

Royal Holloway, University of London

# Statistical Mechanics

PH2610

D.D.Kodwani

2013

Statistical Mechanics is a mathematical tool in large areas of science which involve large systems with numerous variables. Saying that statistical mechanics is a theory of gases is like saying calculus is the theory of planetary orbits. In the most general form it is the same as probability theory, there is no general line drawn between when it is used to describe reality (as this is when it is called statistical mechanics).

In these notes, for simplicity we work in units of Boltzmann constant,  $k$ , equal to 1. Therefore we have two simple relations:

$$T = k_B T$$

$$S = \frac{S}{k_B}$$

That can be used to convert from my units to text book units.

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

## Stationary systems

Systems that have no laws of motion, they simply remain stationary.

### Coin flipping

Consider a fair coin (by fair a mean it is completely symmetric, in its mass distribution etc.). We will disregard the effects of air resistance and any external forces (except gravity of course) during our discussions. The probability of getting a Head is the same as getting a Tail, a half, this shows uniform probability.

This brings us on to the concept of an *a priori* probability. This simply means that we use logically arguments to come up with the probabilities<sup>[1]</sup>, for example a system with N possible outcomes which are mutually exclusive (they do not interact with each other, this is the essence of what I mean by stationary systems), the probability of obtaining one of the possible outcomes is  $1/N$  if the probabilities are uniform (notice that this could also mean that we do not know anything about the system, as we shall see in the coming chapters).

### 6 sided die

Consider a fair die with a six sided symmetry with different colours painted on each of the sides.



Figure 1

In the figure there are of course three other colour on the sides that are not seen. Let's say the six colours are: Red(R), Green (G), Blue (B), Orange (O), Yellow(Y), and Purple (P).

As we did with the coin, we neglect external affects and assume the die is symmetric. Once again by the same logic we arrive at the *a priori* probability of  $1/6$  for any of the colours to be chosen. This might seem trivial and you might ask what the point of discussing another case is, however now consider a die that has instead of purple another red side. Therefore there are now only five colours: R, G, Y, O, and B. so one might assume that from the same logic the probability of picking out the any colour at random will be  $1/5$ . However in the case of die we know this nonsense, as the probability would be  $1/3$  to pick out a red side and  $1/5$  to have another other colour. This shows the nature of how we determine the *a priori* probability; we must in the case look at the symmetry of the sides of the die and not the symmetry in the colour. This will become clearer as we discuss further these topics and generalise this topic.

Another way to find the *a priori* probability is to flip the die a very large number of times, this would also give as the probability stated above (of course in practice we would have to do the experiment an infinite number of times to get the exact probability). During this entire discussion the key component missing is change with time. We have assumed time symmetry (basically means the system is not changing), and this brings us on the next section of Dynamical systems.

## Dynamical Systems

A dynamical system changes with time subject to some laws of motion; these could be Newton’s Laws, Kepler’s Laws etc. (however we may not use this to describe quantum phenomena as quantum mechanics follows a different *logic* to classical physics, mainly because it is not deterministic).

### Laws of motion (LOM)

A law of motion can be thought of simply as an updating process. Given a system in an initial configuration the law just takes it into the next configuration in the next time interval.

#### LOM 1

Imagine dividing up time into equal segments and the die, described above (shown in Figure 1), updating itself in each of the given time segments. We can write a LOM for the die such as if the colour on the top of the die is Red, it goes to Purple, for simplicity I could represent this as  $R \rightarrow P$ . Therefore the LOM I have could be something like:

$$R \rightarrow P$$

$$B \rightarrow G$$

$$Y \rightarrow R$$

$$G \rightarrow O$$

$$O \rightarrow Y$$

$$P \rightarrow B$$

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

This is our LOM 1; it can also be represented by a mapping:

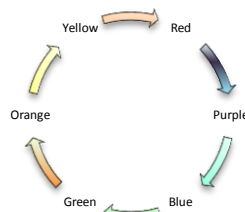


Figure 2

This is a simple die. If we say the die therefore the rate of the cycle above is very fast (the time intervals between adjacent intervals is very short), then the *a priori* probability of picking out any colour of the die is simply 1/6 as we had in the stationary case.

dynamical theory of the changes very rapidly,

#### LOM 2

Now we suppose that our law of motion follows the following rule:

$$R \rightarrow B$$

- $B \rightarrow G$
- $G \rightarrow R$
- $Y \rightarrow P$
- $P \rightarrow O$
- $O \rightarrow Y$

Do you notice anything different about the system? Maybe it will be clearer when we draw the mapping between the configurations:

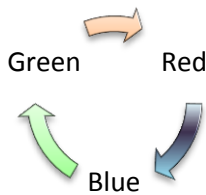


Figure 3

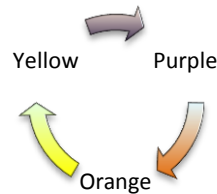


Figure 4

As we can see, the law divides the system into two separate cycles. Therefore we cannot calculate an *a priori* probability for this system if we have no information as to which cycle it is in (this is after we have waited for a sufficiently long time, in this case it has to be longer than the time interval between the changes in the configuration of the system). Suppose the probability of being in either of cycles is a half (they are symmetric), then with the same assumptions as the stationary case we get the probability of any colour to be obtained is again  $1/6$  ( $1/2 * 1/3$ ), however in general this is not the case.

**Conserved Quantities**

This difference can be seen from a different angle as this LOM having a conserved quantity. This is a very important idea in statistical mechanics as all of the probability distributions can only be found with conserved quantities. This might seem an abstract way to go about it, but it will probably make more sense when we talk about some physical examples.

If we define our conserved quantity as  $\mathcal{D}$  and say that it has a value of 1 for cycle 1 (I define this to be the cycle in figure 3) and a value of 0 for cycle 2 (the cycle in figure 4). Mathematically it means:

$$\mathcal{D} = 1, \text{ cycle 1}$$

$$\mathcal{D} = 0, \text{ cycle 2}$$

If I am saying that  $\mathcal{D}$  is a conserved quantity then it will always have the same value, that is to say that if it is 1 then it will remain as that, which clearly means that the configuration of the system will remain in cycle 1, once it has been measured to be in 1 (this is called a closed cycle). This now leads

to an *a priori* probability as the system is now symmetric and finding the system in any of the three configurations has a probability of a 1/3.

In general a conserved quantity means that a system breaks up into different cycles which have different values of the conserved quantity. It is important to remember that when we talk about separate cycles or orbits we are not talking about them in regular coordinate space; it is in phase space (the space of states).

In physics the most important conserved quantity is Energy. Momentum and angular momentum we do not talk about as much in thermo dynamical descriptions, because we are usually talking about gases or liquids contained in containers. Therefore a gas molecule collides with the walls and the momentum of the wall is usually neglected (it's like throwing a tennis ball at a wall and disregarding the recoil of the wall and the earth) therefore conservation of momentum does not hold. Electric charge is another conserved quantity.

I will end this section by giving an example of a conserved quantity applied to a physical system.

### Two Dice

Consider two usual dice with six numbers on them, except this time they interact with each other.

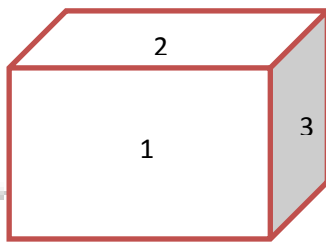


Figure 5

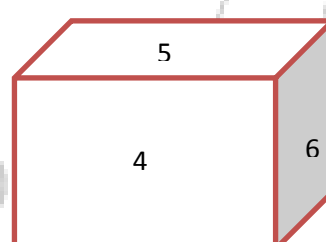


Figure 6

As one flips, so does the other but we constrain the fact that as they flip the sum of the two numbers that appear on top of each die must be equal. That is to say, if 2 and 5 are on top of the die initially then the configuration can be 4&3, 3&4, 6&1 and 1&6. This is a conservation law for the sum of the numbers on top of the dice. We can call this conserved quantity  $\lambda$ , the values that it can have are 2,3,4,5,6,7,8,9,10,11,12. Each of these values has an associated cycle with them with different number of configurations within the cycle. The state(a given value of  $\lambda$ ) which has the most number of possible configurations is the most likely outcome in a random trail. For example 12 has only one possible configuration, therefore it is very constraint and less likely in a random trail (any one has a played monopoly knows this and why it is so hard to get out of jail!).

**Information**

This is probably (pardon the pun) the most important section I will write about and the deepest foundation on SM (it is also the deepest foundation of classical mechanics). We start off by considering a new law of motion:

LOM 3

$$R \rightarrow R$$

$$Y \rightarrow R$$

$$B \rightarrow R$$

$$G \rightarrow R$$

$$O \rightarrow R$$

$$P \rightarrow R$$

This is a valid system under a given LOM, however as we see all the configurations go to one configuration under this LOM. As a map we have:

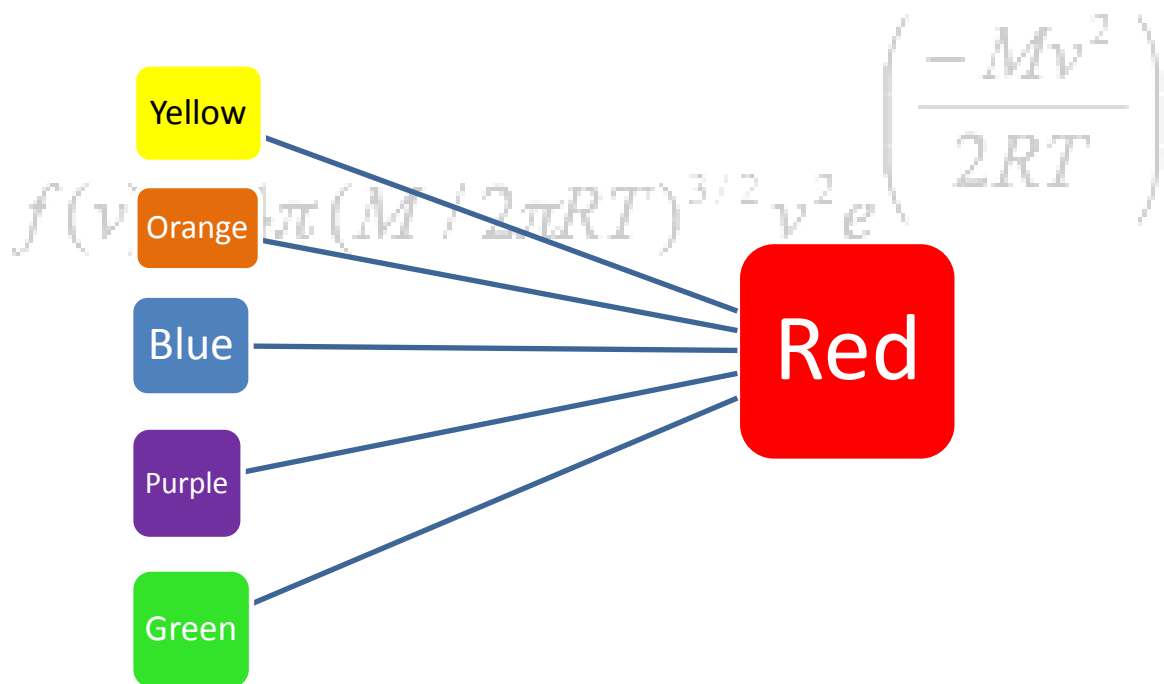


Figure 7

No matter when you sample (as long as it is after the time interval for one process), you will always find R. This system is not representative of any real physical system as it does not conserve information. A real physical system has a property of “conservation of distinction” or “conservation of information” (both mean the same thing). This means that a system must be deterministic, which this system is, and the distinctions don’t merge, which happens in our system.

From figure 7 we can see that every configuration will definitely (with probability 1) lead to another configuration, therefore it is deterministic. However once the configuration has reached red, there is no way to tell what configuration the system was in before. This means some information has been lost, which is not allowed in any physical system, it can be summed up in the statement; "Input gives Output but Output does not give Input". The entire theory of classical physics is based on the assumption of the conservation of information. The theorem that proves this called *Liouville's theorem* (Basically that the mapping between any two configurations is 1 to 1).

A classical system is always deterministic, it only appears statistical because of the large number of variables mean that we simply cannot know enough, but if we could know all the variables we should be able to determine every possible configuration the system will be in at any given point in time.

In essence, the conservation of information is the deepest of the laws of thermodynamics and hence I will call it the -1<sup>st</sup> law of thermodynamics.

#### Example: Single particle

In classical mechanics we usually deal with continuous variables. Therefore we consider a single particle moving through space with a given momentum. The configurations of the particle can be labelled by the position and momentum at any given instance in time. Therefore we can map out its trajectory on a graph, which is known as the phase space of the particle (not be confused with ordinary coordinate space).

The phase space is six dimensional as there are three values to describe its position ( $x$ ,  $y$  and  $z$ ) and three values to describe its momentum ( $p_x$ ,  $p_y$ ,  $p_z$ ).

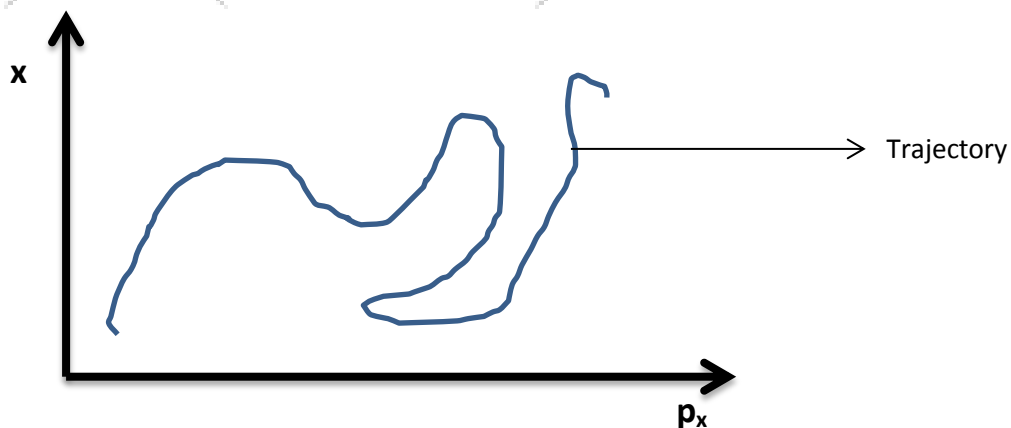


Figure 8

The phase diagram just has two axes that is because I don't know how to draw a 6 dimensional space!

We see that the trajectory can take any path; it can even come back to the same point as it started. What it cannot do is cross over itself at any point, to see why consider the following diagram:



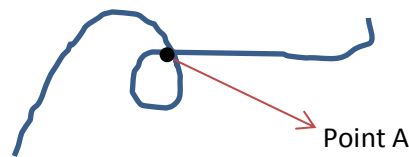


Figure 9

When the particle reaches point A, it has two options it can either go left as we see it or continue to go down and since we have no other information about the particle (other than its trajectory and momentum) it could go either way. This means the trajectory is not deterministic, which is not possible.

What usually happens in a complex system with conserved quantities is that there are several different trajectories that correspond to different values of the conserved quantity.

Example: Harmonic Oscillator

A harmonic oscillator can be drawn in a phase space such as the one in figure 8 and it would look something like this:

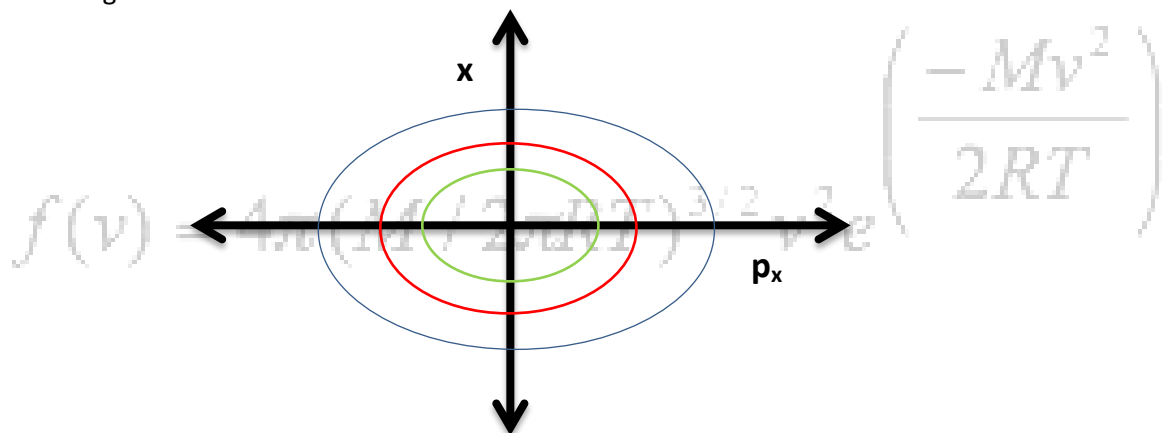


Figure 10

Here the trajectories correspond to different values of the conserved quantity which is energy. It is easy to see that the blue ellipse corresponds to a higher energy value than the other two.

Note: To plot trajectories we need the initial condition of the system as well as the LOM acting on it.

Another trajectory that cannot happen physically is when the two trajectories come infinitely close to together, that is they show asymptotic behaviour. As this would mean that if we wait sufficiently long then they would merge and once again become non deterministic.

I will now briefly explain what Liouville’s theorem is and why it is so important.

## Liouville's Theorem

Suppose we have a phase space with of position and momentum and the space of states for a given system occupy a given volume in this phase space (It is important to note that the volume in this phase space is NOT the volume of coordinate space, infact the volume of a phase space has units of length cubed times momentum cubed which is the unit for action, as supposed to length cubed for the volume of coordinate space):

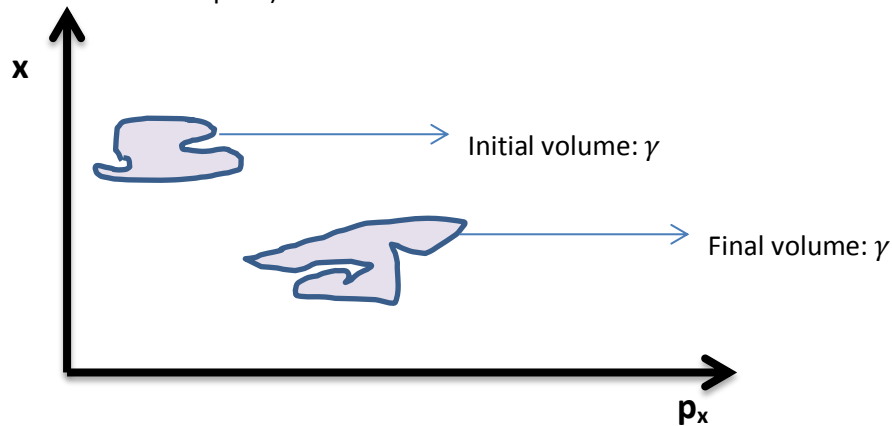


Figure 11

Liouville's theorem states that after any given amount of time the system might have a different configuration of states however the volume occupied by the space of states is always the same, as shown in figure 11. The volume of phase space is conserved, which is the same as stating the  $-1^{\text{st}}$  Law. In reality with complex systems, the space of states usually ends up as some very complicated, fractalated, volume of states such as:

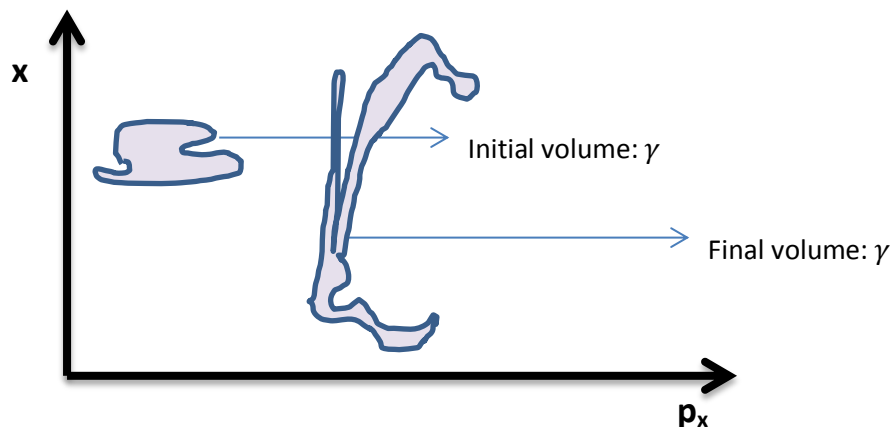


Figure 12

I should stress that even though in the diagram the lines in volume two of figure 11 look like they come asymptotically close, they do not really as that is forbidden, it is just my poor drawing. This is called chaos and is indeed the foundation of chaos theory.

Note: In all of these cases topology is also always conserved.

The axioms for Liouville's theorem are Hamilton and Lagrangian equations of motion and the principle of least action. Subject to a conserved quantity, like energy, the *a priori* probability is uniform in the given phase space. To summarise this theorem; Subject to a conserved quantity, like energy, the *a priori* probability is uniform in the given phase space.

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

## Entropy

We start by discussing probability distributions. They can be either continuous or discrete and are normalised with the following equations:

$$\int P(x)dx = 1, \text{continuous}$$

$$\sum P(i) = 1, \text{discrete}$$

The average value (also known as the expectation value) of a discrete function, say  $F(i)$  is given by:

$$\langle F \rangle = \sum F(i)P(i)$$

Suppose we have a system and all we know about it, is that the configuration of the system is in one of  $m$  states. On a graph it looks like:

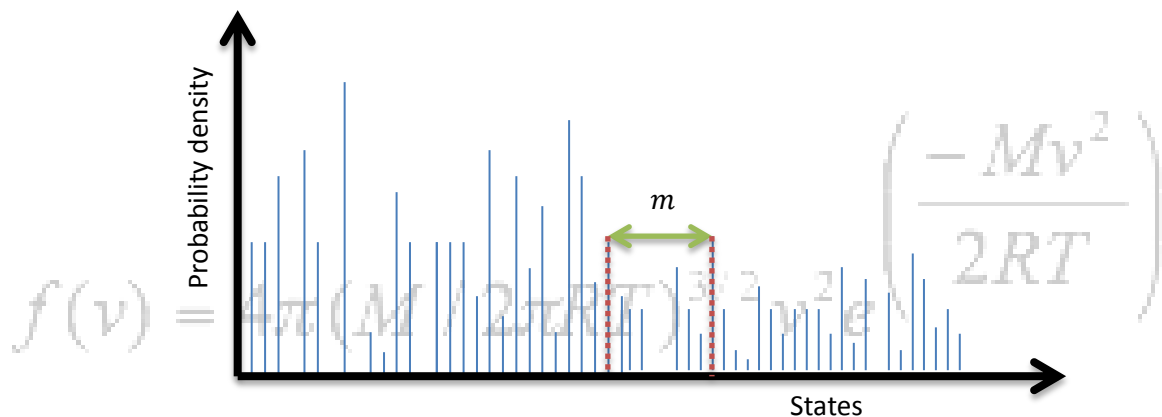


Figure 13

With no other knowledge, all we can say that the probability of finding a given state is  $1/m$  (height of the probability in this region). This can also be seen as the measure of ignorance (how much we don't know about the system), in this case  $m$  is measure of our ignorance. Infact any monotonic function of  $m$  will be a measure of our ignorance. The function we use to define entropy is the logarithm of the  $m$ . Entropy is defined microscopically as:

$$S = \log(m) \text{ or } -\log\left(\frac{1}{m}\right)$$

The base of the logarithm can be chosen arbitrarily depending on what field we are using it in (the change between bases is a trivial one; therefore we can choose which ever one for our convenience). In physics we usually use the base  $e$ , in computer science base of 2 is used as they work with binary systems.

### Example

Suppose we have  $N$  coins, each of which has two possible outcomes heads (H) or tails (T) and suppose we know nothing about the system (complete ignorance).

The numbers of possible configurations are  $2^N$  (as each coin can have two outcomes and there are  $N$  coins). Taking the logarithm we have  $N \log(2)$ , which is the entropy. Now we see that the entropy is proportional to the number of coins itself, therefore we can talk about the entropy that every coin has. In this case the entropy per coin is simply  $\log(2)$ . Entropy is microscopically measured in units of bits.

Another interesting example would be to say; what happens when we know everything about the system?

The answer is of course that there is no ignorance and therefore no entropy. We can also see this from the mathematics, as knowing the space of states simply has one configuration (then one we know it's in) and therefore the entropy becomes proportional to the  $\log(1)$  which is always 0, no matter what the base is.

There is something very odd here, as if you think about it, entropy isn't just something that depends on the system. It is also a measure of how much we know about the system; therefore it seems strange to think of it as a conserved physical quantity (unless we consider ourselves part of the system, like in quantum mechanics a measuring apparatus is considered part of the system it is measuring in an entangled state).

Now we are ready to make a general definition for what entropy is. It depends on the probability distribution we have for a system, which can be affected by what we know about the system as I stated above and the conserved quantities within the system, like in the case of the two dice we saw that each possible value of the conserved variable has its own probability distribution. As we are making a general definition, consider any given probability distribution (discrete or continuous):

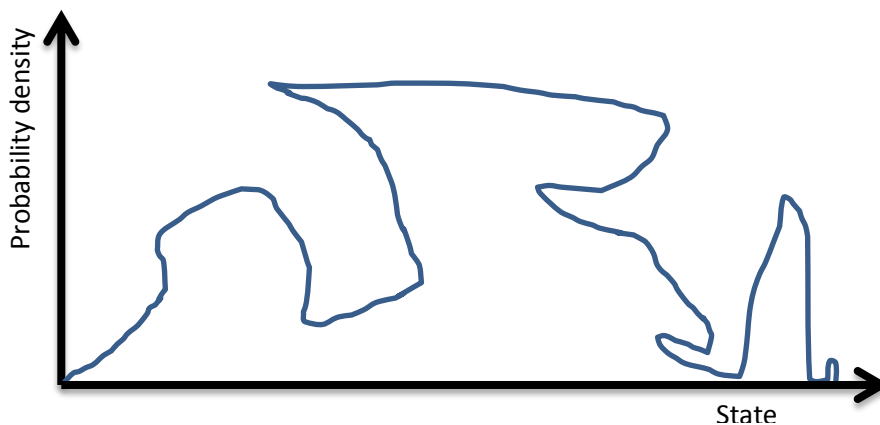


Figure 14

Lets say the function we are interested in is  $\Psi(x)$ . We can calculate the average value of it by:

$$\langle \Psi(x) \rangle = \int \Psi(x)P(x)dx$$

However we know that whatever we know about  $\Psi(x)$  is given by its entropy, therefore we get:

$$S = - \int P(x) \log(P(x)) dx$$

This is the most general form of describing the entropy.

We can apply this to the square distribution described previously. The probability of picking out any given state was  $1/m$ , therefore we obtain the following equation:

$$S = - \int \frac{1}{m} \log\left(\frac{1}{m}\right) dx$$

It is easy to see now that if we integrate over  $m$  (as that is the region in which the probability distribution lies), we get the same relation as we did before for the entropy:

$$S = -\log\left(\frac{1}{m}\right)$$

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left(\frac{-Mv^2}{2RT}\right)}$$

## Temperature

Let's postulate some called thermal equilibrium for a system (of course this is what it means, however we are doing it from first principles) and define it to be:

A property of a small system to contact with a much larger system (such as a heat bath). It cannot be a property of an isolated system.

Consider a heat bath with a small thermometer in it, we label the systems B and A respectively:

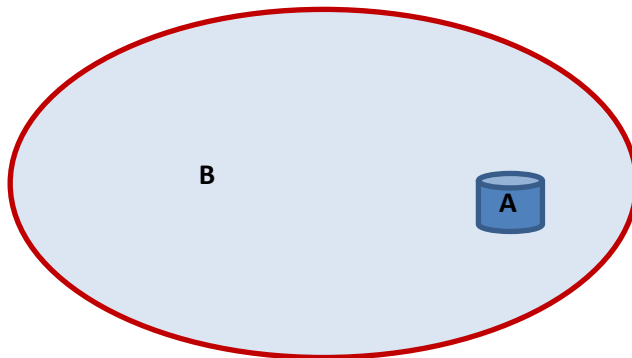


Figure 15

The bigger system B can be thought of as a closed and isolated system, which just means that energy cannot flow in or out of it. The two systems together can be thought of as isolated but not own their own.

The systems A and B can weakly interact with each other (they can exchange energy), therefore if we wait for long enough the systems will each reach the same temperature and this is what we call thermal equilibrium.

If we consider this same situation in a phase space, then the phase space of the combined system has the configuration which has a constant energy, however each of the systems individually can move around with various configurations in there phase space (that could equally mean the different values of energy), for example for system A we have a probability distribution of its states that looks like:



Figure 16

The entropy is given by:

$$S = -\log(P(x))$$

Now we can ask how much energy is needed to change the entropy by 1 bit:

$$\frac{\partial \langle E \rangle}{\partial S}$$

In general this is multiplied by the value of one 1 bit, this quantity given in the equation above is defined as temperature (I must add that all these equations have a factor of  $k_B$  missing for Boltzmann's constant missing, which should be added in front).

This has a striking application when thought about carefully, as it shows' manipulating any bit of information creates a change in temperature. This is a concept widely known in computer science and more specifically information theory. It was first thought about by Rolf Landauer and it holds that "any logically irreversible manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding entropy increase in non-information bearing degrees of freedom of the information processing apparatus or its environment"<sup>[2]</sup>.

Landauer's principle asserts that there is a minimum possible amount of energy required to change one bit of information, known as the *Landauer limit*:

$$T \ln 2$$

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$



## Lagrange Multipliers

In statistical mechanics we are always looking to minimise or maximise (mathematically they are both represented by the same thing) functions of various variables subject to constraints. We are likely to maximise quantities like entropy (in thermal equilibrium entropy is maximised). Constraints could be things like we happen to know the total energy of the system, or the average energy, or the electric charge etc. The method of Lagrange multipliers (also known as method of undetermined coefficients) is a useful way to achieve these points, but it is not the only way.

If we have a function  $F(x, y)$  and we want to find the stationary points of it subject to a given constraint  $G(x, y)$

The method is straightforward:

Multiply the constraint by a constant, usually called  $\lambda$  (the Lagrange multiplier)

$$\lambda G(x, y)$$

Add the constraint to the function

$$F(x, y) + \lambda G(x, y)$$

Differentiate this new function with respect to both  $x$  and  $y$  and set it to zero for the stationary point

$$\frac{\partial F(x, y)}{\partial x} + \lambda \frac{\partial G(x, y)}{\partial x} = 0$$

$$\frac{\partial F(x, y)}{\partial y} + \lambda \frac{\partial G(x, y)}{\partial y} = 0$$

These equations can now be solved in terms of  $\lambda$  and then the value of  $\lambda$  is chosen by fixing the constraint to it.

This might seem an odd way to solve an equation as we just have two equations here and two unknowns, therefore we should be able to solve them by substituting one into the other and it might seem strange that we are introducing yet another variable  $\lambda$ . However in general we could have an equation with  $n$  variables in which case it would be extremely difficult to solve them by substituting them into each other, if we use the method of Lagrange multipliers we simply have to use the same method just with  $\lambda_{n-1}$  multipliers and this is generally easier, however they both should give the same answer.

### Why it works?

In general it can be shown that solving by substitution is essentially the same, consider once again the two functions  $F(x, y)$  and  $G(x, y)$  where  $G(x, y)$  is the constraint:

$$G(x, y) = 0$$

Now imagine solving  $G$  for the variable  $y$ , this will give us  $y$  as a function  $x$ :

$$y = y(x)$$

Now we can substitute this back into the function F:

$$F(x, y) = F(x, y(x))$$

We can differentiate this in the form:

$$\frac{\partial F(x, y(x))}{\partial x} + \frac{\partial F(x, y(x))}{\partial y} \frac{dy}{dx} = 0$$

Now consider the constraint function G(x,y), since it is a constraint, by definition it is constant, therefore:

$$\frac{\partial G(x, y)}{\partial x} dx + \frac{\partial G(x, y)}{\partial y} dy = 0$$

This can be rearranged to give:

$$\frac{dy}{dx} = - \frac{\partial G(x, y)}{\partial x} * \frac{\partial y}{\partial G(x, y)}$$

Now we can substitute this equation into the one above for the function F:

$$\frac{\partial F(x, y)}{\partial x} - \frac{\partial G(x, y)}{\partial x} * \frac{\partial G(x, y)}{\partial y} * \frac{\partial y}{\partial G(x, y)} = 0$$

This can be rearranged as:

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{-\left( \frac{Mv^2}{2RT} \right)}$$

Now if we take the two equations given for the Lagrange multipliers:

$$\frac{\partial F(x, y)}{\partial x} + \lambda \frac{\partial G(x, y)}{\partial x} = 0$$

$$\frac{\partial F(x, y)}{\partial y} + \lambda \frac{\partial G(x, y)}{\partial y} = 0$$

And solve for λ in one of them:

$$\lambda = - \frac{\frac{\partial F(x, y)}{\partial x}}{\frac{\partial G(x, y)}{\partial x}}$$

Substituting this into the other equation we get:

$$\frac{\partial G(x, y)}{\partial y} \frac{\partial F(x, y)}{\partial x} = \frac{\partial F(x, y)}{\partial y} \frac{\partial G(x, y)}{\partial x}$$

Which is the same equation that we obtained from the substitution method

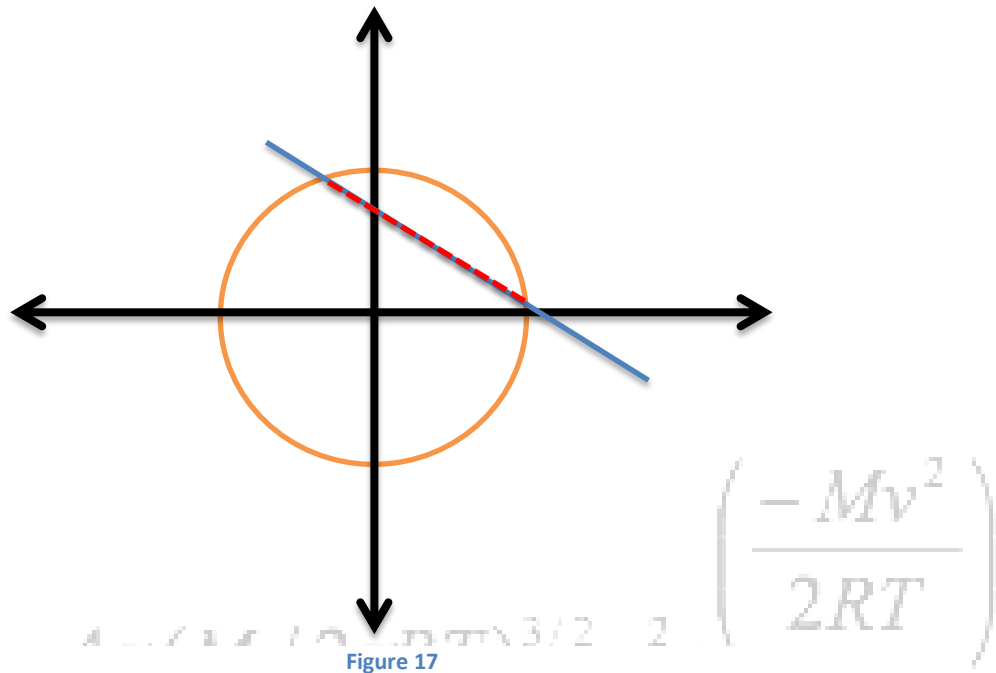
**Example**

Now we can see this in action with an example, consider the functions:

$$F(x, y) = x^2 + y^2$$

$$\text{constraint, } G(x, y): x + 2y = 1$$

I can map this to give some visual representation of the problem we are solving (I have chosen these functions for this reason!)



This map shows the function  $F(x,y)$  as the circle and the constraint represented by the straight line and the solutions we are looking for the points that lie on the red section of the line.

Now we can solve the constraint for  $x$ :

$$x = 1 - 2y$$

This can now be substituted into the equation  $F(x,y)$ :

$$F(x, y) = y^2 + (1 - 2y)^2$$

Differentiating we get:

$$\frac{dF(x, y)}{dy} = 2y - 4(1 - 2y)$$

For stationary points we get this equal to zero (remember  $F$  is the function we are trying to minimise):

$$y = \frac{2}{5}$$

And therefore we get  $x$ :

$$x = \frac{1}{5}$$

Now I will use the Lagrange method, so using the two main equations stated previously we have:

$$\frac{\partial F(x, y)}{\partial x} + \lambda \frac{\partial G(x, y)}{\partial x} = 0$$

$$\frac{\partial F(x, y)}{\partial y} + \lambda \frac{\partial G(x, y)}{\partial y} = 0$$

Substituting for the functions:

$$\frac{\partial F(x, y)}{\partial x} = 2x$$

$$\frac{\partial F(x, y)}{\partial y} = 2y$$

$$\frac{\partial G(x, y)}{\partial x} = 1$$

$$\frac{\partial G(x, y)}{\partial y} = 2$$

Therefore:

$$2x + \lambda = 0$$

$$2y + 2y = 0$$

Solving for  $\lambda$ :

$$x = -\frac{\lambda}{2}$$

$$y = -\lambda$$

Now we substitute  $\lambda$  into the constraint equation:

$$-\frac{\lambda}{2} - 2\lambda - 1 = 0$$

This gives a value of  $\lambda$  as -0.4, which we can substitute back to get the values of x and y:

$$x = \frac{1}{5}$$

$$y = \frac{2}{5}$$

Which we see are the same as we got from the method of substitution.

We can generalise this method for several variables:

If the function we want to minimise is of the form:

$$F(x_1, \dots, x_n)$$

And if we have a set of constraints (they have to be less than the number of variables):

$$G(x_1, \dots, x_{n-1})$$

We can introduce a Lagrange multiplier for each of the variables:

$$\lambda_n$$

Now we form the equations of the form:

$$\frac{\partial F(x_1, \dots, x_n)}{\partial x_1} + \lambda_1 \frac{\partial G(x_1, \dots, x_n)}{\partial x_1} = 0$$

For each of the variables.

Now we follow the same procedure as before and substitute the multipliers to satisfy all of the constraints and use that to work out the values of the variables.

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

### Sterling's approximation

Sterling's approximation is used to give an approximate equation that is generally easier to deal with for the factorial function. The factorial function is defined as:

$$N! = 1 * 2 * 3 * 4 \dots N$$

Now we take the logarithm of both sides:

$$\ln N! = \ln 1 + \ln 2 + \ln 3 + \ln 4 + \dots \dots \ln N$$

It can be seen why this is useful, as the equation shows the right hand side is additive as supposed to multiplicative. A graph can be drawn to give a better representation of the process:

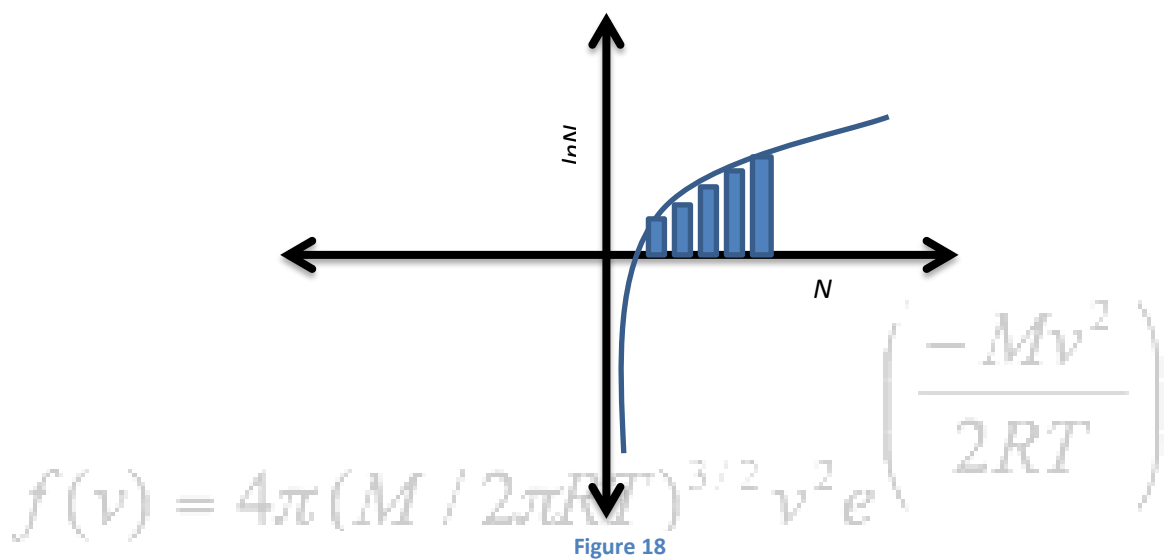


Figure 18

The figure shows that in this method we are essentially integrating under the curve as an approximation for the factorial function:

$$\ln N! = \sum_{x=1}^N \ln x$$

Taking the limit of infinite number of terms in the sum we get:

$$\sum_{x=1}^N \ln x \rightarrow \int_1^N \ln x \, dx$$

Integrating the right hand side we get:

$$\int_1^N \ln x \, dx = N \ln N - N$$

This gives Sterling's approximation:

$$N \ln N - N = N!$$

This can also be rewritten by taking exponentials as:

$$N^N e^{-N} = N!$$

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

### Number of ways to get a particular state

Consider a number of boxes:



Figure 19

Each box can be in a given state (let's not worry about what the state is or represents, in general it could be energy or momentum etc.)

If we have a set of discrete states, say N of them, and we call each individual state  $n_1, n_2, n_3, \dots$  etc. then we can quantify the number of boxes in any given state.

$$\sum_{i=1}^N n_i = N$$

We can get the number of ways there are of getting a given set of boxes in a given number of states very simply by the following equation:

$$\# \text{ of possible configurations} = \frac{N!}{n_1! * n_2! * n_3! * \dots}$$

It can also be written as:

$$\# = \frac{N!}{\prod_i n_i!} \left( \frac{-Mv^2}{2RT} \right)$$

Where # is the number of possible configurations, I will just use the symbol to represent it from now on.

Think of # as a function of  $n_i$ , as  $n_i$  becomes very large we get something like:

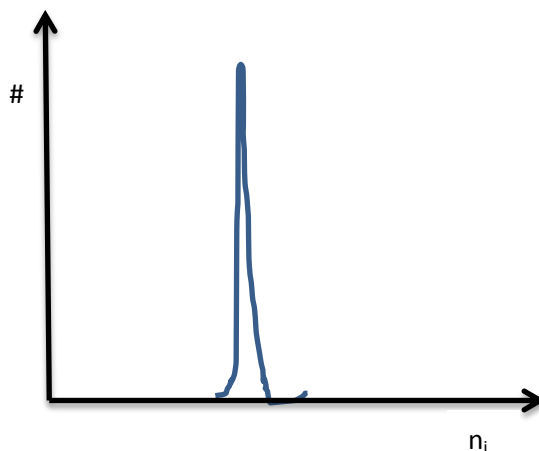


Figure 20



We see the peak is very narrow which means the uncertainty in  $n_i$  is small and we can be sure about the value of  $n_i$  being what it is. For large  $N$  these  $n_i$ 's can become very definitive and we can define the probability as being:

$$p(i) = \frac{n_i}{N}$$

This makes sense as the value of  $p(i)$  is dimensionless and represents the ratio of a value with the total value.

Now we can write down two constraints that follow from the definition of a probability:

$$\sum_i \frac{n_i}{N} = 1$$

This simply follows from the fact that the probabilities always add up to one. The second constraint is a little harder to see. We know that the total energy is:

$$E_T = \sum_i n_i E_i$$

Suppose we also know the average energy, then we can also express the total energy as:

$$E_T = N \langle E \rangle$$

Equating these two we have:

$$f(v) = 4\pi \left( \frac{M}{2RT} \right)^{3/2} v^2 e^{-\frac{Mv^2}{2RT}}$$

If we now divide both sides by the total number of particles we have:

$$\sum_i \frac{n_i E_i}{N} = \frac{N}{N} \langle E \rangle$$

However we know that the probability is simply  $n_i/N$ , therefore:

$$\sum_i p(i) E_i = \langle E \rangle$$

This is our second constraint and is also intuitively pleasing as it is what we have come to expect for any quantity. It might be easier to think of it as the representation of the number of boxes in a given state. As more and more boxes end up being in one state, it is more likely to find a box in that state.

Now we can find the maximum  $\#$  as a function of the  $n_i$ 's, subject to the two new constraints, this is the mathematical problem we have to solve.

When  $\#$  is maximum the log of  $\#$  will also be maximum, it turns out that it is easier to find the log. From the definition of  $\#$  we can take the log of both sides of the equation to get:

$$\log \# = \log N! - \sum \log n_i!$$

This can be rearranged to get it into a more convenient form. First we use sterling's approximation:

$$\log N! = N \ln N - N$$

Therefore:

$$\log \# = N \log N - N - \left( \sum n_i \log n_i + n_i \right)$$

Substitute  $n_i = NP(i)$ :

$$\log \# = N \log N - N - \sum NP(i) \log(P(i)N) + N$$

This can be rearranged as:

$$\log \# = -N \sum P(i) \log P(i)$$

Notice that:

$$\sum P(i) \log P(i) = \text{entropy}, S$$

Therefore we obtain another important and intuitively pleasing result. The entropy we know is related to the probability of states and therefore also gives a measure of the different possible configurations which is kind off observable from the result we obtain.

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{-\frac{Mv^2}{2RT}}$$

## Partition function and its relations

Coming back to the problem, we want to maximise  $\log \#$ , subject to the constraints. We do this using the method of Lagrange multipliers (LM) that we used before.

We add a LM for each of the constraints and then once we have a solution in terms of them we can substitute them back into the constraints to find the value of the multipliers.

$\log \#$  is a function of  $P(i)$  and  $N$  and we want to maximise it w.r.t to the probability (find the state with the maximum probability). So we form the equation of LM:

$$\log \# = -\alpha \sum_i P(i) - \beta \sum_i E_i P(i)$$

$\alpha$  and  $\beta$  are the LM. Now we substitute for  $\log \#$ :

$$-\sum_i P(i) \log P(i) - \alpha \sum_i P(i) - \beta \sum_i E_i P(i) = 0$$

We are trying to maximise therefore  $N$  can be ignored for  $\log \#$ .

Differentiate w.r.t  $P(j)$ , where  $j$  is some specific number:

$$\log P(j) + 1 + \alpha + \beta E_j = 0$$

$$P(i) = e^{-(1+\alpha)} e^{-\beta E_i}$$

We see here that the first term is just constant; the second term is the important term. For now we will define the LM  $\beta$  to be:

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

This will be proven later.

Now we define:

$$e^{-(1+\alpha)} = \frac{1}{Z}$$

Now we can apply the constraints to find the LM:

$$\sum_i \frac{1}{Z} e^{-\beta E_i} = 1$$

Which gives us the relation:

$$Z = \sum_i e^{-\beta E_i}$$

We call this the partition function, applying the second constraint:

$$\sum_i E_i P(i) = \langle E \rangle$$

We can now substitute for the probability to see what we can do:

$$\frac{1}{Z} \sum_i E_i e^{-\beta E_i} = \langle E \rangle$$

After some thought we can see that the equation above can be rewritten as:

$$\frac{1}{Z} \sum_i \frac{d}{d\beta} e^{-\beta E_i} = \langle E \rangle$$

We can take the differential outside the sum and we see that we obtain a very useful relation:

$$-\frac{1}{Z} \frac{d}{d\beta} Z = \langle E \rangle$$

Which we can write more concisely as:

$$-\frac{d \log Z}{d\beta} = \langle E \rangle$$

We can now use the fact that  $\beta$  is  $1/T$  to get:

$$-T \log Z = A$$

This function A must have units of energy (of some sort), we will see that it is in fact the Helmholtz free energy.

Now we have three functions which will turn out to be very useful, Entropy, Energy and Helmholtz function. Consider, first, the entropy:

$$S = - \sum_i P(i) \log P(i)$$

Substituting for P(i):

$$S = - \left( \sum_i \frac{1}{Z} e^{-\beta E_i} \right) (-\beta E_i - \log Z)$$

We can now substitute for the Energy and the Helmholtz function to obtain:

$$S = \beta(E - A)$$

Or in terms of T we have:

$$TS = E - A$$

Or in terms of the partition function:

$$S = \log Z - \beta \frac{d \log Z}{d\beta}$$

### Statistical Fluctuations

Consider a quantity  $X$  which is a function  $i$ , that is for each state in a system  $X$  has a unique value,  $X(i)$ . We want to see how  $X$  fluctuates (varies), this can be defined as:

$$\langle (x - \langle x \rangle)^2 \rangle = \text{variance}, \sigma$$

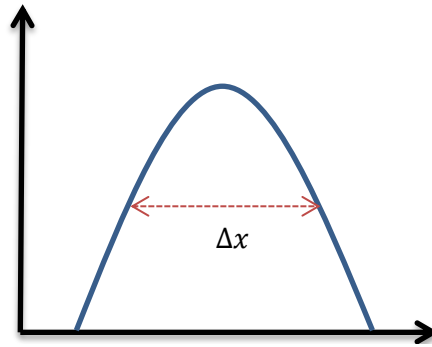


Figure 21

This graph shows the mathematical representation of the variation of our designated variable  $X(i)$ .

The equation of variance can be expanded to give the more familiar form:

$$\sigma = \langle x^2 \rangle - \langle x \rangle^2$$

Variance is defined as the square in the uncertainty:

$$\sigma = \Delta x^2$$

Variance in energy =  $4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$

We can now attempt to find the uncertainty in the energy of a given system. We know that:

$$\langle E \rangle = - \frac{d \log Z}{d\beta}$$

Therefore we can simply square this to get the second term of the variance, the mean value squared. However now we need to calculate the mean of the square for the first term and this is done by:

$$\langle E^2 \rangle = \sum P(i) E_i^2$$

Substituting for  $P(i)$ :

$$\langle E^2 \rangle = \frac{1}{Z} \sum e^{-\beta E_i} E_i^2$$

After some thought we can see that this equation can also be written as:

$$\langle E^2 \rangle = \frac{1}{Z} \frac{d^2 Z}{d\beta^2}$$

Substituting these ingredients into the equation of the variance we obtain:

$$\sigma = \frac{1}{Z} \frac{d^2 Z}{d\beta^2} - \left( \frac{1}{Z} \frac{\partial Z}{\partial \beta} \right)^2$$

Which can be written in a concise form as:

$$\sigma = \frac{\partial^2 \log Z}{\partial \beta^2}$$

We see that this is simply the differential of E energy w.r.t  $\beta$ :

$$\sigma = - \frac{\partial E}{\partial \beta}$$

This can be rewritten in terms of temperature by relating  $\beta$  with temperature:

$$\sigma = \frac{dE}{dT} T^2$$

Notice the change from partial differentials to ordinary differentials; this is because we happen to know that energy is only a function of temperature. We can see that the differential term is simply the heat capacity, hence we are left with a simple relation:

$$\sigma = CT^2$$

I will now remind you that there is a factor of Boltzmann constant missing, as we have been working in units where it is equal to one. With the Boltzmann constant we are left with the equation:

$$\sigma = C k_B T^2$$

Conclusions can be drawn from these results now. We know that C is proportional to the size of the system, in general the size could mean the mass or dimensions but in this case we will simply mean the number of particles in the system N (indeed when you think about it, that's all the size actually means, higher mass comes from more particles as long as the substance has the same particles and larger volume also results in more number of particles as long as the particles are identical).

$$C \propto N$$

The energy itself is also proportional to the number of particles in the system (more degrees of freedom).

$$E \propto N$$

If we rearrange the variance for the uncertainty in energy:

$$\sigma = \Delta E^2$$

Therefore:

$$\Delta E = \sqrt{C} T$$

And therefore:

$$\Delta E \propto \sqrt{N}$$

Therefore we arrive at a very important result; we see that since the change in energy is only proportional to the square root of the number of particles in the system the value of the uncertainty compared to the value of the energy will be small in a system containing a large number of molecules.

$$\frac{\Delta E}{E} \ll 1$$

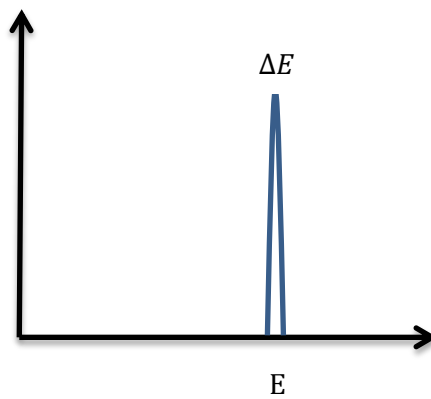


Figure 22

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

## Pressure

We start by discussing the concept of pressure by considering a piston in a container.

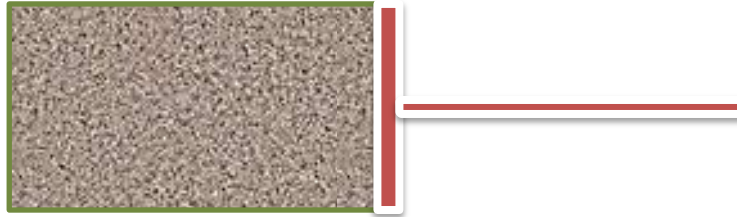


Figure 23

The small dots in the container are the gas molecules, the red block at the end represents the piston.

The pressure on the piston is given by the usual formula of pressure:

$$Pressure = \frac{Force}{Area}$$

Firstly we need to calculate the force on the piston wall. In this discussion we assume the force is only caused by the molecules of gas, however this is not the most general case, as the force can be caused by long range potentials such as gravity or electromagnetism (here we assume there are not external forces). Force is defined in differential form as:

$$F = -\frac{\partial U}{\partial x}$$

Where U is the potential energy.

We can use this for the force on the piston, by taking the energy of the system being the work done by the system on the gas, which can be seen equally as the work done by the gas on the piston. It is important that the motion of the piston is slow, to see why this has to be the condition, we can consider the piston moving away in the container at a very high speed. There will be a certain point when the velocity of the piston is as fast as the velocity of the molecules and in this case there will just be no collisions between the particles at the piston and hence no pressure!

Admitting this is an extreme example, however this shows the limitations of the process. This condition is called the adiabatic condition. By substituting force into the equation for pressure we get the following equation:

$$-\frac{\partial U}{\partial V} = P$$

The negative energy comes by the increasing in energy as the piston is used to compress the gas in the container. Another way to think about why the adiabatic condition is needed is by remembering that the whole concept of pressure is an average concept in the case of a gas. The collisions and force are averaged over time and the number of particles, therefore we need time and hence the slower the piston the moving, the more time we have and hence the more accurate the result is.

### Adiabatic Process



It is a process done slowly and has constant entropy (it comes from quantum mechanics). It can be expressed by considering a system with the discrete energy levels:

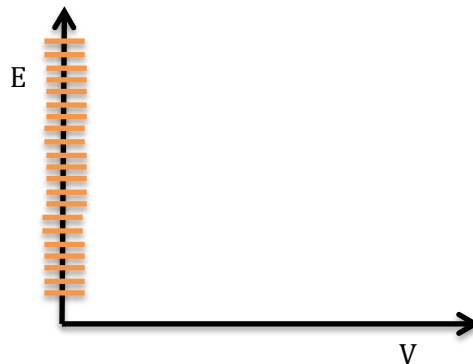


Figure 24

The energy values of the states change as the volume is changed, the energy levels will be closer together if the volume of the container is small and decrease with increasing volume.

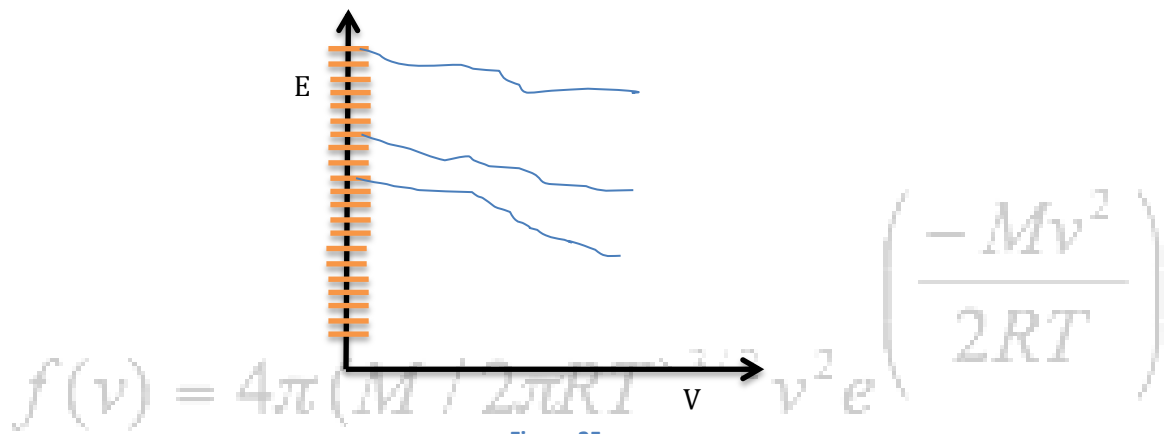


Figure 25

The lines represent how the energy level values change with volume, however they do not appear or disappear or cross each other and hence have the same order as well. Suppose we start of the system in a given state as we vary the control variable (in this case volume, in general it is a variable that we can control, the system remains in the same state even though the state itself varies in the value it has).

Now we suppose that the system starts off in one of three states with a probability of 0.5, 0.25, and 0.25. After a given amount of time the states will change in values, however the probability will still be the same (as we have not gained any more information) and hence the entropy, which is given by:

$$S = - \sum P(i) \log(P(i))$$

Is also the same, it is conserved.

This is expected due to the conservation of information (look back at Liouville's theorem discussed previously).

Now we can use the fact that the entropy is conserved to get the pressure in an adiabatic process. By considering the central equation:

$$dE = PdV - TdS$$

At constant entropy we have:

$$P = \left(\frac{\partial E}{\partial V}\right)_S$$

In general we can have the energy in terms of any control parameter, X, for a given variable, Y. So in general we have:

$$Y = \left(\frac{\partial E}{\partial X}\right)_S$$

We have Y = P and X = V, these pairs of variables are called conjugate thermodynamic variables.

In general entropy is an extremely tough parameter to manipulate; therefore we need a different relation.

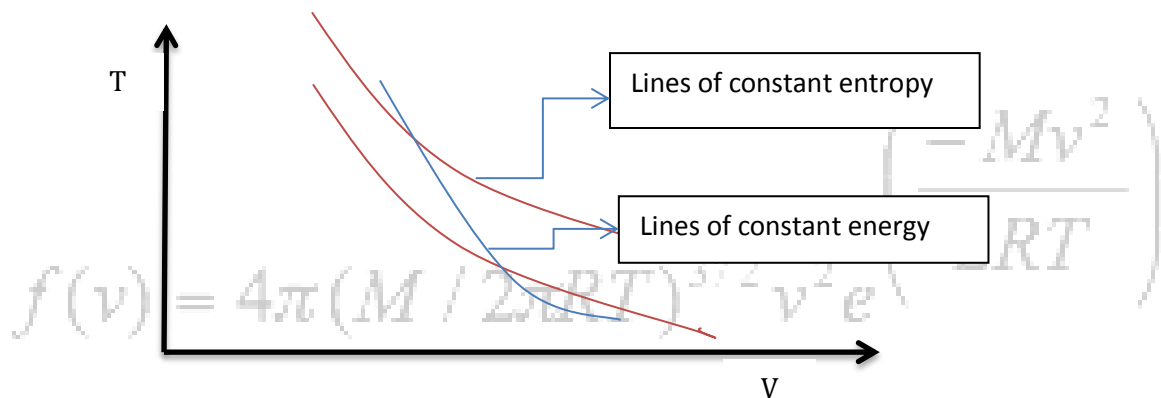


Figure 26

Consider a system with lines given above; if we want to obtain a change in energy we follow the line of constant energy. We split this process up into two stages:

First we go long horizontally at constant T

Then vertically at fixed V (entropy has changed)

This can be mathematically described as:

$$\left(\frac{\partial E}{\partial V}\right)_S = \left(\frac{\partial E}{\partial V}\right)_T + \left(\frac{\partial E}{\partial T}\right)_V$$

Now we note that the second term on the right hand side has with it a corresponding entropy change and by considering the lines of constant entropy (more specifically how to get from one red line to the next) we have:

$$\left(\frac{\partial E}{\partial T}\right)_V = -\left(\frac{\partial E}{\partial S}\right)_V \left(\frac{\partial S}{\partial V}\right)_T$$

Substituting back into the original equation we have:

$$\left(\frac{\partial E}{\partial V}\right)_S = \left(\frac{\partial E}{\partial V}\right)_T - \left(\frac{\partial E}{\partial S}\right)_V \left(\frac{\partial S}{\partial V}\right)_T$$

However we know from the central equation that:

$$\left(\frac{\partial E}{\partial S}\right)_V = -T$$

Hence we arrive at the result:

$$\left(\frac{\partial E}{\partial V}\right)_S = \left(\frac{\partial E}{\partial V}\right)_T - T \left(\frac{\partial S}{\partial V}\right)_T$$

Now, since we are differentiating the left side w.r.t volume at constant temperature we can rewrite the equation above as:

$$\left(\frac{\partial E}{\partial V}\right)_S = \left(\frac{\partial(E - TS)}{\partial V}\right)_T$$

Which simplifies to:

$$\left(\frac{\partial E}{\partial V}\right)_S = \left(\frac{\partial A}{\partial V}\right)_T$$

And using the relation for pressure stated in the beginning we have our final result:

$$p = -\left(\frac{\partial A}{\partial V}\right)_T = -\left(\frac{\partial(\log Z)}{\partial V}\right)_T$$

This is a lot easier to deal with compared to the previous relation we had in terms of entropy.

**Ideal gas**

First of all, we should consider exactly what assumptions we are making in the description of an ideal gas. We assume that all the gas molecules are point like and the volume they occupy is negligible. Another important assumption is that all the particles are identical and they have no interactions between them.

We want to calculate the partition first, once we have that we can calculate all the thermodynamical properties of this system, like the pressure as we can see from the equation we just derived.

The partition function is given by:

$$Z = \sum_i e^{-\beta E_i}$$

This is the sum of the states of the system. If we have N particles, they will show three states of momenta and three states of position each. In general we are dealing with classical systems and therefore the sum turns into an integral as they states become continuous:

$$Z = \int dx^{3N} dp^{3N} e^{-\beta(\sum_i \frac{p_i^2}{2m})}$$

The notation used here is that the superscripts on the differentials show the 3 directions that we have to integrate over and the energy in the exponential term is simple the kinetic energy as we have already stated the system is not under any external forces.

Since the particles are identical if we swap two particles the system is left in an unchanged state overall. So if we have three particles they can be changed around in 6 different ways to leave the overall system unchanged and hence we need to compensate for this factor while finding the partition function. This is done simply by dividing by a factor of N factorial, so our partition function becomes:

$$Z = \frac{1}{N!} \int dx^{3N} dp^{3N} e^{-\beta(\sum_i \frac{p_i^2}{2m})}$$

There is no dependence on position in the integral therefore integrating over dx simply gives the volume element for a particle, therefore we can rewrite our integral as:

$$Z = \frac{V^N}{N!} \int dp^{3N} e^{-\beta(\sum_i \frac{p_i^2}{2m})}$$

The N is the exponent of the volume is there for the number of particles in the system. For one particle we just have:

$$Z_{sp} = \frac{V}{N!} \int dp e^{-\frac{\beta p^2}{2m}} \left( \frac{-MV^3}{2RT} \right)$$

*f(v) = 4π(M/2πRT)<sup>3/2</sup> v<sup>2</sup> e<sup>-Mv<sup>2</sup>/2RT</sup>*

This is a Gaussian integral and has a solution:

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

We want the logarithm, therefore we use sterling's approximation to obtain:

$$\log Z = -\log N! + N \log N - \frac{3N}{2} \log \beta + \frac{3N}{2} \log 2m\pi$$

The energy of the system is defined as:

$$E = -\frac{\partial \log Z}{\partial \beta} = \frac{3}{2} NT$$

Which gives the energy per particle as:

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

Remember that we are working in units in which the Boltzmann constant is set as one, that is why it is missing from the equation above, we can put it back in:

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

Here we see the foundations of the equipartition theorem, as the factor of a half comes from the Gaussian integral and the 3 comes from the 3 direction we account for. So for any given direction of motion the energy is simply ½ kT which is what the equipartition theorem says (note that I say direction however it could be any degree of freedom)

**Pressure for ideal gas**

Now that we have the partition function we can calculate the pressure for an ideal gas by simply differentiating using the result obtained previously:

$$P = - \left( \frac{\partial(\log Z)}{\partial V} \right)_T$$

Substituting the partition function we get:

$$P = \frac{NT}{V}$$

Once again remember there is a Boltzmann constant, and we see that this exactly the same as the ideal gas law that we know.

We can now also calculate the entropy which is defined as:

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

Substituting for E and A:

$$S = -\log Z + \beta \frac{\partial \log Z}{\partial \beta}$$

For an ideal gas this becomes:

$$S = N \log V - \frac{3}{2} N \log \beta - \frac{3}{2} N$$

This is much more complex than the Helmholtz function, therefore we have justified not using the entropy in the very beginning.

An interesting fact can be seen by considering the same container with an ideal gas filled inside it. If we place a light movable piston in the middle of the container (assume there is no friction), then at equilibrium the piston won't move. This means that the force it feels from both sides must be the same and hence so must the pressure.

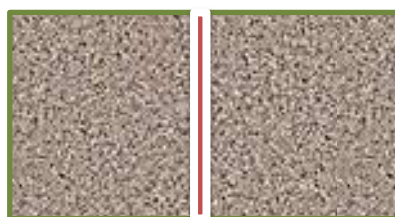


Figure 27

Using our definition of the pressure we have:

$$\left(\frac{\partial A_L}{\partial V_L}\right)_T = -\left(\frac{\partial A_R}{\partial V_R}\right)_T$$

Where  $V_L$  is the volume on the left and  $V_R$  is the volume on the right. For an ideal gas at equilibrium they are both the same, so we can rewrite this equation as:

$$\left(\frac{\partial(A_L + A_R)}{\partial V}\right)_T = 0$$

Hence we see that in equilibrium state, the Helmholtz free energy is minimised, that is at first glance a surprising result as we would (or at least I would) expect the overall energy to be minimised.

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left(\frac{-Mv^2}{2RT}\right)}$$

## Modified Ideal gas

We know take away the assumption that the particles of the gas are point like, instead they have some internal structure and it is this internal structure that shall be the studied in this section.

The effect of this assumption is still maintained, that is we assume that there are no collisions or interactions between the gas molecules (this just means that the density of the gas is very low, which is a fairly good assumption).

Suppose now that we have a molecule made of two atoms and they are connected by a rigid rod (this represents the chemical bond between them):



Figure 28

Now for simplicity we shall consider the molecule as being able to rotate in only one plane (this would have no effect on the physics, as more directions to rotate in can be modelled in the same way and lead to the same effects). The molecule now gas an additional energy of rotation along with its translational kinetic energy, this energy can be excited by collisions and external effects.

To make the calculations easier to follow, we can consider that one atom in the molecule weighs a lot more that the other and hence the heavier one does not rotate much (just like the solar system).

Now if we model the energetics of the system, we have three translational energy components (the motion of the centre of mass of the molecule) of the form:

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{-\frac{Mv^2}{2RT}}$$

Where the M represents the total mass

The rotational energy can be found easily by considering one atom rotating about the other:

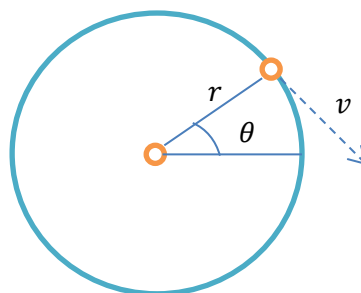


Figure 29

Where r is the radius of rotation and simply the length of the rod we used to represent the bond between the atoms. v is the tangential velocity of the lighter atom moving around the heavier atom.

The rotational energy is then simply:

$$E_R = \frac{1}{2}mv^2$$

However since we are assuming the motion is circular we know that:

$$v = \dot{\theta}r$$

Substituting this into the previous equation we get:

$$E_R = \frac{1}{2}m\dot{\theta}^2r^2$$

From classical mechanics we know:

$$\text{Moment of inertia, } I = mr^2$$

Therefore we arrive at our final result:

$$E_R = \frac{1}{2}I\dot{\theta}^2$$

The angular momentum is defined as:

$$L = I\dot{\theta}$$

Therefore we can rewrite the equation for energy as:

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{-\frac{Mv^2}{2RT}}$$

Now we can write down the total Hamiltonian of the molecule:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{L^2}{2I}$$

Now looking at this from the point of view of statistical mechanics, we see that the overall energy will be a sum of two independent terms, and hence the partition function will be factorable:

$$Z_{sp} = \int e^{-\frac{\beta p^2}{2m}} dx^3 dp^3 \int e^{-\frac{\beta L^2}{2I}} d\theta dL$$

We have done these integrals before and they are the Gaussian integrals, which have the solution of the form:

$$Z_{sp} = V \left( \sqrt{\frac{2m\pi}{\beta}} \right)^3 2\pi \sqrt{\frac{2I\pi}{\beta}}$$

Notice that is no power of three on the last term, this is because we have restricted the rotational motion to a single plane, in general it could be in many planes and as long as they are independent we can add to the power of that term.



This partition can be extended to  $N$  particles simply (as we showed before):

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

However we are generally interested in the log of the partition function:

$$\log Z = -\log N! + N \log V - \frac{3N}{2} \log \beta - \frac{N}{2} \log \beta + \text{constants} \dots$$

Recall the energy as:

$$E = -\frac{\partial \log Z}{\partial \beta}$$

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

This is of course the result of the equipartition function that we studied in the first year.

This is where the problem arises in the classical picture. Consider taking the radius, so the length of the rod connecting the two atoms becoming infinitesimally small, then we should be unable to distinguish between the molecules being point like or not. This would mean that the term adding to the internal energy, which comes from the rotational motion should disappear. However this is not observed (at least at high enough temperatures), and we can extend this argument further if we are making  $r$  very small, we would have the degrees of freedom of the individual electrons of the atoms (which there could be many!) and the degrees of freedom of the nucleus itself which we know is not point like.

So this leads to a major paradox in classical physics as this would mean there are many many degrees of freedom and hence the energy should become vast! (This paradox is closely related to the Ultraviolet catastrophe in black body radiation).

The solution to this problem comes from quantum mechanics. We do not even need to get into the deep theoretical differences to do this; we simply need to assume the fact that the angular momentum is conserved:

$$L = n \hbar$$

If we substitute this relation into the rotational energy:

$$E_R = \frac{n^2 \hbar^2}{2I}$$

The partition function for a single particle now becomes:

$$Z_{sp} = \sum_n e^{-\frac{\beta n^2 \hbar^2}{2I}}$$

This sum is too hard to complete analytically, therefore we have to make some approximations for very high temperatures and very low temperatures.

High Temperatures (low  $\beta$ )

At high temperatures the function will be very changing very slowly and therefore we can approximate the sum to an integral. The function plotted might look something like this:

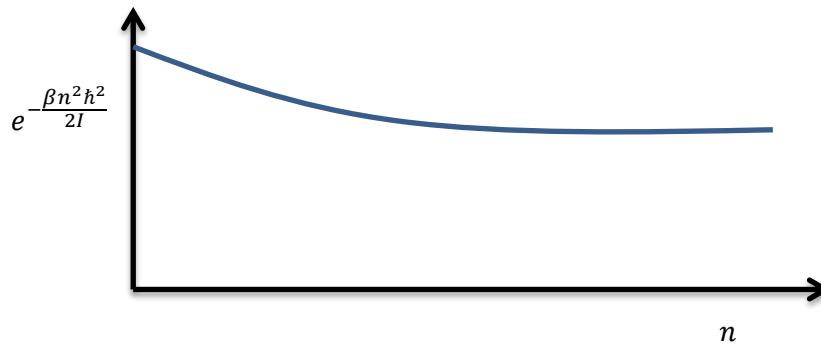


Figure 30

If we replace the sum with an integral we simply get back to the classical regime that we were using earlier and get the exact same integral as we did before, except now we are integrating over  $n$  as supposed to  $L$ , however it makes no difference as they are proportional to each other, and we get the classical answer of  $2NT$  for  $N$  particles.

Low temperature

At low temperatures we have a very rapidly varying function therefore we can't take the integral of the sum as it would not be accurate. However we know that since it is rapidly varying we can expand it and just take the first few terms and they should represent the function to a good approximation.

The curve looks something like this:

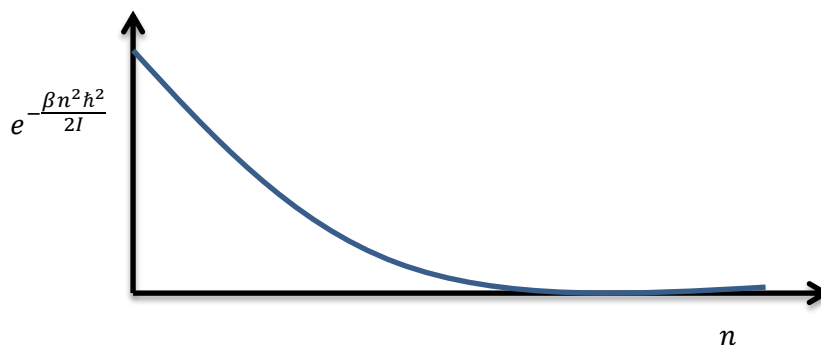


Figure 31

Therefore if we take the first two terms of the sum:

$$Z_{sp} = \sum_n e^{-\frac{\beta n^2 \hbar^2}{2I}}$$

We get:

$$n = 0 \rightarrow Z_{sp} = 1$$

$$n = 1 \rightarrow Z_{sp} = 2e^{-\frac{\beta \hbar^2}{2I}}$$

Therefore we get the approximate single particle partition function as:

$$Z_{sp} = 1 + 2e^{-\frac{\beta \hbar^2}{2I}}$$

This can be extended to N particles as:

$$Z_N = \left(1 + 2e^{-\frac{\beta \hbar^2}{2I}}\right)^N$$

We note that if the temperature is small and  $I$  is small the second term becomes negligibly small. Again we are interested in the log of it:

$$\log Z_N = 2Ne^{-\frac{\beta \hbar^2}{2I}}$$

Therefore the energy is:

$$f(v) = 4\pi \left(\frac{M}{2\pi RT}\right)^{3/2} v^2 e^{-\left(\frac{Mv^2}{2RT}\right)}$$

This does not like the classical formula that we got for the energy; however this is the solution to the molecular structure of the degrees of freedom.

So now the question can be asked about where does the crossover between the classical limit and the quantum behaviour take place. The answer is straight forward, it is when the rotational energy approaches the thermal energy from the equipartition theorem:

$$\frac{1}{2}k_B T = \frac{\hbar^2}{2I}$$

## Black hole thermodynamics

This is a very interesting application I thought of statistical mechanics and thermodynamics and we do not need to know much about the black hole itself to do this.

All we need to know is that it is an object with mass; let's say  $M$ , and a radius, say  $R$ , which represents the Schwarzschild radius (we learnt this in the first year). Consider the black hole having just one control parameter which is energy and that it depends on  $M$  and  $R$ .

We know that  $R$  is:

$$R = \frac{2MG}{c^2}$$

Where  $G$  is Newton's gravitational constant and  $c$  is the speed of light.

Now say we wanted to add 1 bit of entropy into the black hole, to do this we could throw into it a particle and this would also increase the size of the black hole.

We want to know how the radius changes as we add the bits of entropy. Say we throw in photons one by one into a big black hole. We create a photon with precise energy and wavelength to give it one bit of information, however as we threw it into the black hole it will fall at some specific position and this position will be marked by a set of coordinates therefore it would not have just 1 bit of entropy (note I continually change between entropy and information, for this discussion they mean the same thing).

Now we have to assume some facts about the black hole that I will not prove here. Black holes can either absorb or reflect photons (or any other particle) depending on the energy of the photon. The energy of course corresponds to a certain wavelength. It turns out that photons with wavelengths that are shorter than the radius of the black hole are absorbed and the others are reflected. So our photon must have a wavelength less than the radius, however we stated above that this would mean a necessary increase in the entropy. The way to get around it is to simply throw in a photon that has wavelength equal to the radius of the black hole, therefore it has half and half probability of being absorbed or being reflected and therefore the only thing we know is that either its absorbed or reflected and hence corresponding to one bit of entropy.

This can be simply understood by the elementary equations of quantum mechanics, the change in wavelength corresponds to a change in momentum:

$$\lambda = \frac{\hbar}{p}$$

And a change in momentum can be used in the Heisenberg uncertainty relation for position and moment:

$$\Delta x = \frac{\hbar}{\Delta p}$$

The increase in energy is just:

$$E = \frac{hc}{\lambda}$$

The equation for R has a mass in it and we know that mass is equivalent to energy therefore we can use those relations to see that the change in R is:

$$\Delta R = \frac{2\hbar G}{c^3 R}$$

If we move the R from the denominator on to the other side we get a change in area:

$$\Delta A = \frac{2\hbar G}{c^3}$$

This change in area corresponds to change in entropy of one bit. So in general for a change in entropy we have:

$$\Delta A = \frac{2\hbar G}{c^3} S$$

This shows that the entropy will become very large for a black hole and that is intuitively pleasing as the black hole can “hide” a lot of information within it. The quantum mechanical effect on entropy can be seen by the  $\hbar$  in the denominator for entropy and this in fact limits the entropy instead of making it infinite as in classical physics as there is no  $\hbar$ .

Now we can look at the temperature of the black hole. We know that:

$$dE = T dS$$

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

If  $dS$  is just once, corresponding to change in entropy of one bit then the temperature is just the change in the energy. We know the change in energy is simply:

$$dE = \frac{hc}{R} = T$$

We can substitute back for R to obtain:

$$T = \frac{hc^2}{MG}$$

As we stated above M is the same as energy and therefore we have a peculiar result that the temperature is inversely proportional to energy! This is not something we observe very often in nature, except in stars. Decreasing the energy of stars increases their temperature as the gases will start to collapse due to the gravitational pull.

This also means that these objects carry a negative heat capacity! And are very unstable. To see why consider a black hole in thermal equilibrium with its environment. The net transfer of energy is zero however if there is even the slightest of fluctuations in energy, say by taking in a little bit more energy, the temperature of the black hole would then decrease. As this happens the heat from outside will flow in and more energy will lower the temperature even more and this will lead to a run-away

effect until the black hole reaches absolute zero! We know this does not happen, so we have a paradox here that the black hole can never be in thermal equilibrium with its environment.

The temperature that we see here is the temperature observed by someone very far away from the black hole. This is not however the temperature of the black hole itself. That is because the photons have to travel from the black hole (of course they can't come from inside the black hole, strictly speaking they come from the surface of it) and as they do so they lose energy in escaping the gravitational pull of the black hole.

We can use this temperature to calculate the luminosity of the black hole by using the Stefan-Boltzmann law:

$$L \approx AT^4$$

We also know that for a black hole:

$$A \propto R^2$$

$$T \propto \frac{1}{R} \propto \frac{1}{M}$$

Hence the luminosity is:

$$L \propto \frac{1}{M^2}$$

Therefore we see that a big black hole has a small luminosity, and hence radiates slower. This is just the hawking radiation!

$$J(\nu) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} \nu^2 e^{\left( \frac{-M\nu^2}{2RT} \right)}$$

## Second Law of Thermodynamics

In this section I will discuss what the second law is and mainly explain the foundations for where it comes from as it seems to contradict the principles of classical physics at first sight.

In words, the second law of thermodynamics states that any system evolving over time will go from a state of more order to a state of less order. However the fundamentals of classical physics are Newton's laws of motion and according to those the trajectories taken by particles is exact and reversible, but the second law clearly states they are never reversed, so there is a contradiction.

### Box of gas

We start of by considering Newton's laws for a gas in a box and seeing where they lead us. Consider a box in which at time zero we put a bunch of particles in the corner of the box:

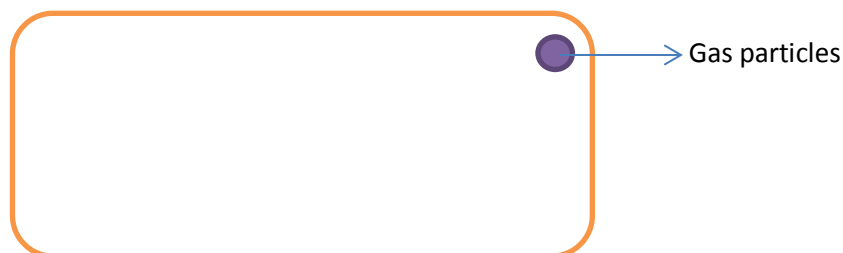


Figure 32

Now the gas particles and the box are taken to be one system. If we let this system evolve over time, the gas will expand and after a given amount of time it will fill the box almost homogeneously:

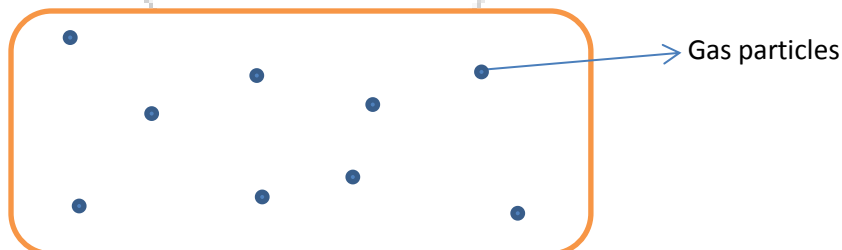


Figure 33

According to Newton's laws of motion the trajectory taken by each of these particles can be known and the gas particles can be taken back, so that the system will be in the initial state again. Hence the box would violate the second law of thermodynamics. Of course this does not happen as the reversing process never happens (in though it is theoretically possible) and the reason for this behaviour lies fundamentally in the idea behind chaos.

To understand this fundamentally we need to understand how systems evolve with time in phase space. As we have already discussed many times in the early chapters, phase space is in general multi-dimensional, we just use two to simplify things.

If we start with a lump in phase space:

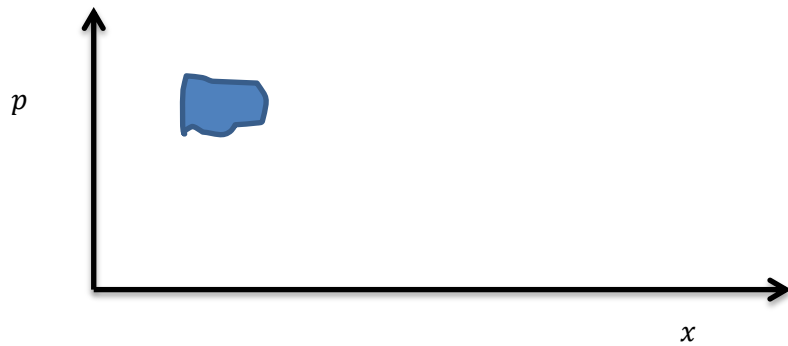


Figure 34

As this lump evolves in time, it forms different shapes and from Liouville's theorem we know that the volume of these lumps always the same, it is worth remembering that the lump we are using know is just a probability distribution.

For simplicity we can say that the probability distribution is 0 outside the boundary of the lump and constant inside it. Now we can apply the equation of entropy to this, as the volume of the lump is the probability distribution we have:

$$S \propto \ln V_s$$

Where  $V_s$  represents the volume of a lump in phase space. However as this is constant the entropy of the system should be constant and again is contradicting the second law of thermodynamics! The reason again is the classical nature of Liouville's theorem.

**Chaos**

Consider an idealised snooker table with no holes:

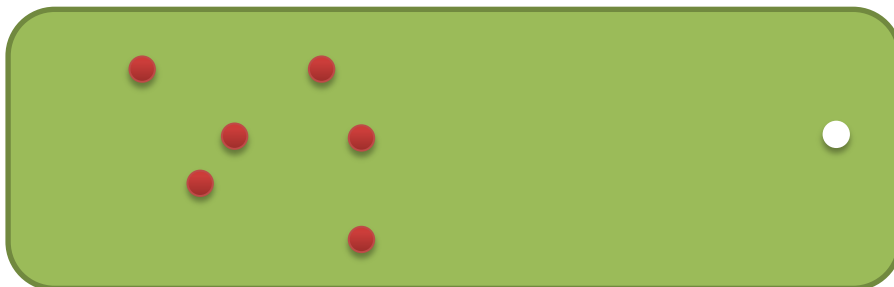


Figure 35

Now imagine hitting the white ball with a specific velocity at a specific angle. According to classical physics the system should be deterministic if we know exactly what the initial conditions are. However if there is even the slightest mistake in the angle or velocity the results could be entirely different. Consider the figure below:



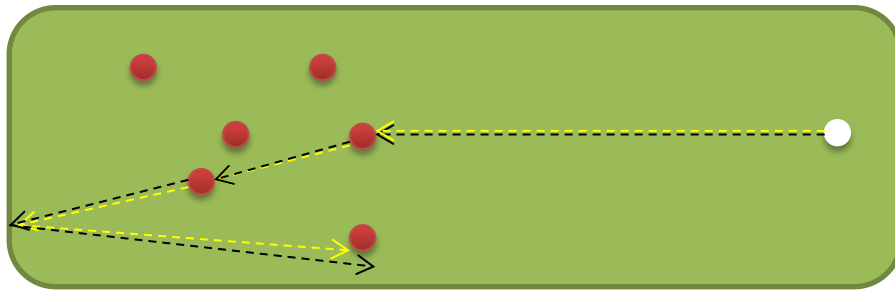


Figure 36

The yellow line shows the trajectory of the balls when the velocity and angle are known exactly and the black line shows the trajectories when the angle is slightly different. The results of this small change can be extremely different over time.

Most systems are complex and chaotic like the snooker table, where the slightest change in initial conditions lead to drastically different results. To look at this in a phase space diagram, we would have something like this:

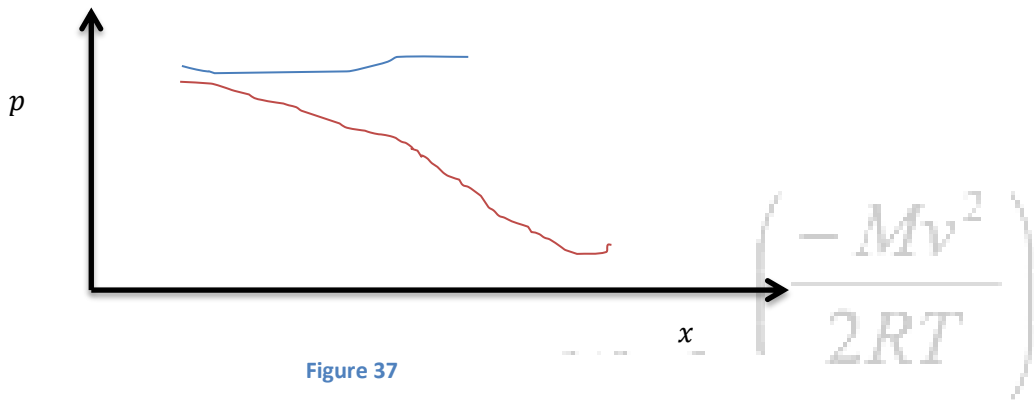


Figure 37

The two lines show trajectories of two systems starting of initially very close together in phase space. If we let the small difference between the trajectories be  $\mathfrak{D}$ , then the difference in the states of the system after some time can be modelled (for a chaotic system) as:

$$\hat{\mathfrak{D}} = \mathfrak{D}e^{\gamma t}$$

Where  $\gamma$  is called the Lyapunov constant and represents the degree of instability of a system. For this small change in the initial conditions what generally happens to a system is that it forms a very fractalated structure in phase space:

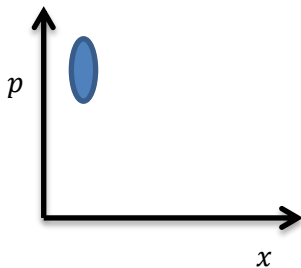


Figure 38

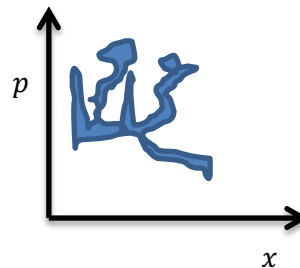


Figure 39

Figure 38 shows the system at initial conditions and figure 39 shows the system after some time has passed.

An interesting point is that the Planck's constant has units of area in phase space, so as the systems become more fractalated there comes a point when it cannot be further fractalated as the lengths of the fractals become of the order of Planck's length.

To see why this fractalated structure for chaotic systems leads to the second law of thermodynamics we have to realise that we can never know exactly the initial conditions. This comes from limitations in measurements as well as the natural limit's from quantum mechanics (in essence quantum mechanics causes chaos, no matter how accurately the measurements are made).

This uncertainty in measurements can be represented by small blobs or boxes in phase space instead of having points:

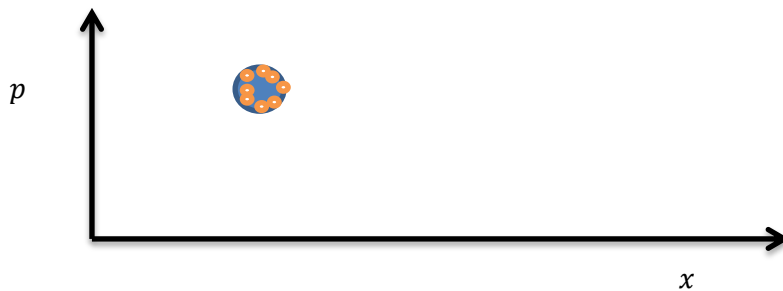


Figure 40

The small orange blobs are the measure of uncertainty in the initial conditions. I haven't drawn all of the blobs but you can imagine the entire volume being full of these blobs. The blobs inside the phase space will slightly increase the volume of the overall phase space, as the blobs on the edge of the volume will extend slightly (as they are no longer considered point like).

This is called coarse grained phase space. After some time we get the fractalated structure like in figure 39, except since it is coarse grained and the fractals become very close and dense, the overall structure of the fractals just becomes one big volume:

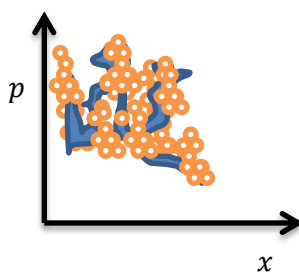


Figure 41

Therefore if we use our definition for entropy like we did previously, we see that it always increases as the volume always increases.

This coarse grained structure also explains why the system cannot be traced back in its trajectory to its original state. The small blobs lead to even more fractalated structure as time increases, we can think of a blob itself being made of many tiny points and that only one of those points actually to the "real" value of the state (strictly speaking these can't be point like, instead have to be small boxes of units  $\hbar$ ).

So we have a very large number of points each leading to fractalated structures and therefore if we trace any point of the fractalated structure it will almost never lead back to the initial state. There will be in general a minuscule number of points that were actually representing the original state of the system. The rest of the points if traced “back” will lead to completely different states as they were never part of the original state and tracing back is just evolving in time. Hence they will actually just go on creating more fractalated structures and making the volume of the states even more complex.

Another important point is that the energy which the initial volume carries will always remain constant. In phase space it is like having a plane of constant energy in which systems evolve (remember such a system must be a closed system, that is, it is not under influence from any external forces).

Chaos also leads to a loss of information (as any weather forecaster will know and we also don't know exactly what the Hamiltonian is. In statistical mechanics and thermodynamics the entropy always takes on this coarse grained structure.

### Harmonic oscillator

In the previous section we discussed a diatomic molecule connected by a rigid rod, we shall now replace that rod with a spring connecting the particle. This time we shall ignore the rotational degrees of freedom for simplicity and consider the potential energy stored in the spring. Just as we did for the modified ideal gas we compute the Hamiltonian:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{kx^2}{2}$$

$k$  is the spring constant, when it is large the spring is stiff and the particle becomes point like. Therefore for low temperatures the internal motion should be frozen out.

We can model the internal vibrations we calculating the partition function for it:

$$Z_{sp} = \int e^{-\frac{\beta p^2}{2m}} e^{-\frac{\beta k x^2}{2}} dx dp$$

Therefore:

$$Z_{sp} = \sqrt{\frac{2\pi m}{\beta}} \sqrt{\frac{2\pi}{\beta k}} = \frac{2\pi}{\beta \omega}$$

Where

$$\omega = \sqrt{\frac{k}{m}}$$

Taking the log we have:

$$\log Z_{sp} = \log\left(\frac{2\pi}{\omega}\right) - \log \beta$$

Now we can calculate the energy:

$$E = -\frac{\partial \log Z}{\partial \beta} = \frac{1}{\beta} = T$$

Remember

$$T = k_B T$$

So the harmonic oscillator always stays the same, so it never gets to being a point particle as we expect from classical physics as this term has no dependence on k.

**Quantum harmonic oscillator**

As was the case with the modified ideal gas, the solution to this problem comes from quantum mechanics and the quantisation of energy as:

$$E = n\hbar\omega$$

Partition function for the quantum mechanical harmonic oscillator becomes:

$$Z_{sp} = \sum_n e^{-\beta\hbar\omega n}$$

This sum is infact a geometric series that converges to the solution:

$$Z_{sp} = \frac{1}{1 - e^{-\beta\hbar\omega}}$$

Therefore for the energy we have:

$$E = \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1}$$

The exponential can be expanded out to give approximate qualitative results:

$$E = \frac{\hbar\omega}{1 + \beta\hbar\omega - 1} = T$$

Which is the same as the classical limit, so is intuitively pleasing. And we see that when the temperature is very small,  $\beta$  is very large the energy tends to zero very rapidly due to the exponential factor and hence the problem of the classical harmonic oscillator is no longer there.

In the argument above we simply outlined the equations that lead to the quantum mechanical solutions. Another way to look at this would be to look at the phase space of the harmonic oscillator system. As we have seen previously, for quantum mechanical behaviour the concept of points in phase space does not exist due to the uncertainty principle. Therefore we have small boxes of area  $\hbar$  as this has units of area in phase space.

However in general this does not have to be a box in phase space, it could be any weird shape as long as the area is  $\hbar$ , so it could be something like:

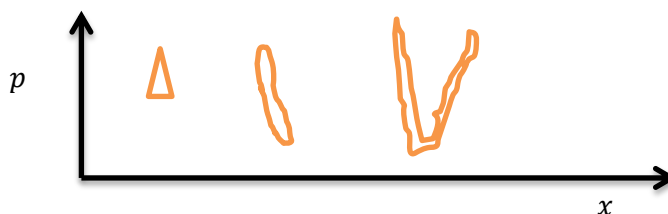


Figure 42

Where all three shapes correspond to the area of  $\hbar$  and as can be seen the shape on the very right side will have  $\Delta x \times \Delta p$  a lot greater than  $\hbar$  compared to the other two. And this clearly satisfies Heisenberg's uncertainty principle.

Now we can write down the Hamiltonian of the system, same as before:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{kx^2}{2}$$

This has a circular shape in phase space, where the area of the circle is  $\hbar$ .

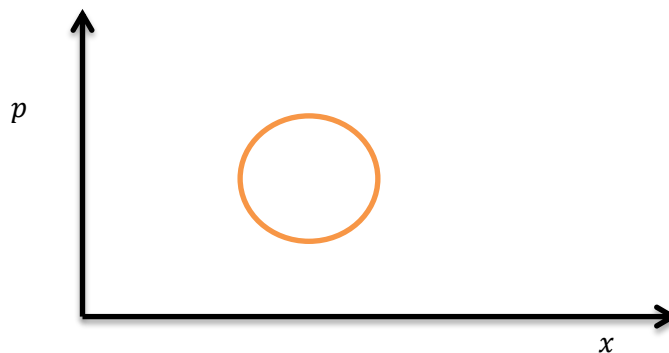


Figure 43

Now classically if we have an increase we would simply get a larger circular which could be of arbitrarily small or large area. However in quantum mechanics we know the energy is quantized as we stated before. Therefore the higher energy states form concentric rings around the circle in the ground state (let's assume the circle in the graph above is in the ground state) with the area between the rings always equalling  $\hbar$ .

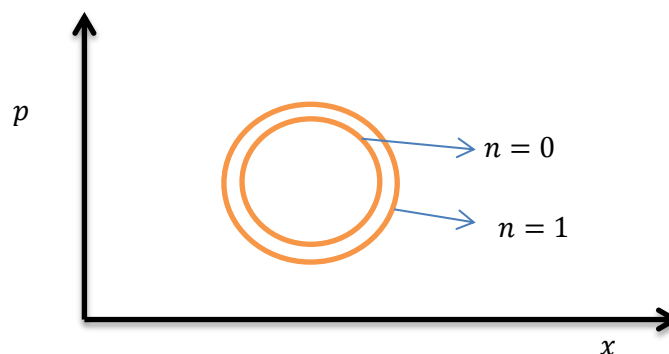


Figure 43

So we can immediately tell that the area between the rings must increasingly get smaller for large  $n$  (this just represents that the energy levels become closer and closer in the traditional energy diagrams).

For an  $n$ th state, the area of the entire disc (not the rings!) is simply  $n\hbar$ . We also know that the area of a circle is

$$A = \pi r^2$$

Therefore equating this to the quantum mechanical area gives:

$$\pi r^2 = n\hbar$$

Therefore the radius is:

$$r = \sqrt{\frac{n\hbar}{\pi}}$$

However now we see that from the geometry of the phase space:

$$r = \sqrt{\frac{p^2}{2m} + kx^2}$$

Therefore we once again obtain the energy being quantised as  $n\hbar$  by equating the equations written above.

In thermal equilibrium these rings have a probability distribution attached to them, which is just the Boltzmann distribution.

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

## Black body radiation

There was a big puzzle attached to the black body radiation at the start of the 20<sup>th</sup> century. To outline the problem, consider a cavity full of radiation within it and it has completely reflective walls. The radiation can exist in different wavelengths; classically there are an infinite number of wavelengths (as the wavelengths can just be divided up in smaller and smaller intervals). The radiation can also be thought of as a collection of harmonic oscillators. In the same sense, that a vibrating violin string can be thought of consisting many harmonic oscillators (by that I mean the vibrations can have different wavelengths).

Now if we consider the system in thermal equilibrium, each harmonic oscillator should have energy  $kT$  according to classical thermodynamics (equipartition theorem). However there are an infinite number of possible oscillation modes (infinite number of possible wavelengths, each corresponding to an energy of  $kT$ ). Therefore in thermal equilibrium the system should have an infinite amount of energy! Of course this was not true, infact experiments showed that there was a minimum wavelength of the system in which the radiation could be, and the results also showed that this wavelength depended on temperature.

Well it was relatively easy to accept that there was a smallest wavelength limit in nature, however the fact that this energy depended on temperature required explaining. However we can now just look back at our discussion on the quantum mechanical harmonic oscillator in which we just saw that the energy is quantised in discrete wavelengths (frequencies, but it's the same thing!).

Before any concrete theory of quantum mechanics was developed, Planck postulated this relation as an explanation to this problem, however he had no strong foundation to support why this would be the case, he simply got the equations right!

These quanta's of radiation that Planck postulated are what we know call photons. Einstein was the one who came up with a complete theory of this in 1905 with the explanation of the photo-electric effect and he knew very well that he has infact showing a particle like behaviour of the light.

We want to calculate the total amount of radiation in the box, so we need a clear description of the radiating oscillators. If we take the box to be completely reflecting, then we have boundary conditions that the oscillations at the walls are zero (analogous to a guitar string).

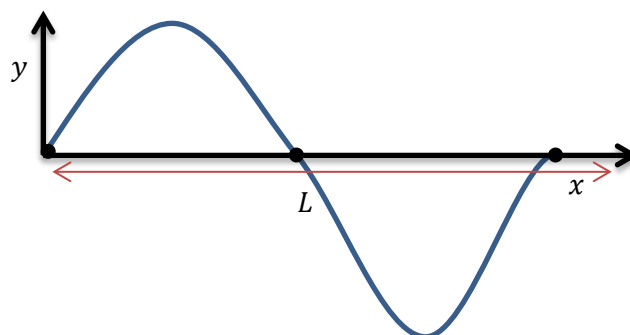


Figure 44: Vibrating motion of a string

The motion of the string can be modelled by solving the wave equation as we have done many times, therefore I will simply quote the result:

$$\phi(x, t) = \sum_m Y_m(t) \sin \frac{m\pi x}{L}$$

Where L is the length of the walls (or any other thing that is used as a boundary) and m is any integer value.

This is for a 1 dimensional wave; this can simply be extended to 3 dimensions as:

$$\Phi = \sum_{m_x, m_y, m_z} Y_{m_x, m_y, m_z}(t) \sin \frac{m_x \pi x}{L} \sin \frac{m_y \pi y}{L} \sin \frac{m_z \pi z}{L}$$

This completely describes the motion for one harmonic oscillator. The frequency is just:

$$\omega = |k|c$$

Where:

$$|k| = \sqrt{k_x^2 + k_y^2 + k_z^2}$$

Now we can find the total energy of the harmonic oscillator (HO) by summing up these waves. Classically we saw that the energy became infinite, as the energy in every HO was the same (1/2 kt).

We saw in the section for QHO, the energy was given by:

$$E = \frac{\omega \hbar}{e^{\beta \omega \hbar} - 1}$$

Substituting for  $\omega$  gives:

$$E = \frac{|k|c\hbar}{e^{\beta|k|c\hbar} - 1}$$

Therefore the total energy is:

$$E_T = \sum_{m_x, m_y, m_z} \frac{|k|c\hbar}{e^{\beta|k|c\hbar} - 1}$$

This is a difficult sum to deal with as it is, therefore we have to think of a different way to go finding the energy. We know that:

$$k = \frac{m\pi}{L}$$

If L is large compared to the wavelength then the distance between neighbouring values of m is small. Therefore if we sum over k instead of m, then we can approximate the sum for an integral:

$$\sum_{m_x, m_y, m_z} \frac{|k|c\hbar}{e^{\beta|k|c\hbar} - 1} = \frac{1}{\Delta k^3} \int \frac{|k|c\hbar}{e^{\beta|k|c\hbar} - 1} dk^3 = E$$



This integral is done by substitution and using a standard integral, therefore I will simply quote the result here:

$$E = \frac{L^3 \pi^2}{\beta^4 c^3 \hbar^3}$$

This is the Stefan-Boltzmann law (we are just missing a 2, since the photon can be polarized in two orthogonal directions), once again the conversion from the quantum to the classical regime occurs when:

$$k_B T \sim \omega \hbar$$

**Radiation pressure**

The pressure is given by:

$$\frac{\partial A}{\partial V_T} = -P$$

A is as usual the Helmholtz function:

$$A = -T \log Z$$

Therefore pressure is:

$$\frac{2}{3} \frac{E}{V} \left( \frac{-Mv^2}{2RT} \right)$$

**Number of photons and entropy**

The number of photons in a given oscillation mode:

$$\frac{1}{e^{\beta \hbar \omega} - 1}$$

Therefore the total number of photons is simply the sum over all oscillation modes:

$$\int \frac{1}{e^{\beta \hbar \omega} - 1} d\omega \sim T^3$$

This integral is a convergent integral

The entropy on the other hand is:

$$S = \frac{E - A}{T}$$

We know that E and A both have units of energy and therefore go as  $T^4$ , hence entropy must go as  $T^3$ . So we see that entropy is the same as the total number of photons, to within a factor of a numerical constant, showing that it is just the information for a given photon.

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

## Magnetic Systems

A magnetic system is a collection of little magnets that align themselves in a strong magnetic field. In general they can point in any direction, further they are like small compasses in 3-D space. They have interactions between them that favour parallel aligned neighbours, these interactions are short range therefore interaction between neighbouring magnets are the only ones that we need to consider. When magnets want to align in the same direction they are known as Ferro magnets, when they prefer to align anti-parallel to each other they are known as anti-ferromagnetic.

At very low temperature the lower energy states are important, and they all have the magnets aligned. The energy from them aligning parallel is now felt as the overall energy of the system is small. We make mathematical models for these magnets as we shall discuss below.

### Non interacting magnets (Z2)

These magnets are modelled as having no interaction between magnets and they can align in only two ways, up or down. There is an energy function that favours neighbouring magnets to be aligned. At low temperatures the magnets either point up or down (let's call them spins from now). The ground state is therefore 2-fold degenerate. A tiny B field in the vicinity of the spins will change the direction of the spins and cause them to align with it; this is called spontaneous symmetry breaking. This means the spins do not align randomly; instead there is a bias to either up or down (whichever the direction of the B field might be). Many systems in statistical mechanics can be modelled like magnets in this way as I shall discuss below.

Consider a fluid (a collection of molecules that are free to move) in a box.

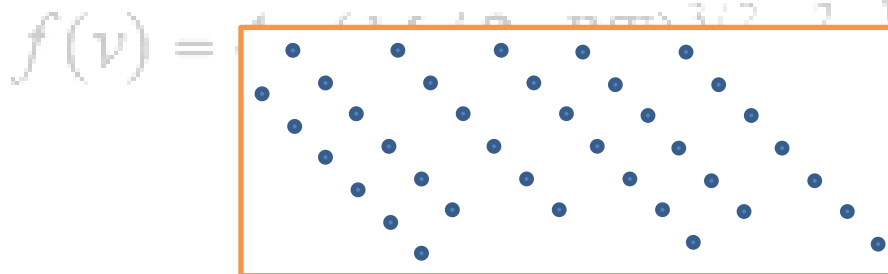


Figure 45: Box of molecules

Now imagine breaking up the box into cells that are so small, they are just the size of one molecule. Only one particle can fit into one cell, therefore each box either has a particle or not, analogue to the spins being up or down. Coming back to the spins, if we apply a small B field point down, the spins will align downwards. If we slowly bring the B field back to zero, the spins should remain pointing down. Now if the B field is pointed in the opposite direction, if spontaneous symmetry breaking (SSB) occurs the spins will face upwards immediately. If this were to happen for the fluid, it would mean the cells would go from being either empty (less dense) or full (more dense) to the opposite, which is an example of a phase transition. The question now is, what corresponds to changing the B field for a gas that will cause the phase transition and this turns out to be the chemical potential energy.

### SSB in mean field approximation

Consider a lattice with atoms in it that can have two configurations:

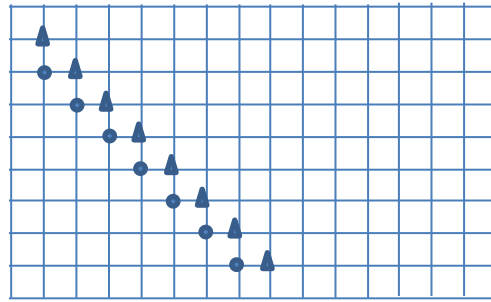


Figure 46: 2-D lattice of atoms

In general we could have a  $\mathcal{D}$  dimensional lattice. For a mean field approximation we are doing what the name suggests, taking the mean over a given field. Therefore the more the number of neighbouring molecules, the better the approximation works (for larger dimensions we have more neighbouring molecules).

Now we can postulate an energy function for the system in Fig 46, let's say that energy is minimised when the symbols are the same for the neighbouring particles:

$$E = -\sigma_i \sigma_j$$

So we see that when the symbols are the same the energy is minimised. Therefore when a sign change occurs we have a change in energy of 2 units (-1->1). We can call this process the breaking of a bond. A broken bond will have two opposite symbols at either end:

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

Figure 47: Broken bond

Broken bonds cost 2 units of energy therefore it favourable for a system to minimise the number of broken bonds, however as we raise the temperature, the number of broken bonds increases.

Now consider a single molecule, its energy is:

$$E_{\sigma_o} = -\sigma_o \sum_{NN} \sigma$$

Where the sum represents the sum over the nearest neighbour molecules. Suppose the spins have an average value:

$$\langle \sigma \rangle = S$$

Where S is between -1, 1, this is called the mean field. Therefore we can write a general equation for the energy:

$$E_{\sigma_o} = -\sigma_o 2\mathcal{D}S$$

Where the  $2\mathcal{D}$  represents the number of nearest neighbours in  $\mathcal{D}$  dimensions. If the system is assumed to be magnetised ( $S \neq 0$ ), the partition function is:

$$Z_{sp} = \sum_{\sigma=-1}^1 e^{\beta 2\mathcal{D}S\sigma} = e^{2\beta\mathcal{D}S} + e^{-2\beta\mathcal{D}S} = 2 \cosh 2\beta\mathcal{D}S$$

Therefore the energy is:

$$\langle E_{\sigma_0} \rangle = -2\mathcal{D}S \tanh 2\beta\mathcal{D}S$$

From the general equation for energy we can write down the average value of the spin:

$$\langle \sigma_0 \rangle = -\frac{\langle E_{\sigma_0} \rangle}{2\mathcal{D}S} = \tanh 2\beta\mathcal{D}S$$

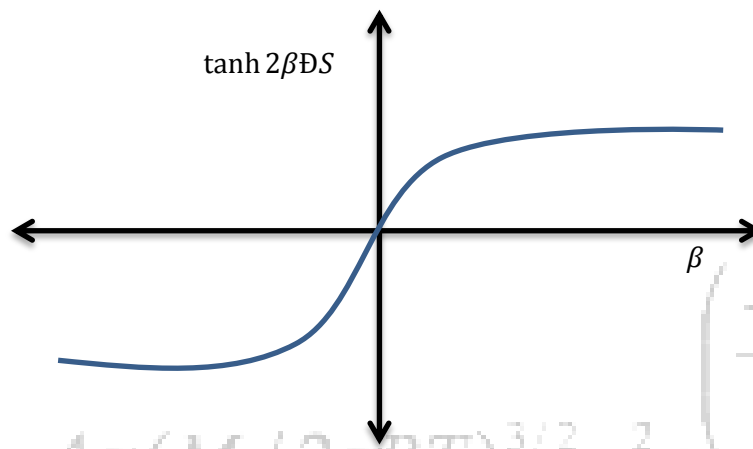


Figure 48

This only makes sense if:

$$\langle \sigma_0 \rangle = S$$

$S$  is known as the magnetisation.

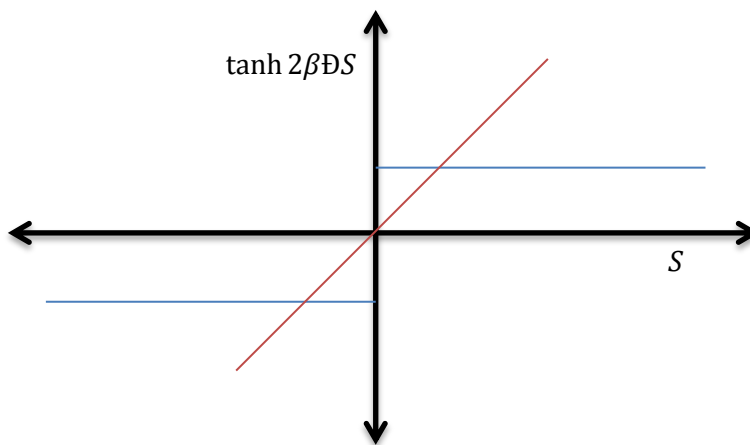


Figure 49

When the temperature is large the curve spreads out in S, the straight red line in the graph representing the magnetisation intersects the curve only once at the origin when the temperature is large (as the curve spreads away from the straight line). However as we decrease the temperature the curve becomes a box function like the one shown in Fig 49 and the magnetisation crosses it 3 times. The two non-zero magnetisations are the average values of the spins. If the system was bias with an external magnetic field, it is only these two solutions that persist.

As the temperature gets to the point of:

$$T = 2DJ$$

The interaction points come down until they reach the origin; this means that the magnetisation is non-zero until it reaches this temperature, also known as the Curie temperature. This temperature is what corresponds to the changing of the magnetic field in a system of spins. This is called the Ising magnet as well as the Z2 magnet. The energy of this magnet is:

$$E_i = -\sigma_i \sigma_j J$$

Where J is a constant and must have units of energy since the sigmas do not have a unit. The basic method for a mean field approximation is commonly used in physics (we used it in PH2510), to summarise:

- If we want to study a particle in a large system with many d.o.f's, then we can average the rest of the d.o.f and then set 1 d.o.f we want to study equal to the rest of the d.o.f average.
- However the main limiting factor for this approximation is that the dimensions of space have to be increased to give accurate results.

Now if we consider flipping two spins in 1-D, the energy input required will be 4 units (remember 2 units for every spin), but the other nearest neighbour will also require an energy of only 4 units and so it will also flip, therefore there is an instability in 1-D systems. However in systems in large dimensions this is not a problem as flipping two spins will cost 4 units of energy however in 2-D there are more neighbours therefore the amount of energy needed in more for the neighbours and therefore there is a stability in 2-D structures.

Now if we consider the phase diagram of the spins against the temperature it looks something like this:

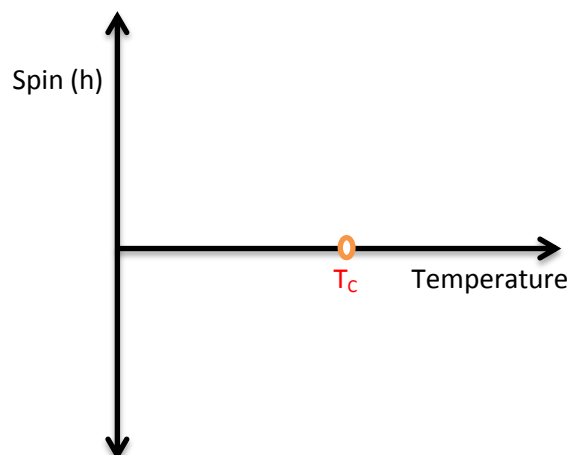


Figure 50: Phase diagram

If  $h$  is less than zero it means that the spin is pointing down. Suppose we apply a  $B$  field that makes the spins align down in the first place, then we slowly remove the field, so the spins are still pointing down. On the phase diagram it would mean that we are approaching the  $h = 0$  line from below. If we assume that there is some magnetisation even when the  $B$  field has gone to zero, there will be a jump discontinuity at the point of  $h = 0$ , as all the spins will jump from pointing down to pointing up. This is an example of a first order phase transition if we are talking about a gas. However this only happens when the temperature is below the critical point (we can make some intuitive sense of this because a higher temperature means the spins or gas molecules will have more kinetic energy and will be able to fluctuate a lot more, hence there will not be an abrupt change in density or the direction of spin), the temperature exceeds the critical point then the phase change is gradual and is known as a second order phase transition.

**Chemical Potential**

Imagine a set of boxes, where each of the boxes can have some energy and the energy can be exchanged by a weak interaction. Each box will have a probability distribution to have certain energy.

Suppose we want to maximise the entropy subject to the constraint that the energy is fixed:

$$\sum_{i=1}^m n_i E_i = \text{constant}$$

So we use the method of Lagrange multipliers, and the probability distribution comes out of the form:

$$f(v) = 4\pi \left( \frac{M}{2\pi RT} \right)^{3/2} v^2 e^{-\beta E} \left( \frac{-Mv^2}{2RT} \right)$$

$PD = \frac{e^{-\beta E}}{Z}$

Suppose that the particles are exchanged between them as well. Then we have to add another constraint:

$$\sum_i n_i = N$$

This constraint basically states that the sum of all the particles in all the different states is the same total number of particles.

We also know that:

$$\frac{n_i}{N} = P(i) \text{ (probability of a box in state } i \text{)}$$

Therefore we can use this to say:

$$\sum_i \frac{n_i}{N} E_i = \frac{E_T}{N}$$

$$\sum_i P(i) E_i = \langle E \rangle$$

The entropy is:

$$S = - \sum_i P(i) \log P(i)$$

Using the Lagrange method a new equation is formed:

$$Y = - \sum_i P(i) \log P(i) + \beta \sum_i P(i) E_i$$

Where  $\beta$  is a Lagrange multiplier, to maximise this it is differentiated and set equal to zero:

$$\frac{\partial Y}{\partial P(i)} = - \sum_i \log P(i) + 1 - \beta E_i = 0$$

And hence we arrive at the probability function:

$$P(i) = \frac{e^{-\beta E_i}}{Z}$$

A similar method for the other constraint gives:

$$P(i, N) = \frac{e^{-\beta E_i} e^{-\beta \mu N}}{Z}$$

Where  $\beta\mu$  is the traditionally called the new Lagrange multiplier. With a little bit of thought it can be seen that  $\beta$  is the variable that controls the temperature and hence the energy in the box and  $\mu$  is the variable that controls the number of particles in the box and is called the potential energy.

Now if we return to the phase diagram in fig 50, if we want to see what happens as we approach the critical temperature at  $h = 0$ , we first need to understand the idea of a correlation length.

Suppose the average of spins at point a and b is given by:

$$\langle \sigma(a)\sigma(b) \rangle$$

This is called the correlation function; at high temperatures the correlation is likely to be zero due to the random fluctuations of the spins. As T is lowered the spins tend to become parallel, hence the correlation function is likely to be positive and approach 1 (as the spins will either align up or down). We expect to lose the same fraction of correlation as we move away from the nearest neighbouring particles. It could be something like:

$$\langle \sigma(a)\sigma(b) \rangle = e^{-m(T)D}$$

D is the distance between spins and m is a temperature dependent constant, plotting this will give:

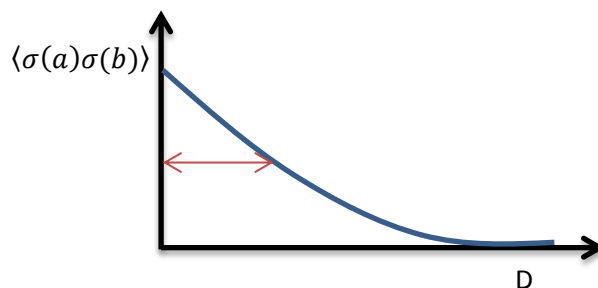


Figure 51: Correlation function varying with distance



The red line on the graph represents the correlation length with varies with temperature (is dependent on the parameter  $m$  that we defined in the exponential term).  $m$  represents the distance that needs to be travelled before the correlation function drops by a factor of  $e$  for a given temperature.

Now imagine going below the critical temperature, now all the bias of spins is in one direction, therefore the correlation length will become infinite and the correlation function will approach a constant value.

As the critical temperature is approached the curve falls of as a power law, as supposed to an exponential:

$$y = \frac{1}{x^n}$$

Suppose we look microscopically at the magnet as we approach the critical temperature, we see:

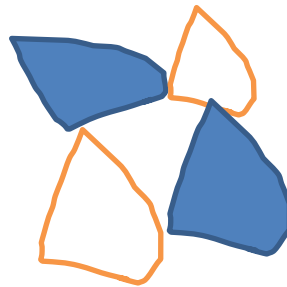


Figure 52: microscopic view of the magnet near critical temperature

$$f(v) = 4\pi(M / \omega\omega_1)^2 v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

We start seeing patches of spins all aligned equally; where the two differed colours represent the two different configurations of the spins. These patches are of course usually approximately the size of the correlation length. It is important to note that these patches don't show all the spins pointing in the same direction, instead they shown patches of a preferred spin. As the temperature approaches the critical temperature, these patches become extremely large.

They effectively cover the entire sample (as the correlation length approaches infinity) and they become extremely faded, meaning it is very difficult to tell between regions of different spin configuration preferences. In fact if we zoom in towards a patch we see a self-similar structure (similar to fractals), meaning that at every scale the structure looks the same. It is also known as scale invariance.

Therefore the structure becomes completely isotropic, which is not expected as we do not expect the correlation length to be the same along the axis and at an angle of  $45^\circ$  for example. Therefore it can be said that at the critical temperature the system loses memory of its original structure. If the system was modified to include the second nearest neighbour interactions aswell, then the system will reach the same state once the critical temperature is obtained, the only difference will be in the value of the critical temperature and when the system is at high temperatures.

## Inflation

Inflation is the exponential expansion of the universe, driven by what is sometimes known as dark energy or vacuum energy or the cosmological constant, essentially it is the name given to the potential energy stored in the scalar field. A process called “reheating” was postulated to take place at the end of inflation. The form of this potential energy is completely unknown and therefore it was simply postulated that it was of the form:

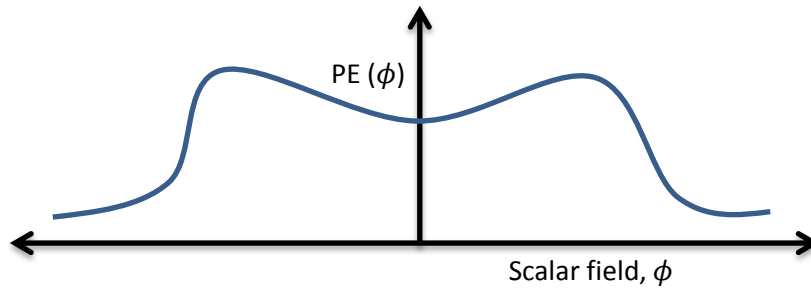


Figure 53: Postulated vacuum energy

The scalar field has a kinetic energy as well as a potential energy:

$$KE = \frac{\partial \phi}{\partial t}$$

$$PE = \frac{\partial \phi}{\partial x}$$

$$\left( \frac{-Mv^2}{2RT} \right)$$

Sometimes the potential energy is just combined as the kinetic energy due to combining of space and time in relativity. However this field also appears to have a potential energy that is independent of any derivatives and that is potential energy shown in fig 53. It is postulated that the universe began in a state of stable equilibrium (the reason is not known) at the point where the scalar field is zero. At this point the potential was postulated to be very large and it got even larger as we moved away from the  $\phi = 0$  point, before reaching a sharp drop. Therefore close to the centre position this system looks like a particle in a potential well, and the universe can, given sufficient time tunnel through the potential barrier just like a particle in a box. Therefore once it tunnels through the potential barrier, the universe drops down the potential hill and over time this potential energy is converted into heat (mechanism is unknown) and this stage is what became known as reheating.

The problem is that unlike a particle in a box, the universe is many many particles in a box! And it is highly unlikely that all the particles in the universe would end up with a higher energy at exactly the same time. It is more likely that a small patches of the universe undergoes tunnelling and if the patch reaches a sufficiently large size it acts as some kind of a catalyst for the rest of the universe to undergo tunnelling an analogy to this is of the behaviour of really cold water (below freezing) that solidifies (undergoes a first order phase transition) when a block of ice is added. However this model turned out not to be true, because the mathematics of inflation showed that the expansion of the universe was much faster than the spreading of the small patches of the universe that had a greater energy could catalyse the tunnelling and therefore this transition would never be completed.

**Adiabatic variation of parameters**

We know the probability distribution of the system being in a given state has the form:

$$P = \frac{e^{-\beta E}}{Z}$$

Now suppose we vary the energy and the system and that the final energy and the initial energy are linearly related.:

$$E_F = a + bE_i$$

The probability distribution then becomes:

$$P = \frac{e^{-\beta(a+bE_i)}}{Z}$$

However this is not very helpful as we just get two new probability distributions. If a small system with only a few d.o.f is given and we change the parameters it will not stay in thermal equilibrium. But if the system is large it does stay in thermal equilibrium to a good approximation. The energy of a large system can be written in terms of the energy density  $\epsilon$ :

$$E_I = \epsilon_I V$$

And for the final energy:

$$E_F = \epsilon_F V$$

Once again  $\epsilon_I$  can be a linear function of  $\epsilon_F$ . As usual for a large system the energy levels become very dense and we can approximate sums with integrals. As the density is large for large energies, the density is a rapidly rising function when plotted against energy.

The partition function can be written as:

$$Z = \int e^{-\beta E_i} \frac{dN}{dE} dE$$

The exponential term is a very rapidly falling function and the number of particles per energy is a very rapidly increasing function (as the density increases for large energies) . Hence the number of energies that are actually available have a very limited range.

Over this range the function must be linear and we expect the density to be a smooth function. Since the number of states are quantised we are simply counting them and therefore the variation in the energy follows Poisson statistics:

$$E = \epsilon V$$

$$\Delta E \sim \sqrt{V}$$

$$\frac{\Delta E}{V} = \frac{1}{\sqrt{V}}$$

In this case when V becomes large  $\epsilon_I$  can be linear with  $\epsilon_F$ . Therefore we must be dealing with a a large system. The probability of a small system being in thermal equilibrium after in a change in adiabatic parameters is very small.

**Gravity and thermodynamics**

A system which interacts solely via the gravitation force will have a negative heat capacity. To see why consider the small object orbiting a large object, as it orbits it will radiate away energy be that in EM waves or gravitational waves. Hence it will start to loose energy and spiral inwards. As it does so it will gain potential energy (which is negative) and kinetic energy. The Virial theorem shows that the increase in kinetic energy is always half of the increase in potential energy, this means that the system will have a negative heat capacity as the thermal energy will be equal to the kinetic energy and hence even though the overall energy of the system will be decreasing the temperature will be increasing! As shown with the black hole thermodynamics, these systems are incredibly unstable.

**Negative temperature**

Ordinarily temperature cannot be negative as the Boltzmann function will have a positive exponent and will not converge on to a value. The only way to have a negative temperature would be to have a system in which the energy is bound somehow. A good example is the system of magnetic spins as the energy can simply have two values. We saw previously that the partition function for this system was:

$$Z = 2\cosh\beta$$

In this case having a negative temperature would make no difference as the function of cosh is symmetric about the origin, if we make a plot of energy against entropy we get something like this:

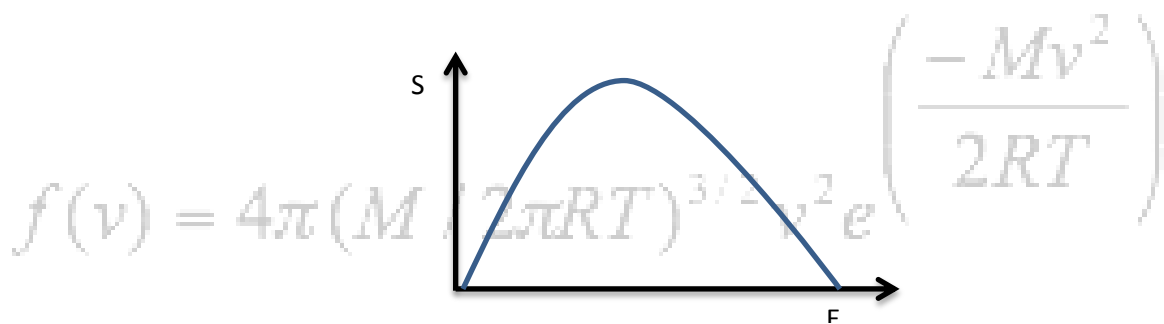


Figure 10

We know that temperature is

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

Which is inverse of the gradient of the curve, so we see that the first part of the curve is increasing hence has a positive temperature, the stationary point has the gradient 0 and hence has an infinite temperature, after the stationary point the curve is decreasing and hence has a negative temperature. In this magnetic system the temperature is negative if more than half of the spins are pointing upwards, this is called population inversion.

## References

[1] <http://farside.ph.utexas.edu/teaching/sm1/lectures/node25.html>

[2] [http://en.wikipedia.org/wiki/Landauer's\\_principle](http://en.wikipedia.org/wiki/Landauer's_principle)

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$

$$f(v) = 4\pi (M / 2\pi RT)^{3/2} v^2 e^{\left( \frac{-Mv^2}{2RT} \right)}$$