

TSP (Thermal and Statistical Physics) Part II Phy

Tai Yingzhe, Tommy (ytt26)

Contents

1	Recap of Thermodynamics	3
1.1	Definitions	3
1.2	Zeroth Law and First Law	4
1.3	van der Waals Gases	6
1.4	Entropy	7
1.5	Chemical Potential	11
1.6	Potentials and Availability	13
1.7	Phase Transitions	16
1.8	Ideal gas mixtures	19
2	Recap of Statistical Mechanics	20
2.1	Definitions	20
2.2	Microcanonical ensemble	21
2.3	Gibbs Entropy	22
2.4	Canonical ensemble	23
2.5	Grand canonical ensemble	26
3	Classical Statistical Mechanics	30
3.1	Phase Space	30
3.2	Ideal Gas of N particles	31
3.3	Equipartition Theorem	32
3.4	Gas of Diatomic Molecules	33
4	Quantum Statistical Mechanics	36
4.1	Density of states	36
4.1.1	Open Boundary Conditions	36
4.1.2	Periodic Boundary conditions	36
4.1.3	Integral conversion	36
4.2	Bosons and Fermions	37
4.3	Ideal Fermi Gas	40
4.4	Ideal Bose Gas	42
4.5	Blackbody radiation	44
4.6	Vibrational Energy of Solids	46
5	Real Gases and Liquids	47
5.1	Virial theorem	47
5.2	Generic interacting gas/liquid	48
5.3	van der Waals gas	51
6	Phase transitions	52
6.1	Mixing and Phase Separation	52
6.2	Ising Model	54
6.3	Landau Theory of Phase Transitions	56
6.3.1	$B = 0$: a continuous phase transition	56
6.3.2	$B \neq 0$: a discontinuous phase transition	57
6.3.3	Close to Critical Point	58
6.4	Critical Exponents and Universality	59

7	Fluctuations in Equilibrium	60
7.1	Fluctuations	60
7.2	Connection with Thermodynamics	60
7.3	Near Critical Points	62
8	Stochastic Physics	63
8.1	Brownian Motion	63
8.2	Diffusion of Free and Confined Particles	63
8.3	Working with probability distributions	65
8.4	Fluctuation-Dissipation Theorem	67

1 Recap of Thermodynamics

1.1 Definitions

Definition 1.1 (Thermodynamic System): A thermodynamic system contains an arbitrary amount of matter and may allow exchange with the surroundings.

Definition 1.2 (Isolated Systems): Isolated systems do not interact in any way with the surroundings. Here, energy and particle number are conserved quantities.

Definition 1.3 (Closed Systems): Closed systems can only exchange energy with their surroundings. Energy is not conserved; if the system is in equilibrium with its surroundings, mean value of energy is related to the temperature of the system or the surroundings.

Definition 1.4 (Open Systems): Open systems can exchange energy and matter with their surroundings; energy and particle number are not conserved; if the system is in equilibrium with its surroundings, mean values of energy and the particle number are related to the temperature and chemical potential of the system or of the surroundings.

Definition 1.5 (Thermodynamic Equilibrium State): Thermodynamic equilibrium state is defined as the one macroscopic state of a system which is automatically attained after a sufficiently long period of time such that the macroscopic properties of the system no longer change with time.

Definition 1.6 (Thermal Equilibrium): A system is said to be in thermal equilibrium if there is no temperature difference between system and surroundings.

Definition 1.7 (Mechanical Equilibrium): A system is said to be in mechanical equilibrium if there is no unbalanced forces acting on any part of the system or the system as a whole.

Definition 1.8 (Diffusive Equilibrium): A system is said to be in diffusive equilibrium if there is no chemical reaction within the system and there is no movement of particles from one part of the system to the other.

Definition 1.9 (Thermodynamic State Functions): Thermodynamic state functions are macroscopic quantities which describe the system. The state functions are only well-defined and measurable when the system is in equilibrium.

Definition 1.10 (Equation of State): Equation of state gives a mathematical relationship between thermodynamic state functions. Equation of state is specified empirically in thermodynamics but is derived theoretically in statistical mechanics.

Definition 1.11 (Extensive Quantities): Extensive quantities are proportional to the amount of matter in a system.

Definition 1.12 (Intensive Quantities): Intensive quantities are independent of the amount of matter in a system.

Definition 1.13 (Conjugate Pair Variables): (T, P, μ) are intensive variables which form a conjugate pair with (S, V, N) , which are extensive variables, respectively.

$$\mu = \frac{\partial U}{\partial N}, \quad T = \frac{\partial U}{\partial S}, \quad P = -\frac{\partial U}{\partial V}$$

If one applies a constraint to the given system, then the corresponding conjugate variable will change as a response.

1.2 Zeroth Law and First Law

Law 1.1 (Zeroth Law of Thermodynamics): If two systems are separately in thermal equilibrium with a third, then they must also be in thermal equilibrium with each other.

Definition 1.14 (Temperature): Temperature, a state function, describes the tendency of a system to exchange energy with other system. An empirical state function whose existence is implied from the zeroth law.

Corollary 1.1: All systems in thermal contact have the same temperature when in thermal equilibrium and no energy is exchanged between them. Otherwise, energy is transferred from the hotter system to the colder system.

Definition 1.15 (Energy): Energy of the system is the sum of two contributions:

- For the macroscopic motion of the system, i.e. the kinetic energy of the motion of the centre of mass of the system, and the potential energy of the system due to the presence of external fields.
- Internal energy of the system: sum of the energy of all the internal degrees of freedom, which include the sum of the kinetic energy of the particles in the system in the reference frame in which the centre of mass is at rest, as well as, the potential energy arising from the interactions of the particles in the system.

The first contribution is ignored in thermal physics. The internal energy is an extensive state function and depends only on the equilibrium state of the system, described by a set of state functions.

Definition 1.16 (Ideal Gas): Ideal gas is a collection of non-interacting point-like particles with equation of state $PV = Nk_B T = nRT$, where $k_B = \frac{R}{N_A}$, where N_A is the Avogadro number. The internal energy of an ideal gas is only dependent on its temperature.

Definition 1.17 (Heat): Heat is the energy flow between a system and its surroundings due to a temperature difference across the wall (which partitions the system from its surroundings), and a finite thermal conductivity of the wall.

Definition 1.18 (Work): Work is any other kind of energy across the wall of the system.

Corollary 1.2: Both heat and work are not state functions and depend on the nature of the process transferring energy to the system.

Law 1.2 (First Law of Thermodynamics): Energy, i.e. internal energy U , is conserved and is transferred between systems as either heat Q or work W

$$dU = \delta Q + \delta W$$

where dU is an exact differential but δQ and δW are inexact differentials. They are conventionally positive if energy flows into the system.

Definition 1.19 (Reversible Process): A process is reversible if it is possible to restore the system and its surroundings to their original conditions.

Definition 1.20 (Quasi-Static Process): A process is quasi-static if it is sufficiently slow such that any intermediate state can be considered as an equilibrium state. They can be specified by a mathematical curve \mathcal{C} as a function of state functions.

Example 1.1. *Examples of quasi-static processes:*

- *Isobaric: constant pressure*
- *Isochoric: constant volume*
- *Isothermal: constant temperature*
- *Adiabatic: zero heat transfer, especially when system is thermally isolated from surroundings.*

Definition 1.21 (Quasi-Static Work): Quasi-static work can be regarded as a product of generalized displacement and their conjugate generalized forces. The former is usually extensive while the latter are intensive.

Example 1.2. For generalized systems, the generalized force and displacement respectively are:

- *Film: Surface Tension and Area*
- *Magnet: Magnetic Field and Magnetization*
- *Dielectric: Electric Field and Polarization*
- *Chemical Reaction: Chemical Potential and Particle Number*

Definition 1.22 (Response Functions): Response functions are used to characterize the macroscopic behaviour of a system and are experimentally measured from the changes of thermodynamic state functions with external probes.

Definition 1.23 (Heat Capacity): Heat capacity is an extensive quantity which describes the change in the temperature of the system resulting from a transfer of heat, i.e. $C = \frac{\partial Q}{\partial T}$. The nature of heat process is to be specified. To make this quantity intensive, we divide it by number of moles or by mass.

Definition 1.24 (Thermal Responses): Thermal responses probe the change in the thermodynamic state function with temperature.

Definition 1.25 (Expansivity): Expansivity measures the fractional change in volume with a change in temperature.

Definition 1.26 (Force Constants or Generalized Susceptibility): Force constants probe the change in the thermodynamic state function with force.

Definition 1.27 (Compressibility): Compressibility is the fractional change in volume when a force is applied, multiplied by a negative sign.

Example 1.3. For ideal gas, $PV = nRT$, we have expansivity $\beta_P = \frac{1}{V}(\frac{\partial V}{\partial T})_P = \frac{1}{T}$ and compressibility $\kappa_T = -\frac{1}{V}(\frac{\partial V}{\partial P})_T = \frac{1}{P}$.

Proposition 1.1: For an ideal gas, the heat capacities satisfy the relation

$$C_P - C_V = nR$$

Example 1.4 (Paramagnetic salt in an applied magnetic field). The work required to change the magnetic moment of a substance from \mathbf{m} to $\mathbf{m} + d\mathbf{m}$ in an applied external field \mathbf{H} conventionally regards only the work needed to change the state of the substance, returning any external sources of field back to zero. A paramagnetic substance with magnetic moment $\mathbf{m}(\mathbf{H})$ at some large distance from a permanent magnet will experience an attractive force

$$(\mathbf{m} \nabla) \mathbf{H}$$

Letting the substance approach the magnet until it experiences a field \mathbf{H}_1 , we can extract work, so that the work done on the substance is

$$W' = -\mu_0 \int_0^{H_1} \mathbf{m}(\mathbf{H}) d\mathbf{H}$$

However, the magnet is still near the substance, so in order to calculate the work required to change the state of the substance alone, we need to pull the magnet away to the previous separation, while keeping the magnetic moment \mathbf{m}_1 constant, i.e.

$$W_1 = \mu_0 \int_0^{H_1} \mathbf{m}_1 dH = \mu_0 m_1 H_1$$

This gives a net work $W = \mu_0 \int \mathbf{H} d\mathbf{m}$.

The difference between the two expressions is the change in the potential energy of the dipole moment in an applied field, $-\mu_0 d(mH)$, which is conventionally not counted as part of the internal energy.

1.3 van der Waals Gases

Definition 1.28 (van der Waals Gas): The equation of state for a van der Waals gas is

$$\left(P - \frac{a}{V_m^2}\right)(V_m - b) = RT$$

where V_m is the molar volume, a parametrizes the strength of the inter-molecular interactions while b is the volume excluded owing to the finite size of the molecules.

Note 1.1 (Derive equation of state). Assuming n moles of gas in volume V , the number of nearest neighbours is proportional to n/V , and so the attractive intermolecular interactions lower the total potential energy by an amount proportional to the number of atoms multiplied by the number of nearest neighbours. Changing V , the energy changes by an amount an^2dV/V^2 . This energy change can be thought of as being due to an effective pressure P_{eff} , i.e. $-P_{\text{eff}}dV$, hence $P_{\text{eff}} = -\frac{a}{V_m^2}$. The pressure that you measure is the sum of the ideal pressure and effective pressure, so $P_{\text{ideal}} = P + \frac{a}{V_m^2}$.

We can estimate b by simple hard sphere model. Each molecule excludes a spherical volume $\frac{4}{3}\pi(2r)^3$ and for one mole, $b = \frac{1}{2}N_A\frac{4}{3}\pi(2r)^3 = \frac{16}{3}N_A\pi r^3$. The factor 0.5 arises because the key situation is when a collision occurs - so one considers how much space is 'off limits' for a pair of atoms.

Note 1.2 (van der Waals' Isotherms). Starting at the low density side (i.e. high volume v per particle), we find behaviour approaching that of an ideal gas. As the density increases, the inter-molecular attraction term N^2a/V^2 causes the isotherm to fall below the ideal gas pressure, until we start to approach the density of closest packing, at which point the hard core repulsion term Nb causes the pressure to skyrocket. On some isotherms there is a very odd region where $\partial p/\partial V$ is positive, a situation which is mechanically unstable.

The isotherms can be obtained via the cubic equation

$$PV^3 - (Pb + RT)V^2 + aV - ab = 0$$

At high T , $PV_m = RT$ is a good approximation. C is a point of inflection, denoting the critical temperature, and the point is on the critical isotherm. For $T < T_c$, the isotherms exhibit a S-shape curve with a minimum and a maximum. There is a region in which the isothermal compressibility $\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T$ is negative, i.e. compressing the gas increases the volume.

The dashed curve with maximum at critical point, represents phase equilibrium between vapor and liquid. The region bounded by the curve is a 2-phase, for small V on the left outside the region is a liquid and for large V on the right, it is a vapor.

Proposition 1.2 (Critical Point): At the critical point, $\frac{P_c V_c}{RT_c} = \frac{3}{8}$.

Proof. Since $P = \frac{RT}{V-b} - \frac{a}{V^2}$ and at the critical point, we require both

$$\left(\frac{\partial P}{\partial V}\right)_T = \frac{-RT}{(V-b)^2} + \frac{2a}{V^3} = 0, \quad \left(\frac{\partial^2 P}{\partial V^2}\right)_T = \frac{2RT}{(V-b)^3} - \frac{6a}{V^4} = 0$$

to be true. This yields $V_c = 3b$ and hence $T_c = \frac{8a}{27Rb}$. The corresponding critical pressure is $P_c = \frac{a}{27b^2}$. The compressibility factor is $\frac{P_c V_c}{RT_c} = \frac{3}{8}$, independent of both a and b . Moreover, at this point, the isothermal compressibility diverges. \square

1.4 Entropy

Definition 1.29 (Heat Reservoirs): A heat reservoir is a thermodynamic system with a heat capacity so large that the temperature of the reservoir does not change when a reasonable amount of heat is added or extracted.

Definition 1.30 (Ideal Heat Engine): A heat engine is a cyclic process in a system that absorbs heat Q_{in} , rejects heat Q_{out} , and a positive amount of work $|W|$ is done by the engine. The heat transferred to the system in one cycle is $Q = -W$ since $\Delta U = 0$. The efficiency of the heat engine is

$$\eta = \frac{|W|}{Q_{\text{in}}} = 1 - \frac{|Q_{\text{out}}|}{|Q_{\text{in}}|}, \quad 0 \leq \eta \leq 1$$

Definition 1.31 (Refrigerator): A refrigerator is a heat engine that is run backwards so that you put work in and cause a heat flow from a cold reservoir to a hot reservoir. Another name is a heat pump.

Law 1.3 (Clausius Statement of Second Law of Thermodynamics): There is no process whose only effect is to accept heat from a colder reservoir and transfer it to a hotter reservoir.

Law 1.4 (Kelvin Statement of Second Law of Thermodynamics): There is no process whose only effect is to accept heat from a single heat reservoir and convert it entirely into work.

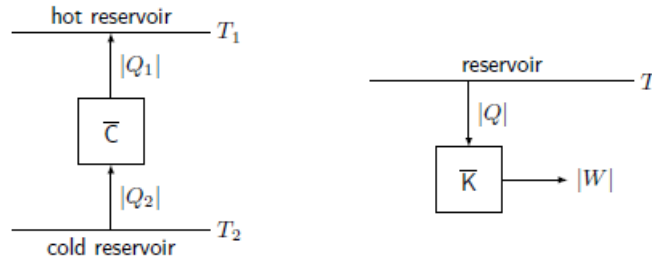


Figure 1: Clausius and Kelvin Engine. [1]

Theorem 1.1: The Kelvin and Clausius statements of the Second Law of Thermodynamics are equivalent.

Proof. We first show that if Clausius statement is not true, then it implies that Kelvin statement is not true. Suppose there exists a Clausius-violating device (accepts $|Q_2|$ heat from cold reservoir and dump the exact same amount to the hotter reservoir), that acts together with a heat engine (accepts heat $|Q_1| - |Q_2|$ from a single heat reservoir at temperature T_1 and converting it entirely into work $|W| = |Q_1| - |Q_2|$). This combined system violates Kelvin's statement since the heat accepted from the hotter reservoir is completely converted to work.

Conversely, suppose there exists a Kelvin-violating device (converting heat $|Q'_1|$ entirely into work $|W|$), that acts together with a refrigerator, accepts heat $|Q_2|$ from a colder reservoir at temperature T_2 and transfer heat $|Q_1|$ to a hotter reservoir at temperature T_1 . The combined system violates Clausius' statement since it takes heat $|Q_2|$ from a colder reservoir and dumps heat $|Q_2| = |Q_1| - |Q'_1| > 0$ onto a hotter reservoir. \square

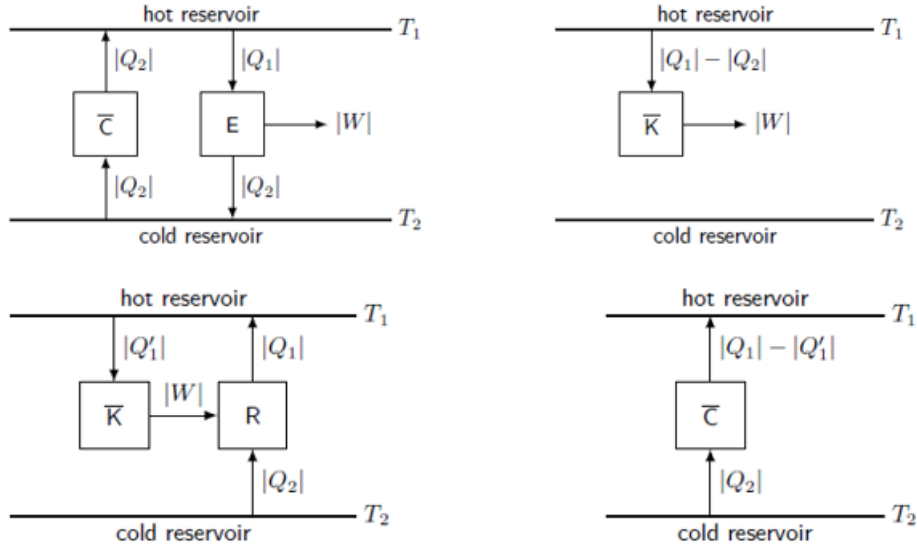


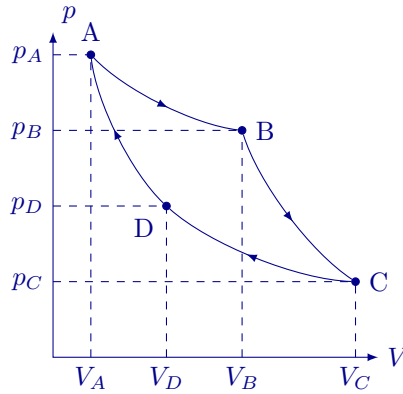
Figure 2: (Top) First part of proof. (Bottom) Second part of proof. [1]

Definition 1.32 (Carnot Engine): Carnot engine is an idealized system that is cyclic and reversible with all of its heat exchanges taking place at a hotter reservoir at temperature T_1 and a colder reservoir at temperature T_2 .

Definition 1.33 (Ideal Refrigerator): An ideal refrigerator is a Carnot cycle run in reverse.

Definition 1.34 (Carnot Cycle): Carnot Cycle consists of 4 processes:

1. A to B: isothermal expansion at T_1
2. B to C: adiabatic expansion to T_2
3. C to D: isothermal compression at T_2
4. D to A: adiabatic compression to T_1 .



Theorem 1.2: The energy efficiency of Carnot engine with ideal gas as a working substance is

$$\eta = 1 - \frac{T_2}{T_1}$$

Proof. Since A to B is an isothermal expansion, C to D is an isothermal compression, then we have

$$|Q_{in}| = Q_{A \rightarrow B} = nRT_1 \log(V_B/V_A)$$

$$|Q_{out}| = -Q_{C \rightarrow D} = -nRT_2 \log(V_D/V_C)$$

Then the efficiency is $\eta = 1 - \frac{|Q_{out}|}{|Q_{in}|} = 1 - \frac{T_2 \log(V_D/V_C)}{T_1 \log(V_B/V_A)}$. Since $TV^{\gamma-1}$ is a constant for adiabatic processes, then $V_C/V_B = V_D/V_A$ and so $\eta = 1 - \frac{T_2}{T_1}$. \square

Corollary 1.3: For Carnot engine, $\frac{|Q_{out}|}{|Q_{in}|} = \frac{T_2}{T_1}$

Theorem 1.3 (Carnot's Theorem): No engine operating between two reservoirs can be more efficient than a Carnot engine operating between the same two reservoirs.

Proof. Suppose we have a hypothetical super-efficient engine SE and Carnot refrigerator CR, such that $\eta_{SE} > \eta_{CR} \implies |Q_1| > |Q'_1|$ and hence $W = |Q'_1| - |Q'_2| = |Q_1| - |Q_2|$ and hence $|Q_1| - |Q'_1| > 0$ is the amount of heat dumped into the hotter reservoir while $|Q_2| - |Q'_2| > 0$ is the amount of heat extracted from the colder reservoir. This violates Clausius statement. \square

Corollary 1.4: All Carnot engines operating between the same two reservoirs have the same efficiency.

Proof. Prove this by setting up $\eta_1 \leq \eta_2$ and $\eta_2 \leq \eta_1$ and pay attention to Carnot's Theorem and Clausius' statement. \square

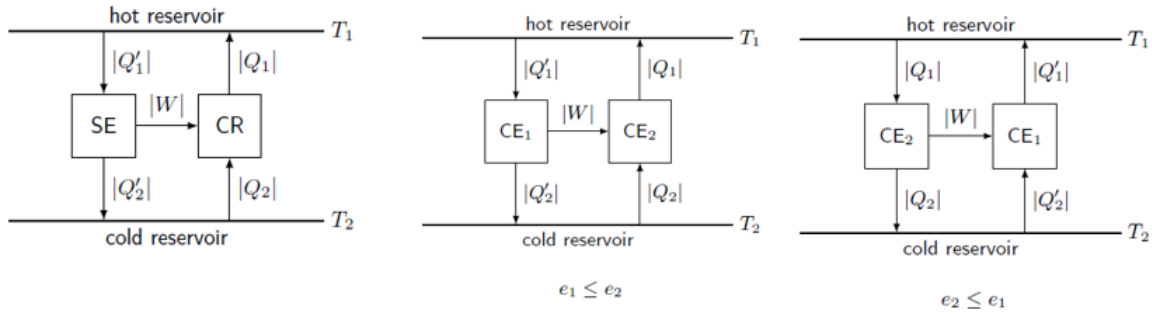


Figure 3: (left) Demonstrates Carnot's Theorem. (centre and right) for Corollary. [1]

Definition 1.35 (Empirical Temperature): Carnot's theorem means that the ratio of the heat transfers into and out of the hot and cold reservoirs is a function of the temperature of the reservoirs only. It is thus imperative to define an empirical temperature to be the ratio of heat transfers, i.e. $\frac{Q_2}{Q_1} = f(T_1, T_2)$ where Q_1 is heat supplied to the cold reservoir T_1 and Q_2 is heat extracted from the hot reservoir T_2 .

Theorem 1.4 (Clausius Inequality): For a general cycle, reversible or not, the following relation is true

$$\oint \frac{dQ}{T} \leq 0$$

Proof. Suppose heat δQ_i at each point of a general cycle is supplied via a Carnot engine, which is connected between a common (all Carnot engines on this cycle) reservoir at temperature T and the reservoir at temperature T_i . Each Carnot engine produces work δW_i and for a Carnot engine, we know from Carnot's theorem that $\frac{\delta Q_i}{T_i} = \frac{\delta Q_i + \delta W_i}{T}$. The total work extracted from the cycle is $\Delta W = \sum_{i \in \text{cycle}} \delta Q_i$. And from Kelvin's statement (we cannot extract work from this assembly), the total work produced per cycle is:

$$\Delta W + \sum_{i \in \text{cycle}} \delta W_i \leq 0 \implies \sum_{i \in \text{cycle}} \delta Q_i + \sum_{i \in \text{cycle}} \delta Q_i \left(\frac{T}{T_i} - 1 \right) \leq 0 \implies \sum_{i \in \text{cycle}} \frac{\delta Q_i}{T_i} \leq 0$$

where $T > 0$. Changing the sum to a closed loop path integral, we obtain the Clausius' inequality. \square

Corollary 1.5: For reversible cycle, $\oint \frac{dQ}{T} = 0$.

Definition 1.36 (Entropy): Entropy, denoted as S , is

$$dS = \frac{\delta q_{rev}}{T}$$

where δq_{rev} is an infinitesimal amount of heat absorbed in a reversible process.

Theorem 1.5 (Second Law of Thermodynamics): An equivalent statement of the Second Law of Thermodynamics states that for a thermally isolated system,

$$dS \geq 0$$

In particular, $dS = 0$ for a reversible process and $dS > 0$ for an irreversible process.

Proof. Consider a loop which contains an irreversible section A to B and a reversible section B to A. By Clausius' inequality, we have

$$dS = \frac{\delta q_{rev}}{T} \geq \frac{\delta q}{T}$$

For a thermally isolated system, we have $\delta q = 0$. So, $dS \geq 0$ in general. \square

Definition 1.37 (Entropy Change of Surroundings and System): Since the surroundings (i.e. a heat reservoir) is much larger than the system, the temperature of the surroundings does not change, despite the heat transfer. This heat transfer, as seen from the point of view of the surroundings, is reversible, hence $\Delta q_{surr} = \Delta q_{surr,rev}$. Moreover, since energy is conserved, $\Delta q_{surr} = -\Delta q_{sys}$

$$\Delta S_{surr} := \frac{\Delta q_{surr,rev}}{T_{surr}} = \frac{\Delta q_{surr}}{T_{surr}} = \frac{-\Delta q_{sys}}{T_{surr}}$$

As for the system, from the definition of entropy:

$$\Delta S_{sys} = \frac{q_{rev,sys}}{T}$$

This is true for both reversible and irreversible process.

Definition 1.38 (Equilibrium): A system is said to be in equilibrium if $dS_{univ} = 0$.

Definition 1.39 (Spontaneity): A spontaneous process occurs when the resultant change in entropy of the Universe is positive, i.e. $dS_{univ} > 0$. Note that

$$\Delta S_{univ} = \Delta S_{surr} + \Delta S_{sys}$$

Note 1.3 (Reversible Expansion Versus Irreversible Expansion). In general, expansion occurs when $p_{internal} > p_{external}$. Whenever there is a constant external pressure imposed on the system of gas, the expansion is said to be irreversible. It is also known as Joule expansion. It is spontaneous and system will tend towards mechanical equilibrium where $p_{internal} = p_{external}$.

In contrast, for a reversible expansion, the external pressure is infinitesimally greater than the internal pressure. Any infinitesimal change in the external pressure will thus change the direction of the process. Such a process is infinitely slow since it is at equilibrium.

Definition 1.40 (Entropy): We can also define the entropy of a system, $S = S(E)$, to be

$$S(E) = \log(\Omega(E))$$

where $\Omega(E)$ is the number of states with energy E .

Definition 1.41 (Thermodynamic Temperature): The thermodynamic temperature T of a system with entropy function $S(E)$ is

$$\frac{1}{T} := \frac{dS(E)}{dE}$$

Note 1.4 (Maximal entropy at equilibrium). Consider an isolated system, partitioned into systems 1 and 2 which are in thermal contact with each other. The subsystems have energies E_1 and E_2 respectively. The total energy is $E = E_1 + E_2$, which is conserved even after the isolated systems come into contact. The number of available states is

$$\Omega(E_1, E_2) = \Omega_1(E_1)\Omega_2(E_2) = e^{S_1(E_1)}e^{S_2(E_2)} \implies S(E_1, E_2) = S_1(E_1) + S_2(E_2) = S_1(E_1) + S_2(E - E_1)$$

so thermal contact allows E_1 to take on any value that state 1 allows, i.e. E_1^i for all possible states labelled by i . The number of available states given fixed energy E is

$$\Omega(E) = \sum_i \exp(S_1(E_1^i) + S_2(E - E_1^i))$$

which will have an extremely high and narrow maximum at a particular energy value, say E_1 . Thus, the sum dominates at this value E_1 , thus maximizing the entropy S .

$$\frac{dS}{dE_1} = \frac{dS_1}{dE_1} + \frac{dS_2}{dE_2} \frac{dE_2}{dE_1} = 0$$

with $E_2 = E - E_1$, so $\frac{dS_1}{dE_1} - \frac{dS_2}{dE_2} = \frac{1}{T_1} - \frac{1}{T_2} = 0$ at equilibrium energy E_1 . Thus, $T_1 = T_2$. The condition for the maximum of S is

$$\frac{d^2S}{dE_1^2} = \frac{d^2}{dE_1^2}(S_1(E_1) + S_2(E - E_1)) = \frac{d}{dE_1} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) = -\frac{1}{T_1^2} \frac{dT_1}{dE_1} - \frac{1}{T_2^2} \frac{dT_2}{dE_2}$$

where $E_2 = E - E_1$. At equilibrium, $T_2 = T_1$, so

$$\frac{d^2S}{dE_1^2} = -\frac{1}{T_1^2} \left(\frac{dT_1}{dE_1} + \frac{dT_2}{dE_2} \right)$$

For maximum S , we require

$$\frac{dT_1}{dE_1} + \frac{dT_2}{dE_2} > 0$$

This will work if $\frac{dT}{dE} > 0$ for every system, i.e. energy increases with temperature. In other words, the slope of $S(E)$ decreases with energy E .

1.5 Chemical Potential

Definition 1.42 (Chemical Potential): The chemical potential μ is the reversible rate of change of internal energy caused by changing mole number n .

Theorem 1.6: Gradient in chemical potential influences particle exchange.

Proof. We have $dU = TdS - pdV + \mu dN \implies dS = \frac{dU}{T} + \frac{pdV}{T} - \frac{\mu dN}{T}$. Writing $dS = (\frac{\partial S}{\partial U})_{N,V} dU + (\frac{\partial S}{\partial V})_{N,U} dV + (\frac{\partial S}{\partial N})_{U,V} dN$ and by comparison, we get $(\frac{\partial S}{\partial N})_{U,V} = -\mu/T$. Consider two systems which are able to exchange particles with each other, but remain isolated from their surroundings. If system 1 loses dN particles, system 2 must gain dN particles and so $dS = (\frac{\mu_1}{T_1} - \frac{\mu_2}{T_2})dN \geq 0$. Assuming $T_1 = T_2$, we find $dN > 0$ when $\mu_1 > \mu_2$. \square

Note 1.5 (Internal Energy as Extensive Function). Since internal energy is extensive, i.e. depends linearly on the size of the system, then $U(\lambda S, \lambda X)$, where X are all other possible extensive variables (could be generalized). Then, $U(\lambda S, \lambda X) = \lambda U(S, X)$. Thus, $U(S, X)$ is a first-order homogeneous function of S and X .

Theorem 1.7 (Gibbs-Duhem Equation): The Gibbs-Duhem equation, in the energy representation, is

$$0 = SdT + \sum_{i=1}^r x_i dF_i + \sum_{i=1}^r N_i d\mu_i$$

where only $r + s$ variables are independent. It states that adjacent equilibrium states cannot differ only in one of these independent variables.

Proof. We invoke Euler's theorem for first-order homogeneous functions which are $f(\{u_i\}) = \lambda f(\{x_i\})$ where $u_i = \lambda x_i \forall i \in [1, n]$. Then, $(\frac{\partial f(\{u_i\})}{\partial \lambda})_{x_i} = f(\{x_i\})$ but $df(\{u_i\}) = \sum_{s=1}^n (\frac{\partial f}{\partial u_s})_{u_k} du_s$ and so

$$\left(\frac{\partial f}{\partial \lambda} \right)_{x_i} = \sum_{i=1}^n \left(\frac{\partial f}{\partial u_i} \right)_{u_i} \left(\frac{\partial u_i}{\partial \lambda} \right)_{x_i} = \sum_{i=1}^n \left(\frac{\partial f}{\partial u_i} \right)_{u_j} x_i$$

This implies $f(\{x_i\}) = \sum_{i=1}^n (\frac{\partial f}{\partial u_i})_{u_j} x_i$ is true for all λ . Take $\lambda = 1$, then $f(\{x_i\}) = \sum_{i=1}^n (\frac{\partial f}{\partial x_i})_{x_j} x_i$. Since $U(S, X)$ is first-order homogeneous, by Euler's theorem, we have $U = TS + Fx + \mu N \implies dU = TdS + SdT + \sum_{i=1}^r (x_i dF_i + F_i dx_i) + \sum_{i=1}^r (\mu_i dN_i + N_i d\mu_i)$ but $dU = TdS + \sum_{i=1}^r F_i dx_i + \sum_{i=1}^r \mu_i dN_i$ and so we recover the desired equation. \square

Corollary 1.6:

$$U = TS - PV + \mu N$$

Proof. Follows from Euler's theorem. □

Proposition 1.3: For a reversible chemical reaction, we must have $\sum_i \nu_i \mu_i = 0$.

Proof. By Gibbs-Duhem equation, we have $d\mu_i = -s_i dT + v_i dp$. The change in chemical potential is thus zero at constant T and P . The variation of G about the equilibrium ratio of chemical components satisfies

$$dG = \sum_i \mu_i dN_i = 0$$

But the dN_i 's are related by matter conservation, so dN_i/ν_i is a constant. We then have

$$\sum_i \nu_i \mu_i = 0$$

where ν_i 's are the ratios, with $\nu > 0$ for reactants and $\nu < 0$ for products. □

Definition 1.43 (Equilibrium constant for concentration):

$$K_c(T) := \prod_i c_i^{\nu_i}$$

Proposition 1.4:

$$\ln K_c(T) = -\frac{1}{k_B T} \sum_i \nu_i \mu_{0,i}(T, P)$$

Proof. Combining $\sum_i \nu_i \mu_i = 0$ and $\mu_i(P_i, T) = \mu_{0,i}(P, T) + k_B T \ln c_i$, we have

$$0 = \sum_i \nu_i \mu_{0,i} + k_B T \sum_i \nu_i \ln c_i = \sum_i \nu_i \mu_{0,i} + k_B T \ln \prod_i c_i^{\nu_i}$$

Hence, $\ln K_c(T) = -\frac{1}{k_B T} \sum_i \nu_i \mu_{0,i}(T, P)$. □

1.6 Potentials and Availability

Availability

Proposition 1.5: Maximizing the total entropy of system and reservoir R with respect to the state of the system is equivalent to minimizing the availability of the system.

Proof. From the principle of maximum entropy, we have

$$0 \leq dS_{\text{tot}} = dS + dS_R = dS + dS_R = dS + \frac{dU_R + P_R dV_R - \mu_R dN_R}{T_R}$$

Maximize S_{tot} is thus equivalent to maximizing the quantity

$$\int (T_R dS - dU - P_R dV + \mu_R dN) = \int (d(T_R S) - dU - d(P_R V) + d(\mu_R N)) = T_R S - U - P_R V + \mu_R N$$

We denote this quantity to be the availability $-A$. The reason for the minus sign is apparent later. \square

Remarks 1.1. The change in availability is written in terms of its natural variables S , V and N :

$$dA = dU - T_R dS + P_R dV - \mu_R dN = (T - T_R) dS - (P - P_R) dV + (\mu - \mu_R) dN$$

Proposition 1.6: The availability is the maximum useful work possible during a process that brings the system into equilibrium with a reservoir R , upon contact.

Proof. The total work extracted is

$$\begin{aligned} -dW &= dU + dU_R \\ &= d(A + T_R S - P_R V + \mu_R N) + T_R dS_R - P_R dV_R + \mu_R dN_R \\ &= dA + T_R dS + T_R dS_R - P_R dV + P_R dV + \mu_R dN - \mu_R dN \\ &= dA + T_R d(S + S_R) \end{aligned}$$

But since $d(S + S_R) \geq 0$, we have $-dW \geq dA$. The maximum work is attained when $S + S_R$ stays constant. \square

Corollary 1.7: The availability of a system in equilibrium with a reservoir is always zero.

Proof. The availability is defined to be the extractable work on attaining equilibrium. This is zero if the system is already in equilibrium. \square

Note 1.6 (Free energy). *Reversibly compress a monatomic ideal gas that is connected to a reservoir. The energy input is in the form of work. The temperature, and hence internal energy is constant. The work flows directly out of the system into the reservoir as heat. If we reverse this reversibly, we recover the energy back. Energy conservation needs to be applied to both the system and the reservoir. The free energy of an isothermal system is the quantity which corresponds to energy conservation in the system and the reservoir.*

Legendre transform

Definition 1.44 (Legendre Transformation): A Legendre transform converts from a function of one set of variables to another function of a conjugate set of variables.

Theorem 1.8: The Legendre transform of $f = f(x_1, \dots, x_n)$ is $g = f - \sum_{i=r+1}^n u_i x_i$, where $u_i = (\frac{\partial f}{\partial x_i})_{x_j}$, such that $g = g(x_1, \dots, x_r, u_{r+1}, \dots, u_n)$ where u_{r+1}, \dots, u_n are conjugate variables to x_{r+1}, \dots, x_n .

Proof. We evaluate dg

$$dg = df - \sum_{i=r+1}^n (u_i dx_i + x_i du_i) = \sum_{i=1}^r u_i dx_i + \sum_{i=r+1}^n (-x_i) du_i$$

g is indeed a function of x_i for $i \in [1, r]$ and u_i for $i \in [r+1, n]$. \square

Thermodynamic potentials

Definition 1.45 (Thermodynamic potentials): Thermodynamic potentials are generated by Legendre transforms. The variables which are differentiated with respect to are the ‘natural variables’ of the corresponding thermodynamic potential.

Remarks 1.2.

1. A thermodynamic potential is minimized in equilibrium under the conditions that its natural variables are held fixed.
2. If one knows any of the thermodynamic potentials as a function of its natural variables, one has complete thermodynamic information about the equilibrium state.
3. The appropriate thermodynamic potential is the property of the system that reflects global energy conservation.

Example 1.5. Since internal energy is the sum of δq and δW , then by the first law of thermodynamics,

$$dU = TdS - PdV$$

where S and V are natural variables of U , i.e. $U = U(S, V) = TS - PV$. We correspondingly define three more thermodynamic potentials: enthalpy H , Gibbs function G and Helmholtz function F :

$$H = U + PV \implies dH = dU + d(PV) = TdS - PdV + PdV + VdP = TdS + VdP$$

$$G = H - TS \implies dG = dH - d(TS) = TdS + VdP - TdS - SdT = -SdT + VdP$$

$$F = U - TS \implies dF = dU - d(TS) = TdS - PdV - TdS - SdT = -SdT - PdV$$

The natural variables are thus $H(S, p)$, $F(T, V)$, $G(P, T)$. Note that here we take N to be fixed. Otherwise, we have $F(T, V, N)$, $G(P, T, N)$, $H(S, P, N)$ and $U(S, V, N)$. Physically, they mean:

- Internal Energy U : energy needed to create a system. To attain equilibrium at constant V , S and N , we minimize U , i.e. $dA = dU$.
- Enthalpy H : energy needed to create a system plus the work needed to make room for it. To attain equilibrium at constant P , S and N , we minimize H , i.e. $dA = dU + PdV = d(U + PV) = dH$.
- Gibbs Free Energy G : energy needed to create a system and make room for it minus the energy you can get from the environment. To attain equilibrium at constant T , P and N , we minimize G , i.e. $dA = dU - TdS + PdV = d(U - TS + PV) = dG$.
- Helmholtz Free Energy F : energy needed to create a system minus the energy you can get from the environment. To attain equilibrium at constant T , V and N , we minimize F , i.e. $dA = dU - TdS = d(U - TS) = dF$.

One last possible thermodynamic potential is the grand potential $\Phi(T, V, \mu) = F - \mu N$. To attain equilibrium at constant T , V and μ , we minimize Φ .

Remarks 1.3. For G , the only extensive natural variable is N , so from Euler’s Theorem, we can write $G = U - TS + PV = \mu N$. Similarly, since the only extensive natural variable for Φ is V , we can write $\Phi = U - TS - \mu N = -PV$.

Corollary 1.8: The thermodynamic potentials can indeed be inter-converted from Legendre Transformation.

Proof. Since $T = (\frac{\partial U}{\partial S})_V$ and $P = -(\frac{\partial U}{\partial V})_S$, then (T, S) and $(P, -V)$ are conjugate variables pair. To construct F from U , we subtract the quantity S times the variable conjugate to S , which is T , i.e. $F = U - TS$. Similarly, to construct H from U , we subtract the quantity V with its conjugate, which is $-P$, i.e. $H = U - (-PV)$. Lastly, to construct G from U , we first convert it to F and then deduct $-PV$. \square

Theorem 1.9 (Maxwell's Relations):

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V, \quad \left(\frac{\partial T}{\partial p}\right)_S = -\left(\frac{\partial V}{\partial S}\right)_p, \quad \left(\frac{\partial p}{\partial T}\right)_V = -\left(\frac{\partial S}{\partial V}\right)_T, \quad \left(\frac{\partial V}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial p}\right)_T$$

Proof. For all four, we use Clairaut's theorem (symmetry of second partial derivatives). The first, second, third and fourth were obtained from U , H , F and G respectively.

Alternatively, one may use Jacobian. Consider a cyclic process that can be described in both the T-S and p-V plane. The internal energy U is a state function and therefore doesn't change in a cycle and since $\oint dU = 0$, we have $\oint pdV = \oint TdS$ and so $\frac{\partial(T,S)}{\partial(p,V)} = 1 \implies \frac{\partial(T,S)}{\partial(x,y)} = \frac{\partial(p,V)}{\partial(x,y)}$ where (x,y) can be taken as (T,p) , (T,V) , (p,S) and (S,V) respectively. \square

Remarks 1.4. In general, there are more than 4 Maxwell's relations. Given a particular thermodynamic potential, expressed in terms of its $(1+r+s)$ natural variables, there are $(1+r+s)(r+s)/2$ distinct pairs of mixed second derivatives yielding $(1+r+s)(r+s)/2$ Maxwell's relations.

Example 1.6. dU gives the heat transfer δq to a system at constant V while dH gives the heat transfer δq to a system at constant p , i.e. energetics of a chemical reaction.

Theorem 1.10 (Gibbs-Helmholtz Equations):

$$U = -T^2 \left(\frac{\partial F}{\partial T} \right)_V, \quad H = -T^2 \left(\frac{\partial G}{\partial T} \right)_p$$

Proof. Using the expressions $S = -(\frac{\partial F}{\partial T})_V$ and $S = -(\frac{\partial G}{\partial T})_p$, we get

$$U = F + TS = F - T \left(\frac{\partial F}{\partial T} \right)_V = -T^2 \left(\frac{\partial F}{\partial T} \right)_V, \quad H = G + TS = G - T \left(\frac{\partial G}{\partial T} \right)_p = -T^2 \left(\frac{\partial G}{\partial T} \right)_p$$

\square

Corollary 1.9: The following statements are true:

- At constant entropy, i.e. reversible adiabatic change, the work done by a system is equal to dU .
- At constant temperature, the maximum amount of energy which can be converted to work is given by $-dF$.
- At constant temperature and pressure, the maximum amount of energy which can be converted to non- pV work is given by dG .

Proof. First one is trivial, since $dU = -pdV$. For second one, since

$$dF = \delta q_{\text{sys}} + \delta W - TdS - SdT$$

and the $\delta q_{\text{sys}} = -\delta q_{\text{res}}$, then at constant temperature (same as that of the reservoir),

$$dF = \delta W - T(dS_{\text{res}} + dS) = \delta W - TdS_{\text{total}} \implies -\delta W = -dF - TdS_{\text{total}} \leq -dF$$

The maximum amount of energy which can be converted to work, occurs when $dS_{\text{total}} = 0$. For third one, since at constant temperature and pressure,

$$dG = \delta q_{\text{sys}} + \delta W + pdV - TdS$$

when the change is reversible, $\delta q_{\text{sys}} = TdS$. Furthermore, $\delta W = -pdV + \delta W_{\text{non}pV}$ and so $\delta G = \delta W_{\text{non}pV}$ and this maximum conversion occurs in a reversible process. \square

Example 1.7. G is extensively used by chemists since it involves constant temperature and pressure. But yet F is extensively used by physicists since in statistical mechanics, F is easier to evaluate. Both allow us to find the entropy change of the surroundings in terms of the properties of the system.

1.7 Phase Transitions

Definition 1.46 (Phase transition): A phase of a thermodynamic system and the states of matter have uniform physical properties. During a phase transition of a given medium, certain properties of the medium change, often discontinuously, as a result of the change of external conditions, such as temperature, pressure, or others.

Remarks 1.5.

1. At a given temperature and pressure, the stable phase will be the one with lower Gibbs free energy.
2. The condition for phase equilibrium: the chemical potential μ of both phases are equal.
3. Recall Gibbs-Duhem relation $d\mu = -sdT + vdP$ where s and v are the entropy and volume per particle respectively. At high temperatures, the phase with highest s have the smallest μ , so more stable since $(\frac{\partial\mu}{\partial T})_P = -\frac{S}{N} = s$. At high pressures, the phase with smallest v have the smallest μ , so more stable since $(\frac{\partial\mu}{\partial P})_T = \frac{V}{N} = v$.

Theorem 1.11 (Clausius-Clapeyron relation): The Clausius-Clapeyron relation is a way of characterizing a discontinuous phase transition (involving a latent heat, with no change in temperature) between two phases of matter of a single constituent. The gradient of the phase boundary on the P - T plane is

$$\frac{dP}{dT} = \frac{L}{T(V_2 - V_1)}$$

Proof. Suppose N_1 particles of phase 1 are in equilibrium with N_2 particles of phase 2. The total Gibbs free energy is $G_{tot} = N_1\mu_1 + N_2\mu_2$. In phase equilibrium, $dG_{tot} = 0 \implies \mu_1 = \mu_2$. At a given temperature and pressure, each coexisting phase has the same chemical potential. We can thus characterize the line of coexistence of two phases to be $\mu_1(P, T) = \mu_2(P, T)$. Along this phase boundary, we have $\mu_1(P+dP, T+dT) = \mu_2(P+dP, T+dT) \implies d\mu_1 = d\mu_2 \implies -s_1dT + v_1dP = -s_2dT + v_2dP$, where s and v are entropy and volume per particle respectively. We then have

$$\frac{dP}{dT} = \frac{(S_2 - S_1)}{(V_2 - V_1)} = \frac{L}{T(V_2 - V_1)}$$

where $L = T(S_2 - S_1)$ is the latent heat evolved from the phase transition. \square

Definition 1.47 (Triple Point): In thermodynamics, the triple point of a substance is the temperature and pressure at which the three phases (gas, liquid, and solid) of that substance coexist in thermodynamic equilibrium.

Definition 1.48 (Critical Point): In thermodynamics, a critical point is the end point of a phase equilibrium curve. At the critical point, defined by a critical temperature T_c and a critical pressure P_c , phase boundaries vanish. One example is the Curie temperature for paramagnets, where $T > T_c$, there is no clear boundary between magnetized up and magnetized down.

Remarks 1.6. It is common to describe various equilibrium states by P - V diagram, where one follows curves of constant entropy (adiabatic changes) or curves of constant temperature (isothermal changes) as they correspond to the system adiabatically isolated or in contact with a heat bath respectively. A phase transition appears as a jump in an isothermal curve. For the isothermal curve of the molar free energy, the slope of one phase is equal the slope of the other (joined by a common straight tangential line).

Note 1.7 (Phase Transition Curves). A phase transition appears as a jump in an isothermal curve. $(T, P(T))$ is a point on the coexistence line, such that $P(T)$ is a solution to the Clausius-Clapeyron equation. $v^{(1)}(T)$ and $v^{(2)}(T)$ are the molar volumes of the two phases with $\Delta v(T)$ being the difference. Where the isothermal curve is continuous within one phase, it has a negative slope since $(\frac{\partial P}{\partial V})_T < 0$ in stable equilibrium. Since $v^{(2)} < v^{(1)}$, the phase 1 is gaseous while phase 2 is liquid. Pressure is derived from free energy, i.e. $P = -\frac{\partial F}{\partial V}$. So

$$F(T, V, n) = nF\left(T, \frac{V}{n}, 1\right) = nf(T, V) \implies P = -\left(\frac{\partial f}{\partial v}\right)_T = P(T, V)$$

We have $-(\frac{\partial f^{(1)}}{\partial v})_T = -(\frac{\partial f^{(2)}}{\partial v})_T$ holds for $(T, P(T))$ on the coexistence line for phases 1 and 2. On this line, we also have equal chemical potentials, i.e. $\mu^{(1)}(T, v^{(1)}) = \mu^{(2)}(T, v^{(2)})$. The corresponding isothermal curve for molar free energy f , have equal slopes at the points $(v^{(2)}, f^{(2)}(T, v^{(2)}))$ and $(v^{(1)}, f^{(1)}(T, v^{(1)}))$. These two points can be connected by a common straight tangential line

$$f^{(2)}(T, v^{(2)}) - f^{(1)}(T, v^{(1)}) = -P(T)[v^{(2)} - v^{(1)}] \implies (f + Pv)^{(1)} = (f + Pv)^{(2)}$$

consistent with $\mu^{(1)} = \mu^{(2)}$. We note that the curvature of the isothermal curve has the wrong sign in some range, i.e. $(\frac{\partial^2 f}{\partial v^2})_T < 0$ or $(\frac{\partial P}{\partial v})_T > 0$. The requirement for $\mu^{(1)} = \mu^{(2)}$ gives rise to the Maxwell construction. Since $[v^{(1)} - v^{(2)}]P(T) = f(T, v^{(2)}) - f(T, v^{(1)})$, we have the two shaded areas above and below the coexistence line respectively to be the same size.

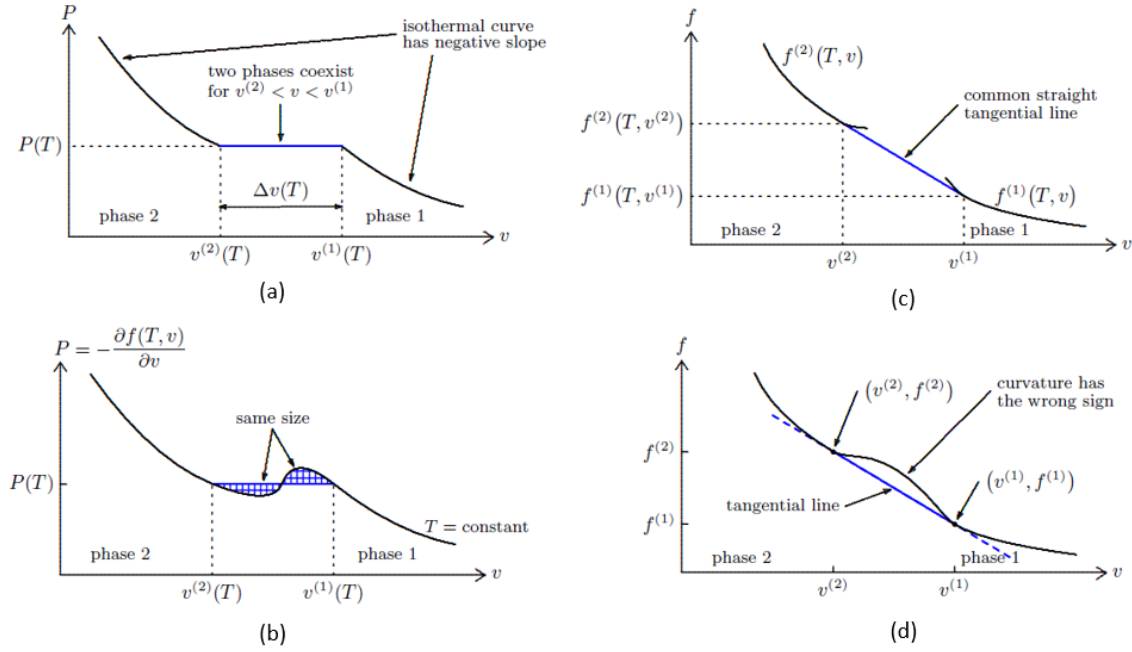


Figure 4: Phase Transition Curves: (a-b) P-v Curve; (c-d) f-v curve

Definition 1.49 (Metastable Phases): It is possible to warm a liquid above boiling point to form metastable superheated liquid. Similarly, it is possible to cool a vapor below boiling point to form metastable supercooled vapor.

Example 1.8. Consider a liquid with p_{liq} in equilibrium with a vapor at pressure p , then $\mu_{liq} = \mu_{vap}$. Suppose $p_{liq} \rightarrow p_{liq} + dp_{liq}$, then $p \rightarrow p + dp$ in order to stay in equilibrium.

$$\left(\frac{\partial \mu_{liq}}{\partial p_{liq}}\right)_T dp_{liq} = \left(\frac{\partial \mu_{vap}}{\partial p}\right)_T dp \implies v_{liq} dp_{liq} = v_{vap} dp \implies p = p_0 e^{v_{liq} \Delta p_{liq} / RT}$$

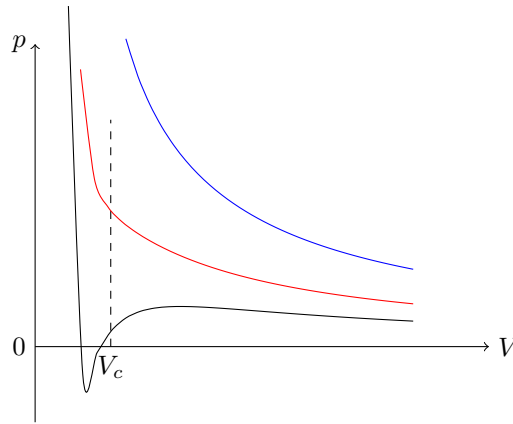
where $\Delta p_{liq} = 2\gamma/r$ (surface tension). This is Kelvin's formula. So small droplets have very high vapor pressure. With the very high pressure, they can evaporate instead of growing. This high pressure stabilizes the vapour, even though it is the thermodynamically unstable phase. The thermodynamic driving force to condense is overcome by the tendency to evaporate.

Similar effect occurs for superheated liquids. As you boil the liquid, any bubble of vapor which does form tends to collapse. Liquid can be kinetically stable above its boiling point, the only bubbles which then do survive are very large ones, and this causes violent bumping. This can be avoided by boiling with small pieces of glasses, providing plenty of nucleation centres for small bubbles to form.

Note 1.8 (Phase Change in van der Waals gas). The Helmholtz function is obtained from $p = -(\frac{\partial F}{\partial V})_T$ and so $F = -RT \log(V - b) - \frac{a}{V} + f(T)$ where $f(T)$ is a function of temperature. The Gibbs function is $G = F + pV = f(T) - RT \log(V - b) - \frac{a}{V} + pV$. G is plotted for $T = 0.9T_c$ (Fig 5) and it was discovered that the Gibbs function becomes multiply valued for certain values of pressure. Since a system held at constant temperature and pressure will minimize G , the system will normally ignore the upper loop, i.e. path BXYB, and proceed from A to B to C as the pressure is reduced.

Moreover, we see that the two points B_1 and B_2 on the curve representing the volume correspond to the single point B on the curve representing G . Since G is the same for these two points, phases corresponding to these two points can be in equilibrium with each other. Liquid is stable in the region AB and gas is stable in the region BC. The line BX represents metastable superheated liquid while the line BY represents metastable supercooled gas. We note that the triangle BXY vanishes for $T > T_c$. When $T > T_c$, the sharp distinction between liquid and gas is lost.

The van der Waals isotherms are as follow: with red ($T = T_c$), blue ($T > T_c$) and black ($T < T_c$).



Theorem 1.12 (Maxwell Construction): Phase existence occurs between two points, when the two areas bounded by the isotherm are equal.

Proof. We have $(\frac{\partial G}{\partial p})_T = V$ and so $G(p_1, T) = G(p_2, T) + \int_{p_2}^{p_1} V dp$. Since $G(p_{B_1}, T) = G(p_{B_2}, T)$, we have $\int_{p_{B_1}}^{p_{B_2}} V dp = 0$. This represents the shaded areas. \square

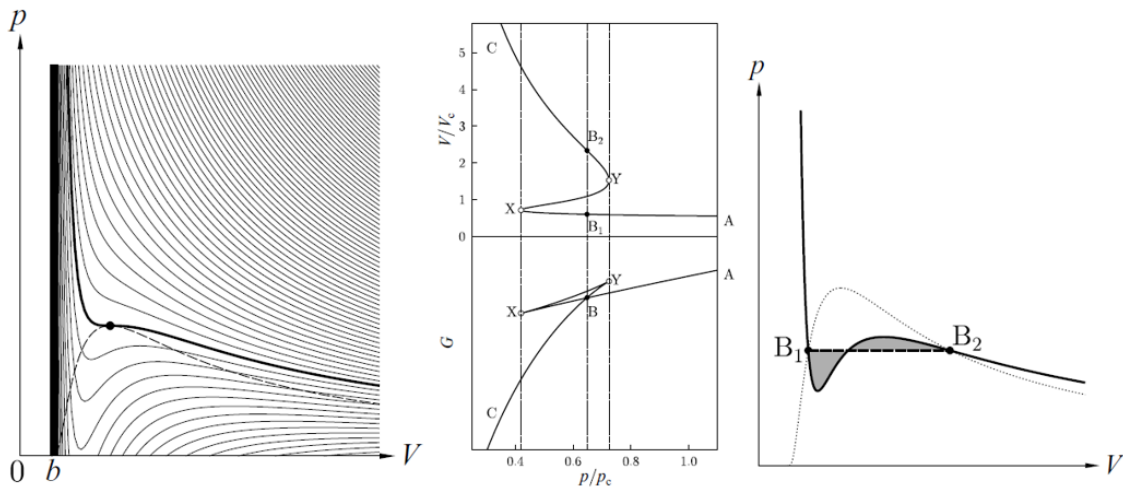


Figure 5: (left) van der Waals Isotherms (thin lines) with the critical isotherm (thick line) and region in which liquid and vapor are in equilibrium (dashed line) and the critical point; (centre) Behaviour of V and G of a van der Waals gas as a function of p at $T = 0.9T_c$; (right) Maxwell construction for van der Waals gas, phase existence occurs between points B_1 and B_2 when the shaded areas are equal.

1.8 Ideal gas mixtures

In a mixture of non-interacting ideal gases, the property of the mixture is just the sum of the independent contributions of each species of ‘component’ of gas.

Theorem 1.13 (Dalton’s law): The total pressure is the sum of the partial pressures, i.e. $p = \sum_i p_i = N_i k_B T / V$, where the entire mixture is at temperature T and volume V .

Proposition 1.7 (Entropy of mixing): The difference between the entropy of an ideal gas mixture and that of a separated out assembly of pure gases each at pressure p is called the entropy of mixing, and it is

$$\Delta S = -N k_B \sum_i (p_i/p) \log(p_i/p)$$

Proof. From first law,

$$dS = \frac{dU}{T} + \frac{pdV}{T} \implies S = C_V \log T + N k_B \log V$$

unique up to some constant. But, T and V are the same for the entire mixture, so we express it in terms of partial pressures (intensive variables), to get

$$S = C_p \log T - N k_B \log p$$

Starting with a collection of vessels all at the same pressure p and containing numbers of particles N_i of gas component i , we could achieve the mixing process by first, an isothermal expansion of each component i to the desired final pressure p_i , followed by joining together all the containers, then isothermally compressing again to the same pressure p to partial pressure p_i , its entropy changes by $\Delta S_i = -N_i k_B \log(p_i/p)$. The difference is entropy is

$$\Delta S = -k_B \sum_i N_i \log(p_i/p) = -N k_B \sum_i c_i \log c_i$$

where $c_i = N_i/N = p_i/p$ is the concentration of the i th component. \square

Proposition 1.8: The chemical potential of the i th component, which is present with concentration c_i in a mixture of ideal gases, differs from the chemical potential for the pure gas at the same total pressure by $k_B T \log c_i$.

Proof. At constant T , we have

$$\left(\frac{\partial G}{\partial p} \right)_T = \left(\frac{\partial \mu N}{\partial p} \right)_T = \frac{k_B T N}{p} \implies \mu(T, p) = \mu(T, p_0) + k_B T \log \frac{p}{p_0}$$

The Gibbs free energy of the mixture is $G = \sum_i G_i$. Using $(\partial G_i / \partial p_i)_{T, N_i} = V$, we have

$$G_i(p_i, T) = G_i(p, T) + \int_p^{p_i} V dp = G_i(p, T) + N_i k_B T \log c_i$$

Since $G_i = \mu_i N_i$, we have $\mu_i(p_i, T) = \mu_{0,i}(p, T) + k_B T \log c_i$. \square

Example 1.9. For a chemical equilibrium, we start with ensuring the Gibbs free energy is a minimum. At constant T and p , $d\mu_i = 0$ from Gibbs-Duhem equation. A variation of G about the equilibrium ratio of chemical components thus satisfy

$$0 = dG = \sum_i \mu_i dN_i = \sum_i \nu_i \mu_i$$

where the dN_i ’s are related by matter conservation. Now, with the equation for μ_i , we have

$$0 = \sum_i \nu_i \mu_{0,i} + k_B T \sum_i \nu_i \log c_i = \sum_i \nu_i \mu_{0,i} + k_B T \log \prod_i c_i^{\nu_i}$$

We can define the equilibrium constant for concentration to be $K_c(T) = \prod_i c_i^{\nu_i}$.

2 Recap of Statistical Mechanics

2.1 Definitions

Note 2.1 (Idea of Statistical Mechanics). *The primary goal of statistical thermodynamics (also known as equilibrium statistical mechanics) is to derive the classical thermodynamics of materials in terms of the properties of their constituent particles and the interactions between them. Physical properties of thermodynamic systems in equilibrium are given by the averages calculated with an appropriate ensemble which assign probabilities to respective microstates. Whereas statistical mechanics proper involves dynamics, here the attention is focused on statistical equilibrium (steady state). Statistical equilibrium does not mean that the particles have stopped moving (mechanical equilibrium), rather, only that the ensemble is not evolving.*

Note 2.2 (Fundamental Postulate of Statistical Mechanics). *The fundamental postulate is the equal a priori probability postulate which states that:*

For an isolated system with an exactly known energy and exactly known composition, the system can be found with equal probability in any microstate consistent with that knowledge.

There are various arguments in favour of the equal a priori probability postulate:

- *Ergodic hypothesis: An ergodic system is one that evolves over time to explore all accessible states: all those with the same energy and composition. In an ergodic system, the microcanonical ensemble is the only possible equilibrium ensemble with fixed energy. This approach has limited applicability, since most systems are not ergodic.*
- *Principle of indifference: In the absence of any further information, we can only assign equal probabilities to each compatible situation.*
- *Maximum information entropy: A more elaborate version of the principle of indifference states that the correct ensemble is the ensemble that is compatible with the known information and that has the largest Gibbs entropy (information entropy).*

Definition 2.1 (Ensemble): *An ensemble is a mental collection of identical copies of the same physical system in which each copy is in a different microstate labelled by α .*

Definition 2.2 (Ensemble Average): *Ensemble average of the physical property X is $\langle X \rangle = \sum_{\alpha} p_{\alpha} X_{\alpha}$ for $0 \leq p_{\alpha} \leq 1$ where p_{α} is the probability of a member in the ensemble in microstate α . Note that due to normalization, $\sum_{\alpha} p_{\alpha} = 1$.*

Theorem 2.1 (Method of Lagrange Multipliers): *Suppose we want to maximize a function of Ω variables, say $f(p_1, \dots, p_{\Omega})$, subject to some constraint that has a general form $g(p_1, \dots, p_{\Omega}) = 0$. The method is to add to the function the constraint multiplied by $-\lambda$ (known as the Lagrange multiplier) and maximize the resulting function unconditionally, with respect to the $\Omega + 1$ number of variables.*

Proof. At the point of maximum of f , $df = \sum_{i=1}^{\Omega} \frac{\partial f}{\partial p_i} dp_i = 0$. But the increments $\{dp_i\}$ are not independent because they are only allowed to change subject to the constraint. So g cannot change, i.e. $dg = 0$. Then, we have say

$$dp_1 = - \sum_{i=2}^{\Omega} \frac{\partial g}{\partial p_i} \left(\frac{\partial g}{\partial p_1} \right)^{-1} dp_i$$

Substitute this back to $df = 0$, we then have $\{dp_2, \dots, dp_{\Omega}\}$ to be now all independent. So, $\frac{\partial f}{\partial p_{\alpha}} - \lambda \frac{\partial g}{\partial p_{\alpha}} = 0$ for $\alpha = 2, \dots, \Omega$. This includes $\alpha = 1$, following the definition of λ . SO we now have $\Omega + 1$ variables and $\Omega + 1$ equations (including the constraint). Essentially, we want to maximize $f - \lambda g$ with respect to $\{p_1, \dots, p_{\Omega}, \lambda\}$ and with no constraints. We can extend this to the case of several constraints. \square

Theorem 2.2 (Stirling's Formula):

$$\log(N!) = N \log(N) - N + O(\log N)$$

2.2 Microcanonical ensemble

Definition 2.3 (Microcanonical ensemble): A microcanonical ensemble is a statistical ensemble where the total energy of the system and the number of particles in the system are each fixed to particular values; each of the members of the ensemble are required to have the same total energy and particle number. The system must remain totally isolated (unable to exchange energy or particles with its environment) in order to stay in statistical equilibrium.

Remarks 2.1.

1. The energy levels of a quantum Hamiltonian of a N -particle system are discrete. For large $N \sim 10^{23}$, we can treat them as a continuum. $\Omega(E)$ counts the number of states with energy between E and $E + \delta E$, where δE is small compared to the accuracy of our measuring apparatus, but large compared to the gap between levels. The gap between levels is far less than the measurement uncertainty δE . Suppose energy is confined in a narrow band, the number of states in band with energy less than E is $\Omega(E) = g(E)\delta E$.
2. We describe the system in terms of a probability distribution over the quantum states, i.e. mixed states. $p(n)$ is not related to the quantum indeterminacy, but due to our ignorance.

Definition 2.4 (Boltzmann Entropy): We define the entropy of a system, $S = S(E)$, to be

$$S(E) = \log(\Omega(E))$$

where $\Omega(E)$ is the number of states with energy E .

Remarks 2.2. With $\Omega(E) = g(E)\delta E$, the entropy is

$$S(E) = \log(\Omega(E)) = \log(g(E)\delta E) = \log(g(E)) + \log(\delta E)$$

where $\log(g(E))$ depends on N , but $\log \delta E$ is independent of N . The effect of δE is negligible.

By the equal a priori probability postulate, every accessible microstate of an isolated system in thermodynamic equilibrium is equally probable, i.e.

$$p_{\alpha}^{m.c.}(E, V, N) = \frac{1}{\Omega(E, V, N)}$$

Proposition 2.1 (Additivity of Entropy): Entropy is additive, i.e. Two systems with Ω_1 and Ω_2 states combined to form a system with $\Omega_1\Omega_2$ states and total entropy

$$S = \log(\Omega_1\Omega_2) = \log(\Omega_1) + \log(\Omega_2) = S_1 + S_2$$

Example 2.1 (Two-State Systems). Consider N identical, weakly interacting, localized particles, each with non-degenerate energy levels, $\varepsilon_n = (-1)^n \varepsilon$ where $n = 1, 2$. Each possible microstate is specified by $\alpha = (n_1, n_2, \dots, n_N)$ where $n_i = 1, 2$ gives energy level occupied by the i -particle; N_i particles occupies ε_i , i.e. $N = N_1 + N_2$ and $E = N_1\varepsilon_1 + N_2\varepsilon_2 = (N_2 - N_1)\varepsilon$. By Stirling's approximation, the entropy is

$$S(E, N) = -\frac{Nk}{2} \left(1 - \frac{E}{N\varepsilon}\right) \log \left(\frac{1}{2} \left(1 - \frac{E}{N\varepsilon}\right)\right) - \frac{Nk}{2} \left(1 + \frac{E}{N\varepsilon}\right) \log \left(\frac{1}{2} \left(1 + \frac{E}{N\varepsilon}\right)\right)$$

where $\Omega(E, N; N_2) = \frac{N!}{N_2!(N-N_2)!}$ with $N_{1,2} = \frac{N}{2} (1 \mp \frac{E}{N\varepsilon})$.

Example 2.2 (Harmonic Oscillators). Consider N identical, weakly interacting, localized harmonic oscillators, each with non-degenerate set of energy levels $\varepsilon_n = n\hbar\omega$ where $n \in \mathbb{Z}^+ \cup \{0\}$. Each possible microstate is specified by $\alpha = (n_1, \dots, n_N)$ where n_i is the energy level occupied by the i -oscillator with energy value $E_i = n_i\hbar\omega$ and N_n is the number of oscillators occupy energy level ε_n . Our constraints are thus $N = \sum_{n=1}^{\infty} N_n$ and $E = \sum_{i=1}^N E_i = \sum_{n=1}^{\infty} n N_n \hbar\omega$. By Stirling's approximation,

$$S(E, N) = -\frac{Ek}{\hbar\omega} \log \left(\frac{E/\hbar\omega}{N + E/\hbar\omega} \right) - Nk \log \left(\frac{N}{N + E/\hbar\omega} \right)$$

where $\Omega(E, N; M) = \frac{(N-1+M)!}{(N-1)!M!}$ such that $M = \sum_{i=1}^N n_i = \frac{E}{\hbar\omega}$.

2.3 Gibbs Entropy

Consider a small system in thermal contact with a reservoir. The fluctuations in the energy of the system will be an appreciable fraction of the mean energy. The states of the system cannot simply be labelled as ‘accessible’ or ‘inaccessible’, we have to think about the probabilities that the system will be in each of its possible states.

Proposition 2.2: The Gibbs Entropy is

$$S = -k_B \sum_n p(n) \log p(n)$$

where $p(n)$ is the probability for the system to attain a state n .

Proof. Apply microcanonical ensemble to the entire collection of W copies of the system, where W is large, in thermal contact with one another. For a particular member of the ensemble, the other $W - 1$ replicas act as a heat reservoirs. The ensemble of replicas is isolated from the surroundings and therefore its internal energy U is constant. Since all of the accessible microstates of the ensemble have internal energy U , then by a priori probability postulate, they are all equally likely. The number of systems in the state $|n\rangle$ is $Wp(n)$. the number of ways to put $p(n)W$ systems into the state $|n\rangle$ for each $|n\rangle$ is $\Omega = \frac{W!}{\prod_n (p(n)W)!}$. Using the Boltzmann entropy,

$$S = k_B \log \Omega = k_B (W \log W - \sum_n W p(n) \log(W p(n))) = -k_B W \sum_n p(n) \log p(n)$$

where we used Stirling’s approximation. Thus, the entropy for a single copy of the system is the probability distribution $p(n)$ over the states. \square

Remarks 2.3.

1. The quantum analog of the Gibbs entropy is

$$S = -\text{Tr } \hat{\rho} \log \hat{\rho}$$

again, see II Principles of Quantum Mechanics.

2. The Gibbs entropy reverts to the Boltzmann entropy if $p(n) = 1/\Omega(E)$, which is applicable for large systems in equilibrium, i.e. For a large system, the probability distribution is very sharply peaked in energy.
3. Gibbs entropy is the most fundamental entropy in equilibrium statistical mechanics, that is applicable to both large and small systems.

Theorem 2.3: Suppose we have no constraint (other than the sum of probabilities), such that the system is isolated, the microstates have equal probabilities, i.e. maximum ignorance.

Proof. Maximizing the Gibbs entropy S_G subject to the constraint $\sum_\alpha p_\alpha = 1$, we have

$$dS_G - \lambda \sum_\alpha dp_\alpha - \left(\sum_\alpha p_\alpha - 1 \right) d\lambda = 0$$

where λ is a Lagrange multiplier. Since $S_G = -\sum_\alpha p_\alpha \log(p_\alpha)$, we have $dS_G = -\sum_\alpha (\log p_\alpha + 1) dp_\alpha$ and we have $-\sum_\alpha (\log p_\alpha + 1 + \lambda) dp_\alpha - (\sum_\alpha p_\alpha - 1) d\lambda = 0$, then $p_\alpha = e^{-(1+\lambda)}$ and $e^{-(1+\lambda)} = \frac{1}{\Omega}$ and so $S_G = \log \Omega$ and $p_\alpha = \frac{1}{\Omega}$. \square

2.4 Canonical ensemble

Definition 2.5 (Canonical Ensemble): Consider a closed system Q that can exchange energy, but not particles, with a large heat reservoir Q' , at an equilibrium temperature T .

What is the probability for finding the system Q in a particular microstate of energy E ? This is particularly important if the system Q is small on microscopic scale. We regard the total system $Q + Q'$ as microcanonical, i.e. with fixed energy E_0 and uncertainty δE .

Proposition 2.3: All states of a canonical ensemble, with the same energy E , have the same probability distribution as a microcanonical probability distribution. The probability for the system to have energy E_r is

$$p_r = \frac{e^{-E_r/T}}{\sum_{r'} e^{-E_r'/T}}$$

Proof. Let total energy of system Q and reservoir Q' be E_0 . At constant entropy S , energy conservation will mean $dE = -PdV$. If Q has energy E , then Q' has energy $E_0 - E$. All states of the total system are equally likely, by principle of equal a priori probability. The number of states with given energy splitting E and $E_0 - E$ is

$$g'(E_0 - E)\delta E = e^{S'(E_0 - E)}$$

where S' is the entropy of heat bath Q' at energy $E_0 - E$, g' is the density of states of the heat bath. Now, since Q' is a reservoir which has a large supply of energy, then $E \ll E_0$, and so expand

$$S'(E_0 - E) = S'(E_0) - \left. \frac{dS'}{dE} \right|_{E=E_0} E$$

where $\frac{\partial S'}{\partial E}(E = E_0) = \frac{1}{k_B T}$ from earlier. The number of states is $e^{S'(E_0) - (E/k_B T)}$ which is directly proportional with $e^{-E/k_B T}$, i.e. the probability of a given microstate for system Q is proportional to $e^{-E/k_B T}$. This is the canonical probability distribution. \square

Definition 2.6 (Partition Function): We define the quantity $Z = \sum_r e^{-E_r/k_B T}$ to be the partition function.

Remarks 2.4. The canonical ensemble is really a special choice of the density matrix

$$\hat{\rho} = e^{-\hat{H}/k_B T} / Z$$

See II Principle of Quantum Mechanics.

Proposition 2.4 (Multiplicative of Partition Function): The partition function is multiplicative.

Proof. Consider two systems, the combined partition function is

$$Z = \sum_{n,m} e^{-E_m^{(1)}/k_B T} e^{-E_n^{(2)}/k_B T} = \sum_m e^{-E_m^{(1)}/k_B T} e^{-E_n^{(2)}/k_B T} = Z_1 Z_2$$

\square

Definition 2.7 (Inverse Temperature): We define β to be the inverse temperature, i.e. $\beta = \frac{1}{k_B T}$.

Corollary 2.1: The average energy and energy fluctuations of a canonical ensemble are respectively

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \log Z, \quad \langle (E - \langle E \rangle)^2 \rangle = -\frac{\partial}{\partial \beta} \langle E \rangle$$

Proof. By definition of expectation,

$$\langle E \rangle = \sum_n p_n E_n = \frac{1}{Z} \sum_n E_n e^{-\beta E_n} = -\frac{\partial}{\partial \beta} \log Z$$

where $\frac{\partial}{\partial \beta} Z = -E_n$. Similarly, we compute the variance by

$$\frac{\partial^2}{\partial \beta^2} \log Z = -\frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 + \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}$$

but $\frac{\partial^2 Z}{\partial \beta^2} = \sum_n E_n^2 e^{-\beta E_n}$ and $\frac{\partial Z}{\partial \beta} = \sum_n E_n e^{-\beta E_n}$ and so $\frac{\partial^2 \log Z}{\partial \beta^2} = \langle E^2 \rangle - \langle E \rangle^2 = \langle (E - \langle E \rangle)^2 \rangle$. \square

Corollary 2.2: The heat capacity of a canonical ensemble is $C_V = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}$.

Proof. Since $E = \langle E \rangle$ (conserved neergy), then

$$C_V = \frac{\partial \langle E \rangle}{\partial T} \bigg|_V = \frac{\partial \beta}{\partial T} \frac{\partial \langle E \rangle}{\partial \beta} = -\frac{1}{k_B T^2} \frac{\partial \langle E \rangle}{\partial \beta}$$

From previous result, we get our desired form for C_V . \square

Remarks 2.5.

1. This is related to the fluctuation-dissipation theorem. Dissipation is quantified by C_V while fluctuation is given by ΔE .
2. We have $\Delta E/E \sim 1/\sqrt{N}$. In the thermodynamic limit, i.e. $N \rightarrow \infty$, E peaks sharply at $\langle E \rangle$.

Theorem 2.4: Suppose we have one other constraint, i.e. maximize entropy subject to a fixed value of mean energy, $\sum_\alpha p_\alpha E_\alpha = U$, then we obtain the partition function for a canonical ensemble.

Proof. Maximizing the Gibbs entropy S_G subjected to both constraints is equivalent to maximizing $S_G - \lambda \left(\sum_\alpha p_\alpha - 1 \right) - \beta \left(\sum_\alpha p_\alpha E_\alpha - U \right)$.

$$dS_G - \lambda \sum_\alpha dp_\alpha - \left(\sum_\alpha p_\alpha - 1 \right) d\lambda - \beta \sum_\alpha E_\alpha dp_\alpha - \left(\sum_\alpha p_\alpha E_\alpha - U \right) d\beta = 0$$

where λ, β are Lagrange multipliers. Using the form for S_G , we get

$$- \sum_\alpha (\log p_\alpha + 1 + \lambda + \beta E_\alpha) dp_\alpha - \left(\sum_\alpha p_\alpha - 1 \right) d\lambda - \left(\sum_\alpha p_\alpha E_\alpha - U \right) d\beta = 0$$

We then have $p_\alpha = e^{-1-\lambda-\beta E_\alpha}$. The normalization constant is obtained from $\sum_\alpha p_\alpha - 1 = 0$ and hence $Z(\beta) = \sum_\alpha e^{-\beta E_\alpha}$ which is the partition function. The desired probability distribution is $p_\alpha = \frac{1}{Z(\beta)} e^{-\beta E_\alpha}$. \square

Proposition 2.5: The entropy in the canonical ensemble is

$$S = k_B \frac{\partial}{\partial T} (T \log Z)$$

Proof.

$$S = -k_B \sum_n p(n) \log(p(n)) = -\frac{k_B}{Z} \sum_n e^{-\beta E_n} \log \frac{e^{-\beta E_n}}{Z} = k_B \frac{\beta}{Z} \sum_n E_n e^{-\beta E_n} + k_B \log Z = k_B \frac{\partial}{\partial T} (T \log Z)$$

Alternatively, $\sum_\alpha p_\alpha E_\alpha = \frac{1}{Z(\beta)} \sum_\alpha E_\alpha e^{-\beta E_\alpha} = -\frac{\partial \log Z}{\partial \beta} = U$. The Gibbs entropy is then rewritten as $S_G = \beta U + \log(Z)$. \square

Corollary 2.3: The free energy of a canonical ensemble in terms of the partition function is

$$F = -k_B T \log Z$$

Proof.

$$F = U - TS = k_B T^2 \frac{\partial}{\partial T} \log Z - k_B T \frac{\partial}{\partial T} (T \log Z) = -k_B T \log Z$$

□

Example 2.3 (Two-States System). Consider N identical, weakly interacting, localized particles, each with non-degenerate energy levels, $\varepsilon_n = (-1)^n \varepsilon$ where $n = 1, 2$. Each possible microstate is specified by $\alpha = (n_1, n_2, \dots, n_N)$ where $n_i = 1, 2$ gives energy level occupied by the i -particle; $E_\alpha = \sum_{i=1}^N E_i(n_i)$, where $E_i(n_i) = (-1)^{n_i} \varepsilon$. The free energy is

$$F(T, N) = -k_B T \log Z = -N k_B T \log \cosh(\varepsilon/kT) - N k_B T \log(2)$$

where $Z(T, N) = \sum_\alpha e^{-\beta E_\alpha} = \prod_{i=1}^N \sum_{n_i=1}^2 e^{(-1)^{n_i+1} \beta \varepsilon} = 2^N (\cosh(\varepsilon/kT))^N$. The internal energy will be

$$U = -\frac{dZ}{d\beta} = -\frac{d}{d\beta} \log(2^N (\cosh(\beta \varepsilon))^N) = -N \frac{1}{2 \cosh(\beta \varepsilon)} \frac{d}{d\beta} 2 \cosh(\beta \varepsilon) = -N \varepsilon \tanh(\beta \varepsilon)$$

The heat capacity will be

$$C = \frac{dU}{dT} = -N \varepsilon (-k_B \beta^2) \frac{d}{d\beta} \tanh(\beta \varepsilon) = N \varepsilon^2 \beta^2 k \operatorname{sech}^2(\beta \varepsilon)$$

For an array of N localized spin 0.5 paramagnetic atoms, $\varepsilon = \mu_B B$ and so we can write

$$C = N k_B (\theta/T)^2 \frac{e^{\theta/T}}{e^{\theta/T} + 1)^2}$$

where $\theta = 2\mu_B B/k_B$. $dC/dT = 0 \implies C_{\max}$ occurs at $\theta/T = 2.4$ and so $T = \frac{5}{6} \frac{\mu_B B}{k_B}$. In fact, $\lim_{T \rightarrow \infty} C = N k_B (\theta/T)^2$ and $\lim_{T \rightarrow 0} C = N k_B (\theta/T)^2 e^{-\theta/T}$.

Example 2.4 (Harmonic Oscillator). Consider N identical, weakly interacting, localized harmonic oscillators, each with non-degenerate set of energy levels $\varepsilon_n = n\hbar\omega$ where $n \in \mathbb{Z}^+ \cup \{0\}$. Each possible microstate is specified by $\alpha = (n_1, \dots, n_N)$ where n_i is the energy level occupied by the i -oscillator with energy value $\varepsilon_i = n_i \hbar\omega$. Note $E_\alpha = \sum_{i=1}^N E_i(n_i)$ where $E_i(n_i) = \varepsilon_i = n_i \hbar\omega$. The free energy is

$$F(T, N) = -k_B T \log Z = N k_B T \log(1 - e^{-\hbar\omega/kT})$$

where $Z(T, N) = \prod_{i=1}^N \sum_{n_i=0}^{\infty} e^{-\beta \hbar\omega n_i} = \left(\frac{1}{1 - e^{-\beta \hbar\omega}}\right)^N$.

2.5 Grand canonical ensemble

In reality, many systems have a variable particle number. For instance:

- system in contact through a permeable membrane;
- system where chemical reaction can occur;
- a subsystem of a gas, i.e. particles can enter and leave;
- electromagnetic radiation, where photon number depends on energy.

Definition 2.8 (Extensive quantities): Extensive quantities scale with particle number N and simultaneously on V .

Example 2.5. U , S and $F = U - TS$ are extensive, i.e. $U = Nf_U(\frac{S}{N}, \frac{V}{N})$ and $F = Nf_F(T, \frac{V}{N})$ for some function f_U and f_F .

Definition 2.9 (Chemical Potential): We define the chemical potential μ to be

$$\mu := \left. \frac{\partial U}{\partial N} \right|_{S,V}$$

where N is a further independent variable of the system, i.e. $dE = TdS - pdV + \mu dN$.

Example 2.6. Quantities such as T , p and μ are intensive since

$$\frac{1}{T} = \frac{\partial S}{\partial U}, \quad p = T \frac{\partial S}{\partial V}, \quad \mu = T \frac{\partial S}{\partial N}$$

are all ratio of two extensive quantities.

Definition 2.10 (Grand Canonical Ensemble): Consider a closed system Q that can exchange energy and particles with a large heat and particle reservoir Q' at an equilibrium temperature T and chemical potential μ . Note that all exchange is between Q and Q' and not the environment.

Definition 2.11 (Grand Partition Function):

$$\mathcal{Z} = \sum_r e^{-\beta(E_r - \mu N_r)}$$

Proposition 2.6: The probability distribution for a grand canonical ensemble is

$$p_r = \frac{e^{-\beta(E_r - \mu N_r)}}{\mathcal{Z}}$$

Proof. We use a similar argument. Let the total energy and particle number be E_0 and N_0 , the system energy and particle number be $E \ll E_0$ and $N \ll N_0$, i.e. dominated by reservoir. Use $dS = \frac{dE}{T} - \frac{\mu dN}{T}$. The entropy of the bath is

$$S'(E_0 - E, N_0 - N) = S'(E_0, N_0) - \frac{E}{T} + \frac{\mu N}{T}$$

The probability of a system being in a microstate with energy E_r , particle number N_r is

$$p_r = \frac{1}{\mathcal{Z}} e^{-\beta E_r + \beta \mu N_r}$$

with r being a state label. The normalization constant is the grand partition function. \square

Remarks 2.6. For macroscopic system, the sum is sharply peaked around certain E and N where there are $e^{S(U,N)}$ number of states, such that $\frac{\partial S}{\partial U}|_N = \frac{1}{T}$ and $\frac{\partial S}{\partial N}|_U = \frac{\mu}{T}$.

Proposition 2.7: The entropy in the grand canonical ensemble is

$$S = k_B \frac{\partial}{\partial T} (T \log \mathcal{Z})$$

Proof.

$$\begin{aligned}
S &= -k_B \sum_n p(n) \log(p(n)) \\
&= -\frac{k_B}{\mathcal{Z}} \sum_n e^{-\beta(E_n - \mu N_n)} \log \frac{e^{-\beta(E_n - \mu N_n)}}{\mathcal{Z}} \\
&= k_B \frac{\beta}{\mathcal{Z}} \sum_n (E_n - \mu N_n) e^{-\beta(E_n - \mu N_n)} + k_B \log \mathcal{Z} \\
&= k_B \frac{\partial}{\partial T} (T \log \mathcal{Z})
\end{aligned}$$

□

Corollary 2.4:

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \log \mathcal{Z}, \quad \langle (N - \langle N \rangle)^2 \rangle = \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial \mu}, \quad \langle E \rangle - \mu \langle N \rangle = -\frac{\partial}{\partial \beta} \log \mathcal{Z}$$

Proof. By definition of expectation,

$$\langle N \rangle = \sum_n p_n N_n = \frac{1}{\mathcal{Z}} \sum_n N_n e^{-\beta(E_n - \mu N_n)} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \log \mathcal{Z}$$

where $\frac{\partial}{\partial \mu} \mathcal{Z} = \beta N_n$. Similarly, we compute the variance by

$$\frac{\partial^2}{\partial \mu^2} \log \mathcal{Z} = -\frac{1}{\mathcal{Z}^2} \left(\frac{\partial \mathcal{Z}}{\partial \mu} \right)^2 + \frac{1}{\mathcal{Z}} \frac{\partial^2 \mathcal{Z}}{\partial \mu^2}$$

but $\frac{\partial^2 \mathcal{Z}}{\partial \mu^2} = \beta^2 \sum_n N_n^2 e^{-\beta(E_n - \mu N_n)}$ and $\frac{\partial \mathcal{Z}}{\partial \beta} = \sum_n \beta N_n e^{-\beta(E_n - \mu N_n)}$ and so $\frac{\partial^2 \log \mathcal{Z}}{\partial \beta^2} = \langle N^2 \rangle - \langle N \rangle^2 = \langle (N - \langle N \rangle)^2 \rangle$. Finally,

$$\langle E \rangle - \mu \langle N \rangle = \sum_n p_n (E_n - \mu N_n) = \frac{1}{\mathcal{Z}} \sum_n (E_n - \mu N_n) e^{-\beta(E_n - \mu N_n)} = -\frac{\partial}{\partial \beta} \log \mathcal{Z}$$

□

Remarks 2.7. We have $\Delta N/N \sim 1/\sqrt{N}$. In the thermodynamic limit, i.e. $N \rightarrow \infty$, N peaks sharply at $\langle N \rangle$.

Theorem 2.5: Suppose we have two other constraints, i.e. maximize entropy subject to a fixed value of mean energy, $\sum_\alpha p_\alpha E_\alpha = U$ and a fixed value of mean particle number, $\sum_\alpha p_\alpha N_\alpha = \langle N \rangle$, then we obtain the grand canonical distribution.

Proof. We need to maximize $S_G - \lambda(\sum_\alpha p_\alpha - 1) - \beta(\sum_\alpha p_\alpha E_\alpha - U) + \beta\mu(\sum_\alpha p_\alpha N_\alpha - \langle N \rangle)$. We then find $\log p_\alpha + 1 + \lambda + \beta E_\alpha - \beta\mu N_\alpha = 0$ and so $p_\alpha = \frac{1}{\mathcal{Z}(\beta, \mu)} e^{-\beta(E_\alpha - \mu N_\alpha)}$ where $\mathcal{Z}(\beta, \mu) = \sum_\alpha e^{-\beta(E_\alpha - \mu N_\alpha)}$ is the grand canonical partition function. □

Definition 2.12 (Grand Potential): We define the grand potential Φ to be the quantity

$$\Phi := U - TS - \mu N = F - \mu N$$

Corollary 2.5:

$$\Phi = -k_B T \log \mathcal{Z}$$

Proof.

$$\Phi = U - \mu N - TS = -\frac{\partial}{\partial \beta} \log \mathcal{Z} - T k_B \frac{\partial}{\partial T} (T \log \mathcal{Z}) = k_B T^2 \frac{\partial}{\partial T} \log \mathcal{Z} - k_B T \frac{\partial}{\partial T} (T \log \mathcal{Z}) = -k_B T \log \mathcal{Z}$$

□

Remarks 2.8.

1. The sum in the grand canonical partition function still converge for large N since $\mu < 0$. Implicitly, particles were conserved when passing from the bath to the system. If not, for example, for photons, we imagine photons coming from the bath. But S' does not depend on their number, so $\mu = 0$.
2. The differential of the grand potential is

$$d\Phi = dE - TdS - SdT - \mu dN - Nd\mu = -SdT - PdV - Nd\mu$$

such that $N = -\frac{\partial\Phi}{\partial\mu}|_{T,V}$. So, the natural variables of Φ are T , V and μ , i.e. $\Phi = \Phi(T, V, \mu)$. There is only one extensive independent variable, namely V , i.e. $\Phi(T, \lambda V, \mu) = \lambda\Phi(T, V, \mu)$. But since

$$\left.\frac{\partial\Phi}{\partial V}\right|_{T,\mu} = -P$$

then $\Phi = -P(T, \mu)V$.

Corollary 2.6: We can recover the canonical distribution from the grand canonical distribution.

Proof. Set $N_\alpha = N \forall \alpha$, we thus have $p_\alpha = e^{-\beta E_\alpha} \frac{e^{\beta \mu N}}{\mathcal{Z}} = \frac{e^{-\beta E_\alpha}}{\mathcal{Z}}$ such that $Z = \sum_\alpha e^{-\beta E_\alpha}$. \square

Example 2.7 (Two State Systems). Consider N identical, weakly interacting, localized particles, each with non-degenerate energy levels, $\epsilon_n = (-1)^n \epsilon$ where $n = 1, 2$. Each possible microstate is specified by $\alpha_{N_S} = (n_1, n_2, \dots, n_{N_S})$ where $n_i = 1, 2$ gives energy level occupied by the i -particle; $E_{N_S, \alpha} = \sum_{i=1}^{N_S} E_i(n_i)$, where $E_i(n_i) = (-1)^{n_i} \epsilon$. The grand potential is

$$\Phi(T, \mu) = -k_B T \log \mathcal{Z}(T, \mu) = k_B T \log \{1 - e^{\beta \mu} [e^{\beta \epsilon} + e^{-\beta \epsilon}]\}$$

where $Z(T, N) = \sum_\alpha e^{-\beta E_{N_S, \alpha}} = \prod_{i=1}^{N_S} \sum_{n_i=1}^2 e^{(-1)^{n_i+1} \beta \epsilon} = (e^{\beta \epsilon} + e^{-\beta \epsilon})^{N_S}$ and hence $\mathcal{Z}(T, \mu) = \sum_{N_S=0}^{\infty} (e^{\beta \mu})^{N_S} Z(T, N_S) = \frac{1}{1 - e^{\beta \mu} (e^{\beta \epsilon} + e^{-\beta \epsilon})}$.

Example 2.8 (Harmonic Oscillators). Consider identical, weakly interacting, localized harmonic oscillators, each with non-degenerate set of energy levels $\epsilon_n = n\hbar\omega$ where $n \in \mathbb{Z}^+ \cup \{0\}$. Each possible microstate is specified by $\alpha_{N_S} = (n_1, \dots, n_{N_S})$ where n_i is the energy level occupied by the i -oscillator with energy value $\epsilon_i = n_i \hbar \omega$. Note $E_{N_S, \alpha} = \sum_{i=1}^{N_S} E_i(n_i)$ where $E_i(n_i) = \epsilon_i = n_i \hbar \omega$. The grand potential is

$$\Phi(T, \mu) = -k_B T \log \mathcal{Z}(T, \mu) = -k_B T \log \left(\frac{1 - e^{-\beta \hbar \omega}}{1 - e^{\beta \mu} - e^{-\beta \hbar \omega}} \right)$$

where $\frac{e^{\beta \mu}}{1 - e^{-\beta \hbar \omega}} < 1$. We have $Z(T, N_S) = \prod_{i=1}^{N_S} \sum_{n_i=0}^{\infty} e^{-\beta \hbar \omega n_i} = \left(\frac{1}{1 - e^{-\beta \hbar \omega}} \right)^{N_S}$. We then have $\mathcal{Z}(T, \mu) = \sum_{N_S=0}^{\infty} (e^{\beta \mu})^{N_S} Z(T, N_S) = \frac{1 - e^{-\beta \hbar \omega}}{1 - e^{\beta \mu} - e^{-\beta \hbar \omega}}$.

Corollary 2.7: Suppose we have numerous species of particles, the grand potential is the sum of that of the individual species.

Proof. Noting that $N = n_1 + n_2 + \dots$, we have

$$\mathcal{Z}(T, \mu, V) = \sum_{N=0}^{\infty} \sum_{n_1, n_2, \dots}^N e^{-\beta(n_1 E_1 + n_2 E_2 + \dots - \mu N)} = \sum_{n_1, n_2, \dots} e^{-\beta(E_1 - \mu)n_1} e^{-\beta(E_2 - \mu)n_2} \dots = \prod_i \mathcal{Z}_i$$

where $\mathcal{Z}_i = \sum_{n_i} e^{-\beta(E_i - \mu)n_i}$. Finally, $\Phi = -k_B T \log \mathcal{Z}$, so $\Phi = -k_B T \log \prod_i \mathcal{Z}_i = -k_B T \sum_i \log \mathcal{Z}_i = \sum_i \Phi_i$. \square

Remarks 2.9. Suppose our system have internal degrees of freedom and external degrees of freedom, which are described by the partition functions Z_{int} and Z_{ext} respectively. If they are independent degrees of freedom, we may multiply to the existing partition function Z .

Example 2.9. Suppose we have a diatomic molecule with internal degrees of freedom (rotational and vibrational) in an external potential. The energy of the molecule is

$$E = E_{\mathbf{k}} + E_{\text{int}} + E_{\text{ext}}, \quad E_{\text{int}} = (n + 0.5)\hbar\omega_0 + \frac{\hbar^2 J(J+1)}{2I}$$

An example of an external potential is gravity. The gravitational field does not affect the internal degrees of freedom because the gravitational potential does not vary significantly over the size of the molecule. But for an external magnetic field, it interacts with the rotational states, lifting their degeneracies, hence affecting the internal degrees of freedom. Consider the former where we have a small volume of gas at height h , the grand partition function for energy level $E_{\mathbf{k}}$ in the classical limit is

$$\begin{aligned} \mathcal{Z}_{\mathbf{k}}(h) &= 1 + \sum_{\text{int}} e^{-\beta(E_{\mathbf{k}} + E_{\text{int}} + E_{\text{ext}} - \mu)} = 1 + Z_{\text{int}} Z_{\text{ext}}(h) e^{-\beta(E_{\mathbf{k}} - \mu)} \\ \implies \Phi_{\mathbf{k}}(h) &= -k_B T \ln \mathcal{Z}_{\mathbf{k}}(h) \approx -k_B T Z_{\text{int}} Z_{\text{ext}}(h) e^{-\beta(E_{\mathbf{k}} - \mu)} \end{aligned}$$

The chemical potential of the gas for fixed particle number N is

$$N = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{T,V} \implies \mu = k_B T \left[\ln \left(\frac{N(h) \lambda^3}{\sigma V} \right) - \ln Z_{\text{int}} - \ln Z_{\text{ext}}(h) \right]$$

where $E_{\text{ext}} = mgh \implies -k_B T \ln Z_{\text{ext}}(h) = mgh$. Now, consider the variation of particle number with height in an ideal classical gas in a gravitational field. In equilibrium, the temperature must be independent of height. In equilibrium μ must also be independent of height, and we see

$$N(h) = N(0) e^{-mgh/k_B T}$$

Proposition 2.8 (Law of mass action): Let the equilibrium constant be $K_N(T) = \prod_i N_i^{\nu_i}$, then

$$K_N(T) = \prod_i (Z_1^i)^{\nu_i}$$

where Z_1 is the single-particle partition function for species i .

Proof. The chemical potential for each species is $\mu_i = k_B T (\ln N_i - \ln Z_1^i)$. where Z_1^i is the single-particle partition function for a single molecule of species i , confined to a volume V , which includes all degrees of freedom. By the condition of chemical equilibrium, $\sum_i \nu_i \mu_i = 0$, we have $K_N(T) = \prod_i N_i^{\nu_i} = \prod_i (Z_1^i)^{\nu_i}$. \square

Example 2.10 (Langmuir adsorption isotherm). Suppose a vapour is in thermal equilibrium with a surface, and that a molecule of the vapour may attach itself to the surface at one of the N_s surface sites, which lowers its energy by an amount ε . The internal degrees of freedom of the molecule can be neglected in the vapour phase, but when the molecule is attached to the surface it vibrates against it, and the partition function of a single adsorbed molecule is $z_s(T)$. The partition function for N adsorbed molecules is $Z_N = \frac{N_s!}{(N_s - N)!} z_s^N e^{-\beta(N\varepsilon)}$. Using Stirling's approximation, we obtain

$$F = -k_B T \ln Z_N = -N\varepsilon - k_B T (N \ln z_s + N_s \ln N_s - (N_s - N) \ln(N_s - N) - N \ln N)$$

The chemical potential is

$$\mu_s = \frac{\partial F}{\partial N} = -\varepsilon - k_B T \left(\ln z_s + \ln \frac{N_s - N}{N} \right) = -\varepsilon + k_B T \ln \frac{\theta}{(1 - \theta) z_s}$$

where $\theta = N/N_s$ is the fraction of the surface sites occupied by molecules. Equivalently, we use the grand partition function is

$$\mathcal{Z} = \sum_i e^{-\beta(E_i - \mu_s N_i)} Z_{\text{int}}^{(i)} = 1 + z_s e^{\beta(\varepsilon + \mu_s)}$$

where each site (the surface sites are independent) has two states: $i = 0$ empty and $i = 1$ full. The corresponding grand potential is $\Phi = -k_B T \ln \mathcal{Z}$ and the average occupation of the site is

$$\theta = \frac{N}{N_s} = - \left(\frac{\partial \Phi}{\partial \mu_s} \right)_{T,V} = \frac{z_s e^{\beta(\varepsilon + \mu_s)}}{1 + z_s e^{\beta(\varepsilon + \mu_s)}} \implies \mu_s = -\varepsilon + k_B T \ln \frac{\theta}{(1 - \theta) z_s}$$

But the chemical potential of an ideal gas (later) is $\mu = k_B T \ln \frac{p \lambda^3}{k_B T}$ and so the Langmuir adsorption isotherm is characterized by $p = \frac{\theta}{1 - \theta} \frac{k_B T}{z_s \lambda^3} e^{-\beta \varepsilon}$.

3 Classical Statistical Mechanics

3.1 Phase Space

Definition 3.1 (Physical space): For a system of N particles moving in 3 dimensions without constraints, the physical space in which the particles move is \mathbb{R}^3 .

Definition 3.2 (Configuration space): The instantaneous configuration of the system is described by the vectors $\mathbf{r}_1, \dots, \mathbf{r}_N$. If there are no constraints, then the $3N$ components of these vectors are independent, and the configuration space (space of all possible configurations) is \mathbb{R}^{3N} . The system is said to have $3N$ degrees of freedom.

Definition 3.3 (Phase space): The future evolution of the system depends on the positions and the velocities. The $6N$ -dimensional vector $(\mathbf{r}_1, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N)^T$ defines a point in phase space, which here is \mathbb{R}^{6N} . The equations of motion define a flow in phase space.

Example 3.1. Consider the example of simple harmonic oscillator in 1D, i.e. $\ddot{x} = -x$. The physical space and configuration space are \mathbb{R} while the phase space is \mathbb{R}^2 . The flow in phase space is clockwise circles, i.e. $\dot{x} = v$ and $\dot{v} = -x$.

Definition 3.4 (Representative Point): A representative point is a single point in phase space represents the state of the whole system (in our case, the precise microstate of a system). In presence of the constraints, these points are confined to some lower dimensional subspace. We regard the initial state of ensemble of systems as corresponding to a density of representative points in phase space.

Definition 3.5 (Phase space density): The phase space density (or probability density) $\rho(\{\mathbf{r}_i, \mathbf{p}_i\})$ equals to the fraction of systems located within an infinitesimal volume $d\Gamma = \prod_{i=1}^N \frac{d^3 r_i d^3 p_i}{(2\pi\hbar)^3}$ surrounding the point $\{\mathbf{r}_i, \mathbf{p}_i\}$.

Theorem 3.1 (Liouville's theorem): The phase-space density is conserved following the flow in phase space. Equilibrium (time-independent) solutions for the phase-space density are possible if ρ is a function of the integrals of motion.

Remarks 3.1.

1. The theorem states that the density in phase space evolves as an incompressible fluid. In another words, the volume occupied by ensemble's representative points do not change.
2. The theorem also says ρ is constant along system trajectories in phase space, which are lines of constant internal energy of the system. In equilibrium, ρ depends only on E , so states with the same energy have the same probability, consistent with the principle of equal a priori probability.

In dealing with a closed system, whose energy is confined to a narrow region between energy U and energy $U + dU$, then all states with energy in this interval are equally likely, while those outside have probability zero. The number of states in this energy shell is directly proportional to $\frac{d\Gamma}{dU} dU$.

Definition 3.6 (Entropy): The definitions for Boltzmann and Gibbs entropy respectively are:

$$S = -k_B \log \frac{d\Gamma(U)}{dU} dU, \quad S = -k_B \int \rho \log \rho d\Gamma$$

Remarks 3.2. Further, if we consider a 'subsystem' plus a 'reservoir' as a large closed system, the Boltzmann distribution is

$$\rho = \frac{e^{-\beta E(\{\mathbf{p}_i, \mathbf{q}_i\})}}{Z_{cl}}, \quad Z_{cl} = \int e^{-\beta E(\{\mathbf{p}_i, \mathbf{q}_i\})} d\Gamma$$

where Z_{cl} is the classical partition function. The phase space arguments will help us in evaluating the partition function sum.

3.2 Ideal Gas of N particles

For the following, we work in the canonical ensemble approach. We aim to reformulate the idea of the partition function in classical mechanics.

Proposition 3.1: Consider an ideal non-interacting gas of N particles in volume V . The gas can be described with phase space coordinates $(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^N$ such that each particle has position $\mathbf{x}^{(n)}$ and momentum $\mathbf{p}^{(n)}$ where $n = 1, 2, \dots, N$. For a single classical particle, the partition function may be defined as the integration over phase space \mathbb{R}^{6N} .

$$Z_1 = \frac{1}{h^3} \int e^{-\beta H(p, q)} d^3 p d^3 q$$

Proof. In one-dimension, the single particle Hamiltonian is $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q})$. The partition function in the canonical ensemble is

$$\begin{aligned} Z_1 &= \sum_n e^{-\beta E_n} \\ &= \sum_n \langle n | e^{-\beta \hat{H}} | n \rangle \\ &= \sum_n \langle n | \int | q \rangle \langle q | e^{-\beta \hat{H}} dq \int | q' \rangle \langle q' | dq' | n \rangle \\ &= \int \langle q | e^{-\beta \hat{H}} | q' \rangle \sum_n \langle q' | n \rangle \langle n | q \rangle dq dq' \\ &= \int \langle q | e^{-\beta \hat{H}} | q' \rangle dq dq' \delta(q' - q) \\ &= \int \langle q | e^{-\beta \hat{H}} | q \rangle dq \end{aligned}$$

In another words, we replaced the sum over the energy eigenstates with an integral over any complete basis eigenstates. We can write Z_1 in a basis-independent fashion.

$$Z_1 = \text{Tr}[e^{-\beta \hat{H}}] = \text{Tr}[e^{-\beta \hat{p}^2/2m} e^{-\beta V(\hat{q})} + O(\hbar)]$$

where $e^{\hat{A}\hat{B}} = e^{\hat{A} + \hat{B} + 0.5[\hat{A}, \hat{B}]}$. If \hat{A} and \hat{B} are canonical variables, then they will give the canonical commutation relation $[\hat{q}, \hat{p}] = i\hbar$. We continue:

$$\begin{aligned} Z_1 &= \int \langle q | e^{-\beta \hat{p}^2/2m} e^{-\beta V(\hat{q})} | q \rangle dq \\ &= \int e^{-\beta V(q)} \langle q | p \rangle \langle p | e^{-\beta \hat{p}^2/2m} | p' \rangle \langle p' | q \rangle dq dp dp' \\ &= \frac{1}{2\pi\hbar} \int e^{-\beta H(p, q)} dp dq \end{aligned}$$

where $\langle q | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipq/\hbar}$. Generalize the result further to 3-dimensions. \square

Proposition 3.2: For an ideal gas (non-interacting) with no internal structure (no rotational or vibrational degree of freedom), then the single particle partition function is

$$Z_1(V, T) = V \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2}$$

Proof.

$$\begin{aligned} Z_1(V, T) &= \frac{1}{(2\pi\hbar)^3} \int e^{-\beta |\mathbf{p}|^2/2m} d^3 q d^3 p \\ &= \frac{V}{(2\pi\hbar)^3} \int e^{-\beta p_x^2/2m} dp_x \int e^{-\beta p_y^2/2m} dp_y \int e^{-\beta p_z^2/2m} dp_z \\ &= V \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \end{aligned}$$

where $\int d^3 q = V$, and $\int e^{-\beta p_i^2/2m} dp_i = \sqrt{2\pi m/\beta}$. \square

Definition 3.7 (de Broglie wavelength): We can define $\lambda := \sqrt{2\pi\hbar^2/mk_B T}$ to be the thermal de Broglie wavelength such that $Z_1 = V/\lambda^3$. If $\lambda^3 \ll V/N$, then quantum effects become important.

3.3 Equipartition Theorem

Theorem 3.2 (Equipartition): Each squared, separable degree of freedom in the Hamiltonian has a mean energy of $\frac{1}{2}k_B T$.

Proof. Consider a Hamiltonian with a separable term that is quadratic in one of the generalized coordinates or momenta, i.e. $H = Aq^2 + \Delta H(Q)$ where Q represents all of the other $6N - 1$ coordinates of the phase space. We can show that the probability distribution of a separable coordinate is independent of all of the other coordinates.

$$\rho(q)dq = \frac{e^{-\beta Aq^2} dq \int e^{-\beta \Delta H} dQ}{\int e^{-\beta Aq^2} dq \int e^{-\beta \Delta H} dQ} = \frac{e^{-\beta Aq^2} dq}{\int e^{-\beta Aq^2} dq} = \frac{e^{-\beta Aq^2} dq}{\sqrt{\pi k_B T/A}}$$

The internal energy is then $U = \langle Aq^2 \rangle + \langle \Delta H(Q) \rangle$, with

$$\langle Aq^2 \rangle = \frac{\int Aq^2 e^{-\beta Aq^2} dq}{\int e^{-\beta Aq^2} dq} = \frac{\sqrt{\pi}/2\beta^{3/2}}{\sqrt{\pi}/\beta^{1/2}} = \frac{1}{2}k_B T$$

□

Remarks 3.3. The equipartition theorem only holds in the classical limit.

Example 3.2. For classical interacting systems of Hamiltonian form

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + V(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

The equipartition theorem still holds irregardless of the form of interactions between particles, as long as the classical limit holds. The probability distribution follows the Maxwell-Boltzmann distribution

$$\rho(v_{x,i})dv = \sqrt{\frac{m}{2\pi k_B T}} e^{-\beta 0.5mv_{x,i}^2} dv$$

The normalized 3-velocity distribution will be the cube of this. Since this is isotropic, the distribution of speeds $v = |\mathbf{v}|$ is

$$\rho(v) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} 4\pi v^2 e^{-mv^2/2k_B T}$$

The mean kinetic energy of atom is

$$\langle v_x^2 \rangle = \sqrt{\frac{m}{2\pi T}} \int_{-\infty}^{\infty} v_x^2 e^{-mv_x^2/2k_B T} dv_x = \frac{k_B T}{m}$$

Hence, for each particle, the mean for $\frac{1}{2}m|\mathbf{v}|^2$ is $\frac{3}{2}k_B T$. This is independent of the interaction potential U . There is $\frac{1}{2}k_B T$ thermal energy per degree of freedom.

Remarks 3.4. This is true for any quadratic degree of freedom, i.e. $\frac{1}{2}k_B T$ energy contribution from each independent degree of freedom. For an ideal (non-interacting) gas, the total energy is $E = \frac{3}{2}Nk_B T$. The gas is approximately ideal at high temperature and low density ($\frac{N}{V} \ll 1$). So, the heat capacity at constant volume is

$$C_V = \left. \frac{dE}{dT} \right|_V = \frac{3}{2}Nk_B$$

3.4 Gas of Diatomic Molecules

Remarks 3.5 (Gibbs paradox). *It may be tempting to directly generalize the single particle partition function z_1 to the N -particle case via $Z = z_1^N$. However, it turns out that the entropy is sensitive to the indistinguishability of the particles. Gibbs considered the following paradox: consider two partitioned regions of the same type of gas. Removing the partition allows gas to mix but since they are of the same type, there should be no increase in entropy due to mixing. A correction factor of $1/N!$ is necessary (number of ways to permute the particle). Hence, for an ideal gas, the partition function is*

$$Z_{\text{ideal}}(N, V, T) = \frac{1}{N!} Z_1^N = \frac{V^N}{\lambda^{3N}} \frac{1}{N!}$$

Proposition 3.3: With Z_{ideal} for a classical gas, one can work out the following macroscopic quantities (using the canonical ensemble):

$$P = \frac{Nk_B T}{V}, \quad E = \frac{3}{2} Nk_B T, \quad S = Nk_B \left[\log \frac{V}{\lambda^3 N} + \frac{5}{2} \right], \quad F = Nk_B T (-\log(V/\lambda^3) + \log N - 1)$$

Proof.

$$P = -\frac{\partial F}{\partial V} = \frac{\partial}{\partial V} k_B T \log Z_{\text{ideal}} = \frac{Nk_B T}{V} \text{ equation of state}$$

$$E = -\frac{\partial}{\partial \beta} \log Z_{\text{ideal}} = \frac{3}{2} Nk_B T \text{ equipartition theorem}$$

$$S = \frac{\partial}{\partial T} T \log Z_{\text{ideal}} = Nk_B \left[\log \frac{V}{N\lambda^3} + \frac{5}{2} \right] \text{ Sackur-Terode equation}$$

$$F = -k_B T \log Z_{\text{ideal}} = -Nk_B T \log \frac{V}{\lambda^3} + Nk_B T \log N - Nk_B T$$

□

Remarks 3.6. *If indistinguishability is not accounted for, we would have incorrectly $S = Nk_B (\log \frac{V}{\lambda^3} + (3/2))$, which is not extensive, i.e. doubling V and N does not double S .*

Proposition 3.4: We may also recover the above results using the grand canonical ensemble. In addition, the chemical potential is

$$\mu = Tk_B \log \frac{\lambda^3 N}{V}$$

Proof. Obtain the grand canonical partition function from the canonical partition function:

$$\mathcal{Z}_{\text{ideal}}(\mu, V, T) = \sum_{N=0}^{\infty} e^{\beta \mu N} Z_{\text{ideal}}(N, V, T) = \exp \left(\frac{e^{\beta \mu} V}{\lambda^3} \right)$$

The particle number is then

$$N = \frac{1}{\beta} \frac{\partial}{\partial \mu} \log \mathcal{Z}_{\text{ideal}} = \frac{e^{\beta \mu} V}{\lambda^3}$$

This then gives the desired expression of μ . The equation of state is obtained via

$$PV = k_B T \log \mathcal{Z}_{\text{ideal}} = k_B T \frac{e^{\beta \mu} V}{\lambda^3} = k_B T N$$

Similar for the rest via $S = k_B \frac{\partial}{\partial T} (T \log \mathcal{Z}_{\text{ideal}})$. □

Remarks 3.7. $\mu = (\frac{\partial E}{\partial N})_{S,V}$ is the energy cost of adding an extra particle at a fixed entropy and volume. However, adding a particle increases the number of ways to share the energy, and so increases entropy. But the entropy is fixed, so $\mu < 0$ for the classical ideal gas. $\mu > 0$ might occur in a suitably strong repulsive gas.

Proposition 3.5: For diatomic molecules, we account for additional independent degrees of freedom, namely two degrees of freedom for rotation, two for vibration, and thus the heat capacity will be in total $\frac{7}{2}N$.

Proof. Model the diatomic molecule to rotate rigidly about the two axes perpendicular to the symmetry axis with moment of inertia I . The Lagrangian, and hence Hamiltonian is

$$\mathcal{L}_{\text{rot}} = \frac{1}{2}I(\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) \implies \mathcal{H}_{\text{rot}} = \dot{\theta}p_{\theta} + \dot{\phi}p_{\phi} - \mathcal{L}_{\text{rot}} = \frac{p_{\theta}^2}{2I} + \frac{p_{\phi}^2}{2I \sin^2 \theta}$$

where $p_{\theta} = I\dot{\theta}$ and $p_{\phi} = I \sin^2 \theta \dot{\phi}$. The single particle partition function Z_1 acquire multiplicative factor

$$\begin{aligned} \int e^{-p_{\theta}^2/2Ik_B T} e^{-p_{\phi}^2/2Ik_B T \sin^2 \theta} \frac{d\theta d\phi dp_{\theta} dp_{\phi}}{(2\pi\hbar)^2} &= \frac{1}{(2\pi\hbar)^2} \int \sqrt{2\pi Ik_B T} \sqrt{2\pi Ik_B T} \sin^2 \theta d\theta d\phi \\ &= \frac{2\pi Ik_B T}{(2\pi\hbar)^2} \frac{\int \sin \theta d\theta d\phi}{4\pi} \\ &= \frac{2Ik_B T}{\hbar^2} \end{aligned}$$

The free energy acquires an additive term $-NTk_B \log \frac{2Ik_B T}{\hbar^2}$. For vibrations, the Hamiltonian is $\mathcal{H}_{\text{vib}} = \frac{p^2}{2\mu} + \frac{1}{2}\mu\omega^2 q^2$. The partition function z acquires a further multiplicative factor

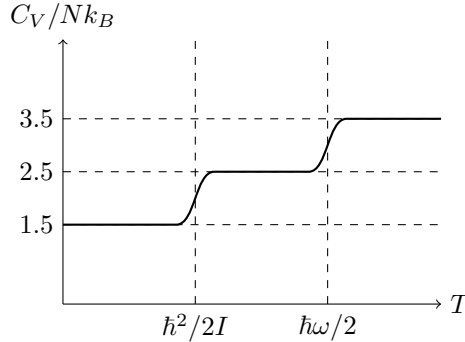
$$\int e^{-p^2/2\mu k_B T} e^{-\mu\omega^2 q^2/2k_B T} \frac{dq dp}{2\pi\hbar} = \frac{1}{2\pi\hbar} \sqrt{2\pi k_B \mu T} \sqrt{\frac{2\pi k_B T}{\mu\omega^2}} = \frac{k_B T}{\hbar\omega}$$

The free energy further acquire an additional term $-Nk_B T \log \frac{k_B T}{\hbar\omega}$ (contributes to another T term for E). The heat capacity will be

$$C_V = -k_B N T f''(T) = -Nk_B T \frac{d^2}{dT^2} \left(-\frac{7}{2} \log T \right) = \frac{7}{2} Nk_B$$

where $f(T) = -\frac{3}{2}k_B T \log \frac{mk_B T}{2\pi\hbar^2} - k_B T \log \frac{k_B T}{\hbar\omega} - k_B T \log \frac{2Ik_B T}{\hbar^2}$. \square

Remarks 3.8. Due to quantum mechanics, the vibrational and rotational modes will freeze out as the temperature is lowered.



In quantum mechanics, the rotational eigenstates are discrete with eigenvalue

$$E_{\text{rot}} = \frac{\hbar^2}{2I} j(j+1), \quad j \in \mathbb{Z}^+ \cup \{0\}$$

which has degeneracy $2j+1$. The single particle partition function is

$$Z_{\text{rot}} = \sum_{j=0}^{\infty} (2j+1) e^{-\beta \hbar^2 j(j+1)/2I}$$

- At high temperatures ($k_B T \gg \frac{\hbar^2}{2I}$): we recover the classical result:

$$Z_{\text{rot}} \approx \int_0^{\infty} (2x+1) e^{-\beta \hbar^2 x(x+1)/2I} dx = \frac{2Ik_B T}{\hbar^2}$$

The corresponding heat capacity is

$$C_{V,\text{rot}} = T \left(\frac{\partial S}{\partial T} \right)_V = -T \left(\frac{\partial^2 F}{\partial T^2} \right)_V = k_B T \left(\frac{\partial^2 T \log Z_{\text{rot}}^N}{\partial T^2} \right)_V = k_B T N \frac{\partial^2}{\partial T^2} T \log \frac{2Ik_B T}{\hbar^2} = k_B T N$$

- At low temperatures ($k_B T \ll \frac{\hbar^2}{2I}$): all states apart from $j = 0$ effectively decouple and $Z_{\text{rot}} \approx 1$, and thus the heat capacity vanishes.

Similarly, the vibrational eigenstates are quantized with eigenvalue

$$E_{\text{vib}} = \hbar\omega(n + 0.5), \quad n \in \mathbb{Z}^+ \cup \{0\}$$

The single particle partition function is then

$$Z_{\text{vib}} = \sum_n e^{-\beta\hbar\omega(n+0.5)} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} = \frac{1}{2 \sinh(\beta\hbar\omega/2)}$$

- At high temperatures $k_B T \gg \hbar\omega$: $Z_{\text{vib}} \approx \frac{1}{\beta\hbar\omega}$. The corresponding heat capacity is

$$C_{V,\text{vib}} = T \left(\frac{\partial S}{\partial T} \right)_V = -T \left(\frac{\partial^2 F}{\partial T^2} \right)_V = k_B T \left(\frac{\partial^2 T \log Z_{\text{vib}}^N}{\partial T^2} \right)_V = k_B T N \frac{\partial^2}{\partial T^2} T \log \frac{k_B T}{\hbar\omega} = k_B T N$$

- At low temperatures ($k_B T \ll \hbar\omega$): $Z_{\text{vib}} \approx e^{-\beta\hbar\omega/2}$ and so the energy converges to the classical limit

$$E_{\text{vib}} = -\frac{\partial}{\partial \beta} \log Z_{\text{vib}} \approx \frac{1}{2} \hbar\omega$$

Similarly, the heat capacity vanishes as $-T \log Z_{\text{vib}}^N = 0.5 T N \hbar\omega / k_B T$ which is independent of temperature.

Remarks 3.9. The limit of a classical ideal gas is $N\lambda^3/V \ll 1$, i.e. the thermal de Broglie volumes associated with the particles do not clash. We can depart from this classical limit in several ways:

- by increasing the density of particles N/V ;
- by increasing the de Broglie wavelength λ , in turn achieved by either lowering the temperatures or by considering lighter particles.

In this case, their wavepackets will start interfering with each other and we are in the quantum regime. Equivalently, the quantum regime occurs when the occupancy of each of the energy levels becomes small and therefore the discreteness becomes relevant, i.e. reciprocal k -space condition. The average occupation of an accessible energy level is

$$\langle n \rangle = \frac{N}{4\pi k_{\text{max}}^3/3} \frac{(2\pi)^3}{V}$$

The crossover occur when $\langle n \rangle \sim 1$. $1/\lambda^3$ is sometimes called the quantum concentration n_Q .

4 Quantum Statistical Mechanics

4.1 Density of states

The density of states appear a lot in this chapter when computing sums over states. Consider an ideal gas trapped in a box of volume $V = L^3$.

4.1.1 Open Boundary Conditions

If we assert the wavefunction must vanish at the walls of the box, then it must be standing waves. The eigenstates and eigenvalues are

$$\psi = \sqrt{\frac{8}{V}} \sin(k_x x) \sin(k_y y) \sin(k_z z), \quad E_{\mathbf{k}} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2)$$

In order to satisfy the boundary conditions, we need $\psi(x_i = 0) = \psi(x_i = L_i) = 0$, $k_i = \frac{n\pi}{L_i}$. The allowed values of the \mathbf{k} -vector form a lattice in \mathbf{k} -space, with volume per point $\Delta k_x \Delta k_y \Delta k_z = \frac{\pi^3}{V}$. The density of allowed \mathbf{k} -vectors will be the inverse of this, multiplied by the spin degeneracy factor g_s .

4.1.2 Periodic Boundary conditions

Now, since $\psi(x_i + L_i) = \psi(x_i)$, the eigenstates must be travelling waves $\psi = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}$. The allowed values of the wavevector \mathbf{k} again form a lattice, but now the spacing between lattice points is $2\pi/L_i$, giving a larger volume per lattice point $(2\pi)^3/V$ and a smaller density $g_s V / (2\pi)^3$. Since we count \mathbf{k} from all eight octants of the 3D wavevector space, whereas for the open boundary conditions, states with wavevector $(\pm k_x, \pm k_y, \pm k_z)$ are actually the same, and so we have to limit our sums to one octant.

4.1.3 Integral conversion

For any macroscopic size box ($\lambda \ll L$), there are many states with $E_n \leq k_B T$ all of which contributes to the sum. Hence, we may approximate the discrete sum to an integral:

$$\sum_n \rightarrow \int d^3 n = \frac{V}{(2\pi)^3} \int d^3 k = \frac{4\pi V}{(2\pi)^3} \int_0^\infty k^2 dk$$

It is thus easier to change variables to $dE = \frac{\hbar^2 k}{m} dk$, so the density of states is

$$\int d^3 n = \int g(E) dE, \quad g(E) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2}$$

where $g(E)dE$ counts the number of states with energy between E and $E + dE$. On the other hand, for massless particles, the density of states is

$$g(E) = \frac{VE^2}{2\pi^2 \hbar^3 c^3}$$

Remarks 4.1. *This conversion is necessary in order to compute the grand potential later. Recall that the full grand potential is summed over all energy states, i.e. $\Phi = \sum_k \Phi_k$. For a continuous non-relativistic gas in 3D, we can cast the discrete sum into an integral over available phase space.*

$$\Phi = \sum_k \Phi_k = \int \Phi(E) \frac{d^3 x d^3 p}{(2\pi\hbar)^3} = \frac{V}{(2\pi\hbar)^3} \int_0^\infty e^{-\beta(E-\mu)} 4\pi p^2 dp = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \sqrt{E} e^{-\beta(E-\mu)} dE$$

which gives $-k_B T V e^{\beta\mu} / \lambda^3$.

Example 4.1. *Consider the partition function of a single particle of mass m in a box of volume V ,*

$$Z_1 = \sum_{k_x, k_y, k_z} e^{-\beta \epsilon_{\mathbf{k}}} = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty e^{-\beta \epsilon} \sqrt{\epsilon} d\epsilon = \frac{V}{4\pi^2} \left(\frac{2mk_B T}{\hbar^2} \right)^{3/2} 2 \int_0^\infty e^{-x^2} x^2 dx = \frac{V}{\lambda^3}$$

where we substituted $x^2 = \epsilon/k_B T$ and the integral is equal to $\sqrt{\pi}/4$. We have $n_Q(T) = 1/\lambda^3$ to be the quantum concentration.

4.2 Bosons and Fermions

If particles are non-interacting and $\lambda \sim (V/N)^{1/3}$ at sufficiently low temperature, then we need to consider quantum statistics. Consider N identical particles, which may be interacting with each other. Due to spin-statistics theorem, the quantum wavefunctions must obey the symmetry property (depending on the quantum statistics)

$$\Psi(\mathbf{x}^1, \dots, \mathbf{x}^i, \dots, \mathbf{x}^j, \dots, \mathbf{x}^N) = \pm \Psi(\mathbf{x}^1, \dots, \mathbf{x}^j, \dots, \mathbf{x}^i, \dots, \mathbf{x}^N)$$

Bosonic (integer spins) wavefunctions are even under exchange, whereas fermionic (half-integer spins) wavefunctions are odd under exchange. We can construct the basis of N -particle states Ψ from independent 1-particle states $\{\psi^1, \psi^2, \dots\}$. We do this by specifying the occupation numbers of N -particle states, constrained by the total number of particles N being fixed, i.e. $\sum_{\alpha} n^{\alpha} = N$. The occupancy for bosons and fermions satisfy respectively $n^{\alpha} \in \mathbb{Z}^+ \cup \{0\}$ and $n^{\alpha} \in \{0, 1\}$. The latter is due to Pauli exclusion principle:

Example 4.2 (Pauli exclusion principle). Consider the two-particle state constructed using ψ^{α} and ψ^{β} :

$$\Psi(\mathbf{x}^1, \mathbf{x}^2) = \psi^{\alpha}(\mathbf{x}^1)\psi^{\beta}(\mathbf{x}^2) \pm \psi^{\alpha}(\mathbf{x}^2)\psi^{\beta}(\mathbf{x}^1), \quad \alpha \neq \beta$$

and Ψ has the required symmetry in accordance to the spin-statistics theorem. Let the occupation numbers be $n^{\alpha} = 1$ and $n^{\beta} = 1$. If $\alpha = \beta$, $\Psi(\mathbf{x}^1, \mathbf{x}^2) = \psi^{\alpha}(\mathbf{x}^1)\psi^{\alpha}(\mathbf{x}^2)$. This is okay for bosons since $n^{\alpha} = 2$ is allowed. But, no state with $n^{\alpha} = 2$ is allowed for fermions.

Example 4.3. Consider two non-interacting quantum particles in a box, in states $|\mathbf{k}_1\rangle$ and $|\mathbf{k}_2\rangle$ with positions \mathbf{r}_1 and \mathbf{r}_2 . Since the particles are indistinguishable, then the state $\Psi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1, \mathbf{r}_2) = e^{i\mathbf{k}_1 \cdot \mathbf{r}_1} e^{i\mathbf{k}_2 \cdot \mathbf{r}_2}$ is different from $\Psi_{\mathbf{k}_2, \mathbf{k}_1}(\mathbf{r}_1, \mathbf{r}_2) = e^{i\mathbf{k}_2 \cdot \mathbf{r}_1} e^{i\mathbf{k}_1 \cdot \mathbf{r}_2}$. The wavefunction is

$$\Psi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \left(e^{i\mathbf{k}_1 \cdot \mathbf{r}_1} e^{i\mathbf{k}_2 \cdot \mathbf{r}_2} |s_1; s_2\rangle \pm e^{i\mathbf{k}_2 \cdot \mathbf{r}_1} e^{i\mathbf{k}_1 \cdot \mathbf{r}_2} |s_2; s_1\rangle \right)$$

where the plus sign refers to bosons, and the minus to fermions. The joint probability density $P_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1, \mathbf{r}_2)$ for the state $\Psi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1, \mathbf{r}_2)$ of finding one particle at \mathbf{r}_1 and the other at \mathbf{r}_2 is

$$\frac{1}{2} (1 + 1 \pm \delta_{s_1, s_2} (e^{-i\mathbf{k}_1 \cdot (\mathbf{r}_1 - \mathbf{r}_2)} e^{-i\mathbf{k}_2 \cdot (\mathbf{r}_2 - \mathbf{r}_1)} + e^{-i\mathbf{k}_2 \cdot (\mathbf{r}_1 - \mathbf{r}_2)} e^{-i\mathbf{k}_1 \cdot (\mathbf{r}_2 - \mathbf{r}_1)})) = (1 \pm \delta_{s_1, s_2} \cos[(\mathbf{k}_1 - \mathbf{k}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2)])$$

which is called the pair-correlation function.

Using the canonical ensemble framework, the partition function is

$$Z = \sum_{\{n_r\}} e^{-\beta \sum_r n_r E_r}$$

where n_r is the number of particles in the state $|r\rangle$, with $\sum_r n_r = N$ being the total number of particles. Imposing quantum statistics on this sum is tricky, so it is much easier if we use a chemical potential μ and allow the total number of particles N to fluctuate, i.e. grand canonical ensemble framework. Treat each one-particle energy level to be in contact with a heat and particle reservoir of temperature T and chemical potential μ . There is no need to constrain total E and N . Essentially, instead of treating the particles as independent thermodynamic systems, we treat the energy levels as independent thermodynamic systems, which contain particles and energy.

Proposition 4.1: The macroscopic grand partition functions for bosons and fermions respectively are:

$$\mathcal{Z}_B = \frac{1}{1 - e^{n\beta(\mu - E)}}, \quad \mathcal{Z}_F = 1 + e^{n(\mu - E)\beta}$$

Proof. For bosons, consider state $|r\rangle$ which may be populated by an arbitrary number of particles:

$$\mathcal{Z}_r^{\text{boson}} = \sum_{n_r} e^{-\beta n_r (E_r - \mu)} = \frac{1}{1 - e^{-\beta(E_r - \mu)}}, \quad n_r \in \mathbb{Z}^+ \cup \{0\}$$

The sum does converge if $E_r - \mu > 0$ which is true $\forall E_r$. But, the ground state is $E_0 = 0$, so we require $\mu < 0$. On the other hand, for fermions, $|r\rangle$ is either occupied or not:

$$\mathcal{Z}_r^{\text{fermion}} = \sum_{n_r=0,1} e^{-\beta n_r (E_r - \mu)} = 1 + e^{-\beta(E_r - \mu)}, \quad n_r \in \{0, 1\}$$

There is no convergence issue in defining \mathcal{Z}_r , so there is no constraint on the sign of μ . \square

Remarks 4.2. We are summing over the number of particles in a given state, rather than summing over the states for a single particle. For photons or phonons (examples of bosons), since the particle number is not conserved, we have to set $\mu = 0$ in order to use the grand canonical ensemble approach.

Proposition 4.2: The Bose-Einstein distribution and Fermi-Dirac distribution are respectively

$$n_B(E) = \frac{1}{e^{\beta(E-\mu)} - 1}, \quad n_F(E) = \frac{1}{e^{\beta(E-\mu)} + 1}$$

Proof. Using results from the grand canonical ensemble: the mean occupation number for bosons and fermions are respectively:

$$\begin{aligned} n_B(E) &= k_B T \frac{\partial}{\partial \mu} \log \mathcal{Z}_B = -k_B T \frac{\partial}{\partial \mu} \log(1 - e^{(\mu-E)/k_B T}) = \frac{1}{e^{(E-\mu)/k_B T} - 1} \\ n_F(E) &= k_B T \frac{\partial}{\partial \mu} \log \mathcal{Z}_F = k_B T \frac{\partial}{\partial \mu} (1 + e^{n(\mu-E)/k_B T}) = \frac{e^{(\mu-E)/k_B T}}{1 + e^{(\mu-E)/k_B T}} = \frac{1}{e^{(E-\mu)/k_B T} + 1} \end{aligned}$$

□

Corollary 4.1: The quantum gas distributions reduce to the Maxwell-Boltzmann distribution.

Proof.

$$n \approx e^{-(E-\mu)/k_B T}, \quad \beta(E-\mu) \gg 1$$

□

Example 4.4. The Maxwell-Boltzmann probability distribution is obtained by

$$P(E)dE = \frac{1}{N} e^{-\beta(E-\mu)} g(E)dE = \frac{1}{N} \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \frac{N\lambda^3}{V} e^{-\beta E} \sqrt{E} dE$$

which gives $P(v)dv = \left(\frac{m}{2\pi k_B T} \right)^{3/2} e^{-\beta m v^2/2} 4\pi v^2 dv$, as expected.

Remarks 4.3. As we will see later:

1. The Bose-Einstein distribution diverges for $E = \mu$, and meaningless for $E < \mu$, where the reservoir floods the energy levels. Hence, we required $\mu < E_0$, where E_0 is the ground state energy. μ is determined by total particle number N , or vice-versa. n_B decreases rapidly with E if T is small. We require $\mu \rightarrow E_0$ as $T \rightarrow 0$ for N to remain fixed. Most particles then occupy ground state E_0 and the energy states similar in energy.
2. $n_F = 0.5$ for $E = \mu$. The steepness of $n_F(E)$ increases as $T \rightarrow 0$. The levels are mostly filled for $E < \mu$ and empty for $E > \mu$. μ can take any value. The Fermi gas becomes degenerate at $T = 0$. All states occupied up to energy $E_F = \mu$ (Fermi Energy) and empty above.

Proposition 4.3: Suppose we have a quasi-continuous density of 1-particle energy levels $g(E)$. The total particle number and total energy are respectively

$$N = \int_{E_0}^{\infty} \frac{g(E)dE}{e^{\beta(E-\mu)} \mp 1}, \quad E = \int_{E_0}^{\infty} \frac{Eg(E)dE}{e^{\beta(E-\mu)} \mp 1}$$

The negative branch is for bosons and positive branch is for fermions.

Remarks 4.4. In general, the two integrals have no closed form, but we can simplify it under certain regimes (as we will see later). For a gas of non-relativistic particles, $E = \frac{p^2}{2m}$, the density of states $g(E)$ is given to be

$$g(E)dE = g_s \frac{V}{(2\pi\hbar)^3} d^3p = \frac{g_s V 4\pi p^2 dp}{(2\pi\hbar)^3} = \frac{g_s V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2} dE := K' E^{1/2} dE$$

where g_s is the degeneracy of the particles.

Proposition 4.4: $PV = \frac{2}{3}E$ is true for bosons and fermions.

Proof. Φ is extensive and a function of T , μ and V , i.e. $\Phi = Vf(T, \mu)$. Recall $P = -\frac{\partial \Phi}{\partial V}|_{T, \mu} = -f(T, \mu)$ and for boson/fermion gases, the grand potential Φ is

$$\Phi = -T \log \mathcal{Z} = -T \int g(E) \log(1 \pm \xi e^{-E/T}) dE = -TK' \int E^{1/2} \log(1 \pm \xi e^{-E/T}) dE$$

where $\xi = e^{\mu/T}$ is defined as the fugacity. Perform an integration by parts:

$$PV = -\Phi = T \log \mathcal{Z} = -TK' \int E^{1/2} \log(1 \pm \xi e^{-\beta E}) dE = K' \frac{2}{3} \int \frac{E^{3/2} dE}{\xi^{-1} e^{\beta E} \pm 1} = \frac{2}{3} E$$

□

Definition 4.1 (Gamma function): The Gamma function is defined to be

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt$$

with the familiar results:

$$\Gamma(n) = (n-1)!, \quad n \in \mathbb{Z}^+, \quad \Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(3/2) = \frac{\sqrt{\pi}}{2}, \quad \Gamma(5/2) = \frac{3\sqrt{\pi}}{4}$$

Proposition 4.5: Let $x = \beta E$ and recall the thermal wavelength to be $\lambda = \sqrt{2\pi\hbar^2\beta/m}$. Then, $\frac{N}{V}$ and $\frac{P}{T}$ of fermions/bosons can be elegantly written as

$$\frac{N}{V} = \frac{g_s}{\lambda^3} f_{3/2}^{F/B}(\xi), \quad \frac{P}{T} = \frac{g_s}{\lambda^3} f_{5/2}^{F/B}(\xi)$$

where

$$f_\nu^F(\xi) := \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}}{\xi^{-1} e^x + 1} dx, \quad f_\nu^B(\xi) := \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}}{\xi^{-1} e^x - 1} dx$$

where $\Gamma(\nu)$ is the Gamma function.

Proof. use the definition of $f_{3/2}^{F/B}(\xi)$, $f_{5/2}^{F/B}(\xi)$ and $\Gamma(3/2) = \sqrt{\pi}/2$, $\Gamma(5/2) = 3\sqrt{\pi}/4$:

$$\frac{N}{V} = \int_0^\infty \frac{g(E) dE}{\xi^{-1} e^{\beta E} \mp 1} = \frac{g_s}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{E^{1/2} dE}{\xi^{-1} e^{\beta E} \mp 1} = \frac{g_s}{4\pi^2} \left(\frac{2m}{\hbar^2 \beta} \right)^{3/2} \int_0^\infty \frac{x^{1/2} dx}{\xi^{-1} e^x \mp 1} = \frac{g_s}{\lambda^3} f_{3/2}^{F/B}(\xi)$$

$$\frac{E}{V} = \int_0^\infty \frac{E g(E) dE}{\xi^{-1} e^{\beta E} \mp 1} = \frac{g_s}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{E^{3/2} dE}{\xi^{-1} e^{\beta E} \mp 1} = \frac{g_s}{4\pi^2} \left(\frac{2m}{\hbar^2 \beta} \right)^{3/2} \frac{1}{\beta} \int_0^\infty \frac{x^{3/2} dx}{\xi^{-1} e^x \mp 1} = \frac{g_s}{\lambda^3} \frac{3}{2\beta} f_{5/2}^{F/B}(\xi)$$

But $PV = \frac{2}{3}E$ for quantum gases, so $\frac{P}{T} = \frac{g_s}{\lambda^3} \frac{2}{3} \frac{3}{2} f_{5/2}^{F/B}(\xi)$. □

Remarks 4.5. When $\xi = 1$, we obtain the Riemann Zeta function

$$f_\nu^B(\xi = 1) = \zeta(\nu) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}}{e^x - 1} dx$$

In general, it is tricky to obtain analytical solutions, hence we shall only consider limiting cases.

4.3 Ideal Fermi Gas

Proposition 4.6: Fermionic particles repel in a low density fermionic gas.

Proof. For low density, expand the Fermi-Dirac function for small ξ :

$$\frac{1}{\xi^{-1}e^x + 1} = \xi e^{-x}(1 - \xi e^{-x} + \dots) \implies f_\nu^F(\xi) = \frac{1}{\Gamma(\nu)} \int_0^\infty x^{\nu-1} (\xi e^{-x} - \xi^2 e^{-2x} + \dots) dx = \xi - \frac{\xi^2}{2^\nu} + \dots$$

This gives

$$\frac{N}{V} = \frac{g_s}{\lambda^3} \left(\xi - \frac{\xi^2}{2^{3/2}} + \dots \right), \quad \frac{P}{T} = \frac{g_s}{\lambda^3} \left(\xi - \frac{\xi^2}{2^{5/2}} + \dots \right)$$

For dilute gas (low density) $\frac{N}{V} \ll \frac{1}{\lambda^3}$, so $\xi \ll 1$. The leading result is $\frac{P}{T} = \frac{N}{V}$ as expected. The next leading term is obtained by eliminating ξ :

$$P = \frac{NT}{V} \left(1 + \frac{1}{2^{5/2}} \frac{\lambda^3}{g_s} \frac{N}{V} \right) = \frac{NT}{V} \left(1 + \frac{\pi^{3/2} \hbar^3}{2m^{3/2} T^{3/2} g_s} \frac{N}{V} \right)$$

The second Virial coefficient $B(T)$ (we will see later) is directly proportional to $T^{-3/2}$. This is small for large T , so $B(T) > 0$, so pressure increase, i.e. fermions repel. \square

Proposition 4.7: Chemical potential is negative for a dilute fermionic gas.

Proof. $\xi = e^{\mu/T} \approx \frac{\lambda^3 N}{g_s V} \ll 1$ for dilute gas, hence $\mu \approx T \log \frac{\lambda^3 N}{g_s V} < 0$ and $n_F(\xi)$ is small for $\xi \geq 0$. As T decreases, λ increases, so μ increases and becomes positive. \square

Remarks 4.6. To keep N fixed, we have to adjust μ as the temperature changes. As the temperature is increased, the chemical potential (originally positive) shifts to lower values, and ultimately becomes negative.

Example 4.5. ^3He is well described by the ideal Fermi gas. It is gaseous down to $T = 3\text{K}$. The atom has two protons, 1 neutron, 2 electrons, so it is a fermion of spin $\frac{1}{2}$. The mass m is small, so λ is relatively large for $T \sim 10\text{ K}$, i.e. quantum effects are significant.

Note 4.1 (Fermi sphere). For electronic gas, the degeneracy is $g_s = 2$ since electrons have spin- $1/2$. At $T = 0$, the Fermi-Dirac distribution is a step function, i.e. energy levels $E < \mu$ is filled while $E > \mu$ is empty. μ is called the Fermi energy E_F . Each fermion we add into the system settle into the lowest available energy state. These are successively filled until we run out of particles. the energy of the last filled state is called the Fermi energy, E_F where μ at $T = 0$ is E_F . Note that μ and T are independent variables in the grand canonical ensemble.

The energy states of the free particles are not localized in position space, but localized in momentum space. Successive fermions sit in states with ever-increasing momentum. Fermions fill out a ball in momentum space with the momentum of the final fermion called the Fermi momentum $k_F = \sqrt{2mE_F}/\hbar$. All states with wavevector $k \leq k_F$ are filled and are said to form the Fermi sphere.

Proposition 4.8: At $T = 0$, electronic gas is degenerate and exhibit degeneracy pressure.

Proof. For a fixed total particle number N and $g_s = 2$, the Fermi momentum satisfy

$$N = \frac{V}{\pi^2 \hbar^3} \int_0^{p_F} p^2 dp = \frac{V p_F^3}{3\pi^2 \hbar^3}$$

For non-relativistic particles, $E_F = \frac{p_F^2}{2m}$, so $N = \frac{V}{3\pi^2 \hbar^3} (2mE_F)^{3/2} \implies E_F = \frac{\hbar^2}{2m} (3\pi^2)^{2/3} \frac{N^{2/3}}{V^{2/3}}$. The result for total particle number is also valid for highly relativistic particles, $E_F = cp_F$. The total energy is

$$E = \frac{V}{\pi^2 \hbar^3} \int_0^{p_F} \frac{p^2}{2m} p^2 dp = \frac{V}{10\pi^2 m \hbar^3} p_F^5 = \frac{V}{10\pi^2 m \hbar^3} (2mE_F)^{5/2} \implies \frac{E}{N} = \frac{3}{10m} (2mE_F) = \frac{3}{5} E_F$$

The pressure is given by

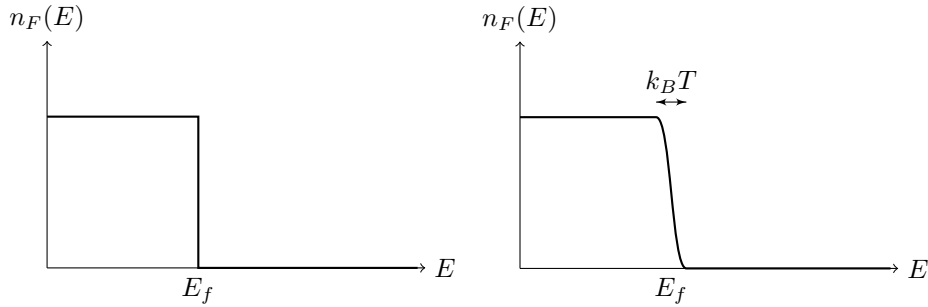
$$PV = \frac{2}{3} E = \frac{2}{5} N E_F \implies P = \frac{\hbar^2}{5m} (3\pi^2)^{2/3} \frac{N^{5/3}}{V^{5/3}} > 0$$

This is the degeneracy pressure. \square

Remarks 4.7. The repulsion of particles in Fermi gas produces pressure, even when $T = 0$ due to Pauli Exclusion Principle. This effect is important for stabilizing white dwarf stars (degenerate electron gas) and neutron stars against gravitational collapse (i.e. becoming black hole).

Proposition 4.9: For metals, the electronic contribution to the heat capacity dominates at low temperatures.

Proof. The density of states $g(E)$ is independent of temperature T . For small T , the slope of n_F at $E = \mu$ is $O(1/T)$ large. Between temperatures 0 and T , the number of particles that move is $O(g(E_F)k_B T)$. As $g(E)$ is approximately constant around E_F , μ does not need to change much to keep N fixed. The increase in total energy between temperatures 0 and T is $O(T^2 g(E_F) k_B^2)$, i.e. $E(T) = E(0) + O(g(E_F) T^2 k_B^2) \implies C = O(g(E_F) T k_B^2)$. \square



Remarks 4.8.

1. A more rigorous proof to the above is via the Sommerfield expansion.

$$\int_0^\infty \frac{g(E)dE}{e^{\beta(E-\mu)} + 1} \approx \int_0^{E_F} g(E)dE + \frac{\pi^2}{6} k_B^2 T^2 g'(E_F) + \frac{7\pi^4}{360} k_B^4 T^4 g^{(3)}(E_F) + \dots$$

To get the grand potential Φ to second order in $k_B T$, we substitute $g(E) = E^{3/2}$ to obtain

$$\Phi(T) = \Phi(0) - \frac{V k_B^2 T^2}{12} \left(\frac{2m}{\hbar^2} \right)^{3/2} E_F^{1/2}$$

and hence the heat capacity at constant volume is

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_{V,\mu} = -T \left(\frac{\partial^2 \Phi}{\partial T^2} \right)_{V,\mu} = T \frac{\partial}{\partial T} \frac{k_B^2 T^2}{3} \frac{V}{2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E_F} = \frac{\pi^2}{3} g(E_F) k_B^2 T$$

2. In solid metal, the contributions to C are $O(T)$ from non-degenerate electrons and $O(T^2)$ from lattice vibrations. The electronic contribution to the heat capacity dominates at low temperatures. The Fermi temperature T_F is very large for metals, typically 50,000 K.
3. We have the density of states to be $g(E) \propto \sqrt{E}$ and so the number is $N(E) = \int g(E)dE \propto \frac{2}{3} E^{3/2}$, and so $g(E_F) = \frac{3N_{tot}}{2E_F}$.

4.4 Ideal Bose Gas

Proposition 4.10: Bosonic particles attract in a low density bosonic gas.

Proof. Set $g_s = 1$. For low density, expand Bose-Einstein function for small ξ :

$$\frac{N}{V} = \frac{1}{\lambda^3} f_{3/2}^B(\xi) = \frac{1}{\lambda^3} \left(\xi + \frac{\xi^2}{2^{3/2}} + \dots \right), \quad \frac{P}{T} = \frac{1}{\lambda^3} f_{5/2}^B(\xi) = \frac{1}{\lambda^3} \left(\xi + \frac{\xi^2}{2^{5/2}} + \dots \right)$$

At low density, $\xi \ll 1$. Eliminating ξ gives

$$P = \frac{NT}{V} \left(1 - \frac{\lambda^3}{2^{5/2}} \frac{N}{V} + \dots \right)$$

The second Virial coefficient (as we will see later) $B(T) = O(T^{-3/2})$ and is negative, i.e. Bosons are attractive. \square

Definition 4.2 (Bose-Einstein Condensation): Bose-Einstein condensation is a condensate in momentum/energy space where a macroscopic number of particles occupy the ground state.

Proposition 4.11: At a critical temperature T_c , we will obtain Bose-Einstein condensation.

Proof. Recall $\frac{N}{V} = \frac{1}{\lambda^3} f_{3/2}^B(\xi)$. As T decreases, $\frac{1}{\lambda^3}$ decreases, so ξ must increase to keep $\frac{N}{V}$ fixed. After ξ reaches to 1, this equation breaks down (see f_{ν}^B integral). The critical temperature T_c is given by

$$\frac{N}{V} = \frac{1}{[\lambda(T = T_c)]^3} f_{3/2}^B(1) = \left(\frac{mk_B T_c}{2\pi\hbar^2} \right)^{3/2} \zeta(1.5) \implies T_c = \frac{2\pi\hbar^2}{k_B m} \left(\frac{1}{\zeta(1.5)} \frac{N}{V} \right)^{2/3}$$

where $\zeta(1.5) = 2.612$. T_c is called the Bose-Einstein Condensation temperature, related to the number density $\frac{N}{V}$. Naively, we see that as we reduce T below T_c , the number of particles N should decrease. This is erroneous since in approximating the finite sum $\sum_k \rightarrow \frac{V(2m)^{3/2}}{4\pi^2\hbar^2} \int E^{1/2} dE$, $E = 0$ does not contribute to the integral. The ‘missing’ states have $E = 0$ which does not contribute to the integral. Using the Bose-Einstein distribution, the ground state occupancy is predicted to be

$$n_B(E = 0) = \frac{1}{\xi^{-1} - 1}$$

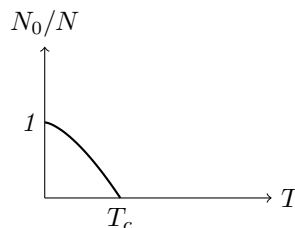
which diverges at $\xi = 0$. As $T \rightarrow T_c$ from above, one crudely approximate by setting $\xi \approx 1 - (1/N)$ so $n_B = O(N)$, i.e. the ground state have macroscopic occupation. \square

Remarks 4.9.

1. At low temperatures, $\mu \approx 0$ and it becomes more negative as temperature increases. $\mu \geq 0$ cannot be true, otherwise the occupation number at $E = \mu$ would become infinite, contradicting our assumption on fixed total N .
2. The correct approach for $T \leq T_c$: separate ground state from the quasi-continuum of positive energy states. Then,

$$\frac{N}{V} = \frac{N_0}{V} + \frac{1}{\lambda^3} f_{3/2}^B(\xi = 1) \implies \frac{N_0}{V} = \frac{N}{V} - 2.612 \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \implies \frac{N_0}{N} = 1 - \left(\frac{T}{T_c} \right)^{3/2}$$

This gives us the fraction of particles in the ground state. At $T = 0$, all particles will be in the ground state.



3. The pressure is given by

$$\frac{P}{T} = \frac{g_s}{\lambda^3} f_{5/2}^B(\xi) = \frac{g_s}{\lambda^3} \zeta(5/2)$$

so P is directly proportional to $T^{5/2}$ for $T \leq T_c$. Earlier, $PV = \frac{2}{3}E$, so E is directly proportional to $VT^{5/2}$. This is true for $T < T_c$ and $\xi \approx 1$.

4. At low temperatures and below T_c , set $\mu = 0$, we have

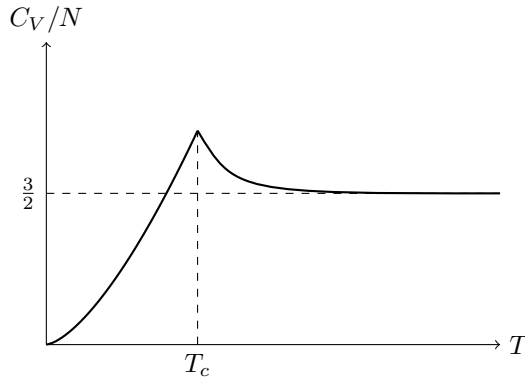
$$U = \frac{2Vk_B T}{\sqrt{\pi}\lambda^3} \int_0^\infty \frac{x^{3/2} dx}{e^x - 1} \approx \frac{2V}{\lambda^3} k_B T \propto T^{5/2}$$

and so the heat capacity is

$$C_V = \left. \frac{\partial E}{\partial T} \right|_V = \frac{5}{2} \frac{k_B V}{\sqrt{2}\lambda^3} \propto T^{3/2}$$

For large T , estimate C_V using virial expansion (see later):

$$E = \frac{3}{2}PV = \frac{3}{2}NT \left(1 - O\left(\frac{N}{V} \frac{1}{T^{3/2}}\right) \right) \Rightarrow C_V = \left. \frac{\partial E}{\partial T} \right|_V = \frac{3}{2}N \left(1 + O\left(\frac{N}{V} \frac{1}{T^{3/2}}\right) \right)$$



C_V is continuous but have a discontinuous derivative at $T = T_c$. We call this a phase transition, and that T_c depends on $\frac{N}{V}$.

Example 4.6. ^4He undergoes transition from normal liquid to a superfluid (related to Bose Einstein condensation, but interactions are really quite strong). Using the density of normal liquid at pressure 1 atm, we predict the Bose Einstein transition temperature $T_c = 3.13$ K. Experimentally, we see a discontinuity in C_V at $T = 2.19$ K. The experimental measurements show a ‘ λ ’ transition at T_c . This is because in our calculations, interactions are not accounted for. Better examples of BEC are dilute gases of heavy atoms in harmonic traps, like Rubidium-87 which have a $T_c \sim 10^{-7}$ K with a macroscopic occupation of $N \sim 10^4$ to 10^7 , i.e. density of 2.5×10^{12} atoms per cubic centimetres.

Note 4.2. The exact description of interacting particles is very difficult. Commonly, we can take a system which is close to its ground state and subject it to a weak external influence, and are interested in the small changes to the system caused by the external influence. These are the excited states or excitations of the system, which interact only weakly with one another. We thus approximate them as non-interacting ideal gases. We interpret their low temperature properties in terms of approximate normal modes, or ‘elementary excitations’. These are characterized by the zero rest mass (infinite de Broglie wavelength, hence always in the quantum regime) and are always bosons (most have no spins).

Example 4.7 (Spin waves). The elementary excitations of a ferromagnetic systems, in which the size of the moments is fixed, are ‘spin waves’. The simplest dispersion relation is $E_k \propto k^2$. The result will be $C \propto T^{3/2}$, which agrees with experimental measurements.

Below, we consider two other examples: photons and phonons.

4.5 Blackbody radiation

Definition 4.3 (Blackbody radiation): Blackbody radiation can be thought of as an ideal gas of photons, with integer angular momentum and obey Bose statistics. To come into equilibrium, they have to interact with a reservoir, which is taken to be the walls of the blackbody cavity.

Proposition 4.12: The chemical potential of blackbody radiation is zero.

Proof. Thermal equilibrium is reached by absorption and emission of photons by the walls, so we cannot take N , the number of photons in the gas, to be constant. μ is obtained by minimizing the Helmholtz free energy, with respect to the variations in the average number of photons, i.e.

$$\left(\frac{\partial F}{\partial N}\right)_{T,V} = 0$$

But this is actually μ by definition of free energy. In this case, the grand potential and the free energy is the same. \square

Proposition 4.13: The Planck distribution $E(\omega)$ is given by

$$E = \int_0^\infty E(\omega) d\omega, \quad E(\omega) = \frac{V\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta\hbar\omega} - 1}$$

Proof. Consider electromagnetic radiation filling a volume $V = L^3$ at temperature T . The radiation modes are $\mathbf{A}e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$ with $\omega = c|\mathbf{k}|$. Again, enforce periodic boundary conditions are $k_x = \frac{2\pi n_x}{L}$, etc. There are two transverse polarization states which are the two directions transverse to the direction of propagation, so the degeneracy $g_s = 2$. Hence, the density of modes in \mathbf{k} -space is $2\frac{V}{(2\pi)^3}$. The number of modes between k and $k + dk$ is

$$g(k)dk = 2\frac{V}{(2\pi)^3}4\pi k^2 dk = \frac{V k^2}{\pi^2} dk$$

Correspondingly, the number of modes between ω and $\omega + d\omega$ is

$$g(\omega)d\omega = \frac{V\omega^2}{\pi^2 c^3} d\omega$$

where we used $\omega = ck$, i.e. relativistic. We start by summing photons with a definite frequency ω . A state with N such photons has energy $E = N\hbar\omega$.

$$Z_\omega = \sum_{N=0}^{\infty} e^{-\beta N\hbar\omega} = \frac{1}{1 - e^{-\beta\hbar\omega}}$$

Sum over all possible frequencies:

$$\ln Z = \int_0^\infty g(\omega) \ln Z_\omega d\omega = -\frac{V}{\pi^2 c^3} \int_0^\infty \omega^2 \ln(1 - e^{-\beta\hbar\omega}) d\omega$$

Using the result from the canonical ensemble, the energy is

$$E = -\frac{\partial}{\partial \beta} \log Z = \frac{V\hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3 d\omega}{e^{\beta\hbar\omega} - 1}$$

Result follows by $E = \int_0^\infty E(\omega) d\omega$. \square

Corollary 4.2 (Wien's displacement law): The Planck distribution peaks at a frequency $\omega_{\max} = \xi T/\hbar$ where ξ satisfies $3 - \xi = 3e^{-\xi}$.

Proof. Find maximum for the Planck distribution $E(\omega)$. \square

Proposition 4.14 (Stefan-Boltzmann Law): The energy flux is given by σT^4 .

Proof. Substitute $x = \beta\hbar\omega$ in the E integral:

$$E = \frac{V}{\pi^2 c^3} \frac{T^4}{\hbar^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{VT^4}{\pi^2 c^3 \hbar^3} \frac{\pi^4}{15}$$

The energy flux is the rate of transfer of energy and can be shown to be $\frac{E}{V} \frac{c}{4}$. The flux of photons is obtained from

$$\frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi/2} c \cos \theta (\sin \theta d\theta) d\phi = \frac{c}{4}$$

where we consider photons travelling perpendicularly away from the object. The result follows. \square

Proposition 4.15: The equation of state of the photon gas is

$$P = \frac{1}{3} \frac{E}{V}$$

Proof. The free energy of a canonical ensemble is

$$\begin{aligned} F &= -T \log Z \\ &= \frac{VT}{\pi^2 c^3} \int_0^\infty \omega^2 \log(1 - e^{-\beta\hbar\omega}) d\omega \\ &= -\frac{V\hbar}{3\pi^2 c^3} \int_0^\infty \frac{\omega^3 e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} d\omega \\ &= -\frac{V\hbar}{3\pi^2 c^3 \beta^4 \hbar^4} \int_0^\infty \frac{x^3}{e^x - 1} dx \\ &= -\frac{V\pi^2 T^4}{45 \hbar^3 c^3} \end{aligned}$$

The pressure is then

$$P = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{E}{3V}$$

or sometimes written as $4\sigma T^4/3c$. This is the radiation pressure. \square

Remarks 4.10.

1. Alternatively, the equation of state can be worked out using the following argument: We have $k = \frac{\pi^2}{15\hbar^3 c^3}$. Write $E = kVT^4$. At constant V ,

$$dE = 4kVT^3 dT = TdS = TV f'(T) dT$$

where $S = Vf(T)$ is extensive. We must have $f'(T) = 4kT^2$ and so $f(T) = \frac{4}{3}kT^3$ since $S(T=0) = 0$ by Third Law. The entropy of blackbody radiation is $S = \frac{4}{3}kVT^3$. Now express E in terms of S and V . We have $E = kV(\frac{4}{3}kV)^{-4/3} S^{4/3}$ which is directly proportional to $V^{-1/3}$. The pressure is

$$P = -\left.\frac{\partial E}{\partial V}\right|_S \implies PV = \frac{1}{3}E$$

Hence, in an adiabatic change (slowly enlarge box with radiation inside), $PV^{4/3}$ is a constant and $TV^{1/3}$ is another constant. Hence, radiation cools as box expands.

2. We can also work out

$$S = -\left(\frac{\partial F}{\partial T}\right)_V = \frac{16V\sigma}{3c} T^3, \quad C_V = \left(\frac{\partial E}{\partial T}\right)_V = \frac{16V\sigma}{c} T^3$$

3. In the classical limit $\hbar\omega \ll T$, we have

$$\frac{1}{e^{\beta\hbar\omega} - 1} \approx \frac{1}{\beta\hbar\omega} \implies E(\omega) \approx \frac{V\omega^2 T}{\pi^2 c^3}$$

which is the Rayleigh-Jeans law for the classical distribution of radiation. This leads to the Ultraviolet catastrophe since $\int_0^\infty E(\omega) d\omega$ diverges.

4. Alternatively, we may compute the total number and energy of photon gas respectively to be:

$$N = \frac{V}{\pi^2 c^3} \int_0^\infty \frac{\omega^2 d\omega}{e^{\hbar\omega/T} - 1} = \frac{VT^3}{\pi^2 \hbar^3 c^3} \int_0^\infty \frac{x^2 dx}{e^x - 1} = \frac{2\zeta(3)T^3}{\pi^2 \hbar^3 c^3} V$$

$$E = \frac{V\hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3 d\omega}{e^{\hbar\omega/T} - 1} = \frac{VT^4}{\pi^2 \hbar^3 c^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^2 T^4}{15 \hbar^3 c^3} V$$

where $\zeta(\nu) = \int_0^\infty \frac{x^{\nu-1}}{e^x - 1} dx$, is the Riemann zeta function such that $\zeta(3) \approx 0.244$ and $\zeta(4) = \frac{\pi^4}{15}$.

4.6 Vibrational Energy of Solids

Definition 4.4 (Phonons): The atoms of a solid interact very strongly with one another. Sound waves in solids come in discrete packets, called phonons.

Consider an atomic crystal with N atoms such that the vibrational modes are $\mathbf{A}e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$. The degeneracy is 3, i.e. $g_s = 3$, with a total of one longitudinal and two transverse polarizations. The frequency ω is related to $\hbar|\mathbf{k}|$ in some complicated way. At long wavelengths, we have sound waves $\omega = v_s|\mathbf{k}|$ where v_s is an effective velocity ($3/v_s^3 = (1/v_L^3) + (2/v_T^3)$ where v_L and v_T are longitudinal and transverse velocities respectively). Assume small speed v_s , the number of modes between frequencies ω and $\omega + d\omega$ is

$$g(\omega)d\omega = \frac{3V\omega^2 d\omega}{2\pi^2 v_s^3}$$

Each mode of frequency ω can be excited to any n -phonon state of energy $n\hbar\omega$.

Proposition 4.16: There is an upper frequency limit, known as the Debye frequency, with $\omega_D = \left(\frac{6\pi^2 N}{V}\right)^{1/3} v_s$.

Proof. Sound waves cannot propagate through the solid if the wavelength of sound is smaller than the atomic spacing, so we have $\lambda = \frac{2\pi c_s}{\omega} < L$, where L is the length of the crystal. Since there is three directions of motion for each of the N atoms, then

$$3N = \int_0^{\omega_D} g(\omega)d\omega = \frac{V\omega_D^3}{2\pi^2 c_s^3}$$

□

Definition 4.5 (Debye temperature): The Debye temperature is the temperature at which the highest frequency phonon starts to become excited, i.e. $T_D = \hbar\omega_D$.

Proposition 4.17: The vibrational heat capacity scales with T^3 for low temperatures and is a constant of $3N$ for high temperatures. The latter is called Dulong-Petit law.

Proof. Since the number of phonons is not conserved, we perform a similar procedure as that of photons.

$$\log Z_{\text{phonon}} = \int_0^{\omega_D} g(\omega) \log Z_\omega d\omega, \quad Z_\omega = \frac{1}{1 - e^{-\beta\hbar\omega}}$$

The energy is then

$$E = \int_0^{\omega_D} \frac{\hbar\omega g(\omega)}{e^{\beta\hbar\omega} - 1} d\omega = \frac{3V\hbar}{2\pi^2 c_s^3} \int_0^{\omega_D} \frac{\omega^3 d\omega}{e^{\beta\hbar\omega} - 1} = \frac{3Vk_B T^4}{2\pi^2 \hbar^3 c_s^3} \int_0^{T_D/T} \frac{x^3}{e^x - 1} dx$$

where we substituted $x = \beta\hbar\omega$. This however has no analytical solution. For low temperatures, $T \ll T_D$, the upper integration limit approaches infinity. Like before, the integral is thus $\frac{\pi^4}{15}$. The heat capacity is then

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V = \frac{2\pi^2 VT^3}{5\hbar^3 c_s^3} = \frac{12\pi^4 N k_B}{5T_D^3} T^3$$

For high temperatures, $T \gg T_D$, we Taylor expand the integrand,

$$\int_0^{T_D/T} \frac{x^3}{e^x - 1} dx \approx \int_0^{T_D/T} x^2 + \dots dx = \frac{T_D^3}{3T^3} + \dots \implies C_V = \frac{Vk_B T_D^3}{2\pi^2 \hbar^3 c_s^3} = 3Nk_B$$

where we invoked the definition of Debye temperature, i.e. $T_D^3 = \hbar^3 \omega_D^3 = \hbar^3 (3N) 2\pi^2 c_s^3 / V$ □

5 Real Gases and Liquids

5.1 Virial theorem

Definition 5.1 (Virial): The classical virial \mathcal{V} is defined as

$$\mathcal{V} := -\frac{1}{2} \sum_i \mathbf{r}_i \cdot \mathbf{f}_i$$

where \mathbf{r}_i and \mathbf{f}_i are respectively, the position of, and the force acting on, the i th particle. From newton's second law, we can write as

$$\mathcal{V} = -\frac{1}{2} \sum_i m_i \frac{d}{dt}(\mathbf{r}_i \cdot \mathbf{v}_i) + \sum_i \frac{1}{2} m_i \mathbf{v}_i^2$$

Example 5.1. Consider a liquid consisting of N particles in a container of volume V . The value of $\mathbf{r}_i \cdot \mathbf{v}_i$ can only fluctuate between finite limits and if we average it over a long period its time derivative must average to zero, which leads to

$$\langle \mathcal{V} = \sum_i \frac{1}{2} m_i \langle v_i^2 \rangle \rangle = \frac{3}{2} N k_B T$$

The mean virial is equal to the mean kinetic energy, which is Clausius' version of virial theorem.

The virial may separate external and internal contributions.

Theorem 5.1 (Virial theorem): Assuming an internal interaction potential ϕ between two particles i and j , that depends only on their separation $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, then the virial theorem is

$$3pV = 2\langle K \rangle - \left\langle \sum_{j>i} r_{ij} \frac{d\phi(r_{ij})}{dr_{ij}} \right\rangle$$

where K is the kinetic energy.

Proof. For the external virial, the force on an element of the wall is $p d\mathbf{A}$. The force on the particles acts inwards and $d\mathbf{A}$ is directed outwards, so we have

$$\langle \mathcal{V}_{\text{ext}} \rangle = \left\langle -\frac{1}{2} \sum_i \mathbf{r}_i \cdot \mathbf{f}_{\text{ext},i} \right\rangle = \frac{1}{2} \int_A p \mathbf{r} \cdot d\mathbf{A} = \frac{p}{2} \int_V \nabla \cdot \mathbf{r} dV = \frac{3}{2} pV$$

For the given interaction potential, the force on particle i from the other $j \neq i$ particles is

$$\mathbf{f}_{\text{int},i} = \sum_{j \neq i} -\frac{\partial \phi(r_{ij})}{\partial \mathbf{r}_i}$$

and using the result:

$$\mathbf{r}_i \cdot \frac{\partial \phi(r_{ij})}{\partial \mathbf{r}_i} = \mathbf{r}_i \cdot \frac{\partial r_{ij}}{\partial \mathbf{r}_i} \frac{\partial \phi(r_{ij})}{\partial r_{ij}} = \mathbf{r}_{ij} \cdot \frac{1}{2r_{ij}} 2(\mathbf{r}_i - \mathbf{r}_j) \frac{\partial \phi(r_{ij})}{\partial r_{ij}}$$

The total internal virial becomes

$$\mathcal{V}_{\text{int}} = \frac{1}{2} \sum_i \sum_{j \neq i} \mathbf{r}_i \cdot \frac{(\mathbf{r}_i - \mathbf{r}_j)}{r_{ij}} \frac{\partial \phi(r_{ij})}{\partial r_{ij}} = \frac{1}{4} \sum_i \sum_{j \neq i} r_{ij} \frac{\partial \phi(r_{ij})}{\partial r_{ij}} = \frac{1}{2} \sum_{j>i} r_{ij} \frac{\partial \phi(r_{ij})}{\partial r_{ij}}$$

where we have noted that we can pair together the terms containing $r_{\alpha\beta}$ and $r_{\beta\alpha}$ and use

$$\mathbf{r}_\alpha \cdot (\mathbf{r}_\alpha - \mathbf{r}_\beta) + \mathbf{r}_\beta \cdot (-\mathbf{r}_\alpha + \mathbf{r}_\beta) = r_{\alpha\beta}^2$$

Upon averaging, we obtain the mean internal virial:

$$\langle \mathcal{V}_{\text{int}} \rangle = \frac{1}{2} \left\langle \sum_{j>i} r_{ij} \frac{\partial \phi(r_{ij})}{\partial r_{ij}} \right\rangle$$

The final result is

$$\langle K \rangle = \langle \mathcal{V}_{\text{ext}} \rangle + \langle \mathcal{V}_{\text{int}} \rangle = \frac{3}{2} pV + \frac{1}{2} \left\langle \sum_{j>i} r_{ij} \frac{\partial \phi(r_{ij})}{\partial r_{ij}} \right\rangle$$

□

Remarks 5.1. Just from $\langle \mathcal{V}_{ext} \rangle = \frac{3}{2}pV$ alone, we may obtain the ideal gas law from the Clausius' version of Virial's theorem.

Example 5.2. Consider a particle moving in a potential $\phi(r) = A/r$, then $r \frac{d\phi}{dr} = -\phi$, and as there is no external virial and hence no pressure, we have

$$\langle K \rangle = -\frac{1}{2}\langle V \rangle$$

where V is the potential energy.

5.2 Generic interacting gas/liquid

Definition 5.2 (Configurational partition function (for interaction potential)): Consider a system of N classical particles of momenta \mathbf{p}_i and positions \mathbf{r}_i where they interact with one another via a potential $\phi(r_{ij})$ which only depend on the separation of point-like particles, then

$$Z_\phi := \int \exp \left(- \sum_{j>i} \beta \phi(r_{ij}) d^3 r_1 \dots d^3 r_N \right)$$

where $\sum_{j>i}$ sums over all particle pairs.

Remarks 5.2. The full Hamiltonian has the form

$$H = \sum_i \frac{p_i^2}{2m} + \sum_{j>i} \phi(\mathbf{r}_{ij})$$

The full partition function is

$$Z = \sum e^{-\beta H(\mathbf{r}, \mathbf{p})} = \frac{1}{N!} \int e^{-\beta H} d^3 r_1 \dots d^3 r_N \frac{d^3 p_1 \dots d^3 p_N}{(2\pi\hbar)^{3N}} = \frac{1}{N!} \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3N/2} Z_\phi$$

Example 5.3 (Lennard-Jones potential). One example of an interaction potential is the Lennard-Jones inter-atomic potential:

$$\phi(r) = 4\epsilon \left[\left(\frac{r_0}{r} \right)^{12} - \left(\frac{r_0}{r} \right)^6 \right]$$

This potential consists of a short ranged hard-core repulsion and a long ranged van der Waals attraction, which decays as $1/r^6$.

Definition 5.3 (Two-particle distribution function): The probability density of finding a particle at \mathbf{r}_1 and another at \mathbf{r}_2 is

$$P_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{N(N-1)}{Z_\phi} \int \exp \left(- \sum_{j>i} \beta \phi(r_{ij}) \right) d^3 r_3 \dots d^3 r_N$$

Definition 5.4 (Radial distribution function): The radial distribution function is

$$g(r_{12}) = \frac{V^2}{N^2} P_2(\mathbf{r}_1, \mathbf{r}_2)$$

where P_2 depends only on radial separation $|\mathbf{r}_1 - \mathbf{r}_2|$.

Example 5.4. Typically, the radial distribution function is small at short distances because of the hard-core repulsion between the atoms. It rises to a peak at the distance of the first shell of neighbours. There is a smaller peak at the second shell but at large distances $g(r)$ goes to unity, which is the value it would have in the absence of interactions (so that $P_2(r) \rightarrow (N/V)^2$). We can understand roughly how $g(r)$ will change when T and V are altered. If the system is compressed we expect the peaks of $g(r)$ to move to smaller separations, and if T is increased we expect the peaks to become less pronounced as the thermal motion smears them out.

Remarks 5.3. Radial distribution functions can be measured using X-ray or neutron diffraction techniques. They can also be calculated using Monte Carlo or Molecular Dynamics methods. Monte Carlo methods are statistical techniques for generating ensembles of configurations, $\{\mathbf{r}_i\}$, distributed according to a known probability distribution. In classical Molecular Dynamics methods the molecules are moved according to Newton's laws in what amounts to a 'computational experiment'. An important modern development is the use of computational quantum mechanical techniques to calculate the interaction energy, which can readily be combined with the classical theory of liquids developed here.

Proposition 5.1:

$$U = \frac{3}{2}Nk_B T + \frac{N^2}{2V} \int_0^\infty \phi(r)g(r)4\pi r^2 dr$$

Proof. The internal energy is

$$U = -\frac{\partial}{\partial \beta} \ln Z = \frac{3}{2}Nk_B T - \frac{\partial}{\partial \beta} \ln Z_\phi$$

where

$$-\frac{\partial}{\partial \beta} \ln Z_\phi = \frac{1}{Z_\phi} \int \sum_{j>i} \phi(r_{ij}) e^{-\sum_{j>i} \beta \phi(r_{ij})} d^3 r_1 \dots d^3 r_N = \frac{N(N-1)}{2Z_\phi} \int \phi(r_{12}) e^{-\sum_{j>i} \beta \phi(r_{ij})} d^3 r_1 \dots d^3 r_N$$

since there are $N(N-1)/2$ terms in the sum. Substitute in the definitions, we obtain our desired result. \square

Remarks 5.4. Unfortunately, T , V and N are not the natural variables of U . So even if we knew $\phi(r)$ and had $g(r)$ as a function of T , V and N , we still do not have enough information to do the complete thermodynamics of liquids.

Corollary 5.1:

$$p = \frac{Nk_B T}{V} - \frac{N^2}{6V^2} \int_0^\infty r \frac{d\phi}{dr} g(r) 4\pi r^2 dr$$

Proof. Use Virial theorem and evaluate $\langle \sum_{j>i} r_{ij} \frac{d\phi(r_{ij})}{dr_{ij}} \rangle$:

$$\left\langle \sum_{j>i} r_{ij} \frac{d\phi(r_{ij})}{dr_{ij}} \right\rangle = \frac{N(N-1)}{2Z_\phi} \int r_{12} \frac{d\phi(r_{12})}{dr_{12}} e^{-\sum_{j>i} \beta \phi(r_{ij})} d^3 r_1 \dots d^3 r_N = \frac{N^2}{2V} \int r \frac{d\phi(r)}{dr} g(r) d^3 r$$

and use the equipartition theorem to get $\langle K \rangle = \frac{3}{2}Nk_B T$. \square

Remarks 5.5. The key quantities in the classical statistical mechanics of liquids are the inter-molecular potential $\phi(r)$ and the radial distribution function $g(r)$. The Lennard-Jones potentials works well for simple liquids, and would normally give a poor description of systems with strong ionic, covalent or metallic bonding. For ionic bonding, one should include the long-range Coulomb forces, for covalent materials one needs directional bonding forces, and in metals the interaction energy is not well described by a sum of pairwise terms. Also, if the molecules are far from spherical or have significant dipole moments the inter-molecular potentials will depend on the relative orientation of the molecules.

Definition 5.5 (Virial expansion): We may expand the radial distribution function as a polynomial of density $n = N/V$. Plugging in the virial equation of state, we obtain the virial expansion:

$$\frac{p}{k_B T} = n + B_2(T)n^2 + B_3(T)n^3 + \dots$$

The m th Virial coefficient, $B_m(T)$ reflects the m -body correlations in the equation of state.

Remarks 5.6. Mayer developed a diagrammatic recipe which in principle allows each coefficient to be calculated. Note that the virial expansion only converges when the distance between molecules is much greater than their size, hence less useful at liquid densities. The virial expansion gives an excellent description of imperfect gases at low densities, when only pair interactions between particles are relevant.

Proposition 5.2 (Second virial coefficient):

$$B_2(T) = \int_0^\infty 2\pi r^2 (1 - e^{-\phi/k_B T}) dr$$

Proof. To calculate $B_2(T)$, we need an approximation to the radial distribution function which is accurate at low densities. The two-particle distribution function can be written as

$$P_2(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) e^{-\beta\phi(r_{12})} \frac{\int e^{-\beta \sum'_{j>i} \phi(r_{ij})} d^3r_3 \dots d^3r_N}{\int e^{-\beta \sum_{j>i} \phi(r_{ij})} d^3r_1 \dots d^3r_N}$$

where the primed summation indicates that we have removed the $\phi(r_{12})$ term. At very low densities, the integrals over particles 3, 4, ..., N in the numerator mostly involve configurations in which these particles are distant from each other and from particles 1 and 2, and therefore we can set all the $\phi(r_{ij})$ inside the integral to zero. In the denominator, particles 1 and 2 are also far apart for almost all configurations, hence again set all the $\phi(r_{ij})$ inside the integral to zero. Therefore

$$P_2(\mathbf{r}_1, \mathbf{r}_2) \approx \frac{N(N-1)}{V^2} e^{-\beta\phi(r_{12})} \implies \lim_{N \rightarrow \infty} g_0 = e^{-\beta\phi(r)}$$

which is as expected: $\phi(r=0) = \infty$, $g(r=0) = 0$ and $\phi(r=\infty) = 0$, $g_0(r=\infty) = 1$. Also, $g_0(r)$ takes its maximum value at the minimum in $\phi(r)$. We substitute $g_0(r)$ into the Virial expansion and integrate by parts:

$$\begin{aligned} \frac{p}{k_B T} &= n - \frac{n^2}{6k_B T} \int_0^\infty 4\pi r^3 \frac{d\phi}{dr} e^{-\phi/k_B T} dr \\ &= n + \frac{n^2}{6} \left\{ \left[4\pi r^3 e^{-\phi/k_B T} \right]_0^\infty - \int_0^\infty 12\pi r^2 e^{-\phi/k_B T} dr \right\} \\ &= n + \frac{n^2}{6} \left\{ \int_0^\infty 12\pi r^2 dr - \int_0^\infty 12\pi r^2 e^{-\phi/k_B T} dr \right\} \\ &= n + n^2 \left\{ \int_0^\infty 2\pi r^2 (1 - e^{-\phi/k_B T}) dr \right\} \end{aligned}$$

where we read off $B_2(T)$. □

Example 5.5. At high temperatures, $B_2(T)$ is dominated by hard core repulsion and is positive and roughly constant. At low temperatures, it is dominated by the van der Waals attraction, and is negative and roughly proportional to $1/T$. Experimental measurements agree with theory well except where quantum effects are important.

Definition 5.6 (Boyle temperature): Boyle temperature T_B is the temperature at which $B_2(T = T_B) = 0$. The contribution from the long-range particle attraction on average exactly compensates the short-range repulsion, leaving effectively ‘non-interacting’ particles. At higher temperature the gas is harder to compress than an ideal gas, and at low temperature it is easier.

Proposition 5.3: Under weak and slowly varying potential limit $\beta\phi(r) \ll 1$, the mean potential energy of a N -particle system is obtained by counting the interacting pairs:

$$U_{\text{pair}} \approx \frac{N^2}{V} k_B T B_2(T)$$

Proof. From counting the interacting pairs, the mean potential energy is

$$U_{\text{pair}} = \sum_{i>j} \phi(r_{ij}) \approx \frac{N^2}{2V} \int \phi(r) d^3r = \frac{N^2}{2} \bar{\phi}$$

where $\frac{1}{2}N^2$ is the number of distinct pairs in the system and the pair potential is an average over space, $\bar{\phi}$. In the limit of $\beta\phi(r) \ll 1$, we have $B_2 \approx \frac{1}{2} \int \beta\phi(r) d^3r$, and hence $U_{\text{pair}} \approx \frac{N^2}{V} k_B T B_2(T)$. □

5.3 van der Waals gas

Proposition 5.4 (Equation of state):

$$p = \frac{Nk_B T}{V - Nb} - \frac{aN^2}{V^2}$$

Proof. Separate the inter-atomic potentials as a sum of a short-range repulsive part $\phi_r(r)$ and an long-range attractive part $\phi_a(r)$:

$$\phi(r) = \phi_r(r) + \phi_a(r)$$

Each particle will feel the long-range attractive potential from many others and we can write

$$\sum_{j>i} \phi_a(r_{ij}) = \frac{1}{2} \sum_{j \neq i} \sum_i \phi_a(r_{ij}) \approx \frac{N(N-1)}{2V} \int \phi_a(r) d^3r \approx -\frac{aN^2}{V}, \quad a := -\frac{1}{2} \int \phi_a(r) d^3r$$

The effect of the short range repulsive part is to exclude the particles from a volume around each of them, and so we can write

$$\int e^{-\sum_{j>i} \beta \phi_r(r_{ij})} d^3r_1 \dots d^3r_N \approx (V - Nb)^N$$

The partition function is then

$$Z = \frac{1}{N!} \left(\frac{m}{2\pi\hbar^2\beta} \right)^{3N/2} \int e^{-\sum_{j>i} \beta [\phi_r(r_{ij}) + \phi_a(r_{ij})]} d^3r_1 \dots d^3r_N \approx \frac{1}{N!} \left(\frac{m}{2\pi\hbar^2\beta} \right)^{3N/2} (V - Nb)^N e^{\beta a N^2 / V}$$

The Helmholtz free energy is

$$F = -k_B T \ln Z = -k_B T \ln \left[\frac{1}{N!} \left(\frac{m}{2\pi\hbar^2\beta} \right)^{3N/2} \right] - k_B T N \ln(V - Nb) - \frac{aN^2}{V}$$

The pressure is then

$$p = -\left(\frac{\partial F}{\partial V} \right)_T = \frac{Nk_B T}{V - Nb} - \frac{aN^2}{V^2}$$

This is the desired van der Waals equation of state. □

Corollary 5.2: The second virial coefficient and Boyle temperature of a van der Waals gas respectively are

$$B_2(T) = b - \frac{a}{k_B T}, \quad T_B = \frac{a}{bk_B}$$

Proof. Approximate the equation of state to be:

$$p \approx \frac{Nk_B T}{V} \left(1 + \frac{Nb}{V} \right) - \frac{aN^2}{V^2} = \frac{Nk_B T}{V} + \frac{N^2}{V^2} (b - \beta a)$$

Result follows immediately. □

6 Phase transitions

6.1 Mixing and Phase Separation

Definition 6.1 (Average interaction energy): We define a non-dimensional parameter χ to be a measure of the average interaction energy.

$$k_B T \chi = (2\epsilon_{12} - \epsilon_{11} - \epsilon_{22})$$

where ϵ_{ij} is the interaction energy between a particle of species i and another of species j .

Proposition 6.1: The potential energy of mixing is

$$U_{\text{mix}} = N k_B T \chi c_1 (1 - c_1)$$

and hence the corresponding free energy is

$$F_{\text{mix}} = N k_B T (c \ln c + (1 - c) \ln(1 - c) + \chi c(1 - c))$$

Proof. We have shown that under a weak and slowly varying potential $\beta\phi(r) \ll 1$, the mean potential energy between any pair is

$$U_{ij} \approx \frac{N_i N_j}{V} k_B T B_2(ij)$$

where B_2 is the second virial coefficient. Limiting ourselves to pair interactions, we have:

$$U_{11} \approx \frac{N_1^2}{V} k_B T B_2(11), \quad U_{12} \approx 2 \frac{N_1 N_2}{V} k_B T B_2(12), \quad U_{22} \approx \frac{N_2^2}{V} k_B T B_2(22)$$

The entropy of mixing is $S_{\text{mix}} = -k_B N \sum_i c_i \ln c_i$. For two components, we have $c_2 = 1 - c_1$. Summing up the pair interactions, the total interaction energy

$$U_{\text{int}} = U_{11} + U_{22} + U_{12} = N(\epsilon_{11} c_1^2 + \epsilon_{22} (1 - c_1)^2 + 2\epsilon_{12} c_1 (1 - c_1))$$

where each energy factor is $\epsilon_{ij} = k_B T B_2(ij)/v_0$. Comparing the unmixed state, where we place all the particles of each species to one side, where they occupy a volume $V_i = v_0 N_i$. Neglecting the energy of the interface between the separated species, we have the unmixed energy to be

$$U_{\text{unmix}} = \frac{N_1^2}{V_1} k_B T B_2(11) + \frac{N_2^2}{V_2} k_B T B_2(22) = N c_1 (\epsilon_{11} c_1 + \epsilon_{22} (1 - c_1))$$

The potential energy of mixing is the difference between the two: $U_{\text{mix}} = N(2\epsilon_{12} - \epsilon_{11} - \epsilon_{22}) c_1 (1 - c_1)$. The free energy is $F_{\text{mix}} = U_{\text{mix}} - T S_{\text{mix}}$. \square

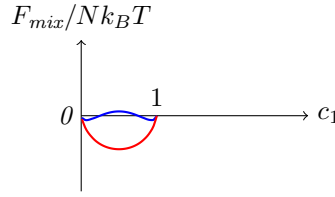
Remarks 6.1. Here, we take a ‘mean field’ approximation. We take an average field, the concentration of particles in this case, rather than a spatially dependent concentration field.

Example 6.1 (Osmotic pressure). The osmotic pressure is the pressure exerted by molecules ‘1’ in solution against a membrane which is impenetrable to them but allows the flow of solvent molecules ‘2’. The equation of state is

$$p_1 = -\frac{\partial F_{\text{mix}}}{\partial V} = -\frac{k_B T}{v_0} [\ln(1 - c) + \chi c^2] \approx \frac{k_B T}{v_0} \left[c + \frac{1}{2} (1 - 2\chi) c^2 \right]$$

where we take $c \ll 1$. The first term is the ideal gas law for species ‘1’ while the second term is the second virial correction, giving $B_2(T) = 0.5(1 - 2\chi)$.

Remarks 6.2. We plot F_{mix} as a function of c_1 . At high temperature (or low χ), the entropy of mixing stabilizes a fully mixed state - at any concentration (in red). Lowering the temperature (or increasing χ) results in a maximum in F_{mix} (in blue). When a two-well structure forms, F can be lowered by separating into two phases, one with $c > 0.5$ and the other with $c < 0.5$. Within the central part of the c -range, $F'' < 0$ and so the mixture is unstable regarding infinitesimal fluctuations. Spinodal decomposition (phase separation occurs spontaneously) thus occurs. Further out, at both extremes of the c -range, $F'' > 0$ and so the mixture is stable against small fluctuations. Phase separation still occurs but nucleation is required.



Proposition 6.2: The free energy of mixing has one minimum at $c = 0.5$ when χ is small or negative, and one maximum, two minima when $\chi > 0$ is large.

Proof. Check numerically for the following:

$$0 = \frac{\partial F_{\text{mix}}}{\partial c} = \chi(1 - 2c) - \ln \frac{1 - c}{c}$$

which is difficult to solve analytically. \square

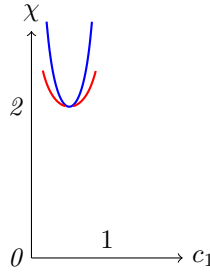
Definition 6.2 (Binodal line): The line of solutions for the minimum of mixing free energy $c^*(\chi)$ is called the binodal line.

Definition 6.3 (Spinodal line): The set of inflection points given by the solution of $\frac{\partial^2 F_{\text{mix}}}{\partial c^2} = 0$ is called the spinodal line, where

$$\frac{\partial^2 F_{\text{mix}}}{\partial c^2} = \frac{1}{(1 - c)c} - 2\chi$$

It represents the boundary of stability.

Remarks 6.3. The plots of the binodal (red) and the spinodal (blue) lines, as $\chi(c)$ variation, are shown. This maps the regions of different behaviour.



Below the binodal line (small χ and negative), at any concentration the mixture is stable: the F_{mix} is convex at any value of c . For large χ , different values of c_1 will give different signs for $F''(c_1)$, i.e. stability. Above the spinodal line, the values of c and χ are such that the mixture is absolutely unstable ($F'' < 0$) and its components separate immediately. The region between both lines is the region of metastability ($F'' > 0$) where one of the phases (mixed or separated) may have a higher free energy than the other, but is prevented from reaching the ground state by an energy barrier. Here, nucleation is required for phase separation.

Corollary 6.1: The critical point, where the region of metastability shrinks to zero as its two boundaries coincide, is at $c = 1/2$.

Proof. The minimum of the spinodal line $\chi_s = \frac{1}{2c(1-c)}$:

$$\frac{\partial \chi_s}{\partial c} = \frac{2c - 1}{2c^2(1 - c)^2} = 0 \implies c = \frac{1}{2} \iff \chi = 2$$

In another words, the mixture has to separate at the lowest value of the effective repulsive interaction χ . \square

Example 6.2 (First-order phase transition). Consider a system at equilibrium at concentration c_A and allow the concentration to fluctuate such that $c = c_A + \phi$ is reached. Assume ϕ is small, then

$$\frac{F_{\text{mix}}(c) - F_{\text{mix}}(c_A)}{Nk_B T} = -\frac{1 - 2\chi c_A(1 - c_A)}{2c_A(1 - c_A)}\phi^2 - \frac{1 - 2c_A}{6c_A^2(1 - c_A)^2}\phi^3 + \frac{1 - 3c_A + 3c_A^2}{12c_A^3(1 - c_A)^3}\phi^4 + \dots$$

where $F_{\text{mix}}/Nk_B T = c \ln c + (1-c) \ln(1-c) + \chi c(1-c)$. The term linear in ϕ has disappeared due to the equilibrium condition. For an arbitrary c_A , increasing χ gives a trajectory in the phase diagram that passes through the region of metastability. A new minimum develops after some χ , corresponds to crossing the binodal line. The new free energy minimum now occurs at a finite $\phi \neq 0$. Here, $F = 0$. There is still an energy barrier separating this new state of increased c from the original c_A . The barrier persists even when the new '1'-rich state has a significant preference in free energy, preventing the actual separation. Only at a substantially higher χ , the barrier finally disappears the state $\phi = 0$ becomes absolutely unstable, and corresponds to crossing the spinodal line, leading to a first order phase transition.

At the critical point, $c_A = 1/2$, we minimize the change in free energy to find ϕ^* :

$$\frac{F_{\text{mix}}(c) - F_{\text{mix}}(c_A)}{Nk_B T} = (2 - \chi)\phi^2 + \frac{4}{3}\phi^4 + \dots \implies \phi^* = \sqrt{\frac{3}{8}(\chi - 2)}$$

6.2 Ising Model

Definition 6.4 (Ising Model): The Ising model is a simple model for a magnet. Consider a lattice in d spatial dimensions with N sites. On each site $i = 1, \dots, N$, there is a discrete variable, called 'spin' which takes values:

$$S_i = +1 = \uparrow, \quad S_i = -1 = \downarrow$$

The system of spins $\{S_i\}$ has energy

$$E = -B \sum_i S_i - J \sum_{\langle i, j \rangle} S_i S_j$$

where $\langle i, j \rangle$ is sum over nearest neighbour pairs (which depend on the dimensions d and the lattice type) and B is the external magnetic field. To minimize E , $B > 0$ will lead to a preferential spin-up configuration while $B < 0$ spin-down. If $J > 0$, the spins preferentially align, i.e. ferromagnet. Conversely, if $J < 0$, the spins prefer to anti-align, i.e. anti-ferromagnet.

We are interested in the physics of the Ising model at a finite temperature T .

Remarks 6.4. A finite temperature T encourages the spins to ignore both the interactions and magnetic field because maximizing entropy becomes more important than minimizing energy. Since there are many more random configurations than aligned configurations, the temperature will tend to mess up the nice ordered states that the interactions and magnetic field have so carefully prepared.

Example 6.3. The 1D Ising model shows a phase transition, but only at exact $T = 0$. It costs very little energy to reverse the direction for a large block of spins on a line. For dimensions greater than one, the Ising model shows a real phase transition as $B \rightarrow 0$ at a temperature proportional to J , into a ferromagnetic phase with non-zero magnetization, while the high-temperature disordered paramagnetic phase has $M = 0$. When $B \neq 0$, the equilibrium value of M is finite and varies smoothly with temperature, and therefore no phase transition occurs.

Definition 6.5 (Mean field theory): Assume the interactions between any particular spins and the other spins can be represented by their average effect. This particular spin is $\sigma_i = \langle \sigma \rangle + \delta \sigma_i$ and assume the fluctuation is small.

Proposition 6.3: Under mean field theory (MFT), the spontaneous magnetization is

$$\langle \sigma \rangle = \tanh \frac{Jq \langle \sigma \rangle}{k_B T}$$

where q is the coordination number.

Proof. The Hamiltonian is

$$H = -m_0 B \sum_i \sigma_i - J \sum_{ij} (\langle \sigma \rangle + \delta \sigma_i)(\langle \sigma \rangle + \delta \sigma_j) = -m_0 B \sum_i \sigma_i - J \sum_{ij} (\langle \sigma \rangle^2 + \langle \sigma \rangle (\delta \sigma_i + \delta \sigma_j) + \delta \sigma_i \delta \sigma_j)$$

Usually, $m_0 = 1$. Drop the term quadratic in fluctuations, and assume that fluctuations are site-independent, then:

$$H \approx -m_0 B \sum_i \sigma_i - \frac{1}{2} z J \sum_i (\langle \sigma \rangle^2 + 2 \langle \sigma \rangle [\sigma_i - \langle \sigma \rangle]) \approx \frac{1}{2} q J N \langle \sigma \rangle^2 - (m_0 B + z J \langle \sigma \rangle) \sum_i \sigma_i$$

where z , the coordination number, is the number of nearest neighbours and the $1/2$ prefactor avoids double-counting, since each bond participates in two spins. The single-particle partition function is

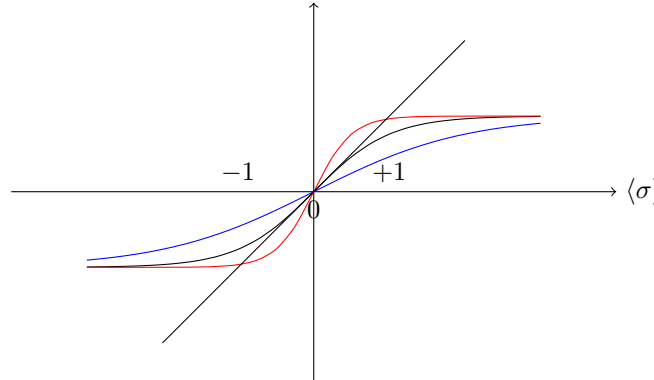
$$Z_1 = \sum_{\sigma=\pm 1} \exp \left[-\frac{q J \langle \sigma \rangle^2}{2 k_B T} - \frac{m_0 B \sigma + J q \sigma \langle \sigma \rangle}{k_B T} \right] = 2 e^{-0.5 \beta z J \langle \sigma \rangle^2} \cosh \frac{m_0 B + J q \langle \sigma \rangle}{k_B T}$$

The Helmholtz free energy per spin $F_1 = -k_B T \ln Z_1$ will then give us the mean magnetic moment per spin $\langle \sigma \rangle$ (related via $m = m_0 \langle \sigma \rangle$):

$$F_1 = -k_B T \ln \left[2 \cosh \frac{m_0 B + J q \langle \sigma \rangle}{k_B T} \right] + \frac{1}{2} q J \langle \sigma \rangle^2 \implies m = - \left(\frac{\partial F_1}{\partial B} \right)_T m_0 = m_0 \tanh \frac{m_0 B + J q \langle \sigma \rangle}{k_B T}$$

The spontaneous magnetization is obtained by setting $B = 0$. □

Graphic solution of the self-consistent mean-field equation $m = \tanh \beta J q m$ for the Ising model, illustrating the qualitatively different behaviour for $T > T_c$ (blue), $T < T_c$ (red), and $T = T_c$ (black). We can see that only when $T < T_c$, we have two other non-zero solutions to $m = \tanh[\beta(B + J q m)]$



Remarks 6.5. *There's a nice intuition behind this mean field result. It can be derived by assuming that each spin experiences an effective magnetic field given by $B_{\text{eff}} = m_0 B + J q \langle \sigma \rangle$, which includes an extra contribution from the spins around it. In this way, the tricky interactions in the Ising model have been replaced by an averaged effective field B_{eff} , which is the highlight of the mean field approximation.*

Remarks 6.6 (Validity of MFT).

1. For $d = 1$: MFT is completely wrong as there is supposed to be no phase transition.
2. For $d = 2, 3$: The basic structure of the phase diagram looks right, but the details near the T_c are wrong.
3. For $d \geq 4$: MFT gives the right answers.

Similar stories for other systems:

1. MFT fails completely for $d \leq d_l$, the lower critical dimension.
2. MFT always works for $d \geq d_u$, the upper critical dimension.
3. What about $d_l < d < d_u$? Often, those d 's are interesting.

6.3 Landau Theory of Phase Transitions

Definition 6.6 (Phase transition): *Conventionally, a phase transition occurs when some quantity changes discontinuously.*

For the Ising model, this order parameter that changes discontinuously is m . Landau theory offers a simple yet effective way to understand phase transitions qualitatively, based on free energy and symmetry (A \mathbb{Z}_2 symmetry when $B = 0$: $m(\mathbf{x}) \rightarrow -m(\mathbf{x})$. If $B \neq 0$, the symmetry should be $B \rightarrow -B \implies m(\mathbf{x}) \rightarrow -m(\mathbf{x})$). Taylor expand the free energy for the Ising model obtained from the mean field approximation:

$$F_1 \approx -k_B T \ln 2 + \frac{qJ}{2k_B T m_0^2} (k_B T - qJ) m^2 + \frac{(qJ)^4}{12(k_B T)^3 m_0^4} m^4 + \dots$$

Rewriting this in a cleaner expression, and dropping the insignificant constant terms:

$$f(m) \approx -Bm + \frac{1}{2} \left(\frac{1}{\beta} - Jq \right) m^2 + \frac{1}{12} \frac{1}{\beta} m^4 + \dots$$

where we introduced the term $-Bm$ when there is an external field B . The equilibrium value of m is m_{min} , as discussed earlier. Landau theory explains how does m_{min} change as we vary the parameters.

Remarks 6.7. *In Ehrenfest's classification scheme, phase transitions were labelled by the lowest derivative of the free energy which is discontinuous at the transition. In the modern classification, phase transitions are classified based on whether a latent heat is involved (only in first-order).*

6.3.1 $B = 0$: a continuous phase transition

Definition 6.7 (Second order phase transition): *At a second order phase transition, the order parameter changes continuously.*

Proposition 6.4: m turns off abruptly (but changes continuously from ± 1 to 0) as we increase T beyond a critical temperature T_c . $m \neq 0$ and $m = 0$ are respectively called ordered and disordered phases.

Proof. When $B = 0$, the free energy is

$$f(m) \approx \frac{1}{2} (T - T_c) m^2 + \frac{1}{12} T m^4 + \dots, \quad T_c = Jq$$

The equilibrium magnetization occurs at the minimum of the free energy:

$$0 = \frac{\partial f(m)}{\partial m} = (T - T_c) m + \frac{1}{3} T m^3 + \dots$$

When $T > T_c$, there is only one possible equilibrium magnetization, which is $m = 0$. This is expected since temperature randomizes the spins. When $T < T_c$, we solve for

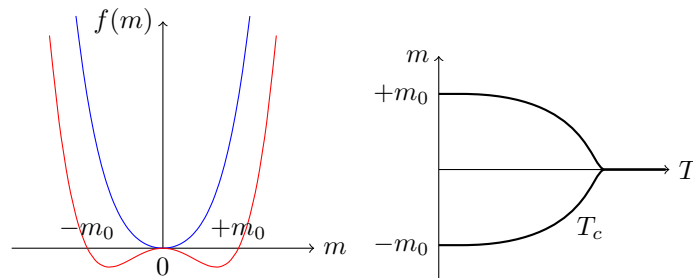
$$(T_c - T) m = \frac{1}{3} m^3 + \dots$$

but we need to find the minimum so

$$0 < \frac{\partial^2 f(m)}{\partial m^2} = T - T_c + T m^2 + \dots$$

As we can see, $m = 0$ is a maximum of the free energy. □

On the left is the graph for $f(m)$ where red and blue curves represent respectively $T < T_c$ and $T > T_c$. On the right is the graph for $m \neq 0$ when $T < T_c$.



Remarks 6.8.

1. The expression

$$m_0(T \sim T_c) = \pm \sqrt{3(T_c - T)/T}$$

is only valid close to T_c when m is small, and higher order terms $O(m^4)$ can be ignored.

2. $f(m)$ is invariant under \mathbb{Z}_2 symmetry, i.e. $m \rightarrow -m$ leaves $f(m)$ invariant. This symmetry is inherited from the $S_i \rightarrow -S_i$ symmetry of the Ising model when $B = 0$.

3. For $T < T_c$, the system chooses one of the two states $m = \pm m_0$, i.e. the \mathbb{Z}_2 symmetry is spontaneously broken. Spontaneous symmetry breaking occurs if the symmetry of a system is not respected by the ground state.

Proposition 6.5: The heat capacity is discontinuous at $T = T_c$.

Proof. The heat capacity (at constant $B = 0$) is a response function, given by

$$C = \frac{\partial \langle E \rangle}{\partial T} = \beta^2 \frac{\partial^2}{\partial \beta^2} \ln Z, \quad \langle E \rangle = -\frac{\partial}{\partial \beta} \ln Z$$

Bearing in mind the various m_{\min} (that minimizes the effective free energy) solutions obtained earlier, in the different regimes, we have

$$f(m_{\min}) = \begin{cases} \frac{1}{2}(T - T_c) \times 0 + \frac{1}{12}T \times 0 = 0 & T > T_c \\ \frac{1}{2}(T - T_c) \frac{3(T_c - T)}{T} + \frac{1}{12}T \frac{9(T_c - T)^2}{T^2} = -\frac{3}{4} \frac{(T_c - T)^2}{T} & T < T_c \end{cases}$$

Using $\ln Z \approx -\beta N f(m_{\min})$ gives the heat capacity per particle to be

$$c = \frac{C}{N} = \frac{\beta^2}{N} \frac{\partial^2}{\partial \beta^2} (-\beta N f(m_{\min})) = \begin{cases} 0 & T \rightarrow T_c^+ \\ \frac{3}{2} & T \rightarrow T_c^- \end{cases}$$

□

Remarks 6.9. The essence of a phase transition is that some quantity is discontinuous. But, everything was determined by the partition function Z , a sum of smooth analytic functions. Yet, we obtained non-analytic behaviour characteristic of a phase transition. In reality, Z is only necessary analytic if the sum is finite, which is not true when $N \rightarrow \infty$. Essentially, discontinuous phase transitions only occur in infinite systems. Similarly, spontaneous symmetry breaking can only occur in infinite systems.

6.3.2 $B \neq 0$: a discontinuous phase transition

Definition 6.8 (First order phase transition): At a first order phase transition, the order parameter changes discontinuously.

Proposition 6.6: For all temperatures T , the sign of m is completely determined by the sign of B .

Proof. The free energy is

$$f(m) \approx -Bm + \frac{1}{2}(T - T_c)m^2 + \frac{1}{12}Tm^4 + \dots$$

Two minima exist for low temperatures - one being metastable and another is a true ground state.

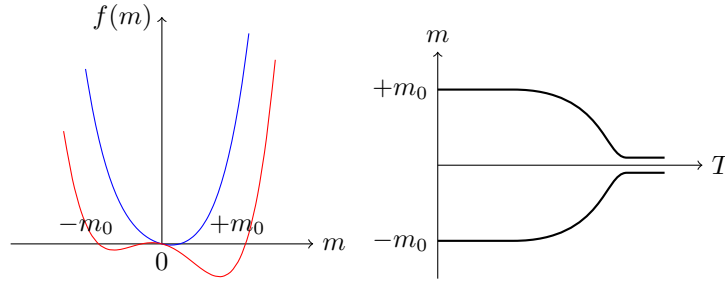
$$0 = \frac{\partial f(m)}{\partial m} = -B + (T - T_c)m + \frac{1}{3}Tm^3 + \dots$$

$$0 < \frac{\partial^2 f(m)}{\partial m^2} = (T - T_c) + Tm^2 + \dots$$

The metastable state disappears if the temperature increases above a certain value - the spinodal point. As we vary temperature T , the magnetization varies smoothly with asymptotic behaviour:

$$\lim_{T \rightarrow \infty} m = \frac{B}{T}, \quad \lim_{T \rightarrow 0} m = \pm 1$$

□



Proposition 6.7: For $T < T_c$, the magnetization m (the order parameter) jumps discontinuously from $-m_0$ to $+m_0$ as B flips from negative to positive. This is characteristic of a first order phase transition.

Proof. Take the red curve from the previous diagram. The position of the true minima depends on the sign of B . If $B = 0$, we obtain two minima, like earlier. \square

Remarks 6.10. Draw B against T . This line of first order phase transitions occurs at $B = 0$ along $T > 0$ axis and ends at a second order phase transition at $T = T_c$. This is referred to as the critical point. For $T > T_c$, we no longer experience a sign change in m if the sign of B changes.

6.3.3 Close to Critical Point

Proposition 6.8: The magnetic susceptibility (another type of response function, i.e. $\chi = \frac{\partial m}{\partial B}|_T$) of the Ising model have the following behaviour near a critical point.

$$\chi \sim \frac{1}{|T - T_c|}$$

Proof. Close to the critical point $T = T_c$,

$$f(m) \approx -Bm + \frac{1}{12}T_c m^4 + \dots \implies \frac{\partial f}{\partial m} = 0 \implies m \sim B^{1/3}, \quad B > 0$$

and when $B < 0$, $m \sim -|B|^{1/3}$. For $T > T_c$,

$$f(m) \approx -Bm + \frac{1}{2}(T - T_c)m^2 + \dots \implies \frac{\partial f}{\partial m} = 0 \implies m \approx \frac{B}{T - T_c} \implies \chi = \frac{\partial m}{\partial B} \sim \frac{1}{T - T_c}, \quad T \rightarrow T_c^+$$

For $T < T_c$, we can write minimum of f as $m = m_0 + \delta m$ for small B , and so to leading order in B/T ,

$$m \approx m_0 + \frac{B}{2(T_c - T)} \implies \chi \sim \frac{1}{T_c - T}, \quad T \rightarrow T_c^-$$

\square

Remarks 6.11. Landau theory assumes that the free energy can be expanded about the transition point in a power series in the order parameter, but this assumption is actually invalid at the critical point because the free energy is singular at a continuous phase transition. Crucially, Landau theory ignores fluctuations in the order parameter.

6.4 Critical Exponents and Universality

Example 6.4 (Critical Exponents). *In our MFT approach, we have seen that various quantities scale as power laws, as we approach the critical point. We first approach the critical point along the temperature axis ($B = 0$):*

$$m \sim (T_c - T)^\beta, \quad \beta = \frac{1}{2}, \quad T < T_c$$

$$c \sim c_\pm |T - T_c|^{-\alpha}, \quad \alpha = 0 \text{ (discontinuity)}$$

$$\chi \sim |T - T_c|^{-\gamma}, \quad \gamma = 1$$

To obtain the final exponent, we vary B at $T = T_c$:

$$m \sim B^{1/\delta}, \quad \delta = 3$$

$\alpha, \beta, \gamma, \delta$ are called critical exponents.

Compare MFT results with the true values for the Ising model in $d = 2$ and 3 :

	MFT	d=2	d=3
α	0	0	0.1101...
β	1/2	1/8	0.3264...
γ	1	7/4	1.2371...
δ	3	15	4.7898...

Remarks 6.12. *The critical exponents are only analytically worked out in $d = 2$. $d = 3$ is a numerical result, and no one has obtained an analytical solution for the 3D Ising model.*

The liquid-gas transition shows many similarities to the Ising model. Again, a line of first-order phase transitions ending at a critical point. At fixed p , the volume per particle $v = V/N$ jumps discontinuously at a liquid-gas first-order phase transition. This suggests v is an order parameter, analogous to m in the Ising model. Using an equation of state (e.g. van der Waals equation in statistical physics) at the critical point:

$$v_{\text{gas}} - v_{\text{liquid}} \sim (T_c - T)^\beta, \quad \beta = \frac{1}{2}, \quad T < T_c$$

$$v_{\text{gas}} - v_{\text{liquid}} \sim (p - p_c)^{1/\delta}, \quad \delta = 3, \quad p \rightarrow p_c^+$$

Analogous to the susceptibility, the compressibility (another response function) is

$$\kappa = -\frac{1}{v} \frac{\partial v}{\partial p} \bigg|_T \sim |T - T_c|^{-\gamma}, \quad \gamma = 1$$

$$C_V \sim C_\pm |T - T_c|^{-\alpha}, \quad \alpha = 0$$

The critical exponents ($\alpha = 0, \beta = 1/2, \gamma = 1, \delta = 3$) of liquid-gas transition are the same as the Ising model using MFT. None of these theoretical values, again, agree with experiment. But yet, the actual experimental measurements of the critical exponents agree with $d = 3$ Ising.

Definition 6.9 (Universality): *The critical point governs behaviour of many different physical systems. All memory of underlying microscopic physics is washed away. If two systems are governed by the same critical point, we say they live in the same universality class.*

Remarks 6.13 (Ising model as a lattice gas). *Consider the same d -dimensional lattice, but now with particles hopping between sites. Let the number of particles on site i be n_i . By the hard core assumption, no more than one particle can be on each site. So $n_i = 0$ and $n_i = 1$ for an empty and filled site respectively. The energy would then be*

$$E = -4J \sum_{\langle i,j \rangle} n_i n_j - \mu \sum_i n_i$$

For $J > 0$, there is an attractive force between sites. μ is the chemical potential. This is equivalent to the Ising model if we identify $S_i = 2n_i - 1$.

7 Fluctuations in Equilibrium

7.1 Fluctuations

Example 7.1. *At second-order phase transitions, fluctuations in some thermodynamic quantities diverge, giving rise to new phenomena. Even far from critical points, spatial fluctuation in the density scatter particles. For instance, the scattering of conduction electrons in metals from density fluctuations of the positively-charged ions on the lattice.*

Definition 7.1 (Variance in canonical ensemble):

$$\langle \Delta x^2 \rangle := \langle x^2 \rangle - \langle x \rangle^2 = \frac{1}{Z} \sum_i x_i^2 e^{-E_i/k_B T} - \left(\frac{1}{Z} \sum_i x_i e^{-E_i/k_B T} \right)^2$$

Proposition 7.1: The variance of x is related to its response function, assuming a linear dependence $E_i = -f x_i$.

Proof. If the linear dependence between the microstate energy and the variable, a generic trick applies

$$\langle x \rangle = \frac{1}{Z} \sum_i x_i e^{f x_i \beta} = \frac{1}{\beta Z} \frac{\partial Z}{\partial f}, \quad \langle x^2 \rangle = \frac{1}{Z} \sum_i x_i^2 e^{f x_i \beta} = \frac{1}{Z \beta^2} \frac{\partial^2 Z}{\partial f^2}$$

Take one more derivative of $\langle x \rangle$:

$$k_B T \frac{\partial \langle x \rangle}{\partial f} = (k_B T)^2 \left[-\frac{1}{Z^2} \left(\frac{\partial Z}{\partial f} \right)^2 + \frac{1}{Z} \frac{\partial^2 Z}{\partial f^2} \right] = -\langle x \rangle^2 + \langle x^2 \rangle = \langle \Delta x^2 \rangle$$

□

Example 7.2. Consider $E_i = -M_i B$, then the mean square fluctuation of the magnetization M is

$$\langle \Delta M^2 \rangle = k_B T \left(\frac{\partial \langle M \rangle}{\partial B} \right)_T$$

For a system of non-interacting spins, $\langle M \rangle = N m_0 \tanh \frac{m_0 B}{k_B T}$, and so

$$\langle \Delta M^2 \rangle = \frac{N m_0^2}{\cosh^2(m_0 B / k_B T)} \implies \frac{\sqrt{\langle \Delta M^2 \rangle}}{\langle M \rangle} \propto \frac{1}{\sqrt{N}}$$

Corollary 7.1: The mean square fluctuation in the internal energy U is proportional to the heat capacity.

Proof.

$$U = \langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}, \quad \langle E^2 \rangle = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}$$

and so

$$\langle \Delta U^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 = \frac{\partial}{\partial \beta} \left[\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right] = k_B T^2 \frac{\partial U}{\partial T}$$

□

7.2 Connection with Thermodynamics

Proposition 7.2: The magnitude of fluctuations in x is related to the availability of the system via

$$\langle \Delta x^2 \rangle := \langle (x - x_0)^2 \rangle = \frac{k_B T_R}{(\partial^2 A / \partial x^2)_{x=x_0}}$$

Proof. Consider a system that can exchange energy and particles with a large reservoir. We wish to calculate the probability distribution $P(x)$ for variable x of the system, when the total internal energy of the system and reservoir is U_{tot} , it is easy to see that $P(x) \propto \Omega_{\text{tot}}(x, U_{\text{tot}})$, where $\Omega_{\text{tot}}(x, U_{\text{tot}})$ is the number of microstates of the system and reservoir in which the variable of the system is equal to x . The entropy of the system and reservoir is given by $S_{\text{tot}}(x, U_{\text{tot}}) = k_B \ln \Omega_{\text{tot}}(x, U_{\text{tot}})$, but we may relate this with the availability via $dS_{\text{tot}} = -dA(x)/T_R$ where T_R is the temperature of the reservoir. hence, $P(x) \propto e^{-A(x)/k_B T_R}$. Expand A about the equilibrium value of x , x_0 :

$$A(x) = A(x_0) + (x - x_0) \left. \frac{\partial A}{\partial x} \right|_{x=x_0} + \frac{1}{2} (x - x_0)^2 \left. \frac{\partial^2 A}{\partial x^2} \right|_{x=x_0} + \dots$$

but the first derivative is zero since it is a minimum, so the probability distribution within the Gaussian approximation is

$$P(x) = \frac{1}{\sqrt{2\pi\langle\Delta x^2\rangle}} \exp \left[-\frac{\Delta x^2}{2\langle\Delta x^2\rangle} \right]$$

where $\langle\Delta x^2\rangle$ follows. □

Corollary 7.2:

$$\langle\Delta x^2\rangle = k_B T \left. \frac{\partial x}{\partial f} \right|_{x=x_0}$$

Proof. The first derivative $\frac{\partial A}{\partial x}$ is equal to the thermodynamic force f conjugate to the variable x . □

Example 7.3 (Fluctuation-dissipation theorem). *Consider the arbitrary conjugate pair of thermodynamic variables (x, f) as a force and the corresponding displacement via the linear response relation $f = \alpha x$, then the mean squared fluctuation of a thermodynamic variable is $\langle\Delta x^2\rangle = k_B T / \alpha$. This is the first example of the fluctuation-dissipation theorem we will see.*

Corollary 7.3: The probability distribution for fluctuations in the internal energy U of a system held at constant volume and particle number but thermally coupled to a reservoir, is a Gaussian.

Proof. At constant V and N , the availability is the free energy:

$$dA = dU - T_R dS := dF \implies \left(\frac{\partial F}{\partial U} \right)_V = 1 - T_R \left(\frac{\partial S}{\partial U} \right)_V = 1 - \frac{T_R}{T}$$

which must be zero in equilibrium, so $T = T_R$. Differentiating again, we obtain

$$\left(\frac{\partial^2 F}{\partial U^2} \right)_V = \frac{T_R}{T^2} \left(\frac{\partial T}{\partial U} \right)_V = \frac{1}{T C_V}$$

where $T_R = T$. Using earlier results, we have

$$\langle\Delta U^2\rangle = \frac{k_B T}{(\partial^2 F / \partial U^2)_V} = k_B T^2 C_V, \quad P(U) = \frac{1}{\sqrt{2\pi k_B T^2 C_V}} \exp \left[-\frac{\Delta U^2}{2k_B T^2 C_V} \right]$$

□

Example 7.4. 1 cm^3 of monatomic ideal gas contains $N = \frac{pV}{k_B T} \sim 2.5 \times 10^{19}$ atoms. Equipartition gives $U = \frac{3}{2} N k_B T = 0.15 \text{ J}$ and the heat capacity at constant volume is $C_V = \frac{3}{2} N k_B = 5.2 \times 10^{-4} \text{ J/K}$, so the root mean square fluctuation is

$$\langle\Delta U^2\rangle^{1/2} \sim T \sqrt{k_B C_V} \sim 2.5 \times 10^{-11} \text{ J}$$

The fractional fluctuation is $\sqrt{\langle\Delta U^2\rangle}/U \sim 1.6 \times 10^{-10}$, which is unmeasurably small.

7.3 Near Critical Points

Example 7.5 (van der Waals gas). *In p - V coordinates, the critical isotherm has a single point where $\frac{dp}{dV} = 0$. Applying the first law, $dU = T_R dS - p dV \implies dA = (p_R - p)dV$, and so differentiating gives*

$$\left(\frac{\partial A}{\partial V}\right)_T = p_R - p, \quad \left(\frac{\partial^2 A}{\partial V^2}\right)_T = -\left(\frac{\partial p}{\partial V}\right)_T = \frac{1}{\kappa_T V}$$

where κ_T is the isothermal compressibility, i.e. the linear response coefficient for the conjugate pair of thermodynamic variables (p, V) . Hence,

$$\langle \Delta V^2 \rangle = k_B T V \kappa_T, \quad P(V) = \frac{1}{\sqrt{2\pi k_B T V \kappa_T}} \exp\left(-\frac{\Delta V^2}{2k_B T V \kappa_T}\right)$$

The fractional volume fluctuation is

$$\frac{\sqrt{\langle \Delta V^2 \rangle}}{V} = \sqrt{k_B T \kappa_T / V}$$

$\kappa_T \propto dV/dP$ diverges at the liquid-vapour critical point, and hence the magnitude of the volume fluctuations. We observe this as a ‘critical opalescence’ phenomena - volume of density fluctuations increases to the point they start scattering light. A similar example is seen in the demixing region of a phase-separating mixture, but in (T, c) .

Example 7.6 (Ising model). *The mean square fluctuation of magnetisation is*

$$\langle \Delta M^2 \rangle = k_B T \frac{\partial M}{\partial B}$$

Evaluated at $B \rightarrow 0$ and in equilibrium $M_0 = 0$ at $T > T_c$, the derivative is equal to the susceptibility χ , the corresponding linear response coefficient of the (M, B) pair. This susceptibility diverges at the critical point and so must the magnitude of fluctuation in magnetisation. This divergence occurs on both sides of the critical point.

8 Stochastic Physics

8.1 Brownian Motion

Definition 8.1 (Brownian motion): Brownian motion refers to the pereptual irregular motions exhibited by small grains or particles of colloid size immersed in a liquid.

Definition 8.2 (Langevin equation): The theory of the Brownian motion of a free particle is described by the Langevin equation:

$$m \frac{d\mathbf{u}}{dt} = -\gamma \mathbf{u} + \xi(t)$$

where \mathbf{u} denotes the vector of particle velocity. The term $-\gamma \mathbf{u}$ represents the kinetic friction experienced by the particle, which is governed by fluid dynamics. The second contribution is fluctuating, which is the force provided by collisions with molecules of surrounding liquid. This stochastic force has ‘white noise’ properties:

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t)^2 \rangle \neq 0, \quad \langle \xi(t_1) \xi(t_2) \rangle = 0 \text{ if } t_1 \neq t_2$$

i.e. zero expectation and complete lack of correlation, i.e. $\langle \xi(t) \xi(t') \rangle = \Gamma \delta(t - t')$ where Γ is a measure of intensity of the stochastic force.

Proposition 8.1:

$$\Gamma = 2k_B T \gamma$$

Proof. The homogeneous dynamical equation $m\dot{u} = -\gamma u$ has the obvious solution $u = u_0 e^{-\gamma t/m}$, which represents the decay of the initial condition $u(t=0) = u_0$. With the added stochastic force, the solution is

$$u(t) = u_0 e^{-\gamma t/m} + \int_0^t e^{-\gamma(t-t')/m} \frac{\xi(t')}{m} dt'$$

The instantaneous value of $u(t)$ cannot be predicted, but its mean square value is easy to determine:

$$\begin{aligned} \langle u^2 \rangle &= \left\langle u_0^2 e^{-2\gamma t/m} + 2u_0 e^{-\gamma t/m} \int_0^t e^{-\gamma(t-t')/m} \frac{\xi(t')}{m} dt' + \int_0^t \int_0^t e^{-\gamma(2t-t_1-t_2)/m} \frac{\xi(t_1) \xi(t_2)}{m^2} dt_1 dt_2 \right\rangle \\ &= u_0^2 e^{-2\gamma t/m} + \frac{1}{m^2} \int_0^t \int_0^t e^{-\gamma(2t-t_1-t_2)/m} \langle \xi(t_1) \xi(t_2) \rangle dt_1 dt_2 \\ &= u_0^2 e^{-2\gamma t/m} + \frac{\Gamma}{m^2} e^{-2\gamma t/m} \int_0^t e^{2\gamma t_1/m} dt_1 \\ &= u_0^2 e^{-2\gamma t/m} + \frac{\Gamma}{2m\gamma} (1 - e^{-2\gamma t/m}) \end{aligned}$$

The cross-term, linear in $\langle \xi \rangle$ is dropped out. In the long-time limit, the decaying exponential vanishes, i.e. the transient dies out. Invoking equipartition theorem,

$$\langle 0.5 m u^2 \rangle = \frac{\Gamma}{4\gamma} = \frac{1}{2} k_B T$$

The desired result follows. □

8.2 Diffusion of Free and Confined Particles

Proposition 8.2: For the simplest 1D example,

$$\langle x^2 \rangle = 2Dt$$

where D is some constant, called the diffusion constant.

Proof. After a long time of observation, $t \gg m/\gamma$, no memory of the initial particle velocity u_0 must remain and the current velocity must be distributed according to the Maxwell distribution.

Inertial effects are irrelevant on long time scales of interest and the particle acceleration \dot{u} must average out. The forces must balance out:

$$0 = -\gamma u + \xi(t) \implies \frac{dx}{dt} = \frac{1}{\gamma} \xi(t) \implies x(t) = \frac{1}{\gamma} \int_0^t \xi(t') dt'$$

Hence, it follows that

$$\langle x^2 \rangle = \frac{1}{\gamma^2} \int_0^t \int_0^t \langle \xi(t_1) \xi(t_2) \rangle dt_1 dt_2 = \frac{\Gamma t}{\gamma^2} = 2Dt, \quad D = \frac{k_B T}{\gamma}$$

□

Corollary 8.1:

$$\langle \mathbf{r}^2 \rangle = 6Dt$$

for 3D diffusion.

Proposition 8.3: For a confined potential of $V = \frac{1}{2}\alpha x^2$, we have instead

$$\langle x^2 \rangle = \frac{k_B T}{\alpha}$$

Proof. At time scales of observation which are much longer than the velocity relation time $\tau = m/\gamma$, inertial effects are again irrelevant and the acceleration averages to zero in the equation of motion for the average position. In this limit, the Langevin equation takes the overdamped form:

$$\gamma \dot{\mathbf{x}} = -\alpha \mathbf{x} + \xi(t)$$

which has the same equation form for a free particle with the inertial term present. The stochastic solution for $x(t)$ is

$$x(t) = x_0 e^{-\alpha t/\gamma} + \int_0^t e^{-\alpha(t-t')/\gamma} \frac{\xi(t')}{\gamma} dt'$$

Again, its mean-squared value is

$$\begin{aligned} \langle x^2 \rangle &= x_0^2 e^{-2\alpha t/\gamma} + \frac{1}{\gamma^2} \int_0^t \int_0^t e^{-\alpha(2t-t_1-t_2)/\gamma} \langle \xi(t_1) \xi(t_2) \rangle dt_1 dt_2 \\ &= x_0^2 e^{-2\alpha t/\gamma} + \frac{\Gamma}{\gamma^2} e^{-2\alpha t/\gamma} \int_0^t e^{2\alpha t_1/\gamma} dt_1 \\ &= x_0^2 e^{-2\alpha t/\gamma} + \frac{\Gamma}{2\alpha\gamma} (1 - e^{-2\alpha t/\gamma}) \end{aligned}$$

A new time scale $\tau_x = \alpha/\gamma$ has emerged, which determines how long does the memory of the initial condition x_0 persists in such a system. Without thermal noise $\xi(t)$, the particle released at $x = x_0$ would come to a rest at the point of its equilibrium $x = 0$ in approximately this length of time. In the long time limit, at $t \gg \tau_x$, we obtain

$$\langle x^2 \rangle = \frac{\Gamma}{2\alpha\gamma} = \frac{k_B T}{\alpha}$$

□

Example 8.1. Dissipation leads to fluctuations in electric circuits. Nyquist's argument is as follow: the transmission line is a blackbody with spectrum

$$E_\omega d\omega = k_B T l d\omega / \pi c$$

Terminate the transmission line with a characteristic impedance R . The termination resistor injects noise power $I^2 R d\omega = \frac{k_B T}{2\pi} \omega$. Model in terms of a voltage source injecting noise:

$$I_\omega = \frac{V_\omega}{2R} \implies \frac{V_\omega^2}{4R^2} R d\omega = k_B T df$$

This gives the Johnson noise $V_f^2 = 4Rk_B T$ injected by the resistor R .

8.3 Working with probability distributions

Proposition 8.4: The probability distribution for a 1-D random walk is

$$P(N_+, N) = 2a\sqrt{\frac{1}{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right)$$

for some D . This is the probability that after N such steps the particle reaches a certain distance x from the starting point. N_+ and N_- are the steps to the right and left respectively, with $x = a(N_+ - N_-)$, where a is the length of an elementary step.

Proof. The probability of finding the configuration with N trials out of which N_+ are ‘successful’ is

$$P(N_+, N) = \frac{N!}{N_+!(N - N_+)!} \frac{1}{2^N} = \sqrt{\frac{N}{2\pi N_+(N - N_+)}} e^{N \ln N - N_+ \ln N_+ - (N - N_+) \ln (N - N_+) - N \ln 2}$$

where we used Stirling formula $N! \approx e^{-N} N^N \sqrt{2\pi N}$. Expand in terms of $\epsilon = x/aN = x\tau/at$, and keeping terms up to second order in the exponential, gives

$$P(N_+, N) = \sqrt{\frac{2}{\pi N}} \sqrt{\frac{1}{1 - (x/aN)^2}} \exp\left[-\frac{x^2}{2a^2 N}\right] = 2a\sqrt{\frac{1}{4\pi Dt}} \exp\left[-\frac{x^2}{4Dt}\right]$$

where we used $D = a^2/2\tau$. □

Remarks 8.1. We can always write $P(N_+, N) = 2aP(x, t)$ for some continuous probability density, such that $P(x, t)dx$ is the probability of finding the particle after time t in the interval $[x, x + dx]$.

Definition 8.3 (Transition probability): Assign a number m to the current position the particle occupies after $N = 1$ steps and denote the probability of this occurrence is $P(m, N + 1)$.

$$P(m, N + 1) = w(m, m - 1)P(m - 1, N) + w(m, m + 1)P(m + 1, N)$$

where $w(m, m + 1)$ and $w(m, m - 1)$ are transition probabilities.

Proposition 8.5 (Free Diffusion equation):

$$\frac{\partial P(x, t)}{\partial t} \propto \frac{\partial^2 P(x, t)}{\partial x^2}$$

Proof. For a free Brownian motion in 1D, we set all w ’s to be $1/2$ and take infinitesimal limit:

$$\begin{aligned} \frac{P(m, N + 1) - P(m, N)}{\tau} &= \frac{1}{\tau} w(m, m - 1)P(m - 1, N) + \frac{1}{\tau} w(m, m + 1)P(m + 1, N) \\ &\quad - \frac{1}{\tau} [w(m - 1, m) + w(m + 1, m)]P(m, N) \\ \frac{\partial P(m, t)}{\partial t} &= \frac{1}{2\tau} [P(m - 1, t) - 2P(m, t) + P(m + 1, t)] \\ &= \frac{1}{2\tau} \frac{\partial^2 P(m, t)}{\partial m^2} = \left(\frac{a^2}{2\tau}\right) \frac{\partial^2 P(x, t)}{\partial x^2} \end{aligned}$$

where we replaced the index m by a continuous variable $m = x/a$. □

Remarks 8.2. Assume the transition probabilities are not equal, independently of where the particle is on the x -scale, we take

$$w(m, m - 1) = \frac{1}{2} - \epsilon, \quad w(m, m + 1) = \frac{1}{2} + \epsilon$$

Then the diffusion equation now becomes

$$\frac{\partial P(m, t)}{\partial t} = \frac{1}{2\tau} \frac{\partial^2 P(m, t)}{\partial m^2} + \frac{2\epsilon}{\tau} \frac{\partial P(m, t)}{\partial m}$$

Often, it is casted in the form $\partial_t P(x, t) = D \partial_x^2 P(x, t) + C \partial_x P(x, t)$.

Proposition 8.6 (Fokker-Planck equation): The full, time-dependent, generalized diffusion equation under an external force $f = -dE/dx$, where $E(x)$ is the corresponding potential energy of the particle at point x , is

$$\frac{\partial P(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left[\frac{1}{\gamma} f(x) P(x, t) \right] + D \frac{\partial^2}{\partial x^2} P(x, t)$$

Proof. First, consider the steady state of this process:

$$0 = D \frac{\partial^2}{\partial x^2} P(x) + C \frac{\partial}{\partial x} P(x) \implies P_{\text{eq}}(x) = P_0 e^{-Cx/D}$$

where at very large x , both P and P' have to be zero. The equilibrium probability distribution must have the Boltzmann form, $P_{\text{eq}} \propto \exp[-E(x)/k_B T]$, with $E(x)$ the corresponding potential energy at point x . So, $C = \frac{1}{\gamma} \frac{dE(x)}{dx}$. Let $f = -dE/dx$ be the external force, then plug this back and differentiate with respect to x to get the full equation:

$$\partial_t P(x, t) = \partial_x \left(\frac{1}{\gamma} \frac{dE}{dx} + D \frac{\partial}{\partial x} \right) P(x, t)$$

as desired. \square

Corollary 8.2: We may cast the Fokker-Planck equation to a continuity equation form:

$$\frac{\partial P(x, t)}{\partial t} = -\frac{\partial J(x, t)}{\partial x}$$

Proof. Comparing the continuity equation and the Fokker-Planck equation, we have

$$J(x, t) = -D \frac{\partial P(x, t)}{\partial x} + \frac{1}{\gamma} f(x) P(x, t) = -D e^{-\beta E(x)} \frac{\partial}{\partial x} \left[e^{\beta E(x)} P(x, t) \right]$$

where $J(x, t)$ represents the diffusion current for the probability $P(x, t)$. \square

Example 8.2 (Kramers' problem). *Kramers' problem describes the escape of particles over potential barriers. Suppose the particles are initially caught in the metastable state at A, and we wish to find the rate at which the particles will escape from this state over the potential (of energy barrier ΔE before reaching a global minimum behind it at B), as a result of thermal motion. When steady current of particles is established in the long-time limit,*

$$J \int_A^B e^{\beta E(s)} ds = -D \left[e^{\beta E(x)} P(x) \right]_A^B$$

Approximate the potential near the metastable equilibrium point A as $E(x) \approx E_A + \frac{1}{2} K_A x^2$. The number of particles in the vicinity of A can be estimated by taking

$$dN_A = P(x_A) e^{-\beta E_A} dx \implies N_A \approx P(x_A) \int_{-\infty}^{\infty} e^{-\beta K_A x^2 / 2} dx = P(x_A) \sqrt{\frac{2\pi k_B T}{K_A}}$$

Assuming the potential well $E_B(x)$ is deep enough, then the steady state current is

$$J \approx \frac{DP(x_A)}{\int_A^B e^{\beta E(s)} ds}$$

and the rate is $J/N_A = D \sqrt{\frac{K_A}{2\pi k_B T}} / \int_A^B e^{\beta E(s)} ds$. The principal contribution to the integral arises only from the very small region near the potential barrier C, with an approximate parabolic form near the maximum, i.e. $E_c \approx \Delta E - 0.5 K_c (x - x_0)^2$.

$$\int_A^B e^{\beta E(s)} ds \approx e^{\Delta E / k_B T} \int_{-\infty}^{\infty} e^{-0.5 \beta K_c (x - x_0)^2} dx = e^{\Delta E / k_B T} \sqrt{\frac{2\pi k_B T}{K_c}}$$

Hence, the rate of particles transit over the barrier (equivalent to the rate of leaving the metastable state A) is

$$\frac{J}{N_A} = D e^{-\Delta E / k_B T} \frac{\sqrt{K_A K_c}}{2\pi k_B T} = \frac{\sqrt{K_A K_c}}{2\pi \gamma} e^{-\Delta E / k_B T}$$

This expression gives the probability, per unit time, that a particle originally in the potential hole A, will escape to a deeper well B by crossing the barrier at C. The prefactor fraction is determined by various microscopic features of the system. The exponential form represents thermal activation.

8.4 Fluctuation-Dissipation Theorem

Proposition 8.7: The fluctuating Langevin force indeed has a uniform power spectrum and that the intensity of the Langevin force is given by the damping constant.

Proof. We start from a damped harmonic oscillator (since the Langevin equation is the equation of motion of a damped simple harmonic oscillator with a stochastic applied force $f(t)$), which has a response function:

$$\alpha_\omega = \frac{1}{-m\omega^2 - i\gamma\omega + k} = \frac{m(\omega_0^2 - \omega^2) + i\gamma\omega}{m^2(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2}$$

Equipartition principle says the mean square displacement should be

$$\frac{1}{2}k\langle x^2 \rangle = \frac{1}{2}k_B T \implies \langle x^2 \rangle = k_B T \frac{1}{k} = k_B T \text{Re}[\alpha_{\omega=0}]$$

The fluctuation dissipation theorem gives us

$$\langle \delta x_\omega^2 \rangle = \frac{2k_B T}{\omega} \text{Im}[\alpha_\omega]$$

where $\delta x = x - \bar{x} = x$ as $\bar{x} = 0$ for our damped oscillator. We have

$$x_\omega = f_\omega \alpha_\omega \implies \langle x_\omega^2 \rangle = \langle |x_\omega|^2 \rangle = \langle |\xi_\omega|^2 \rangle |\alpha_\omega|^2 = \frac{2k_B T}{\omega} \text{Im}[\alpha_\omega] \implies \langle |\xi_\omega|^2 \rangle = \frac{2k_B T}{\omega} \frac{\text{Im}[\alpha_\omega]}{|\alpha_\omega|^2} = 2k_B T \gamma$$

i.e. constant. □

References

- [1] S.J. Blundell and K.M. Blundell. *Concepts in Thermal Physics*. OUP Oxford, 2010.