

Chapter 2

Basics of Statistical Mechanics

The laws of thermodynamics are a powerful tool that can be used to work out whether a process is permitted, or in calculating an equilibrium state of a system. They are based on experiment, and provide the central concept of maximising entropy which is key in determining the equilibrium states. However, it has little physical interpretation. Statistical Physics provides a microscopic view of physical systems, that provides a firm conceptual interpretation of entropy.

2.1 Counting States

In statistical physics it will be important to count the number of ways of arranging a system with a particular fixed constraint. To help with this we will briefly recap some essential tools to do this. We will often use combinatorics to count the number of ways of arranging N objects made up of n_1 of type 1, n_2 of type 2, etc. This is given by

$$\frac{N!}{n_1!n_2!\dots n_m!}. \quad (2.1)$$

For example, suppose we have 200 pieces of cutlery; 100 knives, 80 forks and 20 spoons, then there are

$$\frac{200!}{100!80!20!} \quad (2.2)$$

ways of arranging them in a line. The difficulty with expressions like this is that the numbers involved are very large (and we have yet to put in Avogadro's number of particles $10^{23}!$). Stirling's approximation is extremely useful in calculating large factorials such as these.

State counting and Stirling's approximation

Stirling's approximation gives an estimate for the factorial of a large number N

$$\boxed{\ln N! \approx N \ln N - N.} \quad (2.3)$$

There are many ways to derive this, and indeed to put in higher order corrections to the series. It can be justified as follows. Let us first rewrite the logarithm as a sum, then approximate it as an integral

$$\ln(N!) = \sum_{n=1}^N \ln n \quad (2.4)$$

$$\approx \int_{\frac{1}{2}}^{N+\frac{1}{2}} \ln n dn \quad (2.5)$$

$$= [n \ln n - n]_{\frac{1}{2}}^{N+\frac{1}{2}} \quad (2.6)$$

$$\approx N \ln N - N \quad (2.7)$$

This series can be calculated systematically to higher order in N , if a more accurate approximation is required.

2.2 Most likely macroscopic state

The key idea of statistical physics can be illustrated using the tools we described above with a simple example. The resemblance of this example to some of the calculations of equilibrium of closed systems is no coincidence.

Consider two chambers sharing N particles, with N_1 in the first and N_2 in the second such that $N = N_1 + N_2$ (i.e. we conserve the number of particles). The number of ways of sharing the particles between the chambers is

$$W = \frac{N!}{N_1!(N - N_1)!}. \quad (2.8)$$

If we apply Stirling's approximation to this formula, we obtain the following

$$\ln W \approx N \ln N - N_1 \ln N_1 - (N - N_1) \ln(N - N_1) \quad (2.9)$$

What is the likelihood that all the particles are in one side of the container? Intuitively we know that it is extremely unlikely that all the air molecules in the room are in one particular corner. It would be instructive to see this result from our expression for W . To evaluate this, we make a Taylor expansion of the expression on the right hand side about the point $N/2$ (the expected result of equal sharing), i.e. $N_1 = N/2 + n$ and $N_2 = N/2 - n$, where $2n$ is the difference between the number of particles in the two containers. We find the following

$$\ln W \approx N \ln 2 - \frac{2n^2}{N}. \quad (2.10)$$

We can rewrite this as a unnormalised Normal distribution

$$W \approx 2^N e^{-\frac{2n^2}{N}} \quad (2.11)$$

This indicates that the particles are Gaussian distributed with mean $n = 0$ and standard deviation $\frac{\sqrt{N}}{2}$. For large $N \sim 10^{23}$ we find that there are $N/2$ particles in the two sides with error $\frac{1}{\sqrt{N}} \sim 10^{-12}$ i.e. we find a very sharply peaked distribution about the point $n = 0$! The distribution is extremely narrow with mean at the expected equal sharing value from thermodynamics. This example illustrates that fluctuations in the thermodynamic quantities is proportional to $\frac{1}{\sqrt{N}}$. This is a general feature that we will come back to later. The normal distribution arose in this problem because we used Stirling's approximation in the Binomial distribution. However, the Normal distribution occurs more

widely as a result of the *Central limit theorem*. Essentially this theorem shows that if we add together many random events (random variables) then, subject to some pretty weak assumptions, their sum is Normally distributed. We will use the Normal (also called Gaussian) distribution frequently in this course.

2.3 Boltzmann Entropy

Statistical physics takes a different view of macroscopic systems than thermodynamics. The microscopic details of the system are considered, then averaged over. Typically a system consists of 10^{23} particles that can occupy many different states, and is constantly making transitions between these states. A macroscopic measurement corresponds to an average over the states occupied by the particles.

A microstate of the system corresponds to a particular realisation of the (quantum) state of the system. For example specifying all the velocities and positions of the particles making up a monatomic gas. Boltzmann's great insight was to link the principle of maximum entropy with the idea that the state of thermal equilibrium must also be a state of maximum probability, i.e. maximum likelihood. He postulated the following (unique) expression for the entropy

$$S = k_B \ln W \quad (2.12)$$

where $k_B = 1.3807 \times 10^{-23} \text{ J K}^{-1}$ is Boltzmann's constant and W is the number of microstates consistent with the macroscopic constraints. The entropy is extensive, i.e. the entropy of two systems should combine additively $S_T = S_1 + S_2$. However, the number of microstates is multiplicative, i.e. $W_T = W_1 W_2$. Boltzmann's expression accommodates both of these facts through the properties of logarithms i.e. if we add the entropy for two systems $S_1 + S_2 = k_B \ln(W_1 W_2)$.

Equal *a priori* probability

Suppose we start counting the states in a system with the same internal energy $W(U)$. The principle of equal a priori probability claims that it is equally likely to find the system in any one of these $W(U)$ states.

There are some philosophical questions associated with this principle as in practice there can be no degenerate levels (levels with the same energy) as some external perturbation will lift the degeneracy. However, we shall instead take some tiny energy band (i.e. states with energy between U and $U + dU$) over which the states can be distributed.

Ensembles

The results of calculations in statistical physics are typically averages of quantities, or probabilities that the system will be in a particular state. We will calculate the possible microstates of the system, and use the principle of equal a priori probabilities to calculate a probability of the state of the system. This probability is most easily interpreted in terms of an *ensemble*. We will imagine that we have a very large number of identical systems so that the fraction of systems that are in a particular (non-degenerate) microstate is $p_i = \frac{1}{W}$, where W is the total number of microstates. The alternative (which Boltzmann assumed gave the same result) is a time ensemble where the system is sampled at discrete times, and the probability is interpreted by how often each microstate is found.

2.4 Microcanonical Formalism

In the microcanonical ensemble our system has a fixed amount of energy. We count states that have the same energy (degenerate), or in practice they are very close in energy.

Examples

1. Consider a set of n simple harmonic oscillators with m quanta of energy to divide between them. We can count the (large!) number of microstates on how this can be done as follows. We will represent each quantum as an \times , and the boundary between each oscillator as an $|$. A particular microstate of the system might then be written as

$$\times \times \times || \times || \times | \times$$

This microstate corresponds to a system of 6 oscillators, the first of which has 3 quanta, the next is empty, the next has 1 quantum, the next empty and the last two have 1 quantum each. The number of microstates in the system can be evaluated by counting the number of ways we can arrange the $n - 1$ “|” and the m “ \times ” symbols, i.e.

$$W = \frac{(n - 1 + m)!}{(n - 1)!m!} \quad (2.13)$$

If we apply Stirling’s approximation to this we can rewrite the expression for natural logarithm of W , $\ln W$, in our system of oscillators as

$$\ln W \approx (n - 1 + m) \ln(n - 1 + m) - m \ln(m) - (n - 1) \ln(n - 1) \quad (2.14)$$

2. We will now consider Schottky defects in a crystal. Here atoms in a crystal lattice are displaced from their sites in the crystal and migrate to the surface. Let ϵ be the energy to form one of these defects. Suppose we have N atoms in the crystal and n Schottky defects ($N \gg n$ for the crystal to have any meaning). The $N + n$ lattice sites are divided into two categories N that are occupied by atoms, and n that are occupied by defects. Hence the number of microstates is

$$W = \frac{(N + n)!}{n!N!} \quad (2.15)$$

If we apply Stirling’s approximation we find the entropy is approximately

$$S \approx k_B [(N + n) \ln(N + n) - n \ln n - (N) \ln(N)]. \quad (2.16)$$

We saw that thermodynamically the temperature can be defined as

$$\frac{1}{T} = \frac{\partial S}{\partial U}. \quad (2.17)$$

In this system $U = n\epsilon$ is the energy of the Schottky defects. If we differentiate S with respect to n we find

$$\frac{dS}{dn} = k_B (-\ln n + \ln(N + n)), \quad (2.18)$$

hence the temperature must be

$$\frac{1}{T} = \frac{\partial S}{\partial U} = \frac{dS}{dn} \frac{dn}{dU} = \frac{1}{\epsilon} k_B (-\ln n + \ln(N + n)) \quad (2.19)$$

$$\Rightarrow n = \frac{N}{e^{\epsilon/(k_B T)} - 1} \quad (2.20)$$

This result is only valid for low temperatures $\epsilon \gg k_B T$ where the crystal lattice has very few defects. Hence we can neglect the 1 in the denominator and find

$$n = N e^{-\epsilon/(k_B T)}. \quad (2.21)$$

This result has the usual form of results of a Boltzmann factor that commonly in Statistical Physics.

2.5 Thermal Equilibrium

We now return to thermal equilibrium but from a statistical physics point of view. Suppose we have two systems with energy U_1 and U_2 respectively, and hence number of microstates $W_1(U_1)$ and $W_2(U_2)$. If these two systems are put into contact with each other they can share energy, and will have $U = U_1 + U_2$, and

$$W_T = W_1(U_1)W_2(U - U_1) \quad (2.22)$$

microstates. To calculate the most likely sharing of the energy between the two systems we find the most likely state i.e.

$$\frac{d}{dU_1} \ln W_T = \frac{d}{dU_1} \ln W_1(U_1) + \frac{d}{dU_1} \ln W_2(U - U_2) = 0 \quad (2.23)$$

$$\Rightarrow \frac{d}{dU_1} \ln W_1(U_1) = - \frac{d}{dU_2} \ln W_2(U_2) \quad (2.24)$$

This condition derived from the maximum in the entropy is enables use to define a temperature for Statistical Physics

$$\beta = \frac{d}{dU} \ln W(U) \quad (2.25)$$

If two systems are in thermal equilibrium then they have the same value of β .

We can recover an expression for β by recalling that

$$W = e^{S/k_B}. \quad (2.26)$$

Hence we have

$$\beta = \frac{d}{dU} \frac{S}{k_B} = \frac{1}{k_B T}. \quad (2.27)$$

It is useful to retain β as short hand for this expression.

2.6 Microscopic definitions of entropy

We have so far met Boltzmann's entropy for a closed system $S = k_B \ln W$. Applying this to a system that has fixed energy, so only levels with the same energy are accessible, has some conceptual difficulties. In practice the energy levels are always non-degenerate, as a weak external effect will split the degeneracy. Consequently we must think about states within a small band of energy dU . Another problem relates to the kinetics of the system. Since W is an extremely large number, then this definition must only apply in the very long time limit, so that all these accessible states can be visited. There are a couple of alternative definitions that do not have such deep conceptual problems.

Boltzmann entropy for a partitioned system

Instead of considering the states accessible by the whole system, we can divide up the system into many smaller parts. The characteristic size of these parts will relate to the measurement we are performing, or the time scale. If we are only measuring for a short time, then size of the subsystem should be chosen so that number of states in each subsystem can be explored in the time available. The entropy of the whole system is thus made up of a sum over all the subsystems

$$S = \sum_{i=1}^N k_B \ln W_i(U_i) \quad (2.28)$$

where $\sum_i U_i = U$ the total energy of the system. Each of these subsystems will have a poorly defined energy, and the spread of energies will depend on the size of the subsystems. These fluctuations represent real effects seen in measurements.

The Gibbs entropy

Gibbs' idea of ensembles leads to another definition of entropy. The probability that the system is in a particular microstate labelled by i is p_i . The entropy of the system is then

$$S = -k_B \sum_i p_i \ln p_i \quad (2.29)$$

Suppose that we have many identical systems in the ensemble, out of a total of M . Thus there are $m_j = Mp_j$ of these systems in the particular microstate labelled j . The total number of realising this particular configuration of systems is

$$W = \frac{M!}{m_1! m_2! \dots m_j! \dots} \quad (2.30)$$

If we then use Boltzmann definition of entropy, and Stirling's approximation we find

$$S_M = k_B \ln W = k_B (M \ln M - \sum_j m_j \ln m_j). \quad (2.31)$$

Since $\sum_j m_j = M$, we can rewrite this as

$$S_M = -k_B M \sum_j p_j \ln p_j \quad (2.32)$$

where $p_j = m_j/M$. This is the entropy for the whole ensemble of M systems, hence the entropy for a single system is

$$S = -k_B \sum_j p_j \ln p_j \quad (2.33)$$

where the sum is over the microstates j , and the probability of finding the system in microstate j is p_j . This expression for the entropy can be applied to very small systems to calculate their disorder. It has found application more widely in information theory to quantify the amount of information contained in a message.

Entropy of Mixing

Entropy can be thought of intuitively as a measure of disorder. One illustration of this is the entropy of mixing. Beating an egg is an irreversible process as it is very difficult to unmix the yolk afterwards. The beaten egg has a higher entropy than the separated one. As an example, consider again a volume separated into two equal parts, each of volume V , with $N/2$ white particles on the left, and $N/2$ black particles on the right. The entropy of the unmixed state is

$$W_{\text{black}} = W_{\text{white}} = \frac{V^{N/2}}{(N/2)!} \quad (2.34)$$

$$S_{\text{unmixed}} = 2k_B \ln \frac{V^{N/2}}{(N/2)!} \quad (2.35)$$

If the partition is removed, and the black and white particles allowed to mix the entropy of the mixed state is now

$$W_{\text{black}} = W_{\text{white}} = \frac{(2V)^{N/2}}{(N/2)!} \quad (2.36)$$

$$S_{\text{mixed}} = 2k_B \ln \frac{(2V)^{N/2}}{(N/2)!} \quad (2.37)$$

Hence the change in entropy is

$$\Delta S = S_{\text{mixed}} - S_{\text{unmixed}} = Nk_B \ln 2 \quad (2.38)$$

so the intuitively more disordered mixed state has a higher entropy.