

Lecture 5

Thermal Equilibrium

When a hot and a cold body are placed in thermal contact, that is they can exchange energy, we know that heat always flows from hotter to colder body. Consequently, the temperature of the two bodies changes. The hot body loses temperature and the cold body gains temperature. After sufficiently large time, there will be no more exchange of energy between the two bodies and their temperatures will be same. The two bodies are said to be in "thermal equilibrium" now. So, thermal equilibrium is achieved when there is no exchange of heat, and the temperatures of the two bodies are same. The process that leads to thermal equilibrium is called thermalization.

The thermalization process is irreversible, that is, if you take two bodies and make them in contact, which have the same temperature, they will never go backward, that is away from thermal equilibrium. Thus, thermal

processes denote an arrow of time, called the thermodynamic arrow of time.

If various bodies are in thermal equilibrium with each other, then their temperatures ~~are~~ ^{should be the} same. Indeed, it is true, and the idea is incorporated in zeroth law of thermodynamics:

If two systems A and B are in thermal equilibrium with ~~each other~~, a third system C, then A and B must be in thermal equilibrium with each other. We will review this things in more detail in classical thermodynamics.

Microstates and macrostates

Let's think about an example to visualize the idea of microstate and macrostate. Say, you have 100 number of fair coins, and you tossed them. How many different configurations are possible here?

Let's start with one coin. We have only two possible configurations, each of them equally probable - H and T. If we have two coins, then, for each probable outcome of the first coin, we have two outcomes. The total number of configurations is four — HH, HT, TH, TT. For a ~~the~~^{number} of three coins, you have eight outcomes. So, the number of configurations for N coins, will simply be 2^N . Here, for our 100 coins, you have 2^{100} number of distinct probable configurations, and each of them are equally likely, like before. In terms of power of 10, this will be roughly 10^{30} (check). Since all the configurations are equally probable, the probability of one of them occurring is 10^{-30} . We will call each of the particular configuration a "microstate" of the system.

But, when you would like to measure something in the system, you would probably calculate the number of heads and number of tails after a particular trial. Say, after you tossed 100 coins, you got 55 heads and 45 tails. This configuration (55 heads and 45 tails), we will call a macrostate of the system. Now although all the microstates are equally likely,

$\text{Total number of microstates} = N^{2^N}$

the macrostates are not. For example, out of 100 coin tosses, the ~~probability~~^{number of configurations} that you are going to get 50 heads and 50 tails is = $\frac{100!}{50! 50!} \approx 4 \times 10^{27}$

$$\begin{array}{ccccccc} 53 & " & " & 47 & " & " & \\ 90 & " & " & 10 & " & " & \\ 100 & " & " & 1 & " & " & \end{array} = \frac{100!}{53! 47!} \approx 3 \times 10^{27}$$

$$= \frac{100!}{90! 10!} \approx 10^{13}$$

$$= \frac{100!}{100! 0!} = 1$$

So, the probability that you will get 100 heads and 1 tail is $= \frac{1}{10^{30}} \approx 10^{-30}$. However, the probability of getting

50 heads and 50 tails $= \frac{4 \times 10^{27}}{10^{30}} \approx 4 \times 10^{-3}$, which is quite a good probability. As you number of coins goes up and up, this probability will increase more and more. You can verify that, for 1000 coins, the probability of 50 heads and 50 tails becomes ≈ 0.98 . The thing is, although all the configurations are equally probable, there are just more number of microstates with 50 H and 50 T, which makes this macrostate to have the most probability of occurring.

So, one take away from this example is that, if all the microstates are equally likely, then the most likely microstate that the system will find itself is the microstate that corresponds to the maximum number of microstates.

We could also consider the example of throwing fair dice, say 100 of them. We put them in a cup, shake for sometime, and then throw them on the table. The essential ingredients necessary for an analysis are stated below -

- (i) Specification of the state of the system: We need ~~detailed~~ detailed method for describing the outcome of the experiment. For our current problem, specification of the state of the system requires a statement about which face is up for each of the 100 dice.
- (ii) Statistical ensemble: In principle our problem of dice throwing is deterministic. If you exactly ~~knew~~ knew the position of each dice, how you were shaking and how you are throwing etc., you could use the laws of classical mechanics to find out the exact outcomes of the

throw of 100 dice. But, it is next to impossible to know all these things. Instead, we describe the experiment in terms of probability. Instead of thinking about a single experiment, we consider an "ensemble" consisting of many such experiments, under the same conditions. The set of such identical experiments is what we will call an ensemble. The outcomes of these experiments will be different in general, but we can ask the probability of a particular outcome (say 10 number of 1, 32 number of 2, and so on). We could also ask the probability of 1 occurring. We would simply divide the number of 1 that occurred by the (number of dice \times ensemble number). This procedure shows how to calculate probability experimentally. Our theoretical aim is to predict this probability with ~~a~~ fundamental postulates.

(iii) Basic postulate of "a priori" probability: To make

theoretical progression, we must introduce some basic postulates. Our knowledge about the situation demands that, for a ~~equally~~ regular dice with uniform density, there is nothing in the laws

of mechanics that will prefer one particular face of the dice to others. So, we postulate that each of the faces of the dice has equal probability of landing ~~up~~ up. This is equal "a priori" (that is based on our notion, yet unverified) probability.

The postulate is obviously reasonable, and does not contradict any laws of mechanics. Now, whether the postulate is valid or not has to be decided by making theoretical prediction based on the postulate, and checking this prediction are confirmed by experimental observations. If such predictions are repeatedly verified, the validity of the postulate can be accepted with increasing confidence.

(iv) Probability calculations: Once the basic postulates are adopted, the theory of probability allows us to do theoretical calculations of the probability of outcome for any experiment with these dice. The more these calculations matches with experiments, the more confidence you gain about your postulate.

Statistical formulation of the mechanical problem

Specification of the state of the system

Consider any system of particles, no matter how complicated (a gas, a liquid, a car). The particles in that system can be described by the laws of classical mechanics, and in a finer detail and accuracy by quantum mechanics. Although atoms and molecules are properly described by the laws of quantum mechanics, sometimes classical mechanics gives you an useful approximation.

But, how do we describe the state of the system? Consider a single particle in one dimension. This system is completely described by its position coordinate q and momentum coordinate p . One can represent the situation geometrically by drawing Cartesian axes q and p . Specification of p and q of the particle is equivalent to a point in this two dimensional space (the phase space). As the coordinate and momentum changes, the point moves in the phase space. To describe the situation in terms of possible states, we can divide the phase space into cells of sides ~~$\delta q \cdot \delta p = h_0$~~ , where

where δq and δp gives the subdivision of q and p . \hbar_0 is some small constant having the dimension of angular momentum. The state of the system is denoted by stating that the coordinates lie between some interval q and $q + \delta q$ and momentum lies between p and $p + \delta p$, that is, by stating the representative point (q, p) lies in a particular cell in phase space. The specification of the state becomes more and more precise as $\hbar_0 \rightarrow 0$. Although it is permissible in classical mechanics, there is a restriction in quantum mechanics by Heisenberg uncertainty principle ($\delta q \delta p \geq \hbar$). So, in QM, the smallest possible area of phase space is \hbar . We can extend the idea now in 3D and more number of particles. Such a system will be described by f number of position q_1, q_2, \dots, q_f and f number of momentum p_1, p_2, \dots, p_f coordinates, where f is the number of independent coordinates needed to describe the system, also called the degrees of freedom (for a system of N point particles in 3D is just $3N$). The phase space will now be of dimensional and we would require the volume

of $Sq_1 \dots Sq_f, Sp_1 \dots Sp_f = h_0^f$ for the cell in phase space
We can define the states by asking which cell corresponds
to our system.

There are also other ways of defining the state, like in
terms of energy. You could define a system having
a specified energy, and ask the possible configuration
where atoms have different amount of energies such that
the total energy is the same. The energy idea is also
inherent in the phase space picture, specifying positions
and momenta gives you information about energy.

Statistical ensembles

Since we are going to use probability to describe our
systems, and we do not know the exact specification
about all the constituents, our approach is to ~~is~~
imagine repeating an experiment to measure a property of
the system again and again. In an attempt to formalize this, Josiah Willard Gibbs introduced the concept of
ensemble. This is an idealization in which one considers
making a large number of mental "photocopies" of the
system, each one of which represents a possible
state that the system can be in. The system, in

general will be in different states, and will be characterized by different macroscopic parameters (such as pressure, ~~to~~ volume etc.). We want to find the probability of occurrence of a particular value of ~~#~~ such a parameter, that is ~~number~~^{fraction} of cases in this ensemble the parameter assumes a particular value.

Microstate: As in our discussion, a microstate of a system will be a configuration where you specify the microscopic parameters (position and velocity or perhaps energy) of each and every atoms in the system.

Macrostate: A state of a system that ~~can~~ be specified by the macroscopic properties of the system, such as the pressure, volume or total energy of the system.

A macrostate of a gas with pressure 10^5 Pa and volume 1 m^3 will in general be associated with a very large number of microstates. We are not

going to care about what those microstates are, but rather consider their existence and make theoretical predictions subjected to basic postulates.

Basic postulates

Suppose our system under consideration is isolated and thus exchange no energy with the surroundings. So, the laws of mechanics states that the total energy of the system is conserved. Thus, the system can be characterized by this value of energy. However, there are great many number of states corresponding to this energy, and the system can be in any one of them.

We can consider simple case of a system that is in thermal equilibrium. But what do we mean by equilibrium? Equilibrium means all macroscopic properties of the system such as pressure, total energy, volume, total momentum etc. are independent of time. ~~Given~~ So, the probability of finding the system in any of the states is also independent of time. This is because we are saying the average properties are independent of time, and the averages are taken w.r.t.

some probability distribution. You can certainly say that the averages are time independent, if the probabilities themselves are time independent. \ominus

Now, given that the system is in ~~a~~ equilibrium and isolated, the only information we have is that the system must be in one of its accessible microstates consistent with the constant energy. Now, ^{there is nothing in} the laws of mechanics that would prefer one particular state to other. Hence, it seems reasonable that the system is equally likely to be found in any one of the accessible microstates. This leads us to the fundamental postulate of equilibrium statistical mechanics —

for an

"In a state of thermal equilibrium, and isolated system, all the accessible microstates of the system are equally probable."

This is known as equal a-priori probability. This is a postulate, and can not be proven from more fundamental laws of physics (at least until now). However, the converse can be proven. That is, if at any instant of time, all the microstates are equally probable, then they remain ^{equally} probable for all time, so ^{the system is} ~~they are~~ in equilibrium. This is a consequence

on "Liouville's theorem".

What do we mean by accessible microstates?
Consider a monoatomic gas contains atoms whose energy can range from 0 to ∞ . If there are no interactions, we are talking about kinetic energies only.
If the total energy of the system is 10 J, then no atom can have greater kinetic energy than this.
An atom having 15 J energy and others having different energies is a possible microstate, but not an accessible one.

The approach to equilibrium: Consider a system that is not equally likely to be found in any of the accessible microstates, meaning it is not in equilibrium.
Since the system is not in equilibrium, if the system is at any of the accessible microstate at any one time, it will not remain there indefinitely. Instead, at some other time the system is likely to be found at any of the other ~~not~~ accessible microstates.
The system will in the course of time always make transitions between its various microstates as a result of interactions between the constituent particles.

Now consider a statistical ensemble of such a system.

Suppose, these systems are initially distributed over their accessible states in a particular way, for example, they are only found in a particular subset of the accessible microstates. Now, each of the systems will continuously make transitions among various microstates. Given enough time, ultimately, the systems will be in different microstates with randomly distributed. So, again, the probability of finding the system in any of the microstates becomes equal and we get to the time-independent equilibrium situations. It's like shuffling a deck of cards. If you shuffle enough, each of the cards are equally likely to be at any position in the deck.

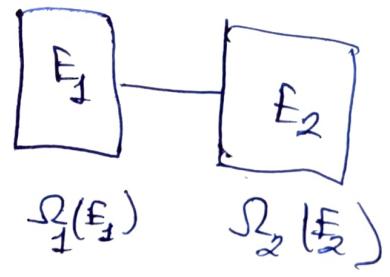
For example consider a box of gas, where all the molecules are at one side of the box, which is done by using a partition. If the partition is removed at t , what will happen at later times? The molecules are not distributed with equal probability over all their accessible microstates, since ^{all} the molecules are localized on the left side, although ~~not~~ the right side is empty, which is perfectly accessible now. It is hugely improbable that this situation will prevail. As a result of collisions with walls and between molecule

The molecules will very quickly redistribute themselves over the whole volume of the box, and the equilibrium will be attained.

Now, since all the microstates are equally likely, and the system explores all possible microstates, the system will most likely to be found in the macrostate that corresponds to the maximum number of microstates. For example, in coin toss, although the distinct microstates are equally probable, the outcome of 100 coin toss will "most likely" be 50 H and 50 T, since that contains maximum number of microstates. So, if you repeat the experiment for large number of time, in a large number of experiment it will be 50H and 50 T. For 1000 coin, the probability is staggeringly high. For a large system, the phrase "most likely" becomes ~~more~~ "absolutely likely". So,

"A system will appear to choose a macroscopic configuration that maximizes the number of microstates."

Consider two large systems that are in thermal contact with each other, but thermally isolated from their surroundings. The first system has energy E_1 and the second system has energy E_2 . The total energy is $E = E_1 + E_2$ which is fixed since the systems can't exchange energy with its surroundings. So, the value of E_1 is enough to describe the macrostate of the system. For a given macrostate, there can be and will be a huge number of accessible microstates. As stated before, we are not going to worry about what those microstates are. Let us assume, the first system can be in any one of the microstates $\Omega_1(E_1)$ and second system can be in any one of the microstates $\Omega_2(E_2)$. Thus the whole system can be in any of the microstates $\Omega_1(E_1)\Omega_2(E_2)$. This is a product since, for a given microstate $n \Omega_1(E_1)$, the second system can be in any one of the $\Omega_2(E_2)$ microstates. So, the total number must be the product of them.



$$\Omega_1(E_1) \quad \Omega_2(E_2)$$

Now, the most probable division

Now, since the systems are in thermal contact and can exchange energy between them, given enough time, they will come into thermal equilibrium, meaning E_1 and E_2 will come to fixed values.

Now, the most probable configuration of E_1 and E_2 will be such that the number $\Omega_1(E_1)\Omega_2(E_2)$ is maximized. We can maximize the ~~system~~ expression with respect to E_1 ,

$$\frac{d}{dE_1} (\Omega_1(E_1)\Omega_2(E_2)) = 0$$

$$\Rightarrow \Omega_2(E_2) \frac{d\Omega_1(E_1)}{dE_1} + \Omega_1(E_1) \frac{d\Omega_2(E_2)}{dE_1} = 0$$

$$\Rightarrow \Omega_2(E_2) \frac{d\Omega_1(E_1)}{dE_1} + \Omega_1(E_1) \frac{d\Omega_2(E_2)}{dE_2} \cdot \frac{dE_2}{dE_1} = 0$$

Since, $E = E_1 + E_2 = \text{constant}$, so,

$$\frac{dE}{dE_1} = \frac{dE_1}{dE_1} + \frac{dE_2}{dE_1} \Rightarrow \frac{dE_2}{dE_1} = -1$$

$$\therefore \frac{1}{\Omega_1} \frac{d\Omega_1}{dE_1} = \frac{1}{\Omega_2} \frac{d\Omega_2}{dE_2}$$

$$\boxed{\therefore \frac{d \ln \Omega_1}{dE_1} = \frac{d \ln \Omega_2}{dE_2}}$$

This is the condition that maximizes the number of microstates. Such division of energy E_1 and E_2 between two system is called being at the same temperature. So, we identify the quantity $\frac{d \ln \Omega}{dE}$ as temperature T so that $T_1 = T_2$. We define the temperature as -

$$\frac{1}{k_B T} = \frac{d \ln \Omega}{dE}$$

where $k_B = 1.38 \times 10^{-23} \text{ J K}^{-1}$ is called the Boltzmann constant.