Phys 359

STATISTICAL MECHANICS

University of Waterloo

Course notes by: TC Fraser
Instructor: Michel Gingras

tcfraser@tcfraser.com

Version: 1.1

Table Of Contents

Page

1 Introduction **5**
1.1 What is Statistical Mechanics . . . 5
1.2 History . . . 5

2 Foundations **5**
2.1 Essence of Statistical Mechanics . . . 5
2.2 Postulate of Statistical Mechanics . . . 6
2.3 Perspective from Coin Tossing . . . 7
2.4 Stirlings Formula and Gaussian Integrals . . . 7
2.4.1 Differentiation Trick . . . 8
2.4.2 Stirling's Formula . . . 8
2.4.3 Gaussian Integrals . . . 9
2.5 Connections between Thermodynamics and Statistical Mechanics . . . 10
2.6 Example of a Physical System with Constraint . . . 11

3 Review of Thermodynamics **11**
3.1 Definitions . . . 11
3.2 Zeroth Law of Thermodynamics . . . 11
3.3 Functions of State . . . 12
3.4 Work . . . 13
3.4.1 Configuration Work . . . 13
3.4.2 Dissipative Work . . . 13
3.4.3 Sign Convention . . . 13
3.4.4 Abiabatic Work . . . 13
3.5 First of Law of Thermodynamics . . . 14
3.5.1 Heat . . . 14
3.5.2 Heat Capacities . . . 14
3.6 Adiabatic Work Continued . . . 15
3.6.1 Gay-Lussac Experiment . . . 16
3.6.2 Joule-Thomson(Kelvin) Experiment . . . 18
3.6.3 Summary of First Law . . . 18
3.7 Second Law of Thermodynamics . . . 18
3.7.1 Heat Engines . . . 18
3.7.2 Carnot Cycle/Engine . . . 19
3.7.3 Second Law Statements . . . 20
3.7.4 Entropy . . . 21
3.7.5 Clausius Inequality . . . 23
3.7.6 Perspective on Entropy . . . 26
3.7.7 Conclusion . . . 27

4 S.M. Basis for T.D. **27**
4.1 Contact Between S.M. & T.D. . . . 27
4.2 Classical Ideal Gas from S.M. . . . 30
4.2.1 Volume . . . 30
4.2.2 Energy & Particle in a Box . . . 31
4.3 Microcanonical Ensemble . . . 32
4.3.1 Solving the Microcanonical Ensemble . . . 33
4.4 Gibb's Paradox . . . 37

5 Ensemble Theory **38**
5.1 Phase Space of Classical System . . . 39
5.1.1 Dynamics & Phase Space . . . 39
5.1.2 Trajectories in Phase Space . . . 40
5.2 Essential Replacement for Ensemble Theory . . . 40
5.2.1 Ergodic Hypothesis . . . 41
5.2.2 Mixing Hypothesis . . . 41
5.2.3 Statistical Ensemble . . . 41
5.2.4 Liouville's Theorem . . . 42
5.3 Microcanonical Ensemble Revisited . . . 44
5.4 Two Examples . . . 46
5.4.1 Compute Entropy . . . 46
5.4.2 Simple 1D Classical Harmonic Oscillator . . . 47
5.4.3 Reflection . . . 48

6 Canonical Ensemble **48**
6.1 Preamble . . . 48
6.2 Tackling the Canonical Ensemble . . . 49
6.2.1 Example . . . 50
6.3 Canonical Ensemble Description . . . 51
6.4 Physical Significance of Various Statistical Variables in the Canonical Ensemble . . . 52
6.4.1 Logarithm of Partition Function . . . 53
6.4.2 Helmholtz Free Energy . . . 53
6.5 Summary of Z . . . 54
6.6 Alternative Form for Z & Degeneracy of States . . . 54
6.7 Classical Systems / Free Particles . . . 55
6.8 Examples of C.E. . . . 58
6.8.1 Two Masses on Springs . . . 58
6.8.2 Rigid Polar Rotator . . . 58
6.8.3 Polarizable Molecule . . . 61
6.8.4 Permanent Electric Dipole . . . 62
6.8.5 Spin in Magnetic Field . . . 63
6.8.6 Two Level Systems . . . 64
6.9 Energy Fluctuations . . . 64
6.10 Equipartition & Virial Theorems . . . 67
6.10.1 Equipartition . . . 67
6.10.2 Virial Theorem . . . 69

7 Grand Canonical Ensemble **70**
7.1 Derivation . . . 70
7.2 Comparison to Canonical Ensemble . . . 71
7.3 Physical Significance of α, β . . . 72
7.4 Grand Canonical Partition Function . . . 73

8 Ideal Quantum Gases **75**
8.1 Microcanonical Ensemble For IQG . . . 75
8.2 Statistical Distributions . . . 76
8.3 Entropy & Distributions . . . 77
8.4 Ideal gas in other Q.M. ensembles . . . 77
8.4.1 Implications of $\langle n_\varepsilon \rangle$ . . . 81

Disclaimer

These notes are intended to be a reference for my future self (TC Fraser). If you the reader find these notes useful 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 (Michel Gingras) 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 class 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,

$$\vec{F}_i = m_i \vec{a}_i = \sum_{i \neq j} \frac{q_i q_j}{4\pi\epsilon_0} \frac{1}{|\vec{r}_{ij}|^2}$$

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

1.2 History

1738 Daniel Bernoulli

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

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

1859 James Clerk Maxwell

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

1884 Josiash 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
• great because they are totally general • relationship's (Maxwell's relations) between c_p, c_v, α, κ	• does not tell us how to compute anything • does not tell us what entropy is

The 2nd law of Thermodynamics reveals $\mathrm{d}U = đQ - đW$ where $đQ = T\mathrm{d}S$.

But what is S and what does it **physically** mean? Boltzmann 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 $\mathrm{d}U, \mathrm{d}N, \mathrm{d}V = 0$
- 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		Microstate				Thermo Probability	True Probability
	N_H	N_T	A	B	C	D		
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 likes get a general expression for Ω were 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,

$$\left(\frac{\mathrm{d}\Omega}{\mathrm{d}N_H}\right|_{N_H=N_H^*} = 0 \qquad \left(\frac{\mathrm{d}^2\Omega}{\mathrm{d}N_H^2}\right|_{N_H=N_H^*} > 0$$

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

2.4 Stirlings Formula and Gaussian Integrals

Consider the integral,

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

This can be evaluated using integration by parts,

$$I = N \int_0^\infty x^{N-1} e^{-x} \mathrm{d}x = \cdots = N! \tag{2.1}$$

2.4.1 Differentiation Trick

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

$$\int_0^\infty e^{-ax}\mathrm{d}x = \left(-\frac{1}{a}e^{-ax}\right|_0^\infty = \frac{1}{a} \tag{2.2}$$

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

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

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

If one explores the Nth derivative of (2.2) with respect to a, you will derive the expression,

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

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),

$$\int_0^\infty x^N e^{-x}\mathrm{d}x = N! \tag{2.4}$$

How can we approximate $N!$ using the left had 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,

$$N! = \int_0^\infty \sqrt{N}e^{-N}e^{N\ln\left(N+\sqrt{N}y\right)}e^{-\sqrt{N}y}\mathrm{d}y$$

The approximation begins by expanding the logarithm for large N,

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

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

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

Thus,

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

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

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

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

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \mathrm{d}y = \sqrt{\frac{\pi}{a}} \tag{2.5}$$

Thus letting $a = 1/2$,

$$N! \approx \sqrt{2\pi N} e^{-N} N^N \tag{2.6}$$

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,

$$\ln(N!) \approx \left(N + \frac{1}{2}\right) \ln(N) - \left(N - \frac{1}{2} \ln(2\pi)\right)$$

But $\frac{1}{2}\ln(2\pi)$ is much smaller than N for very large N,

$$\ln(N!) \approx N \ln N - N \tag{2.7}$$

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,

$$I_x = \int_{-\infty}^{\infty} e^{-ax^2} \mathrm{d}x$$

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

$$I_x I_y = \int_{-\infty}^{\infty} e^{-ax^2} \mathrm{d}x \int_{-\infty}^{\infty} e^{-ay^2} \mathrm{d}y$$

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-a\left(x^2+y^2\right)} \mathrm{d}x \mathrm{d}y$$

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

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

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,

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

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,

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

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 $\mathrm{d}V = 0$,

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

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 Boltzmann'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 $j = 0, 1, 2, 3, \ldots$.

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 $\rho, \mathcal{T}, E$ are all important for the equation of state for a liquid crystal sample.

Equations of state with typically look like

$$f\left(P, V, T\right) = 0$$

Another important notion is the notion of *function of state*. A quantity that depends only on the thermodynamic variables of the system and *not it's 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, $G = g\left(x, y\right)$ where x, y are the thermodynamic variables and G is a function of state analytic everywhere and obeys some properties:

- $\mathrm{d}G = \left(\frac{\partial G}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial G}{\partial y}\right)_x \mathrm{d}y$
- 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

$$\mathrm{d}G = \frac{\partial G}{\partial x}\mathrm{d}x + \frac{\partial G}{\partial y}\mathrm{d}y \tag{3.1}$$

in the case of functions of state like U, $\mathrm{d}U = đQ - đW$ and the inexact differentials and how they relate to exact differentials like $\mathrm{d}V$ and $\mathrm{d}S$? In particular, when can we integrate $\mathrm{d}U$?
Answer: equation (3.1) can be integrated in situations where

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

When equation (3.2) is held for a physical system, one can say that $\mathrm{d}G$ 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_1, y_1). 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

Configuational work is denoted $đW$ where the symbol đ represents that it **is not** and exact differential.

$$đW = \sum_i y_i \mathrm{d}x_i$$

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

3.4.3 Sign Convention

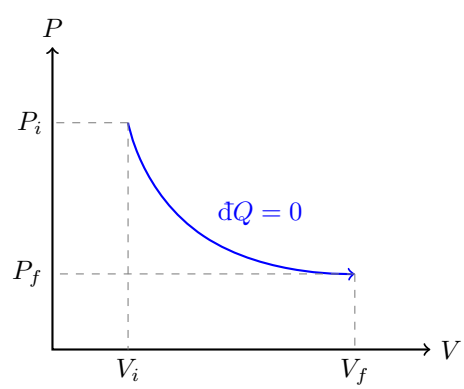
$$đW > 0 \quad \text{work done by the system}$$
$$đW < 0 \quad \text{work done on the system}$$

Note:

- work is **not** a property of the system
- work is **not** a function of state
- integration on a closed loop is not degenerate $\oint đW \neq 0$

3.4.4 Abiabatic 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. Lets define this a 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. Along adiabatic paths, there is an exact differential associated with $đW$. It is called internal energy $\mathrm{d}U$.

3.5 First of Law of Thermodynamics

3.5.1 Heat

In systems such as boiling a pot of water, there is no work performed on the system however the state of the system changes (not boiling to boiling). The mechanism that causes this change of state is heat (Q).

$$\mathrm{d}U = đW + đQ \tag{3.3}$$

Equation (3.3) is using the following convention:

$$\begin{aligned} đQ &:< 0 \quad \text{heat flow } \textbf{in} \\ đQ &:> 0 \quad \text{heat flow } \textbf{out} \\ đW &:< 0 \quad \text{work is done } \textbf{by} \text{ system} \\ đW &:> 0 \quad \text{work is done } \textbf{on} \text{ system} \end{aligned}$$

3.5.2 Heat Capacities

Use the convention that C is the heat capacity (extensive) and c is the specific heat capacity (intensive).

$$c = \frac{C}{n}$$

$$C \equiv \lim_{\Delta T \to 0} \frac{\Delta Q}{\Delta T}$$

Heat capacity C is very important. Measuring C in a metal at low temperatures reveals very important intrinsic properties of the material. Also note that heat capacity is always $C \sim \frac{1}{T}$ and not some other power than -1.

Relationship Between c_v, c_p for an Ideal Gas:

$$PV = nRT$$

Using Boyle's law, work with $n = 1\,\text{mol}$ for convenience. How can we relate C to the properties of the substance? Let us write (noting the '-' convention),

$$\mathrm{d}U = đQ - đW$$

Switch to specific (per mole) quantities,

$$\mathrm{d}u = đq - đw$$

$$\mathrm{d}u = \left(\frac{\partial u}{\partial T}\right)_v \mathrm{d}T + \left(\frac{\partial u}{\partial v}\right)_T \mathrm{d}V$$

$$đq = \mathrm{d}u + đw$$
$$đq = \left(\frac{\partial u}{\partial T}\right)_v \mathrm{d}T + \left\{\left(\frac{\partial u}{\partial v}\right)_T \mathrm{d}V + P\mathrm{d}v\right\}$$

However,

$$c_v = \lim_{\mathrm{d}T\to 0} \frac{đq}{\mathrm{d}T} \quad \text{At constant V}$$

Thus,

$$c_v = \left(\frac{\partial u}{\partial T}\right)_v \tag{3.4}$$

Which gives us,

$$đq = c_v \mathrm{d}T + P\mathrm{d}v$$

Note that this formula uses the fact that,

$$\left(\frac{\partial U}{\partial v}\right)_T = 0 \quad \text{For ideal gas.}$$

Which we will prove momentarily, but can be taken as an experimental fact discovered by Guy-Lussac. Namely U is only a function of T for an ideal gas ($U = U(T)$).

$$đq = c_v \mathrm{d}T + P\mathrm{d}v + (v\mathrm{d}p - v\mathrm{d}p)$$
$$đq = (c_v + R)\,\mathrm{d}T - v\mathrm{d}p$$

Thus,

$$\lim_{\mathrm{d}T\to 0} \left(\frac{đq}{\mathrm{d}T}\right|_{p=\text{const.}} = c_v + R$$

Thus we can define a new quantity,

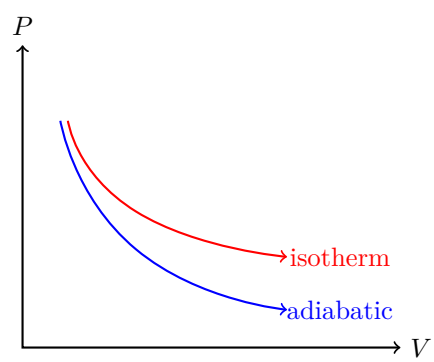
$$c_p \equiv \left(\frac{đq}{\mathrm{d}T}\right|_{p=\text{const.}} = c_v + R$$

For an monoatomic gas ($U = \frac{3}{2}RT$),

$$\gamma \equiv \frac{c_p}{c_v} = \frac{5}{3}$$

3.6 Adiabatic Work Continued

Let us examine an isotherm (red) next to an adiabatic curve (blue),



$$đq = c_v \mathrm{d}T + P\mathrm{d}v$$
$$đq = c_P \mathrm{d}T - v\mathrm{d}P$$

And let's set $đq = 0$ (adiabatic),

$$c_v \mathrm{d}T + P\mathrm{d}v = c_P \mathrm{d}T = v\mathrm{d}P$$

Which gives,

$$\frac{\mathrm{d}P}{P} = -\gamma \frac{\mathrm{d}v}{v} \quad \text{Using } \gamma = c_P/c_v$$

Or integrating along a curve,

$$\int_1^2 \frac{\mathrm{d}P}{P} = -\int_1^2 \gamma \frac{\mathrm{d}v}{v}$$
$$\ln\left(\frac{P_2}{P_1}\right) = -\gamma \ln\left(\frac{v_2}{v_1}\right)$$

Which gives the relationship,

$$P_1 V_1^\gamma = P_2 V_2^\gamma$$

Which holds true along an adiabatic curve. However, $PV = RT$. Thus using the expression $\mathrm{d}U = -P\mathrm{d}V|_{\text{adiabatic}}$ gives the result,

$$W = \left\{\frac{P_2V_2 - P_1V_1}{1-\gamma}\right\}_{\gamma>1}$$

3.6.1 Gay-Lussac Experiment

For an ideal gas,

$$\left(\frac{\partial U}{\partial V}\right)_T = 0$$
$$đQ = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left\{\left(\frac{\partial U}{\partial V}\right)_T + P\right\}\mathrm{d}V$$
$$\frac{đQ}{T} = \frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \frac{1}{T}\left\{\left(\frac{\partial U}{\partial V}\right)_T + P\right\}\mathrm{d}V$$

Employ an integrating factor for the inexact differentials. It will take us from and inexact differential to an exact differential.

Math Interlude:

$$đG = A(x,y)\mathrm{d}x + B(x,y)\mathrm{d}y$$

Multiply G by $\mu(x,y)$ and unknown function of x and y,

$$\mathrm{d}\tilde{G} = \mu \cdot đG$$

Thus,

$$\mathrm{d}\tilde{G} = \mu A\mathrm{d}x + \mu B\mathrm{d}y$$

Making,

$$\frac{\partial\left(\mu A\right)}{\partial y}=\frac{\partial\left(\mu B\right)}{\partial x} \tag{3.5}$$

Example:

$$\mathrm{d}G=\sin(y)\mathrm{d}x+\cos(y)\mathrm{d}y$$
$$\mathrm{d}\tilde{G}=\mu\sin(y)\mathrm{d}x+\mu\cos(y)\mathrm{d}y$$

Let us look for a μ such that $\mu(x,y)=\mu(x)$ and plugging into equation (3.5),

$$\frac{\partial\left(\mu\sin(y)\right)}{\partial y}=\frac{\partial\left(\mu\cos(y)\right)}{\partial x}$$
$$\mu(x)\cos(y)=\frac{\mathrm{d}\mu}{\mathrm{d}x}\cos(y)+\mu\frac{\mathrm{d}\cos y}{\mathrm{d}x}$$
$$\mu(x)\cancel{\cos(y)}=\frac{\mathrm{d}\mu}{\mathrm{d}x}\cancel{\cos(y)}+\mu\cancelto{0}{\frac{\mathrm{d}\cos y}{\mathrm{d}x}}$$

Thus,

$$\mu(x)=\frac{\mathrm{d}u}{\mathrm{d}x}$$

Which gives,

$$\mu(x)=e^x$$

The purpose of introducing the notion of an integrating factor is to reveal that although $đQ$ is an inexact differential, we can introduce an integrating factor $\frac{1}{T}$ that makes the quantity $\frac{đQ}{T}$ an **exact** differential. This quantity is somehow more important that Q as it describes a property of the system. We will see that this quantity is the motivation for entropy. Let us examine this,

$$\frac{đQ}{T}=\underbrace{\frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_V}_{A}\mathrm{d}T+\underbrace{\frac{1}{T}\left\{\left(\frac{\partial U}{\partial V}\right)_T+P\right\}}_{B}\mathrm{d}V$$

And thus,

$$\frac{\partial A}{\partial V}=\frac{\partial B}{\partial T}$$

Gives,

$$\frac{\partial A}{\partial V}=\frac{\partial}{\partial V}\left\{\frac{1}{T}\frac{\partial U}{\partial T}\right\}=\frac{1}{T}\frac{\partial^2 U}{\partial V\partial T}$$
$$\frac{\partial B}{\partial T}=\frac{\partial}{\partial T}\left\{\frac{1}{T}\left(\frac{\partial U}{\partial V}+P\right)\right\}=-\frac{1}{T^2}\left(\frac{\partial U}{\partial V}+P\right)+\frac{1}{T}\frac{\partial^2 U}{\partial T\partial V}+\frac{1}{T}\frac{\partial P}{\partial T}$$

Equating these two expressions (noticing that $\frac{\partial^2 U}{\partial T\partial V}=\frac{\partial^2 U}{\partial V\partial T}$ and that $PV=RT\implies\frac{P}{T^2}=\frac{R}{TV}$),

$$\frac{1}{T}\frac{\partial^2 U}{\partial V\partial T}=-\frac{1}{T^2}\left(\frac{\partial U}{\partial V}+P\right)+\frac{1}{T}\frac{\partial^2 U}{\partial T\partial V}+\frac{1}{T}\frac{\partial P}{\partial T}$$
$$\frac{1}{\cancel{T}}\cancel{\frac{\partial^2 U}{\partial V\partial T}}=-\frac{1}{T^2}\left(\frac{\partial U}{\partial V}+P\right)+\frac{1}{\cancel{T}}\cancel{\frac{\partial^2 U}{\partial T\partial V}}+\frac{1}{T}\frac{\partial P}{\partial T}$$
$$0=-\frac{1}{T^2}\left(\frac{\partial U}{\partial V}+P\right)+\frac{1}{T}\frac{\partial P}{\partial T}$$

$$\frac{1}{T^2}\left(\frac{\partial U}{\partial V}+P\right)=\frac{1}{T}\frac{\partial P}{\partial T}$$
$$\frac{1}{T^2}\frac{\partial U}{\partial V}+\frac{P}{T^2}=\frac{1}{T}\frac{\partial P}{\partial T}$$
$$\frac{1}{T^2}\frac{\partial U}{\partial V}+\frac{\cancel{R}}{\cancel{T}V}=\frac{\cancel{R}}{\cancel{T}V}$$

Therefore Gay-Lussac was destined to find this relationship through experiment:

$$\frac{\partial U}{\partial V}=0$$

3.6.2 Joule-Thomson(Kelvin) Experiment

Using $\frac{1}{T}$ as the integrating factor, one can go through a similar derivation as in section 3.6.1 to understand that the quantity S is a state function of a system with,

$$\mathrm{d}S \equiv \frac{đQ}{T}$$

3.6.3 Summary of First Law

In summary, the first law of thermodynamics is really just conservation of energy with expression,

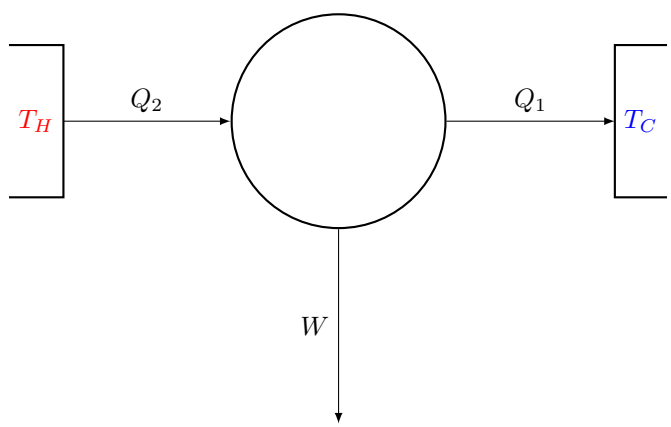
$$\mathrm{d}U = đQ - đW$$

3.7 Second Law of Thermodynamics

Physical laws on the microscopic scale always are found to have an inherent time-reversal symmetry. The physics behaves the same running forward in time and backward in time. However, a semi-paradox of real-world phenomenology seems to suggest that there are some processes can not occur moving backward in time. They still satisfy the conservation of energy, but there is another law that governs these macroscopic phenomena.

3.7.1 Heat Engines

Imagine a an engine E that runs from temperature T_2 to T_1 (two heat baths), where it extracts heat Q_2 and disposes heat Q_1 and does some work W.



The efficiency of the engine is defined as,

$$\eta \equiv \frac{\text{output}}{\text{input}} = \frac{W}{Q_2}$$

Note that on a cyclic process $\Delta U = 0$ since U is a function of state. However it is also equal to,

$$\Delta U = (Q_1 + Q_2) - W$$

Therefore typically the efficiency is given by,

$$\eta = 1 + \frac{Q_1}{Q_2}$$

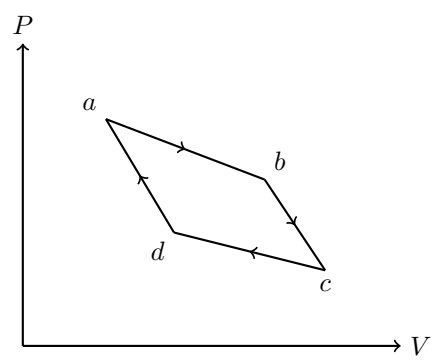
- Q_2 : injected into engine E and is > 0
- Q_1 : extracted from engine E and is < 0

This can be re-written as,

$$\eta = 1 + \frac{Q_1}{-|Q_2|} = 1 - \frac{|Q_1|}{|Q_2|}$$

3.7.2 Carnot Cycle/Engine

It can be shown that no engine is more efficient than a Carnot engine. Upon considering this, one can use entropy to expose the Second Law of Thermodynamics and the natural "arrow of time".



ab	isotherm (T_2)	$Q_2 > 0$	$W^{ab} > 0$
bc	adiabat	$Q = 0$	$W^{bc} > 0$
cd	isotherm (T_1)	$Q_1 < 0$	$W^{cd} < 0$
da	adiabat	$Q = 0$	$W^{da} < 0$

The total work is given by,

$$W_{\text{tot}} = W^{ab} + W^{bc} + W^{cd} + W^{da}$$

1) First the isotherms:

$$W \sim \int P\mathrm{d}V \quad \text{Use } PV = RT \text{ (1 mole)}$$

$$W = \int_1^2 \frac{RT}{V}\mathrm{d}V = RT \ln\left(\frac{V_2}{V_1}\right)$$

ab	$RT_2 \ln\left(\frac{V_b}{V_a}\right) > 0$	$\implies Q_2$
cd	$RT_1 \ln\left(\frac{V_d}{V_c}\right) < 0$	$\implies Q_1$

This is to be expected because we are on isotherms and $\Delta U = 0$.

2) Second the adiabats:
Along an adiabat,

$$PV^\gamma = \text{const.}$$
$$P_1 V_1^\gamma = P_2 V_2^\gamma$$

Furthermore, relating this to $PV = RT$,

$$\frac{RT_1}{V_1} V_1^\gamma = \frac{RT_2}{V_2} V_2^\gamma$$
$$T_1 V_1^{\gamma-1} = T_2 V_2^{\gamma-1}$$

Applying this to our diagram,

$$T_2 V_b^{\gamma-1} = T_1 V_c^{\gamma-1} \quad \text{and} \quad T_1 V_d^{\gamma-1} = T_2 V_a^{\gamma-1}$$

Which after rearrangement yields a constraint for the volumes,

$$\frac{V_b}{V_a} = \frac{V_c}{V_d}$$

Therefore,

$$\eta_{\text{Carnot}} = \eta_C = 1 + \frac{Q_1}{Q_2}$$
$$\eta_C = 1 + \frac{RT_1 \ln(V_d/V_c)}{RT_2 \ln(V_b/V_a)} = 1 - \frac{T_1}{T_2} \frac{\ln(V_c/V_d)}{\ln(V_b/V_c)}$$
$$\eta_C = 1 - \frac{T_1}{T_2}$$

3.7.3 Second Law Statements

1. **Clausius:** You can not build a machine or engine operating in a cycle can be constructed whose sole effect is to transfer heat from a cold to hot.

2. **Kelvin-Planck:** It is impossible to construct a device whose sole effect is to produce work by only absorbing heat from a single reservoir. Two reservoirs are needed.

Now let us use Clausius statement to prove Carnot's theorem as per the Carnot Engine; the most efficient engine is a Carnot engine.

Example:
Suppose we have an engine E that extracts heat from a hot T_2 reservoir $Q_2 = 100\,\text{cal}$, performs $W = 20\,\text{cal}$ of work and then dumps the remaining $Q_1 = 80\,\text{cal}$ calories of heat into a cold reservoir T_1. This is our Carnot engine.

$$\eta_C = \frac{W}{Q_2} = \frac{20}{100} = 20\%$$

Now suppose we had a better engine with $\eta_y > \eta_C$. Namely, this new engine has specs $Q_2 = 100\,\text{cal}, W = 40\,\text{cal}, Q_1 = 60\,\text{cal}$.

$$\eta_y = \frac{W}{Q_2} = \frac{40}{100} = 40\%$$

Now to prove is can not be the case, let us take our Carnot engine and reverse it's processes (making it a fridge), and doubling it's size. Therefore this Carnot fridge will require 40 cal to run. This can be supplied

by our "better" engine. Therefore, this combined engine has the net effect of absorbing 100 cal from a cold reservoir at T_1 and injecting 100 cal into a hot reservoir at temperature T_2, **with no other effect.** This violates the Clausius hypothesis.

Corollary:
All engines operating reversibly between two temperatures T_1 and T_2 have the same efficiency. In other words, the presence of any irreversible process degrades the efficiency.

$$\eta_{\text{irreversible}} < \eta_{\text{Carnot}} \tag{3.6}$$

3.7.4 Entropy

We have found that,

$$\eta_C = 1 + \frac{Q_1}{Q_2} = 1 - \frac{T_1}{T_2}$$

Or equivalently,

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0$$

For a single Carnot engine, the sum of heat at step n over the temperature at step n has sum,

$$\sum_{n=1}^{k} \frac{Q_n}{T_n} = 0 \tag{3.7}$$

Which gives that $\Delta Q/T$ over the whole cycle is zero. Taking the limit as $k \to \infty$, it suggests that the quantity $đQ/T$ is an exact differential. Introduce the label of entropy,

$$\mathrm{d}S = \frac{đQ}{T} \quad \text{with} \quad \oint \mathrm{d}S_{\text{rev}} = 0$$

As you can see, considering a Carnot engine also suggests the presence of a quantity S call entropy. Just as before, S is a state function.

How can we compute entropy?

Example (Constant Volume):
Let's consider heat injected at constant volume, and attempt to calculate ΔS.

$$\Delta U = \Delta Q - \Delta W \quad \text{At constant volume, } \Delta W = 0$$

But we also have $\Delta Q = C_v \Delta T$ where C_v is the heat capacity of the material. This combination yields,

$$\mathrm{d}S = \frac{đQ}{T} = \frac{C_v}{T}\mathrm{d}T$$

$$\int_1^2 \mathrm{d}S = \int_{T_1}^{T_2} \frac{C_v}{T}\mathrm{d}T = c_v \ln\left(\frac{T_2}{T_1}\right)$$

$$(S_2 - S_1) = C_v \ln\left(\frac{T_2}{T_1}\right)$$

Example (Constant Temperature):
Let's consider change of entropy for an isothermal process.

$$S_2 - S_1 = \int_1^2 \frac{đQ}{T}$$

But from the first law we have for constant temperature,

$$\mathrm{d}U = đQ - P\mathrm{d}V = 0$$

Thus,

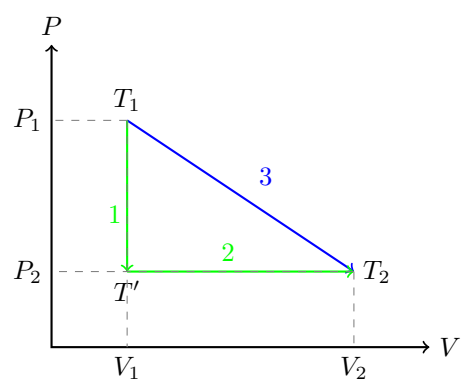
$$đQ = P\mathrm{d}V$$

$$\int_1^2 \mathrm{d}S = \int_1^2 \frac{P}{T}\mathrm{d}V = \int_1^2 \frac{nR}{V}\mathrm{d}V = nR\ln\left(\frac{V_2}{V_1}\right) \tag{3.8}$$

Summarize:

1. No engine is more efficient than Carnot.
2. As a by-produce, we discovered a quantity ΔS defined as:
   - $\Delta S = \frac{\Delta Q}{T}$ ($\mathrm{d}S = \frac{đQ}{T}$ for the infinitesimal)
   - $\mathrm{d}S$: defined at each point in (P, V, T) diagram
   - $\mathrm{d}S$: is exact differential. $S_2 - S_1$ only depends on the **endpoints** not the **path**

Let us show this by comparing two paths,



Along the green path,

$$\Delta S_1 = \int \left(\frac{đQ}{T}\right)_1 = C_v \int_{T_1}^{T'} \frac{\mathrm{d}T}{T} = C_v \ln\left(\frac{T'}{T_1}\right)$$

$$\Delta S_2 = \int \left(\frac{đQ}{T}\right)_2 = C_p \int_{T'}^{T_2} \frac{\mathrm{d}T}{T} = C_p \ln\left(\frac{T_2}{T'}\right)$$

Thus the total change is given by,

$$\begin{aligned}\Delta S &= \Delta S_1 + \Delta S_2 \\ &= C_v \ln\left(\frac{T'}{T_1}\right) + C_p \ln\left(\frac{T_2}{T'}\right)\end{aligned}$$

$$
\begin{aligned}
&= C_v \ln\left(\frac{T'}{T_1}\right) + (C_v + R)\ln\left(\frac{T_2}{T'}\right) \\
&= C_v \ln\left(\frac{T_2}{T_1}\right) + R\ln\left(\frac{T_2}{T'}\right) \\
&= \cancelto{0}{C_v \ln\left(\frac{T_2}{T_1}\right)} + R\ln\left(\frac{T_2}{T'}\right) \\
&= R\ln\left(\frac{T_2}{T'}\right)
\end{aligned}
$$

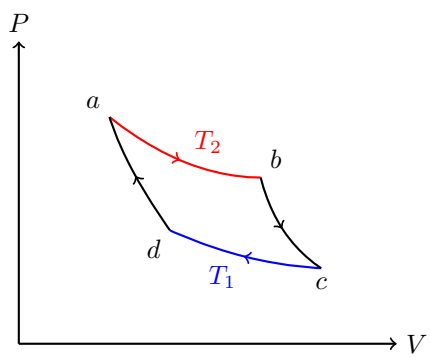
Note the cancellation $T_1 = T_2$ since (3) is an isotherm.

Since path (2) is at constant P

$$\Delta S = R\ln\left(\frac{V_2}{V_1}\right)$$

Which is identical to the change in entropy along the isotherm derived above. Thus confirming that ΔS is independent of path.

3.7.5 Clausius Inequality

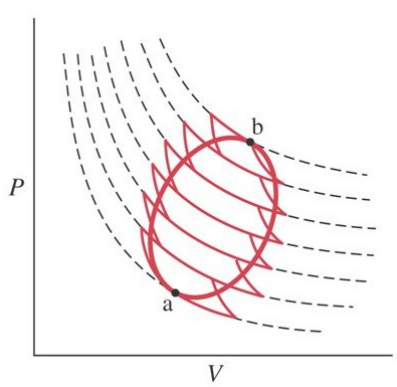


Recall from equation (3.7), we have that for a reversible Carnot cycle,

$$\frac{Q_2}{T_2} + \frac{Q_1}{T_1} = 0$$

Now consider the case of heat extraction and heat disposal along a *tiny Carnot cycle*.

$$\frac{đQ_2}{T_2} + \frac{đQ_1}{T_1} = 0$$



The idea here is to consider any physical process or cycle as being made up of numerous tiny adiabatic curves and isotherms. We *approximate* the real cycle as being composed of a series of reversible processes. In the limit of the number of "switches" goes to infinity, the real engine process is made exactly reversible.

Essentially, we are mapping out the contour of any given cycle by following many isotherms and adiabats. Then for any given sub-process here,

$$\sum_{i=1}^{n} \frac{đQ_i}{T_i} = 0$$

Or in the limit as $n \to \infty$,

$$\oint \frac{đQ_{\text{rev}}}{T} = 0$$

Again we observe that Q_{rev}/T is an **exact** differential and we can call it $\mathrm{d}S$ as done before. S is a state variable.

Any reversible cycle must obey the following equation,

$$\oint \frac{đQ}{T} = 0 \quad \text{For a reversible cycle.}$$

What about for an irreversible cycle? An irreversible cycle is some process that somewhat has a step that can not be performed backward. What we have shown is that the efficiency of an engine with an irreversible cycle η' must be less than that of Carnot η_C (see (3.6)).

$$\eta' < \eta_C$$

Using equation (3.6), we can show that,

$$\frac{Q_1'}{Q_2'} < \left(\frac{Q_1}{Q_2}\right)_{\text{rev Carnot}}$$

Or that,

$$\frac{đQ_1'}{T_1} + \frac{đQ_2'}{T_2} < 0 \quad \text{For an irreversible cycle.}$$

Which lead to the generalization for an infinite number of steps as,

$$\sum_{i=1}^{n} \frac{đQ_i'}{T_i} < 0 \to \oint \frac{đQ'}{T} < 0$$

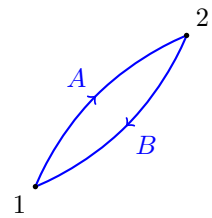
This leads to the Clausius Inequality:

$$\oint \frac{đQ'}{T} \leq 0$$

Which achieves equality only when the **entire** cycle is irreversible. Inequality is always maintained if any part of the cycle is irreversible. This can be understood as another formulation of the Second Law of Thermodynamics.

There are two cases "applications" of Clausius inequality:

1) Consider a single two-step *reversible* cycle



Therefore,

$$\oint \frac{đQ}{T} = 0$$

$$\int_1^2 \left(\frac{đQ}{T}\right)_A + \int_2^1 \left(\frac{đQ}{T}\right)_B = 0$$

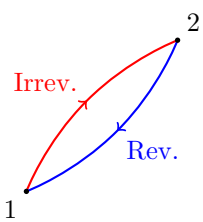
Rearrangement gives,

$$\int_1^2 \left(\frac{đQ}{T}\right)_A = -\int_2^1 \left(\frac{đQ}{T}\right)_B = \int_1^2 \left(\frac{đQ}{T}\right)_B$$

Which implies that entropy is conserved,

$$(S_2 - S_1)_A = (S_2 - S_1)_B$$

2) Consider a two-step *irreversible/reversible* process



Therefore by the Clausius Inequality we have,

$$\oint \frac{đQ}{T} \leq 0$$

$$\int_1^2 \left(\frac{đQ}{T}\right)_{\text{irrev}} + \int_2^1 \left(\frac{đQ}{T}\right)_{\text{rev}} \leq 0$$

$$\int_1^2 \left(\frac{đQ}{T}\right)_{\text{irrev}} \leq -\int_2^1 \left(\frac{đQ}{T}\right)_{\text{rev}} = \int_1^2 \left(\frac{đQ}{T}\right)_{\text{rev}}$$

Thus we can conclude that with $\mathrm{d}S \left(\frac{đQ}{T}\right)_{\text{rev}}$,

$$\int_1^2 \left(\frac{đQ}{T}\right)_{\text{irrev}} \leq \int_1^2 (\mathrm{d}S)_{\text{rev}} = S_2 - S_1$$

Which gives that our entropy increases each cycle,

$$(S_2 - S_1)_{\text{rev}} \geq \int_1^2 \left(\frac{đQ}{T}\right)_{\text{irrev}}$$

Now suppose that our system is undergoing an adiabatic process along the whole path. This means that no heat is exchanges during the process ($đQ = 0$). This gives us,

$$(S_2 - S_1)_{\text{irrev}} \geq 0 \quad \text{For an irreversible, adiabatic process.}$$

In conclusion, the entropy can only increase for an *isolated* (our adiabatic example) system undergoing an irreversible process. This is a restatement of the Second Law of Thermodynamics; entropy always increases for an isolated system. Of course this is all subject to the constraint imposed on the system: N, V, E might be constrained or imposed to be constant. Or more clearly:

At *equilibrium*, entropy must reach its maximum value. Otherwise it would increase because it has to.

3.7.6 Perspective on Entropy

Consider a container with two sub-sections. The left hand side will have a gas and the right hand side will have a vacuum (each side having volume V). At time $t = 0$, the barrier between the two containers is removed and the gas has access to the entire container of volume $2V$. At equilibrium, the gas occupies the entire container. Evidently, this is an irreversible process (the gas does not have the luxury to return to it's original state). Some comments:

- irreversible process
- gas does no work against the vacuum
- no heat exchange
- $\Delta U = \Delta Q - \Delta W = 0$ (Conservation of energy $\implies$ temperature is constant)

This process is irreversible, so there is no way to directly compute it's change in entropy throughout the entire process. However, the entropy can be calculated using properties about the two endpoints and a *reversible* process that connects these endpoints.

The *reversible* process we will consider as a replacement will be letting the gas expand via a piston held initially at the midway point of the container that is controlled externally. We will slowly move the piston until the gas fills the entire container. This new process is consider isothermal that takes use from state (1) (half gas, half vacuum) to state (2) (total gas). Along this process we can compute the change in entropy $S_2 - S_1$ along the reversible path. Turns out we already computed this in equation (3.8).

$$\Delta S = S_2 - S_1 = nR \ln\left(\frac{V_2}{V_1}\right)$$

Let $nR = k_B N_A$ where k_B is the Boltzmann constant and N_A is Avogadro's number. Note that $V_1 = V$ and $V_2 = 2V$,

$$\Delta S = k_B N_A \ln\left(\frac{2V}{V}\right) = k_B \ln\left(2^{N_A}\right)$$

Thus we expect,

$$\Delta S \propto \underbrace{(\text{const.})}_{\text{units } J/K} \times \ln\left(2^N\right)$$

Note that the '2^N' term measures the number of microstates that the system received or now has access to to due to the process. Consider the microstates counted by the possibility that the particle can be on the left or the right of the container.

- All particles on left except 1
  - LLLLL...LLRLLL...LLLLL
  - N ways for this to occur
- All particles on the left except 2
  - LLLLL...LLRLLL...LLRL...LLLLL
  - $N(N-1)/2$ ways for this to occur

The most probable macrostate occurs when the total number of microstates reaches a maximum. Using the Stirling's approximation, this occurs when there is a 50/50 split between the left and right with total number of microstates,

$$\Omega \sim 2^N + \text{small corrections}$$

Note that disordered ("mixed") states (50/50) can be realized in so many more ways than "ordered states".

With this view, entropy can be viewed as a measure of the number of microstates through the rational of the second law. These ideas are what lead Boltzmann to consider the possibility,

$$S \sim \ln \Omega$$

With further corrects and units,

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

Where the constant S_0 can be shown to be zero with the Third Law of Thermodynamics.

3.7.7 Conclusion

The Second Law of Thermodynamics is very much a manifestation of the probability/statistics of the large number of states microscopic objects can realize.

The law $\Delta S \geq 0$ spontaneously for an isolated system means that we must seek microscopic configurations that *maximize* S, as when this is achieved, the system must have reached equilibrium.

4 S.M. Basis for T.D.

Statical Mechanics Basis for Thermodynamics.

4.1 Contact Between S.M. & T.D.

1. Systems where $V \to \infty$ and $N \to \infty$ and the density N/V remains constant
   - Thermodynamically large system
2. Equilibrium Thermodynamics / S.M.
   - Consider two regions $(1, 2)$ of a container separated by a diathermal wall
   - (N_1, V_1, E_1) and (N_2, V_2, E_2) number of particles, volume, energy respectively
   - Interested in total number of microstates
   - Composite system $\Omega^{(0)} = \Omega^{(1)} \times \Omega^{(2)}$

This composite system has $E^{(0)} = E_1 + E_2 =$ constant. We are interested in Ω: the number of microstates. At *any time*, system (1) is just as likely to be in any of its microstates. This is also true for system (2).

$$\Omega^{(0)}\left(E_1, E_2\right) = \Omega_1(E_1) \times \Omega_2(E_2)$$

$$\Omega^{(0)}\left(E_1, E^{(0)} - E_1\right) = \Omega_1(E_1) \times \Omega_2(E^{(0)} - E_1)$$

The most likely macrostate is the one with the most number of microstates. It is important to note that is true at *any time*. To maximize $\Omega^{(0)}$,

$$\frac{\partial \Omega^{(0)}}{\partial E_1} = 0 \implies \left(\frac{\partial \Omega_1}{\partial E_1}\bigg|_{E_1=\bar{E}_1} \Omega_2(\bar{E}_2) + \Omega_1(\bar{E}_1)\left(\frac{\partial \Omega_2}{\partial E_2}\bigg|_{E_2=\bar{E}_2} \frac{\partial E_2}{\partial E_1} = 0\right.\right.$$

Note since $E^{(0)}$ is fixed,

$$\frac{\partial E_2}{\partial E_1} = -1$$

Which reveals,

$$\frac{1}{\Omega_1(E_1)}\left(\frac{\partial \Omega_1}{\partial E_1}\bigg|_{E_1=\bar{E}_1}\right. = \frac{1}{\Omega_2(E_2)}\left(\frac{\partial \Omega_2}{\partial E_2}\bigg|_{E_2=\bar{E}_2}\right.$$

Notice the logarithmic nature of this equation to show that,

$$\underbrace{\left(\frac{\partial \ln \Omega_1}{\partial E_1}\bigg|_{E_1=\bar{E}_1}\right.}_{\beta_1} = \underbrace{\left(\frac{\partial \ln \Omega_2}{\partial E_2}\bigg|_{E_2=\bar{E}_2}\right.}_{\beta_2}$$

Therefore in summary, equilibrium requires $\beta_1 = \beta_2$. Recall that for a reversible process,

$$\mathrm{d}U = đQ - đW$$

Thus,

$$\mathrm{d}E = đQ - đW$$

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V$$

Where the definition of entropy can be expressed,

$$\left(\frac{\partial S}{\partial N, V}\right)_E = T$$

Which gives for system 1,

$$\left(\frac{\partial E_1}{\partial \ln \Omega_1}\right)\left(\frac{\partial \ln \Omega_1}{\partial S_1}\right) = T_1$$

And the analogous expression for system 2,

$$\left(\frac{\partial E_2}{\partial \ln \Omega_2}\right)\left(\frac{\partial \ln \Omega_2}{\partial S_2}\right) = T_2$$

Combining these two systems gives,

$$\frac{\partial S_1}{\partial \ln \Omega_1} = \frac{1}{\beta_1 T_1} \qquad \frac{\partial S_2}{\partial \ln \Omega_2} = \frac{1}{\beta_2 T_2} \tag{4.1}$$

Equations (4.1) should be telling for the following assumption.

Assumption: There exists an intimate relationship between Statistical Mechanics and Thermodynamics for all and any system.
We will propose that,

$$\frac{1}{\beta_1 T_1} = \frac{1}{\beta_2 T_2} = \frac{1}{\beta T} = \text{ constant/universal}$$

Thus,

$$S = \text{constant} \cdot \ln \Omega \tag{4.2}$$

Where the unknown constant is actually Boltzmann's constant, k_B with,

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

What about the relationships between other quantities associated with systems (1) and (2)? First consider volume.

$$\underbrace{\left(\frac{\partial \ln \Omega_1}{\partial V_1}\right|_{N_1, E_1, V_1 = \bar{V}_1}}_{\eta_1} = \underbrace{\left(\frac{\partial \ln \Omega_2}{\partial V_2}\right|_{N_2, E_2, V_2 = \bar{V}_2}}_{\eta_2}$$

Which gives,

$$\eta_1 = \eta_2$$

These new variables describe the equilibrium of the two systems in terms of volume. Is it required to introduce a new constant similar to above to relate volume to entropy? Consider,

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V \tag{4.3}$$

Thus,

$$\frac{\partial S}{\partial E} = \frac{1}{T} \qquad \left(\frac{\partial S}{\partial V}\right)_E = \frac{P}{T}$$

However, we could introduce another term to (4.3) to include flux of particles; a chemical potential term. We could add more terms,

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V + \mu\mathrm{d}N - H\mathrm{d}M$$

Which would produce more relations like,

$$\frac{\partial S}{\partial N} = -\frac{\mu}{T}$$

From this we get,

$$\frac{\partial \ln \Omega}{\partial V} = \eta$$

Which by chain rule,

$$\frac{\partial \ln \Omega}{\partial S}\frac{\partial S}{\partial V} = \eta$$

But we know from (4.1) that $\frac{\partial \ln \Omega}{\partial S} = \beta T$. Therefore,

$$\frac{\partial S}{\partial V} = \frac{1}{\beta T}\eta = \frac{P}{T}$$

Therefore rearrangement gives,

$$\eta = \frac{P}{k_B T}$$

In summary, once we have the result (4.2), with the universal constant k_B, all expected equilibrium between generalized forces must be satisfied. No new constants or relations need to be constructed to explore the relationship between entropy and microscopic quantities.

An equilibrium of energy and volume between systems (1) and (2) it must be that,

$$T_1 = T_2 \qquad P_1 = P_2$$

Furthermore, and equilibrium of energy and number of particles it must be that,

$$T_1 = T_2 \qquad \mu_1 = \mu_2$$

Remark: Equilibrium properties in the extensive (E, V, N) variables of the system translate in an equally in the intensive (T, P, μ) variables.

In summary, the equation $S = k_B \ln \Omega$ recovers all expected descriptions of thermodynamical equilibrium.

4.2 Classical Ideal Gas from S.M.

For this analysis, we will begin with the consideration at $\hbar = 0$. In other words, we will consider the particles to point like and ignore their wave like nature and treat interactions between particles $(V(r_{ij}))$ is very small. Note however, in physics, one cannot claim something is "small" unless the quantity is dimensionless. Everything is small compared to something else. Thus when we say $\hbar$ is small, we mean that the DeBrogile wavelength is small compared to the thermal length scale. Similarly for the potential energies between the particles, $V/k_BT \ll 1$.

4.2.1 Volume

Consider the volume the gas occupies as a fixed volume V. Also notice that it is acceptable at this point to say that the number of microstates is proportional to the volume.

$$\Omega \propto V^N$$

Where N is the number of particles. But how can be compare microstates which are dimensionless to the volume which could have dimensions m^3? In practice we assume,

$$\Omega = \text{ constant } \times \left(\frac{V}{\Lambda^3}\right)^N$$

Where Λ is a length scale of the system. We will return to this idea later when considering quantum gases. Note from above, we have,

$$\left(\frac{\partial S}{\partial V}\right)_E = \frac{P}{T}$$

Combined with $S = k_B \ln \Omega$ and $\Omega = cV^N$,

$$PV = Nk_BT$$

This is quite remarkable. We just derived an expression for an ideal gas using $S = k_B \ln \Omega$ and some simple motivations for the form of Ω. Typically, this ideal gas law is written with,

$$Nk_B = N_A R$$

Where N_A is Avogadro's number 6.02×10^{23} and R was known that $R = 8.31\,\mathrm{J\,mol^{-1}\,K^{-1}}$ which gives a value for,

$$k_B = 1.38 \times 10^{-23}\,\mathrm{J\,K^{-1}}$$

4.2.2 Energy & Particle in a Box

As a reminder, an ideal gas has no interaction $V = 0$ and modeling the particles as particles in a box. As such, they are fundamentally quantum mechanical in nature. For a particle in a box, recall from Q.M. that for a potential,

$$U(x) = \begin{cases} 0 & 0 \leq x \leq L \\ \infty & \text{otherwise} \end{cases}$$

The energy is given by,

$$H = \frac{p^2}{2m}$$

Where,

$$\vec{p} = \frac{\hbar}{i}\vec{\nabla} \qquad H = -\frac{\hbar^2}{2m}\nabla^2$$

The Time Independent Schrödinger Equation (TISE) gives,

$$H\psi = \varepsilon\psi$$

We can use this to find the energy ε,

$$-\frac{\hbar^2}{2m}\nabla^2\psi = \varepsilon\psi$$

Which yields generic solution,

$$\psi = A\sin(kx)$$

Where $kL = n\pi$ and $A = \sqrt{2/L}$ and the energy in the n-th state is given by,

$$\varepsilon = \frac{\hbar^2}{2mL^2}\pi^2 n^2 \quad \text{For 1D box}$$

However, the generalization for 3d space $\mathbb{R}^3$, the energy is determined by three quantum numbers (n_x, n_y, n_z) all strictly positive integers $\mathbb{Z}_{>0}$,

$$\varepsilon = \frac{\hbar^2}{2mL^2}\pi^2\left\{n_x^2 + n_y^2 + n_z^2\right\}$$

This energy ε is characterized by 3 independent integers -for a single particle. For N particles in this box $V = L^3 \subset \mathbb{R}^3$ we will require $3N$ independent numbers. Classically, the Lagrangian of the system is,

$$L = T - V$$

Where T is the kinetic energy and V is the potential energy. The kinetic energy is given by,

$$T = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}$$

Evidently for a single particle, n cannot be zero because then that would yield the trivial solution $\psi(x) = 0$. So why is it that for the 3d example, it is still required that $n_x, n_y, n_z > 0$ and each cannot equal zero? It is because the solution is solved using separation of variables. In any of the components of the solution are the trivial solution $n_i = 0$ for a particular i, the complete solution becomes the trivial solution.

Therefore using this three number n_x, n_y, n_z, we can describe the energy of a given particle for all time. However, for a fixed ε that can be measured, how many distinct set of n_i's can produce this measured energy ε?

$$\frac{\hbar^2}{2mL^2}\pi^2\left\{n_x^2 + n_y^2 + n_z^2\right\} = \varepsilon$$

Since the n_i's are integers, solving this problem for a fixed ε is the same as solving the Diophantine equation,

$$x^2 + y^2 + z^2 = R^2 \tag{4.4}$$

Which geometrically is equivalent to asking the question:

Where does the surface of the sphere intersect with a 3d grid of unit spacing?

Therefore, the total number of microstates for a fixed energy ε is the same as determining the number of solutions to (4.4). Let us call,

$$\varepsilon^* = \frac{2m}{\hbar^2\pi^2} V^{2/3} \varepsilon$$

Where ε^* acts as a dimensionless energy that characterizes the problem ($R^2 = \varepsilon^*$).

$$\Omega = \# \text{ of solutions to (4.4)}$$

In general (like all Diophantine equations), this problem is very hard. It is known as the **microcanonical ensemble** problem.

4.3 Microcanonical Ensemble

The Microcanonical ensemble problem involves finding positive integers n_x, n_y, n_z such that,

$$n_x^2 + n_y^2 + n_z^2 = \frac{2mL^2}{\hbar^2\pi^2}\varepsilon = \frac{2mV^{2/3}}{\hbar^2\pi^2}\varepsilon \equiv \varepsilon^*$$

Classically the total energy energy of the system is given by the sum of the kinetic energies of each of the particles.

$$E = K_1 + K_2 + K_3 + \cdots + K_N$$

Which when broken into each component,

$$E = \left(\frac{p_{1x}^2 + p_{1y}^2 + p_{1z}^2}{2m}\right) + \left(\frac{p_{2x}^2 + p_{2y}^2 + p_{2z}^2}{2m}\right) + \cdots + \left(\frac{p_{Nx}^2 + p_{Ny}^2 + p_{Nz}^2}{2m}\right)$$

For fixed E in general, this constitutes a $3N$ dimensional problem,

$$\sum_{i=1}^{N} \frac{\hbar^2\pi^2}{2mV^{2/3}} \left(n_{ix}^2 + n_{iy}^2 + n_{iz}^2\right) = E$$

Which can be expressed with shorthand notation as,

$$\sum_{r=1}^{3N} \frac{\hbar^2\pi^2}{2mV^{2/3}} n_r^2 = E$$

Where n_r is the r-th degree of freedom of the system. This can be express even more compactly by absorbing constants,

$$\sum_{r=1}^{3N} n_r^2 = E^* \equiv E\frac{2mV^{2/3}}{\hbar^2\pi^2} \tag{4.5}$$

Notice that the term $EV^{2/3}$ completely characterizes the number of microstates Ω as it determines the number of solutions of (4.5). What we notice that without calculating Ω, we can determine the form of entropy.

$$S = S\left(N, V, E\right) = S\left(N, EV^{2/3}\right)$$

For a system undergoing a reversible adiabatic change, we have both $đQ = 0$ which implies that $\mathrm{d}S = 0$ for fixed temperature which implies that entropy is constant. Therefore for a reversible adiabatic process, in order to ensure that entropy is constant, it must be that,

$$V^{2/3} \cdot E = \text{const.} \quad \text{For reversible/adiabatic process}$$

Now using the known expression,

$$P = -\left(\frac{\partial E}{\partial V}\right)_{N,S}$$

And using c as a constant,

$$E = cV^{-2/3}$$

Which gives,

$$PV^{5/3} = \text{constant} \quad \text{For reversible/adiabatic process.}$$

Or more generally, for quantum/classical systems,

$$PV^{\gamma} = \text{constant} \quad \gamma = 5/3 \text{ for monoatomic system} \tag{4.6}$$

Note that (4.6) does not hold for relativistic systems as $E = pc$ where c is the speed of light.

4.3.1 Solving the Microcanonical Ensemble

One will notice that as E^* increases, the number of solution Ω varies wildly. It can be precisely 0 and some values for E^* and very many for other, nearby energies. Instead of calculating only the n_r for which E is precisely given, we will instead compute all the states for which $E' < E$ to simplify the problem. In principle, we want to replace the problem of finding,

$$\Omega(N, V, E)$$

We will instead compute (for 1 particle),

$$\sum_{'}(\varepsilon') = \sum_{\varepsilon' \leq \varepsilon} \Omega(1, V, \varepsilon')$$

Where,

$$\sum_{n_r'}^{3} n_r^2 = \varepsilon^*$$

Which can be converted to a continuous problem of calculating the volume of this sphere,

$$\int_V \mathrm{d}V$$

Which is a problem of finding the volume of a sphere of radius $\sqrt{\varepsilon^*}$. In this case of 1 particle, V has dimension 3. Therefore,

$$V(R) = \frac{4\pi R^3}{3}$$
$$V(\varepsilon^*) = \frac{4\pi}{3}\left(\varepsilon^*\right)^{3/2}$$

Now since we only care about the cases where n_x, n_y, n_z are positive, this quantity *overestimates* by $2^3 = 8$ times the desired amount. Therefore the sum over all states where $0 \leq \varepsilon' \leq \varepsilon^*$ is

$$\sum_{'} = \frac{1}{8}\frac{4\pi}{3}\left(e^*\right)^{3/2} = \frac{\pi}{6}\left(e^*\right)^{3/2}$$

Now we must repeat this calculation for not a 3d sphere but a 3Nd-sphere.

Mathematical Interlude:
In 3-d, a volume element is given by, $\mathrm{d}V = \mathrm{d}x\mathrm{d}y\mathrm{d}z \equiv \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$.

In n-d, a volume element is given by,

$$\mathrm{d}V_n = \prod_{i=1}^{n}\mathrm{d}x_i$$

Therefore in n-d the volume of a sphere is given by,

$$V_A = \int \mathrm{d}^n r = \int \cdots \int \mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n = \int \cdots \int \prod_{i=1}^{n}\mathrm{d}x_i$$

And performing this sum under the constraint that,

$$0 \leq \sum_{i=1}^{n} x_i^2 \leq R^2$$

Dimensions	Volume of Ball
2	πR^2
3	$\frac{4}{3}\pi R^3$
$\cdots$	$\cdots$
n	$c_n R^n$

Therefore, we need to find c_n in general to get the volume of a hypersphere. Since $V_n = c_n R^n$, consider

$$\mathrm{d}V_n = \mathrm{d}\left(c_n R^n\right) = \underbrace{c_n \cdot n \cdot R^{n-1}}_{S_n}\mathrm{d}R$$

Trick: We will take a function that we like for which we can do exactly the integrals in $\mathrm{d}x_i$ (i.e. Cartesian Coordinates) and perform the integral in spherical coordinates.

$$\int f(\{x_i\})\prod_{i=1}^{n}\mathrm{d}x_i = \int f(R)S_n\mathrm{d}R = \int f(R)nc_nR^{n-1}\mathrm{d}R$$

Now make use of Gaussian integrals used in section 2.4.3,

$$f(x) = e^{-x^2}$$

$$f(\{x_i\}) = \prod_{i=1}^{n} e^{-x_i^2}$$

Gives,

$$\int_{-\infty}^{\infty} e^{-x^2}\mathrm{d}x = \sqrt{\pi}$$

Which for generic n with separating the terms x_i gives,

$$\prod_{i=1}^{n}\int_{-\infty}^{\infty} e^{-x_i^2}\mathrm{d}x_i = \int_{-\infty}^{\infty} e^{-x_1^2}\mathrm{d}x_1 \int_{-\infty}^{\infty} e^{-x_2^2}\mathrm{d}x_2 \cdots \int_{-\infty}^{\infty} e^{-x_n^2}\mathrm{d}x_n = \int_{0}^{\infty} f(R)nc_nR^{n-1}\mathrm{d}R$$

The RHS has n terms that each evaluate to $\sqrt{\pi}$,

$$\pi^{\frac{n}{2}} = \int_0^\infty f(R) n c_n R^{n-1} \mathrm{d}R = n c_n \int_0^\infty f(R) R^{n-1} \mathrm{d}R$$

Since $f(R)$ is just a Gaussian in R,

$$\pi^{\frac{n}{2}} = n c_n \int_0^\infty e^{-(\sum_{i=1}^n x_i^2)} R^{n-1} \mathrm{d}R = n c_n \int_0^\infty e^{-R^2} R^{n-1} \mathrm{d}R$$

Using the identity,

$$\int_0^\infty e^{-\alpha y^2} y^\nu \mathrm{d}y \equiv \frac{1}{2\alpha^{\frac{\nu+1}{2}}} \Gamma\left(\frac{\nu+1}{2}\right) \quad \text{For } \nu > -1$$

Where $\Gamma(x)$ is the *Gamma function* which is equivalent to factorials for x integers, we get,

$$\pi^{\frac{n}{2}} = n c_n \frac{1}{2} \Gamma\left(\frac{n}{2}\right)$$

Thus,

$$c_n = \frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!}$$

Therefore,

$$V_n(R) = \frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!} R^n \tag{4.7}$$

Therefore for $n = 3N$ and $R = \sqrt{E^*}$ the total number of states contained inside the hypersphere with radius R is given by,

$$\sum\nolimits_N (N, V, E) = \left(\frac{V}{h^3}\right)^N \frac{(2\pi m E)^{3N/2}}{\left(\frac{3N}{2}\right)!}$$

Where $\sum_N$ doesn't explicitly represent a *sum*. Instead it is the total number of microstates within the hypersphere. It is a *function* of N, V, E. Note this includes the factor of $\frac{1}{2}^{3N}$ to only count the positive integer solutions.

What we want to do is to derive the thermodynamic properties for E. In reality however, no system is perfectly isolated from surroundings. In practice E is not exactly defined. We will construct the idea of Δ or uncertainty in E is only because of the imperfect nature of the system. Assume that the energy is given by $E \pm \frac{\Delta}{2}$. We will also assume that $\Delta \ll E$. Since initially we wanted to know $\Omega(N, V, E)$ we really want to know the the number of states associated with a measure of the energy within the uncertainty. This will be the difference of two $\sum_N$ values.

$$\Gamma_{N,V,\bar{E}=E} = \sum\nolimits_N \left(N, V, E + \frac{\Delta}{2}\right) - \sum\nolimits_N \left(N, V, E - \frac{\Delta}{2}\right) \tag{4.8}$$

Now in the limit that $\Delta/E \to 0$,

$$\lim_{\Delta/E \to 0} \Gamma_{N,V,\bar{E}=E} = \Gamma_{N,V\bar{E}=E} = \lim_{\Delta/E \to 0} \left[\sum\nolimits_N \left(N, V, E + \frac{\Delta}{2}\right) - \sum\nolimits_N \left(N, V, E - \frac{\Delta}{2}\right)\right] = \frac{\partial \sum_N}{\partial E} \cdot \Delta$$

Therefore the total number of states associated with $E = \bar{E}$ is given by,

$$\Gamma_{N,V,\bar{E}=E} = \frac{\partial \sum_N}{\partial E} \cdot \Delta = \frac{3N}{2}\frac{\Delta}{E}\sum_N (N, V, E = \bar{E}) \tag{4.9}$$

Using the Boltzmann equation for entropy and Stirling's approximation with $N \gg 1$ (see section 2.4.2),

$$S = k_B \ln \Gamma_N = k_B \left[N \ln \left\{ \frac{V}{h^3} \left(\frac{4\pi m E}{3N} \right)^{3/2} \right\} + \frac{3N}{2} + \left\{ \ln \left(\frac{3N}{2} \right) + \ln \left(\frac{\Delta}{E} \right) \right\} \right]$$

This gives us the entropy of an ideal gas at energy E within a window $E \pm \Delta/2$.

What thermodynamic properties does this expose?

$$S(N, V, E) = k_B \ln \Gamma_N \approx k_B \left[N \ln \left\{ \frac{V}{h^3} \left(\frac{4\pi m E}{3N} \right)^{3/2} \right\} + \frac{3N}{2} \right] \tag{4.10}$$

Notice that this depends on $\ln(N)$ and also $N^{-3/2}$ which seems to indicate that theses are not extensive variables. We will return to this problem soon. Rearranging for energy,

$$E(S, V, N) = \frac{2h^2 N}{4\pi m V^{2/3}} \exp\left(\frac{2S}{3Nk_B} - 1 \right)$$

Thermodynamically, we have,

$$T = \left(\frac{\partial E}{\partial S} \right)_{V,N}$$

Which reveals,

$$E = N \left(\frac{3}{2} k_B T \right)$$

And the heat capacity at constant volume is given by,

$$c_V = \left(\frac{\partial E}{\partial T} \right)_{N,V} = \frac{3}{2} N k_B = \frac{3}{2} nR$$

And analogously,

$$c_P = T \left(\frac{\partial S}{\partial T} \right)_{N,P} = \frac{5}{3} N k_B$$

Furthermore, thermodynamic pressure is given by,

$$P = -\left(\frac{\partial E}{\partial V} \right)_{N,S} = \frac{2}{3}\frac{E}{V}$$

Combining these two results, we can see that,

$$PV = nRT$$

Which means we have recovered all known results for ideal gases from only the *explicit* form of $\sum_N$ (or Γ_N).

Consider an isothermal change of entropy of entropy, and we recover,

$$S_2 - S_1 = N k_B \ln \left(\frac{V_2}{V_1} \right)$$

Or for a reversible adiabatic change/process,

$$E, T \sim V^{-2/3}$$

$$PV^{\gamma} = \text{const.}$$
$$TV^{\gamma-1} = \text{const.}$$

We also recover that,

$$\mathrm{d}E = -p\mathrm{d}V = -\frac{2}{3}\frac{E}{V}\mathrm{d}V$$

In conclusion we have discovered that measuring and counting microstates $\Omega_N \approx \sum_N, \Gamma_N$ we can then use $S = k_B \ln \Omega$ to recover and describe thermodynamic properties. All that remains is to discuss the problem that V and N do *not* appear to be extensive variables in equation (4.10). This problem is known as **Gibb's Paradox**.

4.4 Gibb's Paradox

How can we fix or understand the origin of the lack of extensivity in the formula $S = k_B \ln \Omega$? We should expect S to scale with V but in (4.10) it does not. Does this mean that (4.10) is somehow wrong?

To expose this paradox, consider a container segmented by a wall into 2 regions with properties N_1, V_1, T_1 and N_2, V_2, T_2 with total volume $V = V_1 + V_2$. Now measure the entropy using (4.10) for each system S_1, S_2 and then adiabatically remove the wall between the segments of the container and recalculate entropy S_T of the combined system. Does $S_T = S_1 + S_2$?
Using $E = \frac{3}{2}Nk_BT$

$$S_i = N_i k_B \ln V_i + \frac{3}{2}N_i k_B \left\{1 + \ln\left(\frac{2\pi m_i k_B T}{h^2}\right)\right\} \quad i = 1, 2$$

While,

$$S_T = \sum_{i=1}^{2} N_i k_B \ln V$$

Consider the case where the particle densities are the same and thus is also the same for the mixture after removing the wall.

$$\frac{N_1}{V_1} = \frac{N_2}{V_2} = \frac{N_1 + N_2}{V_1 + V_2}$$

Therefore the change in entropy is given by,

$$\Delta S = S_T - (S_1 + S_2)$$

Which is explicitly,

$$\Delta S = k_B \left\{N_1 \ln\left(\frac{N_1 + N_1}{N_1}\right) + N_2 \ln\left(\frac{N_1 + N_1}{N_2}\right)\right\} \tag{4.11}$$

Now lets assume $m_1 = m_2$ or that we have identical particles. From this, the macrostate of the system looks like it doesn't change when the wall is removed by we still have by equation (4.11) that,

$$\Delta S > 0$$

This makes no sense because if the particles are the same, we have not *mixed* anything.

The resolution is to write (using Stirling's formula),

$$\Delta S = k_B \left\{\ln\left[(N_1 + N_2)!\right] - \ln\left[N_1!\right] - \ln\left[N_2!\right]\right\}$$

To fix this problem we need to reduce S or Ω by a factor of $N!$.

$$S \text{ reduced by } k_B \ln N!$$
$$S_1 \text{ reduced by } k_B \ln N_1!$$
$$S_2 \text{ reduced by } k_B \ln N_2!$$

This is an *ad hoc* fix. It does not have any profound explanation of justification. The mechanism that describes this, is the notion of indistinguishably. The notion of indistinguishably is needed to be taken into account in order to properly count the number of microstates of a system. Proper analysis of this idea requires the theory of Quantum Statistics. Now we can go back and determine the proper formula for (4.10).

$$S(N,V,E) = Nk_B \ln\left(\frac{V}{N}\right) + \frac{3Nk_B}{2}\left\{\frac{5}{3} + \ln\left(\frac{2\pi m k_B T}{h^2}\right)\right\} \tag{4.12}$$

This is the formula for the entropy of a non-degenerate ideal gas. This is known as the **Sackur-Tetrode equation**. One should note that this equation is quite remarkable. It is composed of only N, V, E and a set of constants. This equation gives a proper extensivity of S.

For more, look at the theory of super-solid helium.

5 Ensemble Theory

Let's step back and review some of the statements we have made. Here are come general facts:

- Every microstate has equal probability of being selected
- As time passes, the system will explore all of the microstates
- Experimental measurements are made of the time averaged of the true, non-stationary value of physical quantities f
  - $f = E, \text{pH}, \text{etc.}$
  - For example, measuring the pH level of a chemical solution will be the average measure over a macroscopic duration of time
  - For systems where the dynamics of the system have time scale much shorter than the measurement frequency, the measured quantity ends of being an averaged value
- Making a "movie" of a physical system (many measurements) and replacing the time average by a "still frame" average
  - $[f]_{\text{time}} = \frac{1}{T}\int_{t_0}^{t_0+T} f(t)\mathrm{d}t$: the square brackets with subscript time indicate an average measure over time
  - $\langle f \rangle_{\text{exp}} = [f]_{\text{time}}$: Natural one should expect the time averaged value to be measure experimentally
  - In doing thus, there must be some sort of averaging of the true real-time value using "snapshots" of f

The question becomes, what are the options for performing this "snapshot" averaging? One could make periodic snapshots at fixed frequency or random intervals or combinations there of. It is the spirit of ensemble theory to determine how to construct an adequate way to compute $\langle f \rangle_{\text{exp}}$ withing explicitly doing $\int f(t)\mathrm{d}t$. Some options include,

- microcanonical ensemble
- canonical ensemble
- grand ensemble

The procedure introduced by Gibb's involves utilizing phase space and classical systems.

Remark by TC Fraser: This is intimately connected to the *Erdős discrepancy problem.* (I think?)

5.1 Phase Space of Classical System

5.1.1 Dynamics & Phase Space

Imagine we have a system of N *distinguishable* particles,

$$i = 1, 2, 3, \ldots, N$$

that lives in the space of 3 dimensions $\mathbb{R}^3$. We will fix for the system N, V, E. This is a microcanonical system/construction. We will also consider for the system that the particles are localized (they are were they are) and they do move (they have dynamics). The position of a particle μ is given by,

$$\vec{r}_\mu = x\hat{x} + y\hat{y} + x\hat{z}$$

And each particle μ has same mass m. The momentum of particle μ is given by,

$$\vec{p}_\mu = m\dot{\vec{r}}_\mu$$

To index the entire system of N particles, we will consider this index map,

$$\vec{r}_\mu = r_{\alpha,\mu} \mapsto r_i \quad \alpha = 1, 2, 3 \quad \mu = 1, 2, \ldots, N \quad i = 1, 2, \ldots, 3N$$

Similarly for $\vec{p}_\mu$

$$\vec{p}_\mu = p_{\alpha,\mu} \mapsto p_i \quad \alpha = 1, 2, 3 \quad \mu = 1, 2, \ldots, N \quad i = 1, 2, \ldots, 3N$$

The product of the two spaces $\{r_i\} \times \{p_i\}$ is a $6N$ dimensional space known as the **phase space** of the system. To determine the dynamics of the system, we need to solve the following problem.

$$\mathcal{L} = K\left(\{\dot{q}_i\}\right) - V\left(\{q_i, \dot{q}_i\}\right)$$

Where $\mathcal{L}$ is the **Lagrangian**, K is the kinetic energy that depends on velocities, and V is the potential energy which depends on position and sometimes velocities. Will will make the switch to the *Hamiltonian* formalism of classical mechanics through a **Legendre Transform**,

$$H = \sum_i^{3N} \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} = \sum_i^{3N} \dot{q}_i p_i - \mathcal{L}$$

Where $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ is the generalized momentum. In principle, all classical problems can be described under this formalism. The "bottle neck" of this formalism is typically solving the equations of motion,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad \dot{q}_i = \frac{\partial H}{\partial p_i} \tag{5.1}$$

For generic physical quantity $f(q, p; t)$ that depends on time, positions an momentum of all the particles, all that would be required is to solve the equations of motion, and predictions could be made. However, this is very *difficult*. Ensemble Theory allows for the construction of proper approximations for generic f.

μ-space vs. Γ-space:
We will now introduce the notion of μ-space. In this $6N$ dimensional space, we assign to each particle μ a single point,

$$\vec{r}_\mu, \vec{p}_\mu \mapsto (r_x, r_y, r_z, p_x, p_y, p_z)$$

For N particles, there are N points in the μ-space which are characterized by $6N$ numbers. The alternative description is phase space (Γ-space). Points in phase space are representative points on the *entire* system at once,

$$\text{state vector} = (q_1, q_2, \ldots, q_{3N}; p_1, p_2, \ldots, p_{3N})$$

At a given time t, there is 1 point in phase space that is $6N$ dimensional. We will subscribe to the notation,

$$(q_i(t), p_i(t)) = (q, p)$$

5.1.2 Trajectories in Phase Space

How can be discuss or describe the time evolution of the system using phase space? The state vector of the system (q, p) changes over time and this describes how each and every particles position and momentum changes with time.

As a function of time, the representative point traces a trajectory which **never** intersects itself. This is because the time evolution of the system is governed by the Hamiltonian H and the solutions to the equations of motion are unique for given initial conditions. If there were intersection, take the intersection point as the initial condition and there would be two potential trajectories available for the system. Note however, that the system can loop back on itself and form cyclic trajectories.

Furthermore because of the fixed energy E condition, there is a loss of 1 degree of freedom. Thus the trajectory is constrained to a $(6N-1)$ dimensional surface or hypersurface. Similar to the microcanonical ensemble problem solved in section 4.3.1 (specifically equation (4.8)), we will treat this hypersurface or hypershell as a volume between the surfaces prescribed by the two energies,

$$E + \frac{\Delta}{2}, E - \frac{\Delta}{2}$$

In general, the representative points (q, p) are bounded. The positions q are bounded by volume and the momentums p by total energy through $K = \frac{p^2}{2m}$.

Remark: One refers to classical solutions for "real particles", as opposed to *virtual particles* (*off shell*), as being *on shell*, meaning they are found *exactly* on the hypersurface defined by fixed energy E.

5.2 Essential Replacement for Ensemble Theory

So what do we mean when discussing "ensembles"? Well one thing to notice is there a issue of difficulty with trying to solve the equations of motion *explicitly*. We would like to instead compute a notion of average,

$$\langle f(q,p)\rangle_{\text{exp}} = \frac{1}{T}\int_{t_0}^{t_0+T} f\left(q(t), p(t)\right) \mathrm{d}t$$

Where T is a duration that is "long enough". The equations of motion for the system are given by Hamiltonian mechanics (5.1),

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

Solving these $3N + 3N = 6N$ equations is done explicitly in molecular dynamics and in the study of galaxy formation by this is in general very difficult. We would like to replace this with something more tractable.

For a given macrostate (N, V, E), we have $6N$ microscopic variables. In this sense the macroscopic description is *incomplete*. We would like to take advantage of this over abundance of "micro-information" and average/integrate it out. What this means is that we need some kind of density or probability description of the microstates.

Another thing to consider is we don't really have E totally precise. Instead we can examine the hypershell of energy. Recall our previous description of a hypershell,

$$\bar{E} - \frac{\Delta}{2} \leq E \leq \bar{E} + \frac{\Delta}{2}$$

5.2.1 Ergodic Hypothesis

With the Ergodic Hypothesis comes both good news and bad news.

Good news	Bad news
It cannot be proven	It cannot be proven

Statement: Given long enough time the representative point of an isolated system will come *arbitrarily close* to any given point on the energy hypershell/phase space. Equivalently, given enough time, the system will explore all possible locations on the $6N - 1$ hypersphere (this is somewhat justified as the trajectories cannot intersect themselves).

Comments & Examples:
For the system of a simple harmonic oscillator,

$$E = \frac{p_1^2}{2} + \frac{q_1^2}{2}$$

The Ergodic hypothesis is satisfied. It explores all points on the ellipse in phase space.

For a system of springs coupled,

$$E = \frac{p_1^2}{2} + \frac{p_2^2}{2} + \frac{q_1^2}{2} + \frac{q_2^2}{2} + q_1 q_2$$

It can be shown that finite regions of phase space are not visited by solution trajectories. In general, the Ergodic Hypothesis is far from trivial or obvious but it is believed to apply in many, many cases. Note that a famous example of the the potential failing of the Ergodic Hypothesis is the case of glass panes, clays, or gels. We don't yet have a description of how these systems behave microscopically.

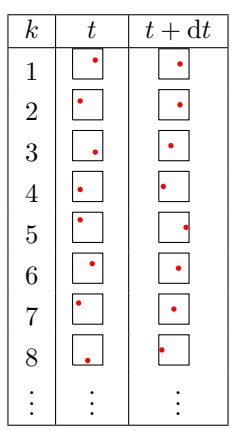
5.2.2 Mixing Hypothesis

Note that for the 3-body problem, we have chaotic solutions. This is also the case for $N > 3$. If we consider a system with many, many degrees of freedom ($N \gg 1$), we should expect that there are many degrees of freedom that will tend to be chaotic and explore all sorts of potential values on-shell. This can be taken as motivation for the Ergodic Hypothesis.

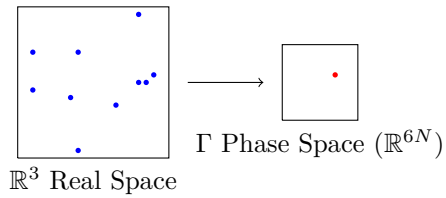
This explosion in time of an initial distribution of representative points suggests we should focus on a notion of density of representative points in phase space. This density will be time-dependent.

5.2.3 Statistical Ensemble

The idea of statistical ensembles is to make *mental copies* of the system of interest. What this means is to ascribe a representative point to each member of the ensemble. Therefore each member of the ensemble will track a representative point. In doing so, we can *replace* the time-averaging of the *real* system by performing an averaging of the long time results of many duplicated and repeated experiment.



Each of these k tracks the location of the representative point over time. These boxes can be thought of "trackers" or "members" of the phase space. They act as a single vector that represents the entire system.



We need an object/quantity that gives us a measure of the distribution representative points.

$$\rho\left(\underbrace{p(t), q(t)}_{6N \text{ variables}}\,; t\right)$$

This is density of representative points per unit of phase space volume $\mathrm{d}^{3N}q\mathrm{d}^{3N}p$. Consider this quantity,

$$\left\{\rho\left(q, p; t\right)\mathrm{d}^{3N}q\mathrm{d}^{3N}p\right\}$$

This gives us the number of points around the $6N$ dimensional point (q, p) in phase space. $\rho(q, p; t)$ symbolizes the way that the members of the ensemble (of representative points) are distributed overall possible microstates (q_i, p_i) at different instants in time.

$$\langle f\rangle_{\text{ensemble}} \equiv \frac{\int f(q, p; t)\rho(q, p; t)\mathrm{d}^{3N}q\mathrm{d}^{3N}p}{\int \rho(q, p; t)\mathrm{d}^{3N}q\mathrm{d}^{3N}p} \tag{5.2}$$

The idea here is to replace the time averaging associated with the real experiment $\langle f\rangle_{\text{exp}}$ to the ensemble average characterized by the ensemble member weight ρ.

$$\langle f\rangle_{\text{exp}} = \langle f\rangle_{\text{ensemble}}$$

This ρ acts a weight for all possible k's that could potentially make up the solution. The goal is to compute this $\rho(q, p; t)$. To do this, we will make use of the fact that ρ is constrained by the equations of motion. We must invoke the time dependence of q, p. Using (5.1),

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

Gives us a corresponding description of $\rho\left(q, p; t\right)$.

Important Remark: If we are concerning ourselves with statical mechanics, we are limiting ourselves the constraint that,

$$\frac{\partial \rho}{\partial t} = 0$$

We have to consider how $\rho(p, q)$ is compatible with the equations of motion. The solution to this is known as **Liouville's Theorem**.

5.2.4 Liouville's Theorem

What are the options or possible forms for $\rho(p, q)$? First remember that each member, namely , of the ensemble *drives* the time evolution of its own representative point. As time marches on, the number of representative points that contribute to the whole set of members of the ensemble evolves. In fact, there is an effective flow of members in the ensemble moving in and out of the shell of solutions.

$$\# : \rho\left(p, q\right)\mathrm{d}^{3N}q\mathrm{d}^{3N}p$$

Thus this integral over this quantity varies in time,

$$\frac{\partial}{\partial t}\int \rho\left(p,q\right)\mathrm{d}^{3N}q\mathrm{d}^{3N}p \tag{5.3}$$

The notion of flow (in/out of membership) describes the flow (in/out) of the phase space. With this notion of flow, there is a conservation law on the representative points; or a flow of representative point density. By analogy, this can be considered an effective "current".

$$\int \vec{J}_{\text{R.P.}}\cdot\hat{n}\mathrm{d}\sigma = \int \rho\left(p,q\right)\vec{v}\cdot\hat{n}\mathrm{d}\sigma \tag{5.4}$$

Where $\vec{v}$ is a velocity of the rate of displacement of representative points.

$$\vec{v} = \left(\dot{q},\dot{p}\right)$$

Here $\vec{v}$ is not a *physical velocity* in the sense of particles moving. It is constructed velocity associated with the time derivative of a representative point. The $\dot{q}$ components are interpreted as the *actual* velocity of each of the particles, and the $\dot{p}$ terms are the *actual* forces acting on a given particle. Instead this vector $\vec{v}$ is a flux of representative points through the surface of the solution shell. The conservation implies that (5.4) balances (5.3).

$$\frac{\partial}{\partial t}\int\limits_{\omega} \rho\left(p,q\right)\mathrm{d}^{3N}q\mathrm{d}^{3N}p = -\int\limits_{\sigma} \rho\left(p,q\right)\vec{v}\cdot\hat{n}\mathrm{d}\sigma$$

Notice the negative sign here; the LHS is positive whenever the density inside ω is **increasing**, indicating the corresponding flux should be negative. Where ω is the phase space volume, and $\sigma = \partial\omega$ is the surface bounding ω. Recognize the use of divergence theorem,

$$\int\limits_{\sigma} \rho\left(p,q\right)\vec{v}\cdot\hat{n}\mathrm{d}\sigma = \int\limits_{\omega} \vec{\nabla}\cdot\left(\rho\vec{v}\right)\mathrm{d}\omega$$

Where the volume element is given by $\mathrm{d}\omega = \mathrm{d}^{3N}q\mathrm{d}^{3N}p$. Combining these results,

$$\frac{\partial}{\partial t}\int\limits_{\omega} \rho\left(p,q\right)\mathrm{d}\omega + \int\limits_{\omega} \vec{\nabla}\cdot\left(\rho\vec{v}\right)\mathrm{d}\omega = 0$$

Since ω is presumed to have no explicit time dependence,

$$\int\limits_{\omega} \frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\left(\rho\vec{v}\right)\mathrm{d}\omega = 0$$

Which by the Reymond-Dubois Lemma gives us what is known as the **Phase Space Continuity Equation**,

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\left(\rho\vec{v}\right) = 0$$

More explicitly breaking up the $6N$ components of the divergence term into the relevant $3N, 3N$ chunks,

$$\frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left(\frac{\partial}{\partial q_i}\left(\rho\dot{q}_i\right) + \frac{\partial}{\partial p_i}\left(\rho\dot{p}_i\right)\right) = 0$$

Which by product rule,

$$\frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial q_i}\dot{q}_i + \frac{\partial\rho}{\partial p_i}\dot{p}_i\right) + \sum_{i=1}^{3N}\left(\rho\frac{\partial\dot{q}_i}{\partial q_i} + \rho\frac{\partial\dot{p}_i}{\partial p_i}\right) = 0$$

Using the equations of motion (5.1),

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial}{\partial q_i}\left(\frac{\partial H}{\partial p_i}\right) = \frac{\partial^2 H}{\partial q_i \partial p_i}$$

Assuming that H is analytical in its components, we can swap order of operations,

$$\frac{\partial^2 H}{\partial q_i \partial p_i} = \frac{\partial^2 H}{\partial p_i \partial q_i} = \frac{\partial}{\partial p_i}\left(\frac{\partial H}{\partial q_i}\right)$$

Therefore by the dynamics of the system,

$$\frac{\partial \dot{q}_i}{\partial q_i} = -\frac{\partial \dot{p}_i}{\partial p_i}$$

Making these terms drop out.

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N}\left(\frac{\partial \rho}{\partial q_i}\dot{q}_i + \frac{\partial \rho}{\partial p_i}\dot{p}_i\right) = 0$$

Again using the equations of motion (5.1),

$$\frac{\partial \rho}{\partial t} + \underbrace{\sum_{i=1}^{3N}\left(\frac{\partial \rho}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i}\frac{\partial H}{\partial q_i}\right)}_{\{\rho, H\}} = 0$$

Where $\{\rho, H\}$ is known as the **Poisson Brackets** for this problem. By analogy in Quantum Mechanics the Poisson brackets become a commutator.

$$\{\rho, H\} \Leftrightarrow [\rho, H]$$

This is known as **Liouville's Theorem**,

$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0$$

Observations:

1. $\frac{\partial \rho}{\partial t} = 0$ is already imposed for the regime of equilibrium statistical mechanics
2. If ρ is a constant, everything becomes zero $\{\rho, H\} = 0$
3. Making $\rho(p, q)$ not explicitly dependent on (p, q) enforces that $\rho = \rho(H(q, p))$
   - If ρ is to have dependence on N, V, T we can makes the choice that $\rho = \exp(-\beta H(q, p))$. In doing so we will show that $\beta = 1/k_b T$ (this is known as the canonical ensemble)

5.3 Microcanonical Ensemble Revisited

As a reminder, we have a hypershell in phase space of energy $E \pm \Delta/2$ with $\rho(q, p)$ is constant on the hypershell. Using out new framework with (5.2) independent of time,

$$\langle f \rangle_{\text{ensemble}} \equiv \frac{\int_\omega f(q,p)\rho(q,p)\mathrm{d}^{3N}q\mathrm{d}^{3N}p}{\int_\omega \rho(q,p)\mathrm{d}^{3N}q\mathrm{d}^{3N}p} \tag{5.5}$$

Where ω indicates out integrals are performed with respect to energy within $E \pm \Delta/2$. Using the notion of time averaging,

$$\langle f \rangle_{\text{exp}} = [f(t)]_{\text{time}} = \frac{1}{T}\int_{t_0}^{t_0+T} f(t)\mathrm{d}t$$

1) The first question that arises: Is the ensemble averaging equal to the time averaging technique?

$$\langle f \rangle_{\text{ens}} \stackrel{?}{=} [f(t)]_{\text{time}}$$

2) The second question that arises: Is $\rho(q,p) =$ constant indicate the probability of finding a representative point in a give volume $\mathrm{d}\omega$ in phase space the same as finding a representative point in a different volume $\mathrm{d}\omega'$? More clearly, does a uniform distribution ρ indicate a uniform probability? A priori equal probability to find the system in any microstate. Up until know, we presumed this to be an axiom or postulate of statistical mechanics (see section 2.2).

Answer to 1): Given the ensemble average $\langle f \rangle_{\text{ens}}$ from (5.5), since the ensemble is stationary $\frac{\partial \rho}{\partial t} = 0$, then nothing is changed doing a time average after an ensemble average.

$$\langle f \rangle = \langle f \rangle_{\text{ens}}$$

Then perform a time average on the ensemble average,

$$\langle f \rangle_{\text{ens}} = [\langle f \rangle_{\text{ens}}]_{\text{time}}$$

As the time average and ensemble average are independent (i.e. the Ergodic Hypothesis is satisfied; see section 5.2.1), swap the order of averaging,

$$\langle f \rangle = [\langle f \rangle_{\text{ens}}]_{\text{time}} = \langle [f]_{\text{time}} \rangle_{\text{ens}}$$

Now notice that the time averaging of any quantity $f(t)$ over a long time T must be the same for each member of the ensemble. Again utilizing the Ergodic hypothesis, any representative point will explore the whole phase space for each member of the ensemble over long enough time. In this is the case, there is *nothing* gained by performing the ensemble averaging after the time averaging. Therefore we can write,

$$\langle [f]_{\text{time}} \rangle_{\text{ens}} = [f]_{\text{time}}$$

Combining these results,

$$[f(t)]_{\text{time}} = \langle f \rangle_{\text{ens}}$$

Therefore the experimental averaging is given by,

$$\langle f \rangle_{\text{exp}} = [f(t)]_{\text{time}} = \langle f \rangle_{\text{ens}}$$

Answer to 2): What is the relationship between $\Gamma(N,V,E)$ the number of microstates for fixed $N, V, \bar{E}$ and ω the volume of phase space (where again $\bar{E} - \Delta/2 < E < \bar{E} + \Delta/2$)? There is a relationship between ω and Γ, namely if you enlarge the phase space volume the total number of potential microstates increases. We should expect,

$$\Gamma \propto \omega$$

However Γ is dimensionless, while ω is not. This was reconciled between through $S = k_B \ln \Gamma$ where k_B caries the appropriate dimensions. Specifically ω has the units $q^{3N} p^{3N}$. This motivates the scaling factor ω_0 such that,

$$\Gamma = \frac{\omega}{\omega_0}$$

The units of ω in a typical case of one particle is given by,

$$[\omega] \sim [q]\,[p] \sim (\mathrm{m})\,(\mathrm{N} \cdot \mathrm{s}) \sim \mathrm{J\,s}$$

Which is equivalent to the units of angular momentum similar to Planck's constant. Therefore we can motivate the factor that $\omega_0 = zh^N$ where z is a dimensionless constant. Recall that the calculation of the Sackur-Tetrode equation (4.12) involved Planck's constant h.

5.4 Two Examples

5.4.1 Compute Entropy

The first example that will be examined is computing S from some phase space integrals involving the previously motivated relationship $S \sim k_B \ln \frac{\omega}{\omega_0}$. We have from the density function $\rho(q,p)$ of representative points,

$$\omega = \underbrace{\int \cdots \int}_{\text{fixed } N,V,E(\pm\Delta/2)} \mathrm{d}^{3N}q\mathrm{d}^{3N}p \cdot \rho(q,p)$$

Examining the case of $\rho = \rho^*$ constant and separating the integrals for position and momentum,

$$\omega = \rho^* \cdot \left\{ \int \cdots \int \mathrm{d}^{3N}q \right\} \left\{ \int \cdots \int \mathrm{d}^{3N}p \right\}$$

Where the $\mathrm{d}^{3N}q$ integral is subject to volume constraints and the $\mathrm{d}^{3N}p$ integral is subject to energy constraints. The position constraint gives,

$$\int \cdots \int \mathrm{d}^{3N}q = V^N$$

Where as the energy constraints give,

$$E - \frac{\Delta}{2} \leq \sum_{i=1}^{3N} \frac{p_i^2}{2m} \leq E + \frac{\Delta}{2}$$

Where i counts the degrees of freedom. We can then perform a change of variables $p \to y$ to get,

$$\int \cdots \int \mathrm{d}^{3N}p = \int \cdots \int \mathrm{d}^{3N}y$$

With modified energy constraint,

$$2m\left(E - \frac{\Delta}{2}\right) \leq \sum_{i=1}^{3N} y_i^2 \leq 2m\left(E + \frac{\Delta}{2}\right)$$

This new integral over y becomes the surface of a $3N$-dimensional hypersphere with thickness given by,

$$\delta y_- = \sqrt{2m\left(E - \frac{\Delta}{2}\right)} \quad \delta y_+ = \sqrt{2m\left(E + \frac{\Delta}{2}\right)} \tag{5.6}$$

In order to determine the surface area of this hypersphere, we will leverage the result computed for the volume of a hypershell before. Therefore the surface area is approximated (for thickness thin enough) that the surface area is given by,

$$\text{volume} = (\text{thickness}) \times (\text{surface area})$$

The thickness of this shell is given by the difference in the bounds given (5.6),

$$\begin{aligned}
\delta y_- &= \sqrt{2m\left(E - \frac{\Delta}{2}\right)} \\
&= \sqrt{2mE}\left(1 - \frac{\Delta}{2E}\right)^{1/2} \\
&= \sqrt{2mE}\left(1 - \frac{\Delta}{4E} + \cdots\right) \quad \text{Taylor expansion } (1+x)^n \approx 1 + nx \\
&= \sqrt{2mE} - \frac{1}{2}\frac{\Delta\sqrt{m}}{\sqrt{2E}}
\end{aligned}$$

Similarly for the upper bound,

$$\delta y_+ = \sqrt{2mE} + \frac{1}{2}\frac{\Delta\sqrt{m}}{\sqrt{2E}}$$

Therefore the thickness is given by,

$$\delta y_+ - \delta y_- = \frac{\Delta\sqrt{m}}{\sqrt{2E}}$$

Using the surface area given by a derivative in (4.7) with $n = 3N$,

$$S_n(R) = \frac{\mathrm{d}V_n(R)}{\mathrm{d}R} = \frac{n\pi^{n/2}}{\left(\frac{n}{2}\right)!}R^{n-1} = \frac{3N\pi^{3N/2}}{\left(\frac{3N}{2}\right)!}R^{3N-1} = \frac{2\pi^{3N/2}}{\left(\frac{3N}{2}-1\right)!}R^{3N-1}$$

In this example, the radius is $R = \sqrt{2mE}$ so the volume of the shell becomes the thickness times the surface area,

$$\text{surface area} = \frac{2\pi^{3N/2}}{\left(\frac{3N}{2}-1\right)!} \times (2mE)^{\frac{3N-1}{2}}$$

$$\omega = V^N \times \Delta\left(\frac{m}{2E}\right)^{1/2} \times \frac{2\pi^{3N/2}}{\left(\frac{3N}{2}-1\right)!} \times (2mE)^{\frac{3N-1}{2}} \tag{5.7}$$

$$\omega = V^N \times \left(\frac{\Delta}{E}\right) \times \frac{(2\pi mE)^{3N/2}}{\left(\frac{3N}{2}-1\right)!}$$

Comparing with (4.9) we get the expected result for ω_0,

$$\frac{\omega}{\Gamma} \equiv \omega_0 = h^{3N}$$

5.4.2 Simple 1D Classical Harmonic Oscillator

Recall that for a 1D classical harmonic oscillator,

$$H(q,p) = \frac{kq^2}{2} + \frac{p^2}{2m} = E$$

Where $U = \frac{kq^2}{2}$ is the potential energy V of the system, $\frac{p^2}{2m}$ is the kinetic energy T of the system and E is the total energy of the system which is a constant of motion. For simplicity, we will use the natural angular frequency ω' of the system as $m(\omega')^2 = k$. Note that this ω' with the prime is different from the phase space volume introduced in (5.7). This can be a source of confusion, just be mindful of the different quantities. Therefore the phase space volume for a simple harmonic oscillator is just the *area* of an ellipse ($A = \pi ab$). The equation of the ellipse is given by,

$$\left(\frac{q^2}{2E/m\omega'^2} + \frac{p^2}{2mE}\right) = 1$$

The solutions to the equations of motion are given by,

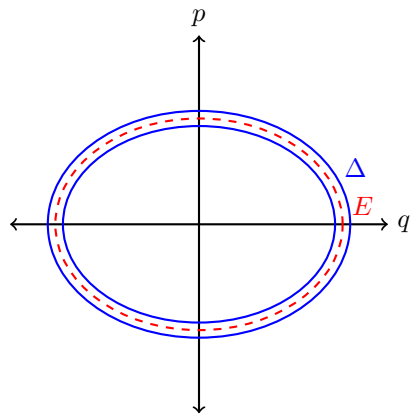
$$q(t) = A\cos(\omega' t + \phi) \quad p(t) = -A\sin(\omega' t + \phi)$$

With area given by,

$$A = \pi\sqrt{2E/m\omega'^2}\sqrt{2mE} = \frac{2\pi E}{\omega'}$$

Therefore subject to the constraints $E - \Delta/2 \leq H(q,p) \leq E + \Delta/2$,

$$\omega = \int \mathrm{d}q\mathrm{d}p = A_{E+\Delta/2} - A_{E-\Delta/2}$$



We are interested in the area between the outer and inner radii. Therefore,

$$\omega = \frac{2\pi\left(E + \frac{\Delta}{2}\right)}{\omega'} - \frac{2\pi\left(E - \frac{\Delta}{2}\right)}{\omega'} = \frac{2\pi\Delta}{\omega'}$$

Noticing that the phase space volume (PSV) contains the characteristic frequency $2\pi/\omega'$.

Now suppose we are interested in determining the "elementary PSV" ω_0. Taking as an initial guess from the quantum mechanical simple harmonic oscillator,

$$E_n = \hbar\omega'\left(n + \frac{1}{2}\right) \implies \Delta E = E_{n+1} - E_n = \hbar\omega'$$

Therefore the elementary phase space volume becomes,

$$(\text{PSV})_{\text{elementary}} = \omega_0 = \frac{2\pi}{\omega'}\hbar\omega' = 2\pi\hbar = h$$

This motivates the elementary phase space volume to be roughly h. Another approach to determining the elementary phase space volume is to use the notion that $E \gg De \gg \hbar\omega'$. Therefore the number of eigenstates that would fit into Δ would be given by, $\# \sim \Delta/\hbar\omega'$. Therefore,

$$(\text{PSV})_{\text{elementary}} \sim \frac{\frac{2\pi}{\omega'}\Delta}{\#} = \frac{\frac{2\pi}{\omega'}\Delta}{\frac{\Delta}{\hbar\omega'}} = 2\pi\hbar \sim h$$

Or alternatively, one can motivate this irreducible unit of phase space volume h by using the uncertainty principle $\Delta x \Delta p \sim \hbar$.

5.4.3 Reflection

Both of the above examples illustrate that computation of phase space volumes are a pain in the neck. We should be looking to *replace* it with something else. If we have the relation $S = k_B \ln(\Gamma)$ which gives,

$$S = k_B \ln(\Gamma) = k_B \ln\left(\frac{\omega}{\omega_0}\right) = k_B \ln(\omega) - \underbrace{k_B \ln(\omega_0)}_{\text{const.}}$$

The pain of computing ω_0 can be ignored or neglected because most thermodynamical variables are related through derivatives of S. In these cases, the constant contribution has zero derivative so it can be skipped.

6 Canonical Ensemble

6.1 Preamble

Up until now we have considered the rather restricted or simple case of systems with fixed N, V, E. This is a natural choice for a first exploration, as it involves isolated systems. Computing microstates and phase

space volume under these restrictions are known as the **microcanonical ensemble**.

There are two comments to be made with regards to the microcanonical ensemble. The first being that E is rarely perfectly fixed, and is rarely known exactly experimentally. The means performing an analysis with fixed energy E is somewhat non-physical. Secondly, no system is perfectly isolated; keeping E fixed in a real system is also artificial and difficult. Doing this would require a measurement device that measures energy directly, and makes adjustments to the system that an experimenter can control (like V, T, N) in order to ensure E is fixed.

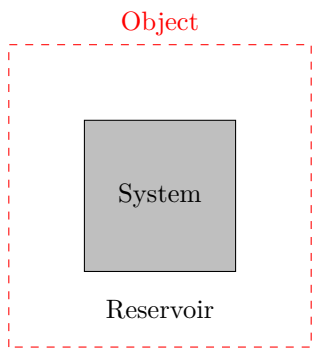
An even easier way of controlling the system, is to fix the temperature. This modifies the ensemble to be restricted by N, V, T. This is known as the **canonical ensemble**. As one would expect, other "ensembles" can be constructed by fixed other combinations of quantities of the system.

6.2 Tackling the Canonical Ensemble

In any ensemble problem the goal is to determine the structure of $\rho(q, p)$. Typically, we should expect ρ to have the form,

$$\rho(q, p) \propto e^{-H(q,p)/k_B T}$$

Where $H(q, p)$ is the Hamiltonian or the total energy of the system.



We will denote the energy of the reservoir $E^{(r)}$ and the energy of the system $E^{(s)}$. The total energy of the object will be constant at $E = E^{(r)} + E^{(s)}$ where typically $E^{(r)} \gg E^{(s)}$. We want to define some microstates for the reservoir and system. Typically one would expect that the reservoir has many more microstates. Let the probability P_j that the probability that the system is in a particular microstate j. The probability P_j is proportional to the number of sates in the reservoir that are consistent with the system being in the microstate j. The reason for this is that since the whole object is isolated, all the microstates between the system and reservoir are equally probable.

If the system is in state j then the number of microstates available to the reservoir is constrained by the fact that the total energy $E = E^{(r)} + E^{(s)}$ is constant. The system with microstate j is said to have energy $E^{(s)} = E_j$. If the system changes microstate and it's own energy changes by an amount ΔE, then the reservoir must change energy by an amount $-\Delta E$. Where ΔE is the change in energy of the system when transiting from microstate j to microstate i. Therefore,

$$\Delta E = E_i - E_j$$

We can then ask the question: What is the ratio of probability that the system is in microstate j vs. microstate i.

$$\frac{P_i}{P_j} = \frac{\Omega^{(s)} \times \Omega^{(r)}\left(E^{(r)} - \Delta E\right)}{\Omega^{(s)} \times \Omega^{(r)}\left(E^{(r)}\right)}$$

But since we are considering the system to be in a particular microstate $\Omega^{(s)} = 1$,

$$\frac{P_i}{P_j} = \frac{\Omega^{(r)}\left(E^{(r)} - \Delta E\right)}{\Omega^{(r)}\left(E^{(r)}\right)} \tag{6.1}$$

Evidently, we can related the number of microstates available to the reservoir to the entropy of the reservoir using $S^{(r)} = k_B \ln \Omega^{(r)}$. Therefore we can rewrite (6.1) using logarithms,

$$\frac{P_i}{P_j} = \exp\left\{\ln\left(\Omega^{(r)}\left(E^{(r)} - \Delta E\right)\right) - \ln\left(\Omega^{(r)}\left(E^{(r)}\right)\right)\right\}$$

Which can be written,

$$\frac{P_i}{P_j} = \exp\left\{\frac{S^{(r)}\left(E^{(r)} - \Delta E\right) - S^{(r)}\left(E^{(r)}\right)}{k_B}\right\}$$

However since $E^{(r)} \gg E^{(s)}$, then $E^{(r)} \gg \Delta E^{(s)}$. A Taylor series expansion gives,

$$S^{(r)}\left(E^{(r)} - \Delta E\right) - S^{(r)}\left(E^{(r)}\right) = -\Delta E \frac{\partial S^{(r)}}{\partial E^{(r)}}$$

But thermodynamically,

$$\frac{1}{T^{(r)}} = \frac{\partial S^{(r)}}{\partial E^{(r)}}$$

Therefore we have the familiar relation (using the fact that in equilibrium $T = T^{(r)} = T^{(s)}$),

$$\frac{P_i}{P_j} = \exp\left\{-\frac{\Delta E}{k_B T}\right\}$$

Or for individual probabilities,

$$P_i \propto \exp\left\{-\frac{E_i}{k_B T}\right\} \tag{6.2}$$

The proportionality sign '$\propto$' is present because this relationship is still missing in general a normalization factor.

6.2.1 Example

Let us perform an example. Consider a system where we have a single spin 1/2 particle in a magnetic field $\vec{B}$. The interaction energy is,

$$H = -\vec{\mu} \cdot \vec{B}$$

Where $\vec{\mu} = -g\mu_B \vec{S}$. Taking the magnetic field to point in one direction $\vec{B} = B\hat{z}$,

$$H = g\mu_B B_z S_z$$

Where S_z can either be $+\frac{1}{2}$ or $-\frac{1}{2}$. Consider E_e corresponding to $S_z = +\frac{1}{2}$ and E_0 corresponding to $S_z = -\frac{1}{2}$. Then we have,

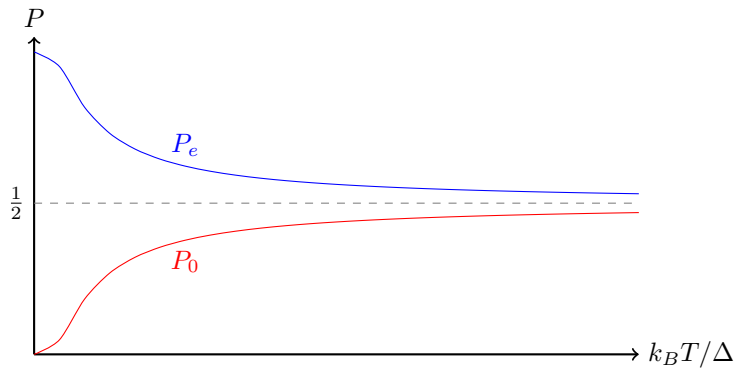
$$P_e \propto e^{-E_e/k_B T} \qquad P_0 \propto e^{-E_0/k_B T}$$

Since $P_e + P_0 = 1$, we get,

$$P_e = \frac{P_e}{P_e + P_0} = \frac{e^{-E_e/k_B T}}{e^{-E_e/k_B T} + e^{-E_0/k_B T}}$$

In general for appropriate choice of $\beta\Delta$,

$$P_e = \frac{e^{-\beta\Delta}}{1 + e^{-\beta\Delta}} \qquad P_0 = \frac{1}{1 + e^{-\beta\Delta}}$$



6.3 Canonical Ensemble Description

Consider an ensemble of identical systems $1, 2, 3, \ldots, \tilde{N}$ that share a total energy $\tilde{E}$. We will label each system with an index r where each system r has energy E_r. We will call n_r the number of the systems in the ensemble that happen to have the same energy E_r. The two restrictions can be written,

$$\sum_r n_r = \tilde{N} \quad \sum_r n_r E_r = \tilde{E}$$

And set of $\{n_r\}$ that satisfies both of these two equations represents a possible "mate"/"type" of distribution of the "systems" in the ensemble. Each of these solution $\{n_r\}$ can be realized in many different ways through permutations of $\{n_r\}$,

$$W(\{n_r\}) = \frac{\tilde{N}!}{n_1! n_2! \cdots n_r!}$$

We aim to find the most probable distribution $W(\{n_r\})$. We can do so by extremizing $\ln W$ with respect to (n_r).

$$\begin{aligned}
\ln W &= \ln\left(\tilde{N}!\right) - \ln\left(\prod_r n_r!\right) \\
&= \ln\left(\tilde{N}!\right) - \sum_r \ln\left(n_r!\right) \quad \text{Logarithm identities} \\
&= \tilde{N}\ln\left(\tilde{N}\right) - \tilde{N} - \sum_r \left[n_r \ln\left(n_r\right) - n_r\right] \quad \text{Stirling's approximation} \\
&= \tilde{N}\ln\left(\tilde{N}\right) + \underbrace{\sum_r n_r - \tilde{N}}_{=0} - \sum_r n_r \ln\left(n_r\right) \\
&= \tilde{N}\ln\left(\tilde{N}\right) - \sum_r n_r \ln\left(n_r\right)
\end{aligned}$$

Therefore varying $\ln W$ with respect to n_r,

$$\delta \ln W = \delta\left\{\tilde{N}\ln\left(\tilde{N}\right) - \sum_r n_r \ln\left(n_r\right)\right\} = -\sum_r \left(\delta n_r \ln n_r + n_r \frac{1}{n_r}\delta n_r\right) \quad \text{Product rule}$$

Therefore,

$$\delta \ln W = -\sum_r \left(\ln n_r + 1\right)\delta n_r \tag{6.3}$$

Setting this equal to zero $0 = \delta \ln W$ gives and the two modified constraints $\sum_r \delta n_r = 0$ and $\sum_r \delta n_r \cdot E_r = 0$, use *Lagrange multipliers* to solve this equation. For each constraint we subscribe two multipliers,

$$\alpha \to \sum_r \delta n_r = 0 \qquad \beta \to \sum_r \delta n_r \cdot E_r = 0$$

Therefore 6.3 becomes,

$$\sum_r \{(-\ln n_r - 1) - \alpha - \beta E_r\} \delta n_r = 0$$

In order to solve this equation, we need for all δn_r,

$$(-\ln n_r - 1) - \alpha - \beta E_r = 0$$

After rearrangement,

$$\ln n_r = -(1+\alpha) - \beta E_r$$

Or that the distribution of n_r must be,

$$n_r = e^{-(1+\alpha)} e^{-\beta E_r}$$

Thus extremizing W with respect to $\{n_r\}$ gives,

$$n_r \propto e^{-\beta E_r}$$

What remains is to show $\beta = 1/k_B T$. The derivation that lead to (6.2) suggests that $\beta = 1/k_B T$. Considering the probability that the system is,

$$P(n_r) \propto e^{-\beta E_r}$$

Normalizing the total probability to be unity gives the value of α,

$$P(n_r) = \frac{e^{-\beta E_r}}{\sum_i e^{-\beta E_i}} \tag{6.4}$$

6.4 Physical Significance of Various Statistical Variables in the Canonical Ensemble

First we shall ask what is the average energy E. Recall that here E is not necessarily fixed. To follow notation from the textbook, we will relabel the average energy of the system U. Then,

$$U = \langle E_r \rangle = \sum_r E_r P(n_r)$$

In terms of notation, it makes no difference to write $P(n_r)$ or $P(E_r)$ because for each n_r there is a corresponding E_r.

$$U = \langle E_r \rangle = \sum_r E_r P(E_r)$$

Therefore using (6.4),

$$U = \frac{\sum_r E_r e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} \tag{6.5}$$

The denominator here is known as the **canonical partition function**.

$$Z \equiv \sum_r e^{-\beta E_r}$$

Example: Suppose we are dealing with a spin system with Hamiltonian given by $\hat{H} = -D\hat{S}_z^2$ with $S = 1$ where D is a difference in energy levels. Then the time independent Schrödinger equation yields,

$$\hat{H} |n\rangle = E |n\rangle$$

With matrix for $\hat{S}_z$,

$$\hat{S}_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

This gives energy levels,

$$E = \begin{cases} -D & (S_z = +1) \\ -D & (S_z = -1) \\ 0 & (S_z = 0) \end{cases}$$

Using this spectrum of energies, all that remains is to plug these states into (6.5) to get,

$$\sum_r e^{-\beta E_r} = e^{\beta D} + e^{\beta D} + e^0 = 2e^{\beta D} + 1$$

Therefore the average internal energy is given by,

$$U = \frac{\sum_r E_r e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} = \frac{-De^{\beta D} - De^{\beta D} + 0e^0}{2e^{\beta D} + 1} = \frac{-2De^{\beta D}}{2e^{\beta D} + 1}$$

Using this we can determine the specific heat of the system,

$$c_V = \left(\frac{\partial U}{\partial T}\right)_V \quad c_N = \frac{\partial U}{\partial T}$$

Specific heat is important because these are quantities that are measurable or tractable to experimentalists. Using chain rule,

$$\frac{\partial U}{\partial T} = \frac{\partial \beta}{\partial T} \cdot \frac{\partial U}{\partial \beta}$$

Where $\beta = 1/k_B T \implies \frac{\partial \beta}{\partial T} = -1/k_B T^2$.

6.4.1 Logarithm of Partition Function

One can convince themselves that,

$$\langle U \rangle = -\frac{\partial \ln Z}{\partial \beta}$$

Namely the proof is as follows,

$$\ln Z = \ln\left(\sum_r e^{-\beta E_r}\right) \implies \frac{\partial \ln Z}{\partial \beta} = -\frac{\sum_r E_r e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} = -\langle U \rangle$$

This means that all that needs to be calculated is Z and the remaining thermodynamical parameters of the system are generated by derivatives of $\ln Z$.

6.4.2 Helmholtz Free Energy

The Helmholtz free energy A (sometimes F) is defined to be,

$$A = U - TS \tag{6.6}$$

Using the first law (with added chemical change),

$$\mathrm{d}U = T\mathrm{d}S - P\mathrm{d}V + \mu\mathrm{d}N$$

Since we have $\mathrm{d}A = \mathrm{d}U - T\mathrm{d}S - S\mathrm{d}T$, it is a function of state like S or U. Combining these two terms yields,

$$\mathrm{d}A = -\cancel{T\mathrm{d}S} - S\mathrm{d}T + \cancel{T\mathrm{d}S} - P\mathrm{d}V + \mu\mathrm{d}N$$
$$\mathrm{d}A = -S\mathrm{d}T - P\mathrm{d}V + \mu\mathrm{d}N$$

This suggests the following relations,

$$P = -\left(\frac{\partial A}{\partial V}\right)_{N,T} \quad S = -\left(\frac{\partial A}{\partial T}\right)_{N,V} \quad \mu = \left(\frac{\partial A}{\partial N}\right)_{V,T} \tag{6.7}$$

A is very useful as its derivatives generate quantities of interest. From (6.6) we obtain,

$$U = A + TS = A - T\frac{\partial A}{\partial T}N, V = -T^2\left\{\frac{\partial}{\partial T}\left(\frac{A}{T}\right)\right\}_{N,V}$$

Furthermore using this equation, it can more suggestively be written as,

$$U = \frac{\partial\left(A/T\right)}{\partial\left(1/T\right)}$$

Moreover since $\beta = 1/k_BT$ one can derive,

$$\ln\left(\sum_r e^{-\beta E_r}\right) = -\frac{A}{k_BT}$$

This analysis leads to the postulate that from $A = -k_BT\ln Z$ the rest of thermodynamics follows.

$$A = -k_BT\ln Z \implies \text{Thermodynamics}$$

If one can compute the partition function Z for a given system, this relationship gives A and then using (6.7) we can obtain quantities of P, S and μ and indirectly T, U etc. So where and how does Z depend on V, T, N? We will need to know this to explore quantities (6.7). Dependence on temperature T is directly related to the definition of Z and $\beta = 1/k_BT$. The dependence on V is determined by how the energy depends on the volume of the system. Recall from a 1d particle in a box $V = L^3$,

$$\varepsilon = \frac{\hbar^2}{2m}\frac{\pi^2}{L^2}n^2 = \frac{\hbar^2}{2m}\frac{\pi^2}{V^{2/3}}n^2$$

Or that $\varepsilon \propto V^{-2/3}$.

6.5 Summary of Z

From this, we have learned that if we can compute the partition function Z and take derivatives, **all physical quantities** can be computed. To obtain the the partition there are a few steps.

1. Determine the Hamiltonian H of the system (active areas of research)
2. Solve the quantum mechanical system $H\psi_n = E_n\psi_n$ (in principle, only E_n is needed and not ψ_n)
3. Compute the sum $\sum_n e^{-\beta E_n}$ (computationally very expensive)

In practice, the first step is very difficult for newly discovered systems. Furthermore, computing the third step precisely using a naive brute force approach is nearly impossible. We need to develop smarter ways to obtain the partition function Z.

6.6 Alternative Form for Z & Degeneracy of States

In any notation the partition function is just a *sum over states*,

$$Z = \sum_n e^{-\beta E_n} = \sum_r e^{-\beta E_r} = \sum_a e^{-\beta E_a} \tag{6.8}$$

Let us apply this to the hydrogen atom. The energy levels for the hydrogen atoms is given by,

$$E_n = -\frac{13.6\,\text{eV}}{n^2}$$

Where n is the principle quantum number. Notice that this n comes from a quantum mechanical consideration of the boundaries conditions for $H\psi_n = E_n\psi_n$ where $\psi = \psi_{nlm}\left(r, \theta, \phi\right)$ is indexed by many number

n, l, m. Writing the states $|nlm\rangle$, the first ground state is written $|1,0,0\rangle$. For $n = 2$, there are four states $|2,0,0\rangle, |2,1,0\rangle, |2,1,-1\rangle, |2,1,1\rangle$. There are many states that share the same energy E_n as $n \to \infty$. Therefore we can compute the partition function,

$$Z = e^{-\beta(-13.6\,\mathrm{eV})} + e^{-\beta\left(-13.6\,\mathrm{eV}/2^2\right)} + e^{-\beta\left(-13.6\,\mathrm{eV}/2^2\right)} + \cdots$$

We can combine the many repeated sums over the same energy level,

$$Z = e^{-\beta(-13.6\,\mathrm{eV})} + 4e^{-\beta\left(-13.6\,\mathrm{eV}/2^2\right)} + \cdots$$

Or using the term **degeneracy of state n**, g_n,

$$Z = \sum_n g_n e^{-\beta E_n} \tag{6.9}$$

It is very important to note that (6.9) is different from (6.8). (6.9) sums over the distinct energies of the system, while (6.8) sums over the states of the system. Another sloppy notation that is invoked is,

$$Z = \sum_\varepsilon g(\varepsilon)\, e^{-\beta\varepsilon}$$

Labels the energy levels not by indices but by the actual energies ε themselves. This is permitted because ε are distinct. If the spacing between the energy levels $\varepsilon, \varepsilon + \Delta\varepsilon$ is small enough, we reduce to an integral,

$$Z = \int_\varepsilon g(\varepsilon)\, e^{-\beta\varepsilon} \mathrm{d}\varepsilon$$

In this case $g(\varepsilon)$ is counting how many states are there between ε and $\varepsilon + \Delta\varepsilon$. In principle, the bounds for ε are from 0 to infinity and Z is a function of temperature T; therefore,

$$Z(T) = \int_0^\infty g(\varepsilon)\, e^{-\beta\varepsilon} \mathrm{d}\varepsilon$$

Note that in the Pathria textbook, the notation used for this quantity is $Q_N(T)$ but we will use the more standard $Z_n(T)$. We can recover in general a discrete case by having $g(E) = \delta(E - E_0) + \delta(E - E_1) + \cdots$ (where δ is the Dirac-Delta function). This gives,

$$Z(T) = \int_0^\infty \left(\delta(E - E_0) + \delta(E - E_1) + \cdots\right) e^{-\beta\varepsilon} \mathrm{d}\varepsilon$$

$$Z(T) = e^{-\beta E_0} + e^{-\beta E_1} + \cdots$$

This function $g(\varepsilon)$ is sometimes called the **density of states**. Using Dirac-Deltas $g(\varepsilon)$ can be applied to continuous and discrete distributions.

6.7 Classical Systems / Free Particles

Recall from the Microcanonical ensemble (5.5) gives the average value for a physical quantity f. Using $\rho = \exp(-\beta H(p,q))$,

$$\langle f \rangle = \frac{\int_\omega f(p,q)\, e^{-\beta H(p,q)} \mathrm{d}^{3N}q \mathrm{d}^{3N}p}{\int_\omega e^{-\beta H(p,q)} \mathrm{d}^{3N}q \mathrm{d}^{3N}p}$$

Where with proper scaling factor h^{3N} and the Gibb's Paradox correct $N!$,

$$\mathrm{d}\omega = \frac{\mathrm{d}^{3N}q \mathrm{d}^{3N}p}{h^{3N} N!}$$

Note that in this course we typically choose our measurable quantity f to be position, energy, momentum, etc. However in principle, we can take f to be **any** measurable quantity.

Example: As an example, let us look at the scenario of free classical particles with Hamiltonian given (most of the time) by,

$$H(p,q) = \text{kinetic energy} + \text{potential energy}$$

More explicitly,

$$T = \sum_{i=1}^{N} \frac{|\vec{p}_i|^2}{2m} \qquad V = \sum_{i>j} V_{ij}(r_{ij})$$

Note that is *not* always possible to write potential energies are functions of the distance between particles r_{ij} directly. Nonetheless it is valid for numerous examples. As an example for particles in magnetic field, with $\vec{\nabla} \cdot \vec{B} = \vec{0}$ we can write $\vec{B}$ as the curl of a vector potential,

$$\vec{B} = \vec{\nabla} \times \vec{A}$$

Using this vector potential, the Hamiltonian can be written,

$$H = \frac{\left|\vec{p} - e\vec{A}\right|^2}{2m}$$

This is a simple example where T and V are coupled together. Moving forward the distance between two distinct particles can be written,

$$|\vec{r}_{ij}|^2 = |\vec{r}_i - \vec{r}_j|^2$$

Where $\vec{r}_i$'s are the generalized position components in the Hamiltonian formalism,

$$\vec{r}_i = (r_{ix}, r_{iy}, r_{iz}) = (q_{ix}, q_{iy}, q_{iz})$$

Computing the partition function,

$$Z = \int_\omega e^{-\beta H} \mathrm{d}\omega = \int_\omega e^{-\beta H} \frac{\mathrm{d}^{3N} q \mathrm{d}^{3N} p}{h^{3N} N!}$$

Subbing in the general Hamiltonian,

$$Z = \int_\omega e^{-\beta\left\{\sum_{i=1}^{N} \frac{|\vec{p}_i|^2}{2m}\right\} - \beta\left\{\sum_{i>j} V_{ij}(r_i, r_j)\right\}} \frac{\mathrm{d}^{3N} q \mathrm{d}^{3N} p}{h^{3N} N!}$$

Decoupling the momentum and position terms (for cases unlike the above where the potential energy **does not** depend on momentum),

$$Z = \int_{\omega_p} e^{-\beta\left\{\sum_{i=1}^{N} \frac{|\vec{p}_i|^2}{2m}\right\}} \frac{\mathrm{d}^{3N} p}{h^{3N} N!} \times \int_{\omega_q} e^{-\beta V(q)} \mathrm{d}^{3N} q$$

Here we have written $\sum_{i=1}^{N} V_{ij}(r_i, r_j) = V(q)$. The second component is often very intractable. The first term can be decomposed more,

$$\int_{\omega_p} e^{-\beta\left\{\sum_{i=1}^{N} \frac{|\vec{p}_i|^2}{2m}\right\}} \frac{\mathrm{d}^{3N} p}{h^{3N} N!} = \int_{\omega_p} e^{-\beta\left(\frac{p_{1,x}^2}{2m} + \cdots + \frac{p_{2,y}^2}{2m} + \cdots\right)} \frac{\mathrm{d}^{3N} p}{h^{3N} N!} = \left[\int_{-\infty}^{\infty} e^{-\beta \frac{p_\alpha^2}{2m}} \mathrm{d}p_\alpha\right]^{3N} \frac{1}{h^{3N} N!}$$

We have seen the integral before in (2.4.3),

$$\int_{-\infty}^{\infty} e^{-ax^2} \mathrm{d}x = \sqrt{\frac{\pi}{a}}$$

Therefore taking the potential energy V be zero (this can be validated if $k_BT \gg V$ which makes the interaction energy negligible compared to the thermal energy),

$$Z_N = \left((2m\pi kT)^{1/2}\right)^{3N} \frac{1}{h^{3N}N!} \times \underbrace{\int_{w_q} e^0 \mathrm{d}^{3N}q}_{V^N}$$

Where V^N is the physical volume to the power of N and it **not** the potential energy. This is bad notation. More suggestively written gives,

$$Z_N = \frac{1}{N!}\left\{\frac{V}{h^3}(2m\pi kT)^{3/2}\right\}^N$$

From this we can look at the Helmholtz free energy,

$$A = -k_BT \ln Z_N$$

And one can recover,

$$PV = Nk_BT$$

As well as the Sackur-Tetrode equation is S is computed.

What about a combination of N_1 particles with mass m_1 and N_2 particles with mass m_2 where $m_2 \neq m_1$. This derivation is modified in such a way that,

$$\iint e^{-ax^2-by^2}\mathrm{d}x\mathrm{d}y = \int e^{-ax^2}\mathrm{d}x \int e^{-ay^2}\mathrm{d}y = \sqrt{\frac{\pi}{a}}\sqrt{\frac{\pi}{b}}$$

The partition function of two groups noninteracting particles is multiplicative.

$$Z = Z_1 \times Z_2$$

Since the entropy is characterized by the logarithm of the partition functions, there corresponding entropies are additive.

Remark: The notation of $\mathrm{d}^{3N}p$ can tend to be confusing. Expanding this out,

$$\mathrm{d}^{3N}p = \mathrm{d}p_1\mathrm{d}p_2\cdots\mathrm{d}p_{3N}$$

Here there are $3N$ terms to consider and often leads to $3N$ integrals. The p_i in this context indicates a particular component of a particular particle. For example, p_i could be the y component of the 53rd particle. Alternatively, when p_i's are written in this way,

$$H = \sum_{i=1}^{N} \frac{|\vec{p}_i|^2}{2m} = \frac{p_{1,x}^2 + p_{1,y}^2 + p_{1,z}^2}{2m} + \cdots + \frac{p_{N,x}^2 + p_{N,y}^2 + p_{N,z}^2}{2m}$$

Each $\vec{p}_i$ represents the total momentum vector of the particle. Be careful of this distinction. It should be evident from the context of the problem.

For example, calculating the canonical partition function is in general written,

$$Z = \frac{1}{h^{3N}N!}\int \mathrm{d}^{3N}q \int \mathrm{d}^{3N}p e^{-\beta H(q,p)}$$

But it is common that $H(p,q) = H(p)$ whenever the potential energy is zero $V = 0$. Then,

$$Z = \frac{V^N}{h^{3N}N!}\int \mathrm{d}^{3N}p e^{-\beta\left(\frac{p_1^2}{2m}+\frac{p_2^2}{2m}+\cdots+\frac{p_{3N}^2}{2m}\right)}$$

$$Z = \frac{V^N}{h^{3N}N!}(2m\pi k_BT)^{3N/2}$$

6.8 Examples of C.E.

6.8.1 Two Masses on Springs

Consider a system of two masses m_1 and m_2 connected by springs to each other and a wall along a linear 1d line. Therefore each mass can only move in along the x-axis. If the spring constants are c_1, c_2 and k with k being between the two masses with x_1 being the position of the first mass and x_2 being the position of the second mass.

$$H = \left(\frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2}\right) + \left(c_1\frac{x_1^2}{2} + c_2\frac{x_2^2}{2} + k\frac{(x_2 - x_1)^2}{2}\right)$$

Therefore the canonical partition function can be computed,

$$Z = \frac{1}{h^2 2!}\iint \mathrm{d}p_1\mathrm{d}p_2 e^{-\beta\left(\frac{p_1^2}{2m_1}+\frac{p_2^2}{2m_2}\right)} \times \iint \mathrm{d}x_1\mathrm{d}x_2 e^{-\beta\left\{c_1\frac{x_1^2}{2}+c_2\frac{x_2^2}{2}+k\frac{(x_2-x_1)^2}{2}\right\}}$$

The kinetic energy contribution is a Gaussian integral (see section 2.4.3) as seen many times before,

$$Z = (2\pi m_1 k_B T)^{1/2}(2\pi m_2 k_B T)^{1/2} \times \iint \mathrm{d}x_1\mathrm{d}x_2 e^{-\beta\left\{c_1\frac{x_1^2}{2}+c_2\frac{x_2^2}{2}+k\frac{(x_2-x_1)^2}{2}\right\}}$$

And the position contribution is also a Gaussian integral, just slightly modified. The solution is left as an exercise to the reader.

Remark: As can be seen in this two examples, even simple problems can lead to complicated integrals.

6.8.2 Rigid Polar Rotator

Consider a dilute gas of H_2C or CO with electric dipole moment $\vec{\mu}$ in an electric field $\vec{E}$. What is the partition function $Z(T)$? One would suspect that is *should* on the general form,

$$Z = \int \cdots \int_\omega e^{-\beta H(\{q,p\})}[\cdots]$$

What is the Hamiltonian? Since the gas is *dilute*, it can be approximated that there are no interactions between the particles. As such, the partition function and Hamiltonian can be determined for a single particle i and then combined together appropriately,

$$Z = \frac{(Z_i)^N}{N!} \qquad H = \sum_{i=1}^n H_i$$

We will hence forth consider looking at a single particle,

$$H_i(p,q) = \frac{|\vec{p}|^2}{2m} - \vec{\mu}\cdot\vec{E}$$

Since $V = -\vec{\mu}\cdot\vec{E}$. For convenience let use take the electric field to be pointing in the $\hat{z}$ direction such that $\vec{E} = E\hat{z}$. Therefore,

$$H_i(p,q) = \frac{|\vec{p}|^2}{2m} - \mu E\cos\theta$$

Where θ is the angle between μ and the $+\hat{z}$ axis. The partition function for an individual molecule is given by,

$$Z_i = \frac{1}{h^5}\int e^{-\beta H(q,p)}\mathrm{d}q\mathrm{d}p$$

Taking the generalized position q to be θ (as is common in many textbooks),

$$Z_i \stackrel{?}{=} \frac{1}{h^5} \int e^{-\beta H(\theta,p)} \mathrm{d}\theta \mathrm{d}p \mathrm{d}p_\theta$$

Notice the $\mathrm{d}p_\theta$ **needs** to be there because for every generalized coordinate has corresponding momentum p_θ. More properly there is dependence on many generalized coordinates. The position of the molecule is given by x, y, z and the generalized linear momentum p_x, p_y, p_z. Also the orientation of the molecule is characterized by θ, ϕ and generalized angular momentum p_θ, p_ϕ. There are 10 coordinates total that the Hamiltonian depends which makes the partition function,

$$Z_i = \underbrace{\iiiint}_{4} \mathrm{d}p_\theta \mathrm{d}p_\phi \mathrm{d}\theta \mathrm{d}\phi \underbrace{\int \cdots \int}_{6} \mathrm{d}x \mathrm{d}y \mathrm{d}z \mathrm{d}p_x \mathrm{d}p_y \mathrm{d}p_z e^{-\beta H(\theta,\phi,p_\theta,p_\phi,x,y,z,p_x,p_y,p_z)}$$

Note, this integral is not decoupled. It is just written this way to illustrate the separate dependence on each of the sets of terms. The Hamiltonian has linear components and rotational components characterized by,

$$H = \frac{|\vec{p}|^2}{2m} - \mu E \cos\theta + H_{\mathrm{rot}}$$

Noting that is still for a single molecule. The rotation terms are given by,

$$H_{\mathrm{rot}} = \frac{m}{2} \left(\underbrace{\ell \dot{\theta}}_{v_\theta} \right)^2 + \frac{m}{2} \left(\underbrace{\ell \sin\theta \dot{\phi}}_{v_\phi} \right)^2$$

As an exercise, make sure these quantities have units of energy. An alternative way of writing this,

$$H_{\mathrm{rot}} = \frac{p_\theta^2}{2m\ell^2} + \frac{p_\phi^2}{2m\ell^2 \sin^2\theta}$$

The total Hamiltonian can be written (still for a single molecule),

$$H = \frac{1}{2m} \left(p_x^2 + p_y^2 + p_z^2 \right) + \left\{ \frac{p_\theta^2}{2m\ell^2} + \frac{p_\phi^2}{2m\ell^2 \sin^2\theta} \right\} - \mu E \cos\theta$$

As such for a single molecule integration over ,

$$Z = \frac{1}{h^5} \left(2\pi m k_B T \right)^{3/2} V \times \{\cdots\}$$

The other six integrals are given by,

$$\int_{-\infty}^{\infty} e^{-\frac{\beta p_\theta^2}{2mr^2}} \mathrm{d}p_\theta = \left(2m\pi r^2 k_B T \right)^{1/2}$$

$$\int_{-\infty}^{\infty} e^{-\frac{\beta p_\phi^2}{2mr^2 \sin^2\theta}} \mathrm{d}p_\phi = \left(2m\pi r^2 k_B T \sin^2\theta \right)^{1/2}$$

Note that the generated $\sin\theta$ term is important here.

$$\int_0^{2\pi} \mathrm{d}\phi = 2\pi$$

$$\int_0^{\pi} \mathrm{d}\theta \sin\theta e^{-\beta(-\mu E \cos\theta)} = ?$$

The canonical partition function is reduced to,

$$Z = \frac{1}{h} V \left(2\pi m k_B T\right)^{3/2} \times \left[\left(2m\pi r^2 k_B T\right)^{1/2}\right]^2 2\pi \int_0^\pi \sin\theta e^{\beta\mu E\cos\theta} \mathrm{d}\theta$$

The last integral can be determined using the substitution $x = \cos\theta$ and $\mathrm{d}x = -\sin\theta \mathrm{d}\theta$,

$$\int_0^\pi \sin\theta e^{\beta\mu E\cos\theta} \mathrm{d}\theta = \int_{-1}^1 e^{\beta\mu E x} \mathrm{d}x = \frac{1}{\beta\mu E} \sinh\left(\beta\mu E\right)$$

Thus,

$$Z = Z_{\text{trans}} \times Z_{\text{rot}} \times \frac{1}{\beta\mu E} \sinh\left(\beta\mu E\right)$$

Now that Z in known, we can compute all thermodynamic properties of the system by taking partial derivatives of Z.

Instead of doing this, let us determine the average polarization of the gas in the z direction $\langle\mu^z\rangle$.

$$\langle\mu^z\rangle = \langle\mu\cos\theta\rangle = \frac{\int_0^{2\pi} \mathrm{d}\phi \int_0^\pi \mu \sin\theta\cos\theta e^{\beta\mu E\cos\theta} \mathrm{d}\theta}{\int_0^{2\pi} \mathrm{d}\phi \int_0^\pi \sin\theta e^{\beta\mu E\cos\theta} \mathrm{d}\theta}$$

Alternatively but completely equivalently this fraction can be reformulated as a derivative,

$$\langle\mu^z\rangle = \frac{1}{\beta}\frac{1}{Z}\frac{\mathrm{d}Z}{\mathrm{d}E} = \frac{1}{\beta}\frac{\partial}{\partial E}\ln Z$$

Evidently, this is more tractable that computing an integral. Therefore,

$$\ln Z = \ln\left(\sinh\left(\beta\mu E\right)\right) - \ln\left(\beta\mu E\right) + \text{const}$$

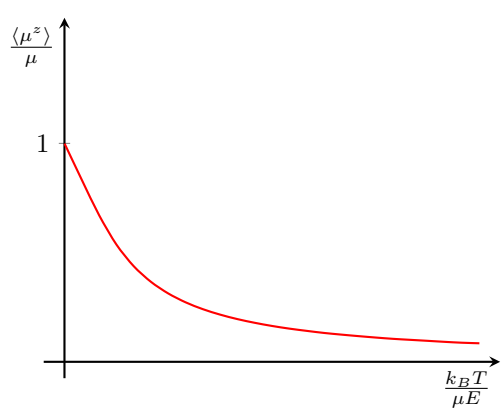
Which yields derivatives,

$$\frac{\partial}{\partial E}\ln Z = \frac{\mu\beta}{\sinh\left(\beta\mu E\right)}\cosh\left(\beta\mu E\right) - \frac{\beta\mu}{\beta\mu E}$$

$$\langle\mu^z\rangle = \frac{\mu\cosh\left(\beta\mu E\right)}{\sinh\left(\beta\mu E\right)} - \frac{\mu}{\beta\mu E}$$

$$\langle\mu^z\rangle = \mu\left\{\coth\left(\beta\mu E\right) - \frac{1}{\beta\mu E}\right\}$$

$$\frac{\langle\mu^z\rangle}{\mu} = \coth\left(\frac{\mu E}{k_B T}\right) - \frac{k_B T}{\mu E}$$



Typically, $\frac{\mu E}{k_B} \approx 0.5\,\text{mK}$.

Response Function: The response function is typically defined,

$$\chi_E = \frac{\partial \langle \mu^z \rangle}{\partial E} = \frac{\mu}{k_B T}$$

This is known as the curie behavior $\frac{\mu}{k_B T}$.

6.8.3 Polarizable Molecule

Consider a polarizable molecule with uni-axial direction $\hat{n}$. The molecule has electric dipole $\vec{\mu}_E$ with makes the Hamiltonian the interaction it has with an electric field $\vec{E}$,

$$H = -\vec{\mu}_E \cdot \vec{E}$$

Similarly to above, the position of the molecule is characterized by x, y, z and its orientation is characterized by θ, ϕ through $\hat{n}$. The dipole is induced by the electric field, and points in the direction of $\vec{n}$.

$$\vec{\mu}_e \propto \left(\vec{E} \cdot \hat{n}\right) \hat{n}$$

More explicitly using the polarizability α,

$$H = -\alpha \left(\vec{E} \cdot \hat{n}\right) \hat{n} \cdot \vec{E} = -\alpha \left(\vec{E} \cdot \hat{n}\right)^2 = -\alpha \vec{E} \cdot [\hat{n}\hat{n}] \cdot \vec{E}$$

Where the term $[\hat{n}\hat{n}]$ is a direct vector product or a rank two tensor (sometimes called dyadic).

$$(\hat{n}\hat{n}) = \begin{bmatrix} n_x & n_y & n_z \end{bmatrix} \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} n_x n_x & n_x n_y & n_x n_z \\ n_y n_x & n_y n_y & n_z n_z \\ n_z n_x & n_z n_y & n_z n_z \end{bmatrix}$$

Taking the angle between $\hat{n}$ and $\vec{E}$ to be θ, the Hamiltonian can be conveniently written

$$H = -\alpha E^2 \cos^2\theta = -\tilde{\alpha} E^2 \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)$$

Where $\tilde{\alpha} = \frac{2}{3}\alpha$. Which gives,

$$\left\langle \cos^2\theta \right\rangle = \int_0^\pi \cos^2\theta \sin\theta \mathrm{d}\theta \mathrm{d}\phi = 4\pi$$

Recognize that the $\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)$ term is the second Legendre polynomial $P_{20} \sim Y_{lm}(\theta, \phi); l = 1, m = 0$. The lesson here is that even for simple systems such as this, the mathematics can become complicated quite quickly. Also consider an electric quadrapole consider a charge distribution such as in N_2,

+ - - +

Where the positive charges represent the nuclei each nitrogen, and the negative charges are the electrons.What does this couple with? This would also involve spherical harmonics. Moving back to our dipole, let Q be the θ dependence on H,

$$Q = \frac{3}{2}\cos^2\theta - \frac{1}{2}$$

Q acts as the effective degree of freedom of this Hamiltonian. The average energy can be determined by knowing the expectation of Q,

$$\langle H \rangle = -\alpha E^2 \langle Q \rangle$$

$$\langle Q \rangle = \frac{1}{Z} \int_0^{2\pi} \mathrm{d}\phi \int_0^\pi Q(\theta) \sin\theta \mathrm{d}\theta e^{-\beta\left[-\alpha E^2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)\right]}$$

All contributions to the partition function are identical to the magnetic dipole case done previously except for the angular contribution

$$Z_{\theta,\phi} = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin\theta \mathrm{d}\theta e^{-\beta\left[-\alpha E^2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)\right]}$$

$$Z_{\theta,\phi} = 2\pi e^{\beta\alpha E^2/2} \int_0^{\pi} \sin\theta \mathrm{d}\theta e^{\beta\alpha E^2 \frac{3}{2}\cos^2\theta}$$

Again making the substitution $x = \cos\theta$,

$$Z_{\theta,\phi} = 2\pi e^{\beta\alpha E^2/2} \int_{-1}^{1} e^{gx^2} \mathrm{d}x$$

Where $g = \frac{3}{2}\alpha E^2 \beta$. The remaining integral is **not** a Gaussian integral. It is known as a **Dawson integral**.

$$D(g) = \int_{-1}^{1} e^{gx^2} \mathrm{d}x$$

Now that Z is known, all of thermodynamics follows by performing derivatives on $\ln Z$. Note that $D(g)$ can't be computed analytically, but the derivatives with respect to temperature and related quantities modify the integrand of $D(g)$ to where it *might* be computable.

As an exercise, calculate $\langle Q \rangle$ in the limit of large temperature. At high temperature, β and thus g is very small. *Hint*: Taylor series expand out $D(p)$ and integrate term by term to desired order.

6.8.4 Permanent Electric Dipole

Consider a permanent electric dipole with Hamiltonian composed of the coupling with the electric field and the rotation kinetic energies,

$$H = H_{\text{rot}} - \vec{\mu} \cdot \vec{E}$$

This dipole molecule will be considered a rigid rotator with moment of inertia I. As before, consider the molecule N_2 as an example molecule.

$$H = \frac{\hbar^2 \hat{L}^2}{2I} - \mu E \cos\theta$$

Where $\hat{L}$ is the angular momentum operator. Noting that the two components of the Hamiltonian here do not commute. This makes solving the equations of motion difficult because it requires a decoupling through diagonalization of the ODEs.

$$\hat{L} \sim \frac{\partial}{\partial\theta} \qquad \cos\theta \sim 1 - \frac{\theta^2}{2}$$

For a free rotor,

$$H_{\text{rot}}\Psi = E\Psi$$

Where $\Psi \sim Y_{lm}(\theta, \phi)$ the eigenfunctions are spherical harmonics. l in this case is quantized $l = 0, 1, 2, \ldots$. The eigenvalues or eigenenergies are determined by l alone,

$$E_l = \frac{\hbar^2}{2I} l(l+1) \qquad l = 0, 1, 2, \cdots$$

Where the m quantum number is bound by l,

$$m = 0, \pm 1, \pm 2, \ldots \pm l$$

As such, there are $(2l+1)$ states that have energy E_l. Summing over the states,

$$Z = \sum_{\text{states}} e^{-\beta\left(\frac{\hbar^2}{2I} l(l+1)\right)}$$

Or summing over the energies,

$$Z = \sum_l (2l+1)\, e^{-\beta\left(\frac{\hbar^2}{2I} l(l+1)\right)}$$

Note that although decreasing the temperature of the system makes lower energy states *more probable*, the number of microstates *increases* as l increases. These two effects are competing effects.
Including the electric field introduces a *Stark effect* analogous to the *Ziemann effect* in magnetic fields. The energy levels of the states are modified by f_{lm},

$$E_{lm} = \frac{\hbar^2}{2I} l\,(l+1) + f_{lm}\,(E)$$

This makes Z way harder to compute. Operationally, everything is the same as in previous examples, but it might require more mathematics. The Stark effect in rigid rotors in found in most Quantum textbooks.

6.8.5 Spin in Magnetic Field

For a permanent magnetic moment in a magnetic field, the Hamiltonian is,

$$H = -\vec{\mu}_m \cdot \vec{B} - g\mu_B \vec{J} \cdot \vec{B}$$

Where g_L is known as the **Landé factor**. For an electron $g_{e^-} \approx 2$. In general,

$$g_L = \frac{3}{2} + \frac{S\,(S+1) - L\,(L+1)}{2J\,(J+1)}$$

For the magnetic field oriented on the z-axis $\vec{B} = B\hat{z}$,

$$H = -g_L \mu_B J^z B$$

Where J^z takes on values $-J, -J+1, \ldots J-1, J$. J^z is known as the magnetic quantum number m_J.

$$Z = \sum_{m_J=-J}^{m_J=J} e^{\beta g_L \mu_B m_J B}$$

As a trick, let $x = e^{-\beta g_L \mu_B B}$

$$Z = \sum_{m=-J}^{m=J} x^{-m}$$

This is just the geometric series (slightly modified) with analytic sum,

$$\sum_{m=-J}^{J} x^{-m} = \frac{x^{-J}\left(x^{2J+1}+1\right)}{x-1}$$

Which gives the partition function the form

$$Z = \frac{\sinh\left(\left(1+\frac{1}{2J}\right)U\right)}{\sinh\left(\frac{U}{2J}\right)}$$

Where the term U is just $U = \beta g_L \mu_B B$. For $J = \frac{1}{2}$,

$$Z_{J=\frac{1}{2}} = 2\cosh\left(U\right)$$

6.8.6 Two Level Systems

Consider a general two level system with two possible energies 0 and ε with corresponding degeneracies g_0 and g_ε.

$$Z = \sum_{\text{states}} e^{-\beta E_{\text{states}}} = \sum_{\text{energies}} g e^{-\beta E}$$

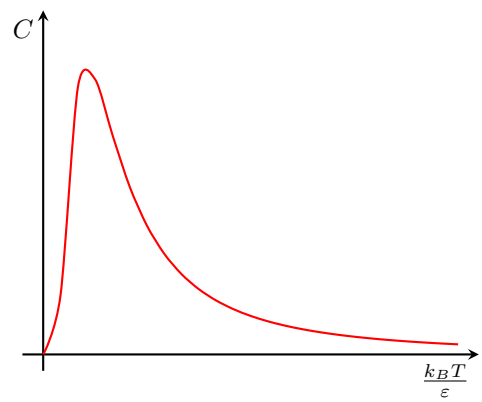
For two levels,

$$Z = g_0 e^{-\beta \cdot 0} + g_\varepsilon e^{-\beta \varepsilon}$$

As an exercise, one can show that the specific heat has behavior,

$$\frac{C}{k_B}(\beta E) = f\left(\frac{g_\varepsilon}{g_0}, e^{-\beta\varepsilon}\right)$$

Specific heat always has the general plot with increasing temperature,



This is known as a **Schottky anomaly**. For low temperature,

$$C \sim e^{-\beta/T}$$

and for high temperatures,

$$C \sim \frac{1}{T^2}$$

6.9 Energy Fluctuations

Now that we have looked at a number of examples using the canonical ensemble, we will begin to develop some more formalism involved with fluctuations of the canonical ensemble. Recall as a summary that the micro-canonical ensemble involved fixing N and V as well as the total energy E or the energy range $E \pm \Delta/2$. The idea was to determine the number of distinct accessible microstates Ω and then computing the entropy of the system, $S = k_B \ln \Omega$. Alternatively, the more realistic scenario is the canonical ensemble. Here we fixed the temperature T and allowed the energy to fluctuate. Then, the probability of finding a system in the ensemble in one of the states states (say E_r) is given by,

$$P_r = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}}$$

Where the denominator is a probability scaling factor and is known as the partition function Z.

$$S = -k_B \sum_r P_r \ln P_r$$

Note the similarity to *Shannon entropy* from information theory. Now what we want to do is to look at *energy fluctuations* in the canonical ensemble. In doing so, we aim to show the approximate equivalency of the two

approaches (microcanonical and canonical). We would **not** like our results for thermodynamically large systems (large N, T) to depend on whether or calculations are done in the microcanonical or canonical ensemble.

Question: For a fixed temperature, the energy is allowed to fluctuate. The question is, how much or how large can these energies fluctuate?

In order to define the fluctuations in energy, we need a notion of where it is fluctuating from. This will be the average energy,

$$U = \langle E \rangle = \frac{\sum_r E_r e^{-\beta E_r}}{\sum_r e^{-\beta E_r}}$$

This has been seen before. It is ensemble average of the energy.

$$\frac{\partial \langle E \rangle}{\partial \beta} = -\frac{\sum_r E_r^2 e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} + \frac{\left(\sum_r E_r e^{-\beta E_r}\right)^2}{\left(\sum_r e^{-\beta E_r}\right)^2}$$

These terms are just ensemble averages themselves,

$$\frac{\partial \langle E \rangle}{\partial \beta} = -\left\langle E^2 \right\rangle + \langle E \rangle^2$$

Note that this is the expectation of the energy deviation,

$$\left\langle (\Delta E)^2 \right\rangle = \left\langle (E - \langle E \rangle)^2 \right\rangle = \left\langle E^2 - 2E \langle E \rangle + \langle E \rangle^2 \right\rangle = \left\langle E^2 \right\rangle - 2 \langle E \rangle \langle E \rangle + \langle E \rangle^2 = \left\langle E^2 \right\rangle - \langle E \rangle^2$$

Therefore,

$$\left\langle (\Delta E)^2 \right\rangle = -\frac{\partial \langle E \rangle}{\partial \beta} = -\frac{\partial U}{\partial \beta}$$

This is just the **variance** of the energy. To take this further, use chain rule,

$$\frac{\partial U}{\partial \beta} = \frac{\partial T}{\partial \beta} \frac{\partial U}{\partial T}$$

With $\beta = 1/k_B T$ we have $T = 1/\beta k_B$,

$$\frac{\partial U}{\partial \beta} = -\frac{1}{\beta^2 k_B} \frac{\partial U}{\partial T} = -k_B T^2 \frac{\partial U}{\partial T}$$

$$\left\langle (\Delta E)^2 \right\rangle = k_B T^2 \frac{\partial U}{\partial T}$$

But the specific heat with respect to constant volume is just $C_V = \left(\frac{\partial U}{\partial T}\right)_V$,

$$\left\langle (\Delta E)^2 \right\rangle = k_B T^2 C_V$$

Therefore the **root-mean-squared** or the fluctuations is just,

$$\frac{\sqrt{\left\langle (\Delta E)^2 \right\rangle}}{\langle E \rangle} = \frac{\sqrt{k_B T^2 C_V}}{U}$$

But $C_V \propto N$ is an extensive variable. Also $U \propto N$ and T doesn't scale with temperature. For very large N, the root-mean-squared fluctuations in energy go as $1/\sqrt{N}$ and thus are negligible for large N. In this limit, the varying energy becomes ultimately zero. Therefore for $N \to \infty$, the canonical ensemble acts like the microcanonical ensemble.

Next we will consider the same example in the case of a *continuous* probability distribution instead of discrete. We now have the interpretation of $P(E)$ as a function of E. Therefore the probability of energy being in a range E to $E + \mathrm{d}E$ is given by $P(E)\,\mathrm{d}E$. Taking $g(E)$ as the density of states (see section 6.6).

$$P(E)\,\mathrm{d}E \propto g(E)\,e^{-\beta E}\mathrm{d}E$$

There are two energy-dependent factors influencing $P(E)$: $g(E)$ and $e^{-\beta E}$. Let's see whether $P(E)$ has an extremum at some point E^*.

$$\frac{\partial}{\partial E}\left(g(E)\,e^{-\beta E}\right|_{E=E^*} = 0$$

Instead of computing this directly, use logarithms. Since ln is monotonically increasing the maximum occurs in the same location E^*.

$$\ln\left(g(E)\,e^{-\beta E}\right) = \ln(g(E)) + \ln\left(e^{-\beta E}\right) = \ln(g(E)) + -\beta E$$

The maximization condition is then,

$$\frac{\partial}{\partial E}\left(\ln\left(g(E)\,e^{-\beta E}\right)\right|_{E=E^*} = 0$$

Which becomes,

$$\frac{\partial}{\partial E}\left(\ln(g(E))\right|_{E=E^*} = \beta \tag{6.10}$$

What is $\ln g(E)$? Recall that entropy is the logarithm of the number of accessible microstates. Since $g(E)$ is the number of microstates for a particular energy E,

$$S(E) = k_B \ln(g(E))$$

Subbing this into (6.10),

$$\frac{1}{k_B}\frac{\partial}{\partial E}\left(S(E)\right|_{E=E^*} = \beta$$

$$\frac{\partial}{\partial E}\left(S(E)\right|_{E=E^*} = \frac{1}{T}$$

But as we know from thermodynamics $\frac{\partial S(E)}{\partial E}\,|_{E=U} = \frac{1}{T}$. This shows us that the ensemble average U is equal to the extremum value E^*.

$$U = E^* \tag{6.11}$$

As such, U and E^* can be taken interchangably. Irrespective of the Hamiltonian of system, the most probable value of its energy is identical to its mean value. Depending on which is more advantageous we will use U or E^*.

Consider $\ln\left[g(E)\,e^{-\beta E}\right] = \ln P(E)$, and Taylor expand this around $E = U$. This can be done as we expect E to not vary wildly from U as it fluctuates.

$$\ln P(E) = \left[\ln e^{-\beta E} + \ln g\right]\Big|_{E=U} + \frac{1}{2}\frac{\partial^2}{\partial E^2}\left[\ln e^{-\beta E} + \ln g\right]\Big|_{E=U}(E-U)^2 + O\left((E-U)^3\right)$$

Ignoring higher order terms,

$$\ln P(E) = -\beta U + \frac{S}{k_B} + \frac{1}{2}\frac{\partial^2}{\partial E^2}\ln g\Big|_{E=U}(E-U)^2$$

The right hand term is given by,

$$\frac{\partial^2}{\partial E^2}\ln g = \frac{1}{k_B}\frac{\partial^2}{\partial E^2}S = \frac{1}{k_B}\frac{\partial}{\partial E}\frac{\partial S}{\partial E} = \frac{1}{k_B}\frac{\partial}{\partial E}\left(\frac{\partial T}{\partial E}\frac{\partial S}{\partial T}\right) = \frac{1}{k_B}\frac{\partial}{\partial E}\left(\frac{1}{T}\right) = -\frac{1}{k_B T^2}\frac{\partial T}{\partial E} = -\frac{1}{k_B T^2 C_V}$$

Which gives,

$$\ln P(E) = -\beta(U - TS) - \frac{1}{2}\frac{1}{k_B T^2 C_V}(E - U)^2$$

Notice the Helmholtz free energy $U - TS$. Looking at $P(E) \propto e^{-\beta E} g(E)$. Then we have,

$$P(E) \sim e^{-\beta(U-TS)} e^{-\frac{(E-U)^2}{2k_B T^2 C_V}}$$

We can then determine the partition function,

$$Z = Z(V,T) \propto e^{-\beta(U-TS)} \int_0^\infty \mathrm{d}E e^{-\frac{(E-U)^2}{2k_B T^2 C_V}}$$

Recall Gaussian integrals 2.4.3,

$$Z \approx e^{-\beta(U-TS)} \sqrt{2\pi k_B T^2 C_V}$$

Recall that $A \equiv -k_B T \ln Z$ so we have,

$$A = (U - TS) - \frac{1}{2} kT \ln\left(2\pi k_B T^2 C_V\right)$$

If $U - TS$ is the Helmholtz free energy, the extra piece is a problem. Consider the high N limit. The logarithm term becomes negligible compared to the leading U term. Hence we have $A \approx U - TS$ derived from stat mech. As such, the microcanonical and canonical ensemble approaches are similar.

In principle, we could derive all of thermodynamics from statistical mechanics. This is left to the reader as an exercise. Since we derived A from the canonical ensemble using S from microcanonical ensemble, the emergence of a consistent thermodynamic relationship establishes that the two approaches are essentially identical.

6.10 Equipartition & Virial Theorems

6.10.1 Equipartition

The Equipartition theorem is really useful to estimate or determine thermodynamic properties of systems by analyzing the degrees of freedom of the system. Consider a Hamiltonian $H(\vec{q}, \vec{p})$. Let's call x_i and x_j any two of the $6N$ generalized coordinates of the system (q, p). In the canonical ensemble we are interested in the expectation,

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle$$

Explicitly,

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = \frac{\int x_i \frac{\partial H}{\partial x_j} e^{-\beta H} \mathrm{d}^{3N} q \mathrm{d}^{3N} p}{\int e^{-\beta H} \mathrm{d}^{3N} q \mathrm{d}^{3N} p}$$

Just considering the numerator requires integration by parts,

$$\int x_i \frac{\partial H}{\partial x_j} e^{-\beta H} \mathrm{d}\omega = \int \mathrm{d}\omega_i \left\{ -\frac{1}{\beta} x_i e^{-\beta H} \Big|_{(x_j)_1}^{(x_j)_2} + \frac{1}{\beta} \int \frac{\partial x_i}{\partial x_j} e^{-\beta H} \mathrm{d}x_j \right\} \mathrm{d}\omega_j$$

Where $(x_j)_1$ and $(x_j)_2$ are the extremal values. $\mathrm{d}w_j$ is the phase space volume excluding $\mathrm{d}x_j$. The first term banishes because H becomes infinite at extremal values. For example, the potential becomes infinite of the kinetic energy becomes infinite at the boundaries. The second term has a δ function.

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}$$

Therefore the numerator becomes,

$$\frac{1}{\beta}\partial_{ij}\int e^{-\beta H}\mathrm{d}\omega_j$$

Thus,

$$\left\langle x_i\frac{\partial H}{\partial x_j}\right\rangle = \frac{\frac{1}{\beta}\partial_{ij}\int e^{-\beta H}\mathrm{d}\omega}{\int e^{-\beta H}\mathrm{d}\omega}$$

After canceling,

$$\left\langle x_i\frac{\partial H}{\partial x_j}\right\rangle = \delta_{ij}k_BT \tag{6.12}$$

Notice that this is independent of the Hamiltonian H.

Considerations: First for $x_i = x_j = p_i$, (6.12) becomes,

$$\left\langle p_i\frac{\partial H}{\partial p_i}\right\rangle = k_BT$$

From Hamiltonian's equations, $\dot{q}_i = \frac{\partial H}{\partial p_i}$,

$$\langle p_i\dot{q}_i\rangle = k_BT \tag{6.13}$$

Similarly using Hamiltonian's other equation $\dot{p}_i = -\frac{\partial H}{\partial q_i}$

$$\langle q_i\dot{p}_i\rangle = -k_BT \tag{6.14}$$

Then adding $i = 1 \to 3N$ for both p_i and q_i,

$$\left\langle \sum_i p_i\dot{q}_i\right\rangle = 3Nk_BT \tag{6.15}$$

$$-\left\langle \sum_i q_i\dot{p}_i\right\rangle = 3Nk_BT \tag{6.16}$$

In many cases, $H(q,p)$ in generalized quadratic form of the coordinates $\{q_i, p_i\}$. A **canonical transformation** can be used to diagonalized H into the following form,

$$H = \sum_j A_jP_j^2 + \sum_j B_jQ_j^2$$

Where A_j, B_j are constants and P_i, Q_i are the corresponding to the transformed coordinates. Under this transformation,

$$\frac{\partial H}{\partial P_j} = A_j2P_j \quad \frac{\partial H}{\partial Q_j} = B_j2Q_j$$

Summing over all j,

$$\sum_j\left(P_j\frac{\partial H}{\partial P_j} + Q_j\frac{\partial H}{\partial Q_j}\right) = \sum_j\left(P_jA_J2P_j + Q_jB_j2Q_j\right) = 2\sum_j\left(A_JP_j^2 + B_jQ_j^2\right) = 2H$$

From equation (6.13) and (6.14),

$$\langle H\rangle = \frac{1}{2}\left(f_A + f_B\right)k_BT = \frac{1}{2}fk_BT$$

Where f_A and f_B sum together to give f are the number of A_i an B_j that are non-zero in the Hamiltonian. *This is a very important result.* Each harmonic (quadratic) term in the Hamiltonian contributes exactly $\frac{1}{2}k_BT$ toward to the energy. As a consequence, each degree of freedom contributes $\frac{1}{2}k_B$ towards the specific heat C_V. This is the classical version of the theorem of Equipartition theorem.

Note that at really low temperatures, this breaks down. Consider *phonons* which are lattice vibrations, each behaving as quantum harmonic oscillators. At low temperatures in certain systems, the energy can become quantized and there are corrections to the above result that can be seen experimentally.

6.10.2 Virial Theorem

Consider equation (6.16),

$$\left\langle \sum_i q_i \dot{p}_i \right\rangle = -3Nk_BT$$

The Virial theorem introduces the notion of generalized force,

$$F_{i,\text{generalized}} = \dot{p}_i$$

Thus (6.16) is the expectation value of the product of coordinate q_i with the force acting on it. It is called the **Virial**.

$$\mathcal{V} = \left\langle \sum_i q_i \dot{p}_i \right\rangle$$

The Virial equation is simply then,

$$\mathcal{V} = -3Nk_BT \tag{6.17}$$

More commonly, it is written as a force integral.

Example: Consider an ideal gas that is dilute enough that it is not interacting with itself but it only interacting with the walls of the container it is in for example. The Virial for this system is thus,

$$\mathcal{V}_0 = \left\langle \sum_i q_i F_i \right\rangle_0 = -P \oiint_S \vec{r} \cdot \mathrm{d}\vec{s}$$

Where the '0' is just to indicate that it is not interacting. Here P is the pressure (uniform) pointing inward on a ball S (hence the minus sign), $\mathrm{d}\vec{s}$ is the area normal pointing outward, and $\vec{r}$ is the location on the surface of the ball S. Applying the divergence theorem, we can relate the surface integral to a volume integral,

$$\oiint_S \vec{r} \cdot \mathrm{d}\vec{s} = \iiint_V \vec{\nabla} \cdot \vec{r} \mathrm{d}V$$

Remember that $\vec{\nabla} \cdot \vec{r}$ is just,

$$\vec{\nabla} \cdot \vec{r} = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial z}\hat{y} + \frac{\partial}{\partial z}\hat{z} \right) \cdot (x\hat{x} + y\hat{y} + z\hat{z}) = 1 + 1 + 1 = 3$$

Thus the Virial is just,

$$\mathcal{V}_0 = -3PV$$

Where V is the volume contained inside the bounding surface S. From this using (6.17),

$$PV = Nk_BT \implies PV = nRT$$

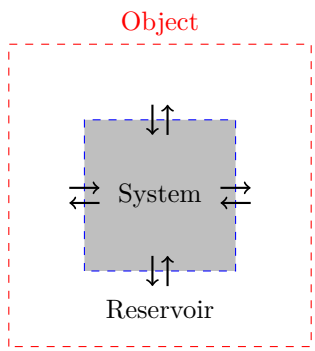
Which is just the ideal gas law. Note the this only works for an ideal gas. A more complicated gas with internal fores would have a more complicated integral for $\mathcal{V}$. From this, the equation of state would become,

$$\frac{P}{\frac{N}{V}k_BT} = 1 + \frac{1}{dNk_BT} \left\langle \sum_{i<j} \vec{F}\left(\vec{f}_{ij}\right) \cdot \vec{r}_{ij} \right\rangle$$

Where d is just the spacing dimension. This is known as the **virial equation of state**. It is typically solved computationally. One can make a career in Chemistry solving the equation of state for different systems.

7 Grand Canonical Ensemble

Recall that microcanonical ensemble was constructed with fixed energy E, the canonical ensemble which has fixed T and energy was allowed to be exchange energy E with the reservoir. The **grand canonical ensemble (GCE)** is simply a generalization of the canonical ensemble where the number of particles N *and* the total energy E is allowed to be shared via fluctuations with the reservoir. Similar to before, we hope to recover thermodynamic relations using this formalism. The derivations of the grand canonical ensemble follow almost precisely the same as it did in the case of the canonical ensemble, with a few caveats.



Now there the interface between the system and reservoir is some sort of permeable membrane that allows for a flow of particles between the system and reservoir.

Observation: Recall that making the switch from the microcanonical ensemble to canonical ensemble made life easier in terms of a procedures/computations because we relaxed the constraint that the system energy E was fixed. The Grand canonical ensemble further relaxes the constraint of N being fixed. This too will further make the computations even *simpler*.

As a first example, consider that a slab of N_2 just above a substrate where there is an exchange of material between the N_2 and the substrate as well as the energy of the two materials as the system evolves. This is an example of a grand canonical ensemble.

Alternatively, an example of a quantum system with Hamiltonian,

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \frac{c}{2}\left(x_{i+1} - x_i\right)^2$$

Can be mapped (with algebra) to the equivalent problem of a system of harmonic oscillators,

$$H = \sum_k = \hbar\omega\left(k\right)\left\{a_k^\dagger a_k + \frac{1}{2}\right\}$$

Where $N_k = a_k^\dagger a_k$ is known as the **number operator**. This number is typically not conserved in most contexts, which further motivates the use of the grand canonical ensemble as lots of systems share the properties of fluctuating energy E and number of particles N.

7.1 Derivation

We will label the number of particles in the in the reservoir $N^{(r)}$ and the energy in the reservoir $E^{(r)}$. The number of particles,energy in the system is $N^{(s)}, E^{(s)}$. Assuming that $N^{(r)} \gg N^{(s)}$ and $E^{(r)} \gg E^{(s)}$ because the reservoir is considered so large. We will assume the total number of particles N and the total energy E to be constant with,

$$N = N^{(r)} + N^{(s)} \qquad E = E^{(r)} + E^{(s)}$$

Note that the fraction of particles in the system has the form,

$$\frac{N^{(s)}}{N} = \left(1 - \frac{N^{(r)}}{N}\right) \ll 1$$

And the fraction of energy in the system is,

$$\frac{E^{(s)}}{E} = \left(1 - \frac{E^{(r)}}{E}\right) \ll 1$$

Similar to the canonical case, if the system is found to be the state $\left(N_i^{(s)}, E_j^{(s)}\right)$, the probability is proportional to the number of microstates, $\Omega^{(r)}\left(N_i^{(r)}, E_j^{(r)}\right)$ that the reservoir has access to,

$$P\left(N_i^{(r)}, E_j^{(r)}\right) \propto \Omega^{(r)}\left(N_i^{(r)}, E_j^{(r)}\right)$$

Well will write this as the exponent of a logarithm to allow for the Taylor series to converge ($P \propto \exp\left(\ln \Omega^{(r)}\right)$). Expanding around the total energy and number of particles E, N to first order,

$$\ln \Omega^{(r)}\left(N^{(r)}, E^{(r)}\right) \approx \ln \Omega^{(r)}(N, E) + \left.\frac{\partial \ln \Omega^{(r)}}{\partial N^{(r)}}\right|_{N^{(r)}=N}\left(N^{(r)} - N\right) + \left.\frac{\partial \ln \Omega^{(r)}}{\partial E^{(r)}}\right|_{E^{(r)}=E}\left(E^{(r)} - E\right)$$

Noting that $E^{(r)} - E = -E^{(s)}$ and $S = k_B \ln \Omega$, the derivative of S with respect to N is given by the chemical potential μ,

$$\frac{\partial S}{\partial N} = -\frac{\mu}{T} \qquad \frac{\partial S}{\partial E} = \frac{1}{T}$$

We obtain,

$$\ln \Omega^{(r)}\left(N^{(r)}, E^{(r)}\right) \approx \ln \Omega^{(r)}(N, E) + \frac{\mu^{(r)}}{k_B T^{(r)}} N^{(s)} - \frac{1}{k_B T^{(r)}} E^{(s)}$$

Here $\mu^{(r)} = \mu^{(s)} = \mu$ and $T^{(r)} = T^{(s)} = T$ since we are considering being in chemical and thermodynamic equilibrium respectively. As such we have obtained that,

$$P_{ij}\left(N_i^{(s)}, E_j^{(s)}\right) \propto \exp\left[-\alpha N_i^{(s)} - \beta E_j^{(s)}\right]$$

Where the constants are defined as,

$$\alpha \equiv -\frac{\mu}{k_B T} \quad \beta = \frac{1}{k_B T} \tag{7.1}$$

Normalizing the probability removes the proportionality. Summing over all accessible states of the system that are allowed $\{i, j\}$,

$$P_{ij}\left(N_i^{(s)}, E_j^{(s)}\right) = \frac{\exp\left[-\alpha N_i^{(s)} - \beta E_j^{(s)}\right]}{\sum_{i,j} \exp\left[-\alpha N_i^{(s)} - \beta E_j^{(s)}\right]}$$

Or dropping the notation of the system,

$$P_{ij}(N_i, E_j) = \frac{\exp\left[-\alpha N_i - \beta E_j\right]}{\sum_{i,j} \exp\left[-\alpha N_i - \beta E_j\right]} \tag{7.2}$$

7.2 Comparison to Canonical Ensemble

We will now consider a system in the grand canonical **ensemble**. Imagine an ensemble of $\mathcal{N}$ identical copies of the system of interest. Together they will share a total energy $\mathcal{N}\bar{E}$ and will also share a total number of particles $\mathcal{N}\bar{N}$. As a manner of notation, we will call $n_{i,j}$ the number of systems that have number of particles N_i and energy E_j. Then we obtain the following relations,

$$\sum_{i,j} n_{i,j} = \mathcal{N} \tag{7.3}$$

$$\sum_{i,j} n_{i,j} N_i = \mathcal{N}\bar{N} \tag{7.4}$$

$$\sum_{i,j} n_{i,j} E_j = \mathcal{N}\bar{E} \tag{7.5}$$

Remark: As a comparison to the canonical ensemble, the new constraint associated with the GCE is the constraint of total number of particles (7.4).

Any set $\{n_{i,j}\}$ of the numbers $n_{i,j}$ that satisfy the above constraints ((7.3), (7.4), (7.5)) represents the possible *modes of distribution* of particles among the ensemble. Any such mode of the distribution can be realized in $W\{n_{i,j}\}$,

$$W\{n_{i,j}\} = \frac{\mathcal{N}!}{\prod_{i,j}(n_{i,j}!)}$$

Where $\mathcal{N}!$ is the number of permutations associated with the copies of the system and $\prod_{i,j}(n_{i,j}!)$ are the internal number of permutations that maintain the microstates of the system. We will again use Lagrange multipliers in order to extremize this distribution to find the *most probable* mode. Just as before the derivation employs that $N \gg 1$. From this one obtains,

$$\frac{n_{i,j}}{\mathcal{N}} = \frac{e^{-\alpha N_i - \beta E_j}}{\sum_{i,j} e^{-\alpha N_i - \beta E_j}}$$

Which is identical to (7.2). This is fitting as $n_{i,j}$ is the number of systems that happen to have N_i particles and E_j energy, and $\mathcal{N}$ is the total number of systems. One can also show that in the thermodynamic limit $n_{\text{most prob}} = \langle n_{i,j} \rangle$. This is precisely the same idea that was invoked to prove (6.11). We will use this result moving forward.

$$n_{i,j} = \langle n_{i,j} \rangle$$

Using this we can compute the average number of particles,

$$\bar{N} = \frac{\sum_{i,j} N_i e^{-\alpha N_i - \beta E_j}}{\sum_{i,j} e^{-\alpha N_i - \beta E_j}}$$

Using the derivative of the logarithm trick performed many times before,

$$\bar{N} = -\frac{\partial}{\partial \alpha} \ln \left\{ \sum_{i,j} e^{-\alpha N_i - \beta E_j} \right\}$$

Similarly, the average energy of the system is determined by a derivative of β.

$$\bar{E} = -\frac{\partial}{\partial \beta} \ln \left\{ \sum_{i,j} e^{-\alpha N_i - \beta E_j} \right\}$$

Just as in the canonical ensemble, $\beta = 1/k_B T$; however what is the physical interpretation/significance of α? Note that this was previously motivated with (7.1) above, but we will re-examine this again now.

7.3 Physical Significance of α, β

For convenience of notation we will define the non-physical quantity η as,

$$\eta \equiv \ln \left(\sum_{i,j} e^{-\alpha N_i - \beta N_j} \right) \tag{7.6}$$

Thermodynamically, α, β, E_j can vary depending on volume. To digest or interpret α, β, we first need to compute $\mathrm{d}\eta$.

$$\mathrm{d}\eta = -\bar{N}\mathrm{d}\alpha - \bar{E}\mathrm{d}\beta - \frac{\beta}{\mathcal{N}} \sum_{i,j} \langle n_{i,j} \rangle \, \mathrm{d}E_j \tag{7.7}$$

Now consider the quantity $\eta + \bar{N}\alpha + \bar{E}\beta$ and its differential,

$$\mathrm{d}\left(\eta + \bar{N}\alpha + \bar{E}\beta\right)$$

By chain rule and (7.7) this can be expressed,

$$\mathrm{d}\left(\eta + \bar{N}\alpha + \bar{E}\beta\right) = \beta \left\{ \frac{\alpha}{\beta}\mathrm{d}\bar{N} + \mathrm{d}\bar{E} - \frac{1}{\mathcal{N}} \sum_{i,j} \langle n_{i,j} \rangle \, \mathrm{d}E_j \right\}$$

Recall from the first law of thermodynamics,

$$\mathrm{d}\bar{E} = \mathrm{d}Q - \mathrm{d}W + \mu \mathrm{d}\bar{N}$$

Comparing term by term,

$$\mu = -\frac{\mu}{\beta} \qquad \mathrm{d}W = -\frac{1}{\mathcal{N}} \sum_{i,j} \langle n_{i,j} \rangle \, \mathrm{d}E_j \qquad \mathrm{d}\left(\eta + \bar{N}\alpha + \bar{E}\beta\right) = \beta \mathrm{d}Q$$

Therefore we obtain that,

$$\alpha = -\frac{\mu}{k_B T}$$

7.4 Grand Canonical Partition Function

Just as found for (7.1). Further examination of the incremental heat,

$$\mathrm{d}\left(\eta + \bar{N}\alpha + \bar{E}\beta\right) = \beta \mathrm{d}Q = \frac{\mathrm{d}S}{k_B}$$

Therefore we have the entropy of the system in terms of $\eta, \alpha, \beta, \bar{N}$ and $\bar{E}$,

$$S = k_B \left(\eta + \bar{N}\alpha + \bar{E}\beta\right)$$

Or rearranged for η,

$$\eta = \frac{S}{k_B} + \frac{\mu \bar{N}}{k_B T} - \frac{\bar{E}}{k_B T} = \frac{ST + \mu\bar{N} - \bar{E}}{k_B T}$$

Recall the Gibb's free energy $G = \bar{E} - TS + PV = \mu\bar{N}$ and the original definition for η (7.8),

$$\eta = \ln\left(\sum_{i,j} e^{-\alpha N_i - \beta N_j}\right) = \frac{PV}{k_B T} \tag{7.8}$$

So η has a physical interpretation after all.

Note that $e^{-\alpha N_i} = \left(e^{-\alpha}\right)^{N_i}$ has the inner term often denoted with a lower case z,

$$z = e^{-\alpha} = e^{\frac{\mu}{k_B T}}$$

Where $z \neq Z$ is **not** the partition function. It is known as the **fugacity**.

Clearly η given by (7.8) plays a crucial role. It can be rewritten as,

$$\eta = \ln\left(\sum_{i,j} e^{-\alpha N_i} e^{-\beta E_j}\right) = \ln\left(\sum_i e^{-\alpha N_i} \sum_j e^{-\beta E_j}\right)$$

The second sum is given by the possible energies of the system E_j. Instead of summing over these arbitrary indices, we can replace this with summing over the states s with E_s depending possibly on the number of particles N_i,

$$\eta = \ln\left(\sum_i e^{-\alpha N_i} \sum_s e^{-\beta E_s(N_i)}\right)$$

Recalling the canonical partition function $Z = \sum_s e^{-\beta E_s}$. Since Z can now depend on the number of particles, we will denote it Z_{N_i}.

$$\eta = \ln\left(\sum_i z^{N_i} Z_{N_i}(V,T)\right)$$

Where Z_{N_i} is the canonical partition function for a given number of particles N_i and it depends on the volume and temperatures of the system. It is often the case that the canonical partition function Z has its label replaced with a Q.

$$\eta = \ln\left(\sum_i z^{N_i} Q_{N_i}(V,T)\right)$$

The term in the logarithm is known as the **grand canonical partition function**.

$$Z_{GC} = \sum_i z^{N_i} Q_{N_i}(V,T)$$

How does it work? Recall that $Q_{N_i}(V,T) = \sum_s e^{-\beta E_s}$ is identically the canonical partition function as done before. Once one obtains this for each possible number of particles N_i, the grand canonical partition function is simply a weighted sum of each of these Q_{N_i}'s. Alternatively, this modification can be viewed as a shifting of energy for the corresponding number of particles,

$$\sum_i e^{-\alpha N_i} \sum_s e^{-\beta E_s(N_i)} = \sum_i \sum_s e^{-\beta(E_s - \mu N_i)}$$

Example: Consider an illuminating example of a system of 3 spins ($S_i = \pm 1, \forall i$) with Hamiltonian,

$$H = J\left(S_1 S_2 + S_2 S_3 + S_3 S_1\right)$$

Where two states have energy $E = 3J$ and 6 states have energy $E = -J$. Letting s run through the states,

$$\sum_{s=1}^{s=8} e^{-\beta E_s} = 2e^{-\beta(3J)} + 6e^{-\beta(-J)}$$

Thus it is more efficient computationally to recognize these repeated values and include the degeneracies in the formula.

Returning to the generic case, instead of summing over the index i,

$$\sum_i e^{\beta \mu N_i} \to \sum_{N_i=0}^{\infty} e^{\beta \mu N_i}$$

We can sum over the actual number of particles N_i. This is precisely analogous to summing over the energies instead of summing over the states as throughly discussed in section 6.6. Altogether,

$$Z_{GC} = \sum_{N_i=0}^{\infty} e^{\beta \mu N_i} \left(\sum_E{}' g(E) e^{-\beta E}\right)_{N_i}$$

Note that there is no corresponding density of states for the number of particles $g(N)$. $g(E)$ corresponds to the number of states that happen to have the same energy E. However if the number of particles are the same, it corresponds to precisely the same state without degeneracies. Moreover notice that the second sum in this formula has a prime to make clear that the sum is made using the constraint that the number of particles is fixed. It behaves, for fixed N, precisely as the *canonical* ensemble.

$$Z_{GC} = \sum_{N=0}^{\infty} z^N Z_C(N)$$

8 Ideal Quantum Gases

We will begin now to discuss ideal quantum gases. Just as in the ideal gas case, we may restrict our setting to just deal with **non-interacting** particles. As such out general Hamiltonian,

$$H = H_{\text{kin}} + \sum_{i>j} V_{ij}$$

May be simply written,

$$H = H_{\text{kin}}$$

This switch can be motivated for the regime of $\beta V = \frac{V}{k_B T} \ll 1$ or in the case of *dilute* gases. Since the particles are quantum mechanical, they are indistinguishable. Examples of these gases include,

- fermions: electrons, protons, neutrons, $S = \frac{1}{2}\hbar$
  - Governed by the Pauli Exclusion principle
- bosons: collective excitations, composite objects, photon, phonon

An example of a composite particle is a hydrogen atom. Under the correct description, a hydrogen atom is composed of two fermions (electron, proton) and their combination can behave as a boson with spin ($S = \frac{1}{2}\hbar$).

8.1 Microcanonical Ensemble For IQG

To begin, we should first look at ideal quantum gases from the perspective of fixed N, V, E (microcanonical ensemble). A non-quantum ideal gas has a fixed energy of (for each particle at temperature T),

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

So how can we modify this expression to deal with *quantum* ideal gases? Recall the results for a particle in a box of length L with Hamiltonian,

$$H_{\text{kin}} = \frac{\hbar^2}{2m} \vec{\nabla}^2$$

With TISE,

$$\frac{\hbar^2}{2m} \vec{\nabla}^2 \psi = E\psi$$

Where states ψ are standing waves within the box. The difference between fermions and bosons is that no two fermions can share exactly the same quantum state while bosons can. For example, only two (one with spin up, one with spin down) electrons can occupy the lowest energy state standing wave $\psi_1 = \sin\left(\frac{\pi x}{L}\right)$. Knowing this, it is clear that the number of configurations for fermions is considerably lower than for bosons.

What we will do is look at an energy spectrum with energy levels ε_i with,

$$\varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_n$$

Where there are g_i states that share that eigenenergy. For each energy level ε_i, we can have n_i particles with that energy. Therefore to characterize the i-th cell, we need $(n_i; \varepsilon_i; g_i)$. Under this picture we have constraints:

1. number constraint $N = \sum_i n_i$
2. energy constraint $E = \sum_i n_i \varepsilon_i$

Now under these constraints, our goal is to compute the number of microstates Ω accessible to the system for fixed N, V, E.

$$\Omega(N, V, E) = \sum_{\{n_i\}}{}' \mathcal{W}\{n_i\}$$

Where $\mathcal{W}\{n_i\}$ is the number of microstates associated with a distribution (of particles) $\{n_i\}$ consistent with constraints $1, 2$. Moreover the prime on the summation symbol Σ' indicates the summation over all these "modes of distribution" that must fulfill the constraints above. What is $\mathcal{W}\{n_i\}$?

$$\mathcal{W}\{n_i\} = \prod_i w(i)$$

Where $w(i)$ is the number of distinct ways in which n_i indistinguishable particles can be distributed among the (distinguishable) g_i levels/states in the i-th cell. The form of $w(i)$ will of course depend on the type of particle in question.

8.2 Statistical Distributions

Bosons: A system of bosons will be called a **Bose-Einstein statistic (BE)**. This is just given by the typical stars and bars result,

$$w_{\text{BE}}(i) = \frac{(n_i + g_i - 1)!}{n_i!\,(g_i - 1)!}$$

Thus making $\mathcal{W}\{n_i\}$ given to be,

$$\mathcal{W}_{\text{BE}}\{n_i\} = \prod_i \frac{(n_i + g_i - 1)!}{n_i!\,(g_i - 1)!}$$

Fermions: By contrast, the same result for a **Fermi Dirac statistic (FD)** is quite different. Any particular fermion has g_i possible states in can occupy for each energy cell ε_i. However, the second particle has less choice. The second particle will have access to $g_i - 1$ states since one is already occupied by the first one. This process continues until the n_i-th particle picks a state,

$$g_i(g_i - 1)(g_i - 2)\cdots(g_i - (n_i - 1)) \quad n_i \leq g_i$$

One can write this more compactly as the difference between two factorials,

$$\frac{g_i!}{(g_i - n_i)!}$$

This gives the number of ways, one can distribute n_i particles into g_i levels with no more than 1 particle per level. However since the particles are indistinguishable we have over-counted by the number of permutations of the n_i particles themselves,

$$w_{\text{FD}} = \frac{1}{n_i!}\frac{g_i!}{(g_i - n_i)!}$$

Making,

$$\mathcal{W}_{\text{FD}}\{n_i\} = \prod_i \frac{1}{n_i!}\frac{g_i!}{(g_i - n_i)!}$$

Maxwell-Boltzmann distribution: Assuming first that the particles are indistinguishable, for the distribution set $\{n_i\}$ we can ask: How many ways can N particles be distributed into n_1 particles in cell 1 and n_2 particles in cell 2 and n_i particles in cell i? This is given by the multinomial distribution,

$$\frac{N!}{n_1! \cdot n_2! \cdots n_i!} = \frac{N!}{\prod_i n_i!}$$

Now with g_i states we can have $g_i^{n_i}$ configurations of particle distribution of n_i particles among g_i states in cell i. Furthermore removing the distinguishability of the N particles requires division by $N!$,

$$\mathcal{W}_{\text{MB}}\{n_i\} = \frac{1}{n_i!}\prod_i g_i^{n_i}$$

For supplementary materials, see Sertagg-Wan Wylen.

8.3 Entropy & Distributions

By with any of these distributions, we can determine the entropy of the system in the usual way,

$$S(N,V,E) = k_B \ln \Omega(N,V,E) = k_B \ln\left(\sum_{\{n_i\}}{}' \mathcal{W}\{n_i\}\right)$$

In the thermodynamic limit $N \gg 1, n_i \gg 1, g_i \gg 1$, etc, the sum $\sum_{\{n_i\}}{}'\mathcal{W}\{n_i\}$ is very well approximated by $\mathcal{W}\{n_i^*\}$ where $\{n_i^*\}$ is the distribution of $\{n_i\}$ that will maximize $\mathcal{W}\{n_i\}$. As such we obtain the approximation,

$$S(N,V,E) = k_B \ln \mathcal{W}\{n_i^*\}$$

Now we have to maximize $\mathcal{W}$ whether it be $\mathcal{W}_{\text{BE}}$, $\mathcal{W}_{\text{FD}}$ or $\mathcal{W}_{\text{MB}}$. Use Lagrange multipliers again and maximize the logarithm of $\mathcal{W}$,

$$\mathrm{d}\ln \mathcal{W}\{n_i\} - \left\{\alpha \sum_i \mathrm{d}n_i + \beta \sum_i \varepsilon_i \mathrm{d}n_i\right\} = 0$$

Plugging in each of the distributions above $\mathcal{W}$,

$$n_i^* = \frac{g_i}{e^{\alpha+\beta\varepsilon_i} + a}$$

Where we have,

Distribution	a
$\mathcal{W}_{\text{BE}}$	-1
$\mathcal{W}_{\text{FD}}$	$+1$
$\mathcal{W}_{\text{MB}}$	0

Remark: One thing to notice is that n_i^* is proportional to g_i. As such, $\frac{n_i^*}{g_i}$ can be interpreted as the most probable number of particles per energy level in the i-th cell. This also suggests we can interpret $\frac{n_i^*}{g_i}$ as the most probable number of particles *even if* we had a single energy level/state ε_i.

Comment: This last statement can be obtained rigorously within the grand canonical ensemble where each single particle energy state is *considered individually* without grouping them in cells with a number g_i much greater than 1 ($g_i \gg 1$) per cell.

What about other ensembles like canonical and grand canonical?

8.4 Ideal gas in other Q.M. ensembles

Canonical: First we will look at the case of the canonical ensemble approach to ideal quantum gases. We will restrict our perspective to gases "free to move"in some volume V subject to their corresponding statistics outlined in section (8.2). We will therefore have in mind energy for states E_s and energy states for single particles ε. The reason for this distinction is that for single, isolated particle alone ($N = 1$),

$$H\psi = -\frac{\hbar^2}{2m}\nabla^2\psi = \varepsilon\psi$$

Since there is only one particle, we are ignoring the interaction potential V_{ij} (mind you we are ignoring them anyway for *ideal* quantum gases). Therefore for particles in a box we have the usual,

$$\varepsilon = \frac{\hbar^2}{2m}\left(k_x^2 + k_y^2 + k_z^2\right) = \sum_{m=1}^{d} \frac{\hbar^2}{2m} k_m^2$$

With wave function state with the general form,

$$\psi\left(x, y, z\right) \sim \frac{1}{\sqrt{L}} e^{i\left(\vec{k}\cdot\vec{r}\right)}$$

Where $\vec{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$. Periodic boundary conditions yield a quantization of energies,

$$\psi\left(x, y, z\right) = \psi\left(x + L, y, z\right) \implies e^{ik_xL} = 1$$

Thus k_xL must be multiple of 2π,

$$k_xL = 0, \pm 2\pi, \pm 4\pi \ldots$$

Notice that $k_x = 0$ would be allowed under the periodic boundary conditions and corresponds to a uniform wavefunction. The same relationships hold for k_y, k_z. So in order to determine the canonical partition function we can utilize k_m to identify our states,

$$Z_N\left(V, T\right) = {\sum_s}' e^{-\beta E_s}$$

Since we have that the energy of the system is $E_s = \sum_\varepsilon n_\varepsilon \varepsilon$ (n_ε is the number of particles with energy ε) we can sum over the energies instead of summing over the states,

$$Z_N\left(V, T\right) = {\sum_{\{n_\varepsilon\}}}' F\left\{n_\varepsilon\right\} e^{-\beta \sum_\varepsilon n_\varepsilon \varepsilon}$$

Where $F\left\{n_\varepsilon\right\}$ is a statistical weight appropriate to the relevant distribution. It characterizes the weight provided for each energy ε. For Bosons, these weights are all 1. The prime on the summation used to indicate that the sum is performed under the constraint that the number of particles is fixed,

$$\sum_\varepsilon n_\varepsilon = N$$

The total energy is also given by the sum of the particles,

$$\sum_\varepsilon n_\varepsilon \varepsilon = E$$

But we don't care about this "constraint" because the total energy is allowed to vary in the canonical ensemble. Therefore we have that,

$$Z_N\left(V, T\right) = {\sum_{\{n_\varepsilon\}}}' F\left\{n_\varepsilon\right\} e^{-\beta \sum_\varepsilon n_\varepsilon \varepsilon}$$

With weights given by,

Distribution	$\mathbf{F\left\{n_\varepsilon\right\}}$
B.E.	1
F.D.	1 if all $n_\varepsilon = 0, 1$ 0 otherwise
M.B.	$\prod_\varepsilon \frac{1}{n_\varepsilon!}$

Where the weighting for the Maxwell-Boltzmann distribution is due to the indistinguishability for each set n_ε. As it turns out, only the Maxwell-Boltzmann case can be done straight forwardly,

$$Z_N(V,T) = \sum_{\{n_\varepsilon\}}{}' \left(\prod_\varepsilon \frac{1}{n_\varepsilon!}\right) e^{-\beta \sum_\varepsilon n_\varepsilon \varepsilon}$$

Using multiplicative properties of exponentials,

$$Z_N(V,T) = \sum_{\{n_\varepsilon\}}{}' \left(\prod_\varepsilon \frac{1}{n_\varepsilon!}\right) \left(\prod_\varepsilon e^{-\beta\varepsilon}\right)^{n_\varepsilon}$$

Under the "prime" constraint that $\sum_\varepsilon n_\varepsilon = N$ it can be shown using the **Multinomial theorem** (generalization of the binomial theorem),

$$Z_N(V,T) = \frac{1}{N!}\left(\sum_\varepsilon e^{-\beta\varepsilon}\right)^N$$

Treating the inner term as the partition function for a single state,

$$Z_N(V,T) = \frac{1}{N!}\left(Z_1(V,T)\right)^N \tag{8.1}$$

Again we can solve for the inner term (the single particle partition function $Z_1(V,T)$),

$$\sum_\varepsilon e^{-\beta\varepsilon} = \sum_{k_x,k_y,k_z} e^{-\beta\frac{\hbar^2 k^2}{2m}}$$

For high temperature we can make the approximation that the discrete sum becomes an integral,

$$\sum_{k_m} \to \int \frac{\mathrm{d}k_m}{L/2\pi}$$

Where $L/2\pi$ is introduced to normalize $\mathrm{d}k_m$ to the correct units. Then we get by utilizing spherical coordinates on k_x, k_y, k_z,

$$Z_1(V,T) = \left(\frac{L}{2\pi}\right)^3 \iiint \mathrm{d}k_x \mathrm{d}k_y \mathrm{d}k_z e^{-\beta\frac{\hbar^2 k^2}{2m}} = \frac{V}{8\pi^3}\int_0^\infty \left(4\pi k^2\right) e^{-\beta\frac{\hbar^2 k^2}{2m}} \mathrm{d}k$$

Gaussian integrals our old friend! The infinite bounds in the integral are a result of the particle having unbounded (by positive) k. We can then write the single particle partition function,

$$Z_1(V,T) = \frac{V}{\lambda_T^3} \qquad \lambda_T = \frac{h}{(2\pi m k_B T)^{1/2}}$$

Making the combined partition function is given by (8.1),

$$Z_N(V,T) = \left(\frac{V}{\lambda_T^3}\right)^N \frac{1}{N!}$$

This is the result for the canonical ensemble. The Grand canonical ensemble partition function immediately follows,

$$Z(z,V,T) = \sum_{N=0}^\infty z^N Q_N(V,T)$$

Where z is $z = e^{-\beta\mu}$ the fugacity as seen before. Notice this is the Taylor series expansion of e^x,

$$Z(z,V,T) = e^{z\frac{V}{\lambda_T^3}}$$

As a reminder, the analogous derivations of Z_N for B.E. and F.D. distributions are not as tractable. *However*, computing Z (the grand canonical partition function) is tractable! In general,

$$Z(z,V,T) = \sum_{N=0}^{\infty} z^N Z_N(N,V,T) = \sum_{N=0}^{\infty} {\sum_{\{n_\varepsilon\}}}' \left[\prod_\varepsilon \left(ze^{-\beta\varepsilon}\right)^{n_\varepsilon} \right]$$

Looks messy! However, now that we have the sum of all N, we can *remove* the constraint of the prime on the summation. This is extremely important. The grand canonical formalism doesn't care about the number of particles being fixed. This gives,

$$Z(z,V,T) = \sum_{n_0}\sum_{n_1}\cdots\sum_{n_\varepsilon} \left(ze^{-\beta\varepsilon}\right)^{n_0}\left(ze^{-\beta\varepsilon}\right)^{n_1}\cdots\left(ze^{-\beta\varepsilon}\right)^{n_\varepsilon}$$

The nested summations are separable,

$$Z(z,V,T) = \sum_{n_0}\left(ze^{-\beta\varepsilon}\right)^{n_0}\cdot\sum_{n_1}\left(ze^{-\beta\varepsilon}\right)^{n_1}\cdots\sum_{n_\varepsilon}\left(ze^{-\beta\varepsilon}\right)^{n_\varepsilon}$$

Here we have just a product of sums with similar structure,

$$Z(z,V,T) = \prod_\varepsilon \left(\sum_{n_\varepsilon}\left(ze^{-\beta\varepsilon}\right)^{n_\varepsilon}\right) \tag{8.2}$$

Recall that for F.D. $n_\varepsilon = 0, 1$ since at most one particle can be in any given state with energy ε which gives,

$$Z_{\text{FD}}(z,V,T) = \prod_\varepsilon \left(1 + ze^{-\beta z}\right)$$

Whereas for B.E., n_ε can take on any value (any number of bosons can share the same state). Then we can utilize the geometric series,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{For } |x| < 1$$

Letting $x = ze^{-\beta\varepsilon} = e^{+\beta\mu}e^{-\beta\varepsilon}$,

$$Z_{\text{BE}}(z,V,T) = \prod_\varepsilon \left(\frac{1}{1-ze^{-\beta\varepsilon}}\right)$$

Suggestively, these two distributions can be written together,

$$Z_{\text{FD;BE}}(z,V,T) = \prod_\varepsilon \left(1 \pm ze^{-\beta\varepsilon}\right)^{\pm 1}$$

Recalling a result we obtained earlier when looking at η in (7.8),

$$PV = k_B T \ln\left(Z(z,V,T)\right) = \pm k_B T \sum_\varepsilon \ln\left(1 \pm ze^{-\beta\varepsilon}\right)$$

Or using $a = \{0, -1, +1\}$ as in section 8.3,

$$PV = \frac{k_B T}{a}\sum_\varepsilon \ln\left(1 + zae^{-\beta\varepsilon}\right)$$

Other results that follow from (8.2) include,

$$\bar{N} = z\left(\frac{\partial \ln Z}{\partial z}\right)_{V,T} = \sum_\varepsilon \frac{1}{z^{-1}e^{\beta\varepsilon} + a}$$

And,

$$\bar{E} = -\left(\frac{\partial \ln Z}{\partial \beta}\right)_{V,T} = \sum_{\varepsilon} \frac{\varepsilon}{z^{-1}e^{\beta\varepsilon} + a} = \sum_{\varepsilon} \varepsilon e^{-\beta(\varepsilon-\mu)}$$

Perhaps most interestingly for a particular ε' (holding all other ε constant),

$$\langle n_{\varepsilon'} \rangle = \frac{1}{Z}\left\{-\frac{1}{\beta}\left(\frac{\partial Z}{\partial \varepsilon'}\right)_{z,T,\ldots\varepsilon}\right\} = \frac{1}{e^{\beta(\varepsilon-\mu)} + a}$$

Which yields,

$$\langle n_{\varepsilon'} \rangle_{\mathrm{BE}} = \frac{1}{e^{\beta(\varepsilon-\mu)} - 1}$$
$$\langle n_{\varepsilon'} \rangle_{\mathrm{FD}} = \frac{1}{e^{\beta(\varepsilon-\mu)} + 1}$$
$$\langle n_{\varepsilon'} \rangle_{\mathrm{MB}} = e^{-\beta(\varepsilon-\mu)}$$

Example: For spin systems, there might be adjusts to the formulas above. Generally, there is a "internal structure"/ multiplicative factor.

$$\bar{E} = \sum_{\varepsilon} \frac{g_s \varepsilon}{e^{\beta(\varepsilon-\mu)} + a} = g_s \left(\frac{L}{2\pi}\right)^3 4\pi \int_0^\infty \frac{\left(\frac{\hbar^2 k^2}{2m}\right) k^2 \mathrm{d}k}{e^{\beta\left(\frac{\hbar^2 k^2}{2m} - \mu\right)} + a}$$

$$\bar{N} = \sum_{\varepsilon} \frac{g_s}{e^{\beta(\varepsilon-\mu)} + a} = g_s \left(\frac{L}{2\pi}\right)^3 4\pi \int_0^\infty \frac{k^2 \mathrm{d}k}{e^{\beta\left(\frac{\hbar^2 k^2}{2m} - \mu\right)} + a}$$

In practice, one needs to invert these integrals to obtain μ for any given system.

8.4.1 Implications of $\langle n_\varepsilon \rangle$

Fermions:
Because the only electrons/fermions capable of being reshuffled by thermal fluctuations are closed to the Fermi energy, one has that all the properties of the metal are controlled by the that *sliver* of electrons for $\varepsilon \approx \varepsilon_F \pm \Delta$ where $\Delta = k_B T$. For real/ordinary metals $T_F \sim 10^4 - 10^5$K. As such $T/T_F \ll 1$. Electrons in metals are essentially/almost at $T = 0^+$.

Because fermions obey the exclusion principle, they exert pressure (called degeneracy pressure). In a white dwarf, it is the Fermi pressure that holds the gravity of the star. There is a limit (the Chandrasekhar limit M_{CH}) when the mass of the white dwarf becomes more massive that M_{CH} where white dwarfs collapse into neutron stars and then beyond that, black holes (quark stars, superconducting cores/color superconductivity).

Bosons:
There is no restriction for Bosons to occupy all the same $\vec{p} \approx 0$ momentum state. What is remarkable is the evolution from "normal gas" to one with $n\delta(\vec{p}) = n(T)\delta(\vec{p})$ appears/proceeds through a *phase transition* called **Bose-Einstein condensate (BEC)**. This was predicted in 1924-1925 by Satyendra, Bose and Einstein and was not observed until June 5th, 1995 where is was found that,

$$n\lambda_T^3 \sim 1 \quad \lambda_T = \frac{h}{(2\pi m k_B T)^{1/2}} \quad T_c \sim 3.3125 \frac{\hbar^2 n^{2/3}}{m k_B}$$

Below this critical temperature, all of the particles merge into a single wave function. Superfluid ^{4}He has $T_c \sim 2.17\,\mathrm{K}$ is a complex form of a BEC. Also ^{3}He (a fermion) is also a superfluid with $T \sim 1\,\mu\mathrm{K}$. This is achieved when two fermions form a **Cooper pair** to form a composite boson that can then be thought of as a BEC. Superconductivity is a form of a BEC.