Matthew Schwartz
Statistical Mechanics, Spring 2021

Lecture 7: Ensembles

1 Introduction

In statistical mechanics, we study the possible microstates of a system. We never know exactly which microstate the system is in. Nor do we care. We are interested only in the behavior of a system based on the possible microstates it could be, that share some macroscopic proporty (like volume V, energy E, or number of particles N). The possible microstates a system could be in are known as the **ensemble** of states for a system. There are different kinds of ensembles.

So far, we have been counting microstates with a fixed number of particles N and a fixed total energy E. We defined Ω as the total number microstates for a system. That is

$$\Omega(E,V,N)=\sum_{\substack{\text{microstates k}\\ \text{with same N,V,E}}} 1 \tag{1}$$

Then $S=k_B \ln\Omega$ is the entropy, and all other thermodynamic quantities follow from S. For an isolated system with N fixed and E fixed the ensemble is known as the **microcanonical ensemble**. In the microcanonical ensemble, the temperature is a derived quantity, with $\frac{1}{T}=\frac{\partial S}{\partial E}$. So far, we have only been using the microcanonical ensemble.

For example, a gas of identical monatomic particles has $\Omega(E,V,N)\sim\frac{1}{N!}V^N E^{\frac{3}{2}N}$. From this we computed the entropy $S=k_B\ln\Omega$ which at large N reduces to the Sackur-Tetrode formula. The temperature is $\frac{1}{T}=\frac{\partial S}{\partial E}=\frac{3}{2}\frac{Nk_B}{E}$ so that $E=\frac{3}{2}Nk_BT$. Also in the microcanonical ensemble we observed that the number of states for which the energy of one degree of freedom is fixed to ε_i is $\Omega(E-\varepsilon_i)$. Thus the probability of such a state is $P_i=\frac{\Omega(E-\varepsilon_i)}{\Omega(E)}\sim e^{-\varepsilon_i/k_BT}$. This is the Boltzmann distribution.

Within the context of the microcanonical ensemble, we also derived the Boltzmann distribution using the principle of maximum entropy. This approach is very general. It uses nothing about the system other than that the total number of degrees of freedom N is large and the total energy is E. To use the maximum entropy principle we counted the number of ways that N particles could be allocated into groups of size n_i with energies ε_i, so that $\sum n_i=N$ and $\sum n_i\varepsilon_i=E$. We found that in the most probable allocation of particles to groups, the probability of finding a particle with energy ε_i was

$$P_i=\frac{1}{Z}e^{-\beta\varepsilon_i} \tag{2}$$

where $Z=\sum_i e^{-\beta\varepsilon_i}$ and $\beta=\frac{1}{k_BT}$.

Sometimes we don't know the total energy, but we know the temperature. This situation is in fact much more common than knowing the energy. For example, in the room you are in, what is the energy of the air molecules? I bet you don't have a clue. But I bet you have a good idea of what the temperature is.

When we fix temperature instead of energy, we have to allow the energy to fluctuate. For example, think of two systems in thermal contact. The thermal contact allows energy to flow in and out of each system, so energy of each system is not fixed. We call a set of microstates with N, V and T fixed but variable E the **canonical ensemble**. In the canonical ensemble, the primary object is not the number of state Ω or the entropy S but rather the **partition function**

$$Z(\beta)=\sum_{\substack{\text{microstates k}\\ \text{with same N,V}}} e^{-\beta E_k} \tag{3}$$

In the partition function, energies of the microstates summed over can vary. Thus the left-hand side, $Z(\beta)$, cannot depend on energy. Instead, it depends on temperature. Once Z is known, it is straightforward to compute the average energy $\langle E\rangle$ and other thermodynamic quantities, as we will see.

In both the microcanonical and canonical ensembles, we fix the volume. We could instead let the volume vary and sum over possible volumes. Allowing the volume to vary gives the **Gibbs ensemble**. In the Gibbs ensemble, the partition function depends on pressure rather than volume, just as the canonical ensemble depended on temperature rather than energy.

In the microcanonical, canonical, and Gibbs ensembles, the number of particles N in the system is fixed. In some situations, we want the number of particles to vary. For example, chemical reactions change the number of each molecule type. So in chemistry we can't fix N. Instead we fix something called the chemical potential, μ. Chemical potential is like a pressure for particle number. Chemical potential is a very important concept, but very difficult to grasp, so we will spend a lot of time understanding it in this lecture and beyond.

When N can vary we use the **grand canonical ensemble**. The main object of interest in the grand canonical ensemble is the **grand partition function**

$$\mathcal{Z} = \sum_{\substack{\text{microstates}k\\ \text{withsame}V}} e^{-\beta E_k} e^{\beta \mu N_k} \tag{4}$$

The grand canonical ensemble is used in chemistry, quantum statistical mechanics, and much of condensed matter physics.

2 Canonical ensemble

In the microcanonical ensemble, we calculated properties of its system by counting the number of microstates at fixed energy. Then, for example, temperature is a derived quantity, $\frac{1}{k_B T} = \frac{\partial \ln \Omega}{\partial E}$. In the canonical ensemble, we fix the temperature T, and the (average) energy becomes the derived quantity.

In order to fix the temperature, it is a useful conceptual trick to imagine our system of interest in thermal contact with a heat reservoir. This means the system and heat reservoir can exchange energy through heat, but no work can be done by the system on the reservoir or vice versa. The point of the reservoir is to make concrete the idea of fixing the temperature and letting the energy fluctuate.

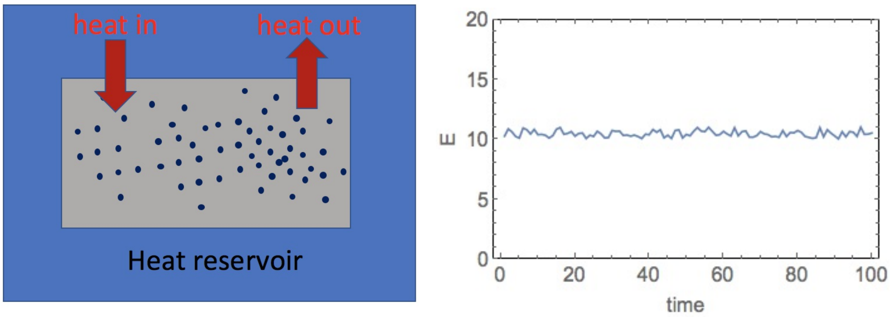


Figure 1. When a system is in thermal contact with a heat reservoir, its temperature is fixed. Its energy fluctuates around its average value.

We do not allow particles to go from the system to the reservoir, only energy. The number of particles in the system can be small – we can have a single atom even – it won't matter. This is important because the canonical ensemble will allow us to discuss systems with a limited number of quantum states, in contrast to the microcanonical ensemble where we really did need to expand at large N to make progress. Although the system can be small, the reservoir does need to be large, so that it has much much more energy than the system. But this is not a constraint, just a conceptual trick, since the reservoir does not actually need to exist. As we will see, the actual computation of the partition function for the canonical ensemble does not involve the reservoir at all.

We would like to know what is the probability of finding the system in a **fixed microstate *k*** with energy E_k? To be clear: every momentum and position of every particle in k is fixed.

Since the system + reservoir is a closed system, the total energy of the system + reservoir is fixed at E_{tot}. Since we have fixed the microstate k of the system, the total number of states is determined only by properties of the reservoir. More precisely, the probability of finding the system in microstate k is proportional to the number of ways of configuring the system + reservoir with the system in microstate k. Since the total energy is fixed, this number is the same as the number of ways of configuring the reservoir with energy $E_{\text{res}} = E_{\text{tot}} - E_k$:

$$P_k = P_{\text{res}}(E_{\text{res}}) = C \times \Omega_{\text{res}}(E_{\text{res}}) = C \times \Omega_{\text{res}}(E_{\text{tot}} - E_k) \tag{5}$$

for some constant C. $\Omega_{\text{res}}(E_{\text{res}})$ is the number of microstates of the reservoir with energy E_{res}.

Now let us use the fact that $E_k \ll E_{\text{res}} \approx E_{\text{tot}}$, which comes from our assumption of a heat reservoir. We can then expand the logarithm of the number of reservoir states around $E_k = 0$:

$$\ln\Omega_{\text{res}}(E_{\text{tot}} - E_k) = \ln\Omega_{\text{res}}(E_{\text{tot}}) - E_k \frac{\partial \ln\Omega_{\text{res}}(E)}{\partial E}\bigg|_{E = E_{\text{tot}}} + \cdots \tag{6}$$

Next we can use that $\frac{\partial \ln\Omega_{\text{res}}(E)}{\partial E} = \beta$ in equilibrium[1], so

$$\ln\Omega_{\text{res}}(E_{\text{tot}} - E_k) = \ln\Omega_{\text{res}}(E_{\text{tot}}) - \beta E_k \tag{7}$$

Exponentiating both sides gives

$$\Omega_{\text{res}}(E_{\text{tot}} - E_k) = \Omega_{\text{res}}(E_{\text{tot}}) e^{-\beta E_k} \tag{8}$$

Then by Eq. (5) we have

$$P_k = \frac{1}{Z} e^{-\beta E_k} \tag{9}$$

for some constant $Z = \frac{1}{C \times \Omega_{\text{res}}(E_{\text{tot}})}$. In the canonical ensemble, we will compute Z not using Ω_{res} but by the shortcut that the probabilities sum to 1, $\sum P_k = 1$.

The formula for P_k we found in Eq (9) is the Boltzmann distribution. Note how much quicker the derivation of the Boltzmann distribution is in the canonical ensemble than in the microcanonical ensemble. In the microcanonical ensemble, we had to count all the states, take the logarithm, expand at large N, express E in terms of T, expand for small ε and simplify. Alternatively, we could use the maximum entropy principle, which still required us to split N particles into m groups, work out the combinatoric factors, take N large, insert Lagrange multipliers, then maximize entropy. In the canonical ensemble, we just hook the system up to a reservoir then "bam!" out pops Boltzmann.

The constant Z is called the **partition function**. Using $\sum P_k = 1$, we find

$$Z = \sum_{\text{microstates } k} e^{-\beta E_k} = \sum_{\text{energies } i} g_i e^{-\beta E_i} \tag{10}$$

In the first sum, we sum over all microstates k, in the second sum, we sum over all possible energies E_i and weight the sum by the **degeneracy** g_i, i.e. the number of microstates with energy E_i.

If the energies are continuous, we write

$$Z = \int g(E) dE e^{-\beta E} \tag{11}$$

where $g(E)$ is called the **density of states**: $g(E)dE$ gives the number of states with energies between E and $E + dE$. The set of energies of a system along with the density of states is called the **spectrum** of a theory.

---

[1] This requires taking the thermodynamic limit (large N) for the reservoir. We do not have to take large N for the system since we are purposefully avoiding ever using Ω_{sys}.

The partition function is an amazingly powerful object. If we know it exactly, we can calculate any thermodynamic property of the system. For example,

$$\langle E\rangle = \sum_k E_k P_k = \frac{1}{Z}\sum_k E_k e^{-\beta E_k} = \frac{1}{Z}\left[-\partial_\beta \sum_k e^{-\beta E_k}\right] = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} \tag{12}$$

So

$$\boxed{\langle E\rangle = -\frac{\partial \mathrm{lnZ}}{\partial \beta}} \tag{13}$$

Thus, knowing the partition function, we can get the expected value for the energy of the system by simply differentiating.

How do we extract the entropy from the partition function. The easiest way is using the Gibbs entropy:

$$S = -k_B \sum_k P_k \ln P_k = -k_B \sum_k \frac{e^{-\beta E_k}}{Z} \ln \frac{e^{-\beta E_k}}{Z} = k_B \sum_k \frac{e^{-\beta E_k}}{Z}(\beta E_k + \ln Z) \tag{14}$$

The first term on the right is the average energy times $k_B \beta = \frac{1}{T}$. For the second term we just use $\sum \frac{1}{Z} e^{-\beta E_k} = \sum P_n = 1$. Therefore

$$S = \frac{\langle E\rangle}{T} + k_B \ln Z \tag{15}$$

We can also write this as $-k_B T \ln Z = \langle E\rangle - TS$. Thus $\ln Z$ is equivalent to the free energy $F = E - TS$. Thinking of the partition function as the free energy will eventually help us understand how to think about it physically. However, we can't do everything all at once. We'll study free energy Lecture 8. In this lecture, we will see how to use the canonical ensemble to calculate Z and how to use Z to derive thermodynamic properties of a system.

An important point about the canonical ensemble is that we use it to derive results about the system only, independent of how it is kept at finite temperature. The partition function is a sum over microstates of the system. P_k is the probability of finding the system in microstate k when it is in equilibrium at a temperature T *no matter what it is in contact with*. We need it to be in contact with something to exchange energy and keep it at finite temperature, but the details of those surroundings are totally irrelevant (except for the temperature). You can see this directly since the surroundings do not appear in the definition of Z.

The reason the canonical ensemble is powerful is because it usually easier to use than the microcanonical ensemble. The essential reason for this is that the microcanonical ensemble involves a constrained sum ($S = k_B \ln\Omega$ is constrained to only include microstates k with $E_k = E$), while the sum in the canonical ensemble is unconstrained. We saw in Section 6 of Lecture 4 how to turn a constrained system into an unconstrained system with an extra variable, in this case β, using Lagrange multipliers. So we replace $\ln\,\Omega(E)$ which is constained at energy E, with $\ln\, Z(\beta) = \ln\Omega - \beta\langle E\rangle$ (i.e. Eq. (15)). Then setting $\frac{\partial \mathrm{lnZ} - E\beta}{\partial \beta} = \langle E\rangle - E = 0$ enforces the Lagrange multiplier constraint. Indeed, the partition function Z is exactly the object we called "Z" in Lecture 4.

Note that we write $\langle E\rangle$ for the expected value of energy, rather than E since $\langle E\rangle$ is calculated rather than fixed from the beginning. The thing $\langle E\rangle$ we compute in Eq. (13) is a function of β, so $\langle E\rangle$ is a derived quantity rather than one fixed from the start. In a real system connected to a heat bath, the temperature would be fixed, but the energy could fluctuate in time around $\langle E\rangle$ as little bits of energy flow in and out of the heat bath. The canonicle ensemble of states is a much bigger set than the microcanonical ensemble – any possible state with any possible energy is included. If the actual system is isolated so its energy does not fluctuate, then we can simply impose the constraint $\langle E\rangle = E$ (equivaletly, we impose the Lagrange-multiplier partial derivative condition). This takes us from an unconstrained canonical-ensemble system to a constrained microcanonical-ensemble system. In fact, this is mostly how we will use the canonical ensemble, to compute equilibrium properties of an isolated system. In such cases, we use $\langle E\rangle$ and E interchangeably, and we can use the extra constraint ($\langle E\rangle = E$) to solve for a relation between T and E. The same relation between E and T can be derived from the microcanonical ensemble or from the canonical ensemble (as we will check when we can).

3 Example 1: monatomic ideal gas

For an ideal monatomic gas with positions q and momenta p, the energy depends only on momenta $E=\sum_j \frac{\vec{p}^2}{2m}$. So

$$Z\approx\int\frac{d^{3N}q d^{3N}p}{(\Delta q)^{3N}(\Delta p)^{3N}}\exp\left[-\beta\sum_j\frac{\vec{p}_j^2}{2m}\right] \tag{16}$$

Here Δp and Δq are the size of phase space regions that we consider minimal. Classical mechanics gives no indication of what we should take for Δq and Δp, and no results that we derive will depend on our choices. As mentioned before, in quantum mechanics, we know to set $\Delta q\Delta p=h$ (see Lecture 10) so let's take this value. Also, recall that for entropy to be extrinsic, we have to count any state in which the same positions and momenta are occupied as the same state. Thus we need to divide the integration by $N!$ for identical particles. This gives

$$Z=\frac{1}{N!}\int\frac{d^{3N}q d^{3N}p}{h^{3N}}\exp\left[-\beta\sum_j\frac{\vec{p}_j^2}{2m}\right] \tag{17}$$

The q integrals trivially give a factor of V^N. The p integrals are the product of $3N$ Gaussian integrals. Each one gives

$$\int_{-\infty}^{\infty}dp e^{-\beta\frac{p^2}{2m}}=\sqrt{\frac{2\pi m}{\beta}} \tag{18}$$

So that

$$Z=\frac{1}{N!}\left(\frac{V}{h^3}\right)^N\left(\frac{2\pi m}{\beta}\right)^{\frac{3}{2}N} \tag{19}$$

Mostly we are interested in this at large N, where $N!\to e^{-N}N^N$ gives

$$\boxed{Z_{\text{monatomic gas}}=e^N\left(\frac{V}{Nh^3}\right)^N\left(\frac{2\pi m}{\beta}\right)^{\frac{3}{2}N}} \tag{20}$$

Once we have Z it is easy to compute the (average) energy:

$$\langle E\rangle=-\frac{\partial \ln Z}{\partial\beta}=-\frac{3}{2}N\frac{\partial}{\partial\beta}\ln\left(\frac{2\pi m}{\beta}\right)=\frac{3}{2}Nk_BT \tag{21}$$

For an isolated system, we then set $\langle E\rangle=E$. This is then in agreement with the result from the equipartition theory (the 3 kinetic degrees of freedom each get $\frac{1}{2}k_BT$ of energy per molecule on average).

Note that this analysis of the ideal gas in the canonical ensemble was a much easier way to compute the average energy than in the microcanonical ensemble, where we had to look at the surface area of a $3N$-dimensional sphere.

3.1 Heat capacity and entropy

Recall that the heat capacity C_V is the amount of heat required to change the temperature at constant volume: $C_V=\left(\frac{Q}{\Delta T}\right)_V=\left(\frac{\partial E}{\partial T}\right)_V$. Recalling that $\beta=\frac{1}{k_BT}$ we have

$$C_V=\frac{\partial\langle E\rangle}{\partial T}=\frac{\partial\langle E\rangle}{\partial\beta}\frac{\partial\beta}{\partial T}=-\frac{1}{k_BT^2}\frac{\partial}{\partial\beta}\left[-\frac{\partial \ln Z}{\partial\beta}\right]=\frac{1}{k_BT^2}\frac{\partial^2\ln Z}{\partial\beta^2} \tag{22}$$

This equation lets us compute the heat capacity directly from the partition function.

Let's check for the monatomic ideal gas. Using Eq. (20) we find that

$$C_V = \frac{1}{k_B T^2}\frac{\partial^2}{\partial\beta^2}\ln\left(\beta^{-\frac{3}{2}N}\right) = \frac{3}{2}N\frac{1}{\beta^2 k_B T^2} = \frac{3}{2}N k_B \tag{23}$$

in agreement with our previous results.[2]

Plugging the partition function for the monatomic ideal gas, Eq. (20) into Eq. (15) for the entropy, we get

$$S = \frac{3}{2}N k_B + k_B \ln\left[e^N\left(\frac{V}{Nh^3}\right)^N\left(\frac{2\pi m}{\beta}\right)^{\frac{3}{2}N}\right] \tag{26}$$

which reduces to:

$$S = N k_B\left[\ln\frac{V}{Nh^3} + \frac{3}{2}\ln[2\pi m k_B T] + \frac{5}{2}\right] \tag{27}$$

Substituting $T = \frac{2}{3Nk_B}E$ this gives back the Sackur-Tetrode equation that we computed with the microcanonical ensemble.

4 Example 2: vibrational modes

Let's work out the canonical ensemble for another system, the vibrational modes of a diatomic molecule. For a diatomic molecule, motion along the axis of the molecule is governed by a potential $V(x)$. The equilibrium position x_0 is where the force vanishes: $F = -V'(x_0) = 0$. Expanding the potential near its minimum (the equilibrium position), $V(x) = V(x_0) + \frac{1}{2}(x - x_0)^2 V''(x_0) + \cdots$ we see that for small deviations from equilibrium, the potential is quadratic. Thus for small displacements it is going to be well modelled by a simple harmonic oscillator with spring constant $k = V''(x_0)$. The oscillation frequency is $\omega = \sqrt{\frac{k}{m}}$.

I assume you studied the quantum mechanics of a simple harmonic oscillator in your QM course. The oscillator has Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{28}$$

The energy eigenstates are

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) \tag{29}$$

where $H_n(z) = (-1)^n e^{z^2}\frac{d^n}{dz^n}(e^{-z^2})$ are the Hermite polynomials. You can check that

$$\left(-\frac{\hbar^2}{2m}\partial_x^2 + \frac{1}{2}m w^2 x^2\right)\psi_n = E_n \psi_n \tag{30}$$

---

[2] It's interesting to write the calculation another way. Note that

$$\frac{\partial}{\partial\beta}\left[-\frac{\partial \ln Z}{\partial\beta}\right] = \frac{\partial}{\partial\beta}\left[\frac{1}{Z}\sum E e^{-\beta E}\right] = -\frac{1}{Z^2}\left(\frac{\partial Z}{\partial\beta}\right)\sum E e^{-\beta E} - \frac{1}{Z}\sum E^2 e^{-\beta E} \tag{24}$$

using that $-\frac{1}{Z}\frac{\partial Z}{\partial\beta} = \langle E\rangle$ we see that this is $-\langle E^2\rangle$. Thus,

$$C_V = -\frac{1}{k_B T^2}\frac{\partial}{\partial\beta}\left[-\frac{\partial \ln Z}{\partial\beta}\right] = \frac{\langle E^2\rangle - \langle E\rangle^2}{k_B T^2} \tag{25}$$

In other words, the heat capacity is given by the RMS energy fluctuations. This tells us that how a system changes when heated can be determined from properties of the system in equilibrium (the RMS energy fluctuations). In other words, to measure the heat capacity, we do not ever have to actually heat up the system. Instead, we can let the system heat up itself through thermal fluctuations away from the mean. This is a special case of a very general and powerful result in statistical physics known as the **fluctuation dissipation theorem**. Another example was our computation of how the drag coefficient in a viscous fluid related to the fluctuations determined by random walks (Brownian motion): if you drag something, the energy dissipates in the same way that statistical fluctuations dissipate.

where

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, \cdots \tag{31}$$

So a harmonic oscillator at rest has energy $E_0 = \frac{1}{2}\hbar\omega$. Each successive mode has $\hbar\omega$ more energy than the previous mode.

Note that for the simple harmonic oscillator, there is only one degree of freedom, so $N = 1$. If we fix the energy E, then we know the exact state of the system, $\Omega = 1$, and there is no ensemble to work with. Thus the microcanonical ensemble is not much use: it doesn't let us answer any questions we would like to ask. For example, we want to know what the typical energy in a vibrational mode is at fixed temperature? If we fix the energy ahead of time, we obviously can't answer this question.

So let us work in the canonical ensemble and compute the partition function for the system. We need to evaluate

$$Z_{\text{vib}} = \sum_n e^{-\beta E_n} = e^{-\frac{\beta}{2}\hbar\omega} \sum_{n=0}^{\infty} e^{-n\beta\hbar\omega} = \frac{1}{2\sinh\left(\frac{\beta}{2}\hbar\omega\right)} \tag{32}$$

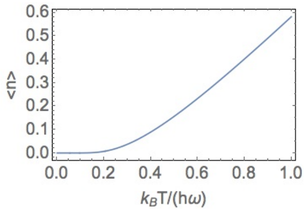
where $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$ is the hyperbolic sine function. In the last step, we have performed the sum over n using $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ with $x = e^{-\beta\hbar\omega}$ and simplified.[3]

With the exact partition function known, we can start computing things. The energy is

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} = -\frac{\hbar\omega}{2}\coth\left(\frac{\beta}{2}\hbar\omega\right) = \hbar\omega\left(\frac{1}{e^{\beta\hbar\omega} - 1} + \frac{1}{2}\right) \tag{33}$$

(Feel free to use mathematica to take derivatives or do it yourself by writing the expression in terms of exponentials and using the chain rule.) Comparing to Eq. (31) we see that the average excitation number is

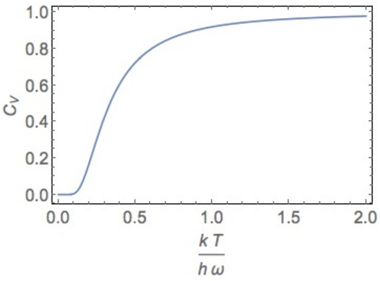
$$\langle n \rangle = \frac{1}{e^{\frac{\hbar\omega}{k_B T}} - 1} = \tag{34}$$



for $k_B T \lesssim \hbar\omega$, $\langle n \rangle \approx 0$ and only the ground state is occupied and from (33), the energy flatlines at its zero point: $E_0 = \frac{1}{2}\hbar\omega$. At higher temperatures, the $\langle E \rangle$ and $\langle n \rangle$ grow linearly with the temperature.

The heat capacity is

$$C_V = \frac{\partial E}{\partial T} = k_B\left(\frac{\hbar\omega}{k_B T}\right)^2 \frac{e^{-\frac{\hbar\omega}{k_B T}}}{\left(1 - e^{-\frac{\hbar\omega}{k_B T}}\right)^2} = \tag{35}$$



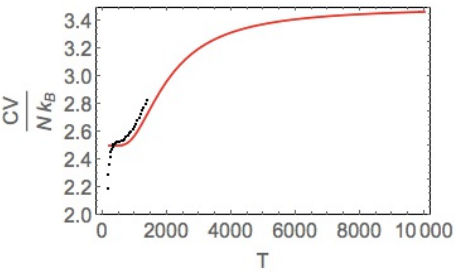
Note that heat capacity is very small until the first energy state can be excited, then it grows linearly.

---

[3] More explicitly, we define $x = e^{-\beta\hbar\omega}$ so that $e^{-n\beta\hbar\omega} = x^n$ and $e^{-\frac{\beta}{2}\hbar\omega} = \sqrt{x}$. Then $Z_{\text{vib}} = \sqrt{x}\sum_{n=0}^{\infty} x^n = \frac{\sqrt{x}}{1-x} = \frac{1}{\frac{1}{\sqrt{x}} - \sqrt{x}} = \frac{1}{e^{\frac{\beta}{2}\hbar\omega} - e^{-\frac{\beta}{2}\hbar\omega}} = \frac{1}{2\sinh\left(\frac{\beta}{2}\hbar\omega\right)}$.

For H_2, the vibrational mode has $\tilde{\nu}_{\text{vib}} = 4342\,\text{cm}^{-1}$ corresponding to $T_{\text{vib}} = \frac{ch\tilde{\nu}}{k_B} = \frac{\hbar\omega}{k_B} = 6300\,K$. So at low energies, the vibrational mode cannot be excited which is why the heat capacity for hydrogen is $C_V = \frac{5}{2}Nk_BT$ rather than $\frac{7}{2}Nk_BT$. We discussed this in Lecture 4, but now we have explained it and can make more precise quantitative predictions of how the heat capacity changes with temperature. Including the kinetic contribution, and a factor of N for the N molecules that can be excited in the vibration mode we see

$$C_V = \frac{5}{2}Nk_B + Nk_B\left(\frac{T_{\text{vib}}}{T}\right)^2 \frac{e^{-\frac{T_{\text{vib}}}{T}}}{\left(1 - e^{-\frac{T_{\text{vib}}}{T}}\right)^2} = \tag{36}$$



This shows how the heat capacity goes up as the vibrational mode starts to be excitable. Note that although the temperature for the vibrational mode is 6300 K, the vibrational mode starts to be excited well below that temperature. The dots are data. We see good agreement! Can you figure out why the heat capacity dies off at low temperature? What do you think explains the small offset of the data from the theory prediction in the plot? We'll eventually produce a calculation in even better agreement with the data, but we need to incorporate quantum indistinguishability to get it right, as we will learn starting in Lecture 10.

We can also compute the partition function for a gas including both kinetic and vibrational motion. Since each momentum and each vibrational excitation is independent for the N particles, we have

$$Z = \frac{1}{N!}\left[\int \frac{d^3q\, d^3p}{h^3} \sum_n e^{-\beta\left(\frac{\vec{p}^2}{2m} + E_n\right)}\right]^N = Z_{\text{monatomic gas}} \times (Z_{\text{vib}})^N \tag{37}$$

This is a very general result: the partition function of a system with multiple possible modes of excitation is the product of the partition functions for the separate excitations. This follows simply from the fact that the total energy is the sum of each excitation energy and the exponential of a sum is the product of the exponentials.

5 Gibbs ensemble

In the microcanonical ensemble, we computed the number of states at a given energy $\Omega(V, N, E)$ and used it to derive the entropy $S(V, N, E) = k_B \ln\Omega(V, N, E)$. In the canonical ensemble, we computed the partition function by summing over Boltzmann factors, $Z(N, V, \beta) = \sum_k e^{-\beta E_k}$. In both cases we have been holding V and N fixed. Now we want to try varying V.

First, let's quickly recall from Section 2 of Lecture 4 why the temperature is the same in any two systems in thermal equilibrium. The quickest way to see this is to recall that entropy is extensive, so a system with energy $E_1 = E$ and another with energy $E_2 = E_{\text{tot}} - E$ has total entropy

$$S_{12}(E_{\text{tot}}, E) = S_1(E) + S_2(E_{\text{tot}} - E) \tag{38}$$

Then the state with maximum entropy is the one where

$$0 = \frac{\partial S_1(E)}{\partial E} + \frac{\partial S_2(E_{\text{tot}} - E)}{\partial E} = \left.\frac{\partial S_1(E_1)}{\partial E_1}\right|_{E_1 = E} - \left.\frac{\partial S_2(E_2)}{\partial E_2}\right|_{E_2 = E_{\text{tot}} - E} = \frac{1}{T_1} - \frac{1}{T_2} \tag{39}$$

where the $\left.\frac{\partial S(x)}{\partial x}\right|_{x = x_0}$ notatino means evaluate the partial derivative at the point $x = x_0$. The minus sign in the second term comes from using the chain rule $\frac{\partial S_2(E_{\text{tot}} - E)}{\partial E} = \frac{\partial E_2}{\partial E}\frac{\partial S_2(E_2)}{\partial E_2} = -\frac{\partial S_2(E_2)}{\partial E_2}$ with $E_2 = E_{\text{tot}} - E$. This is how we showed that $\frac{1}{T} = \frac{\partial S}{\partial E}$ is the same in the two systems in Lecture 4.

Now let's consider an ensemble that lets V vary. This is sometimes called the **Gibbs ensemble**. In the Gibbs ensemble you have two systems in equilibrium that can exchange energy *and* volume. Exchanging volume just means we have a moveable partition in between them. So the total volume is conserved

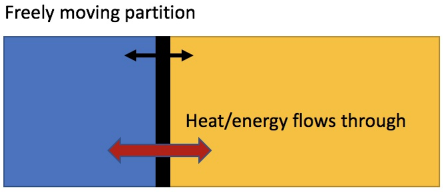


Figure 2. An ensemble where volume is allowed to vary

Now we just apply the same formal argument as in Eqs. (38) and (39): the entropy is the sum of the entropy of the two sides, and the total volume is fixed: $V_1 + V_2 = V_{\text{tot}}$. This implies that

$$S_{12}(E_{\text{tot}}, V_{\text{tot}}, E, V) = S_1(E, V) + S_2(E_{\text{tot}} - E, V_{\text{tot}} - V) \tag{40}$$

And so maximizing entropy, by demanding both the partial derivative with respect to E and the one with respect to V vanish give that the temperature $\frac{1}{T} = \frac{\partial S}{\partial E}$ is the same on both sides (from the E derivative) and that

$$0 = \frac{\partial S_{12}(E_{\text{tot}}, V_{\text{tot}}, E, V)}{\partial V} = \frac{\partial S_1(E, V_1)}{\partial V_1}\bigg|_{V_1 = V} - \frac{\partial S_2(E, V_2)}{\partial V_2}\bigg|_{V_2 = V_{\text{tot}} - V} = \frac{P_1}{T_1} - \frac{P_2}{T_2} \tag{41}$$

In the last step we have *defined* P by $\left(\frac{\partial S}{\partial V}\right)_E = \frac{P}{T}$. The T is added for convenience (so that P will be pressure, as upposed to pressure × temperature). What Eq. (41) shows is that this quantity P is the same for any two systems in equilibrium.

To show that P is the same as what we ordinarily call pressure, all we have to do is compute $\frac{\partial S}{\partial V}$ in some sample system, such as a monatomic ideal gas. Using the entropy of a monatomic ideal gas in the canonical ensemble, Eq. (27), we find

$$\left(\frac{\partial S}{\partial V}\right)_E = \frac{\partial}{\partial V} N k_B \left[\ln\frac{V}{N} + \frac{3}{2}\ln\left(\frac{4\pi m E}{3 N h^2}\right) + \frac{5}{2}\right] = \frac{N k_B}{V} = \frac{P}{T} \tag{42}$$

This explains why we defined P as $\left(\frac{\partial S}{\partial V}\right)_E = \frac{P}{T}$ rather than say $\left(\frac{\partial S}{\partial V}\right)_E = P$.

To summarize, we have established that $\left(\frac{\partial S}{\partial V}\right)_E$ is the same for any two systems in equilibrium. We also already know that $\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_V$ is the same for any two systems in equilibrium. We conclude that the quantity $T\left(\frac{\partial S}{\partial V}\right)_{E,N}$ is the same for any two systems in equilibrium, and give this quantity the name **pressure**:

$$P \equiv T\left(\frac{\partial S}{\partial V}\right)_E \tag{43}$$

There is a unique value for pressure among systems in equilibrium. This is of course consistent with the familiar observation that two gases will equilibrate with equal pressure, since a pressure difference would cause the partition to move. But this way of deriving it, we never had to talk about gas molecules or forces. Using Eq. (43) you can compute the pressure for a solid or photon gas or Bose-Einstein condensate or whatever. For example, one of the important applications of Eq. (43) is the *degeneracy pressure* of a system of identical fermions. Degeneracy pressure is present even at $T = 0$ and is responsible for keeping solids and neutron stars from collapsing. Degeneracy pressure will be studied in lectures 10, 13 and 15.

Note that Eq. (43) works for S computed any way you like it, as $S = k_B \ln\Omega$ in the microcanonical ensemble or $S = \frac{\langle E\rangle}{T} + k_B \ln Z$ in the canonical ensemble. The entropy is the entropy, however you compute it, and the pressure is the pressure.

Now let us consider the total derivative of $S(E,V)$:

$$dS = \left(\frac{\partial S}{\partial E}\right) dE + \left(\frac{\partial S}{\partial V}\right) dV = \frac{1}{T} dE + \frac{P}{T} dV \tag{44}$$

or

$$\boxed{dE = TdS - PdV} \tag{45}$$

This equation is none other than

$$\Delta E = Q - W \tag{46}$$

The change in energy is the heat $Q = TdS$ absorbed minus the work done by changing volume.

You might still be asking, how do we know that the quantity "P" really is "pressure"? Well, we showed that it is for a monatomic ideal gas. And everything in equilibrium has the same "pressure" ($P = \frac{F}{A}$ so if the pressures aren't equal there's a net force, things change, and it's not equilibrium) and the same "P". Thus, by the law of syllogism, it must be "P"= "pressure" for any system. It's the same argument about how we know that T is "temperature".

The Gibbs ensemble is usually just considered a variation on the canonical ensemble. You could in principle try to define a partition function for this ensemble by summing $Z_{\text{GE}} = \sum e^{-\beta E_k - \beta P V_k}$, but then you'd have to be able to compute the volume V_k for a microstate. I don't know of any examples where this is done. The point of the Gibbs ensemble is that thinking of volume varying between systems gives a nice general way to think about pressure, as conjugate to volume and analogous to temperature. That being said, the volume of the actual physical system does not have to have a moveable partition or anything, we're just pretending that it does to do the calculation. The formula $P = T\left(\frac{\partial S}{\partial V}\right)_E$ holds no matter how S is computed, and whether the system is isolated or not.

6 Grand canonical ensemble

Now let's consider systems where the number of particles is not fixed. The basic example for this is chemistry: in a chemical reaction, the number of each molecule species is not conserved. For example, when iron rusts the reaction is

$$3\ \text{Fe} + 4\ \text{H}_2\text{O} \rightarrow \text{Fe}_3\text{O}_4 + 4\ \text{H}_2 \tag{47}$$

Although the number of each type of atom is conserved, the number of each type of *molecule* is not. Other examples with particle non-conservation are the photons that come off of a hot object (blackbody radiation) or radioactive decay, such as $n \rightarrow p^+ + e^-$.

An ensemble where N changes in a system but the total number of particles is conserved among systems is called the **grand canonical ensemble**. Just like in the discussion in the previous section about pressure, the result maximizing entropy implies that

$$\frac{\partial S_1(E,V,N)}{\partial N} = \frac{\partial S_2(E,V,N)}{\partial N} \tag{48}$$

for any two systems in equilibrium that can share "N". As with pressure, we multiply this derivative by T and give this derivative a name

$$\mu \equiv -T\left(\frac{\partial S}{\partial N}\right)_{E,V} \tag{49}$$

This quantity is called the **chemical potential**. In equilibrium, the chemical potential of any two systems that can exchange particles is the same ($\mu_1 = \mu_2$). The minus sign is a convention. It makes the chemical potential negative in most circumstances.

A useful way to think about chemical potential is as a pressure for number density. For example, suppose you have an atom that has two states, a ground state 0 and an excited state 1. In equilibrium, there will be some concentrations $\langle n_0 \rangle$ and $\langle n_1 \rangle$ of the two states, and the two chemical potentials μ_1 and μ_2 will be equal. Since the excited states have more energy, we expect fewer of them, so $\langle n_1 \rangle < \langle n_0 \rangle$. Say we add to the system some more atoms in the ground state. This would push more atoms into the excited state to restore equilibrium. This pushing is due to the "number density pressure" of the chemical potential. Adding to n_0 pushes up μ_0, so $\mu_0 \neq \mu_1$ anymore; the number densities then change until equilibrium is restored.

While there is only one kind of temperature there are lots of chemical potentials: one for every type of conserved N. If there are 4 different types of particles involved, there are 4 chemical potentials. The more general formula is

$$\frac{\partial S_1(E,V,N_1,N_2,\cdots)}{\partial N_1}=-\frac{\mu_1}{T},\qquad \frac{\partial S_1(E,V,N_1,N_2,\cdots)}{\partial N_2}=-\frac{\mu_2}{T},\qquad \cdots \tag{50}$$

You should not think of the chemical potential as being connected to the grand canonical ensemble in any essential way. The chemical potential is property of the system, like pressure or temperature, relevant no matter what statistical system we use to perform the calculation. To see how chemical potential is embedded in the microcanonical ensemble, recall our microcanonical maximum entropy calculation, where we imposed $\Sigma n_i = N$ and $\sum n_i\varepsilon_i = E$ as constraints. Then we maximized entropy by maximizing

$$\frac{S}{k_B}=\ln\Omega=-N\sum_{i=1}^{m} f_i\ln f_i-\alpha\Big(\sum n_i-N\Big)-\beta\Big(\sum n_i\varepsilon_i-E\Big) \tag{51}$$

Since $\frac{\partial\ln\Omega}{\partial E}=\beta$, we identified this Lagrange multiplier β with the usual $\beta=\frac{1}{k_BT}$. Since $\frac{\partial\ln\Omega}{\partial N}=\alpha$ we can now identify $\mu=-\alpha k_BT$ as the chemical potential. Thus given Ω in the microcanonical ensemble, we compute the chemical potential as

$$\mu=-\alpha k_BT=-k_BT\left(\frac{\partial\ln\Omega(E,V,N)}{\partial N}\right)=-T\left(\frac{\partial S}{\partial N}\right)_{E,V} \tag{52}$$

in agreement with Eq. (49).

As in Eq. (45) we can now consider the total derivative of energy, letting E,V and N all vary:

$$dS=\left(\frac{\partial S}{\partial E}\right)dE+\left(\frac{\partial S}{\partial V}\right)dV+\left(\frac{\partial S}{\partial N}\right)dN=\frac{1}{T}dE+\frac{P}{T}dV-\frac{\mu}{T}dN \tag{53}$$

That is,

$$\boxed{dE=TdS-PdV+\mu dN} \tag{54}$$

This implies that

$$\left(\frac{\partial E}{\partial N}\right)_{S,V}=\mu \tag{55}$$

So the chemical potential represents the change in energy when a particle is added at constant V and S. This is *almost* intuitive. Unfortunately the constant-S constraint makes Eq. (55) hard to interpret. Don't worry though, we'll come up with better ways to understand μ in Section 7.

6.1 Grand partition function

As in Section 2 let us now hook a small system up to a reservoir to derive the Boltzmann factor. This time the reservoir should have large energy and large particle number, and both energy and particle number can flow between the system and reservoir. As before, think about picking one microstate k of the system with energy E_k and N_k particles. Once E_k and N are fixed, the total number of microstates is determined only by the states in the reservoir. Eq. (7) becomes

$$\ln\Omega_{\text{res}}(E_{\text{tot}}-E_k,N_{\text{tot}}-N_k)=\ln\Omega_{\text{res}}(E_{\text{tot}},N_{\text{tot}})-\beta E_k+\beta\mu N_k \tag{56}$$

where Eq. (52) was used. This leads to a Boltzmann factor

$$P_k=\frac{1}{\mathcal{Z}}e^{-\beta E_k+\beta\mu N_k} \tag{57}$$

where

$$\mathcal{Z}(V,\beta,\mu)=\sum_k P_k=\sum_k e^{-\beta E_k+\beta\mu N_k} \tag{58}$$

is called the **grand partition function**.

The grand partition function lets us calculate the expected number of particles

$$\langle N\rangle = \sum_k N_k P_k = \frac{1}{\mathcal{Z}}\sum_k N_k e^{-\beta E_k+\beta\mu N_k} = \frac{1}{\beta}\frac{1}{\mathcal{Z}}\frac{\partial\mathcal{Z}}{\partial\mu} = \frac{1}{\beta}\frac{\partial\ln\mathcal{Z}}{\partial\mu} \tag{59}$$

We can also calculate the usual things the partition function lets us calculate, such as the average energy. Taking a derivative with respect to β we get

$$\frac{1}{\mathcal{Z}}\frac{\partial\mathcal{Z}}{\partial\beta} = \sum_k (-E_k+\mu N_k)e^{-\beta E_k+\beta\mu N_k} = -\langle E\rangle + \mu\langle N\rangle \tag{60}$$

so that

$$\langle E\rangle = -\frac{\partial\ln\mathcal{Z}}{\partial\beta} + \frac{\mu}{\beta}\frac{\partial\ln\mathcal{Z}}{\partial\mu} \tag{61}$$

Particle number and chemical potential are conjugate, like pressure and volume. If you know N for a system then you can calculate μ by $\frac{\partial E}{\partial N}$. This is like how if you know the energy for a system, you can calculate temperature from $\frac{1}{T} = \frac{\partial S}{\partial E}$. If you know the chemical potential instead of N, then you can compute average number by $\langle N\rangle = \frac{1}{\beta}\frac{\partial\ln\mathcal{Z}}{\partial\mu}$. This is like how if you know temperature and not the energy, you can compute the average energy form $\langle E\rangle = -\frac{\partial\ln Z}{\partial\beta}$.

Finally, let's compute the entropy, in analogy to Eq. (15). We start with Eq. (14), which goes through to the grand canonical ensemble with $Z\to\mathcal{Z}$ and $\beta E\to\beta(E-\mu N)$:

$$S = k_B\sum \frac{e^{-\beta E_k+\beta\mu N_k}}{\mathcal{Z}}[\beta(E_k-\mu N_k)+\ln\mathcal{Z}] \tag{62}$$

$$= \frac{\langle E\rangle}{T} - \mu\frac{\langle N\rangle}{T} + k_B\ln\mathcal{Z} \tag{63}$$

Thus,

$$\boxed{-k_B T\ln\mathcal{Z} = \langle E\rangle - TS - \mu\langle N\rangle} \tag{64}$$

This will be a useful relation.

7 Chemical potential

One of the most important uses of chemical potential is for phase transitions. In Lecture 9, we"ll see that different phases of matter are characterized by different chemical potentials so that phase boundaries are where the chemical potentials are equal. Unfortunately, we don't have analytical forms for the chemical potential in most real systems, like liquids. So in order to understand phase transitions and other sophisticated uses of chemical potential, we first need to build up some intuition from simple solvable examples.

Remember, μ is independent of which ensemble we use to we calculate it. Although it is a natural object for the grand canonical ensemble, we will work in this section in the microcanonical ensemble since it is somewhat more intuitive. (We'll come back to the grand canonical ensemble and the grand partition function in the next lecture and use it a lot in quantum statistical mechanics.)

7.1 Ideal gas

For a monatomic gas, using the Sackur-Tetrode equation, we find

$$\mu = -T\left.\frac{\partial S}{\partial N}\right|_{E,V} = -T\frac{\partial}{\partial N}\left\{Nk_B\left[\ln\frac{V}{N}+\frac{3}{2}\ln\left(\frac{4\pi mE}{3Nh^2}\right)+\frac{5}{2}\right]\right\} \tag{65}$$

$$= -k_B T\left[\ln\frac{V}{N}+\frac{3}{2}\ln\left(\frac{4\pi mE}{3Nh^2}\right)\right] \tag{66}$$

Note that the $\frac{5}{2}$ has dropped out. Using $E = \frac{3}{2}Nk_BT$ for this gas, we can write this relation in an abbreviated form

$$\mu = k_B T\ln\left[n\left(\frac{h^2}{2\pi m k_B T}\right)^{3/2}\right] = k_B T\ln n\lambda^3 \tag{67}$$

where $n = \frac{N}{V}$ is the number density and

$$\lambda = \frac{h}{\sqrt{2\pi m k_B T}} \tag{68}$$

is called the **thermal de Broglie wavelength** or just **thermal wavelength**. Recall that the de Broglie wavelength is $\lambda = \frac{h}{p}$, with p the momentum. Since the momenta of particles in a gas vary according to the Maxwell-Boltzmann distribution, they have a variety of different de Broglie wavelengths. The thermal de Broglie wavelength is the wavelength of a typical particle, more precisely, one with momentum $p = \sqrt{2\pi m k_B T} = \sqrt{\frac{2\pi}{3}} p_{\text{rms}}$ with p_{rms} the RMS momentum of a gas a temperature T. The de Broglie wavelength is measure of the length scale at which quantum effects become important. If the gas at temperature T is more dense than 1 particle per thermal de Broglie wavelength-cubed, $n > \frac{1}{\lambda^3}$, then quantum statistical mechanics must be used. We will make this correspondance precise in Lecture 10.

To get a feel for typical numbers, consider air at room temperature. The molar mass of air is $29.0 \frac{g}{\text{mol}}$ and density is $\rho = 1.27 \frac{\text{kg}}{m^3}$. So $\lambda = \frac{h}{\sqrt{2\pi m_{N_2} k_B T}} = 1.87 \times 10^{-11} m$ while $n = (3.35 \times 10^{-9} m)^{-3}$. In other words, the typical distance between atoms in air is $d = 3.3\,\text{nm}$ and the thermal de Broglie wavelength is much smaller, $\lambda = 0.02\,\text{nm}$. So in air $n\lambda^3 = 1.7 \times 10^{-7} \ll 1$. This means that $\mu = k_B T \ln n\lambda^3 = -0.39\text{eV}$. Note that the chemical potential of air is negative. Since $\mu = k_B T \ln n\lambda^3$, the chemical potential will always be negative in regions where $n < \frac{1}{\lambda^3}$ and classical statistical mechanics applies.

Solving Eq. (67) for n gives

$$n = \frac{1}{\lambda^3} \exp\left(\frac{\mu}{k_B T}\right) \tag{69}$$

This says that the number density is related exponentially to the chemical potential. Thus if we double the number of particles, $n \to 2n$, the chemical potential goes up by $\mu \to \mu + \ln 2$. As the system gets denser and denser, the chemical potential rises towards 0. When the chemical potential is negative, the gas is not too crowded and can considered dilute, with interparticle interactions ignored (as in a classical ideal gas).

Returning to the analogy between chemical potential like a pressure, suppose you have a system with a concentration gradient. Then the part with more particles will be at higher chemical potential and the part with lower gradient at lower chemical potential, according to Eq. (69). This is why it is called a potential – it is like potential energy, but for particle number.

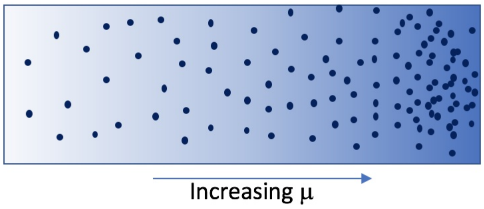


Figure 3. Chemical potential is higher in more dense regions. It is like potential energy for particle number. Particles move from high to low μ, until the μ's are all equal.

7.2 Ground-state energy

In many situations, the lowest energy a particular molecule can have is different for different molecules. There is not a unique way to define the ground state energy; it depends on what we include. For example, we could include chemical bond energy in the ground-state energy, or the rest mass mc^2, or the gravitational potential energy relative to the surface of the earth mgh or relative to the center of the earth $G\frac{Mm}{r}$, and so on. There is always going to be some arbitrariness in the definition of energy. However, once we define the ground-state energy for one thing, we can talk without ambiguity about the ground-state energy of everything.

So let's say that we have established a convention and the ground state energy for a molecule is ε. To be concerete, think of of ε as the chemical-bond energy of the molecule. Then there is a contribution $N\varepsilon$ to the total energy of the N molecules. For a monatomic gas, the total energy is offset from the kinetic energy: $E = E_{\text{kin}} + N\varepsilon$. The ground state energy doesn't affect the number of states, so the functional form of entropy $S(E_{\text{kin}}) = k_B \ln \Omega$ is the same with or without the offset; either way, it counts the number of ways a total amount of kinetic energy is distributed among the momenta of the gas molecules. Since we want to write entropy a function of the total energy, we should the substitute $E_{\text{kin}} = E - N\varepsilon$. Then, in the microcanonical ensemble we have

$$S = N k_B \left[\ln \frac{V}{N} + \frac{3}{2} \ln \left(\frac{4\pi m (E - N\varepsilon)}{3 N h^2} \right) + \frac{5}{2} \right] \tag{70}$$

Of course, if there is just one type of molecule present, we can just choose $\varepsilon = 0$, but including ε explicitly will allow us to discuss systems of molecues with different ground state energies (chemical bond energies, rest masses, etc.)

One can also derive Eq. (70) from the canonical ensemble. Let the partition function without the energy offset (i.e. $\varepsilon = 0$) be Z_0 and so $\langle E_0 \rangle = -\frac{\partial \ln Z_0}{\partial \beta}$ and $S_0(E_0) = \frac{\langle E_0 \rangle}{T} + k_B \ln Z_0$. Then with the energy offset we get $Z = Z_0 e^{-\beta N \varepsilon}$ and so $\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} = \langle E_0 \rangle + N\varepsilon$. Then

$$S = \frac{\langle E \rangle}{T} + k_B \ln Z = \frac{\langle E_0 \rangle}{T} + \frac{N\varepsilon}{T} + k_B \ln Z_0 - k_B \beta N \varepsilon = \frac{\langle E_0 \rangle}{T} + k_B \ln Z_0 = S_0(E_0) = S_0(E - N\varepsilon) \tag{71}$$

Thus using the canonical ensemble we again find that the entropy S with the offset has the functional form as S_0 without the offset; it is only the energy at which we evaluate S that changes. This is in agreement with Eq. (70).

From the entropy, we can compute the chemical potential

$$\mu = -T \left(\frac{\partial S}{\partial N} \right)_{E,V} = k_B T \ln n \lambda^3 + \varepsilon \tag{72}$$

with λ in Eq. (68), so that for an ideal gas with ground-state energy ε

$$n = \frac{1}{\lambda^3} \exp \left(\frac{\mu - \varepsilon}{k_B T} \right) \tag{73}$$

Thus if we were to change the zero-point energy offset $\varepsilon \to \varepsilon + \Delta\varepsilon$, we could compensate for this by shifting $\mu \to \mu + \Delta\varepsilon$. In other words, the chemical potential is measured *relative* to the ground state: only the difference $\mu - \varepsilon$ appears. This is just like only potential energy differences are physical, and why we call the chemical potential a *potential*.

With the energy offset, we can refine our observation about the chemical potential being negative for a classical ideal gas. That observation held when the gas was dilute: $n < \frac{1}{\lambda^3}$, with λ the thermal de Broglie wavelength. Now we see that a more precise statement is that for a classical ideal gas, $\mu - \varepsilon < 0$, i.e. chemical potential is less than the ground state energy ε. Eq. (72), $\mu = k_B T \ln n \lambda^3 + \varepsilon$ says that the chemical potential gets two contributions: one from the density and one from the energy. The density contribution is of entropic origin and depends on how many molecules are in the system. The energetic contribution is due to the internal structure of the molecule and independent of whatever else is going on. Equilibrium, where chemical potentials are equal, comes from a balance between these two contributions. This should become clearer with some examples.

7.3 Chemical reactions

Chemical potentials are useful in situations where particles turn into other types of particles. When there is more than one type of particle in the system (as there typically are when we consider problems involving chemical potentials), we need a different μ for each particle. So Eq. (54) becomes

$$dE = T dS - P dV + \sum \mu_j d N_j \tag{74}$$

As a concrete example, consider the Haber process for the production of ammonia

$$3\ \mathrm{H}_2 + \mathrm{N}_2 \leftrightharpoons 2\ \mathrm{NH}_3 \tag{75}$$

Note that the number of each individual molecule is not conserved, but because the number of hydrogen atoms and nitrogen atoms is conserved, the relative coefficients (3,1 and 2) in Eq. (75) are fixed. In chemistry, the **concentrations** or **molar number densities** of molecule j are denoted as $[j]=\frac{n_j}{N_A}$, with $n_j=\frac{N_j}{V}$ and $N_A=6\times 10^{23}\frac{1}{\mathrm{mol}}$ Avogadro's number. In equilibrium, there will be some relationship among the concentrations $[\mathrm{H}_2]$ of hydrogen, $[\mathrm{N}_2]$ for nitrogen and $[\mathrm{NH}_3]$ for ammonia that we can compute using chemical potentials.

First, we note the reaction Eq. (75) implies that if $[\mathrm{N}_2]$ goes down by one molcule, $[H_2]$ must go down by 3 molecules and $[\mathrm{NH}_3]$ must go up by 2 molecules. That is

$$d[\mathrm{H}_2]=3d[\mathrm{N}_2],\quad d[\mathrm{NH}_3]=-2d[\mathrm{N}_2], \tag{76}$$

As the concentrations change, at fixed volume and fixed total energy, the entropy changes as

$$dS=\frac{\partial S}{\partial[\mathrm{H}_2]}d[\mathrm{H}_2]+\frac{\partial S}{\partial[\mathrm{N}_2]}d[\mathrm{N}_2]+\frac{\partial S}{\partial[\mathrm{NH}_3]}d[\mathrm{NH}_3] \tag{77}$$

Thus, using that $dS=0$ in equilibrium, Eq. (76) and the definiton of the chemical potentials as $\mu_i=\frac{\partial S}{\partial N_i}$, we find[4]

$$3\mu_{H_2}+\mu_{N_2}=2\,\mu_{\mathrm{NH}_3} \tag{78}$$

This constraint among the chemical potentials is a generalization of $\mu_1=\mu_2$ in equilibrium for two systems that can exchange particles. Here there are 3 systems that can exchange particles.

Now, from Eq. (73) we know how to relate the number of particles to the chemical potential for a monatomic ideal gas:

$$[X]=\frac{1}{\lambda^3}\exp\left[-\frac{\varepsilon_X-\mu_X}{k_BT}\right] \tag{79}$$

where ε_X is the ground state energy for molecule X. To get the μ's to drop out, we can take the ratio of concentrations to appropriate powers:

$$\frac{[\mathrm{NH}_3]^2}{[\mathrm{H}_2]^3[\mathrm{N}_2]}\approx\frac{\lambda_{H_2}^9\lambda_{N_2}^3}{\lambda_{\mathrm{NH}_3}^6}\times\exp\left[-\frac{2\varepsilon_{\mathrm{NH}_3}-3\varepsilon_{H_2}-\varepsilon_{N_2}}{k_BT}\right]\underbrace{\exp\left[-\frac{2\,\mu_{\mathrm{NH}_3}-3\mu_{H_2}-\mu_{N_2}}{k_BT}\right]}_{=1} \tag{80}$$

The second exponential is just 1 because of Eq. (78), which is why we chose the powers of $[H_2]$ and $[\mathrm{NH}_3]$ that we did on the left hand side. The $\approx$ means that we are approximating everything as monatomic ideal gases (not a great approximation but it's a start)

The sum of energies is just the net energy change in the reaction, $\Delta\varepsilon$. For the Haber process, which is exothermic, $\Delta\varepsilon=-92.4\,\frac{\mathrm{kJ}}{\mathrm{mol}}$. So

$$\frac{[\mathrm{NH}_3]^2}{[\mathrm{H}_2]^3[\mathrm{N}_2]}\approx\frac{\lambda_{H_2}^9\lambda_{N_2}^3}{\lambda_{\mathrm{NH}_3}^6}\exp\left[-\frac{\Delta\varepsilon}{k_BT}\right]\quad(\text{assuming monatomic gases}) \tag{81}$$

This is special case (for monatomic ideal gases) of the law of mass action. It says that the relative concentrations of reacting molecules in equilibrium are determined by the Boltzmann factor dependent on the change in energy associated with the reaction. This formula arises from a balance between entropic contributions to the chemical potentials on both sides (though their number densities) and energetic contributions (in the exponential factor).

---

[4] Technically speaking, you need a conserved total N to define the chemical potential. Because the numbers of atoms are conserved, there are chemical potentials μ_H and μ_N for them. So what we are doing implicitly above is defining the chemical potentials for the molecules in terms of the atomic chemical potentials $\mu_{H_2}\equiv 2\mu_H$, $\mu_{N_2}\equiv 2\mu_N$ and $\mu_{\mathrm{NH}_3}\equiv\mu_N+3\mu_H$ from which Eq. (78) follows.

We have written explicitly the reminder that this formula assumes that the reactants and products are monatomic gases. This is not a bad assumption in some cases. More generally though, for chemicals reacting, we will need to add corrections to the right hand side. These corrections will be included in the next lecture, where the law of mass action is derived in full.

7.4 Example: matter antimatter asymmetry (optional)

For another example, consider the process of a proton-antiproton annihilation. Antiprotons p^- are anti-particles of protons. They have the same mass as protons but opposite electric charge. Protons and anti-protons can annihilate into photons

$$p^+ + p^- \leftrightharpoons \gamma + \gamma \tag{82}$$

The reverse reaction is photons converting into proton-antiproton pairs. These annihilations and conversions happen constantly when the temperature is well above the threshold energy for pair production

$$k_B T \gg \varepsilon = 2 m_p c^2 = 2\,\mathrm{GeV} \tag{83}$$

We don't care so much about the details of why or how this process occurs, just that it does occur. This threshold temperature is around 2×10^{13} K. So in most systems of physical interest (stars, your body, etc.) this doesn't happen. It did happen however, in the early universe, until 0.01 seconds after the big bang.

Note that while the above reaction conserves the number of protons minus the number of antiprotons, it does not conserve the number of photons. Indeed, other reactions can easily change photon number, such as

$$\gamma + e^- \leftrightharpoons e^- + \gamma + \gamma \tag{84}$$

A more down to earth example is the light in your room – photons are constantly being produced, not absorbed. Eq. (84) implies that

$$\mu_\gamma + \mu_{e^-} = \mu_{e^-} + 2\mu_\gamma \tag{85}$$

In other words, that

$$\mu_\gamma = 0 \tag{86}$$

This is a general property of particles that are not associated with any conserved quantity: their chemical potential vanishes. Then the reaction in Eq. (82) gives

$$\mu_{p^+} + \mu_{p^-} = 0 \tag{87}$$

In addition, it is natural to suppose that all the protons and antiprotons came from processes like $\gamma + \gamma \rightarrow p^+ + p^-$ that produce or remove the same number of protons and antiprotons. This would make the proton/antiproton concentrations equal[5], and their chemical potentials equal too (and hence $\mu_{p^+} = \mu_{p^-} = 0$ by Eq. (87)).

Now, the energy change in the reaction $\gamma + \gamma \rightarrow p^+ + p^-$ is $\Delta\varepsilon = 2 m_p c^2$. Thus, as in Eq. (81), using $T = 3K$ (the temperature of outer space), and treating protons and antiprotons as monatomic ideal gases (an excellent approximation in fact) with a a thermal wavelength $\lambda = \frac{h}{\sqrt{2\pi m k_B (3K)}} \approx 1\,\mathrm{nm}$, we find

$$[p^-] = [p^+] = \frac{1}{\lambda^3} e^{-\frac{\Delta\varepsilon}{2 k_B T}} = \left(\frac{2\pi m_p k_B T}{h}\right)^{3/2} e^{-\frac{m_p c^2}{k_B T}} = 4 \times 10^{-843112945335} \frac{1}{m^3} \approx 0 \tag{88}$$

So this first pass calculation says there shouldn't be any protons *or* antiprotons around at all!

---

[5] Actually, in any unitary quantum field theory the equilibrium concentrations of particles and antiparticles must be the same. This follows from an unbreakable symmetry known as CPT invariance that combines switching particles and antiparticles (C), flipping the particles spins (P) and time-reversal invariance (T).

To refine our calculation, it's important to note that we used equilibrium physics, but since the universe is expanding, equilibrium is not always a good approximation. At some point as the universe expands and cools, the protons and antiprotons become so dilute that they cannot find each other to annihilate. This is called “freeze-out”. The freeze-out temperature is set by when the rate for $p^+p^- \to \gamma\gamma$ is equal to the expansion rate of the universe. The rate for $p^+p^- \to \gamma\gamma$ depends on the proton's scattering cross section, which is approximately its cross-sectional area $\sigma \sim \frac{1}{m_p^2}$, the number density $[p^+]$ and the velocity, which we can take to be given by the Maxwell-Boltzmann average $\langle \frac{1}{2} m_p \vec{v}^2 \rangle = \frac{3}{2} k_B T$. Putting these factors together, the annihilation rate (events per unit time) is:

$$\Gamma_{\text{annihilate}} = n\sigma v = (2\pi m_p k_B T)^{3/2} e^{-\frac{m_p c^2}{k_B T}} \frac{1}{m_p^2} \sqrt{\frac{3k_B T}{m_p}} \tag{89}$$

The expansion rate requires some general relativity. The result is

$$\Gamma_{\text{expansion}} = \frac{k_B^2 T^2}{M_{\text{Pl}}} \tag{90}$$

where $M_{\text{Pl}} = G_N^{-1/2} = 10^{19}$ GeV is Planck's constant. Setting these equal results in a freezeout temperature

$$T_f = 2.4 \times 10^{11} K \tag{91}$$

At this temperature, the proton concentration is not so small:

$$[p^+] = [p^-] \approx \left(\frac{2\pi m_p k_B T_f}{h} \right)^{3/2} e^{-\frac{m_p c^2}{k_B T_f}} = 10^{23} \frac{1}{m^3} \tag{92}$$

However, as the universe continues to expand from T_f down to $3K$ its size scales with temperature so the proton number density gets diluted down to

$$[p^+] = [p^-] \approx 10^{23} \frac{1}{m^3} \left(\frac{3K}{T_f} \right)^3 = 1.68 \times 10^{-10} \frac{1}{m^3} \tag{93}$$

This is the honest-to-goodness prediction of cosmology for the density of protons left over from the big bang.

Unfortunately, the prediction $[p^+] = 10^{-10} \frac{1}{m^2}$ is in stark disagreement with data: the average number density of protons in the universe is actually $[p^+] = 0.26 \frac{1}{m^3}$. This is a problem. In fact, this is one of the great unsolved problems in fundamental physics, called the mystery of **baryogenesis** or the **matter-antimatter asymmetry**.

One possible solution is to set the initial conditions so that $[p^+] \neq [p^-]$ to start with. Once these are set, if all the processes are symmetric in p^+ and p^- then $[p^+] \neq [p^-]$ will persist. Note however, that the universe is currently 10^{26} m wide, and growing. There are 10^{80} more protons than antiprotons in the observable universe today. When the universe is only $10^{-35} m$ across, this would correspond to a shocking number density of $10^{185} \frac{1}{m^3}$. So it would be a little strange to set this enormous asymmetry in the early universe. Moreover, the current cosmological model involves inflation, which produces exponential growth at early times, so whatever initial asymmetry we set would be completely washed away when inflation ends. In other words, it's possible, but would be very unsettling, to solve the baryogenesis problem by tuning the initial conditions.

Another option is to start off symmetric but have processes that are not symmetric between particles and antiparticles. In turns out in the Standard Model of particle physics, there are none: for every way of producing an electron or proton, there is also a way of producing a positron or antiproton with exactly the same rate. In fact, this equality is guaranteed by symmetries (lepton number and baryon number). Moreover, if you made a modification so that the symmetries were violated, then effectively protons could turn into antiprotons. Thus, since protons and antiprotons have the same mass (and value of ε) their chemical potentials would push them towards the same concentrations, which by Eq. (88) is zero. The story is again a little more complicated, since there is inflation, and reheating and the expansion of the universe is not quite quasi-static, and there is actually a super-tiny violation of the symmetry between protons and antiprotons within the Standard Model. Even when you include all these things, it doesn't work, you still get no matter out once the universe cools.

So we are stuck. Why is there so much matter in the universe? Why is there more matter than antimatter? Nobody knows.

8 Partition function and the spectrum (optional)

Some of you may find it illuminating to think about the partition function in a big-picture, more abstract sense (as if it's not abstract enough already!). The following discussion is just included because some students may find it illuminating. It is not required reading for the course.

The partition function is computed by summing over energies. As you probably know, the energies of a system contain a tremendous amount of information. Indeed, in classical mechanics, the energy at a point in phase space is given by the the Hamiltonian function $H(\vec{q}_i, \vec{p}_i, t)$. If you know H, every possible behavior of the system can be determined by solving Hamilton's equations of motion. In quantum mechanics, the same is true: if you know the Hamiltonian operator $\hat{H}(\hat{\vec{q}}_i, \hat{\vec{p}}_i)$, you can determine the time-evolution of the system completely through the Schrödinger equation. Eigenvalues of the Hamiltonian are the energies of the system.

For example, consider the Hamiltonian for interactions among water molecules. We can approximate the Hamiltonian as depending on the distances $R_{\text{ij}} = |\vec{q}_1 - \vec{q}_2|$ between the centers of mass of the molecules. We should find that if two molecules are far away, the energy becomes independent of R_{ij}. If we try to put the molecules on top of each other, it should be impossible, so the energy should blow up. Because of hydrogen bonding, we expect a weak attractive force at intermediate distances with a shallow potential minimum. That is, we expect something roughly like:

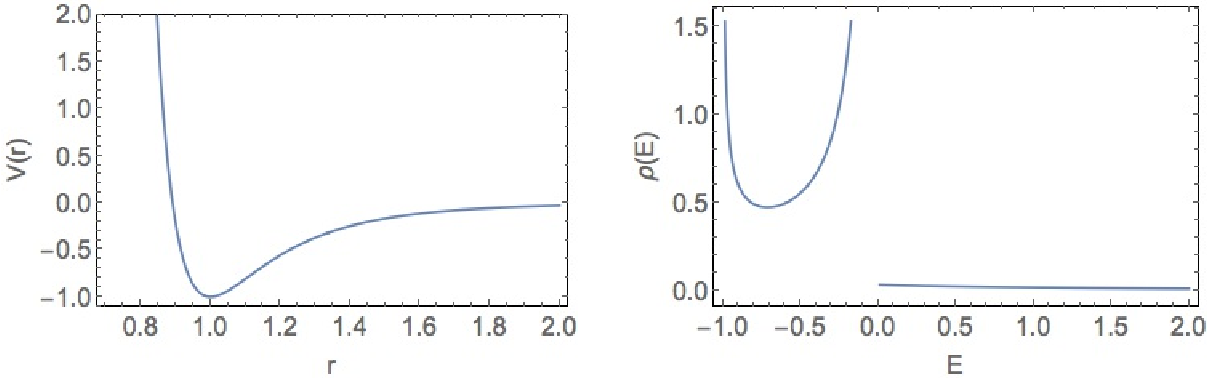


Figure 4. (left) The potential energy between two water molecules is close to a Lennard-Jones potential. The corresponding density of states has singularities at the bound state energy and at zero energy the when the molecules are far apart.

This is called the Lennard-Jones potential. A pairwise potential model like this can explain many the properties of water – surface tension, boiling, freezing, heat capacity, etc. For example, the force from water molecule i is given by $\vec{F} = -\frac{\partial}{\partial \vec{q}_i} H(\vec{q}_i, \vec{p}_i)$. More sophisticated classical models, including 3-body interactions and so on, can explain even more emergent behavior of water. If you know the quantum Hamiltonian $\hat{H}$ exactly, you would be able to determine everything about water exactly (in principle).

If you really want to determine the complete time evolution of a system, you need the full functional form of the classical Hamiltonian, or in a quantum system, the energy eigenvalues ε_i *and* eigenvectors $\psi_i(x)$ of the quantum Hamiltonian. However, you get a long way to understanding the behavior of a system just knowing the spectrum. The density of states for the Lennard Jones potential is shown on the right of Fig. 4. Its singularities indicate the bound state at $E = -1.0$ and the continuum at $E = 0$. The only thing you can't get from the density of states is the bound state distance r_0, since there is no information about position in the density of states (of course, in this case, you could reconstruct the distance from the energy by dimensional analysis). So the spectrum itself has almost all of the information we care about.

In quantum mechanics, the spectrum is the set of eigenvalues. Knowing only the spectrum, we don't have the information contained in the eigenvectors. Most of the time we don't actually care about the eigenvectors. One way to see this is that the eigenvectors are just projections from changing basis, $\psi_i(x) = \langle \varepsilon_i | x \rangle$. In an energy basis, the Hamiltonian is diagonal. Thus if we are interested in basis-independent properties of a system (as we almost always are), the spectrum is sufficient.

The point of the above discussion is to motivate why the spectrum of a system is extremely powerful, and contains almost all the physical information we would care to extract about a system. Now observe that the partition function carries the same information as the spectrum, just represented differently. In fact, the partition function is just the Laplace transform of the spectrum. A Laplace transform is a way of constructing a function $F(\beta)$ from a function $f(E)$ by integrating over E:

$$Z(\beta) = \mathcal{L}[f(E)] = \int_0^\infty dE\, e^{-\beta E} f(E) \tag{94}$$

So the partition function is the Laplace transform of the spectrum $Z = \mathcal{L}[E(q_i, p_i)]$.

Note that a Laplace transform is just a real version of a Fourier transform (take $\beta \to i\beta$). Thus working with the partion function instead of the spectrum is like working in Fourier space, with β representing the frequency. The average over configurations is like the average over time taken to produce a frequency spectrum. Equilibrium corresponds to a particular value of β which is analogous to a particular frequency component dominating, like the resonance on a flute. The shape in Fourier space of the frequency spectrum around a resonance gives the timbre of a note, explaining why a C on a flute sounds different from an C on a trumpet. Thus, in a way, the partition function, through its derivatives at a fixed β, give the timbre of a physical system.

9 Summary

In this lecture, the main conceptual tool in statistical mechanics was introduced: the ensemble. An ensemble is an imaginary collection of systems with different microscopic arrangements and the same fixed macroscopic properties. The properties we consider come in conjugate pairs (E, T), (N, μ) and (V, P), and the ensembles pick one variable from each pair to be independent. The main ensembles are

- The microcanonical ensemble: N, V and E are fixed.
- The canonical ensemble: N, V and T are fixed.
- The grand canonical ensemble: V, T and the chemical potential μ are fixed.
- The Gibbs ensemble (not used much): N, P and T are fixed.

Each ensemble has an natural object that can be used to extract dependent quantities. In the microcanonical ensemble, the object is the number of microstates $\Omega(N, V, E)$ or equivalently the entropy $S = k_B \ln \Omega$. Then

$$\frac{1}{T} = \frac{\partial S}{\partial E}, \quad P = T\frac{\partial S}{SV}, \quad \mu = -T\frac{\partial S}{\partial N} \tag{95}$$

A shortcut to seeing all these relations is the differential relation

$$dE = TdS - PdV + \mu dN \tag{96}$$

In all the ensembles, we just have to *imagine* that the systems are connected to other systems to maintain constant T, P, μ etc. The properties of the systems we calculate hold whether the systems are actually in equilibrium with something else or not. The ensembles as conceptual tools to calculate properties of systems based on known information, independent of their surroundings.

In the canonical ensemble, the natural object is the partition function $Z(T,N,V)=\sum_k e^{-\beta E_k}$ where the sum is over microstates k with E_k their energy. One should think of states in the canonical ensemble as being in thermal equilbrium with something else, like a heat bath. So energy can flow in and out. Since temperature is fixed, not energy, each microstate in the canonical ensemble can have different energy, and the energies can be arbitrarily large. The average of the energies of all the microstates is

$$\langle E\rangle=\frac{1}{Z}\sum_k E_k e^{-\beta E_k}=-\frac{1}{Z}\frac{\partial Z}{\partial\beta}=-\frac{\partial \mathrm{lnZ}}{\partial\beta} \tag{97}$$

This average $\langle E\rangle$ means two things: first, it means we average over all the possible microstates k with all values of E_k, since E can fluctuate due to the thermal contact. Second, it means the average over time of the energy of the actual microstate of the system in thermal contact with a bath. The time average is the ensemble average, due to ergodicity. We can only ever measure the time average, not the ensemble average, but we compute the ensemble average.

In the grand canonical ensemble, the number of particles is not fixed. For example, particles might evaporate, or chemical reactions could occur. The chemical potential μ is the conjugate variable to N, like T is conjugate to E: two systems in equilibrium that can exchange N have the same μ, just like two systems in equilibrium that can exchange E are at the same T.

In the Gibbs ensemble, volume is not fixed, for example, a balloon with gas in it. When two systems in equilibrium can exchange volume, they are at the same pressure. So P is conjugate to V.

The conjugate pairs are not exactly equivalent. Any individual microstate of a system can have N, V and E well defined. The other properties, T,μ,P and S are not defined for a single microstate, only for an ensemble. For example, if we have a gas and specify the positions and momenta of all the molecules, it is a single point in phase space, with no entropy and no temperature. Entropy and temperature are *ensemble* properties. Pressure is an ensemble property too. You might think an isolated system has a pressure. But to measure the pressure, we need to see the force on the walls of the box. This force is microscopically very irregular, so we must average over time to get something sensible. This time average is the same as the ensemble average due to ergodicity. So again we measure the time average but compute the ensemble average, just like temperature.