
Phys 359

STATISTICAL MECHANICS

University of Waterloo

Course notes by: TC Fraser

Instructor: Michel Gingras

tcfraser@tcfraser.com

Version: 1.0

Table Of Contents

	Page
1 Introduction	4
1.1 What is Statistical Mechanics	4
1.2 History	4
2 Foundations	4
2.1 Essence of Statistical Mechanics	4
2.2 Postulate of Statistical Mechanics	5
2.3 Perspective from Coin Tossing	6
2.4 Stirlings Formula and Gaussian Integrals	6
2.4.1 Differentiation Trick	7
2.4.2 Stirling's Formula	7
2.4.3 Gaussian Integrals	8
2.5 Connections between Thermodynamics and Statistical Mechanics	9
2.6 Example of a Physical System with Constraint	10
3 Review of Thermodynamics	10
3.1 Definitions	10
3.2 Zeroth Law of Thermodynamics	10
3.3 Functions of State	10
3.4 Work	11
3.4.1 Configuration Work	12
3.4.2 Dissipative Work	12
3.4.3 Sign Convention	12
3.4.4 Abiabatic Work	12

Disclaimer

These notes are intended to be a reference for my future self (TC Fraser). If you the reader find these notes useful for yourself in any capacity, please feel free to use these notes as you wish, free of charge. However, I do not guarantee their complete accuracy and mistakes are likely present. If you notice any errors please email me at **tcfraser@tcfraser.com**, or contribute directly at **<https://github.com/tcfraser/course-notes>**. If you are the professor of this course and you've managed to stumble upon these notes and would like to make large changes or additions, email me please.

Latest versions of all my course notes are available at **www.tcfraser.com**.

1 Introduction

1.1 What is Statistical Mechanics

Statistical Mechanics is the area of Physics interested in systems with a large number of degrees of freedom n . Note that these variables can be interacting or not.

There are two distinct classes of Statistical Mechanics: equilibrium and non-equilibrium.

The Statistical part of Statistical Mechanics implies that it is inherently a study of probabilities and probability distributions. These laws must still remain fully consistent with physical laws.

Typically, systems are analyzed on a microscopic level. For a system of particles with charges $\{q_i\}$ and their positions $\{\vec{r}_i\}$, the dynamics are governed by the forces acting on each particle,

$$\vec{F}_i = m_i \vec{a}_i = \sum_{i \neq j} \frac{q_i q_j}{4\pi\epsilon_0 |\vec{r}_{ij}|^2}$$

But does labelling the particles really matter? For the case of $N \rightarrow \infty$, the global phenomenology is of interest.

1.2 History

1738 Daniel Bernoulli

- molecules moving in container, they collide with one another
- collisions with walls explain pressure

1850 Gay Lussac, Joule, Thomson (Lord Kelvin), Carnot

1859 James Clerk Maxwell

$$- D(\nu) \sim e^{-\frac{\nu^2}{2k_B T}}$$

1884 Josiah Willard Gibbs

- ensemble averaging

1900 Planck, Einstein, Bose, Pauli, Fermi, Dirac

Today Frontier is in non-equilibrium Statistical Mechanics

- cold atoms
- biology
- quantum information

2 Foundations

2.1 Essence of Statistical Mechanics

Laws of Thermodynamics:

Pros	Cons
<ul style="list-style-type: none"> • great because they are totally general • relationship's (Maxwell's relations) between c_p, c_v, α, κ 	<ul style="list-style-type: none"> • does not tell us how to compute anything • does not tell us what entropy is

The 2nd law of Thermodynamics reveals $dU = \delta Q - \delta W$ where $\delta Q = T dS$.

But what is S and what does it **physically** mean? Boltzman reveals the relation:

$$S = k_B \ln(\Omega)$$

Which we will come back to.

2.2 Postulate of Statistical Mechanics

There is only one postulate of Statistical Mechanics:

For an isolated system in equilibrium, all microstates accessible to the system are equally probable.

In order to digest this postulate, we will require some definitions.

Definitions:

- system
 - part of the universe we care about
 - only weakly coupled to the rest of the universe
 - the dynamics/mechanics are dominated by the internal degrees of freedom and forces
- isolated
 - idealization
 - eliminates all external influences; no force, no energy/heat flux and no particle flux
 - quantities such as the energy, number of particles and volume assumed constant forever $dU, dN, dV = 0$
- equilibrium
 - everything is no-longer changing
- microstate
 - a complete/total description of everything at the microscopic level $\{\vec{r}_i, \vec{p}_i\}$ for each i
- macrostate
 - a description at the macroscopic level in accordance with the external constraints
 - U, P, T, \bar{M}
- equally probable
 - we are dealing with probabilities and statistics
 - microstates are somehow describing probabilistically the properties at the macroscopic level
- accessible
 - consistency with the macroscopic constraints imposed by the conservation laws (fixed energy, fixed number of particles)

Postulate Follow-up:

We assume that the observed/realized macrostate is the one with the most microstates.

2.3 Perspective from Coin Tossing

Consider 4 coins tossed many, many times. What are the microstates describing this system?

Macrostate Label	Macrostate N_H N_T		Microstate A B C D				Thermo Probability	True Probability
1	4	0	H	H	H	H	1	1/16
2	3	1	H	H	H	T	4	4/16
			H	H	T	H		
			H	T	H	H		
			T	H	H	H		
3	2	2	H	H	T	T	6	6/16
			H	T	T	H		
			T	T	H	H		
			T	H	H	T		
			H	T	H	T		
			T	H	T	H		
4	1	3	T	T	T	H	4	4/16
			T	T	H	T		
			T	H	T	T		
			H	T	T	T		
5	0	4	T	T	T	T	1	1/16

We note that the most probable macrostate 3 is the one with the most microstates 6.

How do we deal with very large N, N_H, N_T in order to locate the most likely macrostate? First let's get a general expression for Ω where Ω is the number of microstates. Since $N = N_H + N_T$ and N is considered fixed, there is only one free parameter N_H (taken by choice). Thus Ω can be considered a function of N_H and nothing else.

Recall from probability that the form for Ω is given by,

$$\Omega = \frac{N!}{N_H!(N - N_H)!}$$

The most likely macrostate is given when Ω (the number of microstates) is maximized. This means that we are interested in finding values of N_H , namely N_H^* where,

$$\left(\frac{d\Omega}{dN_H} \right) \Big|_{N_H=N_H^*} = 0 \quad \left(\frac{d^2\Omega}{dN_H^2} \right) \Big|_{N_H=N_H^*} < 0$$

In order to do this, we will need to explore some mathematics ideas.

2.4 Stirlings Formula and Gaussian Integrals

Consider the integral,

$$I = \int_0^{\infty} x^N e^{-x} dx$$

This can be evaluated using integration by parts,

$$I = N \int_0^{\infty} x^{N-1} e^{-x} dx = \dots = N! \quad (2.1)$$

2.4.1 Differentiation Trick

However, integration by parts N times on (2.1) is annoying. There is a nice trick. Notice that,

$$\int_0^{\infty} e^{-ax} dx = \left(-\frac{1}{a} e^{-ax} \right) \Big|_0^{\infty} = \frac{1}{a} \quad (2.2)$$

One can treat a as a *dummy* variable, and examine (2.2)'s derivative with respect to a ,

$$\frac{\partial}{\partial a} \int_0^{\infty} e^{-ax} dx = \int_0^{\infty} \frac{\partial}{\partial a} e^{-ax} dx = \int_0^{\infty} -x e^{-ax} dx = \frac{\partial}{\partial a} \left(\frac{1}{a} \right) = -\frac{1}{a^2}$$

The reason for doing this is to simplify the process of (2.1).

If one explores the N^{th} derivative of (2.2) with respect to a , you will derive the expression,

$$\left[(-1)^N \frac{\partial^N}{\partial a^N} \int_0^{\infty} e^{-ax} dx \right]_{a=1} = N! \quad (2.3)$$

The $(-1)^N$ term is a result of the alternating sign induced by bringing down a $-x$ each time you take a derivative.

2.4.2 Stirling's Formula

Looking back at the integral (2.1),

$$\int_0^{\infty} x^N e^{-x} dx = N! \quad (2.4)$$

How can we approximate $N!$ using the left hand side of (2.4)? To derive Stirling's Formula, we need to make a change of variables $x = N + \sqrt{N}y$. Substituting into (2.4) gives,

$$N! = \int_0^{\infty} \sqrt{N} e^{-N} e^{N \ln(N + \sqrt{N}y)} e^{-\sqrt{N}y} dy$$

The approximation begins by expanding the logarithm for large N ,

$$\ln(N + \sqrt{N}y) = \ln \left(N \left[1 + \frac{y}{\sqrt{N}} \right] \right) = \ln(N) + \ln \left(1 + \frac{y}{\sqrt{N}} \right)$$

Take $\epsilon = \frac{y}{\sqrt{N}} \ll 1$ and apply Taylor series,

$$\ln(1 + \epsilon) \approx \epsilon - \frac{\epsilon^2}{2}$$

Thus,

$$N! \approx \sqrt{N} e^{-N} N^N \int_{-\sqrt{N}}^{\infty} e^{-\frac{y^2}{2}} dy$$

The lower bound can be approximated as ∞ since N is so large,

$$N! \approx \sqrt{N} e^{-N} N^N \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy$$

Notice the remaining integral term. It is called the *Gaussian Integral* and has solution (see [Gaussian Integrals](#)),

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \sqrt{\frac{\pi}{a}} \quad (2.5)$$

Thus letting $a = 1/2$,

$$N! \approx \sqrt{2\pi N} e^{-N} N^N \quad (2.6)$$

Equation (2.6) is known as *Stirling's Formula*. However, there is a much more useful form of Stirling's Formula. It is obtained by taking the logarithm of both sides,

$$\ln(N!) \approx \left(N + \frac{1}{2} \right) \ln(N) - \left(N - \underbrace{\frac{1}{2} \ln(2\pi)}_{\text{small compared to large } N} \right)$$

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

Note that the remaining N is not dropped. This is because for $N \sim 10^{23}$, $N \ln N - N$ and $N \ln N$ differ by about 2%.

Now we can apply this to the problem of maximizing Ω (which is equivalent to maximizing $\ln \Omega$) because the logarithm is monotonically increasing.

$$0 = \frac{\partial \ln \Omega}{\partial N_H} = \frac{\partial}{\partial N_H} \left[\ln \left(\frac{N!}{N_H! (N - N_H)!} \right) \right]$$

Through some manipulation, and applying (2.7), one obtains the expected result,

$$N_H = \frac{N}{2}$$

2.4.3 Gaussian Integrals

Before continuing, we should take a moment to explore how (2.5) is solved. Let,

$$I_x = \int_{-\infty}^{\infty} e^{-ax^2} dx$$

Here comes the trick. Multiply I_x by itself and switch from rectangular coordinates to polar coordinates,

$$I_x I_y = \int_{-\infty}^{\infty} e^{-ax^2} dx \int_{-\infty}^{\infty} e^{-ay^2} dy$$

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-a(x^2+y^2)} dx dy$$

Where we take $\mathbb{R}^2(x, y) \mapsto \mathbb{R}^2(r, \phi)$

$$I^2 = \int_0^{2\pi} \int_0^{\infty} r e^{-ar^2} dr d\phi$$

Which reveals that $I^2 = \pi/a$. Thus,

$$I = \sqrt{\frac{\pi}{a}}$$

2.5 Connections between Thermodynamics and Statistical Mechanics

Consider a lattice of Cu^{2+} atoms. In a lattice the Cu^{2+} atoms are distinguishable because they have unique locations. Now apply an external magnetic field.

$$H_{\text{Zeeman}} = -\vec{\mu} \cdot \vec{B}$$

Recall that $\vec{u} = g\mu_B \vec{s}$ has units J/T where T is tesla. Where for an electron,

$$\mu_B = \frac{e\hbar}{2m} = 9 \times 10^{-24} \text{ JT}^{-1} \quad g \approx 2$$

For $\vec{B} = B\hat{z}$, $H_{\text{Zeeman}} = 2\mu_B B s_z \equiv b s_z$. The splitting of the two spin states $s_z = \pm 1$ for $B = 1 \text{ T}$ has characteristic temperature of,

$$\frac{H_{\text{Zeeman}}}{k_B} = \frac{\varepsilon}{k_B} = \frac{10 \times 10^{-23} \text{ J}}{1.4 \times 10^{23} \text{ JK}^{-1}} \approx 0.6 \text{ K}$$

Now consider N electrons subject to the field \vec{B} where there are N_+ spins “up” and N_- spins “down”. This is completely analogous to the coin flipping example. The total energy of the system is given by,

$$U = -N_- \varepsilon + N_+ \varepsilon$$

Note that $N = N_+ + N_-$ and thus,

$$\frac{U}{N} = \varepsilon - 2\varepsilon \frac{N_-}{N}$$

Constraining U and using the substitution,

$$\frac{N_-}{N} = \frac{1-x}{2} \quad \frac{N_+}{N} = \frac{1+x}{2}$$

Then the microstate measure is given by,

$$\Omega = \frac{N!}{N_+! N_-!}$$

Becomes (after some manipulation as using (2.7))

$$\ln \Omega = -N \left[\left(\frac{1+x}{2} \right) \ln \left(\frac{1+x}{2} \right) + \left(\frac{1-x}{2} \right) \ln \left(\frac{1-x}{2} \right) \right]$$

Now recall that for fixed volume $dV = 0$,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_V$$

But since U depends only on x , we can write,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial x} \right) \left(\frac{\partial x}{\partial U} \right)$$

Thus reveals a slight connection between S the entropy and Ω through x in this example. Further analysis with motivate Boltzman's equation,

$$S = k_B \ln \Omega + S_0$$

2.6 Example of a Physical System with Constraint

Suppose you have 3 particles called A, B, C such that each particle can have $\varepsilon_j = j\varepsilon$ where $j = 0, 1, 2, 3, \dots$

How many microstates are there subject to the constraint that the total energy is 3ε ?

Macrostate Label	Macrostate				Microstate			Thermo Probability	True Probability
	N_0	N_1	N_2	N_3	A	B	C		
1	2	0	0	1	0	0	3ε	3	3/10
					0	3ε	0		
					3ε	0	0		
2	0	1	1	0	0	ε	2ε	6	6/10
					ε	2ε	0		
					2ε	0	ε		
					0	2ε	ε		
					2ε	ε	0		
					ε	0	2ε		
3	0	3	0	0	ε	ε	ε	1	1/10

3 Review of Thermodynamics

3.1 Definitions

Recall Boyle's Law $PV = nRT$.

- processes
 - constant T , isothermal process
 - constant P , isobaric process
 - constant V , isochoric process or isovolumetric process
 - constant S , adiabatic process
 - * Comes from greek *diabatos* which means *to go through*
 - * No heat exchange between system and surroundings
 - * Can also be an approximation for processes that occur really quickly over a short period of time
- processes reversible/irreversible
 - reversible process happens over a number of discrete steps and that are each reversible
 - irreversible processes are like poking a hole in a balloon or a gas expanding in a vacuum
- thermodynamic variables T, P, V, U where U is the *total energy*

3.2 Zeroth Law of Thermodynamics

If systems A and B are in equilibrium with one another and systems B and C are in equilibrium then A is in equilibrium with C .

3.3 Functions of State

Thermodynamic variables are not independent. They are often related by an equation of state. For example $PV = nRT$ for an ideal gas. Other variables might come into an equation of state. For example ρ, T, E are all important for the equation of state for a liquid crystal sample.

Equations of state with typically look like

$$f(P, V, T) = 0$$

Another important notion is the notion of *function of state*. A quantity that depends only on the thermodynamic variables of the system and *not its history*, is called a *function of state*.

We will first focus on $U_{\text{total energy}}$ first and then S_{entropy} as our functions of state.

Mathematically, $G = g(x, y)$ where x, y are the thermodynamic variables and G is a function of state analytic everywhere and obeys some properties:

- $dG = \left(\frac{\partial G}{\partial x}\right)_y dx + \left(\frac{\partial G}{\partial y}\right)_x dy$
- at most values of the thermodynamic variables $g(x, y)$ is “smooth”.
- for example: $\left(\frac{\partial^2 G}{\partial x^2}\right)_y$ or $\left(\frac{\partial^2 G}{\partial y^2}\right)_x$ or $\left(\frac{\partial}{\partial x} \left(\frac{\partial G}{\partial y}\right)_x\right)_y$ are all continuous
- the order of discontinuities is determined by whether or not the system or substance is undergoing transitions of state or not

For functions of state that are analytical everywhere, the order of derivatives is inconsequential.

$$\left(\frac{\partial}{\partial x} \left(\frac{\partial G}{\partial y}\right)\right) = \left(\frac{\partial}{\partial y} \left(\frac{\partial G}{\partial x}\right)\right)$$

What can we say about G in cases where

$$dG = \frac{\partial G}{\partial x} dx + \frac{\partial G}{\partial y} dy \quad (3.1)$$

in the case of functions of state like U , $dU = \delta Q - \delta W$ and the inexact differentials and how they relate to exact differentials like dV and dS ? In particular, when can we integrate dU ?

Answer: equation (3.1) can be integrated in situations where

$$\left(\frac{\partial}{\partial y} \left(\frac{\partial G}{\partial x}\right)_y\right)_x = \left(\frac{\partial}{\partial x} \left(\frac{\partial G}{\partial y}\right)_x\right)_y \quad (3.2)$$

When equation (3.2) is held for a physical system, one can say that dG is an **exact differential**. Review 12,13 in notes on “Review of Thermodynamics”.

The difference in the function $G(x, y)$ between two sufficiently close pairs of points (x_1, y_1) and (x_2, y_2) depends only on the difference in $G(x, y)$ evaluated at those two points.

$$\Delta G = G(x_2, y_2) - G(x_1, y_1)$$

ΔG does not depend on the path from point (x_1, y_1) to (x_2, y_2) . In practice, one can assign such a function G to the values of the thermodynamic variables at the points (x, y) . For example, $U(P, V)$ is such a function of state. It depends only on the description through the thermodynamic variables and not the history.

By counter example, heat Q is not a function of state. No one can say, “that substance has X units of heat in it”.

3.4 Work

There are two types of work. One is called *configuration work* and the other is called *dissipative work*.

3.4.1 Configuration Work

Configurational work is denoted $\mathrm{d}W$ where the symbol d represents that it is **not** an exact differential.

$$\mathrm{d}W = \sum_i y_i \mathrm{d}x_i$$

where y_i is an intensive variable (not proportional to N, V ; examples: pressure, surface tension). It can be thought of as a generalized force. Here $\mathrm{d}x_i$ is the generalized displacement which is an extensive variable.

3.4.2 Dissipative Work

Dissipative work can be thought of as “stirring work”. Examples include a mixer in a liquid or an electrical wire/resistor.

- electrical power:
 - $P = V \cdot I$
 - $\mathrm{d}W_{\text{dis}} = P \cdot \mathrm{d}t = RI^2 \mathrm{d}t$

3.4.3 Sign Convention

$\mathrm{d}W > 0$ work done by the system

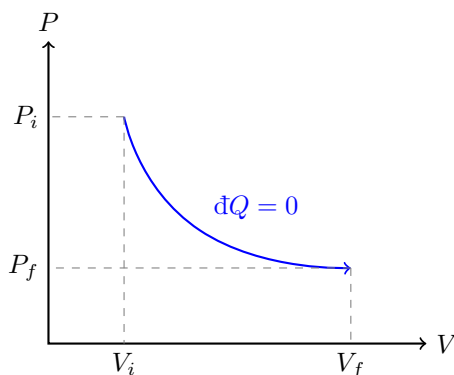
$\mathrm{d}W < 0$ work done on the system

Note:

- work is **not** a property of the system
- work is **not** a function of state
- integration on a closed loop is not degenerate $\oint \mathrm{d}W \neq 0$

3.4.4 Adiabatic Work

Adiabatic work occurs with no heat exchange.



Adiabatic work is done between any two equilibrium states and is *independent of path*. One can define a function of state as the total adiabatic work done on a system. Let's define this as the total internal energy of the system.

$$\mathrm{d}U = -\mathrm{d}W_{\text{adiabatic}}$$

Note the minus ‘-’ sign is due to the fact that the total energy of the system increases when work is done **on** this system.