

ENGSCI YEAR 2 WINTER 2022 NOTES

SELECTED CLASSES AFTER READING WEEK

JAN - APR 2022

BRIAN CHEN

Division of Engineering Science

University of Toronto

<https://chenbrian.ca>

brianchen.chen@mail.utoronto.ca

Contents

I	PHY294	
	Quantum & Thermal Physics	1
1	Quantum Physics	1
1.1	Everything past the first midterm to end of Quantum portion	1
1.1.1	Chapter 3	1
1.1.2	Chapter 4	2
1.1.3	Chapter 5: Spectroscopy	4
2	Statistical Mechanics	5
2.1	Lecture 1: Ideal Gas Law	6
2.2	Lecture 2: Temperature & Ideal Gases & Solids	6
2.2.1	Idealized Solids	8
2.3	Lecture 3: Two State Systems	9
2.4	Lecture 4: Einstein Solids	11
2.5	Lecture 5: Large Numbers	13
2.5.1	Entropy	14
2.6	Lecture 6: Multiplicity of an Ideal Gas	15
2.6.1	Ideal Gas	15
2.7	Lecture 7: What is temperature? (Revisited)	17
2.8	Lecture 8: Thermodynamic Identity	18
2.8.1	Chemical Potential	19
2.9	Lecture 9: Boltzmann Factors	20
2.10	Lecture 10: Average Values	23
2.11	Lecture 11: Maxwell Speed Distribution	26
2.12	Lecture 12: Free Energy	30
2.13	Lecture 13: Gibbs Factor	32
2.14	Lecture 14: Degenerate Fermi Gas	34
2.15	Lecture 15: Black Body Radiation	38
2.15.1	Bosons	38
2.15.2	Black Body Radiation	38
2.15.3	Making sense of it	39
2.15.4	Example: Cosmic background radiation	41
2.15.5	Example: Star formation	43
2.16	Lecture 16: Debye Theory	44
2.16.1	Fermi gas heat capacity	45
2.16.2	Vibrational mode heat capacity	45
2.17	Lecture 17: Bose-Einstein Condensates	48

II ECE259		
Electromagnetism		50
3 Dielectrics		50
3.1 Lecture 15: Boundary Conditions for Dielectrics		51
3.2 Lecture 16: Capacitors		52
3.3 Lecture 17: Electrostatic Potential Energy		53
3.4 Lecture 18: More on Energy and Capacitors		54
3.5 Lecture 19: Poisson's Equation & Uniqueness Theorem		56
3.6 Lecture 20: Method of Images		56
3.6.1 Currents		57
3.7 Lecture 21: Joule's Law and Resistance		59
3.8 Lecture 22: Continuity Equation & BC for current density		60
4 Electromagnetism		61
4.1 Lecture 23: Static Magnetic Fields		61
4.1.1 Magnetic Vector Potential		63
4.2 Lecture 24: Biot-Savart Law		63
4.3 Lecture 25: Biot-Savart Examples		64
4.4 Lecture 26: Ampere's Law		64
4.5 Lecture 27: Ampere's Force Law, Magnetic Torque		65
4.5.1 Magnetic Dipole		65
4.6 Lecture 28: Magnetization		66
4.7 Lecture 29: \vec{H} field and relative permeability		68
4.7.1 Magnetic Materials		69
4.8 Lecture 30: Boundary conditions for magnetic fields		70
4.8.1 Magnetic Circuits		72
4.9 Lecture 31: Inductance		73
4.9.1 Inductance		73
4.10 Lecture 32: Magnetic Energy		74
4.11 Lecture 33: Electromagnetic Induction		75
4.11.1 Chemical Batteries		75
4.11.2 Faraday's Law		77
4.12 Lecture 33: Induction & Faraday's Law Examples		78
4.13 Lecture 34: Transformers & Lenz's Law		78
4.14 Lecture 35: Motional EMF		80
4.15 Lecture 36: Generalized Faraday's Law		81
4.16 Lecture 37: Maxwell's Equations		82
4.16.1 Ampere-Maxwell Law		82

III ECE286		
Probability and Statistics		84

5	Probability Distributions	84
5.1	Lecture 14: Functions of random variables	84
5.2	Lecture 15: More on moment generating functions	84
5.2.1	Linear combinations of random variables	85
6	Sampling	86
6.1	Lecture 16: Sampling	86
6.2	Lecture 17: Random Sampling	88
6.3	Lecture 18: More on sampling	88
7	Midterm 2	90
7.1	Lecture 19 - 25: Midterm 2 Cheat-sheet	90
8	Hypothesis Testing	91
8.1	General concepts	91
8.2	One- and Two-Tailed Tests	92
8.3	P-Value	93
8.4	Usage steps	93
8.5	Goodness of Fit	93
9	Statistical Modelling	94
9.1	Simple Linear Regression	94
9.1.1	Using Maximum Likelihood	95
9.1.2	Variance Analysis & True Regression Line	95
9.1.3	Error Estimation	96
9.2	Support Vector Machines	96
9.3	Markov Chains	98
9.4	Lecture 36: Review	99
9.4.1	Bayes Rule	99
9.4.2	Sampling	100
9.5	Misc	101
9.5.1	Monte-Carlo Integration	101

IV TEP327 Engineering and Law 102

10	Property	102
10.1	Intellectual Property	103
11	Contracts	105
11.1	Drafting and Interpretation of Contracts	108
11.2	Force Majeure	109
11.3	Typical contract terms	109
11.4	Drafting a contract	110
11.5	Other	110
11.6	Agency and Authority	111

11.7 Concurrent liability	111
11.8 Employer's responsibilities	112
11.9 Environmental law	112
11.10 Aboriginal law	112

PHY294

Quantum & Thermal Physics

SECTION 1

Quantum Physics

SUBSECTION 1.1

Everything past the first midterm to end of Quantum portion

Comment

Excuse the mess: these notes were taken the day before the exam

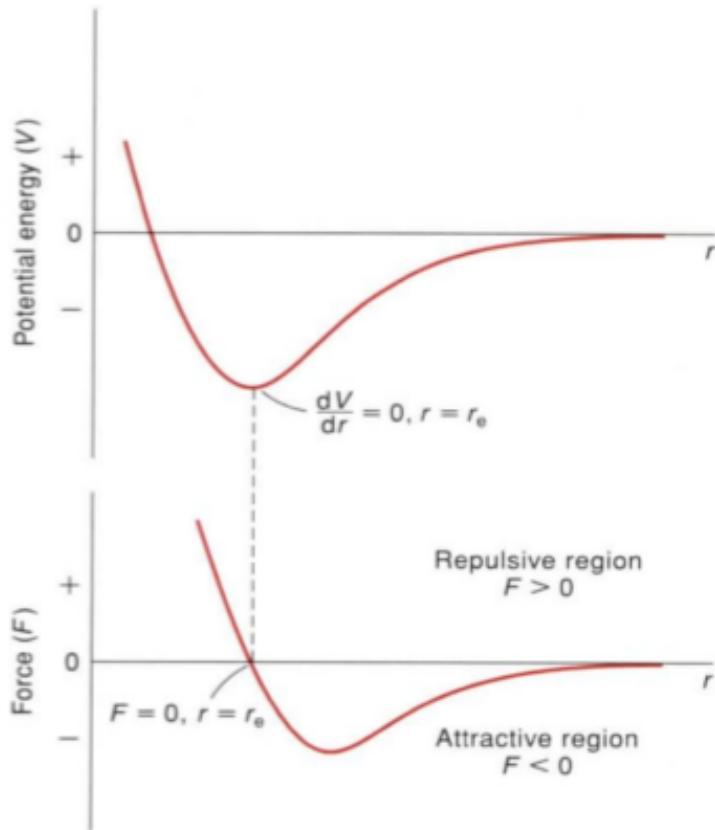
1.1.1 Chapter 3

- electrons when being "shared" become indistinguishable; binding force only depends on time average electron density
- So wavefunctions must satisfy property that electrons are indistinguishable
- Lewis structures/octet rule: $n=2$ shell, lowest energy state.
- Fourier series can be used to approximate wave functions for multielectron atoms [for which the Schrödinger eq cannot be exactly solved]
 - Change in electron density on forming bonds to make new molecules can be written as a superposition of some basis
- Suitable functions must satisfy:
 - Indistinguishable to electron exchange
 - Normalized to 1
 - Must include electron spin
 - Must satisfy Pauli exclusion (antisymmetric to electron exchange)
- Simplest approach: valence bond theory: assume overlap of valence hydrogenic orbitals
- Only really works for s orbitals; predicts incorrect bond angles for p orbitals
- Correction: take linear combinations of valence orbitals (Molecular Orbital Theory): recall: orbital hybridization & bonding/antibonding orbitals, sigma, pi bonding/antibonding pairs.
- Electronegativity: atom's propensity for electrons. Fluorine at 4.0, Cesium at 0.7.
- Hybrid orbital bond wave functions is the linear superposition of valence orbitals

1.1.2 Chapter 4

Scaling atomic interactions to larger systems

- Correspondence principle: there is a scale at which quantum no longer dominates and then we can deal with it classically instead
- Cut off for quantum effects occurs at the rotational motion of molecules.
- Biological molecules are large enough where rotational motions satisfy the classical limit.
- For forces to be bonding, interatomic distances must be smaller than their summative van der Waals radii
- Bond strength: ionic > covalent > metallic. As for intermolecular: ion-dipole > hydrogen bond > dipole-dipole > ion induced dipole > dipole induced dipole > London dispersion
- All potential energy curves have a similar shape like this:

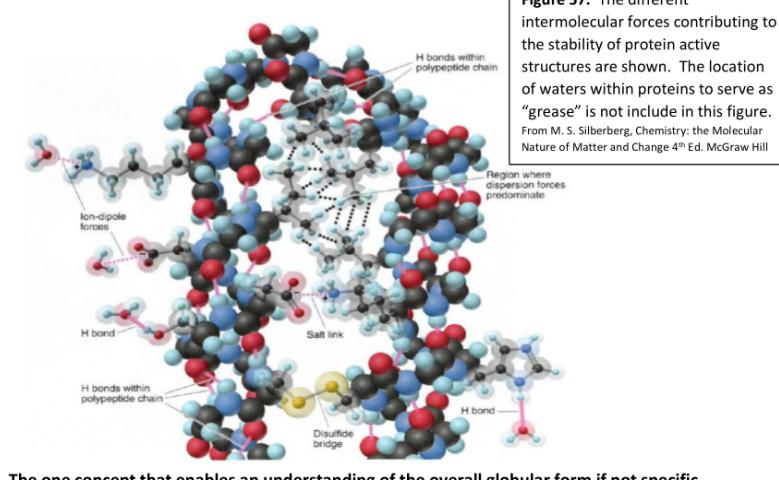


And by inspection we can take the integral/derivative and apply boundary conditions at the 0 point.

- Ionic: attractive term is just Coulomb's law for opp charges
 - Repulsive is harder but has a b/r^6 term between 9 and 12 depending on screening
 - Can find potential and solve for b at the equilibrium bond length via $dU/dr = 0$

- vanderwalls forces can become a signficiant colligative force because it happens with every other molecule and at short range
- H bonds (H-{F, O, N}) allows for imposed directions on molecular interactions.
- H bonds and water; number of h bonds etc are very temperature sensitive
- Water is unique
 - weirdly non viscous dispiste the hbonding
 - Because of H bonding, ice is of lower density than liquid water at low temperature. Water also contracts from 4->0 degrees C.
 - High heat capacity
 - High surface tension (hbond)
 - High idelectric constant
- DNA
 - Specific bonding: AT: 2 Hbond, GC: 3 hbond. G-T only has 1 hbond possible.
 - Just looking at free g/c vs bonded using the energies we find that only about 1/500 GC pairs will have free/broken h bonds. So if we extend to a lot of base paris, it seems to be really selective.
 - Full treatment involves all forces and should not discount the volume of h bonded structure of water
 - TLDR: it is experimentally shown via same shape dna that cant hbond that it isnt just bond energies that cause the really low error rate (it is actually like 1/billion or or something)
- Protiens
 - Really hecking complex (huge molecules from a 20 ammino acid alphabet)

that if we understood all the rules (relationship between the intermolecular forces) we could predict protein structure directly from the primary sequence. The wild card is the effect of water on self assembly.



- A lot of interactions. How can we predict how protiens would actually fold?
- AlphaFold

- Levinthal's paradox

- There are so many degrees of freedoms in proteins. But how do proteins always fold into the desired way? E.g. myoglobin with 100 amino acids, assuming that each one has at least 3 degenerate minima $\Rightarrow 3^{100}$ possible outcomes. That is an absurd number of positions to sample!
- That assumed all DOF are independent. Seems that actually not all of these DOF are independent. Can think of as moving arm; despite all the DOF in arm cells, it is constrained to motion by muscle and joints. Strongly correlated topological units are formed; collective modes.
- Myoglobin is found to have just 4 damped collective helical motions.

1.1.3 Chapter 5: Spectroscopy

Interaction of light with biological molecules, quanta of light

- When light is applied to an electron distribution: force by which light can change distribution is the dipole moment $F = \mu E$. Transfer of energy from light field can only happen if transition between electronic level and associated change in distribution yields an effective dipole moment: selection rule for allowed transitions
- Take Schrödinger for the H atom and then add a time dependent field to it.

$$H = -\frac{\hbar^2}{8\pi^2 m^2} \left(\frac{\partial^2}{\partial x^2} \frac{\partial^2}{\partial y^2} \frac{\partial^2}{\partial z^2} \right) + (U = -\frac{e^2}{4\pi\epsilon} + \mu E(t)) \quad (1.1)$$

- The first term in U is the Coulombic attraction, the 2nd term is the new term which is the action of the time dependent light field.
- Can be treated as a perturbation of hydrogenic solutions
- Just know applied fields lead to change in electron dist that gives a dipole moment.
- Another selection rule: spin: spin cannot change when the electronic level changes.

Vibrational spectroscopy

- Recall the force graph for bonds – there is an equilibrium point and a restoring force: this looks like a spring system!
- Harmonic oscillator is given by Hooke's law and solutions to motion with initial displacement are well known i.e. $x(t) = A \sin 2\pi vt$
- Can treat QM relations as an oscillator with $U = \frac{1}{2}kx^2$ and the mass being the reduced mass $\mu = \frac{m_a m_b}{m_a + m_b}$
- Not easy to solve but just need to remember functional forms of the solutions. Only one 1dof, so only one quantum number associated with solutions: $E = (n + \frac{1}{2})\hbar\omega; n \in \mathbb{Z}$
- This approximation is not perfect; i.e. the actual response is non-linear (recall potential graph) and there are a lot more factors in play. Fairly accurate for $n=0$ to $n=1$, however.
- The IR absorption band for different molecular bonds/atom groups is really nicely spread out and distinct so it can be used to identify substances
- for IR, Most transitions involve $n=0 \rightarrow n=1$. Magnitude of absorption depends on magnitude of dipole moment. In this case (as opposed to electronic), dipole moment is already present in molecule and not induced by E field.

- In order to absorb ir, msut have a dipole moment associated with vib. motions, i.e. IR active mode.

Electronic Spectroscopy

Using higher energy particles to do spectroscopy.

- Absorption of light coupled to electronic dof
- Typically place material in solvent¹ Solvent will cause a red shift in spectrum of any given molecule, called coupling to the bath. Magnitude of effect depends on type of molecule and bath interaction.jj
- In biological systems the bath is the surrounding protien. For example in rhodophin, visual pigment. Active chromophore is retinal. Retinal absorbs in 280-400nm, uv to blue. Retinal in any solvent shifts this to green (530nm).
- Retinal is embedded in opsin protien to geenrate the protien complex rhodopsin which depeates the retinal electron density just the right way to lallow for it to work out properly.
- There are no well defined selection rules for transitions involving larger systems like those in biology. Wave functions do not usually have enough symmetry to give generalized rules. Can do it qualitatively.
 - For light to be absorbed, the transition must involve a change in spatial charge dist. or a dipole moment.
- Selection rule that optical transition cannot involve a change in spin holds true still. The general rule is that the change net electron spin moment $S = \frac{1}{2}n$ must be zero.
- Since things are quantized (incoming and exiting light field will be one quanta of energy less), the transition is effectively instatenous (see: vertical transition line). Therefore nuclei cannot move on the time scale of the transition; so transition probabily must take into account the probability of nuclei being at the same location fro two differenrt states. Connection is the overlap integral of tnuclear wavefn between two levels [these are franck-cordon factors]. Manifests itself as a superposition of vibrational spectra onto electronic surfaces to give rise to vibronic transitions (which are both vibrational and electroinc)

Note: Electronic transitions are vertical or almost vertical lines on such a plot since the electronic transition occurs so rapidly that the internuclear distance can't change much in the process. Vibrational transitions occur between different vibrational levels of the same electronic state. Rotational transitions occur mostly between rotational levels of the same vibrational state, although there are many examples of combination vibration-rotation transitions for light molecules.²

SECTION 2

Statistical Mechanics

SUBSECTION 2.1

Lecture 1: Ideal Gas Law

¹???

²<http://hyperphysics.phy-astr.gsu.edu/hbase/molecule/molec.html>

We begin our discussion on statistical mechanics with the guiding question, "What is temperature". We will first derive the ideal gas law $PV = nRT$, which most should be familiar with.

Imagining a single particle bouncing around elastically in 2D box with speed v_x . If we were to look at the sides of the box we will see intermittent spikes of force experienced [by the box walls],

$$F_x = m \frac{2v_x}{\Delta t_c} \quad (2.1)$$

Over time can get the average force

$$F_x = m \frac{2_x}{\Delta t} \quad (2.2)$$

Which allows us to derive an expression for pressure, which is just force over area

$$P = \frac{F_x}{A} = \frac{mv_x^2}{V} \quad (2.3)$$

For n particles it then becomes

$$P = \sum_{i=1}^N \frac{mv_{x,i}^2}{V} = \frac{m}{V} \sum_{i=1}^N v_{x,i}^2 \quad (2.4)$$

Assuming that that $v_x = v_y = v_z$ ¹, we may extend this to 3 dimensions

¹ This is reasonable, is it not?

$$P = \frac{mv^2}{3} \frac{N}{V} \quad (2.5)$$

Now, compare this expression and that of the ideal gas law to find that temperature is a measure of the average kinetic energy per particle

$$T = \frac{mv^2}{3k} = \frac{2\bar{E}}{3k} \quad (2.6)$$

Note that $\frac{mv^2}{3} = kT \rightarrow \frac{mv^2}{3k} = 2\frac{\bar{E}}{3k}$

SUBSECTION 2.2

Lecture 2: Temperature & Ideal Gases & Solids

Remark Recall, $PV = kT\frac{N}{V}$ where $T = \frac{mv^2}{3k} = \frac{2\bar{E}}{3k}, \bar{E} = \frac{3kT}{2}$ Where T is the experimentally measured temperature

We are still, however, left with a few questions:

1. But how does this apply to things that aren't gasses?
2. And this only describes average speed – what about the distribution?
3. And how can we describe thermal equilibrium?

Definition 1

Heat capacity: the energy required for a unit change in temperature.

Lemma 1

The total energy of an ideal gas at temperature T is given by

$$U = NE = 3 \frac{NkT}{2} \quad (2.7)$$

So, increasing U increases T

Therefore can define heat capacity:

$$C_v \equiv \left. \frac{\partial U}{\partial T} \right|_{V, N, \dots} \quad (2.8)$$

This brings us to the expression for 1 mole of an ideal gas,

$$C_V = \frac{3R}{2} \quad R = 8.314 \text{ J/(K} \cdot \text{mol}) \quad (2.9)$$

Recall $Nk = nR$, because chemistry/physics notation

Theorem 1**Equipartition Theorem**

$$U = N\bar{E} = \frac{3}{2}Nm v^2 = 3N \frac{kT}{2} \quad (2.10)$$

So the total energy is the total quadratic degrees of freedom D multiplied by $\frac{kT}{2}$. For a gas, $D = dN$, where d is the dof ² per particle

Therefore,

$$U = d \frac{nRT}{2} \quad (2.11)$$

Given that

- $\bar{E} = \frac{3}{2}kT$
- There are three degrees of freedom, each of which having kinetic energy $\frac{1}{2}mv_a^2$, where $a \in \{x, y, z\}$

Which gives us

$$C_V = \frac{dR}{2} \quad (2.12)$$

²Degrees of freedom

For an ideal gas

Now, armed with $C_V = \frac{dR}{2}$ we notice that for Helium, Argon, and Neon this works well (and they have dof ≈ 3). Also notice that the ones for which it breaks on are primarily diatomic or massive molecules, e.g. O_2

But there are actually 3 translational degrees of freedom and 3 rotational degrees of freedom. So why don't all gases have $d = 6$? But our prior analysis indicates $d = 3$ for monoatomic and $d = 5$ for diatomic

From quantum mechanics the rotational energy levels for a rigid object are

$$E_J = \frac{\hbar}{2I} J(J+1) \quad (2.13)$$

Where

This is exactly what we said earlier, however, the expression changes depending on the number of degrees of freedom the particle has

- \hbar is the reduced Planck's constant
- J is the quantum number
- $I = \frac{2}{5}mr^2$ is the moment of inertia, in this case we assume a solid sphere

So if we take the Equipartition theorem (Eq. 2.10) to heart, we see that the energy in the rotational modes is entirely irrelevant; it requires temperature at the order of a billion Kelvin to reach the first excited state. At room temperature the modes are "frozen out"

Applying Eq. 2.13 to a diatomic molecule e.g. N_2 we note that the nuclei are separated by a *much* greater distance. If we were to imagine the diatomic molecule it becomes apparent that I is much larger along two axes than one, \Rightarrow only 2 of these modes matter at room temperature and therefore diatomic molecules have 3 + 2 degrees of freedom [at room temperature]. This also has the implication that C_V is temperature-dependent

2.2.1 Idealized Solids

The previous derivation for gases *cannot* be strictly applied to solids.

Assuming a solid is modelled as a bunch of atoms attached with springs, we arrive at $E = \frac{1}{2}mv^2$. But when we look at solids (Fig. 1) we note that they have this plateau effect.

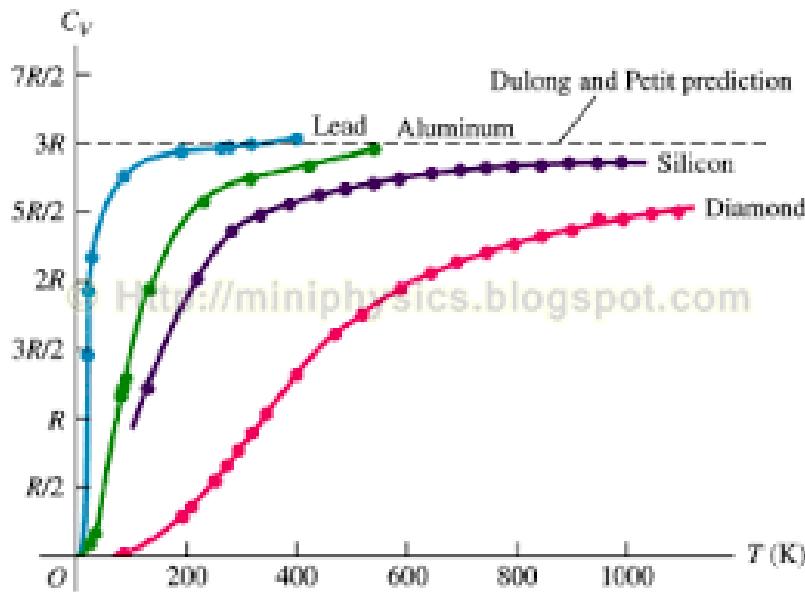


Figure 1. Heat capacity of selected solids

This is explained by there being 3 potential degrees of freedom in the spring³.

But basically what the Equipartition theorem tells us is the energy stored in each quadratic degree of freedom. But we don't always know what the degrees of freedom are. For example how do we know what should we count or not?

Remark Heat capacity at a constant volume is really easy to calculate, but is difficult to measure since solids tend to thermally expand. Heat capacity at a constant press is really easy to measure, but difficult to calculate.

So, to recap:

- The Equipartition theorem seems pretty powerful; it can predict energy per mode as a function of temperature and calculate a bunch of things
- BUT:

³ ...

- We didn't really prove it
- Doesn't really work when quantum levels matter
- 'Degrees of freedom' is not well-defined
- Works only in average energy
- Doesn't explain equilibrium
- Still don't know what entropy is

Let's try to address all of these shortcomings by applying quantum statistical mechanics.

SUBSECTION 2.3

Lecture 3: Two State Systems

Begin with coins; heads/tails with equal likelihood, independent flips.

Definition 2

To describe this better we will define a few terms:

Microstate: a specific configuration; all of which are equally likely, i.e. individual head/tails outcomes

Macrostate: Defined by some combined quantity, possibly containing multiple microstates, i.e. a state w/ 2/3 heads

Multiplicity (Ω): The number of microstates in a macrostate

For example, if we were to flip 100 coins.

- Multiplicity of 0 heads: $\Omega(0) = 1$
- Multiplicity of 1 heads: $\Omega(1) = 100$
- Multiplicity of 2 heads: $\Omega(2) = \frac{100 \cdot 99}{2}$
- Multiplicity of n heads: $\Sigma(N, n) = \binom{N}{n} = \frac{N!}{n! \cdot (N-n)!}$
 - The probability of n heads after N flips: $\frac{\Sigma(N, n)}{\Sigma_{\text{tot}}} = \frac{\Sigma(N, n)}{2^N}$

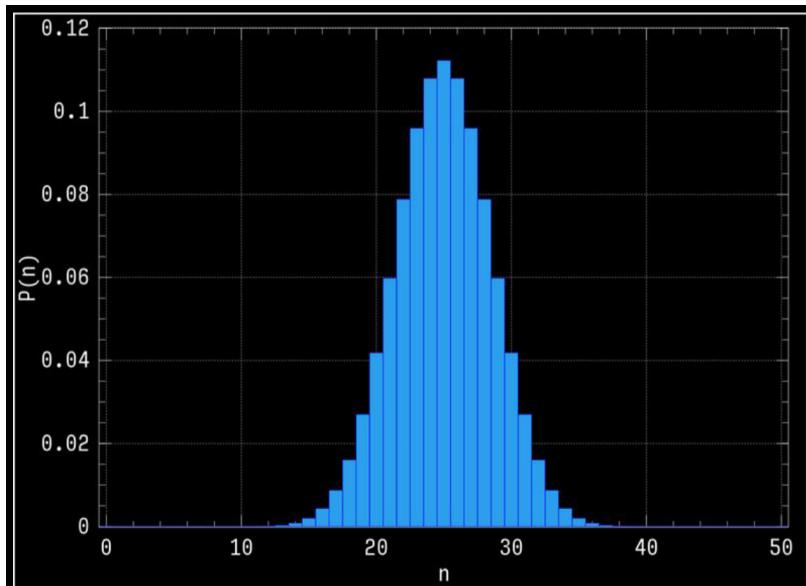


Figure 2. Probability distribution for 50 coin flips

Now, imagine a system of little independent magnetic dipoles that can point up or down in an external magnetic field (Fig. 3) The multiplicity is given by the expression we found before:

$$\Sigma(N, n) = \binom{N}{n} = \frac{N!}{n! \cdot (N - n)!} \quad (2.14)$$

As there are more flips the distribution will become *relatively* more narrow

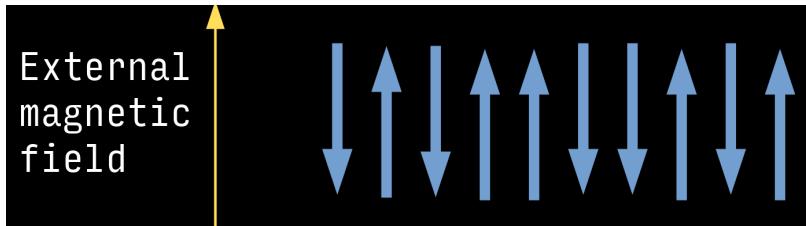
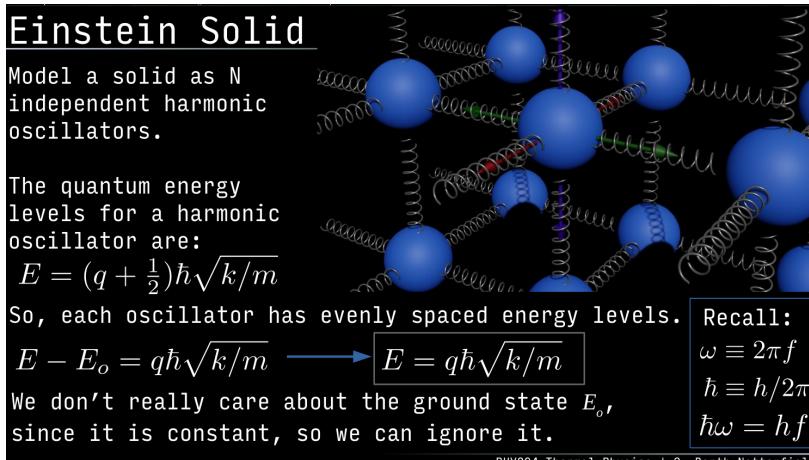


Figure 3. Independent dipoles in field

Equal numbers of dipoles pointed up and down has a far higher multiplicity and therefore is the most probable state although it is not the lowest energy state. On the other hand, the lowest energy state would be all of the dipoles being aligned with the magnetic field, and highest being all anti-aligned. How can resolve this problem?

SUBSECTION 2.4

Lecture 4: Einstein Solids

**Figure 4.** Einstein solid

Moving away from the contrived dipole example we move to the Einstein solid which is basically a bunch of harmonic oscillators as a means to finding out what temperature is.

Comment

Recall: the quantum energy levels of a harmonic oscillator is

$$E = (q + \frac{1}{2})\hbar\sqrt{k/m} \quad (2.15)$$

Where

- $w = 2\pi f$
- $\hbar = \frac{h}{2\pi}$
- $\hbar w = hf$

By inspection we see that each oscillator has evenly spaced energy levels.

$$E - E_o = q\hbar\sqrt{k/m} \longrightarrow q\hbar\sqrt{k/m} \quad (2.16)$$

Each identical independent oscillator has an integer quanta of energy. This means that for a 3-oscillator system there are 6 ways 2 quanta can be distributed is $(2, 0, 0), (0, 2, 0), (0, 0, 2), (1, 1, 0), (0, 1, 1), (1, 0, 1)$ and so on⁴. The **macrostate** is the total energy in the system, $q_{tot} = 2$. The **microstate** is the configuration, i.e. $(1, 1, 0)$. So, the **multiplicity** of the macrostate $q_{tot} = 2$ is $\Omega(2) = 6$. Going onwards $\Omega(3) = 10$ and so forth.

Ground state is a constant so we ignore it

⁴The nice thing about statistical mechanics is that each step is fairly trivial

Theorem 2

The multiplicity for q quanta of energy distributed between N oscillators is

$$\Omega(N, q) = \binom{q + N - 1}{q} = \frac{(q + N - 1)!}{q!(N - 1)!} \quad (2.17)$$

PROOF

Imagine $N = 4$ oscillators and $q = 8$ quanta. Consider the configuration $(1, 3, 0, 4)$. We can draw it like this:



where $*$ is a quanta of energy and $|$ separates the oscillators. Any combination of 8 of $*$ and 3 of $|$ is a valid microstate.

So: There are $q + N - 1$ symbols, q of which are a $*$

TLDR: cute proof! \square

By inspection we find that the multiplicity grows *very* rapidly with increasing N or q .

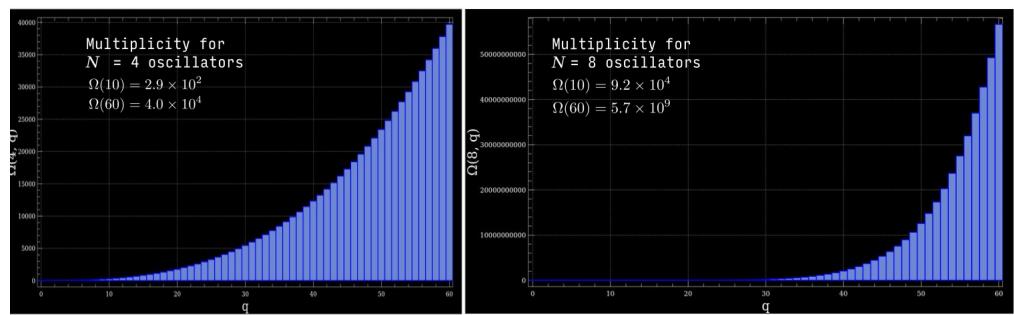


Figure 5. Multiplicity Ω of Einstein solid given N, q

Now, can we understand temperature by making Einstein solids interact with each other? Imagine two Einstein solids with $N = 4$ oscillators each and $q_{tot} = q_1 + q_2$ quanta of energy between them. How do the macrostates and multiplicity look?

Note: if solid 1 has $q_1 = 10$, then solid 2 has $q_2 = 60 - 10 = 50$

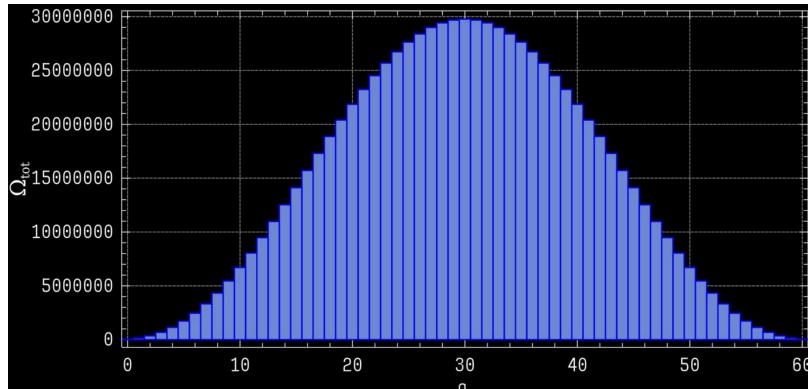
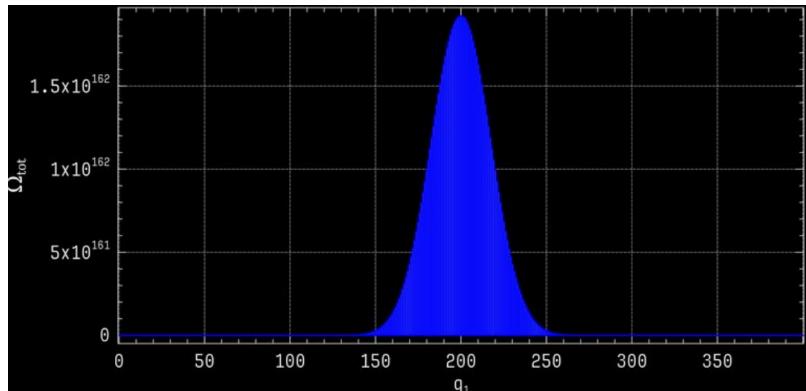


Figure 6. $N_1 = N_2 = 4, q_1 + q_2 = 60, \Omega_{tot} = \Omega(q_1) \times \Omega(60 - q_1)$

Since they share energy, it follows that the total multiplicity of the system is $\Omega_{tot} = \Omega_1 \times \Omega_2$. Plotting this we find that the multiplicity will tend to follow something that looks a bit like a normal distribution, the spread of which depends on the N and q . As this gets larger and larger we gain more and more certainty as to the system's energy – that the system with the greatest multiplicity (most likely) has an energy level which is in-between the energies of the interacting bodies.



Comment This begins to look like a plausible explanation for thermal equilibrium!

SUBSECTION 2.5

Lecture 5: Large Numbers

Let's consider a mole of Einstein solid 1.8×10^{24} at normal temperatures $q = 10^{26}$. If one were to try to calculate what we did in the previous section it would just not work, especially computationally. So let's apply a few methods to model this

Types of numbers

1. Small numbers: e.g. 42 which we can use arithmetic as normal
2. Large numbers: can ignore adding a small number $10^{24} \approx 10^{24} + 24$
3. Very large numbers: can ignore multiplying a very large number by a large number, i.e. $10^{10^{23}} \times 10^{23} \approx 10^{10^{23}}$

Theorem 3

Stirling's approximation

$$N! \approx \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \quad N \gg 1 \quad (2.18)$$

Which also gives us the result (for large N)

$$\ln(N!) = N \ln(N) - N \quad (2.19)$$

Can use this to find $\Omega(N, q)$

$$\begin{aligned} \ln(\Omega) &\approx N \ln\left(\frac{q}{N}\right) + N \quad q \gg N, N \gg 1 \\ \Omega &\approx \left(\frac{eq}{N}\right)^N \end{aligned} \quad (2.20)$$

More formally,

$$\begin{aligned}
\Omega(N, q) &= \binom{q+N-1}{q} \approx \frac{(q+N)!}{q!N!} \\
\ln(\Omega) &\approx \ln((q+N)!) - \ln(q!) - \ln(N!) \\
&\approx (q+N) \ln(q+N) - (q+N) - q \ln(q) + q - N \ln(N) + N \\
&= (q+N) \ln(q+N) - q \ln(q) - N \ln(N) \\
&= (q+N)(\ln(q) + \frac{N}{q}) - q \ln(q) - N \ln(N) \\
\Rightarrow \ln(\Omega) &= N \ln\left(\frac{q}{N}\right) + N + \frac{N^2}{q} \\
\Rightarrow \Omega &= \left(\frac{e^q}{N}\right)^N
\end{aligned} \tag{2.21}$$

Use $\ln(N!) \approx N \ln(N) - N$
 $\ln(1+x) \approx x \quad x \ll 1$
assume $q \gg N$

From our intuition we expect that when we combine two multiplicities that the peak will be at $q_1 = q_2 = q_{tot}/2$...

Let $x = q_1 - \frac{q}{2} = \frac{q}{2} - q_2$

$$\Omega_{tot} = \left(\frac{e}{N}\right)^{2N} \left[\left(\frac{q}{2}\right)^2 - x^2\right]^N \tag{2.22}$$

And then we do some math [textbook 2.26, 2.27] and find that it is a Gaussian

Definition 3

Large N multiplicity of Einstein solids

$$\begin{aligned}
\Omega_{tot} &\approx \Omega_{max} e^{-N(2x/q)^2} \\
\Omega_{max} &= \left(\frac{e}{N}\right)^{2N} \left(\frac{q}{2}\right)^{2N}
\end{aligned} \tag{2.23}$$

And with Stirling approximation we arrive at

$$\Omega \approx \left(\frac{eq_1}{N}\right)^N \tag{2.24}$$

The internal entropy for an Einstein solid is given by

$$E = \varepsilon q \tag{2.25}$$

Where $\varepsilon = \hbar w$, w being the frequency of the oscillations.

And the variance can be found for large N , which becomes increasingly thin for large N . if $N = 10^{24}$, $\sigma = \frac{q}{2 \times 10^{12}}$

$$\sigma = \frac{q}{2\sqrt{N}} \tag{2.26}$$

2.5.1 Entropy

Definition 4

$$S \equiv k \ln(\Omega) \tag{2.27}$$

And we find that systems in thermal contact will tend to be found with the highest entropy; *Entropy increases*.

SUBSECTION 2.6

This is the second law of thermodynamics!

Lecture 6: Multiplicity of an Ideal Gas

Recall the expression for large N multiplicity is given by Eq. 2.24 and for a combination of Einstein solids it is simply $\Omega_{tot} = \Omega_1 \times \Omega_2$

Substituting $x = q_1 - \frac{q}{2}$ we can arrive at the following approximation

$$\approx \Omega_{max} \cdot e^{-x^2/a^2} \quad (2.28)$$

And $a = \frac{q}{\sqrt{2\sqrt{N}}}$ describes the width of the peak (which can be thought of as describing the accuracy).

2.6.1 Ideal Gas

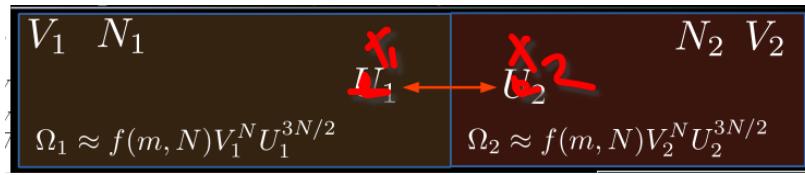
Comment

Recall from the quantum mechanics section of this course that particles in a cubic box have quantized energy

$$E = \frac{h^2}{8m} \left(\frac{n_1^2 + n_2^2 + n_3^2}{V^{2/3}} \right) \quad (2.29)$$

In the textbook it gives us a massive formula for the multiplicity of an ideal gas

$$\Omega_N \approx \frac{1}{N!} \frac{V^{2/3}}{h^{3/N}} \frac{\pi^{3N/2}}{(3N/2)!} (\sqrt{2mU})^3 N = f(m, N) V^N U^{3N/2} \quad (2.30)$$



This is not very satisfying and we will get back to this later. For now let's assume it is correct.

Let's consider a system of two ideal gases in contact. In this case if we allow for energy transfer, i.e. $X_i = U_i \dots$

$$\begin{aligned} \Omega_{tot} &= \Omega_1 \times \Omega_2 \\ \Omega_{tot} &= f()^2 (U_1 U_2)^N (U_1 U_2)^{3N/2} \\ &\approx \Omega_{max} \cdot \exp \left(\frac{-x^2}{a^2} \right) \\ \text{where } x &= U_1 - \frac{U_{tot}}{2} \\ \rightarrow a &= \frac{U}{2\sqrt{3N/2}} \end{aligned} \quad (2.31)$$

And so we get the result that the multiplicity is again a Gaussian with a peak at $x = U_1 - \frac{U}{2}$ and a spread on the order of $\frac{1}{\sqrt{N}}$ as well, which indicates that thermal equilibrium works for ideal gases with energy transfer as well!

Considering a system where we allow for volume transfer, i.e. $X_i = V_i \dots$

$$\begin{aligned}
 \Omega_{tot} &= \Omega_1 \times \Omega_2 \\
 \Omega_{tot} &= f()^2 (V_1 V_2)^N (U_1 U_2)^{3N/2} \\
 &\approx \Omega_{max} \cdot \exp\left(\frac{-x^2}{a^2}\right) \\
 \text{where } x &= V_1 - \frac{V_{tot}}{2} \\
 \rightarrow a &= \frac{V}{2\sqrt{N}}
 \end{aligned} \tag{2.32}$$

We arrive at much the same result as before with the Gaussian and spread.

Considering a system where we allow for particle transfer, i.e. $X_i = N_i$, we get a really ugly expression. I do not want to write this out so here is the screen-shotted lecture slide:

Large N Multiplicity of Ideal Gasses

$U_1 \quad V_1$	N_1	$V_2 \quad U_2$	Poke a hole and let molecules flow...
$\Omega_1 \approx \frac{1}{N_1!} \frac{V_1^{N_1}}{h^{3N_1}} \frac{\pi^{3N_1/2}}{(3N_1/2)!} (\sqrt{2mU_1})^{3N_1}$	N_2	$\Omega_2 \approx \frac{1}{N_2!} \frac{V_2^{N_2}}{h^{3N_2}} \frac{\pi^{3N_2/2}}{(3N_2/2)!} (\sqrt{2mU_2})^{3N_2}$	
$\Omega_{tot} = \Omega_1 \times \Omega_2$	Note that $U_1 = \frac{3}{2}N_1 k T_1$ And use Stirling's approximation		
$\ln(\Omega_{tot}) \approx f(N, m, T, V) - N_2 \ln(N_2) - N_1 \ln(N_1)$	More flailing around...		
$\Omega_{tot} = e^{f(N, T, V, m)} e^{-x^2/c(N)}$	Define $x \equiv N_1 - N/2 = N/2 - N_2$ And use $\ln(1 + x) \approx x$		
	Strongly peaked at $N_1 = N_2$		

The point of all this is that this brings us to the second law of thermodynamics and the concept of entropy which enables us to work with these really big numbers in a much easier way.

Before we do that let's convert our multiplicity expressions to entropy

$$\Omega \approx \left(\frac{eq^N}{N}\right) \quad N = 10^2 2 \Leftrightarrow S = Nk[\ln(\frac{q}{N}) + 1] \tag{2.33}$$

And derive the following property from just log laws

$$\begin{aligned}
 \Omega_{tot} &= \Omega_1 \times \Omega_2 \\
 S_{tot} &= S_1 + S_2
 \end{aligned} \tag{2.34}$$

Which, coupled by the 2nd law of thermodynamics (entropy increases and system will be most stable maximum entropy), we now have a really good way of describing temperature and thermodynamic equilibrium.

Applying this to an ideal gas we can do some ugly math from Eq. 2.30 and Eq. 2.27, use Stirling's approximation and drop some large factors to arrive at the following expression for entropy of an ideal gas

Theorem 4

Entropy of an ideal gas

$$S = Nk \left[\ln \left(\frac{V}{N} \left(\frac{4\pi m U}{3N h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right] \tag{2.35}$$

This is also known as the Sackur-Tetrode equation

Inspecting the equation entropy increases with N, V, m, U , which makes intuitive sense. Note that there is a subtlety in using entropy. For example mixing two different gases; i.e. a gases with different masses, or gases with a different energy; this would cause a positive ΔS . However mixing two *identical* gases would cause no entropy change; $\Delta S = 0$. What to note is that thermal equilibrium is the tendency of a system to enter a more likely state. If they are in the same state; mixing them will not cause any further change in state and therefore positive ΔS .

Entropy is not anything mysterious anymore!

SUBSECTION 2.7

Lecture 7: What is temperature? (Revisited)

This was a short lecture with lots of review. The key takeaway is that thermal equilibrium is found when

$$\frac{\partial S_1}{\partial q_1} = \frac{\partial S_2}{\partial q_2} \quad (2.36)$$

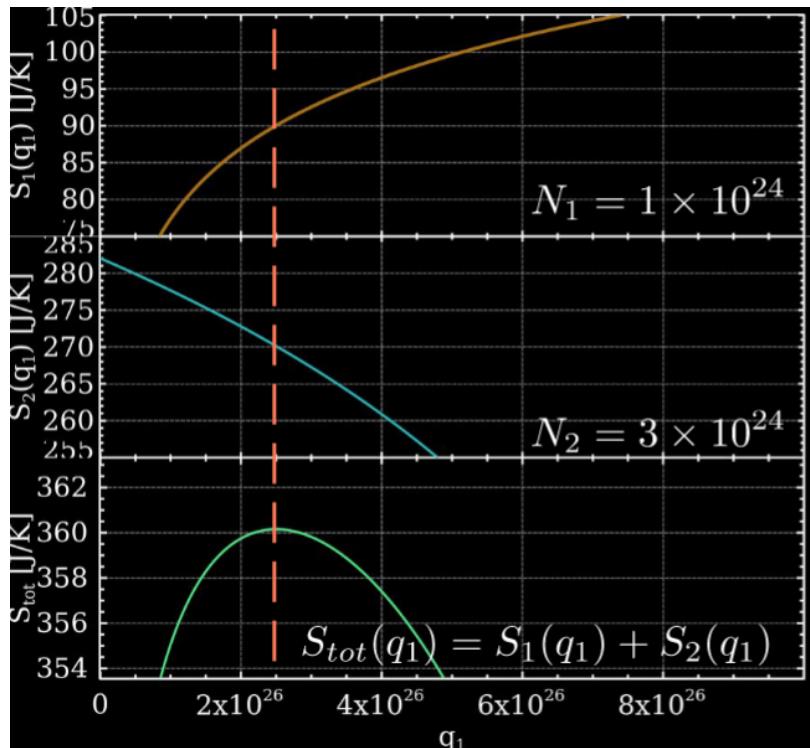


Figure 7. Toy ideal gas system

But $U = qE_o$, so

$$\frac{\partial S_1}{\partial U_1} = \frac{\partial S_2}{\partial U_2} \quad (2.37)$$

The peak is when the derivative of the entropy with respect to energy is the same in both systems.

And so we arrive at the conditions for thermal equilibrium:

1. Temperatures are equal
2. Entropy is maximized (S_{max} is at $\frac{\partial S_1}{\partial U_1} = \frac{\partial S_2}{\partial U_2}$)

Intuitively the peak is where taking energy from one system and putting it in the other does not cause a change in total entropy

But! $\frac{\partial S}{\partial U}$ decreases with increasing U but temperature increases with increasing U . So let's test

$$T \equiv \left(\frac{\partial U}{\partial S} \right)_{N,V} \quad (2.38)$$

We can test it for an ideal gas

$$\begin{aligned} S &= Nk \left[\ln \left(\frac{V}{N} \left(\frac{4\pi m U}{3N\hbar^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right] = Nk \ln \left(U^{\frac{3}{2}} \right) + f(N, V) \\ T &= \left(\frac{\partial S}{\partial U} \right)^{-1} = \left(\frac{2Nk}{2U} \right)^{-1} = \frac{2U}{3Nk} \end{aligned} \quad (2.39)$$

Perhaps this could be temperature? Since they are equal at thermal equilibrium...

Keeping N, V constant

Which brings us to the *very* satisfying result that

$$\Rightarrow U = 3N \frac{kT}{2} \quad (2.40)$$

, which is *exactly* what the equipartition theorem says for an ideal gas! This T is what an ideal gas law thermometer measures – but is applicable to *any* system.

Definition 5

We can therefore define temperature as the partial derivative of energy with respect to entropy

$$T \equiv \frac{\partial U}{\partial S} \text{ or } \frac{1}{T} = \frac{\partial S}{\partial U} \quad (2.41)$$

T^{-1} is how much $S = \ln(\Omega)$ increases when energy is added to a system. This means that at a low temperature putting in a little energy will increase S a lot, and vice-versa for high temperatures.

SUBSECTION 2.8

Lecture 8: Thermodynamic Identity

Comment

Recall: $(T)^{-1}$ is how much $S = k \ln \Omega$ with "everything else" held constant. This means at a low temperature adding energy increases S a lot, and vice versa. This means that when putting a low and high temperature object besides each other, the higher temperature object will more likely give up temperature to the lower temperature one – which is what we expect.

We have a procedure for finding heat capacity

1. Find the multiplicity $\Omega(U, N, V, \dots)$
2. Find the entropy $S = k \ln \Omega$
3. Find $T(U, N, V, \dots) = \left(\frac{\partial S}{\partial U} \right)^{-1}$
4. Solve for $U(T, N, V, \dots)$
5. Find $C_V = \frac{\partial U}{\partial T}$

And we find that this works well for Einstein solids and for real solids e.g. Lead as well. It also works for ideal gases and para-magnets – but other than for these we are kind of stuck;

$$\begin{aligned} U &= \varepsilon q, \Omega \approx \left(\frac{eq}{N} \right)^N \\ &= \left(\frac{eU}{N\varepsilon} \right)^N \\ S &= Nk \left[\ln \frac{U}{N\varepsilon} + 1 \right] \\ T &= \frac{U}{Nk} \\ U &= NkT \\ C_V &= Nk \end{aligned} \quad (2.42)$$

it is *really* difficult to use this tool to find heat capacity for more complex substances. This is where classical thermodynamics comes in

In order to complete the picture we now want to be able to build a better understanding of pressure. Going back to the example in Fig. 7, let's consider allowing volume and energy to change.

$$\text{At } S_{\max} \frac{\partial S_1}{\partial U_1} \frac{\partial S_2}{\partial U_2} \text{ and } \frac{\partial S_1}{\partial V_1} \frac{\partial S_2}{\partial V_2} \quad (2.43)$$

Also, $T_1 = T_2, P_1 = P_2$, where $T \equiv (\frac{\partial S}{\partial U})^{-1}$. We've played this 'game' before with temperature to get the above result – now let's do it with pressure. We note that $\frac{\partial S_1}{\partial V_1} \frac{\partial S_2}{\partial V_2}$ has units of $\frac{N}{m^2 \cdot K}$, and $P_1 = P_2$ has units $\frac{N}{m^2}$. So let's propose

$$P \equiv T \left(\frac{\partial S}{\partial V} \right)_{U,N} \quad (2.44)$$

Plug this into Eq. 2.35 and see if it reproduces the ideal gas law. Which gives us

Definition 6

Pressure

$$P = \frac{TNk}{V} \quad (2.45)$$

And that's the ideal gas law! So Eq. 2.44 is a valid expression for pressure.

2.8.1 Chemical Potential

Let's go back to the example in Fig. 7 yet again but allow for changing volume, energy, and number of particles.

Building upon what we have derived prior for volume and temperature we note that

$$\frac{\partial S_1}{\partial N_1} \frac{\partial S_2}{\partial N_2} \quad (2.46)$$

Definition 7

Chemical Potential is therefore defined as:

$$\mu \equiv T \left(\frac{\partial S}{\partial N} \right)_{U,V} \quad (2.47)$$

At equilibrium total entropy doesn't change; i.e. moving a particle from one side to the other doesn't make an impact

Barth Netterfield: The negative is there for tradition?

Theorem 5

Therefore we may define the **Thermodynamic Identity** as

$$dS = \left(\frac{\partial S}{\partial U} \right)_{N,V} \left(\frac{\partial S}{\partial V} \right)_{N,U} \left(\frac{\partial S}{\partial N} \right)_{U,V} \quad (2.48)$$

And plugging in our definitions for T, P , and μ ,

$$dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN \quad (2.49)$$

With Eq. 2.49 we may integrate over all energies, volumes, number of particles, etc. to get a measure of the entropy and therefore multiplicity in a system! It is still in not a great form to work with so we solve for U to get the following identity

$$dU = TdS - PdV + \mu dN \quad (2.50)$$

Example Let's play with Eq. 2.50 a little bit and find out surprisingly intuitive it is.

1. Keep entropy and volume to be 0 and constant. $\mu = \frac{\partial U}{\partial N}_{S,V}$. This says that chemical potential is the amount of energy you need to add⁵ a particle in order to keep the entropy constant. This makes sense since S is a strong function of N
2. Keep energy and volume to be 0 and constant. $\frac{\mu}{T} = -\frac{\partial S}{\partial N}_{U,V}$. Chemical potential⁶ is the amount the entropy changes when you add a particle (with energy and volume constant). Note that chemical potential is generally negative, so entropy increases while adding particles.
3. Keep energy and entropy to be 0 and constant. $\frac{\mu}{P} = -\frac{\partial V}{\partial N}_{U,S}$. Chemical potential⁷ is the amount the volume must change when a particle is added in order to keep energy and entropy constant
4. Keep energy and number to be 0 and constant. $\frac{P}{T} = -\frac{\partial S}{\partial V}_{U,N}$. Pressure⁸ is the amount by which entropy changes when you increase the volume.

⁵ Chemical potential is usually negative, so you need to take away energy

⁶ divided by temperature

⁷ divided by pressure

⁸ divided by temperature

To recap: we have now have the thermodynamic identity from which most of classical thermodynamics is derived from; what we're missing now is just various free energies.

SUBSECTION 2.9

Lecture 9: Boltzmann Factors

Comment

A summary of key results derived in previous lectures:

Probability of macrostate s :

$$P(s) = \frac{\Omega(s)}{\Omega_{\text{tot}}} \quad (2.51)$$

Entropy (which is maximized at equilibrium) (Eq. 2.27)

$$S \equiv k \ln(\Omega) \quad (2.52)$$

And the thermodynamic identity (Eq. 2.50)

$$dU = TdS - PdV + \mu dN \quad (2.53)$$

Where

$$T \equiv \left(\frac{\partial U}{\partial S}\right)_{N,V}, \quad P \equiv T\left(\frac{\partial S}{\partial V}\right)_{U,N}, \quad \mu \equiv T\left(\frac{\partial S}{\partial N}\right)_{U,V} \quad (2.54)$$

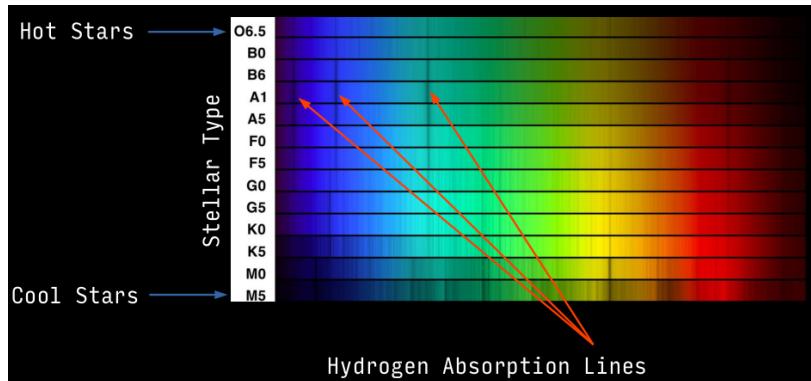


Figure 8. Spectral lines for selected stars. Note absorption lines – how can we use statistical mechanics to explain the behaviour of these hydrogen absorption lines? Why do they appear for hot stars but disappear for cool stars?

Hydrogen absorption lines form the Lyman, Balmer, and Paschen series. The visible ones start at the $n = 2$ shell and so lie in the Balmer series.

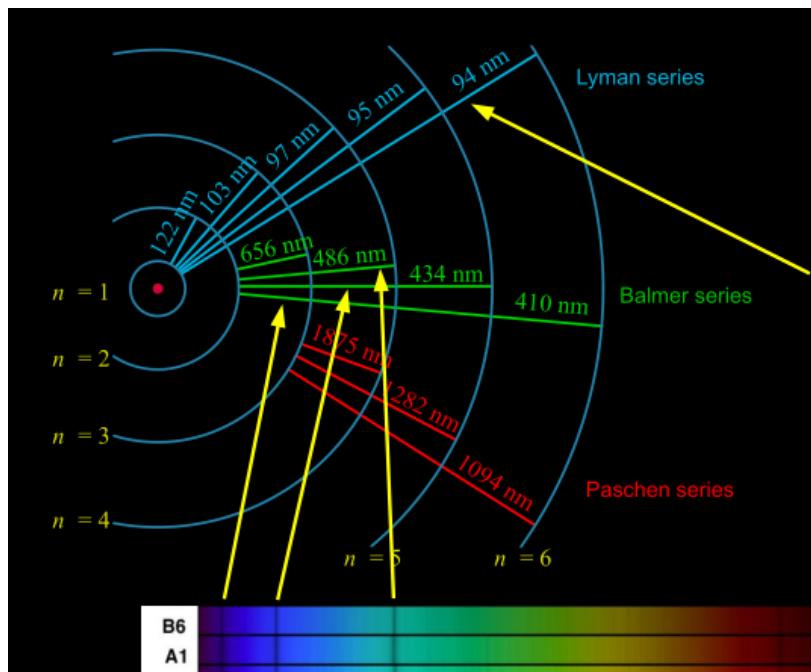


Figure 9. Various spectral series for hydrogen

Therefore cold stars would have largely hydrogen at $n = 1$ and therefore not have visible absorption lines. Can we go more specific and find the *probability* that an atom is in the $n = 2$ state instead of the $n = 1$ state?⁹

Consider the following system of a single atom¹⁰ in the rest of the Sun's atmosphere¹¹. What is the probability that this atom will be in a particular $n = 2$ mode compared to a $n = 1$ mode?¹²

$$\frac{P(s_2)}{P(s_1)} = \frac{\Omega_R(s_2)}{\Omega_R(s_1)} \quad (2.55)$$

⁹ Yes, or else we wouldn't be asking this question

¹⁰ the subsystem

¹¹ The reservoir

¹² Recall multiplicity of energy levels i.e. there are 4 different states with energy level of $n = 2$

Read $\Omega_R(s_i)$ to be the multiplicity of the entire system, sun *and* the atom, *when* the atom is in mode s_i . By definition the multiplicity of a mode is 1. But why would the multiplicity of the entire sun change when the atom is in different modes? Because the energy changes – when energy flows out of the sun reservoir into the atom to increase its energy level causes the sun's multiplicity to go down, and vice-versa.

Rearrange the prior expression a little bit to get it with respect to entropy

$$\begin{aligned} S \equiv k \ln \Omega \rightarrow \Omega &= e^{S/k} \\ \Rightarrow \frac{P(s_2)}{P(s_1)} &= \frac{\Omega_R(s_2)}{\Omega_R(s_1)} = \frac{e^{S_R(s_2)/k}}{e^{S_R(s_1)/k}} = e^{[S_R(s_2) - S_R(s_1)]/k} \end{aligned} \quad (2.56)$$

And then apply the thermodynamic identity¹³, noting that effects on volume is negligible and there is no transfer of particles. Therefore we arrive at

$$\begin{aligned} \frac{dS}{PdV \ll dU} &= \frac{1}{T} (dU + \underbrace{\mu dN}_{dN=0}) \end{aligned}$$

$$\begin{aligned} \Delta S &= \Delta U/T \quad \text{for small } \Delta S \\ &= S_R(s_2) - S_R(s_1) = \frac{1}{T} [\underbrace{U_R(s_2)}_{\text{energy in reservoir during } s_2} - U_R(s_1)] \quad (2.57) \\ &= -\frac{1}{T} [E(s_2) - E(s_1)] \quad (14) \end{aligned}$$

¹⁴ Adding energy to atom is removing from reservoir

Plugging this back into Eq. 2.56 we then get

$$\frac{P(s_2)}{P(s_1)} = e^{[S_R(s_2) - S_R(s_1)]/k} = e^{-\frac{1}{kT}(E(s_2) - E(s_1))} = \frac{e^{-E(s_2)/kT}}{e^{-E(s_1)/kT}} \quad (2.58)$$

Note that E is the energy level of the subsystem, and T is the temperature of the reservoir (not the subsystem). The energy of the subsystem also doesn't make too much sense since it is just a single thing.

Theorem 6

We then define the **Boltzmann Factor**, and the ratio of probabilities of two states is the ratio of their Boltzmann factors.

$$e^{-E(s)/kT} \quad (2.59)$$

And the *absolute probability* is some constant multiplied by the Boltzmann factor, i.e.

$$P(s_i) = \frac{1}{Z} e^{-E(s)/kT} \quad (2.60)$$

The constant Z describes the sum of the Boltzmann factors of all possible microstates, i.e. the partition function

$$Z = \sum_s e^{-E(s)/kT} \quad (2.61)$$

In other words the probability of a particular microstate is the Boltzmann factor of said microstate, divided by the sum of all Boltzmann factors for that system

In the past we talked about each microstate being equally likely. But now not every single subsystem in thermal equilibrium with a larger system is equally likely because the subsystem can give/take energy. So we now have to look at the likelihood that it can give/take energy to get to the higher energy state, which is given by Eq. 2.60

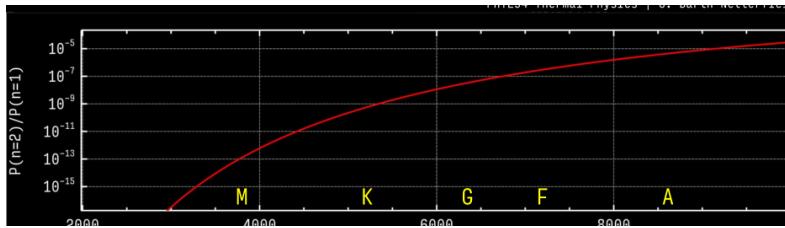
Example Let's try applying this to the Balmer series (Fig. 9). We know that the line strength is depends on the number of hydrogen atoms starting at the $n = 2$ level. The ΔE between $n = 1, 2$ is 10.2eV .

- The multiplicity of $n = 1$ is 1
- The multiplicity of $n = 2$ is 4

Plugging this into Eq. 2.58 we get

$$\frac{P(n=2)}{P(n=1)} = 4e^{-\Delta E/kT} \quad (2.62)$$

Which can be plotted as a function of T



And we can observe that hot stars (i.e. A-stars) are *much* more likely to have hydrogen atoms in the $n = 2$ state, whereas it is *very* unlikely for cool M-stars to have hydrogen atoms in those states; the $n = 2$ level is insufficiently populated for hydrogen absorption lines to appear.

This result makes sense given our observed spectral lines

SUBSECTION 2.10

Lecture 10: Average Values

The neat thing about the Boltzmann factor is that it can be used to find the average values of certain parameters.

Theorem 7

Consider some parameter $X(s)$ that is a function of only the quantum-mechanical mode the subsystem is in, e.g. energy, angular momentum, quantum number, velocity, etc..

$$\bar{X} = \sum_s P(s)X(s) = \frac{1}{Z} \sum_s X(s)e^{-E(s)/kT} \quad (2.63)$$

As an example we can look at the average energy of a system

Example

$$\begin{aligned} \bar{X} &= \sum_s P(s)X(s) = \frac{1}{Z} \sum_s X(s)e^{-E(s)/kT} \\ &= \frac{1}{Z} \sum_s X(s)e^{-E(s)\beta} \end{aligned} \quad (2.64)$$

And then let $X = E$; we're looking for average energy

Let's make the math easier by substituting $\beta = \frac{1}{kT}$

$$\begin{aligned}\overline{E} &= \frac{1}{Z} \sum_s E(s) e^{-E(s)\beta} \\ &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta}\end{aligned}\quad (2.65)$$

Recall how we prior used the equipartition rule and found that the degrees of freedom didn't really quite work out for diatomic molecules. We were able to wave our hands before and just use intuition with the atomic shapes to resolve this problem. Now we have the tools to resolve this problem more formally.

For a diatomic molecule, the energy levels are at $E(j) = j(j+1)\varepsilon$. There is a 2nd quantum number $k \in \{-j, j+1 \dots j-1, j\}$, so the degeneracy is $\Omega_j = 2j+1$

$$Z = \sum_s e^{-E(s)\beta} = \sum_j (2j+1) e^{-j(j+1)\varepsilon\beta} \quad (2.67)$$

Plugging this into what we found earlier for energies,

$$\overline{E} = \frac{1}{Z} \sum_s E(s) e^{-E(s)\beta} = \frac{1}{Z} \sum_j (2j+1) (j(j+1)\varepsilon) e^{-j(j+1)\varepsilon\beta} \quad (2.68)$$

It turns out that $\varepsilon\beta$ is really small at reasonable sizes and we can evaluate this numerically with modern computers.

Example

Let's try to calculate the average energy of H_2

Hydrogen turns out to be a little bit of a weird case where there are two semi-stable states; parahydrogen (both even j) and orthohydrogen (only odd j). So we would have to consider them separately.

Prof. Netterfield wrote some C code (Fig. 10) to calculate the expected energy.

```
#include <math.h>
#include <stdio.h>

#define kb 1.380640E-23
//#define e0 3.44523756E-23
#define h 6.626E-34

#define jmax 1000.0

#define MH 1.6708219E-27
#define RH (7.46-11/2.0)

#define p0 1.0
#define p1 1.0

int main() {
    double Z = 0;
    double Z0 = 0;
    double T;
    double E;
    double beta;
    double e0b;
    double E0b;
    double E0;
    double lastE = 0.0;
    double lastE0 = 0.0;

    double IH2 = MH*RH*RH*2;
    double hbar = h/2.0/M_PI;
    double e0 = hbar*hbar/2.0/IH2;

    for (T = 1.0; T<400.0; T+=1.0) {
        Z = 0.0;
        E = 0.0;
        Z0 = 0.0;
        E0 = 0.0;
        beta = 1.0/kb/T;
        e0b = beta*e0;

        // calculate Z
        for (j=0; j<jmax; j++) {
            Z += p0*(2.0*j+1.0)*exp(-j*(j+1.0)*e0b); // para
            j++;
            Z0 += p0*(2.0*j+1.0)*exp(-j*(j+1.0)*e0b); // ortho
        }

        // calculate E
        for (j=0; j<jmax; j++) {
            E += p0*(2.0*j+1.0)*(j+1.0)*e0b*exp(-j*(j+1.0)*e0b); // para
            j++;
            E0 += p0*(2.0*j+1.0)*j*(j+1.0)*e0b*exp(-j*(j+1.0)*e0b); // ortho
        }

        E = E/Z;
        E0 = E0/Z0;
        printf("%g %g %g %g", T, E, E*beta, (E-lastE)/kb);
        printf("%g %g %g\n", E0, E0*beta, (E0-lastE0)/kb);

        lastE = E;
        lastE0 = E0;
    }
}
```

Figure 10. C code for finding heat capacity for H_2

$$\begin{aligned}X &= \sum_s e^{-E(s)\beta} \\ \frac{\partial Z}{\partial \beta} &= -\sum_s E(s) e^{-E(s)\beta} \\ &= -E(s)Z\end{aligned}\quad (2.66)$$

This sum does converge since for large j the exponential goes to 0

Need to multiply the energy by the multiplicity of the state, $2j+1$

Measurement states $\frac{C_v}{k} = 0.95$ at 298K and $\frac{C_v}{k} = 0$ at 40K. Equipartition predicts $Cv_k = 0.5 * dof$

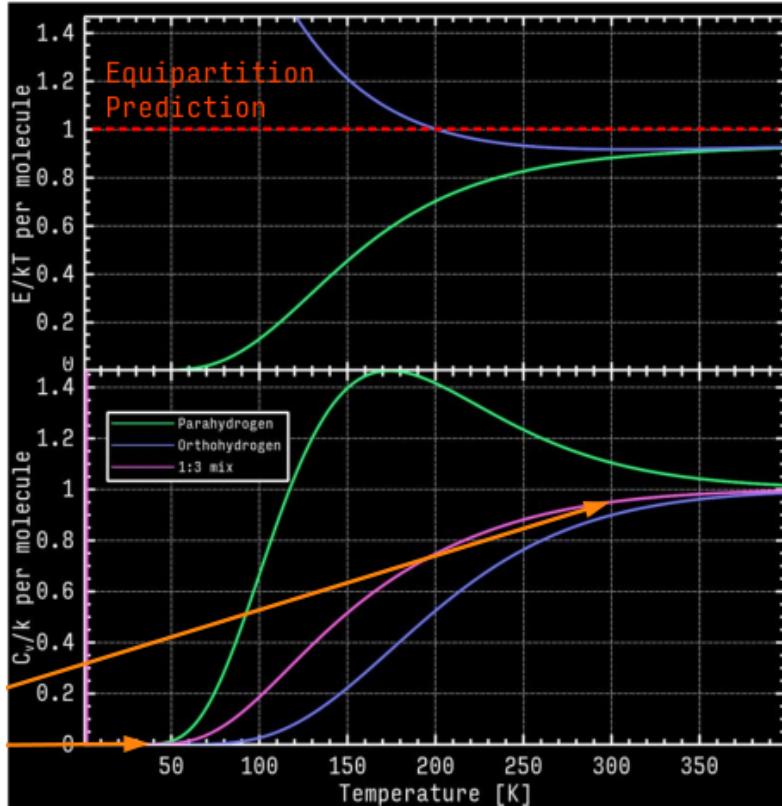


Figure 11. Plot of heat capacity vs temperature predicted through our statistical model

With this in hand we will try to derive a less hand-wavy equipartition theorem.

Consider a quantum system where the energy levels are given by $E_n = Cn^m$ where n is a positive integer and there is no degeneracy in the system. The number of particles and the volume is fixed $\implies \bar{E} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}$.

$$\begin{aligned}
 Z &= \sum_s e^{-E(s)/kT} = \sum_n e^{-Cn^m \beta} \\
 &\approx \int_0^\infty e^{-Cn^m \beta} dn \\
 &= (C\beta)^{-\frac{1}{m}} \int_0^\infty e^{-y^m} dy \\
 &= (CB)^{-\frac{1}{m}} I
 \end{aligned} \tag{2.69}$$

If $\frac{C}{kT} \ll 1$ then many terms are required for the inside of the sum to drop to 0, so let's do approximations. Change of variables $y^m = C\beta n^m$, $n = y(C\beta)^{-\frac{1}{m}}$. Also note that e^{-y^m} is just a constant that depends on m

And therefore we find, after a chunk of math

$$\bar{E} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = \dots \frac{kT}{m} \tag{2.70}$$

Which is exactly what we expect.

- Harmonic oscillator, equally spaced energy levels, $E_n = \varepsilon n$, $m = 1 \implies \bar{E} = kT$
- 1D particle in a box $E_n = \varepsilon n^2$, $m = 2 \approx \frac{kT}{2}$
- 2D particle box: breaks because we don't handle degeneracy.

But this breaks for more complex systems such as a particle in a 2D box. We go back to where we were previously but define the number of modes to be $N_s = Dn^a$.

This gives us this expression for the partition function

$$Z = \sum_s e^{-E(s)/kT} = \sum_n Dn^a e^{-Cn^b \beta} \quad (2.71)$$

Then we do the exact same stuff and then we will arrive at

$$\bar{E} = \frac{(a+1)kT}{b} \quad (2.72)$$

a, b are the same as in Eq. 2.71

So with Eq. 2.72 in hand we try our hand again at some examples

Example

Harmonic oscillator:

$$E_n = \varepsilon n, N_s \propto 1, a = 0, b = 1 \implies \bar{E} = kT \quad (2.73)$$

1D particle in a box

$$E_n = \varepsilon n^2, N_s \propto 1, a = 0, b = 2 \implies \bar{E} = \frac{kT}{2} \quad (2.74)$$

Diatomeric rotation

$$E = \varepsilon(j(j+1)); E_n \propto n^2, N_s \propto n, a = 1, b = 2 \implies \bar{E} = kT \quad (2.75)$$

So now we are no longer counting degrees of freedom but rather the way energy levels increase and the multiplicity changes.

3D particle for a box¹⁵

$$E = n^2, N_s \propto n^2, a = 2, b = 2 \implies \bar{E} = \frac{3kT}{2} \quad (2.76)$$

¹⁵Recall: energy levels given by $n_x^2 + n_y^2 + n_z^2$. Use polar, call it n_r^2

To sum it all up, in this lecture we applied the Boltzmann factor and the partition function Z in order to find the average values of certain parameters. An expression for average energy was then derived using the partition function, and the equipartition theorem was redefined in terms of more rigorous statistical mechanics.

SUBSECTION 2.11

Lecture 11: Maxwell Speed Distribution

Comment

Review of Boltzmann statistics

$$\frac{P(s_2)}{P(s_1)} = \frac{\Omega_R(s_2)}{\Omega_R(s_1)} = \frac{e^{S_R(s_2)/k}}{e^{S_R(s_1)/k}} \xrightarrow{dN=0, dV=0} \frac{e^{-E(s_2)/kT}}{e^{-E(s_1)/kT}} \quad (2.77)$$

Boltzmann factor

$$e^{-E(s)/kT} \quad (2.78)$$

Partition function

$$Z = \sum_s e^{-E(s)/kT} \quad (2.79)$$

And then the probability of a microstate s_1 is

$$P(s_i) = \frac{e^{-E(s_i)/kT}}{Z} \quad (2.80)$$

So we can get the average value of anything using

$$\bar{X} = \frac{1}{Z} \sum_s X(s) e^{-E(s)/kT} \quad (2.81)$$

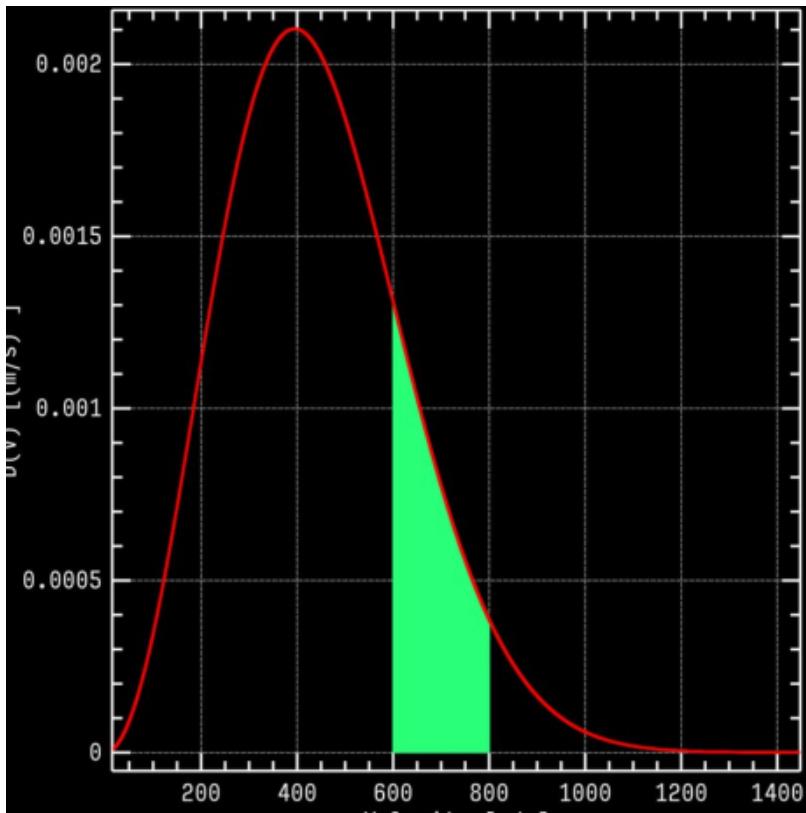
For example, for energy we get

$$\bar{E} = \frac{1}{Z} \sum_s E(s) e^{-E(s)/kT} = \frac{1}{Z} \frac{\partial Z}{\partial \beta} \quad (2.82)$$

Or,

$$\bar{E} = \frac{(a+1)kT}{b} \quad E_n = Cn^b, N_s = Dn^a, \frac{C}{kT} \ll 1 \quad (2.83)$$

For an ideal gas we get $\bar{E} = \frac{1}{2}mv^2 = \frac{3kT}{2}$ So $v_{rms} = \sqrt{\bar{v}^2} = \sqrt{\frac{3kT}{m}}$. But this only gives us the *average* speed. Instead we want to understand the *distribution* of particles instead.



We can actually use this to understand why the moon has no atmosphere, earth has no helium/hydrogen, and Jupiter has both

Figure 12. Let's imagine that this is what the speed distribution for an ideal gas

Our goal is to arrive at something like Fig. 12 where the x-axis describes the velocity and the y-axis describes the *probability density function*. The area under of the curve, i.e. the green area in the above figure gives the *absolute* probability of the speed to be between a range.

Definition 8**Maxwell speed distribution**

$$D(v) = 4\pi \left(\frac{m}{2\pi kT} \right)^{\frac{3}{2}} v^2 e^{-mv^2/2kt} \quad (2.84)$$

PROOF

The molecule has some speed, so

$$\int_0^\infty D(v)dv = 1 \quad (2.85)$$

And just by using our intuition, we can anticipate that $D(v)$ should be proportional to the probability of being in a state with speed v times the density of states with speed v . This is just the Boltzmann factor multiplied by the multiplicity.

$$D(v) \propto e^{-mv^2/2kT} \times N(v) \quad (2.86)$$

Next we have to find something more precise about $N(v)$

For a particle in a cubic box we have $E = E_o(n_x^2 + n_y^2 + n_z^2)$. If we were to define $r^2 = n_x^2 + n_y^2 + n_z^2$, then $E = E_o r^2$. Therefore $E = E_o r^2 = \frac{mv^2}{2} \rightarrow v \propto r$. Since the number of states between r, dr is equal to the volume of the shell with radius r to dr , so $N(r) \propto r^2 \implies N(v) \propto v^2$.

$$D(v) \propto e^{-mv^2/2kT} \times v^2 = Cv^2 e^{-mv^2/2kT} \quad (2.87)$$

With this in hand we can try our hand at the integral

$$\int_0^\infty Cv^2 e^{-mv^2/2kT} dv = 1 \quad (2.88)$$

Apply substitutions $x = v\sqrt{m/2kT}$, $v = x\sqrt{2kT/m}$

$$1 = C \left(\frac{2kT}{m} \right)^{3/2} \int_0^\infty x^2 e^{-x^2} dx \implies C = 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} \quad (2.89)$$

And plugging that into the original expression for $D(v)$ we get the Maxwell speed distribution

$$D(v) = 4\pi \left(\frac{m}{2\pi kT} \right)^{\frac{3}{2}} v^2 e^{-mv^2/2kt} \quad (2.90)$$

□

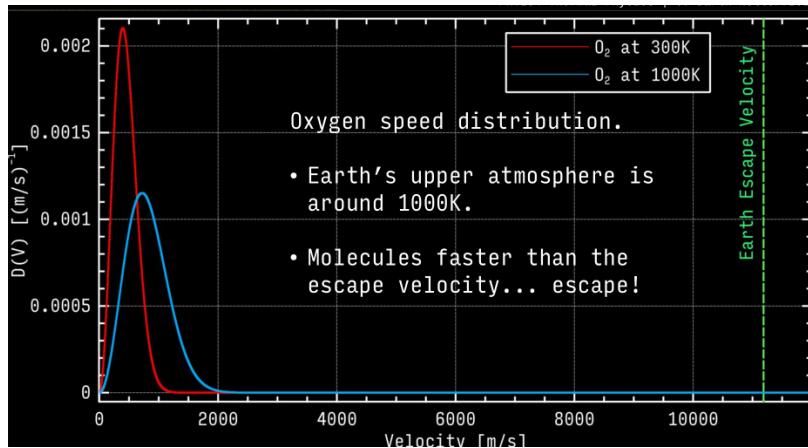
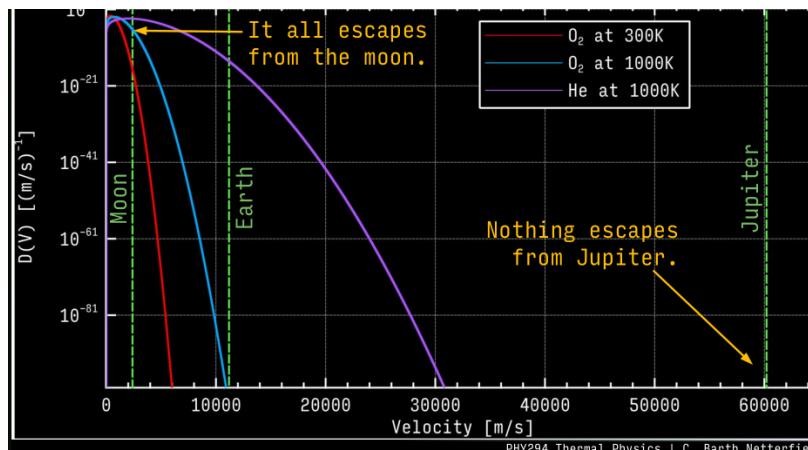


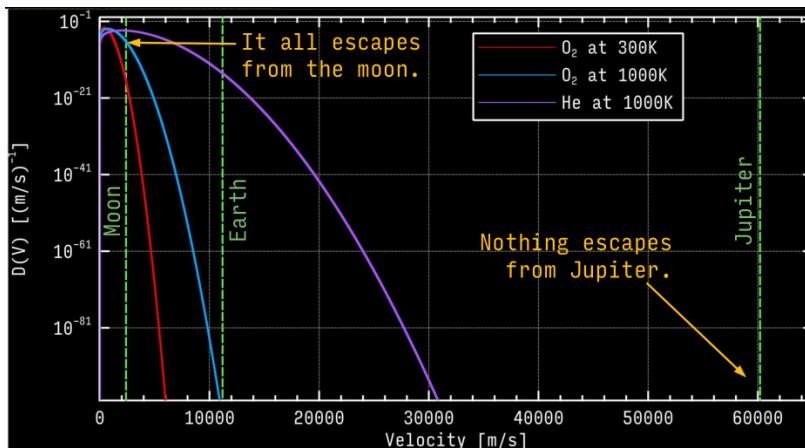
Figure 13. Maxwell distribution for Oxygen at room temperature

We see that a higher temperature gives a wider distribution centered around a higher velocity – makes sense. And for reasonable temperatures, oxygen's thermal velocity is basically always under Earth's escape velocity; what little that *could* escape is entirely insignificant (note log scale). So oxygen doesn't escape thermally.

However for the case of helium some of it exceed the escape velocity and so it leaves the atmosphere.



And we can repeat this for the moon and Jupiter and see that basically everything escapes from the moon and nothing escapes from Jupiter



SUBSECTION 2.12

Lecture 12: Free Energy

Comment

In the previous few lectures we have been working with Boltzmann factors in situations where we can enumerate all of the quantum modes. This enables us to use the partition function (Eq. 2.61) to find the probability of any state. This partition function is also useful in that it gives us an easy way to find the average energy; see Eq. 2.82.

As it turns out there are even more use cases for the partition function.

Previously we looked at diatomic molecules, for which the energy levels and multiplicity are given by:

$$E(j) = j(j+1)\varepsilon \quad \Omega_j = 2j+1 \quad (2.91)$$

And the partition function is therefore

$$Z = \sum_e e^{-E(s)\beta} = \sum_j (2j+1)e^{-j(j+1)\varepsilon\beta} \quad (2.92)$$

There are not many modes since $\varepsilon\beta$ becomes large really easily, so it is trivial to do it computationally. Or when $\varepsilon\beta$ is small we can just do it as a definite integral. So it is fairly straightforwards to calculate Z in any case.

So what more can we do with the partition function? Imagine s_1, s_2 represent the quantum states of two *distinguishable* non-interacting subsystems¹⁶.

As it turns out the combined partition function is just the product of the individual partition functions, i.e.

$$Z = \sum_s e^{-E(s)/kT} = \sum_{s_1} \sum_{s_2} e^{-\beta(E_1(s_1) + E_2(s_2))} = Z_1 Z_2 \quad (2.93)$$

Theorem 8

The partition function for i **distinguishable** non-interacting subsystems is given by

$$Z_{\text{tot}} = \prod_N Z_i \quad (2.94)$$

What if they are indistinguishable? This would just look like over-counting i.e. combinations with repetition.

¹⁶ This could look like s_1 being the translational mode of an O_2 molecule in a box and s_2 being the translational mode for a N_2 molecule in the same box. Or them being the rotational and translational mode of one particle

Theorem 9

The partition function for i **indistinguishable** non-interacting subsystems is given by

$$Z_{\text{tot}} = \frac{\prod_i Z_i}{N!} = \frac{Z_1^N}{N!} \quad (2.95)$$

If they are all indistinguishable then $\prod_i Z_i = Z_1^N$

A new idea that makes the partition function a lot more useful is Helmholtz free energy

Definition 9

Helmholtz Free Energy

$$F \equiv U - TS \quad (2.96)$$

And recall that $dU = TdS - PdV + \mu dN$ (see 2.50), so

$$dF = -SdT - PdV + \mu dN \quad (2.97)$$

Playing the same ‘game’ we did before in Lecture 2.7 with holding things constant, we find that if we can calculate F somehow, then F, P, μ can be found easily. And if we can find $F = F(Z)$, we can do a lot of cool things.

$$S = -\left(\frac{\partial F}{\partial T}\right|_{V,N} \quad P = -\left(\frac{\partial F}{\partial V}\right|_{T,N} \quad \mu = -\left(\frac{\partial F}{\partial N}\right|_{V,T}) \quad (2.98)$$

As it turns out F is a function of Z !

Definition 10

$$F \equiv U - TS = -kT \ln Z \quad (2.99)$$

This can be proved but it is boring and ugly and found in some section in the textbook.

The steps for knowing stuff are then as follows

1. Calculate Z
2. From Z , calculate F
3. From F , calculate C_v, U, S, P, μ

Example

Consider an ideal gas. First we find Z for a cubic box of length L . We can do this because we have prior calculated the energy levels of a particle in a 3D box.

$$E = \varepsilon(n_1^2 + n_2^2 + n_3^2) \quad \varepsilon = \frac{h^2}{8mL^2} \quad (2.100)$$

Now, put this into Eq. 2.82 we get

$$Z_{tr} = \sum_{n_1} \sum_{n_2} \sum_{n_3} e^{-\beta \varepsilon(n_1^2 + n_2^2 + n_3^2)} = \left(\sum_n e^{-\varepsilon \beta n^2} \right)^3 \quad (2.101)$$

It turns out that $\varepsilon \beta \ll 1$ so there are many terms so we can replace it with an integral. $\varepsilon \beta$ would only be large if it were really small or if the box were super small. In that case use computers.

$$Z_{tr} \approx \left(\int_0^\infty e^{-\varepsilon \beta n^2} dn \right)^3 = \frac{1}{\varepsilon \beta} \left(\int_0^\infty e^{-x^2} dx \right)^3 = \frac{V}{h^3} \left(\frac{2\pi m}{\beta} \right)^{3/2} \quad (2.102)$$

So, to summarize, we have

1. Translational modes: $Z_{tr} = \frac{V}{h^3} \frac{2\pi m}{\beta}^{3/2}$

If we can enumerate the quantum levels in a system, and N, V is fixed – but we can enumerate the quantum modes – we can calculate Z and with it a lot about the system

2. Diatomic rotational modes: $Z_{rot} = \sum_j (2j+1) e^{-j(j+1)\varepsilon_r \beta}$
3. Diatomic vibrational modes: $Z_{vib} = \sum_j e^{-(j+\frac{1}{2})\varepsilon_v \beta}$

These are three distinguishable subsystems and therefore we can find $Z_{tot} = Z_{tr}Z_{int}$ for one molecule in a box. If we had N indistinguishable molecules in a box, we would instead have to use Eq. 2.95 since the molecules are indistinguishable.

Define: $Z_{int} = Z_{rot}Z_{vib}$

$$Z_N = \frac{(Z_{tr}Z_{int})^N}{N!} \implies \log Z_N = N \left[\ln V - \frac{3}{2} \ln \beta + \frac{3}{2} \ln \frac{2\pi m}{h^2} + 1 - \underbrace{\ln N}_{\text{Stirling approx}} + \ln Z_{int} \right] \quad (2.103)$$

With Z_N in hand lets figure out other properties of the system.

$$\begin{aligned} U &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \xrightarrow{\text{log laws}} -\frac{\partial \ln Z}{\partial \beta} = \frac{3}{2} N k T + U_{int} \\ C_v &= \frac{\partial U}{\partial T} = \frac{3}{2} N k + \frac{\partial U_{int}}{\partial T} \quad \text{We solved this numerically in Lecture 10} \\ F &= k T \ln Z \\ &= k T N \left[\ln V - \frac{3}{2} \ln \beta + \frac{3}{2} \ln \frac{2\pi m}{h^2} + 1 - \ln N + \ln Z_{int} \right] \\ P &= \left(\frac{\partial F}{\partial x} \right)_{T,N} \implies P = \frac{k T N}{V} \quad \text{This is the ideal gas law!} \\ S &= - \left(\frac{\partial F}{\partial T} \right)_{N,V} = \dots \text{exercise for reader} \\ \mu &= - \left(\frac{\partial F}{\partial x} \right)_{T,V} = \dots \text{as above} \end{aligned} \quad (2.104)$$

And with this we found the ideal gas law with no hand waving! The same method can be applied to entropy or chemical potential of the system. This is incredibly powerful; with just Z we can find out so much about the system. However this is given that the volume does not change and N remains constant. This is with the assumption that the volume is small. What's next? Look into allowing N to change which would help model photons¹⁷.

¹⁷ which are not conserved

SUBSECTION 2.13

Lecture 13: Gibbs Factor

What if we want to allow for particles to come and go? E.g. if we were dealing with a chemical reaction.

Going back to the Boltzmann factor derivation and instead allowing for N to change, we may arrive at the Gibbs factor.

Theorem 10

The Gibbs factor is given by the following

$$e^{[E(s) - \mu N(s)]/kT} \quad (2.105)$$

PROOF

Ratio of probability is given by ratio of reservoir multiplicity in macrostate s_i

$$\frac{P(s_2)}{P(s_1)} = \frac{\Omega_R(s_2)}{\Omega_R(s_1)} \quad (2.106)$$

And from Eq. 2.27 we get $S \equiv \ln \Omega \implies \Omega = e^{S/k}$ and therefore Eq. 2.106 can be rewritten as

$$\frac{e^{S_R(s_2)/k}}{S_R(s_1)/k} = e^{[S_R(s_2) - S_R(s_1)]/k} \quad (2.107)$$

Going back to the thermodynamic identity, we keep μdN and apply $PdV \ll dU$

$$dS = \frac{1}{T}(dU - \mu dN) \quad (2.108)$$

And therefore we find that, for energy and number in the *reservoir* we get the following expressing for small ΔS

$$\Delta S = \frac{\Delta U - \mu \delta N}{T} \quad \text{for small } \Delta S \quad (2.109)$$

And so the change in entropy for the subsystem¹⁸ is given by

$$S_R(s_1) - S_R(s_1) = -\frac{1}{T} [E(s_2) - E(s_1) - \mu N(s_2) + \mu N(s_1)] \quad (2.110)$$

Noting that the above is for the subsystem, i.e. adding energy and particles to the subsystem is removing them from the reservoir.

Expanding on Eq. 2.106 we then get

$$\frac{P(s_2)}{P(s_1)} = \frac{e^{S_R(s_2)/k}}{e^{S_R(s_1)/k}} = \frac{e^{[E(s_2) - \mu N(s_2)]/kT}}{e^{[E(s_1) - \mu N(s_1)]/kT}} \quad (2.111)$$

And therefore the Gibbs factor is given by

$$e^{[E(s) - \mu N(s)]/kT} \quad (2.112)$$

Which is exactly the same as the Boltzmann factor except it takes into account N particles. And the grand partition function \mathcal{Z} is, like for the Boltzmann factor, given by the sum of all Gibbs factors

$$\mathcal{Z} = \sum_{s \in S} e^{[E(s) - \mu N(s)]/kT} \quad (2.113)$$

□

Note that subscript R denotes that reservoir

¹⁸Removing particles from the reservoir is adding them to the subsystem

E and N denotes the energy level and number of the subsystem, while T denotes the temperature of the reservoir

Example

As an example let's try to inspect a chemical process, i.e. looking at how our lungs may interact with carbon monoxide to lead to carbon monoxide posisioning. Let's imagine that in the lungs we have a location that can absorb or not absorb a oxygen molecule. It is at $E = 0$ while nothing is in it, and $E = -0.7eV$ when there is a oxygen bound to it, meaning that it takes energy to break free. The probability of it being occupied is the Gibbs factor for it being occupied over the grand partition function. Since there are only two states (open/closed), it should be easy to calculate.

$$P(O_2) = \frac{e^{-[E(s) - \mu N(s)]/kT}}{\mathcal{Z}} \quad (2.114)$$

At room temperature oxygen exerts roughly 0.2atm of partial pressure. Therefore μ is given

by the following equation derived from the ideal gas law and some fancy stuff. Z_{int} is found by finding the partition function for the rotational modes for oxygen (which is much like what we did for finding the heat capacity of hydrogen prior)

$$\mu = -kT \ln \left(\frac{V}{N} \frac{(2\pi mkT)^{3/2}}{h^3} Z_{int} \right) \approx -0.6\text{eV} \quad (2.115)$$

Gibbs factor in the empty state is given by 1 since if we were to evaluate Eq. 2.105 with $N = E = 0$ we get 1. In the state where there is an O_2 we get $e^{-(0.7\text{eV} - (-0.6\text{eV}))/kT} \approx 40$. And so $\mathcal{Z} = 1 + 50 = 41 \implies P(O_2) = \frac{40}{1+40} = 98\%$.

Repeating 2.115 with CO at 100 times lower pressure we find that $\mu_{CO} = -0.72\text{eV}$. In this system we now have three states; oxygen bound, carbon monoxide bound, and nothing bound. Given that $E = -0.85\text{eV}$ when carbon monoxide is bound to the receptor we repeat the calculations for the Gibbs factor and \mathcal{Z} to find that

1. $Gibbs(\text{empty}) = 1 \implies P(\text{empty}) = \frac{1}{161}$
2. $Gibbs(O_2) = 40 \implies P(O_2) = \frac{40}{161} \approx 0.25$
3. $Gibbs(CO) = 120 = \frac{120}{161} \approx 0.75$

Which means that only a very small amount of carbon monoxide needs to be introduced in order to screw up a lung's ability to absorb oxygen.

SUBSECTION 2.14

Lecture 14: Degenerate Fermi Gas

Definition 11

Fermions: only one particle allowed in a given quantum state, e.g. electrons, protons, neutrons, etc **Bosons:** no limit to number of particles allowed in a given quantum state, e.g. photons, pions

In this lecture we will be investigating Fermions in extreme conditions. Before we do that, as it turns out our previous expression for the number of indistinguishable subsystems (Eq. 2.95)¹⁹ is not quite right.

For example, imagine a quantum system with 2 particles and 3 quantum states. How would we deal with this?

$$^{19} Z_{tot} = \frac{Z_1^N}{N!}$$

$$Z = \sum^{\text{allowed quantum states}} e^{(E(s_1) + E(s_2))\beta} \quad (2.116)$$

If we directly summed over the combinations of 2 particles, 3 states, we would arrive at 9 terms. If we use the expression in Eq. 2.95, we would get 4.5 terms. This is obviously wrong since it is a non-integer.

Indistinguishable Fermions			Indistinguishable Bosons			
s_2	1	2	3	1	2	3
1	-	+	+	1	+	+
2	+	-	+	2	+	+
3	+	+	-	3	+	+

For fermions, there are 3 terms. For bosons, there are 6 terms.

is allowed.
is not allowed
+ is the same state as a +

Figure 14. Counting possible states. Greyed-out states are not allowed

We see that for fermions there are 3 allowed terms. For bosons there are 6 terms. This is basically a result of the different properties of Fermions and Bosons affecting those states in the diagonal²⁰. When n becomes large, the number of terms in the diagonal becomes a lot smaller than the total number of terms, so Eq. 2.95 can be rewritten as

$$Z_{\text{tot}} = \frac{Z_1^N}{N!} \quad Z \gg N \quad (2.117)$$

Definition 12

Fermi Dirac Distribution Function, or how can we deal with fermions when $Z \gg N$ is not true? Instead of thinking about of particles, instead think about the quantum modes instead, i.e. for a given quantum mode, how many particles are in $E = \varepsilon$?

We know that for fermions $N = 0$ or $N = 1$ since a fermion quantum mode can only have one particle. Since the number of particles is changing, we can use Gibbs factors²¹ to work with it.

- Empty $N = 0, E = 0 \Rightarrow 1$
- Full $N = 1, E = \varepsilon \Rightarrow e^{-[\varepsilon - \mu]/kT}$

And so we can find the grand partition function \mathcal{Z} by summing over the Gibbs factors

$$\mathcal{Z} = \sum_{\text{allowed quantum states}} = 1 + e^{-[\varepsilon - \mu]/kT} \quad (2.118)$$

And therefore we can now find the probability of a particle being in the mode, which happens to be the same as that of the *expected* number of particles in the mode

$$P(1) = \frac{1}{\mathcal{Z}} e^{-[\varepsilon - \mu]/kT} \quad (2.119)$$

$$\bar{n} = \sum_n n P(n) = 0 \cdot P(0) + 1 \cdot P(1) = \frac{e^{-[\varepsilon - \mu]/kT}}{1 + e^{-[\varepsilon - \mu]/kT}} = \frac{1}{1 + e^{-[\varepsilon - \mu]/kT}} \quad (2.120)$$

Where μ is the chemical potential of the mode, ε the energy of the mode, and T the temperature.

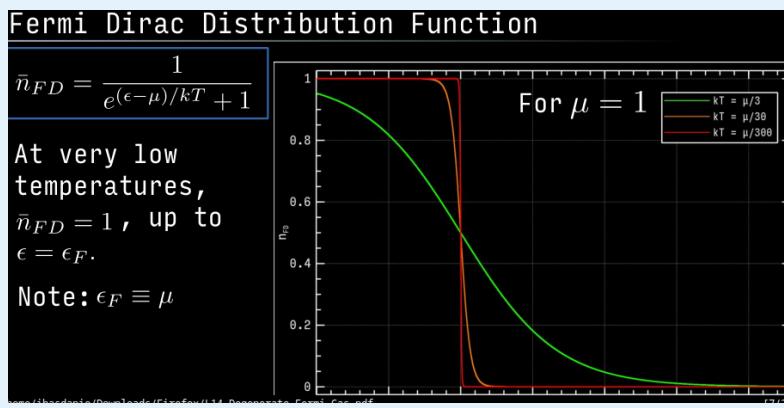


Figure 15. Note temperature dependence; at low temperatures it will fill low energy modes first, up to $\varepsilon_f \equiv \mu$

As an example we will look at neutron stars, which are gravitationally bound objects made only of neutrons. Gravity is holding it together, and neutrons are uncharged so no electrostatic repulsion. Note that neutrons are fermions; this is what's stopping it from collapsing to zero

²⁰Fermions: only one per quantum state, Bosons: multiple in a quantum state

²¹Eq. 2.105

size. Let's use the concept of the Fermi distribution to explain the behaviour of a neutron star.

Example Degenerate Fermi gases: neutron stars

- Assume that kinetic energy is much lesser than the chemical potential $kT \ll \mu$
- These are fermions, so every state is occupied to $\varepsilon = \varepsilon_f = \mu$. In Fig. 15 this would look somewhat like the red line

Assume the energy levels for particles in a box

$$\varepsilon = \varepsilon_0(n_x^2 + n_y^2 + n_z^2) \quad \varepsilon_0 = \frac{h^2}{8mV^2/3} \quad (2.121)$$

Giving N neutrons, this is $U = \sum_s \varepsilon_0(n_x^2 + n_y^2 + n_z^2) = \sum_s \varepsilon_0 n_r^2$ over the lowest N modes. Since N is extremely large, we can replace the sums with integrals.

Before we can get anywhere, however, we must find ε_f first²²

$$\varepsilon_f = \varepsilon(n_{max}) = \varepsilon_0 n_{max}^2 \quad (2.122)$$

What is n_{max} ? N would be the volume of $\frac{1}{8}$ of sphere of radius n_{max} , times two since there is spin up / spin down

$$N = 2 \cdot \frac{4}{8} \cdot \frac{4}{3} \pi n_{max}^3 \rightarrow n_{max} = (3N/\pi)^{(1/3)} \Rightarrow \varepsilon_f = \frac{h^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3} \quad (2.123)$$

Now we know n_{max} as a function of N as well as ε_f . Still need V . Before that let's find U . Replacing the sum with a spherical integral (noting the 2 due to spin) we get

$$\begin{aligned} U &= 2 \int_0^{\pi/2} d\phi \int_0^{\pi/2} \sin \theta d\theta \int_0^{n_{max}} \varepsilon(n) n^2 dn \\ &= \pi \int_0^{n_{max}} \varepsilon_0 n^4 dn = \frac{\pi h^2 n_{max}^5}{40mV^{2/3}} \end{aligned} \quad (2.124)$$

Plug in our expression for n_{max}

$$= \frac{3h^2}{40m} \left(\frac{3}{\pi V} \right)^{\frac{2}{3}} N^{5/3} = \frac{3}{5} N \varepsilon_f$$

V can be determined by finding the pressure first.

$$U = \left(\frac{3h^2}{40m} \left(\frac{3}{\pi} \right)^{\frac{2}{3}} N^{5/3} \right) V^{-2/3} = \frac{3}{5} N \varepsilon_f \quad (2.125)$$

Recalling the thermodynamic identity (Eq. 2.50) we get $P = -(\frac{\partial U}{\partial V})_{S,N}$ we can derive

$$P = \frac{h^2}{20m} \left(\frac{3}{\pi} \right)^{\frac{2}{3}} \left(\frac{N}{V} \right)^{5/3} \quad (2.126)$$

This is not the ideal gas law! It is a goes up faster with V and down faster with N . This is called *degeneracy pressure* and has nothing to do with electrostatic repulsion; only quantum effects.

$$U_k = \left(\frac{3h^2}{40m} \left(\frac{3}{\pi} \right)^{\frac{2}{3}} N^{5/3} \right) V^{-2/3} \quad \underbrace{\quad}_{\text{volume formula}} \quad \frac{3h^2}{40m} \left(\frac{3}{2\pi} \right)^{\frac{2}{3}} \frac{N^{5/3}}{R^2} \quad (2.127)$$

Then we note the boundary condition of this problem; the star will radiate energy until it enters a minimum energy state, i.e. when $U_k + U_g$ is minimized²³
I.e. minimize

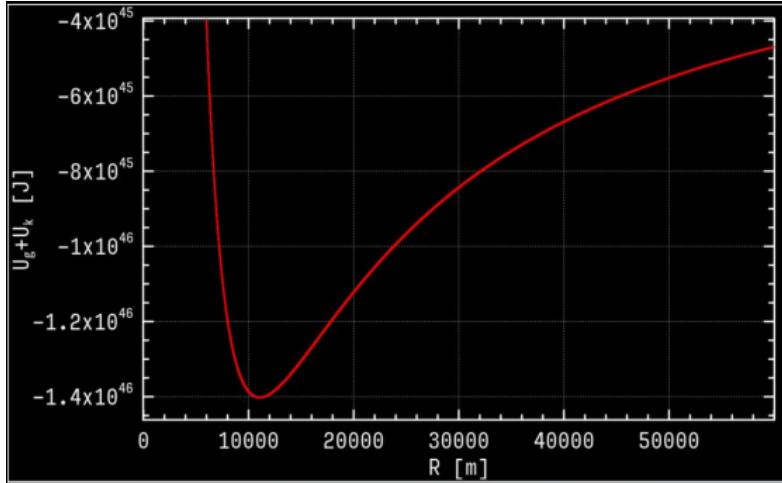
²² ε -Fermi, which denotes the upper limit on energy such that each state below is occupied

Remember, U is the kinetic energy so lets name it that way from now to be more clear

²³Recall: $U_g = -\frac{3GM^2}{5R} = -\frac{-3Gm^2N^2}{5R}$ where M is the stellar mass and m is a neutron mass

$$U = U_k + U_g = \frac{3h^2}{40m} \left(\frac{3}{2\pi}\right)^{\frac{2}{3}} \frac{N^{5/3}}{R^2} - \frac{-3Gm2N^2}{5R} \quad (2.128)$$

If we were to plot this for a reasonable 1.4 solar mass neutron star (which is pretty typical), we find that it has a minimum at about $R = 11\text{km}$, which is pretty neat.



When it is at a minimum,

$$\frac{\partial(U_k + U_g)}{\partial R} = 0 \Rightarrow \frac{\partial U_k}{\partial R} = -\frac{\partial U_g}{\partial R} \quad (2.129)$$

And after some math we get

$$R = \frac{h^2}{4Gm^3N^{1/3}} \left(\frac{3}{2\pi}\right)^{\frac{4}{3}} \quad (2.130)$$

Which for the neutron mass we get $R = 11\text{km}$, which is consistent with real-life observed data. This is *really* cool since we managed to go from a particle in the box and some statistics to finding the radius of a neutron star.

Some observations:

- Adding mass to a neutron star will decrease the radius of the star.
- ε_f increases with mass, but the degeneracy pressure cannot change
- But $\varepsilon_f = \frac{mv^2}{2}$, if it isn't relativistic. So $\rightarrow v = \sqrt{2\varepsilon_f/m} = 0.43c$
- As mass increases the particle speeds get relativistic. So if we get a bit more massive everything we did breaks down and it collapses and becomes a black hole. As it turns out there cannot be neutron stars much larger than a couple of solar masses.
- These approximations are valid for $T \ll \varepsilon/k \approx 1.0 \times 10^{12}\text{K}$. This is OK since our measurements show that the temperatures are below 10^6K
- We also assumed uniform density throughout the star, but it definitely isn't. If we wanted to be more precise we would have to solve the Schrodinger equation repeatedly as we went through the star with the different densities at different radii. But our assumptions still brought us pretty close!

SUBSECTION 2.15

Lecture 15: Black Body Radiation

Comment

What about bosons²⁴? How do we deal with bosons when $Z \gg N$ is not true?

²⁴Recall: bosons can share quantum states

2.15.1 Bosons

Like with Fermions, instead of thinking about particles, we think about quantum modes with $E = n\varepsilon$ $N = n$. For bosons n goes from $0 \rightarrow \infty$

The Gibbs factor for a boson is therefore

$$e^{-n[\varepsilon-\mu]/kT} = (e^{-[\varepsilon-\mu]/kT})^n \quad (2.131)$$

And we do the usual thing for \mathcal{Z} and $P(n)$

$$P(n) \frac{1}{\mathcal{Z}} e^{-n[\varepsilon-\mu]/kT} \quad \mathcal{Z} = \sum_{n=0}^{\infty} (e^{-[\varepsilon-\mu]/kT})^n \quad (2.132)$$

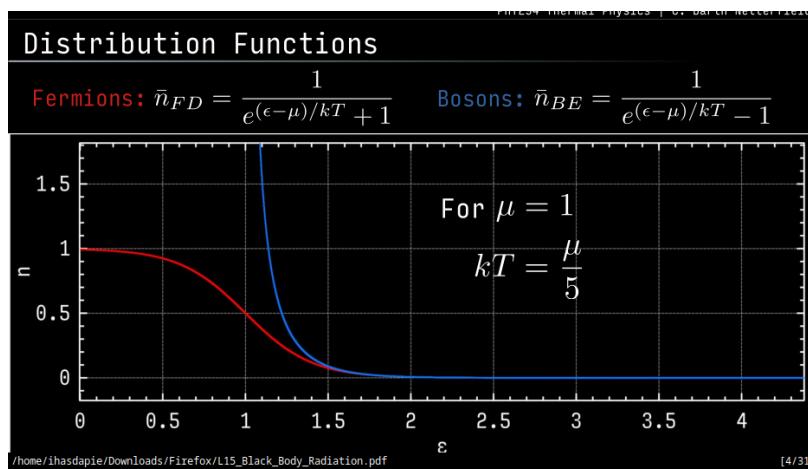
This can be evaluated with knowledge of infinite series $\sum_0^{\infty} x^n = \frac{1}{1-x}$ $x < 1$

$$\mathcal{Z} = \frac{1}{1 - e^{-[\varepsilon-\mu]/kT}} \quad (2.133)$$

And we can do some math to find \bar{n} to get the Bose-Einstein distribution

Definition 13

$$\bar{n}_{BE} = \frac{1}{e^{\frac{\varepsilon-\mu}{kT}} - 1} \quad (2.134)$$



When plotted we find that occupation n heads towards infinity at $\varepsilon = 1$. This makes sense since bosons can share modes n , so they would like to *all* be in the lowest energy state

2.15.2 Black Body Radiation

A black body is a body where all incident light is absorbed; none is reflected. We care because we observe that hot things glow – why?

An easy way to make a black body is to make a box with a small hole in it. The hole is very black, since the light, once it leaves the hole, is not coming back. Ignoring the hole, the energy density in the box can be found by the following procedure

1. Find all energy levels E_s of all photons that could be in the box
2. Find the occupancy \bar{n}_s for these energy levels
3. Apply $U = \sum_s E_s \bar{n}_s$; sum over every state/mode

Step 1: Find energy

Light in the box will take on the form of standing waves. From quantum mechanics we can find their energies using

$$p_x = \frac{h}{\lambda_x} \quad E = |p|c \quad (2.135)$$

So we will need their wavelengths, which by inspection²⁵ is

$$\lambda_x = \frac{2L}{j_x} \Rightarrow p_x = \frac{j_x h}{2L} \quad (2.136)$$

Where j_x is an integer from 1 to ∞ . E is trivial to find now and is just²⁶

$$E = |p|c = \frac{hc}{2L} \sqrt{j_x^2 + j_y^2 + j_z^2} = \frac{jhc}{2L} \quad (2.137)$$

²⁵ Standing wave, closed on both ends...

²⁶ let $j = \sqrt{j_x^2 + j_y^2 + j_z^2}$

Step 2: Find occupancy

Since photons are bosons, the occupancy of quantum modes is described by the Bose-Einstein distribution function (Eq. 2.134). We know ε , energy, for a photon but not the chemical potential μ

By definition^{27 28} $\mu = -T \left(\frac{\partial S}{\partial N} \right)_{U,V}$

But photons are not conserved! So nothing about the reservoir actually changes when we add a photo to the mode. So if the reservoir energy is constant, the reservoir's entropy doesn't change at all when you add a photon to the subsystem – therefore $\mu_{\text{photon}} = 0$.²⁹

So, the Bose-Einstein distribution function which describes the occupancy in this system Eq. 2.134 becomes (since $\mu = 0$)

$$\bar{n}_\lambda = \frac{1}{e^{\varepsilon/kT} - 1} \quad (2.138)$$

Step 3: Find total energy

Applying $U = \sum_s E \bar{n}$, we may build the following expression³⁰

$$U = 2 \sum_{j_x=1}^{\infty} \sum_{j_y=1}^{\infty} \sum_{j_z=1}^{\infty} \frac{jhc}{2L} \frac{1}{e^{jhc/2LkT} - 1} \quad (2.139)$$

And we know everything in it! So we can actually find it. To make our lives easier, turn it into an integral and use spherical coordinates

³¹

$$\begin{aligned} U &= 2 \int_{j_x=0}^{\infty} \int_{j_y=0}^{\infty} \int_{j_z=0}^{\infty} \frac{jhc}{2L} \frac{1}{e^{jhc/2LkT} - 1} dj_x dj_y dj_z \\ U &= 2 \int_{j=0}^{\infty} \int_{j_\theta=0}^{\pi/2} \int_{j_\phi=0}^{\pi/2} \frac{jhc}{2L} \frac{1}{e^{jhc/2LkT} - 1} j^2 \sin \theta dj d\theta d\phi \end{aligned} \quad (2.140)$$

Rewriting this in terms of $\varepsilon = \frac{jhc}{2L}$, we arrive at

$$U = \int_0^{\infty} \frac{8\pi L^3 \varepsilon^3}{(hc)^3 e^{\varepsilon/kT} - 1} d\varepsilon \quad (2.141)$$

²⁷ T is the reservoir temperature

²⁸ The partial denotes the change in reservoir entropy when photons are added/removed while energy and volume are kept constant

²⁹ We do not need to 'pay' when we make a photon – only need to pay for energy in it

³⁰ The 2 term is due to the two polarizations of transverse waves

³¹ We do something sketchy by changing the lower bound to 0 but if we did the math we would find that the value between 0-1 is really small and negligible.

2.15.3 Making sense of it

We may pull the $V = L^3$ term for simplicity

$$\frac{U}{V} = \int_0^{\infty} \frac{8\pi \varepsilon^3}{(hc)^3 e^{\varepsilon/kT} - 1} d\varepsilon \quad (2.142)$$

This shows the energy density of photos in an enclosed box with uniform wall temperature.

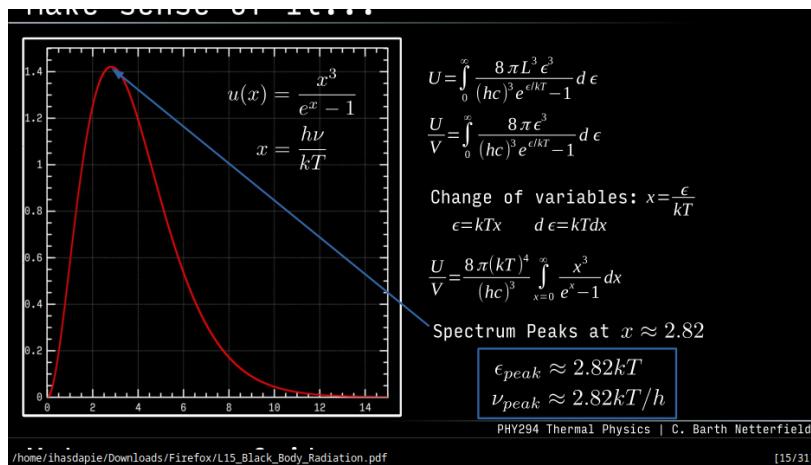


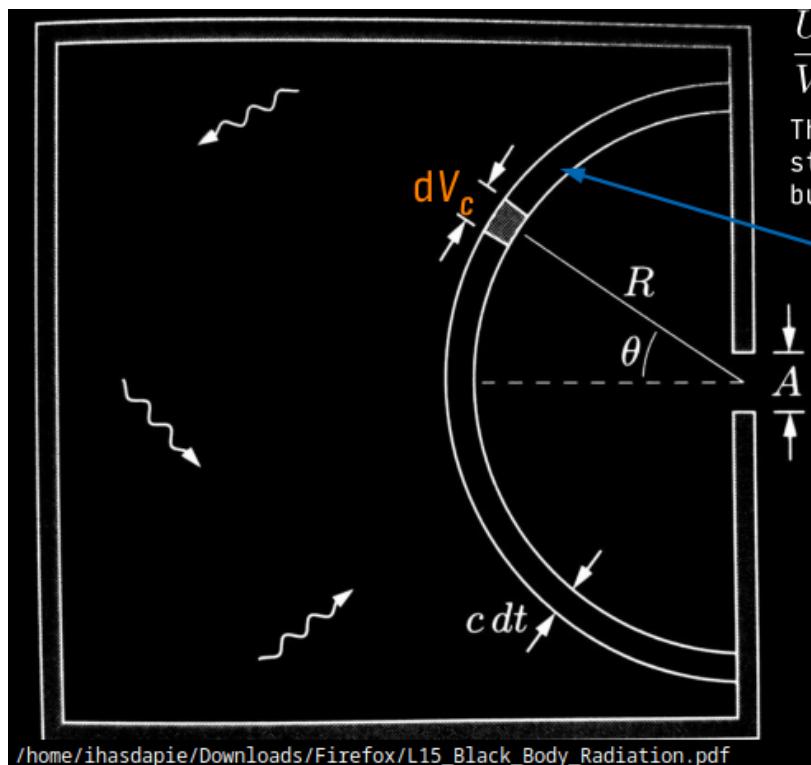
Figure 16. Light spectrum inside a box

Using a change of variables $x = \frac{\epsilon}{kT}$ we can write in a form with a nicer integral to solve in order to find an expression for the energy density of light inside *any* sealed box

$$\frac{U}{V} = \frac{8\pi (kT)^4}{(hc)^3} \int_{x=0}^\infty \frac{x^3}{e^x - 1} dx \xrightarrow{\int_{x=0}^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}} = \frac{8\pi^5 k^4}{15(hc)^3} T^4 \quad (2.143)$$

And we observe that this will peak at $x \approx 2.82$ which corresponds to $\epsilon_{peak} \approx 2.82kT$ and $\nu_{peak} \approx \frac{kT}{h}$

Now, what if we let the light out?



Let's poke an arbitrarily small hole into the box³². Thinking of the photons as rays instead of waves now, we note that any photon leaving the hole must travel through the region \mathbf{c} we defined. With this we define the following:

- $E_{\mathbf{c}}$: the energy in \mathbf{c} right now
- \mathcal{P} : probability of a photon in $d\mathbf{c}$

The photons that are in \mathbf{c} that escape are in region \mathbf{c} for time dt , so the energy that escapes per dt is $E_{\mathbf{c}}\mathcal{P}$.

$$E_{\mathbf{c}} = \frac{U}{V}V_{\mathbf{c}} \Rightarrow dE_{\mathbf{c}} = \frac{U}{V}dV_{\mathbf{c}} \quad \text{energy in } d\mathbf{c} \text{ right now} \quad (2.144)$$

The volume differential is given by

$$dV_{\mathbf{c}} = (Rd\theta) \times (R\sin\theta d\phi) \times (cdT) \quad (2.145)$$

And so

$$dE_{\mathbf{c}} = \frac{U}{V}dV_{\mathbf{c}} = \frac{U}{V}cdtR^22\sin\theta d\theta d\phi \quad (2.146)$$

We are interested in the probability of a photon in $d\mathbf{c}$ escaping without bouncing first. This would be the ratio of the area of the opening, viewed from θ to the area of the sphere centered on $d\mathbf{c}$

$$\mathcal{P}(\theta) = \frac{A \cos \theta}{4\pi R^2} \quad (2.147)$$

The energy in $dV_{\mathbf{c}}$ that escapes per unit time is given by the integral of $\mathcal{P}dE_{\mathbf{c}}$

$$E_{esc} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi/2} \frac{A \cos \theta}{4\pi R^2} \frac{U}{V} cdtR^2 \sin\theta d\theta d\phi = \frac{AU}{4V} cdt \quad (2.148)$$

And it is convenient to integrate again to get an expression for power to remove the unit-time

$$P = \frac{E_{esc}}{dt} = c \frac{AU}{4V} \quad (2.149)$$

And if we divide through with area, we find that

$$\frac{P}{A} = \frac{cU}{4V} = \frac{2\pi^5 k^4}{15h^3 c^2} T^4 = \sigma T^4 \quad (2.150)$$

So if we have a box, and it is dark³³, and it has a small hole in it, it is a light source that emits light carrying power equal to σT^4 . Note how the luminosity grows extremely fast with temperature. This turns out to be true for any black object, not just boxes with small holes in them.

³³ and not reflective!

2.15.4 Example: Cosmic background radiation

The universe is always expanding. Also, because of how light has a finite speed limit, the further away we look the further back in time we look. And since it is expanding, it is also cooling. It also happens to be that the universe is transparent. But if we look far enough, we can peek at when it wasn't [transparent], i.e. when the universe was a plasma – which is opaque and black! This sounds familiar – this plasma is a black body!

As it turns out we can look at it; this is the cosmic microwave background radiation.

This radiation is extremely red-shifted (Doppler shift) because of how the universe is expanding

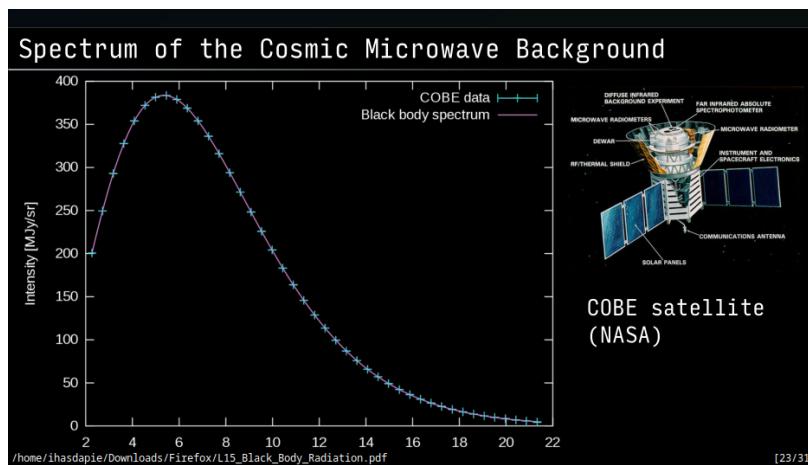


Figure 17. Observed cosmic microwave background radiation matches our model of black body radiation perfectly

And here are some pictures of Prof. Netterfield's balloons, which are a really cool low-cost alternative conventional telescopes.

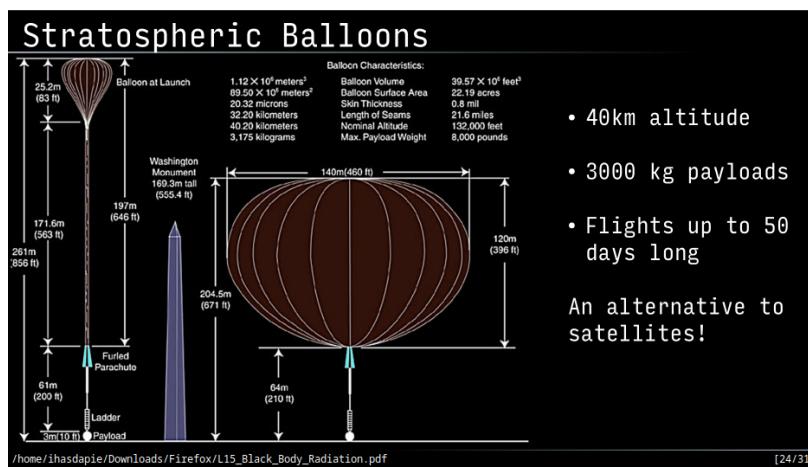


Figure 18. Launch from Antarctica

The pioneers used to ride these babies for miles

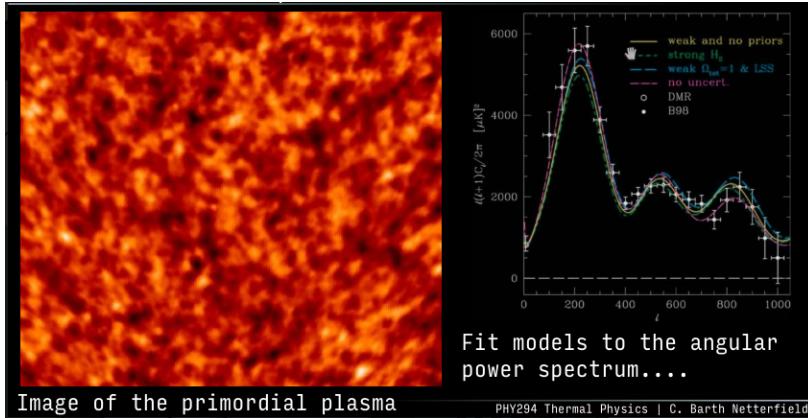


Figure 19. [very] enhanced image of the cosmic background radiation; lines up basically exactly with what we predict

And with this the model that best fits is that the universe is about 70% dark energy, 25% dark matter, and 5% normal matter, and that it is 13.7 billion years old. But the coolest result is that this shows that the universe is geometrically euclidean. Prior, Euclid's axioms of geometry may or may not have been true for our universe, depending on its energy density; it was not something that we could be absolutely sure about before. But now we know that, in our universe, two parallel lines will remain parallel forever.

2.15.5 Example: Star formation

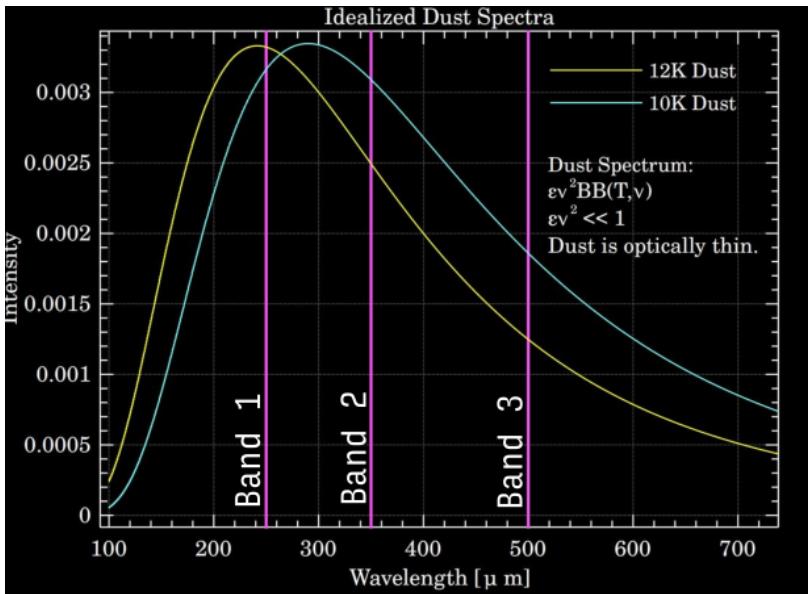
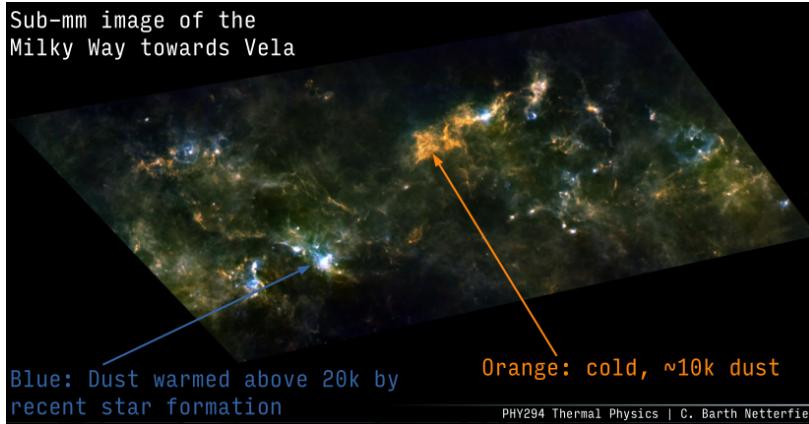


Figure 20. By taking images in bands, the temperature of interstellar dust can be determined (peaks are determined by the temperature of the black body)



And so we can trace where stars have recently been formed by finding the temperature of dust in interstellar space

SUBSECTION 2.1.1

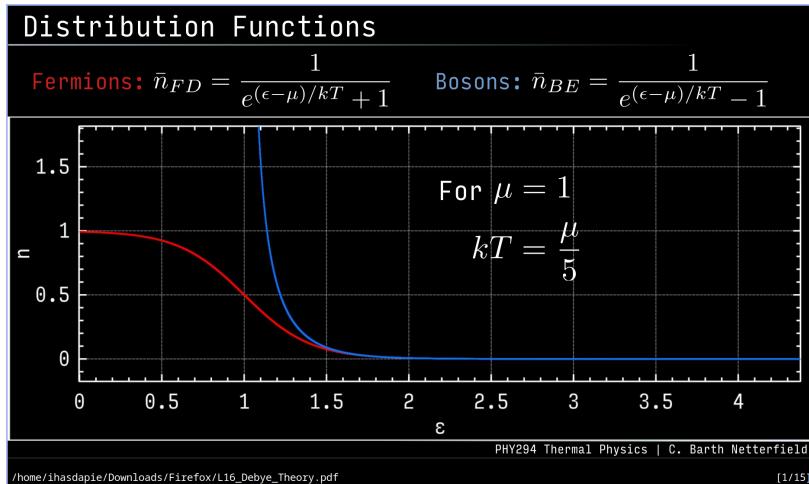
Lecture 16: Debye Theory

Comment

Grinding through some questions with both boson and fermion distributions

Recall the distributions for fermions and bosons:

$$\bar{n}_{FD}(\text{fermion}) = \frac{1}{e^{\frac{\epsilon-\mu}{kT}} + 1} \quad \bar{n}_{BE}(\text{boson}) = \frac{1}{e^{\frac{\epsilon-\mu}{kT}} - 1} \quad (2.151)$$



Today we will look into how we can actually calculate μ for a real-world system through looking at the heat capacity of a metal.

Recall what we know about kinetic energy and heat capacity

$$C_v = \frac{\partial U}{\partial T}_V \quad U = \sum_s \epsilon_s \bar{n} \quad (2.152)$$

So we will need to list out all of the quantum modes, multiply it by the appropriate distribution³⁴, and then take derivative with respect of temperature.

³⁴Fermion or Boson

As it turns out there are two ways to store energy in a metal; as a Fermi gas in the conduction electrons, and the vibrational modes in the solid.

2.16.1 Fermi gas heat capacity

$$\bar{n}_{FD} = \frac{1}{e^{\frac{\varepsilon-\mu}{kT}} + 1} \quad (2.153)$$

At very low temperatures, $\bar{n}_{FD} = 1$, up to $\varepsilon = \varepsilon_f = \mu$. At $T = 0$, $\frac{3h^2}{40m}(\frac{3}{\pi V})^{\frac{2}{3}}N^{5/3}$
So we can't assume the curve that looks like a step function anymore, it would look more like the green line (Fig. 15).

Recall particle in a box, $\varepsilon = \varepsilon_o(n_x^2 + n_y^2 + n_z^2)$ and $\varepsilon_o = \frac{h^2}{8mV^2/3}$

To calculate U (which is the internal energy of the Fermi gas of the electrons) we need to find all the quantum modes, and the sum over the energy in the quantum mode multiplied by its occupation (\bar{n})

$$U = \sum_s \varepsilon_s n_{FD} = 2 \sum_{N_x} \sum_{N_y} \sum_{N_z} \varepsilon \frac{1}{e^{\frac{\varepsilon-\mu}{kT}} + 1} \quad (2.154)$$

Because N is large we can turn this into an integral use spherical coordinates

$$U = \sum_s \varepsilon_s n_{FD} = 2 \int_0^\infty \int_0^{\pi/2} \int_0^{\pi/2} \varepsilon \frac{1}{e^{\frac{\varepsilon-\mu}{kT}} + 1} n^2 \sin \theta dnd\theta d\phi \quad (2.155)$$

And solving this integral we get

$$U(\mu, T) = \frac{\pi}{2\varepsilon_o^{3/2}} \int_{\varepsilon=0}^\infty \frac{\sqrt{\varepsilon}}{e^{\frac{\varepsilon-\mu}{kT}} + 1} d\varepsilon \quad (2.156)$$

This is a function of T , but we don't know what μ is. We can solve for N to find $N(\mu, T)$

$$N(\mu, T) = \sum_s n_{FD} = 2 \sum_{N_x} \sum_{N_y} \sum_{N_z} \varepsilon \frac{1}{e^{\frac{\varepsilon-\mu}{kT}} + 1} \dots = \frac{\pi}{2\varepsilon_o^{3/2}} \int_{\varepsilon=0}^\infty \frac{\sqrt{\varepsilon}}{e^{\frac{\varepsilon-\mu}{kT}} + 1} d\varepsilon \quad (2.157)$$

So, since we know what N is, we would have to solve it numerically with a computer by varying μ until the following function is true to our known values of N . So:

- Numerically solve $N(\mu, T)$ for $\mu(T), \mu(T + \Delta T)$
- Numerically solve $U(\mu, T)$ for $T, \Delta T$

And then we can find $C_v = \frac{\Delta U}{\Delta T}$ and we find that at low T , heat capacity is simply proportional to T ,

$$C_v = \gamma T \quad (2.158)$$

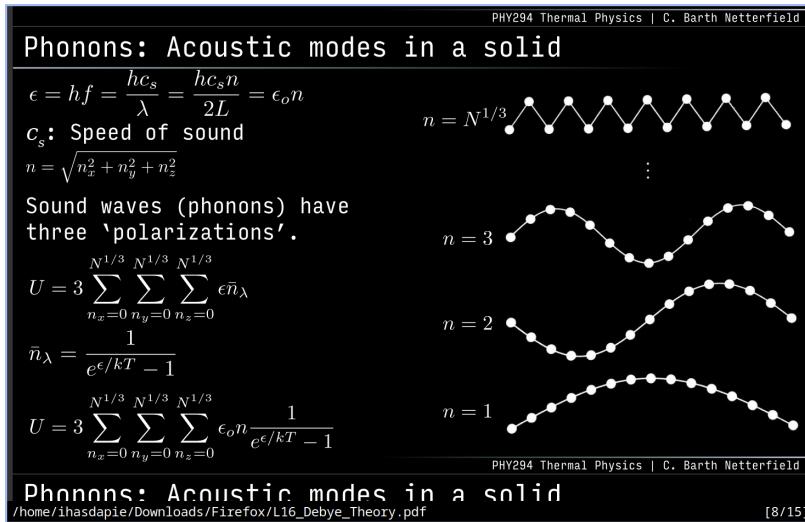
2.16.2 Vibrational mode heat capacity

In the past we used Einstein solids to model it. However this isn't the best model since the particles are not necessarily independent; in actuality the particles are coupled together and the vibrations will propagate instead of being the independent harmonic oscillators we used before – this looks like sound waves! Model this as acoustic waves.

We know that energy is given by $\varepsilon = hf$, so we can expand on it a bit to arrive at

$$\varepsilon = \frac{hc_s n}{2L} = \varepsilon_o n \quad (2.159)$$

Where c_s is the speed of sound.



What is reasonable for the highest frequency possible? We will say that the highest frequency is based on the shortest wavelength, which is based on the interatomic spacing; so approximately $N^{1/3}$

As it turns out, sound waves (phonons³⁵) have three polarizations; compression, transverse, and spatial orientation. Transverse waves only have two polarizations.

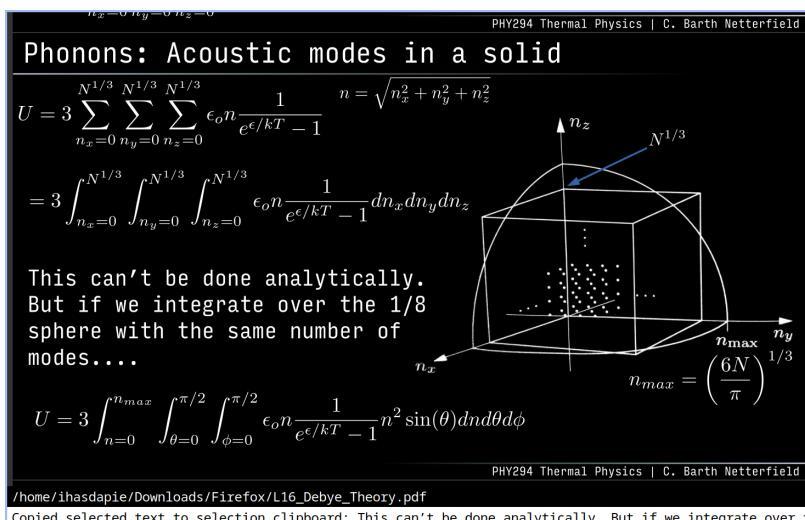
³⁵These are bosons

$$U = 3 \sum_{n_x=0}^{N^{1/3}} \sum_{n_y=0}^{N^{1/3}} \sum_{n_z=0}^{N^{1/3}} \epsilon \bar{n}_\lambda \quad (2.160)$$

And then we do the integral thing, note that $\bar{n}_\lambda = \frac{1}{e^{\epsilon/kT} - 1}$, and try to turn them into spherical integrals to make life better

$$U = 3 \int_{n_x=0}^{N^{1/3}} \int_{n_y=0}^{N^{1/3}} \int_{n_z=0}^{N^{1/3}} \epsilon \bar{n}_\lambda \frac{1}{e^{\epsilon/kT} - 1} dn_x dn_y dn_z \quad (2.161)$$

This can't be done analytically, and we also can't make it into a sphere since this is actually a cube.



But if we do an evil approximation thing and pretend like it is a 1/8 of a sphere with the same number of modes...we get at something that is alright at low temperatures

$$U = 3 \int_{n=0}^{N_{max}} \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{\pi/2} \varepsilon n \frac{1}{e^{\varepsilon/kT} - 1} n^2 \sin \theta dnd\theta d\phi \quad (2.162)$$

Where $n_{max} = \left(\frac{6N}{\pi}\right)^{1/3}$

And then do some math which I am not going to write out

PROBLEMS: ACOUSTIC MODES IN A SOLID

$$U = 3 \int_{n=0}^{n_{max}} \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{\pi/2} \varepsilon_o n \frac{1}{e^{\varepsilon_o n/kT} - 1} n^2 \sin(\theta) dnd\theta d\phi$$

$$n_{max} = \left(\frac{6N}{\pi}\right)^{1/3}$$

$$U = \frac{3\pi}{2} \int_{n=0}^{n_{max}} \varepsilon_o n^3 \frac{1}{e^{\varepsilon_o n/kT} - 1} dn$$

$$x \equiv \frac{\varepsilon_o}{kT} n \quad n = \frac{kT}{\varepsilon_o} x \quad x_{max} \equiv \frac{\varepsilon_o}{kT} n_{max} = \frac{\varepsilon_o}{kT} \left(\frac{6N}{\pi}\right)^{1/3} \equiv \frac{T_D}{T}$$

$$U = \frac{9NkT^4}{T_D^3} \int_0^{T_D/T} \frac{x^3}{e^x - 1} dx \quad T_D \equiv \frac{\varepsilon_o}{k} \left(\frac{6N}{\pi}\right)^{1/3} = \frac{hc_s}{2Lk} \left(\frac{6N}{\pi}\right)^{1/3}$$

PHY294 Thermal Physics | C. Barth Netterfield

/home/ihasdapie/Downloads/Firefox/L16_Debye_Theory.pdf [10/15]

To find that

$$U = \frac{9NkT^4}{T_D^3} \int_0^{T_D/T} \frac{x^3}{e^x - 1} dx \quad (2.163)$$

And now basically we can basically just take two limits

- $T \ll T_D$:

$$U \approx \frac{9NkT^4}{T_D^3} \int_0^{\infty} \frac{x^3}{e^x - 1} dx = \frac{3\pi^4}{4} \frac{NkT^4}{T_D^3} \quad (2.164)$$

And so $C_v = \frac{12\pi^4}{5} \left(\frac{T}{T_D}\right)^3 N_k$

- $T \gg T_D$ Assume $x \ll 1$ and $e^x \approx 1 + x$

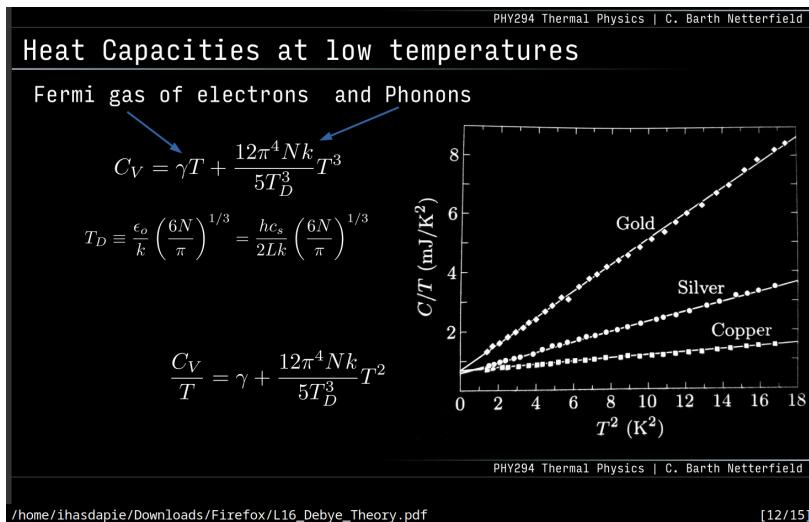
$$U \approx \frac{9NkT^4}{T_D^3} \int_0^{\infty} x^2 dx = \frac{3NkT^4}{T_D^3} (T_D/T)^3 \quad (2.165)$$

And so we get $U = 3NkT$ and $C_v = 3Nk$

And so we can sum up the results from the Fermi gas derivation previously to form a nice expression for C_v for metals at low temperatures

$$C_v = \gamma T + \frac{12\pi^4 N k}{5 T_D^3} T^3 \quad (2.166)$$

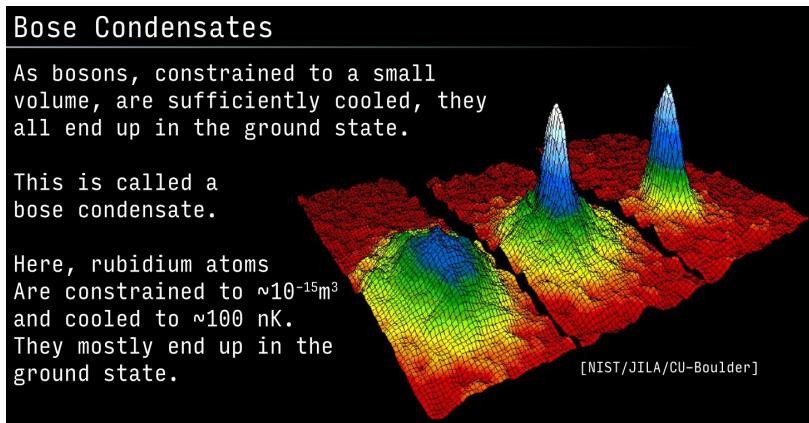
And it falls in line with experimental data as well!



SUBSECTION 2.17

Lecture 17: Bose-Einstein Condensates

We observe that bosons, when constrained to a small volume are cooled down enough, all take on the same ground state.



If we think about a single particle and use Boltzmann statistics, we could write out the Boltzmann factors like the following

$$\begin{aligned} \varepsilon &= \varepsilon(n_x^2 + n_y^2 + n_z^2) & \varepsilon_o &= \frac{\hbar^2}{8mV^{2/3}} \\ Z &= \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} e^{(n_x^2 + n_y^2 + n_z^2)\varepsilon_o/kT} & (2.167) \\ \Rightarrow P(1,1,1) &= e^{-3\varepsilon_o/kT}/Z \end{aligned}$$

For a single particle at $T = 100 \text{ nK}$, $V = 3 \cdot 10^{-16} \text{ m}^3$ this converges by $n_{max} = 100$, which results in $Z = 1228$ and $P(1,1,1) \approx 8 \cdot 10^{-4}$, which is a small chance (not what we expect).

Instead let's try this using distributions and occupation of quantum modes. Taking 2.134

There is a missing negative here somewhere...

$$\bar{n}_{BE} = \frac{1}{e^{\frac{\varepsilon - \mu}{kT}} - 1} \quad (2.168)$$

We just need to find μ , plug it in, and find the occupation. We know N ,

$$N = \sum_s \bar{n}_{BE} = \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} \frac{1}{e^{\frac{\varepsilon - \mu}{kT}} - 1} \quad (2.169)$$

So we can solve numerically with a computer program and find that for $N = 10000$, $T = 100nK$, $\mu = 2.56952 \times 10^{-32}$ For this the occupancy of the ground state is 7206, or about 72% of them, so Bose condensates should form! This is an interesting property of bosons; when only one particle is present, it is unlikely that it is in the ground state, but when we have many particles, it is very likely for many of them to be in the ground state. So, Bose condensates should, and *do*³⁶ form.

³⁶ *experimentally*

ECE259

Electromagnetism

SECTION 3

Dielectrics

Comment

A few useful things from before reading week:

- ϵ_0 : permittivity of free space
- ϵ_r : relative permittivity of the material
- ϵ is the *actual* permittivity of the material

And the three are related by

$$\epsilon_r = \epsilon / \epsilon_0 \quad (3.1)$$

The \vec{D} field is defined as

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P} \quad (3.2)$$

Noting that for a linear, isotropic³⁷ material

$$\begin{aligned} \vec{D} &= \epsilon_0 \vec{E} + \vec{P} \\ &\xrightarrow{\text{linear, isotropic}} = \epsilon_0 \vec{E} (1 + \chi_e) \\ &= \epsilon_0 \epsilon_r \vec{E} = \epsilon \vec{E} \end{aligned} \quad (3.3)$$

³⁷ isotropic: identical in all directions

In free space the generalized Gauss's law is given by

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho_p}{\epsilon_0} \quad \oint_s \vec{E} \cdot d\vec{s} = \frac{Q_{\text{end}}}{\epsilon_0} \quad (3.4)$$

And through a medium it is given by

$$\vec{\nabla} \cdot (\underbrace{\epsilon_0 \vec{E} + \vec{P}}_{\vec{D}}) = \rho_v \quad (3.5)$$

This note ρ_v ; this is for a volume charge density

When working with surface charge densities we have the following expressions

$$\vec{a}_n \cdot (\vec{D}_1 - \vec{D}_1) = \rho_s \quad (3.6)$$

Which leads us to the following fundamental postulates of electrostatics, in integral and differential form³⁸

$$\vec{\nabla} \cdot \vec{D} = \rho_v \Leftrightarrow \oint_s \vec{D} \cdot d\vec{s} = Q \quad (3.7)$$

³⁸ \vec{P} is the polarization vector

$$\vec{\nabla} \times \vec{E} = 0 \Leftrightarrow \oint_l \vec{E} \cdot d\vec{l} = 0 \quad (3.8)$$

\vec{P} is the measure of polarization in a dielectric material when in an electric field. It can be thought of as the electric dipole moment induced in the material per unit volume.

$$\vec{P} = \lim_{\Delta v \rightarrow 0} \frac{\sum_{k=1}^n \Delta v \vec{P}_k}{\Delta v} \quad \left[\frac{C}{m^2} \right] \quad (3.9)$$

$$\vec{E} = \epsilon_0 \chi_e \vec{E} \quad (3.10)$$

Voltage in differential and integral forms:

$$V = \int_l \vec{E} \cdot d\vec{l} \quad \vec{E} = -\vec{\nabla} V \quad (3.11)$$

And here is a really nice summary of vector calculus things

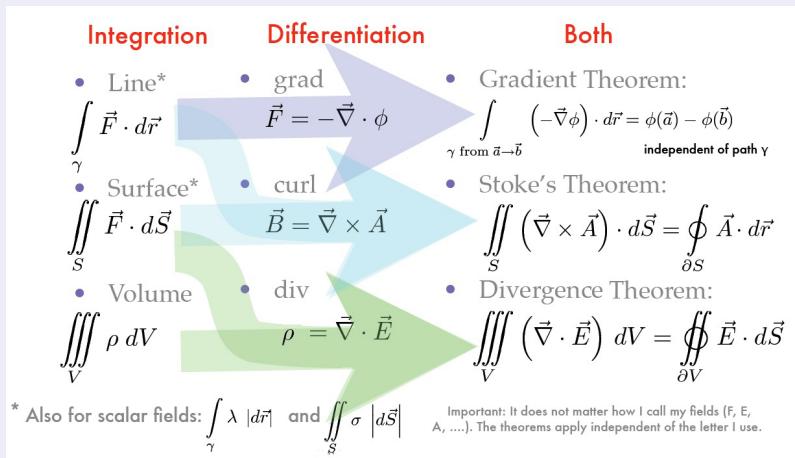


Figure 21. source

SUBSECTION 3.1

Lecture 15: Boundary Conditions for Dielectrics

Remark Recall, at a conductor/free-space interface, $E_+ = 0, E_n = \frac{\rho_s}{\epsilon_0}$

Consider an interface between two generic dielectrics (Fig. 22).

Theorem 11

$$\oint_c \vec{E} \cdot d\vec{l} = 0 \quad (3.12)$$

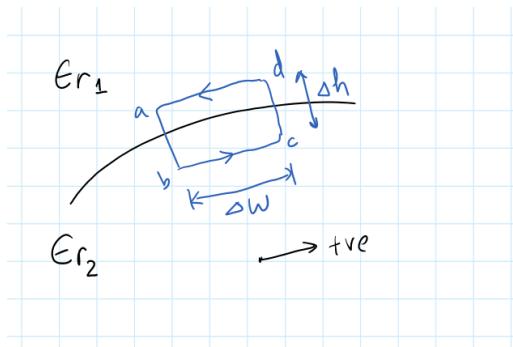


Figure 22. Dielectric interface

Intuition The two integrals parallel to the dielectrics will cancel out, and so will those perpendicular.

This implies that the tangential component of the \vec{E} field is continuous across the boundary,

$$E_{1t} = E_{2t} \quad (3.13)$$

However we get a bit of a different result when working in 3 dimensions where the interface is a surface instead of a line.

Theorem 12

$$\oint_S \vec{D} \cdot d\vec{s} = \rho_S \Delta \rightarrow (\vec{D}_1 - \vec{D}_2) \cdot \vec{a}_{n2} = \rho_S \quad (3.14)$$

a_n denotes normal component

Note: $\vec{D} = \epsilon_0 \vec{E} + \vec{p} = \dots \epsilon \vec{E}$

PROOF As $\Delta h \rightarrow 0$

$$\begin{aligned} \oint_S \vec{D} \cdot d\vec{s} &= \int_{top} \vec{D}_1 \cdot d\vec{s} + \int_{bottom} \vec{D}_2 \cdot d\vec{s} \\ &= \vec{D}_1 \cdot \vec{a}_{n2} \Delta S + \vec{D}_2 \cdot \vec{a}_{n1} \Delta S \\ &= \vec{D}_1 \cdot \vec{a}_{n2} \Delta S - \vec{D}_2 \cdot \vec{a}_{n2} \Delta S \\ &= \rho_S \Delta S \end{aligned} \quad (3.15)$$

□

Solving these problems usually involves finding the tangential and normal components through Eq. 3.13 and 3.14 then applying Pythagoras.

SUBSECTION 3.2

Lecture 16: Capacitors

Definition 14

A capacitor is a device consisting of two isolated conductors for storing energy in the form of electrostatic potential energy. An isolated conductor can also have "capacitance" if the other conductor is far away. The charge of a capacitance refers to the charge on one conductor. The energy stored in a capacitor is equal to the energy it takes to charge a capacitor from a discharged state to a charged state.

A capacitor's *capacitance*³⁹ is defined as

³⁹Capacitance is actually independent of Q and V and is dependent only on the physical attributes of [the capacitor]

$$C = \frac{Q}{V} \quad (3.16)$$

And has units $[C] = \frac{C}{V} F$ [Farads]

Capacitance is calculated as follows

1. choose a coordinate system
2. Assume $+Q / -Q$ on the conductors
3. Find \vec{E} from Q distribution
4. Find $V_{AB} = \int_A^B \vec{E} \cdot d\vec{l}$ where A carries the negative charge and B carries positive.
5. Apply $C = \frac{Q}{V}$

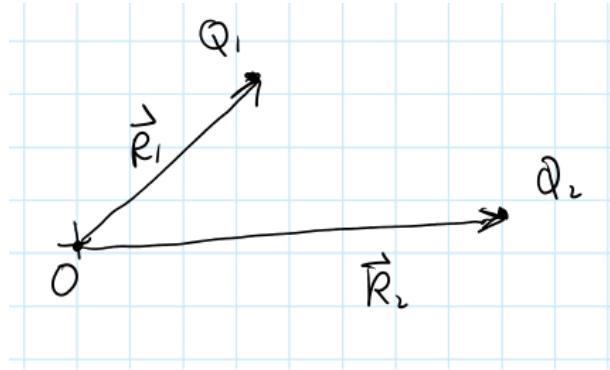
SUBSECTION 3.3

Lecture 17: Electrostatic Potential Energy

Remark Capacitors store energy because it takes energy to arrange charges in a particular way, i.e. separating positive and negative charges

1. How can we calculate the energy stored in a charge distribution?
2. How is this energy related to field quantities \vec{D} & \vec{E} ?

Consider two point charges Q_1 and Q_2



Assume Q_1 was there first and Q_2 is brought from $\infty \rightarrow \vec{R}_2$.

$$W_2 = Q_2 V_{\vec{R}_2} \quad (3.17) \quad \text{Potential at } \vec{R}_2 \text{ due to the field of } Q_1$$

And same vice-versa for Q_2

$$W_1 = Q_1 V_{\vec{R}_1} \quad (3.18) \quad \text{Potential at } \vec{R}_1 \text{ due to the field of } Q_2$$

Then the expressions for voltage may be built to get an expression for work just in terms of Q, R

$$V_{\vec{R}_2} = \frac{Q_1}{4\pi\epsilon_0 |\vec{R}_2 - \vec{R}_1|} \quad V_{\vec{R}_1} = \frac{Q_2}{4\pi\epsilon_0 |\vec{R}_1 - \vec{R}_2|} \quad (3.19)$$

$$W_2 = \frac{Q_2 Q_1}{4\pi\epsilon_0 |\vec{R}_2 - \vec{R}_1|} \quad (3.20)$$

But $W_2 = Q_1 \cdot V_{\vec{R}_1} = Q_2 V_{\vec{R}_2}$, so

$$W_2 = \frac{1}{2} \left(\frac{Q_2 Q_1}{4\pi\epsilon_0 |\vec{R}_2 - \vec{R}_1|} + \frac{Q_2 Q_1}{4\pi\epsilon_0 |\vec{R}_1 - \vec{R}_2|} \right) \quad \text{Note the subscripts on } \vec{R} \quad (3.21)$$

What if we introduced a third point charge Q_3 and brought it from ∞ to \vec{R}_3 ? The additional energy required is

$$\Delta W = Q_3 V_{\vec{R}_3} = Q_3 \left(\frac{Q_2 Q_1}{4\pi\epsilon_0 |\vec{R}_3 - \vec{R}_1|} + \frac{Q_2 Q_1}{4\pi\epsilon_0 |\vec{R}_3 - \vec{R}_2|} \right) \quad (3.22)$$

And then the total energy becomes

$$\begin{aligned} W_3 &= W_2 + \Delta W \\ &= \frac{1}{2} (Q_1 \left(\frac{Q_2}{4\pi\epsilon_0 |\vec{R}_1 - \vec{R}_2|} + \frac{Q_3}{4\pi\epsilon_0 |\vec{R}_1 - \vec{R}_3|} \right) \\ &\quad Q_2 \left(\frac{Q_1}{4\pi\epsilon_0 |\vec{R}_2 - \vec{R}_1|} + \frac{Q_3}{4\pi\epsilon_0 |\vec{R}_2 - \vec{R}_3|} \right) \\ &\quad Q_3 \left(\frac{Q_1}{4\pi\epsilon_0 |\vec{R}_3 - \vec{R}_1|} + \frac{Q_2}{4\pi\epsilon_0 |\vec{R}_3 - \vec{R}_2|} \right)) \\ &= \frac{1}{2} [Q_1 V_1 + Q_2 V_2 + Q_3 V_3] \end{aligned} \quad (3.23)$$

Theorem 13

In general, the interaction energy is

$$W_e = \frac{1}{2} \sum_{k=1}^N Q_k V_k = \frac{1}{2} \sum_{k=1}^N Q_k \underbrace{\left(\frac{1}{4\pi\epsilon_0} \sum_{j=1, j \neq k}^N \frac{Q_j}{|\vec{R}_k - \vec{R}_j|} \right)}_{V_k} \quad (3.24)$$

Note:

1. W_e can be negative
2. The expression here is only about interaction energy, not counting the energy to assemble the charges.

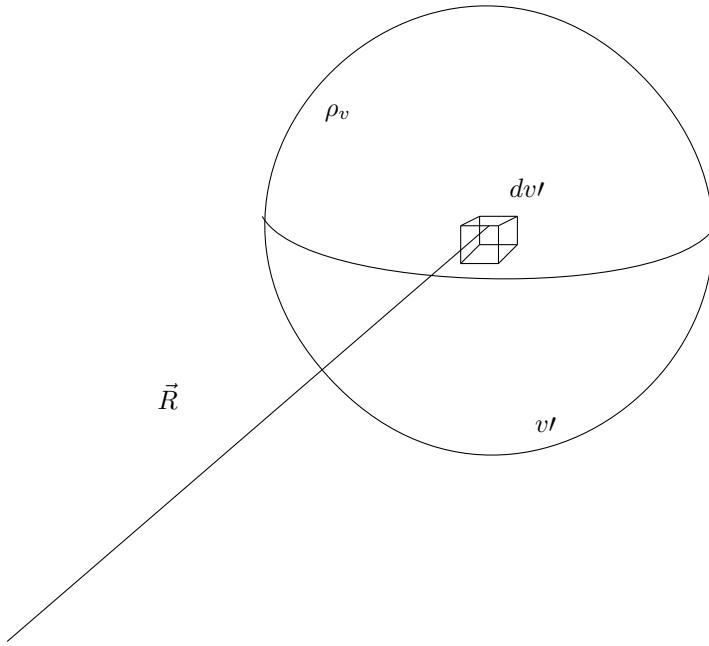
SUBSECTION 3.4

Lecture 18: More on Energy and Capacitors

Remark Recall, the potential energy stored in the system is given by

$$W_e = \frac{1}{2} \int_{v'} \rho_v V(\vec{R}) dv' = \frac{1}{2} \int_{\mathbb{R}^3} \vec{D} \cdot \vec{E} dv \quad (3.25) \quad \text{Can extend } v' \text{ to entire space}$$

Applying generalized gauss's law and performing an integral across the space we find that



$$\begin{aligned}
 W_e &= \frac{1}{2} \int_{v'} \rho_v V(\vec{R}) dv' \\
 &= \frac{1}{2} \int_{v'} (\vec{\nabla} \cdot \vec{D}) V(\vec{R}) dv' \\
 &= \frac{1}{2} \int_{v'} \vec{\nabla} \cdot (\vec{D} V) dv' - \frac{1}{2} \int_{v'} \vec{\nabla} \cdot \vec{D} V dv' \\
 &= \frac{1}{2} \oint_{S'} \vec{D} V \cdot d\vec{s} + \frac{1}{2} \int_{V'} \vec{D} \cdot \vec{E} dv' \\
 \Rightarrow W_e &= \frac{1}{2} \int_v \vec{D} \cdot \vec{E} dv
 \end{aligned} \tag{3.26}$$

Apply divergence theorem
The first term $\rightarrow 0$ as $R \rightarrow \infty$

Therefore the energy density can be defined as

Definition 15

Given energy density w_e with units $[J/m^3]$

$$W_e = \int_v w_e dv \quad w_e = \frac{1}{2} \vec{D} \cdot \vec{E} \tag{3.27}$$

Remark Things to know about capacitors

- Stacking them puts them in series. This looks like the inverse of resistors, i.e.

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} \dots \tag{3.28}$$

- Putting them besides each other puts them in parallel.

$$C = C_1 + C_2 \dots \tag{3.29}$$

SUBSECTION 3.5

Lecture 19: Poisson's Equation & Uniqueness Theorem

Comment

Previously we worked with problems like 'Given a charge density ρ , find \vec{D} , \vec{E} , V '. However a more typical problem we may face in the real world is: 'Given a charge distribution and potential known at *some* points, how can we find the fields?'

Definition 16

Poisson's Equation

$$\nabla^2 V = -\frac{\rho}{\epsilon} \quad (3.30)$$

Which is derived by substituting the definition of \vec{D} , Eq. 3.3 and Eq. 3.11 into the differential form of Gauss's law, Eq. 3.7:

$$\vec{\nabla} \cdot (\epsilon(-\vec{\nabla}V)) = \rho \quad (3.31)$$

And for no charge or $\rho_v = 0$, we have a special case⁴⁰: the 'Laplace equation'

$$\nabla^2 V = 0 \quad (3.32)$$

⁴⁰This is often found in non-charge-surface areas, e.g. between capacitor plates

Using this we can solve for potential as a boundary value problem. For a surface S enclosing a volume v with ρ_v , solving the following gives us the potential difference V inside v

$$\begin{cases} \nabla \cdot (\epsilon \nabla V) = \rho_v & \text{in } v \\ V_s = V_o \end{cases} \quad (3.33)$$

Theorem 14

Uniqueness Theorem: There is only one solution to Poisson's equation⁴¹ for a given set of sources and boundary conditions

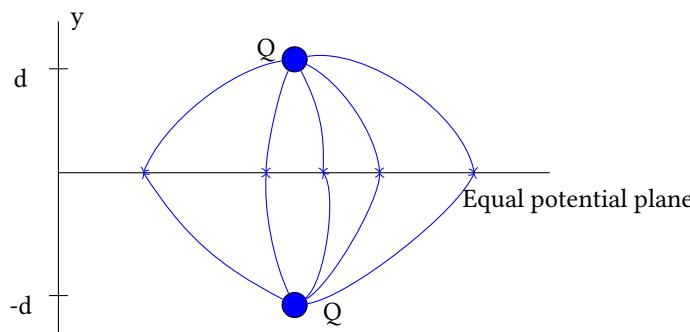
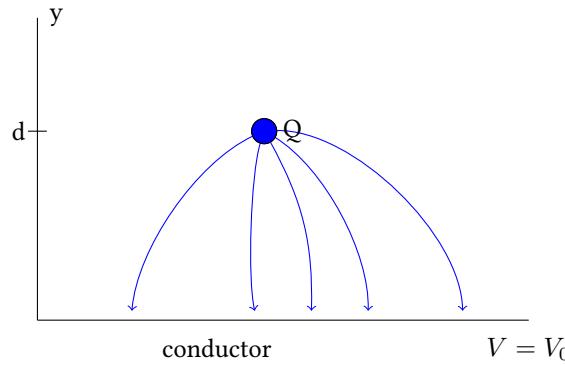
⁴¹of which the Laplace equation is a special case

SUBSECTION 3.6

Lecture 20: Method of Images

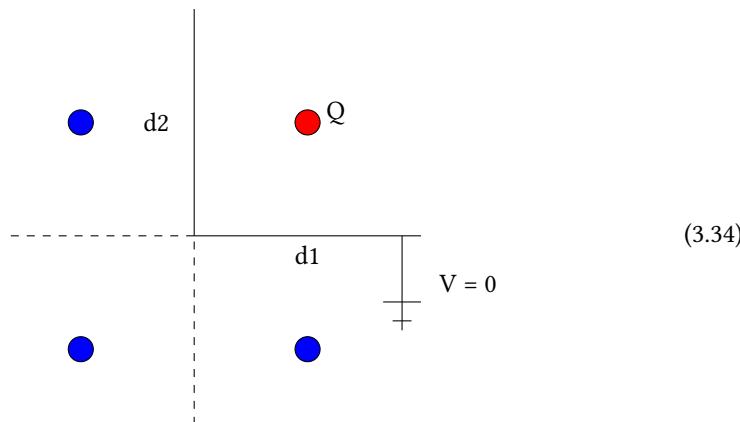
Remark

TLDR: Add imaginary charges such that the field lines meet perpendicular to the planes/lines of interest. Note that this operates on individual charges, not dipoles.



And for the upper-half space ($y > 0$) the two cases have the same source distribution and the same boundary condition on the conducting surface.

Example A charge Q is at (d_1, d_2) relative to two conductor walls. Find force on charge Q due to induced charges on conductor surfaces.



We have to be careful about where to put these virtual charges. They are arranged this way so that there is a right angle at the conductors; that's why we don't just put a single charge in the bottom left quadrant.

The force can be found as follows:

$$\vec{F} = \frac{Q}{4\pi\epsilon} \left[\frac{-Q}{2d_1^2} \hat{x} + \frac{-Q}{2d_2^2} \hat{y} + \frac{Q(2d_1 \hat{x} + 2d_2 \hat{y})}{[(2d_1)^2 + (2d_2)^2]^{3/2}} \right] \quad (3.35)$$

3.6.1 Currents

There are a few types of currents:

- Conduction current: In conductors and semiconductors due to the motion of e^- and holes h ⁴²
- Electrolytic current: motion of ions in solutions
- Convection current: motion of e^- or ions in vacuum

In conductors and semiconductors the free charge is

$$e = 1.6 \times 10^{-19} C \quad (3.36)$$

In the presence of an external field E ,

$$\begin{cases} \vec{E} = 0 & \text{random motion of free charges, current} = 0 \\ \vec{E} > 0 & \vec{\mu}_e = -\mu_o \vec{E}^{43} \end{cases} \quad (3.37)$$

⁴² We will primarily focus on conduction currents

Definition 17

General integral form for current

$$I = \int_S \rho_e \vec{\mu}_e \cdot d\vec{s} \quad (3.38)$$

PROOF Current is the amount of charge through a surface S per unit time

$$d\vec{S} = \vec{a}_n dS$$

The free charge in the volume of

$$\vec{\mu} \cdot \vec{a}_n dt dS \quad (3.39)$$

Will pass S in dt time

$$dq = -N_e e \vec{\mu}_e \cdot \vec{a}_n dt dS \quad (3.40)$$

Apply definition of current

$$dI = \frac{dq}{dt} = -N_e e \vec{\mu}_e \cdot \vec{a}_n dt dS = \rho_e \vec{\mu}_e \cdot d\vec{s} \quad (3.41)$$

And integrate!

$$I = \int_S \rho_e \vec{\mu}_e \cdot d\vec{s} \quad (3.42)$$

The change in current volume is the number of electrons per unit volume times volume

This integral form can be a little awkward to work with so we define current density J

Definition 18 Define current density to be \vec{J} such that

$$I = \int_S \vec{J} \cdot d\vec{s} \quad (3.43)$$

And therefore the volume current density \vec{J} with units $\frac{A}{m^2}$

$$\vec{J} = \rho_e \vec{\mu}_e \quad (3.44)$$

- \vec{J} is a vector and direction is the same as $\vec{\mu}_e$ and is a point function
- I is a scalar and is not a point function

Now ohm's law can be defined more rigorously in point/microscopic form

$$\begin{aligned}\vec{J} &= \rho_e \vec{\mu}_e = -\rho_o \mu_e \vec{E} \\ \Rightarrow J &= \sigma \vec{E}\end{aligned}\tag{3.45}$$

σ denotes conductivity

SUBSECTION 3.7

Lecture 21: Joule's Law and Resistance

Comment

Previously we discussed perfect conductors. But what if that was not the case?

Theorem 15

Joule's Law

$$P = \int_v dp = \int_v \vec{E} \cdot \vec{J} dv\tag{3.46}$$

Intuition: Work done by E-field is turned into kinetic energy of e^- and produces heat

PROOF

Work done by E field on one e^- : $\Delta W = -e \vec{E} \Delta l$

Power per unit time: $p = \frac{\Delta W}{\Delta t} = -e \vec{E} \frac{\Delta l}{\Delta t} = -e \vec{E} \cdot \vec{\mu}_e$

$$dp = N_e (-e) \vec{E} \cdot \vec{\mu}_e dv = \rho_v \vec{E} \cdot dv \Rightarrow dp = \vec{E} \cdot \vec{J} dv\tag{3.47}$$

Therefore the dissipated power density $\frac{dp}{dv} = \vec{E} \cdot \vec{J}$

And then one can integrate over the volume to find the total dissipated power; joule's law

$$P = \int_v dp = \int_v \vec{E} \cdot \vec{J} dv\tag{3.48}$$

Noting that we can take the volume integral with respect to a surface and a length, i.e. $dv = dl ds$, we can write this integral as

$$P = \underbrace{\left(\int_l \vec{E} \cdot d\vec{l} \right)}_{\text{voltage drop}} \underbrace{\left(\int_S \vec{J} \cdot d\vec{s} \right)}_{\text{current}}\tag{3.49}$$

Recall: μ is the average drift velocity

Recall: $\rho_v \vec{\mu}_e = \vec{J}$

This enables us to define resistance R .

Definition 19

Resistance

$$R = \frac{V}{I} = \frac{\int_l \vec{E} \cdot d\vec{l}}{\int_S \vec{J} \cdot d\vec{s}} = \frac{\int_l \vec{E} \cdot d\vec{l}}{\int_S \sigma \vec{E} \cdot d\vec{s}}\tag{3.50}$$

Resistance is independent of the voltage and current that passes through it and instead is dependent on the physical properties of the resistor

Example

For example, assuming uniform \vec{E}, \vec{J} ,

$$R = \frac{\int_l \vec{E} \cdot d\vec{l}}{\int_S \vec{J} \cdot d\vec{s}} = \frac{El}{\sigma E \cdot S} = \frac{l}{\sigma S}\tag{3.51}$$

Observe that resistance is proportional to the resistor length and dimensions l, S as well as

| the material properties σ .

The steps to calculate resistance are as follows:

1. Choose coordinate system
2. Assume V_s : potential drop between terminals
3. Find \vec{E} from V . This can get complicated; need to start with the Laplace equation $\nabla^2 V = 0$ and solve for \vec{E} everywhere using the fact that $\vec{E} = -\nabla V$
4. Find current: $I = \int_S \vec{J} \cdot d\vec{s} = \int_S \sigma \vec{E} \cdot d\vec{s}$
5. Apply definition $R = \frac{V}{I}$

This makes intuitive sense, no?

SUBSECTION 3.8

Lecture 22: Continuity Equation & BC for current density

Definition 20

The continuity equation is just an extension of the principle of conservation of charge, i.e. that charge cannot be created nor destroyed. As such it describes the flow of current and charge through the surface of a differential volume element, stating that the flow of charge out of the volume must equal the decrease in charge in the volume and vice-versa.

$$I_{\text{out}} = \oint_s \cdot d\vec{s} = -\frac{dQ}{dt} = -\frac{d}{dt} \int_v \rho dv \quad (3.52)$$

Or in differential form,

$$\nabla \cdot J = -\frac{\partial \rho}{\partial t} \quad \left[\frac{A}{m^3} \right] \quad (3.53)$$

- I_{out} : current out of the volume
- $-\frac{dQ}{dt}$: rate of change in charge inside the volume

PROOF

$$\oint_s \vec{J} \cdot d\vec{s} \xrightarrow{\text{divergence theorem}} \int_v \vec{\nabla} \cdot \vec{J} dv = -\int_v \frac{\partial \rho}{\partial t} dv \quad (3.54)$$

Partial differentiation is used now because the charge density ρ can be a function of time as well as space. \square

Recall: the divergence theorem is a mathematical statement of the physical fact that, in the absence of the creation or destruction of matter, the density within a region of space can change only by having it flow into or away from the region through its boundary⁴⁴

$$\iiint_V (\nabla \cdot F) dV = \iint_S (F \cdot n) dS \quad (3.55)$$

We may then extend on the work done in 3.54 to find

$$\nabla \cdot J = -\frac{\partial \rho}{\partial t} \quad \left[\frac{A}{m^3} \right] \quad (3.56)$$

In steady state this gives Kirchhoff's current law;

$$\sum_j I_j = 0 \quad (3.57)$$

The current leaving the region is the total outward flux of the current density vector through a surface S . Recall: flux is a vector field through a surface, i.e. $\int_S \vec{F} dS$ where S is a surface and \vec{F} a vector field. Often this can be solved with the help of the divergence theorem

⁴⁴<https://mathworld.wolfram.com/DivergenceTheorem.html>

wolfram.com/

DivergenceTheorem.html

For a steady current the charge density is time-constant therefore $\nabla \cdot J = 0$.

Now we are prepared to answer the question: how long does it take for the circuit to enter a steady-state condition?

$$\begin{aligned}
 \vec{\nabla} \cdot \vec{J} &= -\frac{d\rho}{dt} & \text{(Eq. 3.53)} \\
 \vec{J} &= \sigma \vec{E} & \text{(Eq. 3.45)} \\
 \rightarrow \vec{\nabla} \cdot \sigma \vec{E} &= -\frac{d\rho}{dt} \\
 \frac{\rho}{\varepsilon} \sigma &= -\frac{d\rho}{dt} \\
 \int \sigma dt &= -\varepsilon \int \frac{1}{\rho} d\rho \\
 \implies \rho &= \rho_0 e^{-\sigma \frac{t}{\varepsilon}}
 \end{aligned} \tag{3.58}$$

And if we define $\tau = \frac{\varepsilon}{\sigma}$, we now have an expression for the *relaxation time*, or the time to which it takes to reach steady-state. For a good conductor⁴⁵, the transient time is super brief; in the attosecond region. This means that ρ can be considered zero in the interior of a conductor. As for a good insulator the relaxation time tends to be in the order of hours or days.

When solving problems at steady state the boundary conditions to apply are:

1. $\vec{\nabla} \cdot \vec{J} = 0$
2. $\vec{\nabla} \times \frac{\vec{J}}{\sigma} = 0$

As we know that the normal component of a divergence-less vector field is continuous, from 1) above we get $J_{1n} = J_{2n}$. Likewise the tangential component of a curl-free vector field is continuous, therefore from 2) we arrive at $\frac{J_{1t}}{J_{2t}} = \frac{\sigma_1}{\sigma_2}$

⁴⁵ large $\sigma, \varepsilon \approx \varepsilon_0$

SECTION 4

Electromagnetism

SUBSECTION 4.1

Lecture 23: Static Magnetic Fields

Whereas electric fields are conservative, magnetic fields are non-conservative in the presence of currents or time-varying electric fields

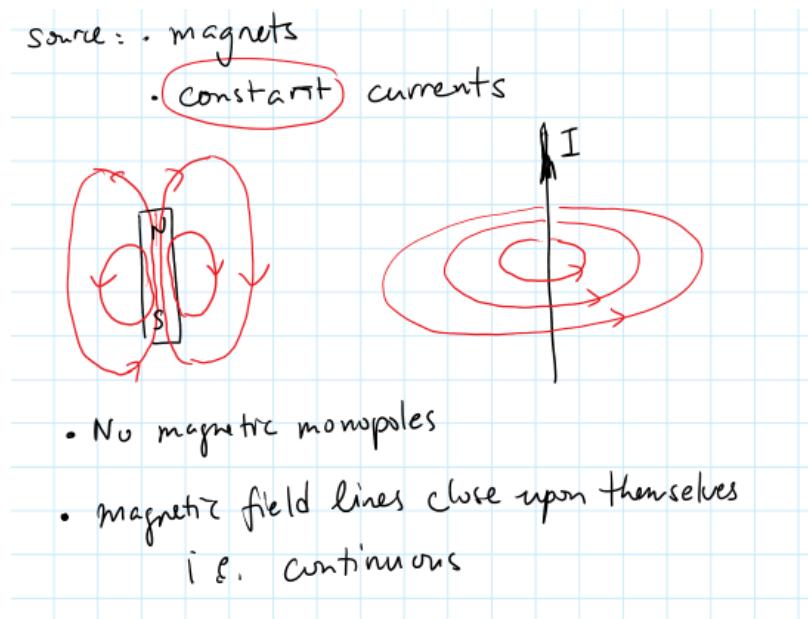


Figure 23. Magnetic fields. Note that there can be no magnetic monopoles and that magnetic field lines are continuous. Direction can be given by the right-hand-rule, i.e. for a current point thumb in direction of current and then the curl of fingers will give direction of magnetic field.

A test charge q in a magnetic field \vec{B} is given by

$$F_m = q\mu \times \vec{B} \quad (4.1)$$

And the total *electromagnetic* force in a charge q is then given by

\vec{B} , $[\frac{m}{C \frac{m}{s}} = T]$ is the magnetic flux density, and \vec{H} is the magnetic field intensity. $\vec{\mu}$ gives the velocity of the moving charge

Definition 21

Lorentz Force Equation

$$F = F_e + F_m = q(\vec{E} + \mu \times \vec{B}) \quad (4.2)$$

In order to study steady magnetic fields in free space we are only concerned with the magnetic flux density \vec{B} .

Theorem 16

Two fundamental postulates for magnetic fields in free space:

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (4.3)$$

$$\vec{\nabla} \times \vec{B} = \mu_o \vec{J} \quad (4.4)$$

Where $\mu_o = 4\pi \times 10^{-7} [\frac{H}{m}]$ is the permeability of free space.

Taking the volume integral of Eq. 4.3 and applying the divergence theorem we find that

$$\oint_s B \cdot ds = 0 \Leftrightarrow \vec{\nabla} \cdot \vec{B} = 0 \quad (4.5)$$

This further affirms that isolated magnetic charges cannot exist; there are no magnetic flow sources and magnetic flux lines always close upon themselves. Eq. 4.5 states that the total outward flux through any closed surface is 0.

Extending upon that we note that the divergence of the curl of a vector field is zero so therefore manipulating Eq. 4.4 can result in the familiar expression $\vec{\nabla} \cdot \vec{J} = 0$ which we found in the previous lecture

If we integrate over both sides of Eq. 4.4 and then apply Stoke's theorem we arrive at Ampere's circuital law

$$\int_s \nabla \times \vec{B} d\vec{s} = \int \mu_o \vec{J} d\vec{s} \implies \oint_c \vec{B} d\vec{l} = \underbrace{\mu_o I}_{\text{Current enclosed by path}} \quad (4.6)$$

Which states that the circulation of the magnetic flux in free space is equal to μ_o times to total current flowing through the surface. It is useful in determining \vec{B} caused by I given a closed path around the current with constant magnitude of \vec{B}

4.1.1 Magnetic Vector Potential

We assume \vec{B} to be divergence free⁴⁶, therefore \vec{B} can be expressed as the curl of another vector field \vec{A} , the *vector magnetic potential*

$$\vec{B} = \nabla \times \vec{A} \quad \left[\frac{Wb}{m} \right] \quad (4.7)$$

With this and some math-foo we get

Definition 22 Vector Poisson's Equation

$$\nabla^2 \vec{J} = -\mu_o \vec{J} \quad (4.8)$$

Comment

Recall: the vector Poisson equation is relatable to the Poisson equation (Eq. 3.30) from electrostatics. The Poisson equation had solutions of the form

$$V = \frac{1}{4\pi\epsilon_o} \int_{v'} \frac{P(\vec{R}')}{|\vec{R} - \vec{R}'|} dv' \quad (4.9)$$

Similar to the Poisson equation the vector Poisson equation has solutions of form

$$\vec{A}(\vec{R}) = \frac{\mu_o}{4\pi} \int_{v'} \frac{\vec{R}(\vec{R}')}{|\vec{R} - \vec{R}'|} \quad (4.10)$$

Noting that $\vec{J} \rightarrow \vec{A} \rightarrow \vec{B} = \vec{\nabla} \times \vec{A}$

SUBSECTION 4.2

Lecture 24: Biot-Savart Law

In most cases we are concerned with currents in wires.

For a thin wire with cross-sectional area S , $dv' = S dl'$, so along the wire we have

$$\vec{J} dv' = JS dl' = Idl' \quad (4.11)$$

Plugging this into the vector Poisson equation solution (Eq. 4.10) we find that, along the wire,

$$\vec{A}(\vec{R}) = \frac{\mu_o}{4\pi} \int_{L'} \frac{Idl'}{|\vec{R} - \vec{R}'|} \quad (4.12)$$

And if we take $B = \vec{\nabla} \times \vec{A}$ and do some math we arrive at the Biot-Savart law, which is comparable to Coulomb's law. Note the difference in that the bottom $|\vec{R} - \vec{R}'|$ term which is cubed for Biot-Savart and squared for Coulomb.

Definition 23

Biot-Savart Law

$$\vec{B}(\vec{R}) = \frac{\mu_0 I}{4\pi} \oint_{c'} \frac{d\vec{l} \times a_{\vec{R}-\vec{R}'} \cdot \hat{n}}{|\vec{R} - \vec{R}'|^3} \quad (4.13)$$

A procedure for solving problems with the Biot-Savart law is as follows

1. Choose a coordinate system
2. Write an expression for $d\vec{l}$
3. Write expressions for \vec{R}, \vec{R}'
4. Write out Biot-Savart Law (Eq. 4.13) and evaluate the integral

SUBSECTION 4.3

Lecture 25: Biot-Savart Examples

No notes for this lecture.

SUBSECTION 4.4

Lecture 26: Ampere's Law

Definition 24

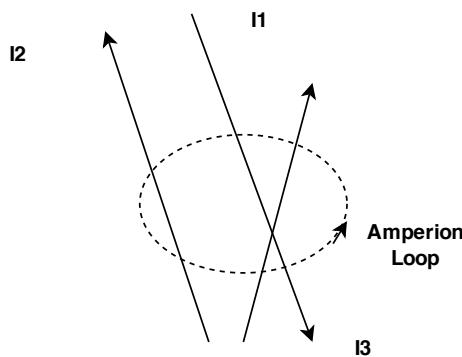
Ampere's Law

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} \Leftrightarrow \oint_c \vec{B} d\vec{l} = \mu_0 I_{\text{net current enclosed by loop}} \quad (4.14)$$

closed line integral around Amperion loop

Notes:

- Don't confuse \vec{B} field direction with the direction of the loop; it is independently specified
- Net current enclosed by loop depends on the loop direction



In the example above the integral is, using right hand rule for the direction,

$$\oint_c \vec{B} d\vec{l} = (I_1 + I_2 - I_3) \mu_0$$

SUBSECTION 4.5

Lecture 27: Ampere's Force Law, Magnetic Torque

Consider an uniform \vec{B} field. The force on a charge dq that will pass through a cross-section of the wire in dt time is given by

$$d\vec{F} = dq\vec{\mu} \times \vec{B} = dq \frac{d\vec{l}}{dt} \times \vec{B} = I d\vec{l} \times \vec{B} \quad (4.15)$$

Integrating we arrive at the following:

$$\vec{F}_m = \underbrace{q\vec{\mu} \times \vec{B}}_{\text{moving charge in } \vec{B} \text{ field}} \Leftrightarrow d\vec{F}_m = \underbrace{Id\vec{l} \times \vec{B}}_{\text{current in a wire in } \vec{B} \text{ field}} \Leftrightarrow \vec{F}_m = \underbrace{I\vec{L} \times \vec{B}}_{\text{uniform } \vec{B} \text{ along } L \text{ and constant direction (along } L \text{) } d\vec{L} \times \vec{B}} \quad (4.16)$$

Definition 25

Ampere's Force Law

Consider two loops C_1, C_2 which carry currents I_1, I_2 . To find the total force \vec{F}_{12} exerted on C_2 due to the \vec{B}_1 field caused by the current in C_1 we can just do an integral.

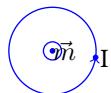
$$\vec{F}_{12} = \oint_{C_2} d\vec{F}_{12} = \oint_{C_2} I_2 d\vec{l}_2 \times \vec{B}_1(\vec{R}_2) \quad (4.17)$$

And if we were to repeat this with the other side we would find that $\vec{F}_{12} = -\vec{F}_{21}$

Extending on this, we may define magnetic torque to be the torque exerted by a magnetic field which tries to align the normal vector of a loop of current with the magnetic field.

$$\vec{T} = \vec{m} \times \vec{B} \quad (4.18)$$

4.5.1 Magnetic Dipole



Given a small loop current, we have a magnetic dipole. In this case the magnetic dipole is pointing out of the page⁴⁷

$$\vec{m} = I\pi b^2_{\text{area}} \vec{a}_z \quad (4.19)$$

If we put this in an external magnetic field, for example for a \vec{B} out of the page (same direction as the magnetic field), the forces experienced by every dl of the wire carrying the current will face outwards, i.e. having no torque. If the \vec{B} field is from left to right, the left and

⁴⁷ Right hand rule, curl fingers in current direction

right sides of the wire will experience forces, namely one going outside of the page on the left side and into the page on the right. This would cause a torque until the loop is aligned with the magnetic field.

For example,

$$\begin{aligned}
 d\vec{l} &= bd\phi\vec{a}_\phi \\
 d\vec{F}_m &= Id\vec{l} \times \vec{B} \\
 &= Ibd\phi(-\sin\theta\vec{a}_x + \cos\phi\vec{a}_y) \times B\vec{a}_x \\
 &= IBb\cos\phi d\phi(-\vec{a}_z) \\
 \Rightarrow dT &= \vec{r} \times d\vec{F}_m \\
 &= b\vec{a}_r \times d\vec{F}_m \\
 &= IBb^2[\cos^2\phi d\phi\vec{a}_y - \sin\phi\cos\phi d\phi\vec{a}_x]
 \end{aligned} \tag{4.20}$$

And just do an integral to find the torque

$$\vec{T} = \int d\vec{T} = IBb^2 \int_0^{2\pi} [\cos^2\phi d\phi\vec{a}_y - \sin\phi\cos\phi d\phi\vec{a}_x] = IBb^2\pi\vec{a}_y \tag{4.21}$$

Noting that πb^2 is the area, which multiplied by I is the magnitude of the magnetic moment. So it is also

$$= B \cdot |\vec{m}| \vec{a}_y \tag{4.22}$$

And from this one can sort of see that this is an alternative proof of what we found earlier (Eq. 4.18) [this isn't really a formal proof or anything] that magnetic torque is given by

$$\vec{T} = \vec{m} \times \vec{B} \tag{4.23}$$

The magnetic potential and magnetic field are related to the dipole by

$$\vec{A}(R) = \frac{\mu_0 \vec{m} \times \vec{a}_r}{4\pi R^2} \tag{4.24}$$

$$\vec{B}(R) = \frac{\mu_0 |\vec{m}|}{4\pi R^3} (2\cos\theta\vec{a}_R + \sin\theta\vec{a}_\theta) \tag{4.25}$$

These are comparable to voltage and electric fields in electrostatics

SUBSECTION 4.6

Lecture 28: Magnetization

Intuition:

- Atoms can be modelled by magnetic dipoles, since we can⁴⁸ think of them as little current (electron) loops around a nuclei.
- These atoms are randomly oriented in the absence of a magnetic field
 - No net magnetic dipole moment
- In the presence of a magnetic field, each atom will experience a magnetic torque
- The magnetic torque will cause the dipoles to orient uniformly such that they are all aligned with the \vec{B} field.
- These now-aligned dipoles will contribute to the magnetic field

⁴⁸ crudely

Definition 26

Magnetization Vector \vec{M}

$$\vec{M} = \lim_{\Delta\nu \rightarrow 0} \frac{\sum_{k=1}^{n\Delta V} \vec{m}_k}{\Delta\nu} \quad (4.26)$$

- \vec{M} is the volume density of the magnetic dipole moment (point function)
- The top of the fraction is the sum of all microscopic dipoles in a small volume
- $n\Delta V$ is the total number of atoms in the volume, n being the # of atoms per unit volume

However this isn't super practical in well, practice.

Consider a volume v' . In this volume we have a small volume dv' , which has a magnetization vector $\vec{M}(\vec{R}')$. We are concerned with the contribution to \vec{A} and \vec{B} from $d\vec{m}$.

The idea is that $d\vec{m} = \vec{M}dv'$ produces magnetic vector potential which we can find using Eq. 4.24

$$d\vec{A} = \frac{\mu_0 \vec{M} \times \vec{a}_R}{4\pi |\vec{R} - \vec{R}'|^2} dv' \quad (4.27)$$

As it turns out, $\frac{\vec{a}_R}{4\pi |\vec{R} - \vec{R}'|^2} = \vec{\nabla}'(\vec{R} - \vec{R}')^{-1}$

So we can make this more complicated expression that has an interesting physical interpretation

$$d\vec{A} = \frac{\mu_0}{4\pi} \left[\left(\frac{\vec{\nabla}' \times \vec{\mu}}{|\vec{R} - \vec{R}'|} \right) dv' + \left(-\vec{\nabla}' \times \frac{\vec{M}}{|\vec{R} - \vec{R}'|} \right) dv' \right] \quad (4.28)$$

And this is just two volume integrals with which we can apply Stoke's theorem

$$\begin{aligned} A &= \int d\vec{A} \\ &= \frac{\mu_0}{4\pi} \left[\int_{v'} \left(\frac{\vec{\nabla}' \times \vec{\mu}}{|\vec{R} - \vec{R}'|} \right) + \oint_{s'} \left(\frac{\vec{M} \times \vec{a}'_n}{|\vec{R} - \vec{R}'|} \right) ds' \right] \end{aligned} \quad (4.29)$$

The first term of which can be interpreted as the vector potential produced by a volume current density

$$\vec{J}_m = \nabla \times \vec{M} \quad \left[\frac{A}{m^2} \right] \quad (4.30)$$

The second term can be seen as the vector potential produced by a surface current density

$$\vec{J}_{ms} = \vec{M} \times \vec{a}'_n \quad \left[\frac{A}{m} \right] \quad (4.31)$$

These are called 'equivalent magnetization current densities' since there is no real current flowing; they arise as result of the material being magnetized. So we could calculate and construct a equivalent magnetization current to inspect magnetization behaviour.

SUBSECTION 4.7

Lecture 29: \vec{H} field and relative permeability

Comment

In the past when we worked with electric fields inside materials we worked with a new field \vec{D} so that our math would not be encumbered [as much] by the induced fields. The \vec{H} field does the same but for magnetic fields.

Recall, in free space we have Gauss's law

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (4.32)$$

And Ampere's law

$$\vec{\nabla} \times \vec{B} = \mu_o \vec{J} \quad (4.33)$$

In a material that has been magnetized, we have the equivalent current density⁴⁹

$$\vec{J}_m = \vec{\nabla} \times \vec{M} \quad (4.34)$$

Therefore

$$\vec{\nabla} \times \vec{B} = \mu_o (\vec{J} + \vec{J}_m) = \mu_o \vec{J} + \mu_o \vec{\nabla} \times \vec{M} \quad (4.35)$$

And doing some manipulation we get

$$\vec{J} = \vec{\nabla} \times \left(\frac{\vec{B}}{\mu_o} - \vec{M} \right) \quad (4.36)$$

So we now have a quantity that only depends on the real current density, which is convenient.

⁴⁹ Today I realized that I should be using `\curl` instead of `\vec \nabla` so I will be remembering to use them now

Definition 27

Magnetic Field Intensity \vec{H}

$$\vec{H} = \frac{\vec{B}}{\mu_o} - \vec{M} \quad \left[\frac{A}{m} \right] \quad (4.37)$$

Which gives us generalized Ampere's law for magnetostatics

$$\vec{\nabla} \times \vec{H} = \vec{J} \quad (4.38)$$

And in integral form,

$$\int_S (\vec{\nabla} \times \vec{H}) dS = \int_S \vec{J} dS = I \xrightarrow{\text{stokes theorem}} \oint_C \vec{H} \cdot d\vec{l} = I \quad (4.39)$$

It is also worthwhile noting that Eq. 4.37 can be rewritten as

$$\vec{B} = \mu_o \vec{H} + \mu_o \vec{M} \quad (4.40)$$

Also, in a linear isotropic material⁵⁰

$$\vec{M} = \chi_m \vec{H} \quad (4.41)$$

Where χ_m is unitless magnetic susceptibility⁵¹.
Therefore

$$\vec{B} = \mu_o \vec{H} + \mu \chi_m \vec{H} = \mu_o (1 + \chi_m) \vec{H} \quad (4.42)$$

If we let relative permeability $\mu_r = 1 + \chi_m$, then

⁵⁰ which is all we're dealing with in this course anyways

⁵¹ Like the electric susceptibility we covered in electrostatics

This looks basically identical to how \vec{D} fields work in electrostatics

$$\vec{B} = \mu_0 \mu_r \vec{H} = \mu \vec{H} \quad (4.43)$$

Where μ is the absolute permeability

A general sequence for how we can solve a lot of problems related to these concepts is as follows

1. Knowing I , find \vec{H}
2. Using the material properties, find \vec{B}
3. Using \vec{H} , \vec{M} , can find \vec{M}
4. Take the curl of \vec{M} to find the equivalent volume current density \vec{J}_m
5. Cross-product \vec{J}_m with surface normal to find the surface equivalent current density \vec{J}_{ms}

4.7.1 Magnetic Materials

- Most materials are not magnetizable, i.e. $\mu_r \approx 1$.
- Paramagnetic materials, i.e. $\mu_r \gtrsim 1$.
- Diamagnetic $\mu_r \lesssim 1$, interesting since would actually flip the other way than a regular magnetic material. But effect is usually minimal
- Ferromagnetic $\mu_r \gg 1$ e.g. iron, cobalt, nickel, etc. $\chi_m \approx 10^3$
 - Highly temperature dependent; there is a curie temperature, above which a magnet will become paramagnetic. So heating a material can cause demagnetization.

$\mu_r = 1 + \chi_m$, for paramagnetic $\chi_m \approx 10^{-5}$

Ferromagnetic materials are unique in that they have been experimentally confirmed to be composed of many small⁵² domains that are fully magnetized together but relative to [other domains], scattered. These domains are separated by domain walls of about 100 atoms. When an external field is applied to the material the walls move as to increase the size of the domains, increasing magnetic flux density. This is reversible for weak fields, but for larger fields it is irreversible and the domains will rotate in the direction of the \vec{B} field, resulting 'hysteresis' in magnetization lagging behind the field producing it.

⁵² micron to mm scale

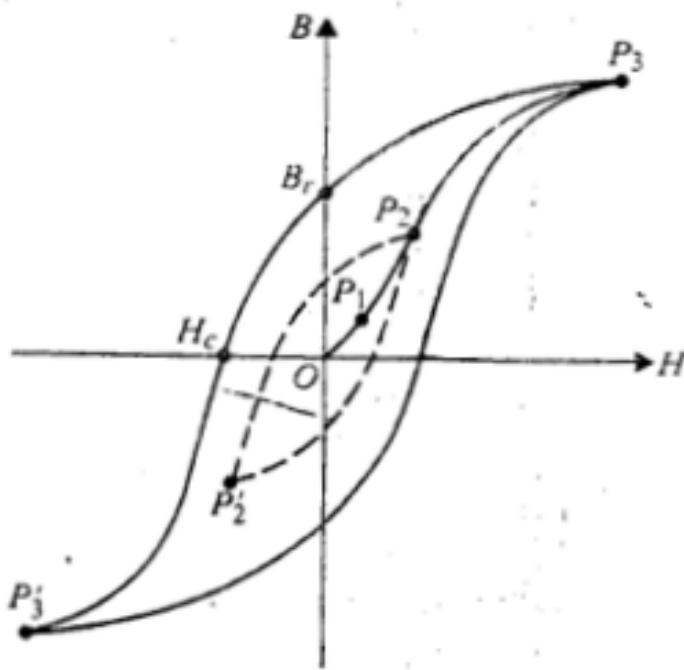


Figure 24. Hysteresis loop in $B - H$ plane for ferromagnetic material

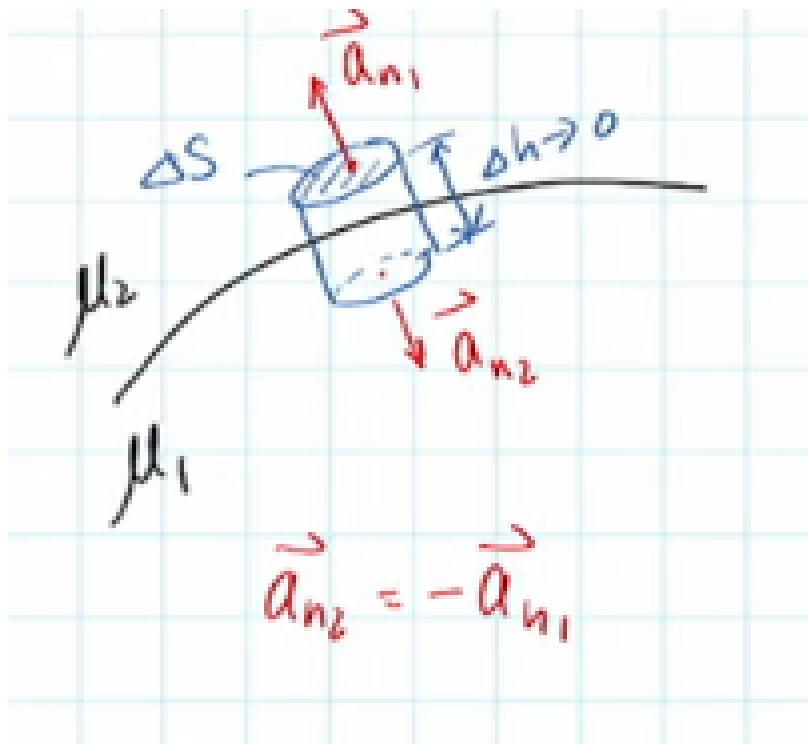
SUBSECTION 4.8

Lecture 30: Boundary conditions for magnetic fields

Given an interface between two materials with different permeabilities,

The **normal component** can be inspected by taking Gauss's law (Eq. 4.5)

$$\text{When } \Delta h \rightarrow 0 \quad \vec{B}_2 \cdot \vec{a}_{n_1} \Delta S + \vec{B}_1 \cdot \vec{a}_{n_2} \Delta S = 0 \quad (4.44)$$



Which implies that the normal component of the \vec{B} field is continuous.

$$B_{1n} = B_{2n} \quad (4.45)$$

And in a linear isotropic material

$$\mu_1 H_{1n} = \mu_2 H_{2n} \quad (4.46)$$

As for the **tangential component**, we can use Ampere's law (Eq. 4.14)

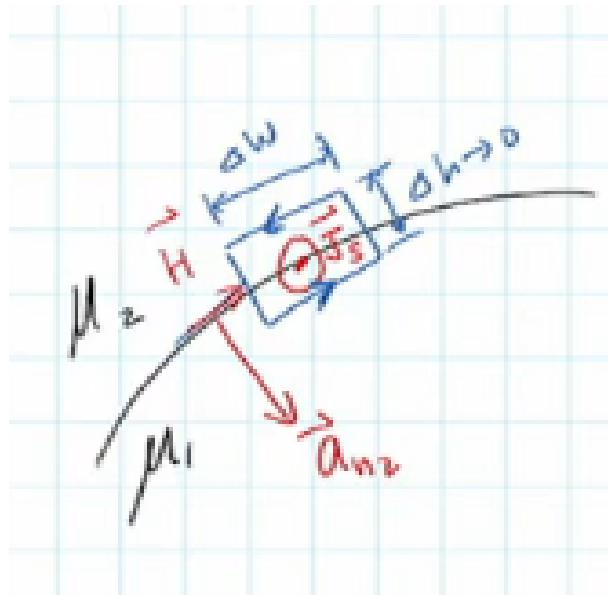


Figure 25.

$$\oint_c \vec{H} \cdot d\vec{l} = I \Leftrightarrow \vec{\nabla} \times \vec{H} = \vec{J} \quad (4.47)$$

So when $\Delta h \rightarrow 0$, $H_{1t} \Delta W - H_{2t} \Delta W = \Delta W J_s$
So in scalar and vector form,

$$H_{1t} - H_{2t} = J_s \Leftrightarrow \vec{a}_{n2} \times (\vec{H}_1 - \vec{H}_2) = \vec{J}_s \quad (4.48)$$

Note: if σ_1, σ_2 are finite (i.e. real physical material), $J_s = 0$ so $H_{1t} = H_{2t}$. This would only not be in the case in perfect conductors that are magnetic or superconductors.

When h becomes small there can only be a surface current

Table 1. Summary of boundary conditions

	Electrostatics	Magnetostatics
Continuous at interface	$E_{1t} = E_{2t}$ $(\vec{\nabla} \times \vec{E} = 0)$	$B_{1n} = B_{2n}$ $\vec{\nabla} \cdot \vec{B} = 0$ $H_{1t} = H_{2t}$ (for finite σ) $\mu_1 H_{1n} = \mu_2 H_{2n}$
Non-continuous at interface	$\vec{a}_{n2} \cdot (\vec{D}_1 - \vec{D}_2) = \rho_s$ $\vec{\nabla} \times \vec{D} = \rho$	$\vec{a}_{n2} \cdot (\vec{H}_1 - \vec{H}_2) = \vec{J}_s$ $\vec{\nabla} \times \vec{H} = \vec{J}$
Current densities		$J_{1n} = J_{2n}$ $\vec{\nabla} \cdot \vec{J} = 0$ (under steady state) $\frac{J_{1t}}{J_{2t}} = \frac{\sigma_1}{\sigma_2}$ $\vec{\nabla} \times \left(\frac{\vec{J}}{\sigma}\right) = 0$

Review 1

4.8.1 Magnetic Circuits

We may think of magnetic systems as analogous to electric circuits.

Given a magnetic core, which is some material which has a much larger than that of the

surrounding space $\mu_c \gg \mu_0$ which will confine \vec{B} very well, making it behave like a current so to speak⁵³.

⁵³ Current is analogous to the \vec{B} field, the integral of which gives the magnetic flux Φ_m

Table 2. Electric vs Magnetic Circuits

Electric	Magnetic
Voltage V (electromotive force)	$V_m = N \cdot I$ (magnetomotive force)
Current I	Φ_m magnetic flux
Resistance R	\mathcal{R} reluctance (resistance to letting a \vec{B} field through)

An example of where this magnetic circuit model can be useful is for doing magnetic tape data storage. Or for modelling transformers which scale a voltage up and down through a \vec{B} field and two coils. Adding a third coil makes it a three-phase transformer. We will discuss transformers more later.

SUBSECTION 4.9

Lecture 31: Inductance

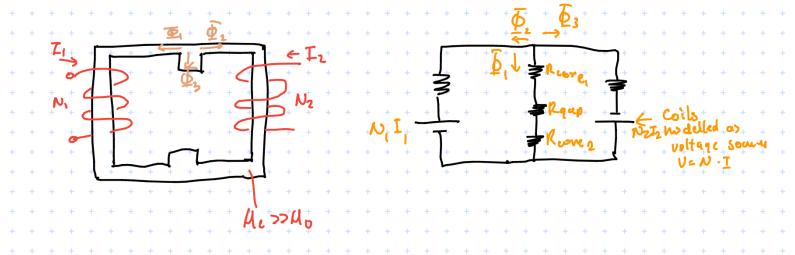


Figure 26. Modelling magnetic circuit with electric circuit

Recall Table 2 that we map current to flux⁵⁴. Reluctances can be built up as a function of the material and the surroundings. Coils are modelled as voltage sources of magnitude NI with polarity $+\rightarrow-$ in direction opposing the \vec{B} field⁵⁵

⁵⁴ sum of fluxes at node must be 0 (like current)

⁵⁵ Right hand rule

4.9.1 Inductance

Consider two current-carrying loops C_1 and C_2 creating two surfaces S_1, S_2 .

The magnetic flux through S_2 caused by I_1 is given by

$$\Phi_{12} = \int_{S_2} \vec{B}_1 d\vec{S}_2 \quad (4.49)$$

We note that Φ_{12} is directly proportional to I_1 , so we can define inductance L to be

Definition 28 Mutual Inductance

$$\Phi_{12} = L_{12} \cdot I_1 \quad (4.50)$$

Units of Henry $H = \frac{Wb}{A}$.

W denotes Weber

Mutual inductance gives a sense of how coupled the two circuits are. If we make these fluxes change over time a current would be induced in the 2nd current.

Looking at the flux through the first circuit Φ_{11} caused by itself we see that it is again proportional to the current I_1 .

We now define the self inductance as

$$\Phi_{11} = \int_{S_1} \vec{B}_1 \cdot d\vec{S}_1 \quad (4.51)$$

Definition 29

Self Inductance

$$\Phi_{11} = L_{11} \cdot I_1 \quad (4.52)$$

Where L_{11} is the self inductance of the current-carrying loop.

Looking at L_{12} and L_{11} we can note that L_{11} will always be larger than or equal to that of L_{12} just because all of the flux lines will flow through the first loop but only a few will pass through the second.

Note that inductance (both self and mutual) is given by

$$L_{ab} = \frac{\Phi_{ab}}{I_a} \quad (4.53)$$

SUBSECTION 4.10

Lecture 32: Magnetic Energy

Comment

Recall: there is energy stored in a system of electric charges. Likewise, there is energy stored in a system of magnetic fields.

Definition 30

Stored Magnetic Energy

$$W_m = \frac{1}{2} \int_{\mathbb{R}^3} \vec{B} \cdot \vec{H} dv \quad (4.54)$$

Note: this looks like Eq. 3.25

It follows⁵⁶ that the energy stored in an inductor is

$$W_m = \frac{1}{2} L I^2 \quad (4.55)$$

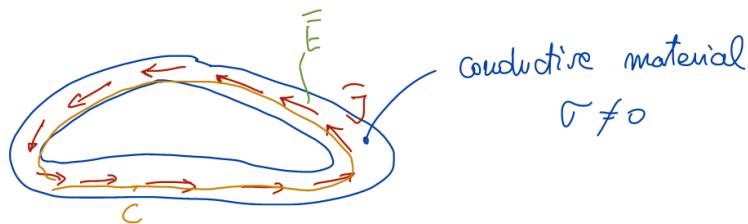
⁵⁶according to the prof, idk the derivation

This gives us another way to find inductance which can be more convenient in some circumstances e.g. inductance per unit length of a coaxial cable

1. Assume current I
2. Find \vec{B}
3. Find $W_m = \frac{1}{2} \int \vec{B} \cdot \vec{H} dv$
4. $L = 2W_m/I^2$

SUBSECTION 4.11

Lecture 33: Electromagnetic Induction



If we had a current in a circuit driven by an electric field, \vec{E} would have to be pointing in the same direction as \vec{J} . As such, an electric field \vec{E} must be non-conservative for it to drive a current through a circuit, i.e. $\oint_C \vec{E} \cdot d\vec{l} \neq 0$. Another way to think of this is that work must be done i.e. energy must be provided for the current to flow; there is going to be some resistance in the circuit.

Recall: any \vec{E} made by a system of static charges is conservative. Therefore in order to build a non-conservative \vec{E} we need voltage sources e.g. chemical batteries, photovoltaic cells, and induction.

4.11.1 Chemical Batteries

Consider a battery.

$$\vec{E} = \vec{E}_{\text{conservative}} + \vec{E}_{\text{non-conservative}} \quad (4.56)$$

The non-conservative part due to chemical forces \vec{F}_{chem} is sketched in purple. The electric field of this would be pointing upwards.

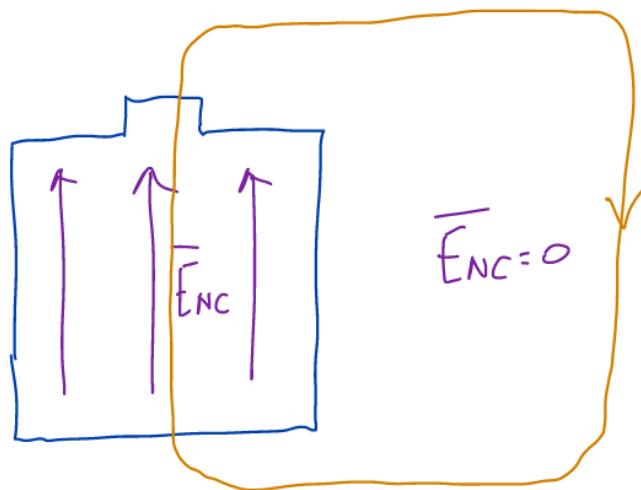
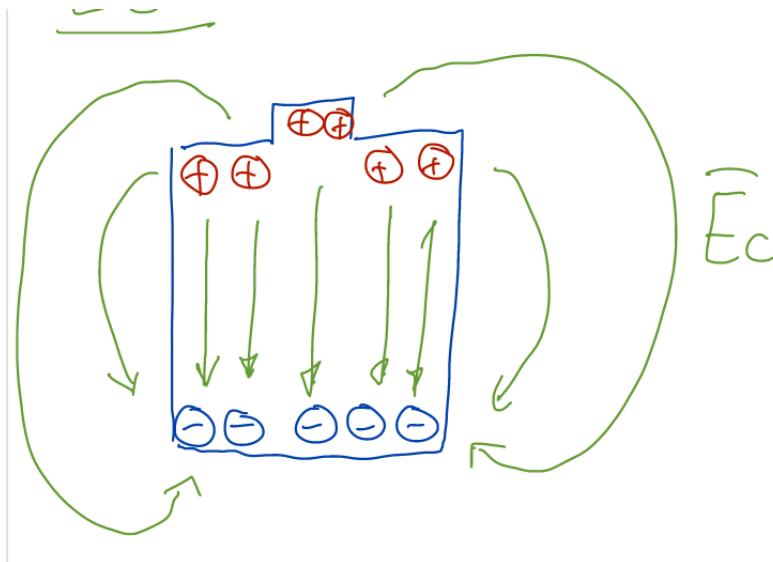
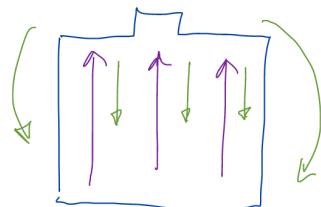


Figure 27. It is easy to see why this is non-conservative; making a closed path through and outside the battery (where there would be no chemical reactions causing a \vec{F}_{chem}) and taking the integral of the field would *definitely* be nonzero

On the other hand, a conservative field is created by the electric field of the separated charges.



$$\vec{E} = \vec{E}_{NC} + \vec{E}_c$$



\vec{E}_c grows until it becomes $\vec{E}_c = -\vec{E}_{NC}$

In the battery \vec{E}_c will grow until it becomes equal and opposite to the non-conservative field. Therefore at equilibrium the total field inside the battery is $\vec{E} = 0$ and the outside field is $\vec{E} = \vec{E}_c$.

The nub side of a battery is positive and the flat side is negative

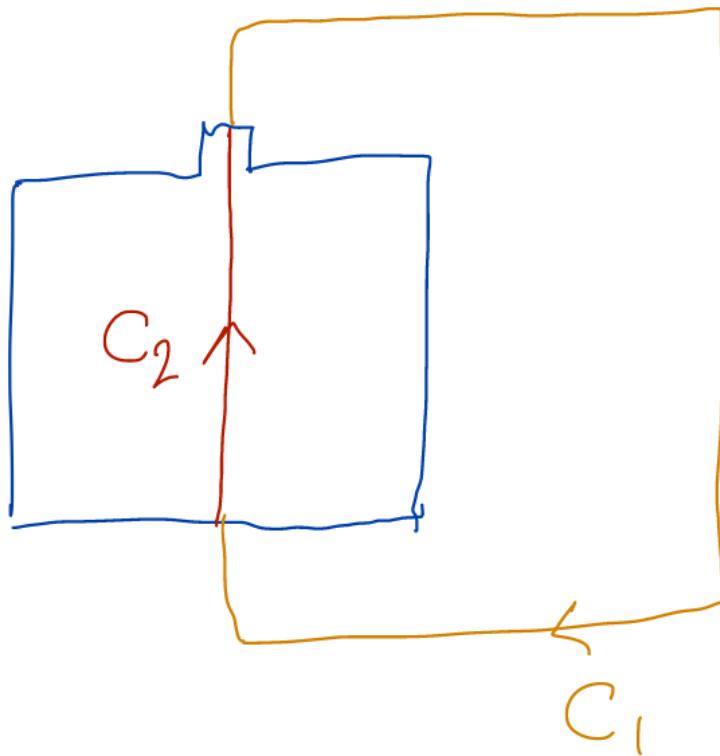
Definition 31 Electromotive Force

$$\mathcal{V}_{emf} \equiv \oint_c \frac{\vec{F} \cdot d\vec{l}}{q} \quad \left[\frac{J}{C} = W \right] \quad (4.57)$$

Think of it as a test charge along a path c . There will be a force done on the test charge, which is given by the top half of the fraction. Then we normalize it with the q used in the test.

Alternatively we may apply Lorentz's law (Eq. 4.2) and arrive at the following:

$$\mathcal{V}_{emf} = \oint_c (\vec{\mu} \times \vec{B}) \cdot d\vec{l} - \int_S \frac{\partial B}{\partial t} dS \quad (4.58)$$



Or we can just integrate over the non-conservative part only (c_2) Since integrating over the non-conservative part in the battery example above is the same as integrating over the outside portion of the conservative path c_1 .

$$\int_{c_1+c_2} \vec{E} \cdot d\vec{l} = \int_{c_2} \vec{E}_{nc} \cdot d\vec{l} = \int_{c_1} \vec{E}_c \cdot d\vec{l} \quad (4.59)$$

$$\mathcal{V}_{emf} = \oint_c \vec{E}_{nc} \cdot d\vec{l} \quad (4.60)$$

Which is just the voltage across the terminals! This is different than voltages, though; voltages are for conservative fields and are what appears to be the voltage from the outside of the battery; \mathcal{V}_{emf} quantifies the strength of the non-conservative field inside the battery. The distinction is made in order to understand non-conservative fields better.

4.11.2 Faraday's Law

Definition 32

Faraday's Law

$$\underbrace{\oint_c \vec{E} \cdot d\vec{l}}_{V_{emf}} = -\frac{d}{dt} \underbrace{\int_S \vec{B} \cdot d\vec{S}}_{\text{magnetic flux through } S} \quad (4.61)$$

Any time that there is a time-varying \vec{B} field flowing through a surface, a V_{emf} is induced. So we can induce emf using a time-varying \vec{B} or using motion

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (4.62)$$

Recall: time-varying \vec{E} field also induces current (Refer to Maxwell's equations)

SUBSECTION 4.12

Lecture 33: Induction & Faraday's Law Examples

Comment

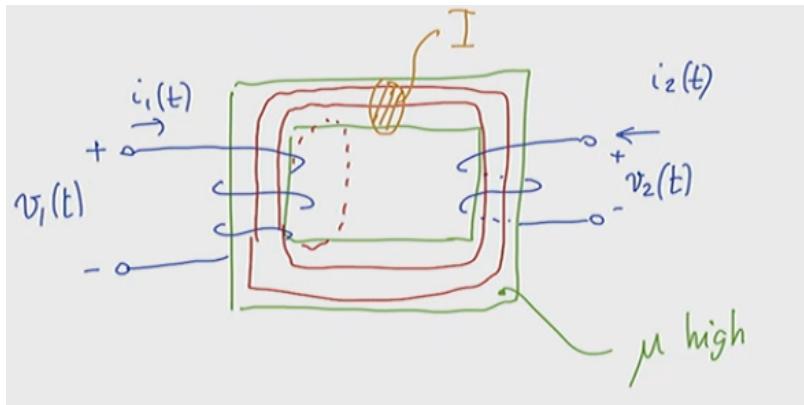
Examples from this lecture are not typeset in the interest of my own sanity.

- A induced V_{emf} may be represented in a circuit as a voltage source with $V = V_{emf}$. The polarity of the circuit is given by the right hand rule
- It's also worthwhile noting that in these time-varying conditions it is now possible for a wire to have a voltage; when we worked with statics it was assumed that no voltage can develop over a good conductor.
- Faraday's law also explains how inductors⁵⁷ work. As inductors are basically just a loop of wire, putting a time-varying current through it produces a \vec{B} . This B field produces a flux $\Phi_m(t)$ which, since it is time varying, induces a $V_{emf} = -\frac{d}{dt}\Phi_m(t) - L$

SUBSECTION 4.13

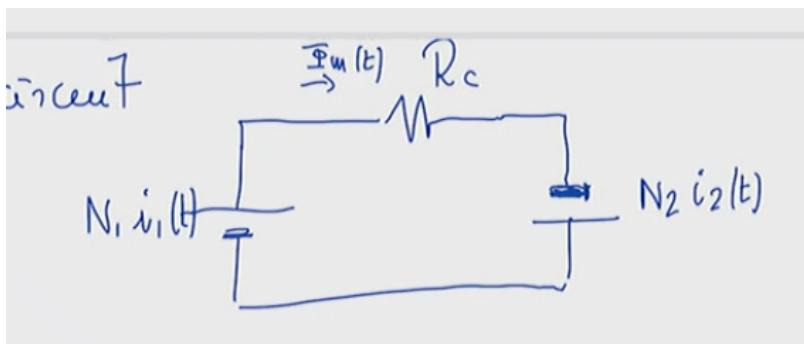
Lecture 34: Transformers & Lenz's Law

A transformer is a device that couples circuits through magnetic flux. This magnetic flux is typically confined to a 'magnetic circuit' by means of a material with $\mu \gg \mu_0$.



In order to solve transformer problems we must first find the magnetic flux that flows through the circuit.

It is convenient to represent the magnetic circuit as a corresponding electric one.



- A coil is represented as a voltage source with $V = N_i I_i(t)$

Point thumb in direction of B field. The voltage will oppose the direction indicated by the curl of the fingers, i.e. the direction of the palm will be the negative side of the voltage source and fingers the positive side.

- A reluctance is represented by a resistance with $R_c = \frac{l_c}{\mu_c \cdot S}$
- And flux is analogous to current.

Using this we can find $\Phi_m(t)$ at each coil using Faraday's law. Note that since we have N_i turns, the total flux is $N_i \Phi_m(t)$ And it follows by Faraday's law that

$$V_{emf} = -\frac{d}{dt} N_1 \Phi_m(t) \quad (4.63)$$

From this we can build $v(t)$ from V_{emf} , which is (note RHR). And then apply the same thing to both sides of the transformer, after which this becomes just a circuit problem. A neat result is that it turns out that $\frac{v_2(t)}{v_1(t)} = \frac{N_2}{N_1} \Rightarrow v_2(t) = \frac{N_2}{N_1} v_1(t)$; so this makes for a really easy way to scale voltages up and down. Transformers are important because it is easier transmit high power through large voltages than large currents⁵⁸. This explains why wall power is in AC; you can't have transformers with DC current.

⁵⁸ $P = IV = I^2 R$

Representing this in terms of the other quantities, we may find that

$$v_1(t) = \frac{N_1^2}{R_c} \frac{di_1(t)}{dt} + \frac{N_1 N_2}{R_c} \frac{di_2(t)}{dt} \quad (4.64)$$

And for the other side,

$$v_2(t) = \frac{N_2^2}{R_c} \frac{di_2(t)}{dt} + \frac{N_1 N_2}{R_c} \frac{di_1(t)}{dt} \quad (4.65)$$

This is a complete relation between the currents and voltages in the transformer. This is similar to how inductors work; in fact we can write it as

$$v_1(t) = \underbrace{\frac{N_1^2}{R_c} \frac{di_1(t)}{dt}}_{L_{11}} + \underbrace{\frac{N_1 N_2}{R_c} \frac{di_2(t)}{dt}}_{L_{12}} \quad (4.66)$$

$$v_2(t) = \underbrace{\frac{N_1 N_2}{R_c} \frac{di_1(t)}{dt}}_{L_{21}} + \underbrace{\frac{N_2^2}{R_c} \frac{di_2(t)}{dt}}_{L_{22}} \quad (4.67)$$

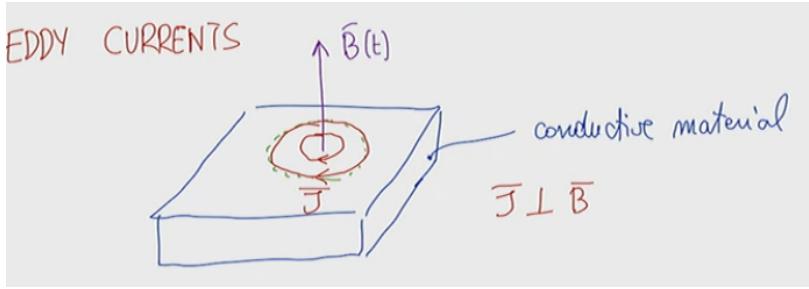
Which looks like a sort of two-dimensional inductor or some sort of evil inductor matrix – which it is. We can think of it as two inductors each with their own self inductance L_{11}, L_{22} , which is coupled by a mutual inductance $M = L_{12} = L_{21}$

Definition 33

Lenz's Law The direction of an electric current induced in a conductor will be such that the magnetic field created by the induced current will oppose changes in the initial magnetic field.

An example of where this can be useful is in the case of eddy currents. Suppose we have a piece of conductive material with some external $\vec{B}(t)$. If we were to apply Faraday's law to this material we note that a current will be induced in the material.

This won't tell us anything more than Faraday's, but will give us more insight about the direction of V_{emf} . Note that the direction of the current is dependent on whether the external $\vec{B}(t)$ is increasing or decreasing



This is the basis for induction cooktops (the induced eddy currents (think about the rings...) produces heat in the material) and eddy-current braking. This is why in actual transformers the conductor is split up into thin laminated layers in order to improve their efficiency since less energy will be wasted into heat.

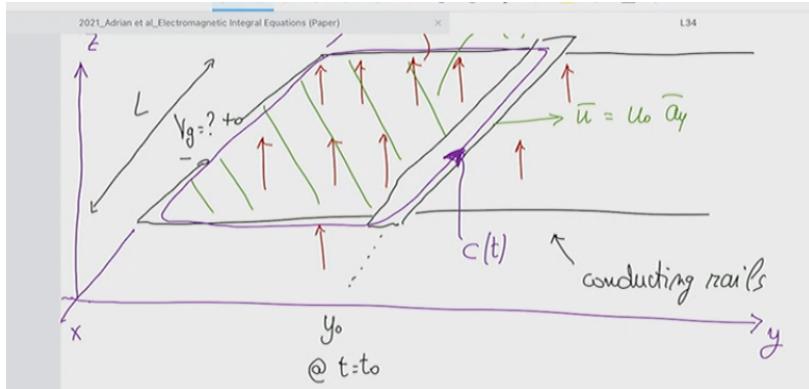
SUBSECTION 4.14

Lecture 35: Motional EMF

Comment

Motional EMF was actually split between two lectures. I have grouped them under one heading and will be omitting most examples.

Whereas transformer EMF depended on $\vec{B}(t)$ being directly time-modulated, motional EMF is the result of the area over which the closed loop encloses changing over time.



In this example we have a moving bar across two charged rails.
The area $S(t)$ is therefore

$$S(t) = L(y_0 + \mu_o \cdot t) \quad (4.68)$$

The is then

$$\Phi_m(t) = \int_{S(T)} \vec{B} \cdot d\vec{S} = B_o S(t) = B_o L(y_0 + \mu_o \cdot t) \quad (4.69)$$

So Faraday's law can be applied to find V_{emf}

$$V_{emf} = -\frac{d}{dt} \Phi_m = -B_o L \mu_o \quad (4.70)$$

And solving the problem we insert a equivalent voltage source V_{emf} .

$$V_g = +V_{emf} = -B_o L \mu_o \quad (4.71)$$

And we find that this is linearly moving generator produces a V_g that is constant over time.

Note that if we look at I by putting down a test resistance the current will be such that⁵⁹ the induced magnetic field will produce a force⁶⁰ that opposes $\vec{\mu}$. Therefore work must be done on the bar to produce a current – which is turned into heat in the resistor or routed to do useful work.

SUBSECTION 4.15

Lecture 36: Generalized Faraday's Law

Consider a moving circuit in a time-varying \vec{B} field. For example, if we had a charge q moving through a time-varying \vec{B} field,

- If observed by a stationary observer, $\vec{F} = q\vec{E} + q\vec{u} \times \vec{B}$
- If observed by an observer moving alongside the circuit, $\vec{F} = q\vec{E}'$
 - Where $\vec{E}' = \vec{E} + \vec{\mu} \times \vec{B}$

If we have a moving circuit (i.e. think of a loop carrying current moving with velocity $\vec{\mu}$), the EMF induced is given by

$$\oint_c \vec{E}' \cdot d\vec{l} = \oint_c (\vec{E} + \vec{\mu} \times \vec{B}) \cdot d\vec{l} = \oint_c \vec{E} \cdot d\vec{l} + \oint_c (\vec{\mu} \times \vec{B}) \cdot d\vec{l} \quad (4.72)$$

Which we can observe to be the sum of transformer and motional EMF.

$$V_{emf} = \oint_c \vec{E}' \cdot d\vec{l} = - \underbrace{\int_s \frac{\partial \vec{B}}{\partial t} \cdot d\vec{s}}_{\text{transformer}} + \underbrace{\oint_c (\vec{\mu} \times \vec{B}) \cdot d\vec{l}}_{\text{motional}} \quad (4.73)$$

V_{emf} is independent of frame of reference, but the transformer V_{tr} and motional V_m do change (though their sum is constant).

We can show that the negative ratio of the charge of flux through the closed loop is independent of the frame of reference

$$-\frac{d}{dt} \underbrace{\int_s \vec{B} \cdot d\vec{s}}_{\Phi(t)} = - \int_s \frac{\partial \vec{B}}{\partial t} \cdot d\vec{s} + \oint_c (\vec{\mu} \times \vec{B}) \cdot d\vec{l} = -\frac{d\Phi}{dt} \quad (4.74)$$

Which gives us Faraday's law

$$\oint_c \vec{E}' \cdot d\vec{l} = -\frac{d}{dt} \int_s \vec{B} \cdot d\vec{s} = -\frac{d\Phi}{dt} \quad (4.75)$$

⁵⁹ Recall: Lenz's law

⁶⁰ Recall: Lorentz's Law (Eq. 4.2)

Only portions of the circuit c that 'cut' the \vec{B} lines because of motion $\vec{\mu}$ contribute to motional V_{emf}

SUBSECTION 4.16

Lecture 37: Maxwell's Equations

Comment

We now have collected all of Maxwell's Equations which describe all electromagnetic phenomena; we have yet to find anything that contradicts them. It works across all scales and is compatible with relativity. Quantum variants of the equations also exist to describe EM phenomena with regards to quantum mechanics. These quantum variants look the exact same as the classical Maxwell equations except E, H, D, B become operators instead of fields. These would then be solved with the Schrödinger equation⁶¹.

Combining these equations, namely Faraday's law and the Ampere-Maxwell equation we get a 2nd order differential equation with wave solutions – which describes EM wave behaviour. I guess the point of this is to impress upon me the significance of these equations we learnt.

⁶¹I.e. at quantum scales we want to know how to be dealing with an electron as a wave instead and how the current or charge probability distributions would look

Theorem 17

Maxwell's Equations Gauss's Law:

$$\nabla \cdot \vec{D} = \rho \Leftrightarrow \nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \Leftrightarrow \int_S \vec{E} \cdot d\vec{S} = \frac{Q}{\epsilon_0} \Leftrightarrow \iiint_V \nabla \cdot \vec{E} dV = \frac{Q}{\epsilon_0} \quad (4.76)$$

Gauss's Law for Magnetism

$$\nabla \cdot \vec{B} = 0 \Leftrightarrow \iint_S \vec{B} \cdot d\vec{S} = 0 \quad (4.77)$$

Faraday's Law of induction:

$$V_{emf} = \nabla \times \vec{E} = -\frac{d\vec{B}}{dt} \Leftrightarrow \oint_C \vec{E} \cdot d\vec{l} = -\frac{d}{dt} \int_S \vec{B} \cdot d\vec{S} \quad (4.78)$$

Ampere-Maxwell law:

$$\nabla \times \vec{H} = \vec{J} + \frac{d\vec{D}}{dt} \quad (4.79)$$

All of these can be combined with the following constitutive equations that account for material behaviour

$$\vec{D} = f(\vec{E}) \quad (4.80)$$

$$\vec{B} = g(\vec{H}) \quad (4.81)$$

And charge conservation via the continuity equation

$$\nabla \cdot \vec{J} + \frac{d\rho}{dt} = 0 \quad (4.82)$$

4.16.1 Ampere-Maxwell Law**Ampere-Maxwell Law**

We had previously seen generalized Ampere's law for magnetostatics (Eq. 4.38, 4.39), $\nabla \times \vec{H} = \vec{J}$. This works for statics and breaks for time-varying fields because it is violates charge conservation. This can be observed by taking the divergence of both sides

$$\nabla \cdot (\nabla \times \vec{H}) = \nabla \cdot (\nabla \times \vec{J}) \quad (4.83)$$

The left side becomes 0 since the divergence of curl is zero. But then $\nabla \cdot \vec{J} = 0$, which is

Recall: charge enclosed within closed surface

impossible because of the continuity equation which states that $\nabla \cdot \vec{J} + \frac{dp}{dt} = 0$ (Eq. 3.53). So this brings us to the final form of the Ampere-Maxwell law

$$\nabla \times \vec{H} = \vec{J} + \frac{d\vec{D}}{dt} \Leftrightarrow \oint_c \vec{H} \cdot d\vec{l} = \int_S \vec{J} \cdot d\vec{S} + \int_S \frac{d\vec{D}}{dt} \cdot d\vec{S} \quad (4.84)$$

Which is experimentally verified and validated. Intuitively we may think of this as there being two things that produce a magnetic field; a current and the time derivative of \vec{D} .

It is also useful to define a new term, 'displacement current density' to represent the time derivative of \vec{D}

$$\vec{J}_d = \frac{d\vec{D}}{dt} = \epsilon_0 \frac{d\vec{E}}{dt} + \frac{d\vec{P}}{dt} \quad (4.85)$$

So we write

$$\vec{J}_d = \frac{d\vec{D}}{dt} = \epsilon \frac{d\vec{E}}{dt} \quad (4.86)$$

Recall: $\vec{D} = \epsilon_0 \vec{E} + \vec{P}$ (Eq. 3.2). For most materials the relationship between \vec{D} and \vec{E} is linear i.e. $\vec{D} = \epsilon \vec{E}$

ECE286

Probability and Statistics

SECTION 5

Probability Distributions

SUBSECTION 5.1

Lecture 14: Functions of random variables

In the discrete case, given X with PMF $f(x)$, we can define an *invertible* function Y where $Y = u(X)$, therefore can write $X = u^{-1}(Y)$. If $g(y)$ is the distribution of Y

$$\begin{aligned} g(y) &= P(Y = y) \\ &= P(u^{-1}(Y) = u^{-1}(y)) \\ &= f(u^{-1}(y)) \end{aligned} \tag{5.1}$$

In the continuous case we may arrive at

$$g(y) = f(u^{-1}(y)) \left| \frac{du^{-1}(y)}{dy} \right| \tag{5.2}$$

Definition 35

The r^{th} moment about the origin of the random variable X is

$$\mu'_r = E[X^r] = \begin{cases} \sum_x x^r f(x) & X \text{ discrete} \\ \int_{-\infty}^{\infty} x^r f(x) dx & X \text{ continuous} \end{cases} \tag{5.3}$$

This result is derived through the Leibniz integral rule,

$$g(y) = \frac{d}{dy} \int_{-\infty}^{u^{-1}(y)} f(t) dt$$

- The mean is the first moment
- For variance, $\sigma^2 = E[X^2] - \mu^2 \rightarrow \sigma^2 = \mu'_2 - \mu^2$

Definition 36

The moment-generating function of X is defined as

$$\mu'_r = E[X^{tX}] = \begin{cases} \sum_x e^{tx} f(x) & X \text{ discrete} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & X \text{ continuous} \end{cases} \tag{5.4}$$

In general

$$\mu'_r = \left. \frac{d^r M_X(t)}{dt^r} \right|_{t=0} \tag{5.5}$$

SUBSECTION 5.2

Lecture 15: More on moment generating functions

By definition and completing the square,

$$\begin{aligned}
M_X(T) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{t2\pi\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{t2\pi\sigma} e^{-\frac{x^2-2(x-\mu)^2+\mu^2}{2\sigma^2}}} dx \\
&= e^{\frac{2\mu t+t^2\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(x-(\mu+t\sigma^2))^2}{2\sigma^2}\right\} dx \\
&= e^{\frac{\mu t+t^2\sigma^2}{2}} \quad (5.6)
\end{aligned}$$

The integrand is just a normal PDF and thus integrates to one

5.2.1 Linear combinations of random variables

Given random variable X with distribution $f(x)$. What is the distribution, $h(y)$, of $Y = aX$?

Definition 37

Probability distribution of linear combination of random variables

Discrete:

$$\begin{aligned}
P(X = x) = f(x) \rightarrow h(y) &= P(Y = y) \\
&= P(aX = y) \\
&= f\left(\frac{y}{a}\right)
\end{aligned} \quad (5.7)$$

Continuous:

$$\begin{aligned}
F(y) &= P(Y \leq y) \\
&= P(X \leq \frac{y}{a}) \\
&= \int_{-\infty}^{\frac{y}{a}} f(t) dt \\
\int_{-\infty}^{\frac{y}{a}} f(t) dt &= \int_{-\infty}^y \frac{1}{a} f\left(\frac{s}{a}\right) ds \\
\Rightarrow h(y) &= \frac{1}{|a|} f\left(\frac{y}{a}\right)
\end{aligned} \quad (5.8)$$

u-sub; let $s = at$

Extending from this, let's assume that X has the MGF⁶² $M_X(t)$. Then,

⁶² Moment Generating Function

$$\begin{aligned}
M_Y(t) &= \int_{-\infty}^{\infty} e^{ty} h() \\
&= \frac{1}{|a|} \int_{-\infty}^{\infty} e^{ty} f\left(\frac{y}{a}\right) dy \\
&= \frac{1}{|a|} \int_{-\infty}^{\infty} e^{taz} f(z) a dz \\
&= \int_{-\infty}^{\infty} e^{taz} f(z) dz \\
&= M_X(at)
\end{aligned} \quad (5.9)$$

u-sub $z = \frac{y}{a}$

More generally,

$$M_{aX}(t) = M_X(at) \quad (5.10)$$

When working with more than one random variables, i.e. independent RV X, Y with distributions $f(x), g(y)$, the distribution of $Z = X + Y, h(z)$ is given by

$$P(Z = z) = P(X + Y = z) = \sum_w P(X = w)(Y = z - w) \Rightarrow h(z) = \sum_{w=-\infty}^{\infty} f(w)g(z - w) \quad (5.11)$$

And for the continuous case we arrive at the convolution integral,

$$h(z) = \int_{-\infty}^{\infty} f(w)g(z - w)dw \quad (5.12)$$

SECTION 6

Sampling

SUBSECTION 6.1

Lecture 16: Sampling

Given sample data x_1, \dots, x_n , each x_i is the realization of an RV⁶³, $X_i, i = 1, \dots, n$

Definition 38

Given the mean,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (6.1)$$

the **sample mean** of a random variable X is defined as

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (6.2)$$

Definition 39

The **sample median** is

$$x_m = \begin{cases} \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} & \text{if } n \text{ even} \\ \frac{x_{\frac{n+1}{2}}}{2} & \text{if } n \text{ is odd} \end{cases} \quad (6.3)$$

And **mode** is familiar as the most frequently occurring value

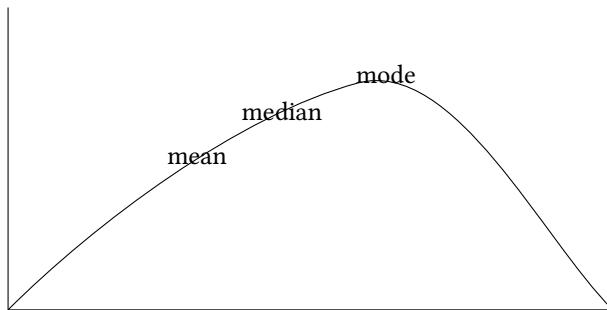


Figure 28. Median, mode, mean labelled on a skew distribution

⁶³ Say each X_i is a Bernoulli random variable

Definition 40

Sample variance

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (6.4)$$

And the corresponding standard deviation ⁶⁴⁶⁴Is this correct??

$$s = \sqrt{s^2} \quad (6.5)$$

Lemma 2

Consider

$$E[\bar{X}] = E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n E[X_i] \quad (6.6)$$

Assuming that X_i have the same distribution, $f(x_i)$, and mean μ ,

$$\begin{aligned} &= \frac{1}{n} \sum_{i=1}^n \mu \\ &= \frac{1}{n} n\mu \\ &= \mu \end{aligned} \quad (6.7)$$

- the X_i are identically distributed
- \bar{X} is an unbiased estimator of μ

We then have

$$\begin{aligned} S^2 &= \frac{1}{n-1} \sum_{i=1}^n X_i^2 + \bar{X}^2 - 2\bar{X}\bar{X} \\ &= \frac{1}{n-1} (n\bar{X}^2 - 2n\bar{X}^2 + \sum_{i=1}^n X_i^2) \\ &= \frac{1}{n-1} (-n\bar{X}^2 + \sum_{i=1}^n X_i^2) \quad (6.8) \end{aligned}$$

We are now ready to take the expectation of S^2 , which isRecall: if Y has mean, variance μ, σ^2 , then $\sigma^2 = E[Y^2] - \mu^2$

$$\begin{aligned} E[S^2] &= \frac{1}{n-1} (-nE[\bar{X}^2] + \sum_{i=1}^n E[X_i^2]) \\ &= \frac{1}{n-1} (-n(E[\bar{X}^2] + \text{var}[\bar{X}]) + \sum_{i=1}^n \sigma^2 + \mu^2) \\ &= \frac{1}{n-1} (-n(\mu^2 + \text{var}[\frac{1}{n} \sum_{i=1}^n X_i]) + n(\sigma^2 + \mu^2)) \\ &\quad \vdots \\ &= \frac{n-1}{n-1} \sigma^2 = \sigma^2 \quad (6.9) \end{aligned}$$

So we get

$$S^2 = \sum_{i=1}^n (X_i - \bar{X}_i)^2 \quad (6.10)$$

So we have an unbiased estimator of variance

SUBSECTION 6.2

Lecture 17: Random Sampling

Comment

The following from Lecture 17 - Lecture 25 will be *very* minimal since I am cramming for an exam

- Observation is a realization of an RV
- A sample is a subset of a population
- The normal distribution of a sample is the same as that of the RV
- A statistic is a function of a sample, e.g. mean/median/mode. A sample is considered biased if it consistently over/under estimates the statistic
- Probability distribution of a statistic: sampling distribution

Some special properties of a normal distribution:

- if X_1, X_2 are normal with means μ_1, μ_2 then the distribution of $X_1 + X_2$ is normal with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$
- If X is normal with μ, σ^2 , then $\frac{X}{n}$ has $\frac{\mu}{n}, \frac{\sigma^2}{n^2}$
- If X_1, \dots, X_n are normal all with μ, σ^2 then \bar{X} has $\mu, \frac{\sigma^2}{n}$

Does not hold for other distributions

Definition 41

Central Limit Theorem If X is the mean of a random sample of size n taken from a population with mean μ and finite variance σ^2 , then the limiting form of the distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (6.11)$$

As $n \rightarrow \infty$, is the standard normal distribution $n(z; 0, 1)$.

SUBSECTION 6.3

Lecture 18: More on sampling

More notes on CLT

- Standard deviation of $\bar{X} \approx \sigma/\sqrt{n}$, so it shrinks at rate \sqrt{n} . This makes sense since we can expect a bigger sample size to enable better estimates.
- CLT lets us estimate how good that estimate is and draw a confidence interval

Whereas the CLT was derived from the sampling distribution of \bar{X} , inspection of the sampling distribution of S^2 yields the χ^2 distribution

Definition 42

If S^2 is the variance of a random variable of size n taken from a normal population having the variance σ^2 , then the statistic

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n = \frac{(X_i - \bar{X})^2}{\sigma^2} \quad (6.12)$$

is distributed as a chi-squared distribution with $v = n - 1$ degrees of freedom.

There is one more degree of freedom if we use μ instead

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n = \frac{(X_i - \mu)^2}{\sigma^2} \quad (6.13)$$

Recall: μ is the mean of the entire sample space, \bar{X} is the mean of the sample

There are also a few other distributions to consider

Definition 43

***t*-distribution**

$$h(t) = \frac{\Gamma[(v+1)/2]}{\Gamma[v/2]\sqrt{v\pi}} \left(1 + \frac{t^2}{v}\right)^{-(v+1)/2} \quad (6.14)$$

Where $v, t \in \mathbb{R}$

Consider the statistic

$$T = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (6.15)$$

Where⁶⁵

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \quad (6.16)$$

⁶⁵This also requires that the X_1, X_2, \dots, X_n are normal

- If the sample is large⁶⁶, S is close to σ and T follows normal
- If sample is smaller then the *t*-distribution is much more accurate
- If we knew σ exactly then we'd have a normal distribution; instead we estimate S^2
- Often used in problems that deal with inference about the population mean or involving comparative samples
- The use of the *t* distribution and sample size do not relate to the CLT

⁶⁶ ≥ 30

Definition 44

***F*-distribution**

Whereas the *t* distribution is useful for comparisons between sample *means*, the *F* distribution is useful for comparisons between sample *variances*.

F is defined as the ratio of two independent chi-squared RV, each divided by its number of degrees of freedom

$$F = \frac{U/v_1}{V/v_2} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \quad (6.17)$$

And the *F* distribution is given by

$$h(f) = \begin{cases} \frac{\Gamma[(v_2+v_2)/2](v_1/v_2)^{v_1/2}}{\Gamma[v_1/2]\Gamma[v_2/2]} & \text{if } f > 0 \\ 0 & \text{if } f \leq 0 \end{cases} \quad (6.18)$$

SECTION 7

Midterm 2

SUBSECTION 7.1

Lecture 19 - 25: Midterm 2 Cheat-sheet

The following cheat-sheet was created in lieu of about two weeks of lecture. I may typeset the more important excerpts from it. It largely covers confidence, prediction, and tolerance intervals for different scenarios as well as maximum likelihood estimators. As an aside 50% of this midterm was on the *Poisson distribution* which had practically no relevance to this part of the course ...but what's done is done :')

SECTION 8

Hypothesis Testing

Comment

This class from Lecture 25 → the end covers hypothesis testing through chapters 10.1 → 10.14 of the textbook[1] and some simple linear regression models for chapters 11.1 → 11.2

SUBSECTION 8.1

General concepts

We may often want to test things e.g. figuring out based on sample data whether if coffee drinking would increase the risk of cancer. These conjectures can be put into the form of a statistical hypothesis for which we may test.

Definition 45

A **statistical hypothesis** is an conjecture concerning one or more populations

A few notes about the relationship of probability with hypothesis testing

- We may never know if a statistical hypothesis is absolutely true or false – but we can take a random sample from the population of interest to examine if it supports/does not support the hypothesis.
- Whereas accepting a hypothesis does not rule out other probabilities, *refuting* one tends to all but ‘rule’ out the hypothesis; that firm conclusion is established when a hypothesis is rejected.
 - The formal statement of the hypothesis is important

Definition 46

Null hypothesis H_0 : any hypothesis we wish to test.

Alternative hypothesis H_1 : the hypothesis that is accepted on rejection of the null hypothesis.

Usually H_1 is the question to be answered or theory to be tested. Therefore H_0 is usually the logical complement of H_1 .

This leads to two conclusions this test can make:

1. reject H_0 in favour of H_1 because of sufficient evidence in data
2. fail to reject H_0 in favour of H_1 because of insufficient evidence in data

There are a few more terms to introduce

Definition 47

- **Test statistic** X : the statistic on which we base the decision on
- **Critical Region** : All possible values of X for which the decision criteria is satisfied⁶⁷
- **Critical Value**: the value of X in the critical region which is closest to not satisfying the decision criteria
- **Type 1 error**: Rejection of the null hypothesis when it is true.
- **Type 2 error**: Non-rejection of the null hypothesis when it is false
- **Level of Significance**: the probability, α , of a Type 1 error

⁶⁷ Doesn't this look like a confidence interval?

- The probability of committing a type 2 error (β) is impossible to compute if we don't have a specific alternative hypothesis.⁶⁸

Generally we want to minimize α, β in tests.

There are a few important properties to a hypothesis test

1. Type 1 and Type 2 errors are related. Generally decreasing one increases the other.
 - For example β can be reduced by increasing the size of the critical region – which would have the effect of increasing α .
2. The size of the critical region⁶⁹ can always be reduced by adjusting the critical values
3. Increasing the sample size will decrease the probability of committing both type 1 and type 2 error.
 - For large enough n we may use the normal or T distributions and use a Z test to determine α, β .
4. If the null hypothesis is false, β is the maximum when the true value of a parameter approaches the hypothesized value. The greater the distance between the true and hypothesized value the smaller β will be

Definition 48

The **power** of a test is the probability of rejecting H_0 given that a specific alternative is true.

SUBSECTION 8.2

One- and Two-Tailed Tests



Figure 29. No, not these tailed beats

Comment

A **one-tailed** test is one where the critical region lies entirely in the left or right side. Likewise, a **two-tailed** test is one where the critical region is split in two parts.

Remark H_0 is often stated using the equality sign $=$. This implies that the alternative 'does not reject H_0 ' can take on multiple values. So whether or not if a one-tailed or two-tailed test is used will depend on the picked alternative.

⁶⁸ Note that this number can appear to be large if we don't restrict the number of alternative hypotheses.

⁶⁹ And the probability of committing a type 1 error

A one-tailed test could be used to test if a vaccine is effective for at least 2 years. A two-tailed test can be used to test if a product weight is within a given error.

SUBSECTION 8.3

P-Value

It is customary to pick $\alpha = 0.05$ or 0.01 and then select the critical region accordingly. For example, if we are using a two-tailed test with $\alpha = 0.05$ and the test involves the standard normal distribution, the critical region would be given by $-1.96 < z < 1.96$, where $1.96 = z_{0.025}$. The result of the test can then indicate if the value of the test statistic is significant or not. However just pre-selecting α does not account for test statistics that are 'close' to the critical region. If to extend on the prior example, if a value $z_i = 1.87$ is observed, the value is not significant, however, the risk of committing a type 1 error⁷⁰ is roughly $P = 2P(Z > 1.87) = 0.0614$ which can be very useful to know. The P value is designed to give the user an alternative to just a reject/do not reject conclusion; it gives the user information about when the z -value falls within the *ordinary* critical region.

⁷⁰ Rejection of H_0 when it is true

Definition 49

The **P -value** is the lowest level of significance at which the observed value of the test statistic is significant.

SUBSECTION 8.4

Usage steps

Approach to fixed α testing:

1. State the null and alternative hypotheses
2. Choose a fixed significance level α
3. Choose an appropriate test statistic and establish the critical region based on α
4. Reject H_0 if the computed test statistic is in the critical region. Otherwise do not reject.
5. Draw conclusions

Approach to P -value testing:

- State the null and alternative hypotheses
- Choose an appropriate test statistic
- Compute P -value based on computed value of test statistic
- Use judgement based on P and knowledge of the system

SUBSECTION 8.5

Goodness of Fit

Given

- Discrete RV with $i = 1, \dots, k$ outcomes
- n trials
- $e_i = nP(i)$ expected frequency of outcome i
- o_i observed frequency of outcome i , O_i is the RV

We may define a statistic for the goodness of fit:

$$\tilde{\chi}^2 = \sum_{i=1}^k \frac{(O_i - e_i)^2}{e_i} \quad (8.1)$$

$\tilde{\chi}^2$ follows an approximately chi-squared distribution with $v = k - 1$ degrees of freedom. A small $\tilde{\chi}^2$ is a good fit.

Now, applying a hypothesis/confidence interval test,

Let H_0 be that the distribution is $P(i), i = 1, \dots, k$ with frequencies e_i, \dots, e_k .

Let

$$\alpha = \int_{\chi_\alpha^2}^{\infty} f(y; v) dy = 1 - F(\chi_\alpha^2) \quad (8.2)$$

Then, $\chi_\alpha^2 = F^{-1}(1 - \alpha)$. So given the observed values

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \quad (8.3)$$

The critical region for rejecting H_0 is given by $\chi^2 > \chi_\alpha^2$

F is the CDF of the chi-squared distribution.

SECTION 9

Statistical Modelling

Comment

In the most basic case statistical modelling can be thought of finding a function $y = f(x)$ such that it best models a sample set of data points (x_i, y_i) . Thought there are various ways of doing this (numerical, graphical, analytical, etc.), they generally follow the form of defining an error function

$$\mathcal{E} = J(y, f(x)) \quad (9.1)$$

With which the *best* model may be quantified, then minimizing it to arrive at the best model

SUBSECTION 9.1

Simple Linear Regression

Idea: choose a linear function $y = ax + b$. If we let the proposed model to be $h(x)$, then the error e_i would be given by $e_i = y_i - h(x_i)$. We can either use that error directly or note that we would want to penalize larger e_i -s more than smaller ones, the intuition for which being that a best fit would not necessarily exactly hug the data points⁷¹ but would rather capture the trends. Therefore we propose

$$\mathcal{E} = \sum_{i=1}^n \frac{e_i^2}{i} = \sum_{i=1}^n (y_i - h(x_i))^2 = \sum_{i=1}^n (y_i - ax_i - b)^2 \quad (9.2)$$

We may then find $\frac{d\mathcal{E}}{da}, \frac{d\mathcal{E}}{db}$ and set to zero in order to find a, b that minimizes b

$$\frac{d\mathcal{E}}{da} = \sum_{i=1}^n \frac{d}{da} (y_i - ax_i - b)^2 = \sum_{i=1}^n -2x_i(y_i - ax_i - b) = 0 \quad (9.3)$$

Rearranging, we find the ‘normal equations’

⁷¹ overfitting

Linear regression with least squares has an easy analytical solution; more complex modelling with more dimensionality would be solved numerically

$$\begin{aligned} nb + a \sum_{i=1}^n x_i &= \sum_{i=1}^n y_i \\ b \sum_{i=1}^n x_i + a \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i y_i \end{aligned} \tag{9.4}$$

And if we take the mean x_i, y_i to be \bar{x}, \bar{y} respectively, we may solve to find that

$$\begin{aligned} a &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})y_i}{(x_i - \bar{x})^2} \\ b &= \frac{\sum_{i=1}^n y_i - a \sum_{i=1}^n y_i}{n} = \bar{y} - a\bar{x} \end{aligned} \tag{9.5}$$

9.1.1 Using Maximum Likelihood

Recall maximum likelihood estimators

- sample x_1, \dots, x_n
- Joint pdf with parameter θ ; $f(x_1, \dots, x_n; \theta)$
- Likelihood function $L(x_1, \dots, x_n; \theta) = f(x_1; \theta)f(x_2; \theta) \dots f(x_n; \theta)$
- Maximum likelihood estimator: $\hat{\theta} = \max L(x_1, \dots, x_n; \theta)$

So, noting that for the linear regression case errors e_i is a realization of a random variable E_i with $\mu = 0, \sigma^2$. This implies each y_i has $\mu = ax_i - b, \sigma^2$

So, we may write⁷²

$$L(e_1, \dots, e_n; a, b) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e_i^{-\frac{e_i^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}} \prod_{i=1}^n e_i^{-\frac{(y_i - ax_i - b)^2}{2\sigma^2}} \tag{9.6}$$

And maximizing the likelihood L over a, b would give the least squares solution

9.1.2 Variance Analysis & True Regression Line

If we take a step back we may realize that each x_i, y_i are actually realizations of a random variable X_i, Y_i ⁷³. This means that the a, b are *realizations* of A, B , which are in turn estimators of α, β , the 'true regression coefficients'.

We would like to know the expectation and variance of A, B to find out if it is an unbiased estimator.

Note that this is a simple convex quadratic (can just plot it), so finding an inflection point would just give the minimum error

⁷² Since in this case the parameter θ is (a, b)

⁷³ with the same mean/variance

$$\begin{aligned} E[Y_i] &= \alpha x + \beta \implies \mu_A = E[A] \\ &= E\left[\frac{\sum_{i=1}^n (x_i - \bar{x})Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}\right] \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})E[Y_i]}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})(\alpha x_i + \beta)}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \alpha \end{aligned} \tag{9.7}$$

And a similar analysis can be carried out to find the following:

$$\sigma_B = \beta \quad (9.8)$$

$$\sigma_A = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (9.9)$$

$$\sigma_A = \frac{\sigma^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \quad (9.10)$$

And therefore we may conclude that A, B are unbiased estimators of α, β .

9.1.3 Error Estimation

Comment Given $y_i = ax_i + b + e_i$, where e_i is a realization of a random variable E_i with σ^2 , then σ^2 is the error

Define the statistic⁷⁴

$$S^2 = \frac{\sum_{i=1}^n E_i^2}{n - 2} \quad (9.11)$$

Demoting the regression estimate to be \hat{y} such that $e_i = y_i - \hat{y}$, the realization of the S^2 would be

$$\begin{aligned} s^2 &= \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2} \\ s^2 &= \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 - \alpha \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 2} \end{aligned} \quad (9.12)$$

Note: Total squared error is $\sum_{i=1}^n e_i^2$

⁷⁴ This is an unbiased estimator of σ^2

SUBSECTION 9.2

Support Vector Machines

The key ingredient in a SVM is the hyperplane, which is completely denoted by the normal vector \vec{w} , where $w \in \mathbb{R}^4$ denotes the orientation of the hyperplane.

We may also write any hyperplane in the following format:

$$w^T x - b = 0 \quad (9.13)$$

Classification:

Given paired training data, we want to find a hyperplane (or hyperplanes) such that we can divide the data into classes.

$$(X_i, Y_i) \quad i = 1, \dots, n, X_i \in \mathbb{R}^n, Y_i \in \{-1, 1\} \quad (9.14)$$

Graphically, we may plot the training data onto a plane and attempt to draw a hyperplane $w^T x - b = 0$ such that we divide the classes.

Here there are only two classes for the sake of this derivation, but realistically there would be more than just 2 classes.

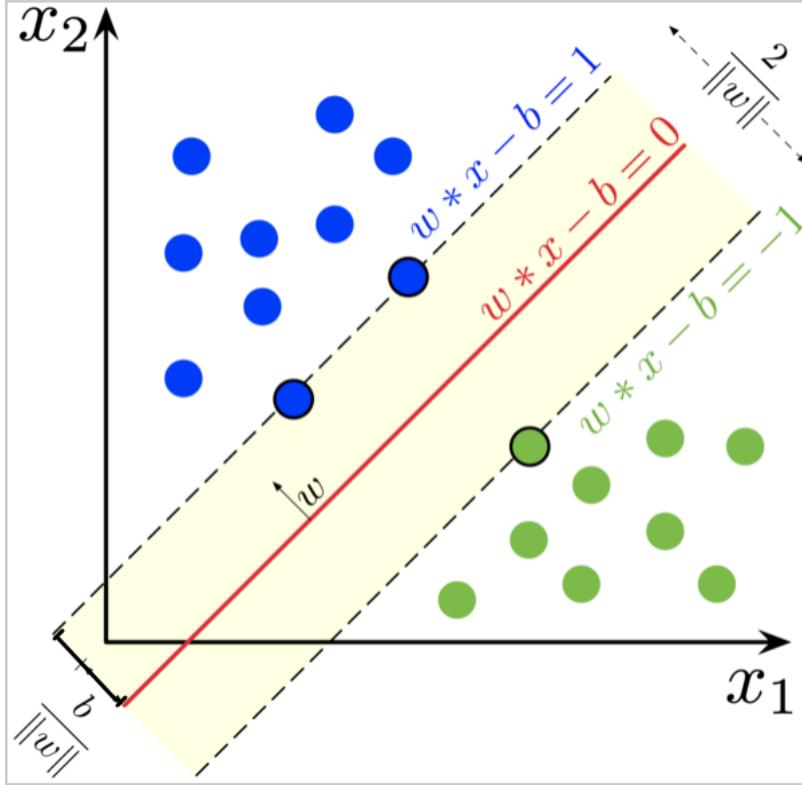


Figure 30. <https://medium.com/@dr.sunhongyu/machine-learning-c-svm-support-vector-machine-simple-example-deff5d55d43e>

From which we arrive at three cases

1. $w^T x - b = 0$ on HP
2. $w^T x - b > 0$ for $Y_i = 1$
3. $w^T x - b < 0$ for $Y_i = -1$

Therefore we may optimize w, b over the training data in order to create a predictor which, when given a new X_i , we may evaluate it for $w^T x - b$ to predict if Y_i is -1 or $+1$

Our optimization heuristic is to pick a hyperplane w such that it maximizes the distance between the hyperplane and the data. We may now draw three parallel equidistant hyperplanes with equations

1. $w^T x - b = 1$
2. $w^T x - b = 0$
3. $w^T x - b = -1$

We may now define a quantity b , for which we aim to maximize through the manipulation of quantities w, b .

To compute the distance b between the hyperplanes $w^T x - b = 1$ $w^T x - b = -1$ we note that if we translate it to the origin, i.e. $w^T x = 0$ $w^T x^* = 1, b$ is just the magnitude of x^*

$$b = |x^*| = \frac{1}{|w|} \quad (9.15)$$

As for non-linear tasks there are a bajillion approaches, one that Josh Taylor suggested that is popular and simple is to use kernel methods, just use a neural network, or even transform the data into a space where it is linear and use linear methods.

And therefore maximizing b means minimizing $|w|$. We are also not interested in a trivial solution e.g. $w = 0$, so we restrict it with

$$Y_i(w^T x_i - b) \geq 0 \quad i = 1, \dots, m \quad (9.16)$$

And the nice thing is we can square this to arrive at a convex quadratic program, which is really easy to solve for huge numbers of parameters.

SUBSECTION 9.3

Markov Chains

Recall: LTI systems.

- state x_k
- Transition matrix A
- E.g. for dynamics $x_{k+1} = Ax_k$

Markov chains take that to the next level

Given:

- n states
- $p_i(k)$ probability of being in state i at time k

In the past we've worked with statics. Markov chains are the easiest way for us to put dynamics and probability together

$$\sum_{i=1}^n p_i(k) = 1 \quad \forall k \quad (9.17)$$

Let P_{ij} be the probability of going to state j from state i . Then, as we have

$$\sum_{j=1}^n P_{ij} = 1 \quad (9.18)$$

We may find that

$$P_i(k+1) = \sum_{j=1}^n p_j(k) P_{ji} \quad \forall i = 1, \dots, n \quad (9.19)$$

Markov chains also have a really nice graphical interpretation as sort of a FSM with *probabilities* of moving from one state to the next.

As it turns out a lot of these Markov chains can be represented as matrices.

The vector form of the probability mass function is

$$p(k) = \begin{pmatrix} p_1(k) \\ \vdots \\ p_n(k) \end{pmatrix} \quad (9.20)$$

Set

This is for a linear time-invariant Markov chain

$$M = \begin{bmatrix} p_{11} & \dots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \dots & p_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n} \quad (9.21)$$

And we know that $M_{ij} = P_{ji}$, so we have $p(k+1) = Mp(k)$. So we have now represented the Markov chain graph as a matrix.

Consider

$$\begin{aligned}
 p_i(k+1) &= M_i p(k) \quad \text{ith row of } M \\
 &= [P_{1i}, P_{2i}, \dots, P_{ni}] \begin{pmatrix} p_n(k) \\ \vdots \\ p_n(k) \end{pmatrix} \\
 &= \sum_{j=1}^n p_j(k) P_{ji} \\
 \implies p(k+s) &= M^s p(k)
 \end{aligned} \tag{9.22}$$

Example Suppose we have a thermostat. A useful way to model the heating system is as a Markov chain with two states; on (1) and off (0). Giving it some numbers, if it is on, then it will have a $p_{11} = 0.8$ probability of staying on and $p_{10} = 0.2$ probability of turning off. If it is off, it has a $p_{00} = 0.3$ probability of staying off and $p_{01} = 0.7$ of turning on.

So,

$$M = \begin{bmatrix} 0.3 & 0.2 \\ 0.7 & 0.8 \end{bmatrix} \tag{9.23}$$

So given this transition matrix we can move along forwards in time by taking the exponent.

Given q is an eigenvector of eigenvalue 1 and

$$q = \lim_{k \rightarrow \infty} M^k q \tag{9.25}$$

So q describes the steady state behavior of the system.

Another cool thing that we can do is with M^T

Recall: $\sum_{j=1}^n P_{ij} = 1$.

We know that identity $\mathbb{I} = M^T \mathbb{I}$

So \mathbb{I} is an eigenvector of M^T

Define:

- **Irreducible** Markov chain: can always go from one state to the other
- **Reducible** Markov chain: there are steady state states that the system can get stuck in

This was something applied in Google's page rank algorithm. They assumed that each website was a state, and that each outgoing/incoming link had equal probability. The problem was that they had some websites without outgoing links. This made the network reducible, which isn't great. So they added a small probability to allow for any page to go to any other page, which made it irreducible. These steady state probabilities were then calculated and then their initial page ranks were derived.

SUBSECTION 9.4

Lecture 36: Review

9.4.1 Bayes Rule

Example Test for disease: want to find $P(D|P)$ where

- D is probability of having disease
- P is probability of a positive test

Recall: Eigenvectors

Definition of eigenvector:
Suppose $q \in \mathbb{R}^n$ such that

$$q = Mq \tag{9.24}$$

Then q is an eigenvector of M

- \bar{D} is probability of not having the disease

$$\begin{aligned}
 P(D|T) &= \frac{P(P|D)P(D)}{P(P)} \\
 &= \frac{P(P|D)P(D)}{P(P \cap D) + P(P \cap \bar{D})} \\
 &= \frac{P(P|D)P(D)}{P(P|D)P(D) + P(P|\bar{D})P(\bar{D})}
 \end{aligned} \tag{9.26}$$

The denominator can be interpreted as the sum of probability of a false positive and probability of a true positive.

9.4.2 Sampling

- Assume that we have a really big population of size N which is too big for us to measure everyone in it
- Assume that the population has a distribution $f(x)$
- A sample of size n is taken, where $n \ll N$
- Can think of the sample as a set of random variables X_1, X_2, \dots, X_n , each of which has the distribution⁷⁵ $f(x)$
- The realized values of the random variables are x_1, x_2, \dots, x_n

⁷⁵ Assume that they are independently and identically distributed; IID

We would now want to glean some insights into the sample.

Sample mean:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \tag{9.27}$$

Sample variance:

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x - \bar{x})^2 \tag{9.28}$$

From the sample mean we may arrive at the Central Limit Theorem which states that for the following statistic Z ,

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \tag{9.29}$$

Which states that as $n \rightarrow \infty$, the PDF of Z approaches a normal distribution $n(z; 0, 1)$

Confidence intervals are given by

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \text{alpha}$$

100(1 - α)% gives the confidence percentage. The values for $z_{\alpha/2}$ can be found in a table.

Rearranging to get it in terms of the sample we find that

$$P\left(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha \tag{9.30}$$

For cases where we don't know σ^2 , we can use the T distribution. This, however, requires that $f(x)$ must be normal. We don't have a great distribution for cases where we don't know σ^2 , n is small, and f is non-normal

- Prediction intervals: building a confidence interval for our next observation.

The nice thing about CLT is that it works for any distribution $f(x)$, as long as we know the variance σ^2 , or if $n \geq 30$ so that use the sample variance

- Tolerance limit: build $\bar{x} = ks$, finding k such that the tolerance limit, e.g. 95%, of the population will fall within this range.

SUBSECTION 9.5

Misc

A few lectures were given on topics not covered in the textbook. Aside from Monte-Carlo integration, a lecture was given on optimization via statistical modelling of supply/demand for energy systems which is kind of a rehash of the news-vendor problem: https://en.wikipedia.org/wiki/Newsvendor_model

9.5.1 Monte-Carlo Integration*Comment*

Handwritten notes were taken. For more, see https://en.wikipedia.org/wiki/Monte_Carlo_integration

TEP327

Engineering and Law

Torts can be identified by

- Elements of negligence can be identified
- There exists a duty of care
- Was the standard of duty of care breached?
- Did that breach of standard cause damage to the plaintiff?

Comment

Must be careful about the standard of care met or not met. For example in the in-class bike example one could argue that in designing the gear the engineer should expect that the gear could be commercialized and mass-produced later on for mass-market. On the other hand it is reasonable to argue that having to design the gear to be used by unskilled 16-year-olds is an unreasonable and outside the standard of care.

Trespass is strict liability. Ignorance of the law is no excuse.

SECTION 10

Property

- There are a number of types of property, i.e. intellectual property, personal (usually regarding money), etc
- Real property
 - Usu. owned by default by government (derived from old english law)
 - Gov. transerrs property rights to private owners; sell, lease, etc (fee simple)
 - Gov. retains a few rights on real property, usu. oil and mineral rights
- Joint ownership: death of one partner causes other to own 100%. Tenancy in common/joint tenancy: death of a partner does not cahnge ownership; usu. determined by wills
- Land interest is divisible; different people can own rights to different things on a piece of land. Or many people can own right to lease a property but not necessarily ownership of the land
- Rights associated with real property:
 - Mortgage: property as security for a debt
 - License: Contractual right to use property in some way
 - Restrictive covenant: conditions of use for that property
 - Easement: right to use property in some way
 - Lien: right to register debt against land (must pay off people you owe with money from selling property before taking it for yourself)

- Lease: right to occupy the property
- Profit à prendre; right to extract value from property

Mineral, oil, gas rights:

- Mineral rights are crown or "freehold"
- Crown rights can be provincial or federal and can grant license to third party to extract resources
- Rights to resources generally excluded from fee simple property rights
- Freehold mineral rights were granted with early land grants
- Resource rights usu. include a right to access to the land in order to extract resources
- Land is registered via the registry system/land title system. Registry system is old and only about 10% of land is in it, is complicated, and led to a lot of squatters. Land title system (ontario) is easy to use.
- **chattels** are tangible personal property called 'goods'. No registry for ownership but is registration of security interests (which demonstrates rights based on order of registration).

SUBSECTION 10.1

Intellectual Property

- A bundle of right that protect expression of ideas, inventions, symbols, and more
- Governed by federal legislation and international treaties
- Has a few different types.
- Copyright
 - Must be an original product of skill, permanent, and published
 - Ephemeral works e.g. live works have no copyright
 - Most things are copyrightable.
 - Copyright protection is valid for about 50 years after death, soon to be 70.
 - Protection is not dependent on registration, but copyrights can be registered
 - Copyright can be assigned by contract, but moral rights cannot be [although they can be waived]
- Trade secrets
 - What todo if you invent something?
 - Keep your mouth shut
 - Not everything is patentable and not all things are patentable
 - Trade secret is information that is kept secret with an industrial or commercial application
 - Once the secret is out it's not possible to make it a secret again
 - Reverse engineering is a totally valid way to discover a trade secret
- Patent
 - A right granted by the crown to a monopoly

- Term of patent in Canada is 20 years, non-renewable
- A balance between the interests of the inventor and that of society
- A few things that cannot be patented; scientific principles, abstract theorems, higher life forms (animals).
- A patent must be new, useful, and non-obvious
- Must pass the obviousness test: 1) Identify a PSITA (Personal skilled in the art) 2) Identify inventive concept 3) Find difference between prior art and inventive concept 4) Ask: Would PSITA find this obvious?
- A patent will have a specification: a description of the procedure/invention, how to reproduce it, etc.
- ... and claims
- Patents are filed with the Canadian Intellectual Property office and the federal court always has jurisdiction
 - * Clear and explicit definition of the subject matter
 - * Will inform the scope of the rights granted
- Infringement is "Use by someone other than the patent holder which interferes in whole or in part with the patent holder's monopoly" (Monsanto Canada Inc. v. Schmeiser, 2004 SCC 34)
- Usu. more to do with the right to monopoly than the IP itself.
- Some exceptions: use of the invention for development and submission of information required by law, fair dealing (experimentation without eye to profit or fraudulent purpose), dedication (patent owner can ignore monopoly), exhaustion (once patent is sold, patent holder may no longer control use)

This is the case in Canada but not really in other places.

This also implies a patent can only have a single inventive concept

- Trademark
 - A sign or combination of signs that is used by a person for sake of distinguishing or so as to distinguish their goods and services from those of others
 - A certification mark
 - **trade name:** means the name under which any business is carried on, whether or not it is a corporation, partnership, or individual
 - 15 years of protection, is renewable
 - Certain marks are prohibited from being a trademark; royal coat of arms, Canadian flag, etc.
- Industrial Designs and Integrated Circuit Topographies: 10 years and protects the design, look, and configuration of ICs, industrial design. Registered of topographies designed by minister from people in the department of industry

Summary

- Copyright: inherent, no registry. Need to provide proof.
- Trademark: registry
- Patent: registry
- Topography & industrial design: registry. Unknown how strong the protection is so far.

In all cases there is no protection for the idea; the invention or creation itself is protected. Let's say that you want to patent your invention.

- Search registered patents for prior art. This is a difficult process, usu. patent agents will do this
- EPO (European) is usually a good place to start because their website doesn't suck. Canada/US have one but it isn't great to work with.

For trademark, search registry to make sure that there is no prior registry. Note that one can't trademark a name, but it can be trademarked for an item if it can be proved to be distinctive; Dyson vacuums, Apple computers, etc.

Trade secret: no protection if competitor reverse-engineers it. As trade secrets are not public, it is not prior art and a competitor could reverse engineer it (which is totally valid).

Know-how: personal to you or corporation; i.e. a plumber's experience, or a boat-builder's special way of bending ribs. Secret, substantial, identified, and valuable; non-patented practical information which has an economic value not publicly accessible. This is different from a trade secret because just telling someone the know-how won't give them the competitive advantage.

SECTION 11

Contracts

- Sets out rights, responsibilities, and liabilities of parties
- Allocates risks between parties
- Enforceable
- Do not have to be written
- Can have express/implied terms
- Only parties privy to contract can enforce terms

For there to be a contract, there must be:

- Offer and acceptance
- Capacity to contract (People are in the right state to enter a contract)
- Consideration
- Intention to create legal obligations
- Lawful purpose

The offer and acceptance process can be messy

- Invitations to treat/offer to treat are calls for offers (e.g. ads); they are not offers unto themselves
- A counter offer is a rejection of an offer

Capacity to contract

- A person under law: partnerships may contract, sole proprietorships contract in the name of their sole proprietor
- Must have capacity
 - Age of majority

- Incapacitated due to alcohol/drugs
- Mentally unfit
- Ostensible authority: where one partner contracts, other is implicitly involved as well even if they are not involved.

Consideration

- The contract must exchange something of value (the "consideration")
- Consideration is essential for there to be a contract
- No consideration = no contract; instead it will be a gift or merely an unenforceable promise
- Considerations do not have to be of equal value
- Services and forbearance are valid consideration

Intention to create a contract

- Must intent to create a contract
- An agreement to agree is not a contract
- Letters of intent or memorandums of understanding (MOU) are not contracts; they are intentions to form a contract later on
- MOUs can become contracts under the right circumstances

Professional service agreements

- Define the scope of work really well
- Define anticipated result/outcome as well as cost/value
- Should contain limitation of liability
- May contain alternative remedies

Terminating contracts

- Voiding: making them as if the contract never existed.
 - Misrepresentation
 - * Doesn't need to be fundamental; something being done in order to induce a party into a contract. Can have many levels e.g. negligent, innocent, fraud. Each level has its own level of consequences. Damages usually available for non-innocent misrepresentation
 - Distress
 - Unconscionable; somebody being taken advantage of. A contract so unfair, oppressive, and one-sided that the court finds it repugnant to enforce it. E.g. an extorted pre-nup, coercion to sign a contract, etc. Pressure must be extreme as to essentially remove the free-will of the person to contract.
 - Frustration; terms not anticipated and cannot be fulfilled. I.e contract for steel and then the factory blows up before it can be fulfilled.
 - * Usually something that's like an act-of-god; natural disasters, etc.
 - * In Engineering contracts we often incorporate "Force majeure clauses" which spell out the consequences of unlikely events

- * Must be careful about what is covered and not by force majeure
- Mistake: both parties must be mistaken and the mistake must be 'material'/fundamental – and must be made at the time of the contract being signed. Mistake is an 'equitable' doctrine.⁷⁶ For example if the seller tries to sell a property and turns out they didn't actually own it.
 - * A mistake as to the identity of the other contract
 - * A mistake as to the subject of the other contract
 - * A mistake to the existence of the subject matter of the contract
 - * A mistake to the legal character of the transaction involved, making the contract not valid
- Breaking: ending the contract but acting as if the contract did exist prior.
- Waiver: when one party relieves the other contracting party from performing all or part of the contract
- Estoppel: 'to be stopped'; can prevent a party from bringing forwards a claim
- Changes to contracts: must be agreed upon by all parties⁷⁷. If original contract sets out how it can be changed, changes must be done in line with that procedure.
- Breach & remedy
 - Contract breached when one party does not perform their end
 - If parties cannot negotiate a settlement (usually described in contract), courts can be involved
 - Monetary damages are usually the most common remedy
 - Contracts can be breached in a few ways. Inability to perform the contract does not justify a breach. Inadvertence, i.e. lack of intention to break the contract does not act as a defense of contract breach.
 - Anticipatory breach; one contracting party informs the other does not intend to perform all or part of its contractual obligations. Upon receipt of such notice the injured party can accept that the contract has already been breached and sue for damages at at ti. Can be defended through litigation, remoteness, causation; i.e. innocent party should take all reasonable steps to minimize loss, damages are foreseeable when contract was formed, and only , or they can wait until the actual breach. When should one accept breach? Pick whichever one works best i.e. minimizes costs.
- Remedies: typically monetary compensation ordered by court for the at fault party to pay. Usually damages are sought
- Damages: usually calculated as the compensation required to place the non-breaching party in the position they would've been but for the breach of contract. Can be defended through litigation, remoteness, causation; i.e. innocent party should take all reasonable steps to minimize loss (litigation), damages are foreseeable when contract was formed (remoteness), and only losses as a result of the breach of contract (nothing speculative) (causation).
 - Consequential (Compensatory): reimburse innocent party for any indirect damages outside of those specifically mentioned in the contract. For example loss of business due to failure to provide.
 - Expectation (Compensatory): provide the innocent party the amounts they expected to obtain from the contract, done based upon contractual terms or market value

⁷⁶ This is a historical artefact of when there were legal courts (which emphasized application of law) and equitable courts where judgement was made as per fairness. Mistakes are an equitable doctrine made out of 'natural law'

⁷⁷ keep it in writing...no sticky notes

- Liquidated: Generally spelt out in contract. Must be a genuine pre-estimate of incurred damages. Must be reasonable. I.e. 10k/month if this project is not done.
- Specific performance
- Performance under protest: what if disagreement over terms of contract? Performing a contract under protest and reserve the right to claim later.
- Exclusion: excludes certain damages or remedies available to innocent party to the contract. Often found in insurance.
- Liability: limits liability but does not entirely excludes it
- Consequential damages clause: often in commercial clauses; restricts ability to claim for lost profits or other damages indirectly related to claim being made; usually used in construction to make the damages more reasonable
- Contract termination: can be ended by complete performance, frustration, mutual agreement, or breach. A party may intentionally breach a contract.
- Termination clauses: specifies acts that justify termination of the contract
- No contract but performed work – ‘Quantum Work’: the amount it is worth. Awarded for work already done but there may not be contract. This is at the mercy of the court. Often this is found in elder care. Shouldn’t be something found in engineering.

SUBSECTION 11.1

Drafting and Interpretation of Contracts

Comment

Professional service agreements

- Define the scope of work **really well**
- Define anticipated result/outcome
- Set value: fixed price, cost plus, or a percentage of actual costs
- Limitation of liability clause

Be really careful about standard form contracts!!! Address the following issues in detail

- Liability
- Site inspections and observations
- Termination
- Any IP issues

Use standard form contracts as a guideline for what you may have missed but don’t forget that there’s no standard project/contract! It’s ultimately your contract so look-at-it-carefully!

Do not implicitly trust standard form contracts

Generally speaking make sure that these clauses are there and that they are OK.

- Scope of work
- Contract time
- Contract price

- How Changes to work/design/time to be done
- Damages in event of breach
- Bonuses for completion better than anticipated
- Warranties
- Termination
- Indemnification (protection from liability)
- Exclusions
- Limitations
- Dispute Resolution mechanisms

Contracts are about assuming risk in order to get a benefit, and that you want to limit the risks that one takes.

SUBSECTION 11.2

Force Majeure

- Something unexpected that can be anticipated, i.e. earthquakes in an earthquake rich region
- Relying on force majeure requires that the event to be *entirely* out of each party's control.
 - Courts will determine if 1) if the change was great enough to strike at the root of the contract and 2) that was it really out of control and if the party would be reasonable able to make up for it.
- Evolves to describe the anticipated risks at the time of the contract
- In the past most contracts in the west did not address pandemics – but now they do (after COVID)
- If it were foreseeable at the time of event (courts will look into knowledge at the time to determine) then it must be negotiated. Contracts are living things – must look into things as we go along and then change them as we go along.
- Example of a force majeure case: Falk enterprises Ltd. V. McLaine 1994: defendant company rented a tent and it got destroyed by rain. And then there was excessive rain which destroyed the tent. Courts ruled that excess rain in Halifax was foreseeable at the time of contract, therefore force majeure is not a valid defense.
- Party seeking to rely on force majeure must make reasonable attempts to mitigate damages caused by force majeure. Courts look to standard of commercially reasonable.

SUBSECTION 11.3

Typical contract terms

- Scope: Be really specific about this! Especially for us engineers since only we can really understand this.
- Contract time
- Contract price

- How changes to work/design/time to be done
- Damages in event of breach
- Bonuses for completion better than anticipated
- Warranties
- Termination
- Indemnification (protection from liability)
- Exclusions
- Limitations
- Dispute Resolution mechanisms

Remember: contracts can be ended by complete performance, frustration, mutual agreement, or breach. A party may intentionally breach a contract, which can be cheaper.

SUBSECTION 11.4

Drafting a contract

Principles of interpretation

- Contract is understood as a whole. Will look at extrinsic evidence if the contract is not sufficient.
- Courts will interpret contracts objectively. Will read it according to plain and ordinary meaning (golden rule). Special meaning of words will be used only if the context requires it.
- A part of the contract must be read in a way that makes another irrelevant. So there can be readings that don't make a lot of sense but are still valid because the court will read it in such a way in order to make both of them relevant.
- Contract is the primary evidence. If the contract is complete then extrinsic evidence will not be considered.

SUBSECTION 11.5

Other

Agency

- A representation by an agent to a third party cannot create any type of agency
- Agency can be limited in scope
- Agents are not personally liable for the obligations they commit on behalf of their principals. However if they go outside of their scope then they can be held liable.

Indemnity

- Insurance policies are at their heart contracts of indemnity
- Many indemnity clauses simply reflect the common law that each party is responsible for the harm it causes but are added to agreements to help avoid litigation and to provide contractual remedies

- Professional engineers have to carry general liability for their practice. (E&O coverage, Errors and Omissions)

Indemnity:

- Occurrence: Covered for life, during the time it is covered. For example insurance against flood damage between 1999-2010; can claim it in 2020 as well.
- Errors and omissions: Covered only during time it is covered (depends on claims made). Often have to carry errors and omissions after retirement because people can still sue you for errors you have made prior.

Bids are always final (unless common law of mistake applies). Courts are usually reluctant to withdraw bids.

SUBSECTION 11.6

Agency and Authority

Engineers usually work as an agent for someone else; a client. Agency relationships can have either actual or apparent authority. The person we talk to on behalf of the client does not know the degree of agency we have; if we bind the client to something without the authority to do so it incurs misrepresentation.

SUBSECTION 11.7

Concurrent liability

Duties in tort and contract can co-exist. This doesn't mean that they can sue for 2x the value, but rather they come after you with 'double the heads'. However if the contract defines the nature and scope of the duty of care, it governs the relationship (Central Trust v Rafuse [1986]). It is also good to put in limitation periods, which are useful if you want to not get sued after you retire. These limitation periods are usually about 2 years after the date of discovery (commencement of running as per when they could have *reasonably* found out)

There can be a few methods of resolving a dispute; negotiation, mediation, arbitration, and court.

Negotiations may be privileged (any communications between parties and lawyers can be kept private) Mediation: a third party mediation facilitates resolution. Usually privileged if no settlement is reached. Arbitration private trial where parties set the rules. May be less expensive than litigation, but is not always the case. Can be binding/non-binding, depending. Arbitrators must be paid. Can take a long time because parties must agree on arbitration and the rules of procedure. Litigation: use of the court system to resolve disputes. May be time-consuming and expensive; really expensive; must fly people over, hotels, etc etc.

Only one that is not privileged is litigation court. Recall: usually civil trials (balance of probabilities)

Procedure:

- Pleadings
- Mandatory mediation (usually try to end it here)
- Disclosure of documents
- Discovery
- Pre-trial conference (judge meets with parties, gives memo with positions, etc)
- Trial
- Appeal (maybe)

Usually losing party pays the winner's costs, costs ordered payable are usually not all of it (usually about 2/3 of the actual amount)

Situations where engineers would get involved with: labour law (unions and stuff) and employment law (individual employees). Employment law would probably be the only one that would directly involve us⁷⁸ Employment is generally governed by contract law, although some legislation may apply e.g. minimum wage, benefits, etc. Implied terms can exist, e.g. duty of loyalty/competence, employer's duty to give adequate notice, employee's obligation to mitigate after a breach has occurred. Common terms for restrictive covenants include limiting what one can do after the contract is terminated, i.e. working for competitors, soliciting customers, etc. Courts are usually really narrow in their scope of authority for restrictive covenants; favour the employee.

Independent contractors don't have benefits and other things, or duty of confidence or loyalty etc. So if you show a independent contractor e.g. a plumber a trade secret, it is gone. Telling an independent contractor from an employee: apply a test: usually on the control the employer has on the worker; i.e. whether the worker provides his own equipment, risk taken by worker, etc

Vicarious liability: makes an employer liable for the work done by the employee, provided that it is being done in the course and scope of the employment.

SUBSECTION 11.8

Employer's responsibilities

- Employer standards and labour act
- Must protect against expected threats (unexpected OK, but subject to redefinition)
- Human rights; prevents discrimination, harassment, etc

SUBSECTION 11.9

Environmental law

- Mix of common and statutory law
- Audits may be required
- Must assist if there is an investigation
- EPA offences are akin to strict liability with only defence being showing that one was duly diligent
- Adjacent properties may be involved in ESA (Environmental site assessment) due to pollution migration

SUBSECTION 11.10

Aboriginal law

- Must understand treaty rights
- As occupiers they are treated as nations with the crown and formed agreements with nations
- Can contract with them; they are an entity
- There exists a duty to consult

⁷⁸But we should be respectful of unions as well!

Although doctors and lawyers are more visible, engineers ultimately make the decisions, though it is a little hidden.

References

- [1] R.E. Walpole, R.H. Myers, S.L. Myers and K. Ye, Probability & Statistics for Engineers & Scientists, 9th ed., Pearson, Inc., 2011. ISBN-13 978-0-13-411585-6.