

One Week Faculty Development Programme on
Engineering Physics
(19 - 24, August 2013)

Chaitanya Bharathi Institute of Technology,
Gandipet, Hyderabad 75

Statistical Mechanics

K. P. N. Murthy

School of Physics,
University of Hyderabad

23 August 2013



Acknowledgements

Thanks to

Dr. N. Amrutha Reddy,

Professor and Head,

and

Dr. B. Linga Reddy,

Associate Professor,

Department of Physics,
Chaitanya Bharathi Institute of Technology,
Gandipet, Hyderabad

for the invitation



Plan of the TALK

I am going to talk about

- the aim of statistical mechanics
- the methods of statistical mechanics based on Gibbs' ensembles and
- if time permits some recent developments.

Whenever and wherever required, I shall introduce the relevant concepts of thermodynamics

Please interrupt me if and when you do not follow what I am saying



Aim of Statistical Mechanics

- to synthesize **macroscopic** behaviour from **microscopic** behaviour.

micro-MACRO synthesis

There are **two issues**:

- The first is simple :
 - Statistical mechanics helps us **calculate macroscopic properties** from the properties of the microscopic constituents and their interactions
 - e.g. predict the properties of water like its density, pressure, heat capacity, latent heat of fusion, latent heat of evaporation, at what temperature does it freeze, at what temperature does it boil, *etc.* - from the properties of water molecules and the knowledge about how the water molecules interact with each other.



time asymmetric macro from time symmetric micro

- The second is rather ambitious
- the second agenda is to derive the **time asymmetry** in the macroscopic behaviour - arrow of time given by the direction of increasing entropy - from the **time symmetric** or **time-reversal invariant** microscopic equations - of Newton, of Schrödinger, of Heisenberg, of Maxwell, of Einstein - special relativity *etc.*
- statistical mechanics has been **eminently successful** in the first agenda.
- In fact this is precisely why, we are studying statistical mechanics
- statistical mechanics has not yet succeeded in the second agenda;
- this is the reason why statistical mechanics remains still interesting and challenging.
- if time permits I shall talk about the issues concerning the second agenda toward the end.



What is a micro world ? and what is a MACRO world ?

micro-MACRO synthesis

MICROSCOPIC CONSTITUENTS

MACROSCOPIC OBJECT

neutrons and protons

nuclei

nuclei and electrons

atoms and molecules

atoms and molecule

solids, liquids and gases

atomic magnets

ferromagnets

monomers

polymer chain

amino acids

protein molecule

sun, planets and satellites

solar system

solar systems

galaxy

galaxies

universe

men, women children and monkeys

society



Micro world obeys

- Classical mechanics : your teacher would have talked about
 - Newton's equations of motion
 - Euler-Lagrange equations
 - Hamilton - Jacobi equations
 - **jargon**
 - position; momentum; constraints; Lagrangian; degrees of freedom; potential; Hamiltonian; force; dynamical trajectory; phase space; Poisson bracket; *etc.*



- In Quantum Mechanics, you would have learnt of
 - Schrödinger equation
 - Heisenberg equation

jargon

wave function; quantum numbers; Hermitian operators; eigenvalues; eigenvectors; uncertainty principle; commutators; *etc.*

- Electrodynamics • Relativity • *etc.*

What are the principal and general characteristics of a microscopic law ?



micro world : determinism and time-reversal invariance

- **DETERMINISM**
- and
- **TIME REVERSAL INVARIANCE.**

Determinism

Entire past and the entire future is frozen in the present

- know of $\vec{x}(t = t_0)$ and $\vec{p}(t = t_0)$; then the classical mechanics machinery shall tell you $\vec{x}(t)$ and $\vec{p}(t)$ for all $t < t_0$ and for all $t > t_0$.



Determinism and time-reversal-invariance - an example

- Solar system is a good example:
 - specify the positions and momenta of all the planets at any instant of time, say now.
 - Newton's equations will tell us where all the planets shall be, an year from now; or where they were some four thousand years ago.

Time Reversal Invariance

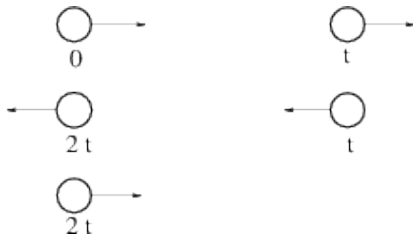
Microscopic laws do not distinguish the future from the past; the microscopic equations are invariant under the transformation of

$$t \rightarrow -t.$$



How do we know of time-reversal-invariance ?

- We can not reverse the time in the laboratory! Then how do we implement time reversal?
 - start at time say zero; move to time say t ;
 - reverse the momentum; move from time t to $2t$;
 - reverse the momentum.



- if you arrive at the same phase-space point you started with, then the dynamics is time-reversal invariant.



time reversal invariance

- Run the solar system forward for an hour; then run it backward for an hour; the planets will all end up exactly where they started.
- Time reversal invariance has a **more general** and **subtle** meaning: Technically,

Newton's equations are good in
either directions of time

- take a movie of classical mechanical trajectory;
- screen the movie in the reverse
 - You will not find anything funny.
 - what you see will be a perfectly valid trajectory



charge in an electric field

- Motion of a charge q in an electric field

$$\vec{E} = -\nabla\phi$$
$$m\frac{d^2\vec{r}}{dt^2} = q\vec{E}(\vec{r})$$

- if $\vec{r}(t)$ is a solution of the above equations, then so is $\vec{r}(-t)$
- the equations are second order in time; the two changes of sign coming from $t \rightarrow -t$ cancel.
- In other words, take a movie of the dynamics of the charge obeying the above equations of motion and screen it in the reverse; what you see is also a physically allowed dynamics.



moving charge in a magnetic field

- What happens if there are magnetic forces?

$$m \frac{d^2 \vec{r}}{dt^2} = q \vec{E}(\vec{r}) + \frac{q}{c} \frac{d\vec{r}}{dt} \times \vec{B}(\vec{r})$$

- the equations of motion include a first order time derivative which changes sign under time reversal.
- e.g. in a constant magnetic field, the sense of circular motion (clock wise or anti-clock wise) is determined by the nature of the charge and not by the initial conditions; hence the time-reversed motion will be in the opposite sense.
- However, if we make the replacement of $\vec{B} \rightarrow -\vec{B}$ we restore time reversal invariance.



Macro World

- the world we see around; the world of trees; of rocks; of houses; of tables; of chalk pieces; of walls; of seas; of beer; *etc.*
- these objects have microscopic constituents;
- these microscopic constituents are usually completely hidden from our view.
- indeed, when we study an object at a particular length scale, we can ignore its structure on smaller length scales by averaging over these length scales. *e.g.* when we study solids we ignore the nuclei of the atoms that make up the solid.



Laws of the macro world : THERMODYNAMICS

First law : $dU = \delta q + \delta W$

Second law : $dS \geq 0$

- **jargon:**

- temperature; pressure; volume; energy; entropy; chemical potential; intensive variables; extensive variables; equation of state; phase diagrams; heat capacity; *etc.*

What is the principal character of a macroscopic behaviour ?

Macroscopic behaviour is *not time reversal invariant*.

There is a definite direction of time ...

the direction of increasing entropy.



Can you **undo** what you have **done** ?

- **Question:** When do we say, an act can be reversed ?
- **Answer:** When we can undo all the consequences of the act and restore everything as they were before the act.
 - you buy a shirt from a shop;
 - back at home, you discover your wife does not like it;
 - the next day you go to the shop and return the shirt;
 - the shop keeper takes the shirt back and returns you the money.
 - You have undone whatever you did.
 - We say the act of buying a shirt from the shop is reversible
- so what ?



How do we detect irreversibility ?

- when something is reversible, then, the reversed behaviour will also happen as often as the forward behaviour;
- you will find in the shop, people buying shirts - the forward act
- you will also find in the shop, people returning the shirts and getting back their money - the reverse act
- When we observe often, a phenomenon as well as its reversal, then we say the phenomenon is REVERSIBLE.
- breaking of a glass mug is something we often see;
- its reversal - broken glass pieces of the mug assembling to form a mug - we do not see at all - at least I have not seen
- hence breaking of a glass mug is an irreversible phenomenon
- most of the macroscopic phenomena are irreversible



Omar Khayyam (1048 - 1131)

*The Moving Finger writes;
and, having writ, moves on
nor all your Piety nor Wit
Shall lure it back
to cancel half a Line,
Nor all your Tears
wash out a Word of it.*

Omar Khayyam (1048-1131)
Scientist-Poet, Persia (now Iran)

- Time reversed **macroscopic** behaviour is **never** seen.
- Time reversed microscopic behaviour is **often** seen.



micro-MACRO dichotomy

The two pillars of theoretical physics,

Newtonian mechanics

and

thermodynamics

seem to stand in contradiction.

- WHY ?
- How do we comprehend this micro-MACRO dichotomy?
- Answer: Through the machinery of **Statistical Mechanics**
- Hence without much ado, let us get on with the task of learning the methods of statistical mechanics.



Why are we uncomfortable with Statistical Mechanics

- Normally, we resort to statistics only
 - when we want to hide our ignorance
 - when we want to hide our inability make predictions
- for example when a coin is tossed we can not tell whether it will show up *Heads* or *Tails*
- May be, e can argue that in principle, Newton's equations should help us make a correct prediction
- In principle, we can write down Newton's equations accounting for all the forces acting on the coin, solve the equation, may be numerically, and arrive at the outcome of the toss
- such an enterprise feasible : is it feasible ?



unpredictability - is it an inherent property ?

- may be ... certain phenomena are **inherently** unpredictable
- perhaps, the outcome of the toss of a coin is inherently unpredictable - that is why we say
 " as random as the toss of a coin"
- the outcome perhaps depends sensitively on several factors like
 - the torque and force applied,
 - the air friction,
 - the wind flow,
 - the mass distribution in the coin
 - *etc*
- we may not know these factors adequately precisely to make a reasonable prediction
- may the dynamics enhances the errors in the initial conditions exponentially, thus destroying completely our ability to make predictions, despite the evolution being deterministic



chaos implies unpredictability

With the advent of nonlinear dynamics and chaos, we now know, that
determinism
does not necessarily imply
predictability

- (non-linear) chaotic dynamical systems are deterministic;
- but they evolve **unpredictably**
- In the light of the above, we take a simple attitude that the result of a toss, can be *Heads* or *Tails* with 50-50 probability



microstate

- Consider an experiment - real or imagined which has more than one outcome
- Example - 1
 - toss a coin
 - it has two out comes : *Heads* - H and *Tails* - T
 - these are the microstates of a single coin
 - the set $\{H, T\}$ is called microstate space; mathematicians call it sample space
- Example - 2
 - toss a dice
 - it has six outcomes : 1 - 6
 - these six outcomes constitute a microstate space or sample space



microstate : a few more examples

• Example - 3

- toss a coin twice (independently) or equivalently toss two (identical) coins (independently)
- microstate space or sample consists of four outcomes :
 $\{(HH), (HT), (TH), (TT)\}$

• Example - 4

- toss N (independent and identical) coins
- How many outcomes are possible ? Answer : 2^N
- each outcome is a string of symbols H and T
- e.g. a string like $(HHTHHHTTTHHHTT)$ is an outcome when $N = 13$.
- There are $2^{13} = 8192$ microstates for the experiment of tossing 13 coins (identical and independent)



microstate : yet another example

• Example - 5

- Consider a gas in a container
- the gas consists of molecules of the order of say 10^{25} , in number
- Each molecule can be specified by giving six numbers : three for its position and another three for its momentum.
- there are N molecules. Hence we require a string of $6N$ numbers to specify a microstate of the system at any given time.
- all possible strings of $6N$ numbers constitute the microstate space for the system
- we can view it as an experiment of organizing N molecules in a container with some constraints (on energy, for an **isolated system**, on temperature for a **closed system** or temperature and chemical potential for an **open system**)
- there are several possible outcomes, of the experiment, which constitute the microstate space



MACRO state

- Consider a subset of the micro state space
- all the microstates belonging to this subset have the same macroscopic properties
 - Consider tossing of three coins.
 - there are eight microstates
 - consider microstates with two *Heads* (and hence one *Tails*)
 - there are three microstates having this macroscopic property
 - why do I call this a macroscopic property ?
 - It is because this property does not depend on the microscopic details - like whether the first coin is *Heads*, the second is *Heads* and the third is *Tails*.
 - This property depends on the macrostate in which there are two *Heads* (it doesn't matter which two of the three coins show up *Heads*) and which one shows up *Tails*



Tossing of N coins

- Let n denote the number of *Heads* in a toss of N coins.
- n can take values from 0 to N .
- Thus there are 2^N microstates and $N + 1$ macrostates.
- Let $\Omega(n; N)$ denote all the microstates having n *Heads* (and hence $(N - n)$ *Tails*)
- Let $\widehat{\Omega}(n; N)$ denote the number of microstates associated with the macrostate n , where $0 \leq n \leq N$.
- In other words $\widehat{\Omega}(n; N)$ is the cardinality of the set $\Omega(n; N)$.
- It is easy to derive an expression for $\widehat{\Omega}(n; N)$, see next slide.



Derivation of an expression for $\widehat{\Omega}(n; N)$

- Take a single microstate $\omega \in \Omega(n; N)$
- From this single microstate we can produce $n!$ permutations by permuting n Heads;
- from each of the $n!$ permutations, we can produce $(N - n)!$ permutations, by permuting $(N - n)$ Tails.
- We can carry out the above exercise, for each microstate belonging to the set $\Omega(n; N)$
- Thus we can produce $\widehat{\Omega}(n; N)n!(N - n)!$ permutations.



$\hat{\Omega}(n; N)$

- it is easy to see

$$\hat{\Omega}(n; N) n! (N - n)! = N!$$

- thus we get,

$$\hat{\Omega}(n; N) = \frac{N!}{n! (N - n)!}$$

- First we verify,

$$\sum_{n=0}^N \hat{\Omega}(n; N) = \sum_{n=0}^N \frac{N!}{n! (N - n)!} = 2^N$$



Probability of a macrostate

- Let us say all the microstates are equally probable
 - in a coin tossing experiment, this statement is obvious if all the coins are fair coins
 - for a single toss, *Heads* and *Tails* are equally probable
 - hence, for N tosses, all the 2^N microstates are equally probable
- Physicists have a fancy name for it : **ERGODICITY**
- Let $P(n; N)$ denote the probability that the macrostate is n
- then we have,

$$P(n; N) = \frac{\hat{\Omega}(n; N)}{\sum_{n=0}^N \hat{\Omega}(n; N)} = \frac{1}{2^N} \frac{N!}{n!(N-n)!}$$



What is the macrostate we observe ?

- The important question is when we carry out measurements on the system what is the value of the macroscopic property we will observe
- Imagine N coins jumping up and falling flat *i.e.* get tossed, once in every 10^{-15} seconds.
- We are going to observe the system for say one microsecond : 10^{-6} seconds.
- During the observation time the system would have visited some 10^9 microstates.
- what we observe as the macroscopic property n is a quantity averaged over 10^9 microstates.
- we call this time average
- Let us now talk about ensemble average....
- What is an ensemble ?



Ensemble : Maxwell's mischief

- How does a mathematician calculate the average of a random variable n which can take values n_1, n_2, \dots with probabilities p_1, p_2, \dots ?

$$\langle n \rangle = \sum_i n_i p_i$$

- How does a physicist calculate average of a quantity ?
- for example, I want to know the average height of a person in this room;
- how do I calculate it ?
- Find the height of every person;
- let us say there are N persons in this room;
- let h_i denote the height of the i -th person

$$\langle h \rangle = \frac{1}{N} \sum_{i=1}^N h_i$$



Maxwell's ensemble

- take the sum of the heights of all the persons in this room;
- divide the sum by the number of persons in this room;
- what you get is the average height of a person in this room; simple
- Maxwell thought that this is what a physicist would understand as average and this is how he would calculate it
- He had a poor opinion about physicists' ability to comprehend the average as defined by mathematicians, involving
 - random variable n ,
 - the possible values of n denoted by n_1, n_2, \dots ,
 - the probabilities p_1, p_2, \dots and
 - sum of the product $n_i p_i$ over all i



Construction of an ensemble, knowing the probabilities

- Consider tossing of a coin
- the microstate space is $\{H, T\}$;
- Let $P(H) = 3/4$ and $P(T) = 1/4$: the coin is biased
- Let $X(\omega)$ be random variable, that attaches $+1$ to H and -1 to T
- $\langle x \rangle = (+1) \times 0.75 + (-1) \times 0.25 = +0.5$
- A possible ensemble is $\{H, H, H, T\}$
- Ensemble average is $\langle x \rangle = (+1 + 1 + 1 - 1)/4 = 0.5$
- Thus an ensemble is a collection of microstates of the system (*i.e.* outcomes of the experiment)
- each microstate is repeated as often in the ensemble as to reflect its probability.



size of an ensemble

- In the above example the size of the ensemble is 4. H occurs three times because the probability of H is $3/4$; $Tails$ occurs only once since $P(T) = 1/4$
- Of course $\{H, T, H, H, H, T, H, H\}$ is a possible ensemble because $P(H) = 6/8 = 3/4$ and $P(T) = 2/8 = 1/4$
- The size of the microstate space is fixed by the nature of the system or the nature of the experiment
- the size of the ensemble is arbitrary; usually we take the size of the ensemble as large as possible to ensure that each microstate is represented in the ensemble in proper proportion to reflect its probability.
- if we have a fair coin, then $P(H) = P(T) = 1/2$
- then the sample space $\{H, T\}$ itself is a candidate for an ensemble.



Calculating probabilities from the ensemble

- we are given the ensemble; for example let us say all the N persons present in this room constitute an ensemble
- we calculate probability distribution of the heights as follows
- Let us say the height can vary from say 4.5 feet to 6.5 feet.
- we divide this range into a set of intervals each of width say Δh
- for example we can take $\Delta h = 0.1$ which gives twenty equal intervals
- we find how many persons have heights in the i -th interval: n_i
- then $p_i = n_i/N$, where $N = n_1 + n_2 + \dots$
- we will construct an ensemble of closed systems on physical grounds and from this ensemble we shall calculate the probability for a closed system to be in any of its microstate.



How does a mathematician calculate average ?

- Let us return to the tossing of N fair coins and the determination of $\langle n \rangle$.
- Following mathematicians, we can calculate the average by explicitly summing the quantity $n \times P(n; N)$, over all possible values the random variable n can take; Note $n = 0, 1, 2, \dots, n$.

$$\begin{aligned}\langle n \rangle &= \sum_{n=0}^N n P(n, N) \\ &= \frac{\sum_{n=0}^N n \times \widehat{\Omega}(n; N)}{\sum_{n=0}^N \widehat{\Omega}(n; N)} \\ &= \frac{1}{2^N} \sum_{n=0}^N n \frac{N!}{n! (N-n)!} \\ &= \frac{N}{2}\end{aligned}$$



Ensemble average

- We can get the same answer by constructing an ensemble of microstates
- Let the size of the ensemble be denoted by M ;
- The size of the ensemble is such that a microstate is repeated as often in the ensemble as to reflect its probability
 - for example M can be taken as the inverse of the probability of the rarest microstate;
 - in the experiment of tossing N fair coins the ensemble size can be taken as 2^N , since all microstates are equally probable
 - instead if $P(H) = 0.1$, then the probability of all the N coins showing up H is the smallest and equals $(0.1)^N$;
 - Inverse of this can be taken as the size of the ensemble
 - thus $M = 10^N$



... ensemble average

- Let n_i be the value of the macroscopic property of the i th microstate of the ensemble
- Ensemble average of n is given by

$$\langle n \rangle = \frac{1}{M} \sum_{i=1}^M n_i$$

- In Markov Chain Monte Carlo methods we construct a sample microstates by sampling from the known equilibrium probability distribution of the system; the Monte Carlo sample is a subset of the ensemble; we calculate averages over the Monte Carlo sample



Maximum entropy state

- Let the N coins be fair coins
- we can attach an entropy to each macrostate given by the logarithm of the number of microstates associated with it
- all microstates are equally probable

$$\begin{aligned} S(n) &= k_B \ln \hat{\Omega}(n; N) \\ &= k_N \ln \left(\frac{N!}{n!(N-n)!} \right) \end{aligned}$$

- We can say that the value of n for which $S(n)$ is maximum shall be the value observed in experiment;
- let us denote this value as \hat{n}
- It is readily seen $\hat{n} = \langle n \rangle$



most probable value

- Since for this problem, we know exactly the probability distribution of the random variable n , we can ask the following question
- What is the probability \mathcal{P} , that n takes a value out side a small interval around $\langle n \rangle = \hat{n} = N/2$?
- Let us take the interval as

$$\frac{N}{2} - \epsilon \frac{N}{2}, \quad \frac{N}{2} + \epsilon \frac{N}{2}$$

where ϵ is a small fraction.

- take $\epsilon = 0.001$



fluctuations about the most probable value

| N | \mathcal{P} |
|--------|-------------------------|
| 10^4 | 0.317 |
| 10^5 | 0.002 |
| 10^6 | 1.5×10^{-23} |
| 10^8 | 2.7×10^{-2174} |

- When the number of coins is a hundred million (10^8) the probability that the random variable n - the number of *Heads*, will take a value within $\pm 1\%$ around $N/2$ is very close to unity; it is $1 - 2.7 \times 10^{-2174}$
- notice 10^8 is a small number for Avagadro
- Avagadro handles numbers of the order of 10^{23} and above
- Imagine the situation when we have 10^{23} to 10^{25} coins.
- in macroscopic systems we deal with this kind of number of atoms or molecules



microcanonical ensemble

- for a single coin, there are only two microstates
- for a single atom, the number of microstates is very large, as we shall see soon
- it is precisely because of this a macroscopic system has a definite value for its macroscopic property, despite statistics or randomness employed for its description
- This is precisely the reason that statistical mechanics is able to predict macroscopic (thermodynamic) properties.
- a typical macroscopic object: a glass of water: it contains nearly 10^{25} water molecules.
- each molecule for complete specification requires three position and three momentum coordinates; in the spirit of coarse graining we ignore the structure of a molecules and the structure of its constituents.



Ergodicity

- we need therefore $\sim 6 \times 10^{25}$ numbers to specify the macroscopic object.
- the system at a microscopic level is indeed hyperactive; the molecules are incessantly in motion;
- the system is switching from one microstate to another, all the time;
- each molecule moves as per the diktats of the Newtonian mechanics.
- the entire system of N molecules is specified by a point in a $6N$ dimensional phase space.
- the point traces out a trajectory in the phase space as per Newton's laws.
- consider an isolated system in equilibrium
- let us say that we are assured that the (isolated and equilibrium) system visits all the regions of the phase space accessible to it ' uniformly ' ;



micro canonical ensemble averaging

- in other words the system spends equal duration of time in equal regions of the phase space.
- experimentally measured macroscopic property can be considered as a time average over the phase space trajectory traced by the system during a measurement.
- then, we can completely ignore the dynamics and calculate a macroscopic property as an average over a static micro canonical ensemble (micro states) - with *all the microstates being equally probable*.
- this is called averaging over a microcanonical ensemble.



Equilibrium system is stable

- notice that despite tremendous activity at a microscopic level, the macroscopic properties of the equilibrium system remain robust and unchanging.
 - water boils at 100 degree Celsius at standard pressure.
 - ice melts at zero degree Celsius.
 - pressure times volume remains the same for a given quantity of an ideal gas at constant temperature.
 - mercury expands upon heating.
 - when two bodies come into thermal contact, the hotter body cools and the colder warms up until both attain the same temperature.



Why is an equilibrium system, stable ?

- a few macroscopic descriptors like volume V , pressure P , entropy S , temperature T , density ρ etc. are adequate to describe the macroscopic state of the system.
- what is the reason for the observed robustness of a macroscopic behaviour ?
 - let us say you are determining a macroscopic property O , of a system by observing it over a duration of time say τ ;
 - what you measure can be thought of as an average of the macroscopic property over the N microstates visited by the system during the experimental observation time τ . Let us denote it by \overline{O}_N ; it is quite clear \overline{O}_N is a random variable.



Fluctuations are inversely proportional to system size

- if N is very large and if (we assume that) the microstates visited are all independent of each other, then \overline{O}_N has a Gaussian distribution; this is a consequence of the Central Limit Theorem.
- more importantly, the variance of the Gaussian distribution is of the order of the inverse of N and hence is small for large N ;
- since N is invariably large, we find a macroscopic property robust and unchanging with time.
- e.g. pressure is average momentum transferred to the wall (of the container) by the colliding molecules of the fluid. This is the simplest of micro-macro connections.
- in fact statistical mechanics establishes several such micro-macro connections.



The Central Limit Theorem

- $\{X_i\}$ are identically distributed independent random variables with mean μ and finite variance σ^2 .
Let $\rho(x)$ denote its probability density function.
- The characteristic function of X is

$$\begin{aligned}\Phi_X(k) &= \int_{-\infty}^{\infty} dx \exp[+ikx] \rho(x) \\ &= \exp \left[ik\mu - \frac{k^2}{2} \sigma^2 + \sum_{n=3}^{\infty} \frac{(ik)^n}{n!} \zeta_n \right]\end{aligned}$$

- In the above ζ_n is the n -th cumulant of X ;

$$\zeta_1 = \mu ; \quad \zeta_2 = \sigma^2.$$



Central Limit Theorem continued

$$Y = \frac{1}{N}(X_1 + X_2 + \cdots + X_N)$$

- The characteristic function of the random variable Y is:

$$\begin{aligned}\Phi_Y(k) &= \left[\Phi_X \left(k \rightarrow \frac{k}{N} \right) \right]^N \\ &= \exp \left[ik\mu - \frac{k^2 \sigma^2}{2 N} + \sum_{n=3}^{\infty} \frac{(ik)^n}{n!} \frac{\zeta_n}{N^{n-1}} \right] \\ &\underset{N \rightarrow \infty}{\sim} \exp \left[ik\mu - \frac{k^2 \sigma^2}{2 N} \right]\end{aligned}$$

- The above is a Gaussian whose Fourier inverse is also a Gaussian with mean μ and variance σ^2/N .



Central Limit Theorem ... Continued

- For $N \rightarrow \infty$, we have strictly

$$\Phi_Y(k) = \exp[i\mu k]$$

whose Fourier inverse is

$$\delta(y - \mu);$$

- Y is not any more random; it is deterministic;
- The macroscopic variable is robust indeed.
- For N adequately large, the fluctuations of Y is inversely proportional to \sqrt{N} and hence is small.



Ludwig Eduard Boltzmann

(20 Feb. 1844 - 5 Oct. 1906)

$$S = k_B \log(\hat{\Omega})$$

S is entropy

k_B is called the Boltzmann constant:

1.381×10^{-23} Joules per degree Kelvin or
 8.671×10^{-5} electron-volt per degree Kelvin

$\hat{\Omega}$ is the number of microstates
accessible to the system

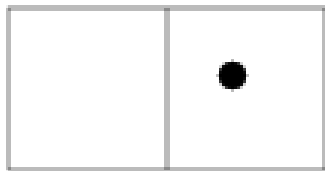
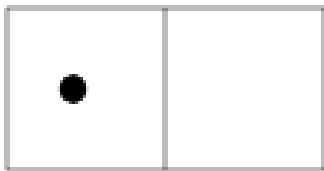


Toy Model : One particle in a box with two portions

- Consider the following experiments:
 - take a particle and place it in a box; left or right portion.

Question : How many ways can you do this?

Answer : 2; $\hat{\Omega} = 2$; $S = k_B \log 2$

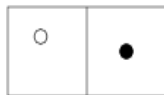
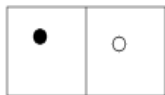
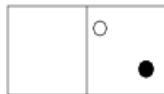
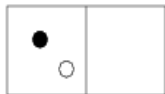


Toy model: two particles in a box of two portions

- take two particles (non interacting) and place them in the box independently.

Question : How many ways can you do this?

Answer : 4; $\hat{\Omega} = 4$; $S = 2 k_B \log 2$



Toy model : N particles in a box with two portions

- take N particles and place them in the box independently;

Question : How many ways can you do this?

Answer : 2^N ; $\hat{\Omega} = 2^N$; $S = N k_B \log 2$

- In the above experiments the box was divided into two equal parts. Consider an experiment of placing N particles independently in a box divided into 3 equal parts.

Question : How many ways can you do this?

Answer : 3^N ; $\hat{\Omega} = 3^N$; $S = N k_B \log 3$



toy model: N particles in a box of V/ϵ portions

- Consider an experiment of placing N molecules independently in a box divided into V/ϵ equal parts, where ϵ is the volume of the tiny cube

Question : How many ways can you do this?

$$\text{Answer : } \hat{\Omega} = \left(\frac{V}{\epsilon}\right)^N$$

$$S = k_B \log \hat{\Omega} = Nk_B \log V - Nk_B \ln \epsilon$$

$$\frac{\partial S}{\partial V} = \frac{Nk_B}{V} = \frac{P}{T} \quad (\text{see box below})$$

$$P V = N k_B T$$

- We have derived the ideal gas law!



S and dS

$$\begin{aligned} S &\equiv S(U, V, N) \\ dS &= \left(\frac{\partial S}{\partial U}\right)_{V, N} dU + \left(\frac{\partial S}{\partial V}\right)_{U, N} dV + \left(\frac{\partial S}{\partial N}\right)_{U, V} dN \\ &= \left(\frac{1}{T}\right) dU + \left(\frac{P}{T}\right) dV + \left(-\frac{\mu}{T}\right) dN \end{aligned}$$

- What more? Consider a thermodynamic process in which only the volume changes and all other thermodynamic variables remain the same. We have,



dS for the toy model

$$S(V) = k_B N \ln V - N k_B \ln \epsilon$$
$$dS = \frac{\partial S}{\partial V} dV = \frac{k_B N}{V} dV$$

- Let us say that the infinitesimal change in volume has occurred because of a quasi-static reversible isothermal process during which the system draws δq_R of energy in the form of (reversible) heat from the heat reservoir.
- We have $PdV = \delta q_R$, which implies that,

$$dS = \frac{Nk_B}{PV} \delta q_R$$

- Replace PV by $Nk_B T$ (ideal gas law which we have already derived) in the above expression.



Boltzmann entropy and Clausius' entropy are equivalent

- We get,

$$dS = \frac{d^*q_R}{T}$$

which is precisely the definition for change in entropy given by Clausius.

- Thus Clausius' entropy readily follows from Boltzmann's entropy.
- Boltzmann gives an expression for the absolute entropy. Clausius gives an expression for change in entropy.
- Both Boltzmann entropy and Clausius' entropy are defined for an equilibrium system.



Gibbs' paradox and Boltzmann counting

- We derived all the above expressions starting from the following,

$$\hat{\Omega} = V^N; S = Nk_B \log V$$

- Is this entropy extensive?
- No; it is not: *Check this out*
- called **Gibbs' paradox**.
- What should we do to make it extensive ?
- Divide $\hat{\Omega}$ by $N!$; the resulting entropy is extensive.
- This was suggested by Boltzmann for taking care of *indistinguishability*.



Micro Canonical ensemble

- The correct full expression for $\hat{\Omega}$ as a function of U , V and N is given by the Sackur-Tetrode equation,

$$\hat{\Omega}(U, V, N) = V^N \frac{1}{h^{3N}} \frac{1}{N!} \left(\sqrt{2\pi m} \right)^{3N} \frac{U^{3N/2}}{\Gamma\left(\frac{3N}{2} + 1\right)}$$

- h^{3N} is the (phase space volume) unit we use to measure the volume of the $6N$ dimensional phase space;
- this choice of unit volume is inspired by the Heisenberg uncertainty principle: $\Delta x \Delta p \geq h$
- We take the ' minimum uncertainty ' volume of the phase space, $\prod_{i=1}^N \prod_{k=1}^3 \Delta x_i^{(k)} \Delta p_i^{(k)} = h^{3N}$
- $\Gamma(y)$ is the usual gamma function defined by, $\Gamma(y) = \int_0^\infty \exp(-x) x^{y-1} dx$; $\Gamma(\nu + 1) = \nu !$



Boltzmann-Gibbs-Shannon Entropy

Josiah Wilard Gibbs (1839 - 1903)
Claude Elwood Shannon (1916 - 2001)

- Consider an isolated system; let Ω_{IS} denote the set of all possible microstates accessible to the system; let $\hat{\Omega}_{IS}$ denote the number of microstates.

$$\begin{aligned} S &= k_B \log \hat{\Omega}_{IS} \\ &= k_B \frac{1}{\hat{\Omega}_{IS}} \hat{\Omega}_{IS} \log \hat{\Omega}_{IS} \\ &= -k_B \sum_{i=1}^{\hat{\Omega}_{IS}} \frac{1}{\hat{\Omega}_{IS}} \log \left(\frac{1}{\hat{\Omega}_{IS}} \right) \end{aligned}$$



Boltzmann-Gibbs-Shanon Entropy

- This can be written in more suggestive form,

$$S = -k_B \sum_{i=1}^{\hat{\Omega}_{IS}} p_i \log p_i$$

where, $p_i = 1/\hat{\Omega}_{IS} \forall i$:

'all microstates are equally probable'

- We can generalize and say that the above gives an expression for entropy even when the probabilities are not the same for all microstates. thus we get the Boltzmann-Gibbs-Shanon entropy.



Ensemble \longrightarrow probabilities

- Consider a closed system;
- the system is in contact with a heat bath with which it exchanges only energy; it does not exchange matter;
- Let N be the number of molecules in the system, V the volume and T the temperature
- Let us index the microstates of the system as $i = 1, 2, \dots$ and the corresponding energies as $\epsilon_1, \epsilon_2, \dots$.
- Aim is to calculate P_i - the probability of microstate i .
- we shall calculate P_i by constructing an ensemble



Method of most probable distribution

Construct a canonical ensemble; calculate the probability of microstate i

- Assemble a very large number of mental copies of the system such that adjacent systems are in thermal contact with each other.
- Isolate the entire assembly of systems
- A macrostate of the isolated system is described by specifying a string of numbers $\{a_1, a_2, \dots\}$.
- a_i is the number of members of ensemble in microstate i
- we have the constraints : $\sum_j a_j = \mathcal{A}$; and $\sum_j a_j \epsilon_j = \mathcal{E}$
- \mathcal{A} is the number of elements in the ensemble;
- \mathcal{E} is the total energy of the isolated system; this energy is constant



Method of most probable distribution

- We have

$$\hat{\Omega}(a_1, a_2, \dots) = \frac{\mathcal{A}!}{a_1! a_2! \dots}$$
$$P_i = \frac{\sum_{\{a_1, a_2, \dots\}} (a_i / \mathcal{A}) \hat{\Omega}(a_1, a_2, \dots)}{\sum_{\{a_1, a_2, \dots\}} \hat{\Omega}(a_1, a_2, \dots)}$$
$$\sum_j a_j = \mathcal{A}; \quad \sum_j a_j \epsilon_j = \mathcal{E}$$

- The isolated system is going to be found in that macro state $\{a_1^*, a_2^*, \dots\}$ for which the entropy $k_B \ln \hat{\Omega}(a_1, a_2, \dots)$ is maximum.
- then

$$P_i = \frac{a_i^*}{\mathcal{A}}$$

- the problem reduces to finding $\{a_1^*, a_2^*, \dots\}$ that maximizes the entropy under the two constraints



Lagrange undetermined multiplier

- Introduce a Lagrange multiplier λ for the first constraint, a Lagrange multiplier β for the second constraint and maximize the entropy. We get

$$P_i = \frac{a_i^*}{\mathcal{A}} = \frac{1}{Q} \exp(-\beta \epsilon_i)$$

$$Q(T, V, N) = \sum_j \exp[-\beta \epsilon_j(V, N)]$$



canonical partition function to thermodynamics

- Thermodynamic internal energy U is given by the canonical ensemble average of statistical mechanical energy $\langle E \rangle$, see below.

$$U = \langle E \rangle = \frac{1}{Q} \sum_i E_i \exp(-\beta E_i) = -\frac{1}{Q} \frac{\partial Q}{\partial \beta}$$

- We can show that

$$\sigma^2 = \langle E^2 \rangle - \langle E \rangle^2 = k_B T^2 \left(\frac{\partial U}{\partial T} \right)_V = k_B T^2 C_V$$
$$F(T, V, N) = -k_B T \ln Q(T, V, N)$$

- Thus once we know the partition function for a macroscopic system, we can calculate all its macroscopic (thermodynamic) properties.



Open system and Grand canonical ensemble

- For an open system, which exchanges energy as well as matter with the bath, we have Ω_{OS} denoting the set of all microstates; the probabilities are given by,

$$p(\mathcal{C}) = Q^{-1} \exp \left[-\beta E(\mathcal{C}) + \beta \mu N(\mathcal{C}) \right]$$

- In the above $\mathcal{C} \in \Omega_{OS}$; μ is the chemical potential; Q is the grand-canonical partition function; $N(\mathcal{C})$ is the number of particles in the system when in microstate \mathcal{C} .

$$Q(T, V, \mu) = \sum_{\mathcal{C} \in \Omega_{OS}} \exp \left[-\beta E(\mathcal{C}) + \beta \mu N(\mathcal{C}) \right]$$



Microscopic Interpretation of Work and Heat

ϵ_j : energy of the microstate j

p_j : probability of the microstate j

$$\begin{aligned}U &= \sum_j p_j \epsilon_j \\dU &= \sum_j \left[\frac{\partial U}{\partial p_j} dp_j + \frac{\partial U}{\partial \epsilon_j} d\epsilon_j \right] \\&= \sum_j \left[\epsilon_j dp_j + p_j d\epsilon_j \right] \\&= \sum_j \epsilon_j dp_j + \sum_j p_j d\epsilon_j \\&= \quad \quad \quad d\mathcal{Q} \quad + \quad \quad \quad d\mathcal{W}\end{aligned}$$



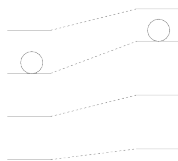
Microscopic interpretation of work and heat

- $\sum_j \epsilon_j dp_j$ refers to Heat;
 - This term refers to the energy exchanged by the system with the heat bath in the form of heat.
 - during this process, the energy of a microstate does not change.
 - only the occupation probabilities $\{p_i\}$ change.
- $\sum_j p_j d\epsilon_j$ refers to Work done on the system.
 - during this process the occupation probabilities - $\{p_j\}$ do not change.
 - only the energy of the microstates change.
 - e.g. when we change volume (boundary conditions) the energy eigenvalues of the system change.



Work Term

$$\left(\sum_j p_j d\epsilon_j \right)$$



- the system remains in the same microstate; only the energy of the microstate changes taking the system along with it.



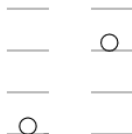
Work in statistical mechanics

$$\begin{aligned}\sum_j p_j d\epsilon_j &= \sum_j p_j \left(\frac{\partial \epsilon_j}{\partial V} \right)_{\{p_j\}} dV \\ &= \left(\frac{\partial}{\partial V} \sum_j p_j \epsilon_j \right)_{\{p_j\}} dV \\ &= \left(\frac{\partial U}{\partial V} \right)_{S,N} dV \\ &= -P dV\end{aligned}$$



Heat in statistical mechanics

Heat Term ($\sum_j \epsilon_j dp_j$)



- The energies of the microstates do not change; the energy put in (or extracted out) **in the form of heat** induces the system to make a transition from one microstate to another of higher (or lower) energy; such a transition could occur even otherwise by a spontaneous fluctuation.



Heat in statistical mechanics - closed system

$$S = -k_B \sum_j p_j \log p_j; \quad dS = -k_B \sum_j [dp_j + dp_j \log p_j]$$

$$\left[\sum_j dp_j = 0 \right]; \quad dS = -k_B \sum_j dp_j \log p_j$$

$$\left[p_j = \exp[-\beta \epsilon_j] / Z \right]; \quad \log p_j = -\beta \epsilon_j - \log Z$$

$$dS = k_B \sum_j dp_j [\beta \epsilon_j + \log Z]$$

$$\left[\sum_j dp_j \log Z = 0 \right] \quad dS = k_B \beta \sum_j dp_j \epsilon_j;$$

$$TdS = \sum_j dp_j \epsilon_j = d'q_R$$



entropy in statistical mechanics

- Isolated system : $S = k_B \ln \hat{\Omega}$
- systems that are not isolated : e.g. closed systems, open systems
- consider a general system with micro states denoted by $\{1, 2, \dots\}$ and corresponding probabilities $\{p_1, p_2, \dots\}$
- construct an ensemble of (macroscopically) identical copies of the system
- a_i denotes the number of members of the ensemble in micro state i
- $\mathcal{A} = \sum_i a_i$ denotes the total number of members of the ensemble
- each a_i is arbitrarily large; hence \mathcal{A} is also arbitrarily large



entropy for a system with $\{p_i : i = 1, 2, \dots\}$

- we have $p_i = a_i/\mathcal{A} \forall i$

$$\widehat{\Omega}(a_1, a_2, \dots) = \frac{\mathcal{A}!}{\prod_i a_i!}$$

- a natural definition of entropy is then

$$\begin{aligned} S(a_1, a_2, \dots) &= k_B \ln \frac{\mathcal{A}!}{\prod_i a_i!} \\ &= k_B \left(\mathcal{A} \ln \mathcal{A} - \mathcal{A} - \sum_i a_i \ln a_i + \sum_i a_i \right) \\ &= k_B \left(\mathcal{A} \ln \mathcal{A} - \sum_i a_i \ln a_i \right) \end{aligned}$$



entropy of a system with $\{p_i : i = 1, 2, \dots\}$

$$\begin{aligned} S(a_1, a_2, \dots) &= -k_B \sum_i [a_i \ln a_i - a_i \ln \mathcal{A}] \\ &= -k_B \sum_i a_i \ln \left(\frac{a_i}{\mathcal{A}} \right) \\ &= -k_B \mathcal{A} \sum_i \left(\frac{a_i}{\mathcal{A}} \right) \ln \left(\frac{a_i}{\mathcal{A}} \right) \\ &= -\mathcal{A} k_B \sum_i p_i \ln p_i \end{aligned}$$

- the above is the entropy of \mathcal{A} systems;
- therefore the entropy of one system in the ensemble is $S/\mathcal{A} = -k_B \sum_i p_i \ln p_i$.



Some very recent developments

Switching Experiment

- Consider an irreversible process in which we switch the value of a thermodynamic variable λ from $0 \rightarrow 1$
- We start the switching when the system is in equilibrium with the surroundings at inverse temperature β .
- Let τ be the switching time
- during the switching process and at the end of the switching process the system need not be in equilibrium
- the system can go far from equilibrium
- Let $\{W_i : i = 1, N\}$ be the values of W observed in N switching experiments, all carried out with the same protocol



Work fluctuations

- Work fluctuation theorem says, for experiments carried out with an irreversible protocol, we can

$$\frac{1}{N} \sum_{i=1}^N \exp(-\beta W) = \exp(-\beta \Delta F)$$

- on the left we have a quantity measured in a nonequilibrium process
- on the right we have an equilibrium quantity
- an equilibrium quantity can be obtained from measurements from irreversible processes
- this is a remarkable finding



Heat and entropy fluctuations

- there are other equalities relating non-equilibrium measurements to equilibrium quantities based on
 - entropy and
 - heat fluctuations
- there are attempts to define entropy for nonequilibrium systems from Chaos theory
- some success has been obtained for steady state systems
- these constitute recent trends in thermodynamics
- these constitute recent developments in thermodynamics



dissipation is proportional to fluctuations

- exponential is a convex function : $\langle e^{-x} \rangle > e^{\langle x \rangle}$
- we recover the second law inequality from work fluctuation theorem $\langle W \rangle > \Delta F$ for an irreversible process
- in the reversible limit $\tau \rightarrow \infty$ we recover conventional thermodynamics relation $\langle W \rangle = \Delta F$
- For τ large but not infinity, we get

$$\langle W \rangle - \Delta F = \frac{1}{2} \beta \sigma_W^2$$

dissipation is proportional to fluctuation



fluctuation dissipation theorem

- Let

$$p = \int_{-\infty}^{W_R} \rho(W, \tau) dW$$

- in the above $\rho(W, \tau)$ describes the work ensemble of non-equilibrium experiments all carried out with the same protocol over time duration τ
- in the reversible limit there is no dissipation $W_d = 0$
- in the reversible limit we also have $\sigma_W^2 = 0$
- Fluctuation dissipation theorem says

$$\sigma_W^2 \propto W_d$$



(Maxwell's) demons don't die

- the above implies $\sigma_W \gg W_d$
- after a bit of algebra we find that when the process becomes more and more reversible $p \rightarrow 1/2$
- p is the probability of second law violation at microscopic level
- the result that $p = 1/2$ in the reversible limit is paradoxical
- Is it that we have Maxwell's demon reborn ?
- these are open questions and we require fresh thinking from young minds



Thanks and Good bye

THANKS

