_E = \frac{\hbar^2}{2\pi m k_B} \frac{1}{[(2s+1) \frac{V}{N} \zeta(\frac{3}{2})]^{2/3}}$$

For sufficiently large $T > T_E$, the ratio $\frac{N_0}{N} = 0$. At $T = 0$, $\frac{N_0}{N} = 1$. Neither of these properties are remarkable. The amazing thing is that for $T < T_E$, this ratio is between 0 and 1! At $T = T_E$, we have

$$\gamma\left(\frac{3}{2}\right) (2s+1) \frac{1}{n \lambda_{TE}^3} = 1$$

so

$$n \lambda_{TE}^3 \approx \frac{0.38}{2s+1}$$

so condensation happens when n exceeds this (huge) value. Let's do some comparisons. For an ideal gas at STP (say nitrogen, with mass $m = 4.65 \times 10^{-26}$ kg), we can work out

$$\lambda_{\text{th}} = \frac{\hbar}{\sqrt{2\pi m k_B T}} = 1.9 \times 10^{-10} \text{ m}$$

The density is $n = \frac{N}{V} = \frac{P}{k_B T} = \frac{10^5 \text{ Pa}}{4.1 \times 10^{-31} \text{ J}} = 2.4 \times 10^{25} \text{ m}^{-3}$ so

$$n \lambda_{\text{th}}^3 = 1.6 \times 10^{-7} \ll 1$$

To see BEC, we need to either increase n or decrease T . We can't really do the first thing, because increasing the density increases the interactions between the particles (ideal gasses don't exist, but they work to good approximation with low density). For example, common BECs are made with sodium atoms at $T = 2\mu\text{K}$. $m_{\text{Na}} = 32\text{u}$ and $\lambda_{\text{th}} = 2.6 \times 10^{-7}\text{m}$ so λ_{th}^3 is bigger by a factor of 2.5×10^{12} .

$$n_{\text{BEC}} \approx \frac{0.38}{(2s+1)\lambda_{\text{th}}^3} = 7.3 \times 10^{18} \frac{\text{particles}}{\text{m}^3}$$

This is three million times less dense than air! In any other situation, this “condensate” would be called a vacuum.

What are the properties of a BEC below T_E ? What can we actually say about the system in this special situation? This is actually the easy part of the problem, because the fugacity is approximately $z = 1$ with fantastic accuracy. This makes evaluating the polylog really easy.

$$\frac{PV}{k_B T} = -\beta\Omega = (2s+1) \frac{V}{\lambda_{\text{th}}^3} \underbrace{L_{5/2}(1)}_{1.341\dots}$$

so

$$P = (2s+1)\gamma\left(\frac{5}{2}\right) \frac{k_B T}{\lambda_{\text{th}}^3} \propto T^{5/2}$$

This is independent of V ! This makes some neat things happen, like $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right) = \infty$ —BECs are infinitely compressible. We can also write

$$P = k_B T \left(\frac{T}{T_E}\right)^{3/2} \frac{N}{V} \frac{\zeta\left(\frac{5}{2}\right)}{\gamma\left(\frac{3}{2}\right)}$$

so

$$c_V = \frac{1}{N} \frac{\partial U}{\partial T} \Big|_V \underset{U=\frac{3}{2}PV}{=} \frac{3}{2} \frac{V}{N} \frac{\partial P}{\partial T} \Big|_V = \frac{15}{4} \frac{\zeta(5/2)}{\zeta(3/2)} k_B \left(\frac{T}{T_E}\right)^{3/2} \approx 1.926 k_B \left(\frac{T}{T_E}\right)^{3/2}$$

The classical limit of c_V is $\frac{3}{2}k_B T^{3/2}$, so at the Einstein temperature, the graph of $c_V(T)$ has a kink. As it turns out, $c_V(T \gtrsim T_E)$ is really hard to calculate.

LECTURE 38: THE IDEAL FERMI GAS

Friday, April 24, 2020

0.35 The Ideal Fermi Gas

For high temperatures, $f_+(\epsilon - \mu)$ is asymptotic to $e^{-\beta(\epsilon-\mu)}$. At low temperature, it is asymptotic to $1 - e^{\beta(\epsilon-\mu)}$. At $\epsilon = \mu$, it has a slope of $-\frac{1}{4}\beta$ and the distribution function equals 1/2. Now let's see how we can learn some facts about this distribution.

$$N = \sum_{\alpha} \langle n_{\alpha} \rangle = \sum_{\alpha} \frac{1}{e^{\beta(\epsilon_{\alpha}-\mu)} + 1} = \int d\epsilon \frac{D(\epsilon)}{e^{\beta(\epsilon-\mu)} + 1}$$

$$U = \sum_{\alpha} \epsilon_{\alpha} \langle n_{\alpha} \rangle = \sum_{\alpha} \frac{\epsilon_{\alpha}}{e^{\beta(\epsilon_{\alpha}-\mu)} + 1} = \int d\epsilon \frac{\epsilon D(\epsilon)}{e^{\beta(\epsilon-\mu)} + 1}$$

We can additionally define the Fermi energy:

$$\epsilon_F \equiv \lim_{T \rightarrow 0} \mu(T, N)$$

Note that this is not the point where the distribution is 1/2, that point is μ . Instead, the Fermi energy lies halfway between the highest occupied state and the lowest unoccupied state.

Observe that

$$\frac{1}{1+x} + \frac{1}{1+\frac{1}{x}} = 1$$

This implies

$$\frac{1}{e^{\beta(\epsilon-\mu)} + 1} = 1 - \frac{1}{e^{-\beta(\epsilon-\mu)} + 1}$$

Let's imagine a two-state system with energies $E_1 = \epsilon$ and $E_0 = 0$. We can easily calculate

$$\begin{aligned} N &= \frac{1}{e^{\beta(E_0-\mu)} + 1} + \frac{1}{e^{\beta(E_1-\mu)} + 1} \\ &= \frac{1}{e^{-\beta\mu} + 1} + \frac{1}{e^{\beta(\epsilon-\mu)} + 1} \\ &= 1 - \frac{1}{e^{\beta\mu} + 1} + \frac{1}{e^{\beta(\epsilon-\mu)+1}} \end{aligned}$$

Say we have one particle in the system. Then

$$\frac{1}{e^{\beta\mu} + 1} = \frac{1}{e^{\beta(\epsilon-\mu)} + 1}$$

or

$$\beta\mu = \beta(\epsilon - \mu) \implies \mu = \frac{\epsilon}{2}$$

Now imagine a continuous spectrum of energies:

$$D(\epsilon) = (2s+1) \left(\frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \epsilon^{\frac{d}{2}-1}$$

We can calculate the number of particles by taking $T = 0$ and integrating up to the Fermi energy (let's also say these are spin- $\frac{1}{2}$ particles).

$$N = \int_0^{\epsilon_F} d\epsilon 2(\cdots) \epsilon^{\frac{d}{2}-1} = 2 \left(\underbrace{\frac{\sqrt{2\pi m \epsilon_F}}{h}}_{k_B T_F = \epsilon_F} \right)^d \frac{V}{\Gamma(\frac{d}{2}+1)}$$

In three dimensions,

$$N = 2 \left(\frac{\sqrt{2\pi m \epsilon_F}}{h} \right)^3 \frac{V}{\frac{3}{4}\sqrt{\pi}} = \frac{V}{3\pi^2} \left(\frac{2m}{\hbar} \right)^{3/2} \epsilon_F^{3/2}$$

where

$$\epsilon_F = \frac{\hbar^2}{2m} (3\pi^2 \underbrace{n}_{N/V})^{2/3}$$

Also, at $T = 0$,

$$\begin{aligned} U &= \int_0^{\epsilon_F} d\epsilon \epsilon D(\epsilon) \\ &= \int_0^{\epsilon_F} d\epsilon 2 \left(\frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \epsilon^{d/2} \\ &= 2 \left(\frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})(\frac{d}{2}+1)} \epsilon^{\frac{d}{2}+1} \end{aligned}$$

Therefore

$$\frac{U}{N} = \frac{d}{d+2} \epsilon_F = \begin{cases} \frac{3}{5} \epsilon_F & d=3 \\ \frac{1}{2} \epsilon_F & d=2 \\ \frac{1}{3} \epsilon_F & d=1 \end{cases}$$

Also recall that we found $U = \frac{d}{2}PV$, so we can now calculate the pressure (again, at $T = 0$):

$$P = \frac{2}{d} \frac{U}{V} = 2 \left(\frac{\sqrt{2\pi m}}{h} \right)^d \frac{\epsilon_F^{\frac{d}{2}+1}}{\Gamma(\frac{d}{2})(\frac{d}{2}+1)}$$

We also worked out that $\epsilon_F \propto n^{2/d}$ so $P \propto \epsilon_F^{d/2+1} \propto n^{1+2/d} \propto V^{-(1+\frac{2}{d})}$. This pressure is called the Fermi pressure, and what's amazing is that it's nonzero at $T = 0$, unlike the regular ideal gas. We can also calculate $\kappa_T = -\frac{1}{V} \frac{\partial P}{\partial V}|_T$ or equivalently, the isothermal bulk modulus

$$\begin{aligned} K &= -V \left. \frac{\partial P}{\partial V} \right|_T = -V \left(-1 - \frac{2}{d} \right) \frac{P}{V} = \left(1 + \frac{2}{d} \right) P = \left(1 + \frac{2}{d} \right) \frac{PV}{N} \frac{N}{V} \\ &= \left(1 + \frac{2}{d} \right) \frac{\frac{2}{d} U}{N} \frac{N}{V} \\ &= \left(1 + \frac{2}{d} \right) \frac{2}{d} \frac{d}{d+2} \epsilon_F \frac{N}{V} \\ &= \frac{2}{d} \epsilon_F \frac{N}{V} \end{aligned}$$

This says that the bulk modulus of an ideal Fermi gas at $T = 0$ is equal to $2/d$ times the Fermi energy divided by the volume per particle.

For small T , the thermodynamics is determined by $D(\epsilon)$ in the vicinity of ϵ_F . Generally, we want to calculate

$$I = \int d\epsilon g(\epsilon) f_+(\epsilon - \mu)$$

The answer is the Sommerfeld Expansion, developed by Arnold Sommerfeld, and we will discuss this in the next lecture.

LECTURE 39: THE SOMMERFELD EXPANSION

Monday, April 27, 2020

At the end of the last class, we wanted to evaluate

$$I_N = \int d\epsilon g(\epsilon) \mathcal{F}_+(\epsilon - \mu)$$

- The first assumption we will make is $\lim_{\epsilon \rightarrow -\infty} g(\epsilon) = 0$ so that there are no states below the ground state.
- The second we will make is that $g(\epsilon) \propto \epsilon^\alpha$ as $\epsilon \rightarrow \infty$.
- The third is that $g(\epsilon)$ is sufficiently smooth at $\epsilon = \epsilon_F$.
- Finally, we will define $G(\epsilon) = \int_{-\infty}^{\epsilon} d\epsilon' g(\epsilon')$.

Now consider first for a fixed μ and $\tilde{f}(x) = \frac{1}{e^x + 1}$:

$$\begin{aligned} I &= \int d\epsilon G'(\epsilon) \tilde{\mathcal{F}} \left(\frac{\epsilon - \mu}{k_B T} \right) \\ &= - \int d\epsilon G(\epsilon) \tilde{\mathcal{F}}' \left(\frac{\epsilon - \mu}{k_B T} \right) \frac{1}{k_B T} \\ &= -dx \underbrace{G(\mu + k_B T x)}_{G(\mu) + x k_B T G'(\mu) + \frac{1}{2} (x k_B T)^2 G''(\mu)} \tilde{\mathcal{F}}'(x) \quad x \equiv \frac{\epsilon - \mu}{k_B T} \\ &= -G(\mu) \underbrace{\int dx \tilde{\mathcal{F}}'(x)}_{-1} - k_B T G'(\mu) \underbrace{\int dx x \tilde{\mathcal{F}}'(x)}_0 - \frac{1}{2} (k_B T)^2 G''(\mu) \underbrace{\int dx x^2 \tilde{\mathcal{F}}'(x)}_{-\pi^2/3} + \dots \\ &= G(\mu) + \frac{\pi^2}{6} (k_B T)^2 G''(\mu) + \mathcal{O}(T^4) \end{aligned}$$

We can think of this as a low- T expansion of the FD-distribution at fixed μ :

$$\mathcal{F}_+(\epsilon - \mu) = \Theta(\mu - \epsilon) - \frac{\pi^2}{6}(k_B T)^2 \delta'(\epsilon - \mu) + \mathcal{O}(T^4)$$

Let's apply this. Suppose $W(\epsilon) = \int_{-\infty}^{\epsilon} d\epsilon' D(\epsilon')$:

$$N = \int_{-\infty}^{\infty} d\epsilon D(\epsilon) \mathcal{F}_+(\epsilon - \mu) = W(\mu) + \frac{\pi^2}{6}(k_B T)^2 W''(\mu) + \mathcal{O}(T^4)$$

Now we “just” need to solve this for $\mu(N)$. This is terribly hard, but here’s a nice solution. We are going to write μ as a series expansion in $k_B T$, insert this into the right-hand side, and then expand again for small $k_B T$! Finally, we compare coefficients of $(k_B T)^n$.

$$\begin{aligned} \mu &= \epsilon_F + \mu_1 k_B T + \mu_2 (k_B T)^2 + \dots \\ N &= W(\epsilon_F + \mu_1 k_B T + \mu_2 (k_B T)^2 + \dots) + \frac{\pi^2}{6}(k_B T)^2 W''(\epsilon_F + \mu_1 k_B T + \mu_2 (k_B T)^2 + \dots) + \dots \\ &= W(\epsilon_F) + W'(\epsilon_F) (\mu_1 k_B T + \mu_2 (k_B T)^2 + \dots) + \frac{1}{2} W''(\epsilon_F) (\mu_1 k_B T + \mu_2 (k_B T)^2 + \dots)^2 + \dots \\ &\quad + \frac{\pi^2}{6}(k_B T)^2 [W''(\epsilon_F) + W'''(\epsilon_F) (\mu_1 k_B T + \mu_2 (k_B T)^2 + \dots) + \dots] \\ &= \underbrace{W(\epsilon_F)}_N + k_B T \underbrace{[W'(\epsilon_F) \mu_1]}_{\mu_1=0} + (k_B T)^2 \underbrace{\left[W'(\epsilon_F) \mu_2 + \frac{1}{2} W''(\epsilon_F) \mu_1^2 + \frac{\pi^2}{6} W''(\epsilon_F) \right]}_{\mu_2 = -\frac{\pi^2}{6} \frac{W''(\epsilon_F)}{W'(\epsilon_F)} = -\frac{\pi^2}{6} \frac{D'(\epsilon_F)}{D(\epsilon_F)}} + \dots \\ \implies \mu(T, N) &= \epsilon_F - \frac{\pi^2}{6}(k_B T)^2 \frac{D'(\epsilon_F)}{D(\epsilon_F)} + \dots \end{aligned}$$

We can insist this is our expression for $\mathcal{F}_+(\epsilon - \mu)$:

$$\begin{aligned} \mathcal{F}_+(\epsilon - \mu) &= \Theta\left(\epsilon_F - \frac{\pi^2}{6}(k_B T)^2 \frac{D'}{D} + \dots - \epsilon\right) - \frac{\pi^2}{6}(k_B T)^2 \delta'\left(\epsilon - \epsilon_F + \frac{\pi^2}{6}(k_B T)^2 \frac{D'}{D} + \dots\right) + \mathcal{O}(T^4) \\ &= \Theta(\epsilon_F) + \Theta'(\epsilon_F) \left(-\frac{\pi^2}{6}(k_B T)^2 \frac{D'}{D} + \dots\right) + \dots - \frac{\pi^2}{6}(k_B T)^2 (\epsilon - \epsilon_F) + \dots \\ &= \Theta(\epsilon_F - \mu) - \frac{\pi^2}{6}(k_B T)^2 \left[\frac{D'(\epsilon_F)}{D(\epsilon_F)} \delta(\epsilon - \epsilon_F) + \delta'(\epsilon - \epsilon_F)\right] + \dots \end{aligned}$$

We can use this expansion in our potential expression:

$$\begin{aligned} \Omega(T, \mu) &= - \int d\epsilon W(\epsilon) \mathcal{F}_+(\epsilon - \mu) \\ &= - \int_{-\infty}^{\mu} d\epsilon' W(\epsilon') - \frac{\pi^2}{6}(k_B T)^2 D(\mu) + \mathcal{O}(T^4) \end{aligned}$$

$$\begin{aligned} F(T, N) &= \max_{\mu} \{\Omega(T, \mu) - \mu N\} = \Omega(T, \mu(T, N)) \\ &= - \int_{-\infty}^{\mu} d\epsilon' W(\epsilon') - \frac{\pi^2}{6}(k_B T)^2 D(\mu) + N\mu \\ &= - \int_{-\infty}^{\epsilon_F - \frac{\pi^2}{6}(k_B T)^2 \frac{D'}{D}} d\epsilon W(\epsilon) - \frac{\pi^2}{6}(k_B T)^2 D(\epsilon_F) + N \left(\epsilon_F - \frac{\pi^2}{6}(k_B T)^2 \frac{D'}{D} \dots \right) \dots \\ &= - \int_{-\infty}^{\epsilon_F} d\epsilon W(\epsilon) - \left(-\frac{\pi^2}{6}(k_B T)^2 \frac{D'}{D}\right) W(\epsilon_F) - \frac{\pi^2}{6}(k_B T)^2 D(\epsilon_F) - N\epsilon_F - \frac{\pi^2}{6}(k_B T)^2 \frac{D'}{D} N + \dots \\ &= - \int_{-\infty}^{\epsilon_F} d\epsilon \int_{-\infty}^{\epsilon} d\epsilon' D(\epsilon') - \frac{\pi^2}{6}(k_B T)^2 D(\epsilon_F) + N\epsilon_F + \dots \end{aligned}$$

$$\begin{aligned}
&= - \int_{-\infty}^{\epsilon_F} d\epsilon' \int_{\epsilon'}^{\epsilon_F} D(\epsilon') + (\text{the other terms}) \\
&= - \int_{-\infty}^{\epsilon_F} d\epsilon' D(\epsilon')(\epsilon_F - \epsilon') + (\text{the other terms}) \\
&= -\epsilon_F N + U_0 + (\text{the other terms})
\end{aligned}$$

so

$$\begin{aligned}
F(T, N) &= U_0 - \frac{\pi^2}{6}(k_B T)^2 D(\epsilon_F) + \mathcal{O}(T^4) \\
S(T, N) &= -\frac{\partial F(T, N)}{\partial T} = \frac{\pi^2}{3} k_B^2 T D(\epsilon_F) + \mathcal{O}(T^3) \\
U(T, N) &= F(T, N) + TS(T, N) = U_0 + \frac{\pi^2}{6}(k_B T)^2 D(\epsilon_F) + \mathcal{O}(T^4) \\
c_V(T, N) &= \frac{\partial U}{\partial T} = \frac{\pi^2}{3} k_B^2 T D(\epsilon_F) + \mathcal{O}(T^3)
\end{aligned}$$

This is linear in T !

Conclusion

This concludes the lecture series taught by Dr. Markus Deserno during the Spring of 2020. I hope these notes will prove useful to current and future students of the course. Please note that by lecture number, I am missing lectures 30 and 32, but there are a few more classes which I missed but for which I forgot to update the lecture number. These classes are mostly at the end of the course (and the missing dates will be clear from the headers of each lecture). I believe the content covered on those missing days was mostly on blackbody radiation. I cannot guarantee the accuracy of these notes, since they were all typed (quickly) during class with minimal proofreading afterwards.

The textbook used for this class was “An Introduction to Statistical Mechanics and Thermodynamics” by Robert H. Swendsen.

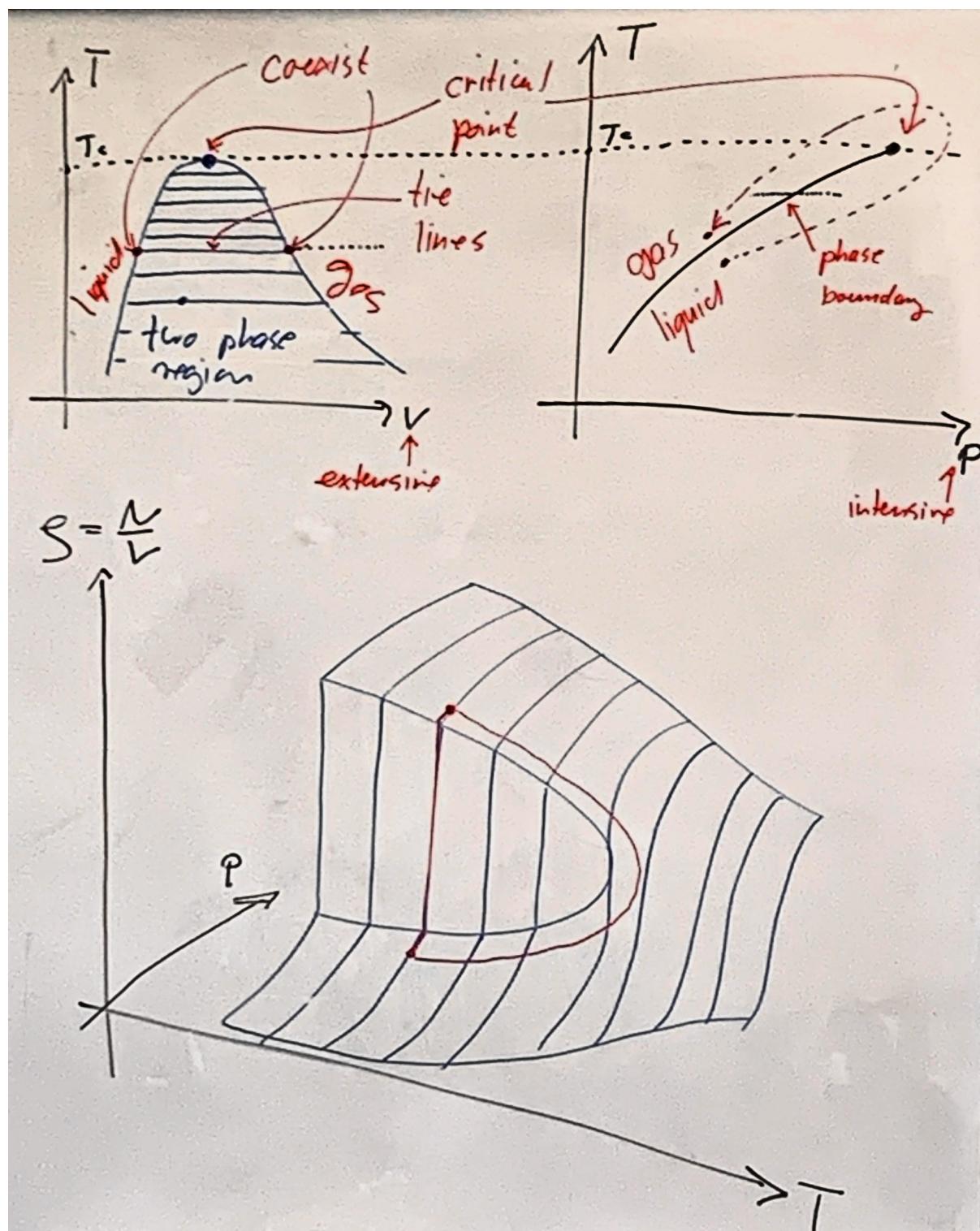


Figure 0.25.1: A typical phase diagram

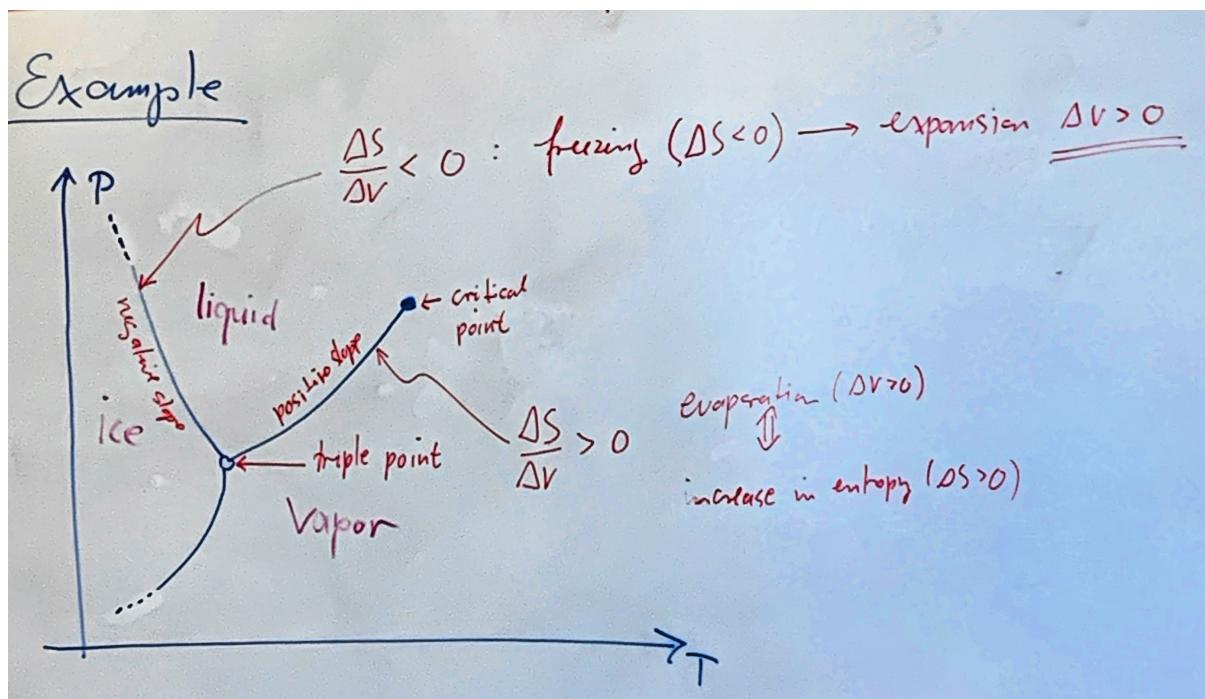


Figure 0.25.2: Phase Diagram of Water



Figure 0.28.1: Plot of an Isotherm through the Critical Point

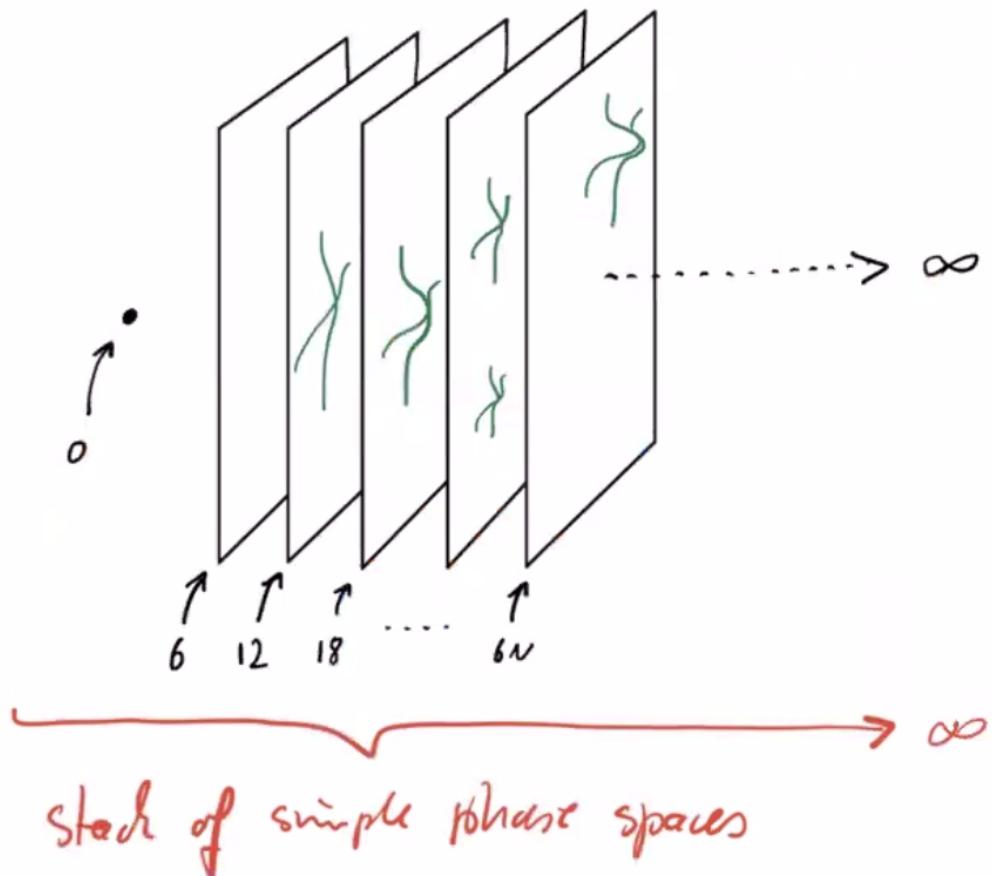


Figure 0.28.2: The Phase Space of the Grand Canonical Ensemble

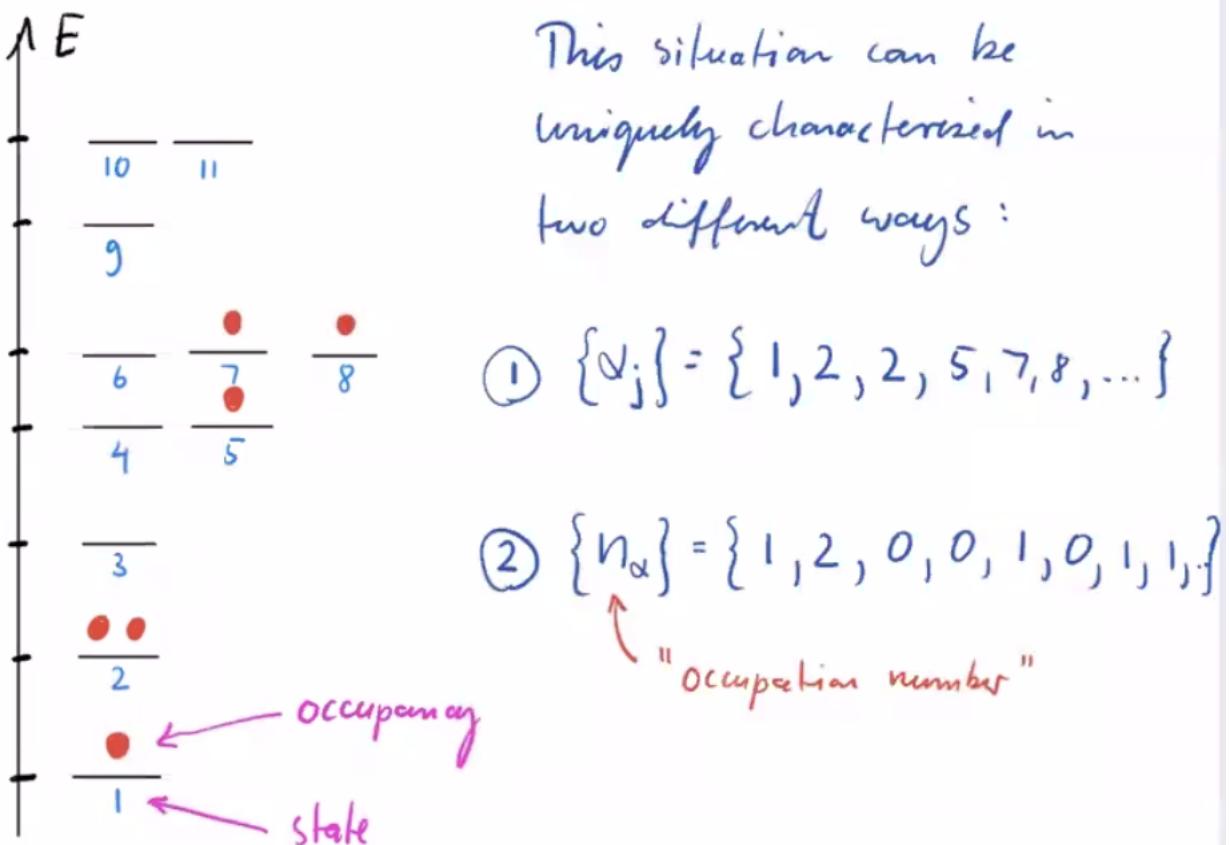


Figure 0.32.1: Using Occupancy Number to label states