Before I start with true content of the module, I go more into depth about some of the aspects of classical thermodynamics and statistical mechanics.

1 Classical Thermodynamics

In thermodynamics, we are trying to describe systems which are composed of many individual particles, with hope that the overall behaviour of the particles can be described by a small number of observable variables.
A special case of these systems are systems in equilibrium, for which we assume now macroscopic variables can change over time and no macroscopic currents are present in the system.
The variables we use to describe the can be usually characterized as two categories - extensive and intensive. Extensive variables are directly proportional to the system size. These are for example volume, energy or number of particles in the system. The intensive variables, on the other hand, are independent of the system size. An example of intensive variable could be temperature, pressure or chemical potential.
These two types create so called conjugate pairs, as will be discussed later.

1.1 Laws of Thermodynamics

The thermodynamics are based on set of laws that are either very fundamental physical concepts or statistically almost observations. These laws are now presented

1.1.1 Zeroth Law

Suppose that we have three systems - A, B and C. If we create connection between A and B and observe that they are in equilibrium (no change in observables and no appearance of currents), then disconnect them and connect A and C to observe equilibrated state as well, then we can infer that B and C are in equilibrium with each other.
This can be formulated as the transitive property of the equilibrium state, and can be used to define some observables. For example, if we connect the systems so that they can exchange energy in form of heat, we can define an intensive property of the system called temperature and say that two systems are in equilibrium if they have the same temperature. If we connect the systems so that they can exchange particles, we can define chemical potential and say that two systems are in equilibrium if the chemical potential of the system is the same etc.

1.1.2 First Law

There is a variable called the internal energy of the system U which is only a function of the current state of the system. This means that the energy of the system does not depend on the path the system took in the past - U is memoryless. In fact, in thermodynamics, we define several so called variables of state (or functions of state, or just state variables) which all share the property that they describe memoryless properties of the system.
The internal energy of the system must obey the energy conservation. Therefore, if there is a flux of energy into the system δQ and the system does work δW between two very similar states, then the infinitesimal change in the internal energy is

$$dU = \delta Q - \delta W \tag{1}$$

Here, we write δ instead of d as the energy flux and work done are not state variables - they can depend on the path the system takes.
The work done depends especially on the type of the system. For example, if the system does work by pressure, we can consider a pressure force acting on some boundary of the system of area dA, so that the this boundary moves distance dx. The force due to the pressure on dA is $dF = pdA$ and the work done by this expansion is $\delta W = dFdx = pdAdx = pdV$. Notice that p here is an intensive variable and V an extensive variable.
Now consider a system which has chemical potential μ which traps the particles inside the system. We define chemical potential as the energy the particle gains when it moves to the system $\Delta E_p = \mu$. Since μ traps particles, μ has to be negative. Then, if during some process the system releases dN particles, the system must have done work $\delta W = -\mu dN$ to move the particles out. The minus sign ensures that work done is a positive quantity in this case, as the system has to do work to move the particles. Again, notice that μ is intensive and N is an extensive variable.

Lastly, we could imagine a system with magnetic moment $\vec{m}$. If we increase the surrounding magnetic induction field by $d\vec{B}$, we change the energy of the system by $-\vec{m} \cdot d\vec{B}$, which is equivalent as the system doing work $\vec{m} \cdot d\vec{B}$.
Similarly, we could come up with other processes by which the system can do work, and we would always find a pair of extensive and intensive variable describing the amount of work done. This is one way how to define conjugate pairs of state variables - together, they describe the work done by the system. Of course, generally, we can have combined processes, for which δW is the sum of the work done by different pairs of conjugate variables. I.e. for a process where both the number of particles and the volume of the system changes, but there is no flow δQ of heat into the system, we could write

$$dU = -\delta W = -pdV - \vec{m} \cdot d\vec{B}$$

1.1.3 Second Law

The second law defines another state variable called the entropy S. Entropy relates to the amount of non-useful energy stored in the system. By non-useful energy I mean mostly the energy of heat, which is the energy of the disordered motion of the particles. If we have two systems, the useful energy which can be extracted from the other one however does not depend just on the amount of heat stored in the first system, but also on the temperature difference between the systems. This leads to the definition of S dependent not just on heat but also on the temperature.
Entropy is defined in terms of infinitesimal change as

$$dS \geq \frac{\delta Q}{T} \tag{2}$$

where the equality appears for the case of reversible processes. The reversible process means a process in which each consecutive state is in the equilibrium with the previous one.
Then, for reversible processes, we can write

$$\delta Q = TdS$$

and then we can rewrite the first law of thermodynamics as

$$dU = TdS - pdV + \mu dN - \vec{m} \cdot d\vec{B} \tag{3}$$

From (2), we can see that if we increase the size of the system by λ, the heat flow into the system increases as λ as well (the heat energy stored is an extensive variable), while the temperature does not change (it is an intensive variable). This implies that S is an extensive variable as well.
This also means that until the system reaches the state when all processes are in equilibrium, entropy has to increase. Therefore, systems evolve by principle of maximizing entropy.

1.1.4 Third Law

Third law of thermodynamics states that as we approach absolute zero temperature, the entropy of the system approaches zero as well, i.e.

$$\lim_{T\to 0} S = 0$$

We will not discuss this law much in this module.

1.2 Legendre Transformations

Suppose we have determined the entropy as a function of some extensive state variable, i.e. $S = S(V)$. This would help us describe the behaviour of the system in terms of its volume, but we might be more interested in describing the system in terms of its pressure, p. Our first guess would be simply find p as a function of V and then find $S = S(V(p))$ to find $S(p)$. We can find p from (3) as

$$p = T\left(\frac{\partial S}{\partial V}\right)_{N,\vec{B},U}$$

We could then invert this equation to find V as a function of p and substitute this back into the S definition to get $S(p)$. However, this process is unrecoverable in a sense that once we translate to $S(p)$, we might loose some information about S.

1.2.1 Loss of Information

To illustrate this, consider an an entropy in form

$$S(V, N) = k_B \log(\alpha V + \gamma N)$$

where α, β and k_B are constants. Suppose also that the system is described by a state equation of form

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

Which evaluates to

$$\frac{\alpha k_B}{\alpha V + \beta N} = C(V, N)\frac{\beta k_B}{\alpha V + \beta N}$$
$$\alpha = \beta C(V, N)$$

Now, we might want to describe S in terms of p. From first law of thermodynamics in zero magnetic field, we know

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{k_B}{\alpha V + \gamma N}\alpha$$
$$p = \frac{\alpha k_B T}{\alpha V + \gamma N}$$

Therefore, we can write

$$S = k_B \log\left(\frac{\alpha k_B T}{p}\right)$$

But, now we have lost any indication of N from this new S. Therefore, we cannot determine the equation of state without the knowledge of $p(V)$, or we might even make a mistake and determine the equation of state as

$$\frac{\partial S}{\partial V} = C(V, N)\frac{\partial S}{\partial N} = 0$$

as we would see that our new $S(p)$ does not depend on S. In order to prevent this loss of information and possibility of desinterpreting the physics, it is more useful to formulate a new state equation for the system we describe by p, which will use a new function of state F, different from S, that will prevent the loss of information. This function F is called the Legendre transform of S.

The basic problem is that as we take the derivative to determine p, any information about the offset of S in a S-V diagram is lost. We can however retain the information about the offset through the specification of the intersect of a line tangent to S-V curve for constant N with the S axis. This is shown in Fig. 1

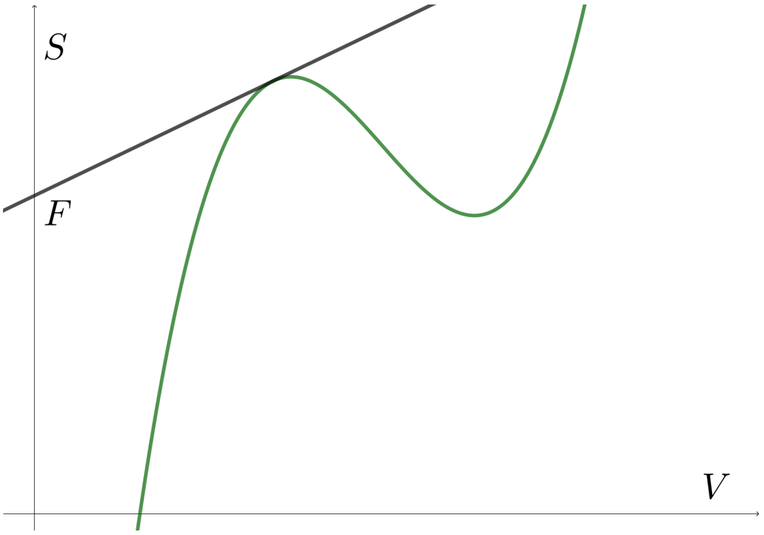

Figure 1: Legendre transform of S is F, marked as the intercept of S axis for a specific V. This makes F function of the same variables as S, while also suitable for being used as a function of derivative of S.

We can see that the tangent line satisfies

$$S = \frac{\partial S}{\partial V}V + F = \frac{p}{T}V + F$$

$$F = S - \frac{p}{T}V = S - V\frac{\partial S}{\partial V} \tag{4}$$

Now, doing the substitution for $S = S(p)$ and $V = V(p)$, we have

$$F = k_B \log\left(\frac{\alpha k_B T}{p}\right) - \frac{p}{T}\left(\frac{k_B T}{p} - \frac{\gamma}{\alpha}N\right) = k_B \log\left(\frac{\alpha k_B T}{p}\right) - k_B + \frac{\gamma p}{\alpha T}N$$

We have now carried out the Legendre trasform of S to F for variable $\frac{p}{T} = \frac{\partial S}{\partial V}$ We already see that there is N present in the new function. Furthermore, we see that

$$\frac{\partial F}{\partial N} = \frac{\gamma p}{\alpha T} = \frac{\partial S}{\partial N}$$

and therefore, we have conserved the information in this new function. Furthermore, we can still determine the state equation, as

$$\frac{\partial S}{\partial V} = \frac{\partial}{\partial V}\left(\frac{p}{T}V + F(p,N)\right) = \frac{p}{T} = C\frac{\partial S}{\partial N} = C\frac{\partial F}{\partial N}$$

And therefore we have not lost any information about the system during this transition.
Inversly, we can calculate S from F if we notice that

$$pV = TS - TF$$

and so

$$\frac{\partial (pV)}{\partial p} = V = \frac{\partial (TS - TF)}{\partial p} = -T\frac{\partial F}{\partial p}$$

$$\frac{-V}{T} = \frac{\partial F}{\partial p}$$

Therefore, we can calculate Legendre transform of F, lets call it G now, due to variable $\frac{-V}{T} = \frac{\partial F}{\partial p}$ as

$$G = F - p\left(\frac{-V}{T}\right) = F + p\frac{V}{T} = S = F - p\frac{\partial F}{\partial p}$$

Therefore, we can get S back from knowledge of $F(p,N)$ only. We can check that

$$G = F - p\frac{\partial F}{\partial p} = F - p\left(k_B\frac{p}{\alpha k_B T}\frac{(-\alpha k_B T)}{p^2} + \frac{\gamma N}{\alpha T}\right) = F - \left(-k_B + \frac{\gamma p}{\alpha T}N\right) = k_B \log\left(\frac{\alpha k_B T}{p}\right)$$

which we know is equal to the entropy S.

1.2.2 Formal Legendre Transforms

If we want to transform function $f(x)$ to its Legendre transform $h(p)$ where $p = \frac{\partial f}{\partial x}$, we write

$$h(p) = f - x\frac{\partial f}{\partial x}$$

The inverse transform is symmetrical, i.e.

$$f(x) = h - p\frac{\partial h}{\partial p}$$

which can be seen by taking derivative with respect to p of the first equation

$$\frac{\partial h}{\partial p} = -x$$

and then substituting for x in the first equation.

Simply by applying the definition, we can derive the formula for transform of function of more than one variable as follows. Lets have function $f(x,y)$, function transformed from x to p called $g(p,y)$ and transform of this function from y to q called $h(p,q)$. From definition

$$g(p,y) = f - x\frac{\partial f}{\partial x}$$

and transforming g

$$h(p,q) = g - y\frac{\partial g}{\partial y} = f - x\frac{\partial f}{\partial x} - y\frac{\partial}{\partial y}\left(f - x\frac{\partial f}{\partial x}\right) = f - x\frac{\partial f}{\partial x} - y\frac{\partial f}{\partial y} - y\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Remembering that $p = \frac{\partial f}{\partial x}$, this function has no explicit dependence on y, so

$$h(p,q) = f - x\frac{\partial f}{\partial x} - y\frac{\partial f}{\partial y} \tag{5}$$

We can perhaps recognize the form of this expression from Hamiltonian mechanics, where we had that the Hamiltonian is given in terms of Lagrangian as

$$H = \sum_i q_i \frac{\partial \mathscr{L}}{\partial \dot{q}_i} - \mathscr{L}$$

which is essentially the minus version of the Legendre transform of the Lagrangian from canonical velocities $\dot{q}_i$ to canonical momenta $\frac{\partial \mathscr{L}}{\partial \dot{q}_i}$.

1.3 Fundamental Equation of Thermodynamics

First law of thermodynamics suggest that we can write the internal energy as function of state variables as follows

$$U = U(S,V,N)$$

and dependence on other variables (T,p,μ) can be determined by the Legendre transformations of U and we will include $-\vec{m}\cdot\vec{B}$ directly to U afterwards.
Since λ and all its arguments are extensive variables, if we increase the system size by ratio λ, we must have

$$\lambda U = U(\lambda S, \lambda V, \lambda N)$$

Taking total derivative with λ leads to

$$U = \left(\frac{\partial U}{\partial S}\right)_{V,N}\frac{d(\lambda S)}{d\lambda} + \left(\frac{\partial U}{\partial V}\right)_{S,N}\frac{d(\lambda V)}{d\lambda} + \left(\frac{\partial U}{\partial N}\right)_{S,V}\frac{d(\lambda N)}{d\lambda} = \left(\frac{\partial U}{\partial S}\right)_{V,N}S + \left(\frac{\partial U}{\partial V}\right)_{S,N}V + \left(\frac{\partial U}{\partial N}\right)_{S,V}N$$

Comparing this to the first law of thermodynamics,

$$dU = TdS - pdV + \mu dN$$

we have

$$U = TS - pV + \mu N \tag{6}$$

This is called the fundamental equation of thermodynamics. It is from here that we can include other types of energy or work into our thermodynamical description easily, for example, we know that the internal energy is changed by $-\vec{m}\cdot\vec{B}$ by presence of magnetic moment and magnetic field, so that

$$U = TS - pV + \mu N - \vec{m}\cdot\vec{B}$$

Similarly, if we had more than one species of particles, we could include them as

$$U = TS - pV + \sum_j \mu_j N_j - \vec{m}\cdot\vec{B}$$

Comparing this with the first law of thermodynamics, we have a conditon on the differentials of state variables

$$0 = SdT - Vdp + \sum_j N_j d\mu_j - \vec{B}\cdot d\vec{m}$$

especially, in zero field, this becomes the Gibbs-Duhem equation

$$SdT - Vdp + \sum_j N_j d\mu_j = 0 \tag{7}$$

which essentially sets a boundary conditions on state variables in the phase space.

1.3.1 Fundamental Thermodynamic Relation Out of Equilibrium

Out of the equilibrium, we have

$$dS \geq \frac{\delta Q}{T}$$

and therefore $\delta Q \leq TdS$ and so

$$dU = \delta Q - \delta W \leq TdS - pdV + \mu dN$$

Hence

$$T \geq \left(\frac{\partial U}{\partial S}\right)_{V,N}$$
$$-p \geq \left(\frac{\partial U}{\partial V}\right)_{S,N}$$
$$\mu \geq \left(\frac{\partial U}{\partial N}\right)_{S,V}$$

And so

$$U = \left(\frac{\partial U}{\partial S}\right)_{V,N} S + \left(\frac{\partial U}{\partial V}\right)_{S,N} V + \left(\frac{\partial U}{\partial N}\right)_{S,V} N \leq TS - pV + \mu N$$
$$U \leq TS - pV + \mu N$$

1.4 Thermodynamical Potentials

Specific situations might require a specific set of variables to describe a certain system. This is especially true when some of the variables are in fact constants by the nature of the system.
For example, consider a system where the volume, number of particles and temperature are fixed (so called canonical ensemble). Instead of using $U = U(S,V,N)$, we might want to switch to a system where the function prescribing the thermodynamics of the system will depend on (T,V,N), because all of these are constants and therefore this function of state will have to be conserved by all processes in the system as well. Therefore, we need to carry out Legendre transformation to

$$F = U - S\left(\frac{\partial U}{\partial S}\right)_{V,N} = U - TS$$

We can quickly check that the differential is

$$dF = dU - TdS - SdT = TdS - pdV + \mu dN - TdS - SdT = -SdT - pdV + \mu dN$$

and since T, V and N are constants, we get the expected result that $dF = 0$. In a non-equilibrium state, we would have

$$dU \leq TdS - pdV + \mu dN$$

and so

$$dF = dU - TdS \leq -SdT - pdV + \mu dN = 0$$

$$dF \leq 0 \tag{8}$$

Therefore, the system will move so that F is minimized. F here is called the thermodynamical potential of the system, and in this specific case it is called the Helmholtz free energy.
We could of course choose different conditions on which variables are constant, and some of these are now listed here.

1.4.1 Microcanonical Ensemble

Here, we suppose that there is no heat flow in the system, which in equilibrium means

$$\delta Q = TdS = 0$$

which is true when we set S constant. Furthermore, we restrict the systems volume and number of particles. Therefore, we see that $U(S,V,N)$ is a constant, and first law states

$$dU = TdS - pdV + \mu dN = 0$$

Again, in non-equilibrium state, $TdS \geq 0$ and $dU \leq 0$ - the internal energy in this system is minimized and U is the potential of this system, ruling the thermodynamics of it.

1.4.2 Pressure Reservoir

Suppose now that we have an isolated system that is kept at constant pressure. This means that the pressure, number of particles and heat flow (and thus entropy) are constants. We therefore search for function $H(S,p,N)$ which will be a potential of the system. By Legendre transform

$$H = U - V\left(\frac{\partial U}{\partial V}\right)_{S,N} = U + pV$$

is the potential of the system. It is called the enthalpy of the system. Again, it is true that at equilibrium, $dH = 0$, otherwise $dH \leq 0$.

1.4.3 Sealed Container

Consider now a system which has a fixed particle number, but can change volume and exchange heat with some reservoir, keeping its pressure and temperature constant. This means that we search for $G(T,p,N)$. Legendre transform U to get

$$G = U - S\left(\frac{\partial U}{\partial S}\right)_{V,N} - V\left(\frac{\partial U}{\partial V}\right)_{S,N} = U - ST + pV$$

This potential is called the Gibbs free energy of the system. Again, $dG \leq 0$ where the equality appears in equilibrium.

1.4.4 Grand Canonical Potential

Lastly, consider a system at fixed T and V, but which is able to exchange particles with a reservoir, which fixes the chemical potential μ. We search for $\Phi(T,V,\mu)$, which leads to

$$\Phi = U - S\left(\frac{\partial U}{\partial S}\right)_{V,N} - N\left(\frac{\partial U}{\partial N}\right)_{S,V} = U - ST - \mu N$$

This is called the grand potential, or also the Landau potential.

1.5 Maxwell Relations

So far, we have been extensively using the first derivatives of the potentials to determine state variables. If we carefuly take second derivatives of potentials, we can then relate first derivatives of some state variables to each other. For example, we know that in a macrocanonical ensemble.

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

and

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

Hence,

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

Since we can exchange the order of partial derivatives, we then have

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

So, we derived that

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

which is a relation not at all obvious from the specification of the system.
In general, if we have differential of the thermodynamical potential in form

$$d\Phi = \sum_i Y_i dX_i$$

We can write Maxwell relations as

$$\forall i,j, i \neq j : \left(\frac{\partial Y_i}{\partial X_j}\right)_{\forall k \neq j} = \left(\frac{\partial Y_j}{\partial X_i}\right)_{\forall k \neq i} \tag{9}$$

1.5.1 Graphical Aid for Finding Maxwell Relations

Suppose that we are given a potential Φ

$$d\Phi = AdX + BdY + CdZ$$

and we are supposed to find

$$\left(\frac{\partial C}{\partial X}\right)_{Y,Z}$$

We underline the symbols that are given in the potential differential, i.e.

$$d\Phi = Ad\underline{X} + BdY + \underline{C}dZ$$

We then draw lines over letters so that each term in the differntial is either with no lines or with both an overline and an underline, i.e.

$$d\Phi = \overline{A}d\underline{X} + BdY + \underline{C}d\overline{Z}$$

The symbols with the overline correspond to the identical Maxwell expression. Therefore, in this case

$$\left(\frac{\partial C}{\partial X}\right)_{Y,Z} = \left(\frac{\partial A}{\partial Z}\right)_{X,Y}$$

The parameters kept constant are always the parameters that appear in the differential form in the differential of the potential. These correspond to the variables that are kept constant by the potential.
To demonstrate this on a physical example, consider finding

$$\left(\frac{\partial N}{\partial V}\right)_{\mu,T}$$

we see that variables kept constant should be V, μ, T, which means that we are dealing with the grand canonical potential. In the differential form

$$d\Phi = dU - TdS - SdT - \mu dN - Nd\mu = -SdT - \underline{N}d\overline{\mu} - \overline{p}d\underline{V}$$

Hence

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

2 Statistical Mechanics

Statistical mechanics is a subject that tries to describe the behaviour of systems with low energy that have a only a limited number of states they can occupy. Instead of reaching a perfect equilibrium state, we assume that the system goes through many somewhat equivalent equilibrium state with probability of existing in a state depends on the distance of the state from the equilibrium and the number of states that correspond to this same distance from equilibrium.
The equivalent states which the system continuously goes through are called the microstates, and the number of microstates is usually dictated by function Ω, which is a function of relevant state variables of the system.
To create a link to classical thermodynamics, we need to define some macroscopic variable in terms of the microstate. The variable chosen is the entropy, which is defined as

$$S = -k_B \sum_{\omega} P(\omega) \ln(P(\omega)) \tag{10}$$

where the sum runs over all microstates ω and $P(\omega)$ is the probability that the state is in a given microstate. This is called the Gibbs definition of entropy.
We then apply the principle of maximizing the entropy to find the equilibrium states. To maximize the entropy, we need to maximize S under the constraint that

$$\sum_{\omega} P(\omega) = 1$$

in order for the probability distribution to be normalized. Using Lagrange multipliers method for some specific microstate γ

$$\frac{\partial}{\partial P(\gamma)}\left(\sum_{\omega} P(\omega) - 1\right) = \lambda \frac{\partial}{\partial P(\gamma)}\left(-k_B \sum_{\omega} P(\omega)\ln(P(\omega))\right)$$

where λ is the Lagrange multiplier. Then

$$1 = -k_B\lambda(\ln(P(\gamma)) + 1)$$

$$1 + k_B\lambda = k_B\lambda\ln(P(\gamma))$$

$$P(\gamma) = e^{\frac{1+k_B\lambda}{k_B\lambda}}$$

To determine λ, we look at the constraint

$$1 = \sum_{\omega} P(\omega) = \sum_{\omega} e^{\frac{1}{k_B\lambda}+1} = e^{\frac{1}{k_B\lambda}+1}\sum_{\omega} 1 = \Omega e^{\frac{1}{k_B\lambda}+1}$$

$$0 = \ln\Omega + \frac{1}{k_B\lambda} + 1$$

$$k_B\lambda = \frac{-1}{\ln\Omega + 1}$$

Hence

$$P(\gamma) = e^{\frac{1}{k_B\lambda}+1} = e^{-\ln\Omega - 1 + 1} = \frac{1}{\Omega}$$

And therefore, the entropy in equilibrium is

$$S = -k_B \sum_{\omega} \frac{1}{\Omega}\ln\frac{1}{\Omega} = k_B \frac{\ln\Omega}{\Omega}\sum_{\omega} 1$$

$$S = k_B \ln\Omega \tag{11}$$

This is the Boltzmann entropy.

2.1 Ensembles in Statistical Physics

In statistical physics, ensembles are specific systems with reservoirs that behave in a well defined way, which usually restricts the changes in some of the state variables. Statistical physics then tries to use different moments of the probability distribution $P(\omega)$ to determine the state functions and variables. For example, the expected internal energy of the system is given by

$$U = \sum_{\omega} U_\omega P(\omega)$$

where U_ω is the energy in a given microstate. However, it shows that the system is usually better described by a different function than the distribution function itself, but rather by the function describing the normalization of $P(\omega)$. This function is called the partition function, and we can use algebraic and calculus manipulation to determine different quantities from the partition function. Of course, partition function differs for different ensembles, as they have different probability distributions $P(\omega)$. Some of the ensembles and their partition functions are now discussed.

2.1.1 Microcanonical Ensemble

In a microcanonical ensemble, we have two systems of similar size, one with energy E_1 and number of microstates $\Omega_1(E_1)$, other with energy E_2 and number of microstates $\Omega(E_2)$. The number of total microstates available to the system as a whole is $\Omega = \Omega_1\Omega_2$, and so the entropy of the whole system is

$$S = k_B\ln\Omega = k_B\ln\Omega_1 + k_B\ln\Omega_2$$

In equilibrium, the entropy will be maximized, so

$$\frac{\partial S}{\partial E_1} = 0 = \frac{\partial S}{\partial E_2}$$

We can assume that the internal energy of the system is conserved, i.e.

$$U = E_1 + E_2$$

is constant, which means that

$$dU = 0 = dE_1 + dE_2$$

$$\frac{dE_1}{dE_2} = -1$$

The conditions for maximization of entropy then become

$$k_B \frac{1}{\Omega_1}\frac{d\Omega_1}{dE_1} + k_B \frac{1}{\Omega_2}\frac{d\Omega_2}{dE_1} = \frac{k_B}{\Omega_1}\frac{d\Omega_1}{dE_1} - \frac{k_B}{\Omega_2}\frac{d\Omega_2}{dE_2} = 0$$

with the second condition being exactly equivalent. Therefore

$$\frac{k_B}{\Omega_1}\frac{d\Omega_1}{dE_1} = \frac{k_B}{\Omega_2}\frac{d\Omega_2}{dE_2}$$

Both sides of this equation are functions of the properties of one system only. Therefore, property of the system θ

$$\theta = \frac{k_B}{\Omega}\frac{d\Omega}{dE}$$

is the same in both systems if they are in equilibrium. With reference to zeroth law of thermodynamics, we then know that for two systems in equilibrium, their temperature T is the same. Therefore, we know that $\theta = \theta(T)$ will also be the same in both systems. Using dimensional analysis, we can determine that simplest θ is

$$\theta = \frac{1}{T}$$

And hence we have a definition of temperature

$$T = \frac{\Omega}{k_B}\frac{dE}{d\Omega}$$

2.1.2 Canonical Ensemble

Suppose now that one of the systems in the case above is much bigger than the other system. Therefore energy of the smaller system, lets say E_1, is always much smaller than energy E_2 and the number of microstates Ω and thus entropy S are both dominated by large system 2. We can therefore say that the entropy in equilibrium, when it is maximised, is

$$S = k_B \ln \Omega_1 + k_B \ln \Omega_2 \approx k_B \ln \Omega_2 \approx k_N \ln \Omega$$

is function of Ω only and therefore a function of U only. Suppose that we now transfer the small energy E_1 into the first reservoir. Taylor expanding the entropy, we have

$$S(E - E_1) = k_B \ln \Omega(E - E_1) \approx k_B \ln \Omega(E) - k_B \frac{1}{\Omega}\frac{d\Omega}{dE} E_1$$

Remembering the definition of temperature, we have

$$k_B \ln \Omega(E - E_1) = k_B \ln \Omega(E) - \frac{E_1}{T}$$

We can state that the probability of this event occuring is proportional to the ratio of number of microstates when all the energy is in the system 2 to the number of microstates when energy E_1 is transferred to system 1, i.e.

$$P(E_1) = C\frac{\Omega(E - E_1)}{\Omega(E)}$$

where C is some constant. Comparing this to the entropy approximation, we have

$$\Omega(E-E_1)=\Omega(E)e^{-\frac{E_1}{k_BT}}$$

writing $\beta=\frac{1}{k_BT}$ and $\frac{1}{Z}=C$, we have

$$P(E_1)=C\frac{\Omega(E-E_1)}{\Omega(E)}=C\frac{\Omega(E)}{\Omega(E)}e^{-\beta E_1}=\frac{e^{-\beta E_1}}{Z}$$

This means that the probability of system 1 occuring in a certian state is completly independent of system 2 properties, as long as the system 2 is big enough. Here, Z is called the partition function. If we have a certain energies E_i the system 1 can take, we have normalization condition

$$1=\sum_i P(E_i)=\sum_i\frac{e^{-\beta E_i}}{Z}$$

$$Z=\sum_i e^{-\beta E_i}$$

and therefore we have determined the partition function. From it, we can derive other properties of the system. Mainly, the expected value of internal energy in system 1 U (changing notation from U reffering to total energy of system 1 and 2) is

$$U=\sum_i E_iP(E_i)=\frac{1}{Z}\sum_i E_ie^{-\beta E_i}=\frac{-1}{Z}\frac{\partial Z}{\partial\beta}$$

Using the Gibbs definition of entropy (10), we can determine the entropy in system 1 alone as (again, changing notation to S instead of S_1)

$$S=-k_B\sum_i P(E_i)\ln P(E_i)=-k_B\sum_i P(E_i)\left(-\beta E_i-\ln Z\right)=$$

$$=k_B\sum_i\beta E_iP(E_i)+k_B\ln Z\sum_i P(E_i)=k_B\ln Z+k_B\beta U=k_B\ln Z-\frac{1}{TZ}\frac{\partial Z}{\partial\beta}$$

Since this is a system which has constant volume and number of particles, but it can interchange energy with a big reservoir at nearly constant temperature, the classical potential for this system would be the Helmholtz free energy. From its definition, we know

$$F=U-TS$$

which in terms of the partition function is

$$F=\frac{-1}{Z}\frac{\partial Z}{\partial\beta}-k_BT\ln Z+\frac{1}{Z}\frac{\partial Z}{\partial\beta}=-k_BT\ln Z$$

Now, since we know the classical thermodynamic potential, we can derive all important properties of the system from it using the differential form of the potential, e.g. the magnetic moment of the system is

$$m=-\left(\frac{\partial F}{\partial B}\right)_{T,V}$$

and similarly for other properties.
One thing we should notice is that the partition function is equal to

$$Z=e^{-\beta F}$$

2.1.3 Grand-Canonical Ensemble

The approach we will use here will be exactly the same as before, but we now allow for the exchange of particles with the big reservoir. The total number of microstates therefore becomes a function of the number of particles as well, and we have

$$\ln\Omega(E-E_i,N-N_i)\approx\ln\Omega(E,N)-\frac{1}{\Omega}\frac{\partial\Omega}{\partial E}E_i-\frac{1}{\Omega}\frac{\partial\Omega}{\partial N}N_i=\ln\Omega(E,N)-\frac{E_i}{k_BT}-\frac{1}{\Omega}\frac{\partial\Omega}{\partial E}\frac{\partial E}{\partial N}N_i$$

But, remembering the definition of chemical potential

$$\mu = -\frac{\partial E}{\partial N}$$

where the minus sign occurs because we are moving particles from the big reservoir to the small reservoir and E refers to the energy of the big reservoir. Hence

$$\Omega(E - E_i, N - N_i) = \Omega(E, N)e^{-\beta(E_i - \mu N_i)}$$

which leads to

$$P(E_i, N_i) = \frac{e^{-\beta(E_i-\mu N_i)}}{\mathscr{Z}}$$

where $\mathscr{Z}$ is the grand-partition function, with

$$\mathscr{Z} = \sum_i e^{-\beta(E_i-\mu N_i)} = \sum_i e^{\beta\mu N_i} Z(N_i)$$

where $Z(N_i)$ is the canonical partition function for the system containing N_i particles. Again, we can derive that

$$N = \sum_i N_i P(E_i, N_i) = \sum_i N_i \frac{e^{-\beta E_i} e^{\beta\mu N_i}}{\mathscr{Z}} = \frac{1}{\beta\mathscr{Z}}\frac{\partial}{\partial\mu}\sum_i e^{-\beta E_i} e^{\beta\mu N_i} = \frac{1}{\beta\mathscr{Z}}\frac{\partial\mathscr{Z}}{\partial\mu}$$

The energy can be derived as

$$U = \sum_i E_i P(E_i, N_i) = \frac{1}{\mathscr{Z}}\sum_i E_i e^{-\beta(E_i-\mu N_i)} = \frac{1}{\mathscr{Z}}\left(-\frac{\partial\mathscr{Z}}{\partial\beta} + \mu\sum_i N_i e^{-\beta(E_i-\mu N_i)}\right) =$$

$$= -\frac{1}{\mathscr{Z}}\left(\frac{\partial\mathscr{Z}}{\partial\beta} - \mu N\mathscr{Z}\right) = -\frac{1}{\mathscr{Z}}\left(\frac{\partial\mathscr{Z}}{\partial\beta} - \frac{\mu}{\beta}\frac{\partial\mathscr{Z}}{\partial\mu}\right) = -\frac{1}{\mathscr{Z}}\frac{\partial\mathscr{Z}}{\partial\beta} + \frac{\mu}{\beta\mathscr{Z}}\frac{\partial\mathscr{Z}}{\partial\mu}$$

The entropy follows from Gibbs entropy

$$S = -k_B\sum_i P(E_i, N_i)\ln(P(E_i, N_i)) = -k_B\sum_i P(E_i, N_i)\left(-\beta E_i + \beta\mu N_i - \ln\mathscr{Z}\right) =$$

$$= k_B\beta\sum_i E_i P(E_i, N_i) - k_B\beta\mu\sum_i N_i P(E_i, N_i) + k_B\ln\mathscr{Z}\sum_i P(E_i, N_i) =$$

$$= k_B\beta U - k_B\beta\mu N + k_B\ln\mathscr{Z}$$

Hence the grand-potential Φ is

$$\Phi = U - TS - \mu N = U - k_B T\beta U + k_B T\beta\mu N - k_B T\ln\mathscr{Z} + \mu N = -k_B T\ln\mathscr{Z}$$

Therefore, we again have

$$\mathscr{Z} = e^{-\beta\Phi}$$

which now clearly starts to become a pattern.

2.2 General Ensemble

Consider now an ensemble where the small system can exchange energy E and a set of extensive variables X_i with the reservoir. This fixes the temperature and intensive variables Y_i which are conjugate variables to X_i. Therefore, the classical potential Φ refering to the small system dynamics will be obtained by Legendre transformation of U as

$$\Phi = U - TS - \sum_i X_i Y_i$$

where I assumed that the extensive-intensive pair is coupled by differential relation

$$Y_i = \left(\frac{\partial U}{\partial X_i}\right)_{j\neq i}$$

where all other extensive variables are kept constant.

The number of microstates will then depend on all the extensive variables that are free to change and on the energy, i.e.

$$\Omega = \Omega(E, \vec{X})$$

where $(\vec{X})_i = X_i$ is the vector of extensive variables. The entropy is dominated by the large system, with extensive variables vector $\vec{X}_0$ and energy E_0. If we then Taylor expand the expression for entropy

$$S(E_0 - E, \vec{X}_0 - \vec{X}) = S(E_0, \vec{X}_0) - \left(\frac{\partial S}{\partial U}\right) E - \sum_i \left(\frac{\partial S}{\partial X_i}\right)_{j \neq i} X_i$$

We can write

$$\left(\frac{\partial S}{\partial X_i}\right)_{j \neq i} = \left(\frac{\partial S}{\partial U}\right)\left(\frac{\partial U}{\partial X_i}\right)_{j \neq i} = \frac{-1}{T} Y_i$$

where I used first law of thermodynamics to determine the derivative with respect to internal energy. We must remember the minus sign, as the energy U here corresponds to the energy leaving the big system. Therefore, using Boltzmann equation for maximised entropy (11)

$$k_B \ln \Omega(E_0 - E, \vec{X}_0 - \vec{X}) = k_B \ln \Omega(E_0, \vec{X}_0) - \frac{E}{T} + \sum_i \frac{1}{T} Y_i X_i$$

$$\ln \Omega(E_0 - E, \vec{X}_0 - \vec{X}) = \ln \Omega(E_0, \vec{X}_0) - \frac{E}{k_B T} + \sum_i \frac{1}{k_B T} Y_i X_i = \ln \Omega(E_0, \vec{X}_0) - \beta E + \beta \sum_i Y_i X_i$$

Therefore, the probability of observing the given state is

$$P(E^\alpha, \vec{X}^\alpha) = \frac{1}{Z} \frac{\Omega(E_0 - E^\alpha, \vec{X}_0 - \vec{X}^\alpha)}{\Omega(E_0, \vec{X}_0)} = \frac{e^{-\beta\left(E^\alpha - \sum_i X_i^\alpha Y_i\right)}}{Z}$$

where α indexes different microstates (and intensive variables Y_i are constant across these states). The partition function can be expressed as

$$Z = \sum_\alpha e^{-\beta\left(E^\alpha - \sum_i X_i^\alpha Y_i\right)}$$

The entropy of the small system can be found from the Gibbs entropy as

$$S = -k_B \sum_\alpha P(\alpha) \ln(P(\alpha)) = -k_B \sum_\alpha P(\alpha) \left(-\beta E^\alpha + \beta \sum_i X_i^\alpha Y_i - \ln Z\right) =$$

$$= k_B \ln Z \sum_\alpha P(\alpha) + \frac{1}{T} \sum_\alpha E^\alpha P(\alpha) - \frac{1}{T} \sum_i Y_i \sum_\alpha P(\alpha) X_i^\alpha$$

Here, we can recognize the expectation values

$$\sum_\alpha E^\alpha P(\alpha) = U$$

$$\sum_\alpha X_i^\alpha P(\alpha) = X_i$$

and therefore

$$S = k_B \ln Z + \frac{U}{T} - \frac{1}{T} \sum_i Y_i X_i$$

And therefore, the potential of this ensemble is

$$\Phi = U - TS - \sum_i X_i Y_i = U - k_B T \ln Z - U + \sum_i X_i Y_i - \sum_i X_i Y_i$$

And we have therefore that for arbitrary ensemble potential

$$\Phi = -k_B T \ln Z \tag{12}$$

and

$$Z = e^{-\beta \Phi}$$

2.2.1 Infinite Many States

If some of the variables X_j can span infinitely many states, then the sum in the partition function becomes an integral. The partition function is

$$Z\sum_{\alpha} e^{-\beta E^{\alpha}+\beta\sum_i X_i^{\alpha} Y_i} = \left(\sum_{\alpha} e^{\beta X_j^{\alpha} Y_j} e^{-\beta E^{\alpha}+\beta\sum_{i\neq j} X_i^{\alpha} Y_i}\right)$$

If we order the sum in such a way that we sum over increments of X_j, we can then write, in the limit where the increments in X_j are infinitesimal

$$Z = \frac{1}{X_j^0}\int_{\forall X_j} dX_j e^{\beta X_j Y_j} \sum_{\gamma(X_j)} e^{-\beta E^{\gamma(X_j)}+\beta\sum_{i\neq j} X_i^{\gamma(X_j)} Y_i}$$

where X_j^0 is some scale for X_j values and the integral runs over all permisible values of X_j and $\gamma(X_j)$ are the indices of microstates available at X_j.
For example, for a system governed by the Gibbs free energy

$$G = U - TS + pV$$

the volume V can grow in arbitrarily small steps. Here, the intensive variable conjugate to V is in fact $-p$, so we have

$$Z = \frac{1}{V_0}\int_0^{\infty} dV e^{-\beta pV} \sum_{\gamma(V)} e^{-\beta E^{\gamma(V)}}$$

3 Phase Transitions

Phase transitions happen when the entropy of some specific arrangement of particles in the system becomes higher than entropy of some other arrangement. Typical examples are the ice melting to water or water boiling to vapor. Usually, they are visualized via the phase diagrams, which show areas in which a specific phase is preffered. The boundaries of these areas are called the coexistence lines, and when the system is setup on these lines, the number of phases that meet here can coexist.

3.1 Conditions on Phase Transitions

Suppose we have a system of two phases. Each of the phases can exchange particles and energy with each other and can take up volume of the other. This means that in equilibrium, the corresponding conjugate variables of these extensive variables has to be kept constant. In the case of energy, we have to relate it to entropy to see that temperature T is the composite variable, otherwise pressure p and chemical potential μ are clear.
Therefore, for phase 1 and phase 2, we have

$$p_1 = p_2$$
$$\mu_1 = \mu_2$$
$$T_1 = T_2$$

To check our reasoning, we can maximize entropy $S = S(U,V,N) = S_1(U_1,V_1,N_1) + S_2(U_2,V_2,N_2)$ as

$$dS = \left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,N_1} dU_1 + \left(\frac{\partial S_1}{\partial V_1}\right)_{U_1,N_1} dV_1 + \left(\frac{\partial S_1}{\partial N_1}\right)_{U_1,V_1} dN_1 +$$
$$+ \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,N_2} dU_2 + \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2,N_2} dV_2 + \left(\frac{\partial S_2}{\partial N_2}\right)_{U_2,V_2} dN_2$$

Since the phases can only exchange the extensive properties with each other, we have

$$dU_2 = -dU_1, dN_2 = -dN_1, dV_1 = -dV_2$$

and therefore

$$dS = \left[\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,N_1} - \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,N_2}\right] dU_1 + \left[\left(\frac{\partial S_1}{\partial V_1}\right)_{U_1,N_1} - \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2,N_2}\right] dV_1 +$$

$$+\left[\left(\frac{\partial S_1}{\partial N_1}\right)_{U_1,V_1}-\left(\frac{\partial S_2}{\partial N_2}\right)_{U_2,V_2}\right]dN_1$$

From the first law of thermodynamics, we have

$$\left(\frac{\partial S}{\partial U}\right)_{V,N}=\frac{1}{T}$$

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

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

and so

$$dS=\left(\frac{1}{T_1}-\frac{1}{T_2}\right)dU_1+\left(\frac{p_1}{T_1}-\frac{p_2}{T_2}\right)dV_1+\left(\frac{\mu_2}{T_2}-\frac{\mu_1}{T_1}\right)dN_1$$

Since we are at equilibrium, entropy is maximized and we require $dS = 0$, which leads to 3 equations on the intesive variables

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

$$\frac{p_1}{T_1}=\frac{p_2}{T_2}$$

$$\frac{\mu_1}{T_1}=\frac{\mu_2}{T_2}$$

which are equivalent with the equations previously stated.
Furthermore, since we require that entropy is maximized, we need that the second differential of the entropy, d^2S, is negative. This derivation is very long, hence I only present the result for specific case of (Gibbs energy)-like behaviour for each phase (when even there is no more particle flow between the phases)

$$\left(\frac{\partial V}{\partial p}\right)_{T,\{N\}_i}\leq 0$$

$$\left(\frac{\partial S}{\partial T}\right)_{p,\{N\}_i}\geq 0$$

$$\forall i:\left(\frac{\partial \mu}{\partial N_i}\right)_{p,T}\geq 0$$

where $\{N_i\}$ is the set of all N_i, if there is more than 1 species of particles present. These conditions set the requirements for the stability of the phase equilibrium - first one ensures stability against pressure fluctuations, second ensures stability against temperature fluctuations and third stability against particle number fluctuations.

3.1.1 Number of Coexisting Phases

If we have $i \in \{1, 2, 3, ...l\}$ indexing chemical species in the system, the chemical potentials at coexistence of phases indexed by $j \in \{1, 2, 3, ..., n\}$ satisfy

$$\mu_{i,j}(p,T,\{N_{k\neq i}\})=\mu_{i,m\neq j}(p,T,\{N_{k\neq i}\})=\mu_{i,p\neq\{m,j\}}(p,T,\{N_{k\neq i}\})=...$$

This is a set of n equations with $l + 2$ unknowns, which means that the solution exists up to $n = l + 2$ phases. This means that we can have equilibrium of up to $l + 2$ phases in the system containing l different chemical species.

3.1.2 Triple Point of Water

As water contains only one chemical species, up to 3 phases can exist at the same time. For the case when we have 3 phases, p and T are exactly set by the requirements given by stability of equilibrium. This occurs at $p = 6 \times 10^{-3}$ atm and $T = 0K$ where the zero is exact - the Kelvin scale is defined with reference to the triple point of water.

3.1.3 Ends of Coexistence Lines

In some phase diagrams, we might see that the coexistence lines disappear - we no longer observe the jump between the phases, rather we observe a continuous change of properties of the system. For example, for water above 374^{o}C and 218 atm, we no longer observe disctinct difference between vapour and liquid. This is called the critical point of water, and region with higher pressure and temperature is calle the supercritical region. The region with higher temperature but lower pressure is called the supercritical fluid region, but there is no distinct phase transition between the two regions. The phase transition that occurs there is continuous, and these types of transitions will be discussed later.

3.2 First Order Phase Transitions

Consider now phase transitions in a Gibbs-like ensemble (isobaric and isothermal). First, we will derive some specific cases of variables for this ensemble.

3.2.1 Variable Concentrations in Gibbs Ensemble

Gibbs free energy is given by Legendre transform of U to varaibles p and T which are fixed by reservoir as

$$G = U - TS + pV$$

Using fundamental equation of thermodynamics (6), we have

$$G = \mu N$$

or for more particle species

$$G = \sum_i \mu_i N_i$$

The differential of the Gibbs free energy is

$$dG = dU - TdS - SdT + pdV + Vdp = Vdp - SdT + \mu dN$$

So, the changing variables are defined as

$$\mu = \left(\frac{\partial G}{\partial N}\right)_{T,p}$$

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N}$$

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

Using the expression for μ, we can substitute this back to expression for G to find

$$G = N\mu = N\left(\frac{\partial G}{\partial N}\right)_{T,p} = N\mathscr{G}$$

where $\mathscr{G}$ is the Gibbs free energy concentration - Gibbs free energy per particle. We can then substitute this definition of Gibbs energy back to expressions for S and V

$$S = -\left(\frac{\partial}{\partial T}\right)_{p,N}\left[N\left(\frac{\partial G}{\partial N}\right)_{T,p}\right]$$

Since the outer derivative is taken at constant N, we can move N in front of the outer derivative to get

$$S = -N\left(\frac{\partial}{\partial T}\right)_{p,N}\left(\frac{\partial G}{\partial N}\right)_{T,p}$$

Swapping the order of partial derivatives leads to

$$S = -N\left(\frac{\partial}{\partial N}\right)_{T,p}\left(\frac{\partial G}{\partial T}\right)_{p,N} = N\left(\frac{\partial S}{\partial N}\right)_{T,p} = N\mathscr{S}$$

where $\mathscr{S}$ is the entropy concentration - entropy per particle.
Similarly

$$V = \left(\frac{\partial}{\partial p}\right)_{T,N}\left[N\left(\frac{\partial G}{\partial N}\right)_{p,T}\right] = N\left(\frac{\partial V}{\partial N}\right)_{p,T} = N\mathscr{V}$$

with $\mathscr{V}$ representing the volume per particle.
Finally, we can invoke Maxwell relations to find

$$\mathscr{S} = \left(\frac{\partial S}{\partial N}\right)_{T,p} = -\left(\frac{\partial \mu}{\partial T}\right)_{N,p}$$

as

$$dG = -\underline{S}d\overline{T} + Vdp + \overline{\mu}d\underline{N}$$

Similarly

$$\mathscr{V} = \left(\frac{\partial V}{\partial N}\right)_{T,p} = \left(\frac{\partial \mu}{\partial p}\right)_{T,N}$$

3.2.2 Characterization of First Order Transitions

In first order transitions, the derivative of the chemical potential with temperature is discontinuous - we observe a sharp edge in the function $\mu = \mu(T)$ at the phase transition. This means that there is a finite step in the entropy per particle

$$\Delta\mathscr{S} = \left(\frac{\partial \mu_1}{\partial T}\right)_{N,p} - \left(\frac{\partial \mu_2}{\partial T}\right)_{N,p}$$

for the change from phase 1 to phase 2 (the signs have swapped due to definition of $\mathscr{S}$).
Hence, across the transition, there is a change in the total entropy

$$\Delta S = N\Delta\mathscr{S}$$

which leads to a latent heat that needs to be provided for the transition to occur, given from the second law of thermodynamics (2)

$$\Delta Q = T_C \Delta S = T_C N \Delta\mathscr{S}$$

where T_C is the temperature of the transition. During the transition, we therefore have to provide finite amount of heat for no observable change in temperature - the heat capacity $C = \frac{\partial Q}{\partial T}$ diverges at the phase transition T_C.
We should finally note that T_C will be a function of pressure. Therefore, we can see that $\mathscr{V}$ will be discontinuous as well at the phase transition, and we will have finite change of volume for no change in pressure, meaning that the isothermal compressibility of the gas will diverge.

3.3 Continuous Phase Transitions

In continuous phase transitions, $\mathscr{S}$ and $\mathscr{V}$ do not change discountinuously, but change very rapidly at the phase transition, usually with the second derivative of μ diverging at the phase transition (i.e. first derivative of S with T diverging at the phase transition). The illustration is presented in Fig. 2 and Fig. 3.
Continuous phase transitions represent a state of matter when small changes in p or T lead to big changes in properties of the system. Therefore, they are usually accompanied by emergence of some long scale oscillations and correlations in the system, as will be discussed later.
Furthermore, we can in some sense see continuous phase transitions as a special kind of first order phase transition where the gap $\Delta\mathscr{S}$ just goes to 0. In this way, we can imagine how for example the water-vapour phase transition becomes a continuous phase transition above the critical point - the entropy gap could exactly close on the critical point.

3.4 p-V diagrams and Metastability

In contrast with $p - T$ diagrams, $p - \mathscr{V}$ diagrams (where $\mathscr{V} = \frac{V}{N}$) contain the coexistence regions rather than coexistence lines. If a system is within a coexistence region on $p - \mathscr{V}$ diagram, it can either be in an inherently unstable state, where any arbitrarily small perturbation will offset it to separate into two phases, or in a metastable region, where the perturbation must reach some finite magnitude to cause the separation of phases.

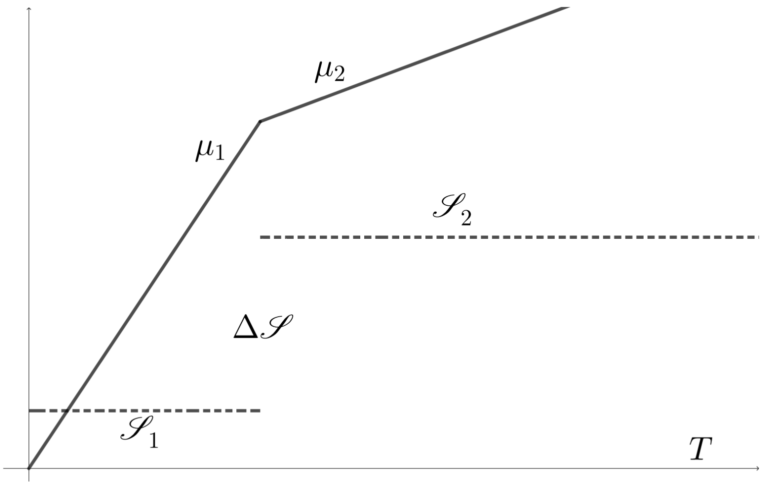

Figure 2: First order phase transition. Entropy per particle $\mathscr{S}$ is discontinuous across the phase transition. Here, I used $\mu = \alpha T$ for $T < T_C$ and $\mu = \beta T$ for $T \geq T_C$. The whole $\mathscr{S}$ is shifted upwards.

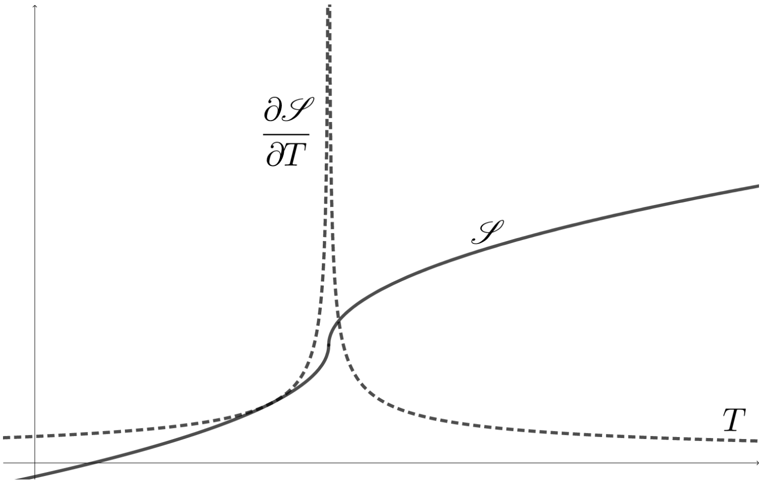

Figure 3: Second order phase transition. At T_C, the derivative diverges while $\mathscr{S}$ remains continuous. In this case, I used $\mathscr{S} = \sqrt{T - T_C}$ for $T \geq T_C$ and $\mathscr{S} = \sqrt{T_C - T}$ for $T < T_C$.

In order to bring the system to a metastable state, we must minimize the perturbations in the system, so that it cannot separate phases. This is usually done by cooling/heating the system very slowly. Once in the metastable region, if the system is to separate, it does so by the so called lever rule.

3.4.1 Lever Rule

Suppose we have a system at constant pressure p which has been brought to a metastable state with volume per particle $\mathscr{V}_m$, which is larger than the largest volume per particle of the denser phase $\mathscr{V}_1$ and smaller than smallest volume per particle of a lighter phase $\mathscr{V}_2$. If the phases are to separate, what will be the ratio of particles in each phase? We can interpret the metastable state as a mixture of both phases. The total number of particles in the state is $N = N_1 + N_2$, and the total volume of the system is $V = V_1 + V_2 = \mathscr{V}_1 N_1 + \mathscr{V}_2 N_2$ Therefore, the metastable concentration is

$$\mathscr{V}_m = \frac{V}{N} = \frac{\mathscr{V}_1 N_1 + \mathscr{V}_2 N_2}{N_1 + N_2}$$

$$\mathscr{V}_m (N_1 + N_2) = \mathscr{V}_1 N_1 + \mathscr{V}_2 N_2$$

$$N_1 (\mathscr{V}_m - \mathscr{V}_1) = N_2 (\mathscr{V}_2 - \mathscr{V}_m)$$

$$\frac{N_1}{N_2} = \frac{\mathscr{V}_2 - \mathscr{V}_m}{\mathscr{V}_m - \mathscr{V}_1} \tag{13}$$

Usually, the metastable states are less stable the further away from $\mathscr{V}_2$ or $\mathscr{V}_1$ we go. Therefore, the coexistence regions have two defined subregions. The outer boundary of the coexistence region is the binodal line, which marks the beginning of the metastable region. Further inwards, there is a spinodal line, which marks the start of truly unstable region, where the phases cannot be metastable. The stability of the state decreases as we approach the spinodal line.

3.5 General Continuous Phase Transitions

In many experimental situations, we observe a similar dependence for seemingly different phenomena. For example, we observe that any binodal line between gas and liquid can be characterized as

$$\frac{\rho_{liquid} + \rho_{gas}}{2\rho_C} = 1 + \frac{3}{4}\left(1 - \frac{T}{T_C}\right)$$

where T_C is the critical temperature of the transition and ρ are the mass densities of the liquid, gas and at the critical transition.
This lead to a hypothesis that all critical phenomena share a distinct set of properties which are very general and independent of the sytem we try to describe.

3.5.1 Properties of Continuous Phase Transitions

Continuous phase transitions exhibit very large long range fluctuations and correlations as we are closing to the critical point. These fluctuations have typically a set dimensionless length connected with them. Depending on the dimensionless length and the speed of divergence of these fluctuations as we approach the critical point, we separate the continuous transitions into what is called the universality classes.
Each universality class has a critical exponent that characterizes the divergence of the correlations and behaviour of other properties close to the critical point.

3.6 Spontaneous Symmetry Breaking

A specific set of continuous phase transitions is the spontaneous symmetry breaking. During this transition, a certain symmetry is broken - an asymmetric state is chosen by the system without any outside interference. Usually, we say that certain order parameter, which describes how this symmetry is broken, emerges at the SSB transition.
For example, if we cool a ferromagnet below certain critical temperature T_C, it chooses one direction for the magnetisation $\vec{M}$, and thus breaks rotational symmetry of the original system, which had $\vec{M} = 0$.
In SSB transitions, the nature of the symmetry also plays a big role. For continuous symmetries (translation, rotation), arbitrarily small fluctuations lead to long range correlations during the transition, which leads to observation of so called Goldstone modes, which have size dependent on dimensionless properties of the system. Discrete symmetries (mirror image, rotation by a set number of degrees) do not share this property.

3.6.1 Ising Model

To illustrate the concepts of spontaneous symmetry breaking, we will discuss the mean field approximation of the Ising model of ferromagnetism. In the Ising model, we have a cubic lattice of atoms with spins $S_i = \pm 1$ in d dimensions (effectively a crystal in $d+1$ dimensions where all the spins are aligned along one axis, reducing the dimensionality of the problem to d). This means that each atom has $2d$ neighbours. Suppose that there is N of these atoms. Then the number of unique nearest-neighbour pairs is $N_{nn} = Nd$, i.e. one half of the all possible pairs.
The energy of each microstate of the magnet is given by

$$E = -J \sum_{<i,j>} S_i S_j - B \sum_i S_i \tag{14}$$

where the first sum runs over all nearest neighbour pairs. Here, since S_i is dimensionless, J and B both have the units of energy. If we wanted to dimensionalize back to SI units, we can divide B by a Bohr magneton, which will give S in units of magnetic moment. Right now, we will keep working in dimensionless variant.
Suppose the system is kept at constant temperature and volume and number of particles - it form a canonical ensemble. The partition function is then

$$Z = \sum_{\vec{S}} e^{-\beta E(\vec{S})} = \sum_{\vec{S}} e^{\beta\left(J \sum_{<i,j>} S_i S_j + B \sum_i S_i\right)}$$

where $\vec{S}$ is the vector giving value of each S_i - it is the microstate of the system.
This is generally hard to solve. It has been solved in 1D and 2D, but in higher dimensions, not yet. Therefore, we apply approximations, with mean field being the standard one. In the mean field approximation, we expect that spins have some overall magnetisation $m =< S >$ (here, m is dimensionless). The mean field helps us reduce the energy expression as

$$E = -J \sum_{<i,j>} (m + (S_i - m))(m + (S_j - m)) - B \sum_i S_i =$$

$$= -J \sum_{<i,j>} \left(m^2 + m\left[S_i - m + S_j - m\right] + (S_i - m)(S_j - m)\right) - B \sum_i S_i =$$

$$= -J \sum_{<i,j>} \left(-m^2 + m(S_i + S_j) + (S_i - m)(S_j - m)\right) - B \sum_i S_i$$

The first term in the first sum evaluates to

$$\sum_{<i,j>} -m^2 = -m^2 \sum_{<i,j>} 1 = -Ndm^2$$

as there is Nd nearest neighbours. The second term evaluates to

$$\sum_{<i,j>} m(S_i + S_j) = m \sum_{<i,j>} S_i + m \sum_{<i,j>} S_j = md \sum_i S_i + md \sum_j S_j = 2md \sum_i S_i$$

as each of the spins have d unique nearest neighbours (half of $2d$). The third term goes to zero

$$\sum_{<i,j>} (S_i - m)(S_j - m) = 0$$

As this means that over large distances, the field is essentially the mean field m, which is our requirement for mean field theory to apply. We will later see that this term correponds to the correlations of the spins, and its negligence in some cases leads to a failure of mean field theories.
With these terms evaluated, the energy becomes

$$E = -J(-Ndm^2 + 2md \sum_i S_i) - B \sum_i S_i$$

$$E = JNdm^2 - (2Jmd + B) \sum_i S_i \tag{15}$$

Therefore, the partition function is

$$Z = \sum_{\vec{S}} e^{-\beta(JNdm^2 - (2Jmd+B)\sum_i S_i)} = e^{-\beta JNdm^2} \sum_{\vec{S}} e^{\beta(2Jmd+B)\sum_i S_i} =$$

$$= e^{-\beta JNdm^2} \sum_{\vec{S}} \prod_i e^{\beta(2Jmd+B)S_i}$$

Suppose that $\vec{S} = (S_1, S_2, ..., S_N) = (S_1, \vec{S}_2^R) = (S_1, S_2, \vec{S}_3^R)$ where I defined $\vec{S}_i^3 = (S_i, S_{i+1}, ..., S_N)$ as the remainder of the vector starting with the ith element. Then, we can order the sum in the following way

$$\sum_{\vec{S}} \prod_i e^{\beta(2Jmd+B)S_i} = \sigma = \sum_{\vec{S}_1^R} \prod_{i=1}^N e^{\beta(2Jmd+B)S_i} =$$

$$= \sum_{S_1=1, \vec{S}_2^R} e^{\beta(2Jmd+B)S_1} \prod_{i=2}^N e^{\beta(2Jmd+B)S_i} + \sum_{S_1=-1, \vec{S}_2^R} e^{-\beta(2Jmd+B)S_1} \prod_{i=2}^N e^{\beta(2Jmd+B)S_i}$$

where we took into account both possibilities for S_1 and add these terms together first. This becomes

$$\sigma = \sum_{\vec{S}_2^R} e^{\beta(2Jmd+B)} \prod_{i=2}^N e^{\beta(2Jmd+B)S_i} + \sum_{\vec{S}_2^R} e^{-\beta(2Jmd+B)} \prod_{i=2}^N e^{\beta(2Jmd+B)S_i} =$$

$$= e^{\beta(2Jmd+B)} \sum_{\vec{S}_2^R} \prod_{i=2}^{N} e^{\beta(2Jmd+B)S_i} + e^{-\beta(2Jmd+B)} \sum_{\vec{S}_2^R} \prod_{i=2}^{N} e^{\beta(2Jmd+B)S_i} =$$

$$= \left(e^{\beta(2Jmd+B)} + e^{-\beta(2Jmd+B)}\right) \sum_{\vec{S}_2^R} \prod_{i=2}^{N} e^{\beta(2Jmd+B)S_i} = 2\cosh(\beta(2Jmd+B)) \sum_{\vec{S}_2^R} \prod_{i=2}^{N} e^{\beta(2Jmd+B)S_i}$$

We have therefore shown that

$$\sigma = \sum_{\vec{S}_1^R} \prod_{i=1}^{N} e^{\beta(2Jmd+B)S_i} = 2\cosh(\beta(2Jmd+B)) \sum_{\vec{S}_2^R} \prod_{i=2}^{N} e^{\beta(2Jmd+B)S_i}$$

We therefore might be able to see that we have factored out part of the sum. In fact, we can similarly show that for any j

$$\sum_{\vec{S}_j^R} \prod_{i=j}^{N} e^{\beta(2Jmd+B)S_i} = e^{\beta(2Jmd+B)} \sum_{\vec{S}_{j+1}^R} \prod_{i=j+1} e^{\beta(2Jmd+B)S_i} + e^{-\beta(2Jmd+B)} \sum_{\vec{S}_{j+1}^R} \prod_{i=j+1} e^{\beta(2Jmd+B)S_i} =$$

$$= 2\cosh(\beta(2Jmd+B)) \sum_{\vec{S}_{j+1}^R} \prod_{i=j+1} e^{\beta(2Jmd+B)S_i}$$

Hence, we can infer that

$$\sigma = \sum_{\vec{S}_1^R} \prod_{i=1}^{N} e^{\beta(2Jmd+B)S_i} = 2\cosh(\beta(2Jmd+B)) \sum_{\vec{S}_2^R} \prod_{i=2}^{N} e^{\beta(2Jmd+B)S_i} =$$

$$= (2\cosh(\beta(2Jmd+B)))^2 \sum_{\vec{S}_3^R} \prod_{i=3}^{N} e^{\beta(2Jmd+B)S_i} = \dots = (2\cosh(\beta(2Jmd+B)))^{N-1} \sum_{\vec{S}_N^R} \prod_{i=N}^{N} e^{\beta(2Jmd+B)S_i} =$$

$$= (2\cosh(\beta(2Jmd+B)))^{N-1} \left(e^{\beta(2Jmd+B)} + e^{-\beta(2Jmd+B)}\right) = (2\cosh(\beta(2Jmd+B)))^{N}$$

Therefore, we finally have the partition function

$$Z = e^{-\beta JNdm^2}\sigma = e^{-\beta JNdm^2} \left(2\cosh\left(\beta(2Jmd+B)\right)\right)^N \tag{16}$$

The thermodynamic potential is the Helmholtz free energy as this is a canonical ensemble. So, we have

$$F = -k_BT\ln Z = -k_BT\left(-\beta JNdm^2 + \ln\left((2\cosh(\beta(2Jdm+B)))^N\right)\right) =$$

$$= JNdm^2 - Nk_BT\ln\left(\cosh(\beta(2Jdm+B))\right) - Nk_BT\ln 2$$

In the presence of magnetic field and total magnetisation M, the internal energy of the system is modified as $dU \to dU - MdB$, and therefore the differential of the Helmholtz free energy is

$$dF = dU - TdS - SdT = -SdT - pdV + \mu dN - MdB$$

Hence, as N and V are constant, we have

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

which in this case leads to

$$M = Nk_BT\frac{1}{\cosh(\beta(2Jdm+B))}\sinh(\beta(2Jdm+B))\beta = N\tanh(\beta(2Jdm+B))$$

We have said that m is the average magnetisation $< S >$, and so we can write $M = mN$, and so we have

$$m = \tanh(\beta(2Jdm+B)) \tag{17}$$

This is a self-consistency criterion for the Ising model, which will help us determine what are the favourable values of magnetisation m the system takes. We are particularly interested in the case when $B = 0$, and therefore the problem should be symmetrical and unless the symmetry is broken, the system should not choose a specific m, as there is no preffered direction.
In the case when $B = 0$, we have

$$m = \tanh(2Jdm\beta)$$

We can clearly see that one solution is at $m = 0$. However, there can be two other solutions. As $\lim_{x\to\pm\infty} \tanh(x) = \pm 1$, if at 0 the tangent to $\tanh(2Jdm\beta)$ has gradient smaller than gradient of m, which is 1, we will have two other solutions, as shown in Fig. 4

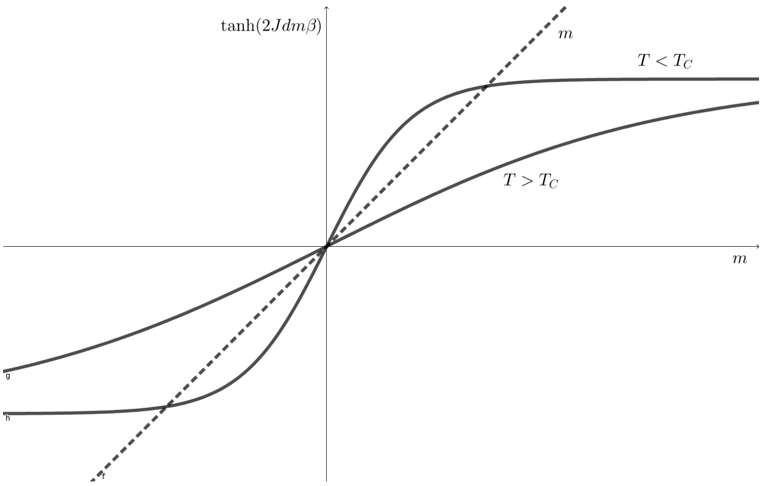

Figure 4: The magnetisation in the Ising model. The dashed line is the line of m, the full lines are lines of $\tanh(2Jdm\beta)$. We see that the number of solutions depends on the slope of $\tanh(2Jdm\beta)$ at $m = 0$.

The gradient of tangent to $\tanh(2Jdm\beta)$ can be obtained as the derivative at 0

$$g(m) = \frac{\partial(\tanh(2Jdm\beta))}{\partial m} = \frac{1}{\cosh^2(2Jdm\beta)} 2Jd\beta$$

At $m = 0$

$$g(0) = 2Jd\beta = \frac{2Jd}{k_B T}$$

Hence, if $g(0) < 1$, we have only one solution with $m = 0$, but if $g(0) > 1$, we have three solutions. The new solutions start to occur at $g(0) = 1$, which corresponds to temperature

$$T_C = \frac{2Jd}{k_B} \tag{18}$$

Here, we give the critical temperature a special name - the Curie temperature T_C.
Finally, to figure out in which of these three solutions the system chooses to be, we need to evaluate the first and second derivative of F to find which solution corresponds to the minimum of F. The first derivative with respect to m is

$$\frac{\partial F}{\partial m} = 2JNdm - Nk_B T \tanh(\beta(2Jdm))2Jd\beta =$$
$$= 2JNdm - 2JNd \tanh(\beta(2Jdm)) = 2JNd\,(m - \tanh(\beta(2Jdm)))$$

we can therefore see that the $\frac{\partial F}{\partial m} = 0$ when $m = \tanh(\beta(2Jdm))$ - these are all the solutions we require from self-consistency - all of these extremise F. The question remains, which are maxima and which are minima. The second derivative is

$$\frac{\partial^2 F}{\partial m^2} = 2JNd\left(1 - \frac{2Jd\beta}{\cosh^2(\beta 2Jdm)}\right)$$

We can use hyperbolic identity

$$\frac{1}{\cosh^2(\beta 2Jdm)} = 1 - \tanh^2(\beta 2Jdm)$$

And from the self-consistency criterion (17), we find that at $B = 0$, for which we are solving the problem

$$\frac{1}{\cosh^2(\beta 2Jdm)} = 1 - m^2$$

And so we have

$$\frac{\partial^2 F}{\partial m^2} = 2JNd\left(1 - 2Jd\beta(1-m^2)\right)$$

We are searching for minima, so we are looking at which m the second derivative will be positive (corresponding to minimum of F), i.e.

$$0 < 2JNd\left(1 - 2Jd\beta(1-m^2)\right)$$

$$0 < 1 - 2Jd\beta + 2Jd\beta m^2$$

Using $T_C = \frac{2Jd}{k_B}$, we have

$$\frac{T_C}{T}m^2 > \frac{T_C}{T} - 1$$

$$m^2 > 1 - \frac{T}{T_C}$$

We can see that for $m = 0$, this applies when

$$0 > 1 - \frac{T}{T_C}$$

$$\frac{T}{T_C} > 1$$

We see that $m = 0$ is the minimum of the system for $T > T_C$. Conversly, if $m \neq 0$, we have $m^2 > 0$, which can only happen when $T < T_C$. Therefore, we have the expected behaviour - spontaneous magnetization below T_C and no magnetization above T_C. Finally, we can see that by substitution of $m = 0$ for $T < T_C$ into the original equation, we get that m becomes the local maximum of F.

As a remark, in a non-zero field, the behaviour is the same, but the Curie temperature T_C is shifted to $T_C \to T_C + \frac{B}{k_B}$, i.e. it grows, because it is easier for spins to overcome thermal fluctuations and magnetize.

3.6.2 Ising Model Close to Transition

The question now is how exactly does the state $m = 0$ transition from being the minimum of F to being local maximum of F? We can write F for small m

$$F = JNdm^2 - Nk_BT\ln(\cosh(\beta 2Jdm)) - Nk_BT\ln 2$$

Using $\cosh(x) \approx 1 + \frac{x^2}{2} + \frac{x^4}{24}$ for small x

$$F \approx JNdm^2 - Nk_BT\ln\left(1 + \frac{1}{2}(2\beta Jdm)^2 + \frac{1}{24}(2\beta Jdm)^4\right) - Nk_BT\ln 2$$

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

$$F \approx JNdm^2 - Nk_BT\ln 2 - Nk_BT\left(\frac{1}{2}(2\beta Jdm)^2 + \frac{1}{24}(2\beta Jdm)^4 - \frac{1}{2}\left(\frac{(2\beta Jdm)^2}{2} + \frac{(2\beta Jdm)^4}{24}\right)^2\right) \approx$$

$$\approx JNdm^2 - Nk_BT\ln 2 - Nk_BT\left(\frac{1}{2}(2\beta Jdm)^2 + \frac{1}{24}(2\beta Jdm)^4 - \frac{1}{2}\frac{(2\beta Jdm)^4}{4}\right) \approx$$

$$\approx JNdm^2 - Nk_BT\ln 2 - \frac{1}{2}Nk_BT(2\beta Jd)^2m^2 + \frac{N}{12}k_BT(2\beta Jd)^4m^4$$

Remembering that $T_C = \frac{2Jd}{k_B}$ (18), we have

$$F \approx -Nk_BT\ln 2 + \frac{N}{2}k_BT_Cm^2 - \frac{N}{2}k_BT\left(\frac{T_C}{T}\right)^2 m^2 + \frac{N}{12}k_BT\left(\frac{T}{T_C}\right)^4 m^4$$

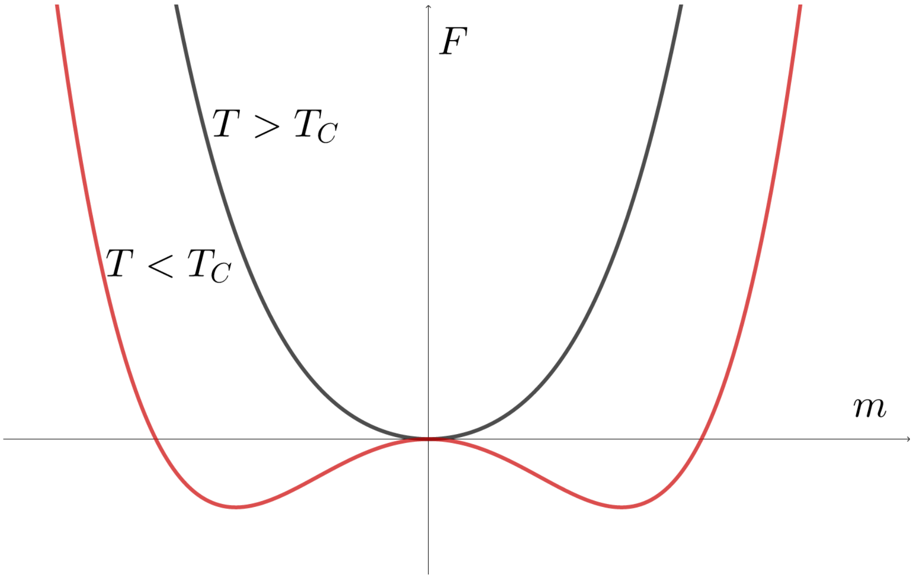

Figure 5: Spontaneous symmetry breaking as consequence of the form of the thermodynamic potential in terms of order parameter. For the case of Ising model, the order parameter is the magnetization m and phase transition occurs at Curie temperature T_C.

Hence

$$F = -Nk_BT\ln 2 + \frac{N}{2}k_B\frac{T}{T_C}(T-T_C)m^2 + \frac{N}{12}k_BT\left(\frac{T}{T_C}\right)^4 m^4 \tag{19}$$

The shape of this function depends largly on the sign of the coefficient before m^2. If the coefficient is positive, then $m = 0$ is the minimum, while if the coefficient is negative, $m = 0$ is local maximum, as shown in Fig. 5

Therefore, we have figured out how the transition from maximum to minimum occurs. In fact, Ginzburg and Landau created a theory that this applies generally for a symmetry breaking transitions, as will be discussed later.

3.6.3 Caveats of Ising Model

The exact solution of the Ising model in 1D suggest that there is no spontaneous magnetization in 1D. Where does the meanfield theory than make the mistake? It is in neglecting the correlation term. As it turns out, in 1D, if we have enough energy to flip one spin to anti-aligned position, we can flip the spins neighbouring to this one without any energy cost. Therefore, we can create arbitrarily long spin flips with a small amount of energy. This means that in 1D, we have divergent correlation length between the spin fluctuations, and therefore we cannot disregard the correlation term $\sum_{<i,j>}(S_i - m)(S_j - m)$.

It is a general property of mean field theories that we have a certain lower critical and upper critical number of dimensions d_C and D_C, respectivelly. Below d_C, the mean field theory does not provide neither qualitative nor quantitative description of the system. Between d_C and D_C, the qualitative behaviour can be distinguished, but quantitative values are wrong. Above D_C, both qualitative and quantitative descriptions are accurate. For the Ising model, $d_C = 1$ and $D_C = 4$.

3.7 Ginzburg-Landau Spontaneous Symmetry Breaking Phase Transitions

In this formalism, we have a free energy that has to be symmetrical due to properties of the system in high symmetry phase. In the high symmetry phase, the symmetrical state of the system is also preffered. However, as we lower temperature towards critical temperature T_C, the symmetrical state of the system becomes unstable, even though the free energy is still symmetrical in the original symmetry. Therefore, system spontaneously decays into the new minima of the free energy, breaking the symmetry.

This break in symmetry is characterized by the emergence of symmetry breaking order parameter, generally some η. The general free energy of the system susceptible to SSB (spontaneous symmetry breaking) can be then expressed as

$$\Phi = \Phi_0(T,Y) + \Phi_2(T,Y)(T - T_C)\eta^2 + \Phi_4(T,Y)\eta^4 \tag{20}$$

where Y are some other intensive properties of the system. Hence, the symmetrical state is preffered as long as $T > T_C$, asymmetry is chosen when $T < T_C$.
The value of the order parameter after the symmetry breaking is given by

$$\frac{\partial\Phi}{\partial\eta} = -2\Phi_2(T_C - T)\eta + 4\Phi_4\eta^3 = 0$$

since $\eta = 0$ is not the minimum, we are left with two other solutions

$$4\Phi_4\eta^2 = 2\Phi_2(T_C - T)$$

$$\eta = \pm\sqrt{\frac{\Phi_2}{2\Phi_4}(T_C - T)}$$

This applies generally for many cases of SSB.

3.7.1 Asymmetric Free Energy

We can also observe emergence of order parameter η in asymmetric free energies of type

$$\Phi = \Phi_0 + \Phi_2(T - T_C)\eta^2 + \Phi_3\eta^3 + \Phi_4\eta^4$$

This energy has three local extrema

$$\frac{\partial\Phi}{\partial\eta} = 2\Phi_2(T - T_C)\eta + 3\Phi_3\eta^2 + 4\Phi_4\eta^3 = 0$$

one at zero (again), and other two appear at

$$2\Phi_2(T - T_C) + 3\Phi_3\eta + 4\Phi_4\eta^2 = 0$$

$$\eta = \frac{-3\Phi_3 \pm \sqrt{9\Phi_3^2 - 32\Phi_2\Phi_4(T - T_C)}}{8\Phi_4}$$

Importantly, one of these extrema is a minimum which becomes global minimum before T reaches T_C. Therefore the system will tend to change to this minimum in a discontinuous jump across some free energy barrier, if the fluctuations provide enough energy. This is typical for the first order phase transitions, which can be described like this. To illustrate this, consider Fig. 6

4 Polymers

Polymers are long, usually 1D molecules that are on the troublesome mesoscale of lengths - too big to be approximated as a continuum and too big to be treated quantum-mechanically. Here, we will look at ideal polymers in a canonical ensemble. The ideal approximation here means that the polymers are non-interacting - the internal energy of the polymer is zero. Therefore, the polymer dynamics are driven solely by the principle of entropy maximization.
We assume that the polymers bonds are chosen in a random direction between each monomer, with some probability distribution. For ideal polymer, we assume that this probability distribution is same for all the bonds (memoryless property). Furthermore, we will assume here that the polymer is linear - every monomer is connected at most by two bonds, i.e. no branches are created.
To describe the polymer, we specify the locations of each monomer with reference to the edge monomer with which the polymer chain starts. The reference position of the ith monomer is given as

$$\vec{r}_i = \sum_{j=1}^{i} \vec{u}_j$$

where $\vec{u}_j$ is the vector corresponding to the jth bond - $\vec{u}_j = \vec{u}_{j+1} - \vec{u}_j$. We define that the starting monomer is at $\vec{r}_0 = \vec{0}$.
Usually, to characterize the polymer, we use several quantities. One is the overall average displacement $< \vec{r}_N >$, where N is the number of bonds ($N + 1$ is the number of monomers). Then, we have the average distance squared between start and end $< \vec{r}_N^2 >$, the centre of mass displacement $\vec{R}_{CM} = \frac{1}{N}\sum_i \vec{r}_i$ and radius of gyration $R_G^2 =< \frac{1}{N}\sum_i(\vec{r}_i - \vec{R}_{CM})^2 >$. Typically, we also talk about the contour length - the total length of the monomer $L = \sum_i |\vec{u}_i|$. Here, $< x >$ means the mean value of variable x.
We will now try to determine these quantities in some basic models of polymers.

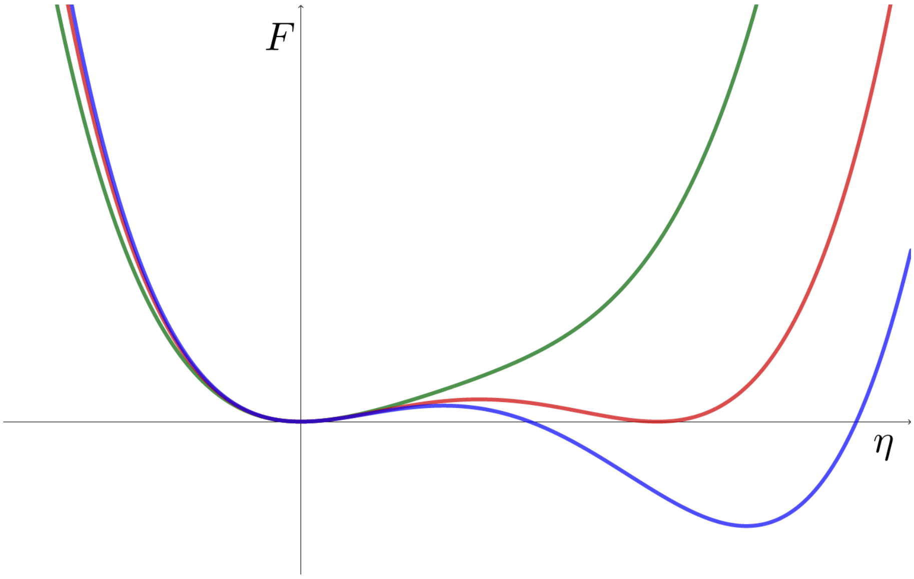

Figure 6: Spontaneous symmetry breaking in first order transition does not occur, because symmetry is broken by the free energy itself, and the system discountinuously jumps over the energy barrier if the fluctuations are high enough. If they are not, the system remains in a metastable state. The green line is for $T > T_C$, red line is for $T = T_C$ and green line is for $T < T_C$.

4.1 Freely Jointed Chain

For the freely jointed chain, we assume that the direction of each bond is chosen completly randomly, while the bond length is constant. This means that the bonds can potentailly cross each other - this is truly a simple model. If the bonds are chosen completly randomly, the direction of each bond is uncorrelated to the others, which means that

$$< \vec{u}_i \cdot \vec{u}_j >= \delta_{ij} < \vec{u}_i^2 >$$

Since the bond length is constant, we can express each bond as

$$\vec{u}_i = l\hat{e}_i$$

where $\hat{e}_i$ is a random unit vector and l is the bond length. Then

$$< \vec{u}_i >= l < \hat{e}_i >= 0$$

$$< \vec{u}_i^2 >= l^2 < \hat{e}_i^2 >= l^2 < 1 >= l^2$$

Hence, we have

$$< \vec{u}_i \cdot \vec{u}_j >= l^2 \delta_{ij}$$

The distance of the from start to end is

$$\vec{r}_N = \sum_{i=1}^{N} \vec{u}_i$$

Hence the average value is

$$< \vec{r}_N >= \left\langle \sum_{i=1}^{N} \vec{u}_i \right\rangle$$

Since the bonds are independently chosen from the distribution, we can say that

$$< \vec{r}_N >= \left\langle \sum_{i=1}^{N} \vec{u}_i \right\rangle = \sum_{i=1}^{N} < \vec{u}_i >= \vec{0}$$

The average square distance is

$$< (\vec{r}_N)^2 >= \left\langle \left(\sum_{i=1}^{N} \vec{u}_i \right) \cdot \left(\sum_{j=1}^{N} \vec{u}_j \right) \right\rangle = \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \vec{u}_i \cdot \vec{u}_j \right\rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} < \vec{u}_i \cdot \vec{u}_j >=$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} l^2\delta_{ij} = \sum_{i=1}^{N} l^2 = l^2 N$$

Hence, the root mean square distance between the start and the end of the polymer is $l\sqrt{N}$. The typical $\sqrt{N}$ dependence is something we will observe again and again in the following sections.
Radius of gyration is

$$R_G^2 = \left\langle \frac{1}{N}\sum_{i=1}^{N}(\vec{r}_i - \vec{R}_{CM})^2 \right\rangle = \left\langle \frac{1}{N}\sum_{i=1}^{N}\left((\vec{r}_i)^2 - 2\vec{r}_i\cdot\vec{R}_{CM} + (\vec{R}_{CM})^2\right) \right\rangle =$$

$$= \left\langle \frac{1}{N}\sum_{i=1}^{N}(\vec{r}_i)^2 - 2\vec{R}_{CM}\cdot\frac{1}{N}\sum_{i=1}^{N}\vec{r}_i + (\vec{R}_{CM})^2 \right\rangle = \left\langle \frac{1}{N}\sum_{i=1}^{N}(\vec{r}_i)^2 - 2\vec{R}_{CM}\cdot\vec{R}_{CM} + (\vec{R}_{CM})^2 \right\rangle =$$

$$= \left\langle \frac{1}{N}\sum_{i=1}^{N}(\vec{r}_i)^2 - (\vec{R}_{CM})^2 \right\rangle = \frac{1}{2}\left\langle \frac{2}{N}\sum_{i=1}^{N}(\vec{r}_i)^2 - \frac{2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\vec{r}_i\cdot\vec{r}_j \right\rangle$$

since $\sum_{j=1}^{N}(\vec{r}_i)^2 = N(\vec{r}_i^2)$, we have

$$R_G^2 = \frac{1}{2}\left\langle \frac{2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{2}(\vec{r}_i)^2 - \frac{2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\vec{r}_i\cdot\vec{r}_j \right\rangle =$$

$$= \frac{1}{2N^2}\left\langle \sum_{i=1}^{N}\sum_{j=1}^{N}(\vec{r}_i)^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}(\vec{r}_j)^2 - \sum_{i=1}^{N}\sum_{j=1}^{N}2(\vec{r}_i\cdot\vec{r}_j) \right\rangle =$$

$$= \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\langle(\vec{r}_i - \vec{r}_j)^2\right\rangle$$

Here, we can use that

$$\vec{r}_i - \vec{r}_j = \sum_{m=1}^{i}\vec{u}_m - \sum_{n=1}^{j}\vec{u}_n = \sum_{m=j+1}^{i}\vec{u}_m$$

and so

$$\left\langle(\vec{r}_i - \vec{r}_j)^2\right\rangle = \sum_{m=j+1}^{i}\sum_{n=j+1}^{i}\langle\vec{u}_m\cdot\vec{u}_n\rangle = l^2\sum_{m=j+1}^{i}\sum_{n=j+1}^{i}\delta_{mn} = l^2(i-j-1+1) = l^2(i-j)$$

And therefore

$$R_G^2 = \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\langle(\vec{r}_i - \vec{r}_j)^2\right\rangle = \frac{l^2}{2N^2}\left\langle\sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)\right\rangle$$

We can move the one half into the expectation value by doing the summation only over half of the terms, which can be achieved by turning the limit of the second sum to i, i.e.

$$R_G^2 = \frac{l^2}{N^2}\left\langle\sum_{i=1}^{N}\sum_{j=1}^{i}(i-j)\right\rangle = \frac{l^2}{N^2}\left\langle\sum_{i=1}^{N}i^2 - \sum_{i=1}^{N}\sum_{j=1}^{i}j\right\rangle$$

we can use that $\sum_{j=1}^{N} j = \frac{N(N+1)}{2}$ to find

$$R_G^2 = \frac{l^2}{N^2}\left\langle\sum_{i=1}^{N}i^2 - \sum_{i=1}^{N}\frac{i(i+1)}{2}\right\rangle = \frac{l^2}{N^2}\left\langle\frac{1}{2}\sum_{i=1}^{N}i^2 - \frac{1}{2}\sum_{i=1}^{N}i\right\rangle$$

Using $\sum_{i=1}^{N} i^2 = \frac{N(N+1)(2N+1)}{6}$, we have

$$R_G^2 = \frac{l^2}{2N^2}\left(\frac{N(N+1)(2N+1)}{6} - \frac{N(N+1)}{2}\right) = \frac{l^2(N+1)}{12N}(2N+1-3) =$$

$$= \frac{l^2(N+1)(N-1)}{6N} = \frac{l^2N}{6} - \frac{l^2}{6N}$$

and for big N, which is common for polymers,

$$R_G = \frac{l\sqrt{N}}{\sqrt{6}} \tag{21}$$

4.2 Gaussian Chain

Suppose we have a chain of completly randmoly oriented bonds with Gaussian distribution in length, which is

$$p_b(\vec{u}_m) = \frac{e^{-\frac{1}{2}\frac{|\vec{u}_m|^2}{l^2/3}}}{\sqrt{\left(2\pi\frac{l^2}{3}\right)^3}}$$

where l^2 is the variance of bond length, and the factor of $\frac{1}{3}$ ensures that the variance of $<(\vec{u}_m)^2>=<u_{mx}^2>+<u_{my}^2>+<u_{mz}^2>$ evaluates to l^3.
We can express the probability that the mth step taken us $\vec{u}_n$ as

$$p(\vec{u}_n) = \int d^3u_m p_b(\vec{u}_m)\delta(\vec{u}_m - \vec{u}_n)$$

and if we are interested in N consecutive steps that will lead us to displacement vector $\vec{r}_N$, we can express this as

$$p(\vec{r}_N) = \int d^3u_1 p_b(\vec{u}_1)\int d^3u_2 p_b(\vec{u}_2)...\int d^3u_N p_b(\vec{u}_N)\delta\left(\vec{r}_N - \sum_{i=1}^{N}\vec{u}_i\right)$$

We can notice that this is a convolution. A convolution of functions f and g is defined as

$$(f\circ g)(y) = \int dx f(x)g(y-x)$$

Hence, convolution of three functions is

$$(h\circ(f\circ g))(y) = \int dz h(z)(f\circ g)(y-z) = \int dz h(z)\int dx f(x)g(y-z-x) =$$
$$= \int dz h(z)\int dx f(x)g(y-x-z) = (h\circ f\circ g)(y)$$

where we can drop the brackets as we can see that the convolution is commutative. Hence, general convolution of functions $f_1(x_1), ..., f_N(x_N)$ with function g is

$$(f_1\circ f_2\circ ...\circ f_N\circ g)(y) = \int dx_1 f(x_1)\int dx_2 f(x_2)...\int dx_N f(x_N)g\left(y - \sum_{i=1}^{N}x_N\right)$$

Therefore

$$p(\vec{r}_N) = (p_b\circ p_b\circ ...\circ p_N\circ\delta)(\vec{r}_N)$$

We can remember that a particular property of the Fourier transform is that it turns transforms of convolutions into products of transforms of the functions. Therefore, the Fourier transform of $p(\vec{r}_N)$ is

$$\tilde{p}(\vec{q}_N) = \left(\tilde{p}_b(\vec{q}_N)\right)^N\tilde{\delta}(\vec{q}_N)$$

where $\vec{q}$ is the conjugate variable of. The Fourier transform of δ function is

$$\tilde{\delta}(\vec{r}_n) = \int d^3r_N e^{-i\vec{r}_n\cdot\vec{q}_n}\delta(\vec{r}_N) = 1$$

Therefore, we have

$$\tilde{p}(\vec{q}_N) = \left(\tilde{p}_b(\vec{q}_N)\right)^N$$

The Fourier transform of a Gaussian distribution is a Gausiann distribution with the inverse variance and without the normalization constant, i.e.

$$\tilde{p}_b = e^{-\frac{1}{2}\frac{|\vec{q}_N|^2}{3}l^2}$$

Hence

$$\tilde{p}(\vec{q}_N) = e^{-\frac{1}{2}\frac{|q_N|^2}{3}l^2N}$$

Therefore, the inverse Fourier transform will give us the Gaussian again, this time as

$$p(\vec{r}_N) = \frac{e^{-\frac{1}{2}\frac{|r_N|^2}{Nl^2/3}}}{\sqrt{\left(2\pi\frac{Nl^2}{3}\right)^3}}$$

which leads to $<\vec{r}_N>=0$ and $<\vec{r}_N^2>=Nl^2$ - same as before.

4.3 Lattice Chain

Suppose we have a 3D lattice on which the monomers can exist. Lets start building our polymer from monomers, starting at position $\vec{x} = 0$. In each building step, we can move along the lattice in either direction with eqaul probability. We are searching for probability distribution $p(\vec{x}, N)$ which gives the probability that a polymer of length N ends at $\vec{x}$. To find it, we will consider how p changes as we change N, i.e. lets try to find $p(\vec{x}, N+1) - p(\vec{x}, N)$. The probability that a polymer of length $N+1$ ends at $\vec{x}$ is given by the difference of (probability of all paths leading to neighbouring lattice points in N steps that will lead to $\vec{x}$ in $N+1$st step) and the probability of paths leading to the lattice point $\vec{x}$ in N steps. This is because if the path ends at $\vec{x}$ in N steps, the polymer in the next step must move away, and therefore cannot end at $\vec{x}$. For a simple cubic lattice, each lattice point has 6 nearest neighbours. Therefore the change that the next building step from a lattice point neighbouring $\vec{x}$ will result into going to $\vec{x}$ is $\frac{1}{6}$. The probability of polymer ending at neighbouring lattice point at Nth step is simply given by the probability distribution. For a cubic lattice with length l, the first probability in the difference is therefore

$$\frac{1}{6} \times \Big(p(\vec{x} - l\hat{i}, N) + p(\vec{x} + l\hat{i}, N) + p(\vec{x} - l\hat{j}, N) + p(\vec{x} + l\hat{j}, N) + p(\vec{x} - l\hat{k}, N) + p(\vec{x} + l\hat{k}, N) \Big)$$

The second probability is simply $p(\vec{x}, N)$. Therefore, the total probability change in the event when polymer ends at $\vec{x}$ in $N+1$st step is

$$p(\vec{x}, N+1) - p(\vec{x}, N) =$$

$$= \frac{1}{6} \times \Big(p(\vec{x} - l\hat{i}, N) + p(\vec{x} + l\hat{i}, N) + p(\vec{x} - l\hat{j}, N) + p(\vec{x} + l\hat{j}, N) + p(\vec{x} - l\hat{k}, N) + p(\vec{x} + l\hat{k}, N) \Big) - p(\vec{x}, N) =$$

$$= \frac{1}{6} \Big(p(\vec{x} + l\hat{i}, N) - 2p(\vec{x}, N) + p(\vec{x} - l\hat{i}, N) + \ldots \Big)$$

We can recognize these expressions as numerical schemes for second derivatives. Therefore, in the limit of small l, we can write

$$p(\vec{x}, N+1) - p(\vec{x}, N) = \frac{1}{6} \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right) l^2 = \frac{l^2}{6} \nabla^2 p$$

Furthermore, we can rewrite that for big N

$$\frac{\partial p}{\partial N} = \frac{l^2}{6} \nabla^2 p$$

This is a diffusion equation, which can be solved given the boundary and initial conditions. Typically, we could use $p(\vec{x}, 0) = \delta(\vec{x})$ as the initial condition, because a polymer with zero bonds is just a monomer sitting at the origin. Furthermore, the boundary conditions should be that $p(\infty, N) = 0$, as the polymer has only a finite contour length Nl. This leads to Green's function that is Gaussian, and given that our initial condition is a point source, we again have a Gaussian probability distribution

$$p(\vec{x}, N) = \frac{e^{-\frac{1}{2} \frac{|\vec{x}|^2}{Nl^2/3}}}{\sqrt{(2\pi \frac{Nl^2}{3})^3}}$$

The reason why we always encounter the normal distribution is the central limit theorem - any distribution that is sampled in big numbers eventually tends towards the normal (Gaussian) distribution.

4.4 Force on Polymers

The polymer is in the canonical ensemble and we assume that it has no internal energy U. Therefore, the Helmholtz free energy is

$$F = U - TS = -TS$$

Using Boltzmann definition of entropy

$$S = k_B \ln \Omega$$

for a given state when the polymer has N bonds and ends at $\vec{r}_N$, we have

$$F = -TS(\vec{r}_N, N) = -k_B T \ln \Omega(\vec{r}_N, N) = -k_B T \ln(p(\vec{r}_N, N)\Omega_0)$$

where Ω_0 is the total number of available microstates to the polymer. Therefore, we can write

$$F = -k_BT\ln p(\vec{r}_N, N) - k_BT\ln\Omega_0$$

If we exert a force on the polymer, we are changing its free energy by doing work. The force $\vec{T}$ is given in terms of change of energy as

$$\vec{T} = -\nabla F = -\frac{\partial F}{\partial|\vec{r}_N|}\hat{r}$$

where I used the fact that $p(\vec{r}_N, N)$ depends only on the size of $\vec{r}_N$ in our case. For Gaussian distribution

$$F = -k_BT\left(-\frac{3}{2}\frac{|\vec{r}_n|^2}{Nl^2}\right) + F_0$$

where $F_0 = -k_BT\ln\Omega_0 + k_BT\frac{3}{2}\ln\left(2\pi\frac{Nl^2}{3}\right)$ is the term constant with $|\vec{r}_n|$. Therefore

$$\vec{T} = -\frac{3k_BT}{Nl^2}|\vec{r}_n|\hat{r} = -\frac{3k_BT}{Nl^2}\vec{r}_n$$

therefore the polymer behaves like a spring.
This description have some good qualitative predictions - for example, as we heat up the polymer, the spring constant rises, so the polymer that has a force acting on it tends to reduce in size when we heat it. However, it does allow for unrestricted displacement $\vec{r}_N$. We know that the maximum value of the displacement should be the countour length Nl.
This can be in fact build into the theory via the statistical mechanics approach.

4.4.1 Contour Length Enforcement

We can define the energy stored in the polymer from the expression derived above approximately as

$$E = -\vec{T}\cdot\vec{r}_N = tl\sum_{i=1}^{N}\hat{t}\cdot\hat{u}_i = tl\sum_{i=1}^{N}\cos(\theta_i)$$

where t is the magnitude of the force and θ_i is the angle between the ith bond of the polymer and the vector of the force. The partition function is then

$$Z = \sum_{\gamma(\vec{\theta})} e^{-\beta tl(\sum_{i=1}^{N}\cos(\theta_i))}$$

where $\vec{\theta}$ is a vector of all θ_i for a given microstate. Since we allow for all orientations of each bond, we have a continuum of microstates, and therefore, we have to do integration. We can choose our scale to be 1 in this case, as we will be integrating the full solid angles, and therefore, we have

$$Z = \iint_{4\pi}\sin\theta_1 d\theta_1 d\phi_1\iint_{4\pi}\sin\theta_2 d\theta_2 d\phi\iint_{4\pi}...\iint_{4\pi}\sin\theta_N d\theta_N d\phi_N e^{-\beta tl(\sum_{i=1}^{N}\cos(\theta_i))}$$

where each integration runs over the full solid angle. We can recognize that

$$\sin\theta_i d\theta_i d\phi_i = -d(\cos\theta_i)d\phi$$

and that

$$e^{-\beta tl(\sum_{i=1}^{N}\cos(\theta_i))} = \prod_{i=1}^{N}e^{-\beta tl\cos(\theta_i)}$$

and therefore we can factorize the partition function as

$$Z = \prod_{i=1}^{N}\iint_{4\pi}e^{-\beta tl\cos(\theta_i)}(-1)d(\cos\theta_i)d\phi = (2\pi)^N(-1)^N\prod_{i=1}^{N}\int_1^{-1}e^{-\beta tlz}dz =$$

$$= (2\pi)^N(-1)^N\prod_{i=1}^{N}\left[\left(e^{\beta tl} - e^{-\beta tl}\right)\frac{1}{\beta tl}\right] = (-1)^N\left(\frac{4\pi\sinh(\beta tl)}{\beta tl}\right)^N = (-4\pi)^N\,\mathrm{sinhc}^N(\beta tl)$$

Therefore, the free energy is

$$F = -k_BT\ln Z = -k_BT\ln(\sinh^N(\beta tl)) + k_BT\ln((\beta tl)^N) + F_0 = -Nk_BT\ln(\sinh(\beta tl)) + Nk_BT\ln(\beta tl) + F_0$$

where $F_0 = -k_BT(-4\pi)^N$ is a part independent of t. From our original expression for F, we can expect that the expectation value of the extension $< |\vec{r}_N| >$ will be equal to

$$< |\vec{r}_N| > = -\frac{\partial F}{\partial t} = Nl\coth(\beta tl) - \frac{N}{\beta t} = Nl\left[\coth(\beta tl) - \frac{1}{\beta tl}\right]$$

So

$$< |\vec{r}_N| > = Nl\left(\coth(\beta tl) - \frac{1}{\beta tl}\right) \tag{22}$$

This means that as $t \to \infty$, $< |\vec{r}_N| > \to 1$, while around zero

$$\coth(\beta tl) \approx \frac{1}{\beta tl} + \frac{1}{3}\beta tl$$

and so

$$< |\vec{r}_N| > \approx \frac{1}{3}N\beta tl^2$$

$$t = \frac{3k_BT}{Nl^2} < |\vec{r}_N| >$$

which is what we derived earlier. We therefore have a self-consistent theory of the polymer under a force that enforces the contour length as the maximum displacement of the polymer.

4.5 Polymer Close to a Wall

If a polymer is close to a wall that it cannot enter, we need to manipulate p in order to reflect these boundary conditions. In the simple case of 1D, we still have the polymer probability density governed by the diffusion equation, only this time it has form

$$\frac{\partial p}{\partial n} = \frac{l^2}{2}\frac{\partial^2 p}{\partial x^2}$$

If the boundary is a plane at $x = 0$, we can use method of mirror charges to enforce the boundary conditions. Lets say the polymer starts at $x = x_0 > 0$. If there was no boundary present, the probability density would be simply the Green's function in 1D for this problem, i.e.

$$p(x, N, x_0) = \frac{1}{(2\pi Nl^2)^{\frac{1}{2}}}e^{-\frac{(x-x_0)^2}{2Nl^2}}$$

where the factors of $1/3$ disappear as we are in 1D, and so does the exponent in the normalization factor change to $\frac{1}{2}$. We can enforce $p(0, N) = 0$ by setting mirror distribution as

$$p(x, N) = p(x, N, x_0) - p(x, N, -x_0)$$

If we assume that x_0 is small, this can be approximated as

$$p(x, N) \approx 2x_0\frac{\partial p}{\partial x_0}\Big|_{x_0=0} = 2x_0\frac{1}{(2\pi Nl^2)^{\frac{1}{2}}}\frac{1}{Nl^2}xe^{-\frac{x^2}{2Nl^2}} = \sqrt{\frac{2}{\pi}}\frac{1}{(Nl^2)^{\frac{3}{2}}}xx_0e^{-\frac{x^2}{2Nl^2}}$$

This distribution is not normalized, but can be normalized by setting

$$1 = \int_0^\infty p(x, N)dx$$

I will not do the normalization here, but I will state that the first moment of the distribution is

$$< x > = \sqrt{\frac{\pi}{2}}l\sqrt{N}$$

hence the polymer tends to get away from the wall. This is why so called polymer brushes are created.

5 Brownian Motion

There are two perspectives for the Brownian motion - the macroscopic perspective, using the Fokker-Planck equation, and the microscopic perspective, using Langevin equation. We will briefly discuss both.

5.1 Fokker-Planck Equation

Assume there is a concentration of particles $c(\vec{x},t)$ free to diffuse through some medium. The concetration is normalised as

$$\int_V d^3c(\vec{x},t) = N(V,t)$$

where $N(V,t)$ is the total number of particles present at time t within volume V.
Suppose that in the bulk of the material, the particles cannot be destroyed. Therefore, the continuity equation applies for them

$$\frac{\partial c}{\partial t} = -\nabla \cdot \vec{j}$$

where $\vec{j}(\vec{x},t)$ is the flux density of particles.
During the diffusion, we experimentally know that the particles tend to diffuse away from the high concentration regions. We can try to incorporate this behaviour by making the flux density proportional to the opposite of the gradient of c, i.e.

$$\vec{j} = -D\nabla c \tag{23}$$

This is called the Fick's first law. Applying it to the continuity equation leads to

$$\frac{\partial c}{\partial t} = D\nabla \cdot \nabla c = D\nabla^2 c \tag{24}$$

This is again the diffusion equation. However, besides the contribution to $\vec{j}$ due to gradient of concentration, there migh also be external forces acting on the particles. Lets suppose that the collisions of the particles within the medium are so frequent that the force only manages to accelerate the particles towards some drift velocity

$$\vec{v} = \frac{\vec{F}}{\gamma}$$

where γ is the drag coefficient. We should note that in typical instances of Brownian motion, this steady state is reached within microseconds from the force application. Hence, we will have an extra flux of particles due to the drift

$$\vec{j}_D = c\vec{v} = c\frac{\vec{F}}{\gamma}$$

For a conservative force, we can express it as a gradient of a scalar potential, i.e.

$$\vec{j}_D = -c\frac{\nabla V}{\gamma}$$

where $\vec{F} = -\nabla V$ and V is the potential. Hence, the total flux density is

$$\vec{j} = -D\nabla c - c\frac{\nabla V}{\gamma}$$

which leads to modified diffusion equation

$$\frac{\partial c}{\partial t} = D\nabla \cdot (\nabla c + \frac{c}{\gamma}\nabla V) \tag{25}$$

This is the Fokker-Planck equation. It is an equation quite difficult to solve, but we will use it mainly to determine the diffusion coefficient D.
We expect that at equilibrium, the concentration profile of the particles will copy the Boltzmann distribution in space due to the potential V. Therefore, in equilibrium, we expect

$$c = c_0 e^{-\beta V(\vec{r})}$$

where c_0 normalizes the distribution. In equilibrium, the concentration does not change anymore and there are no currents. This means that

$$\vec{j} = -D\nabla c - \frac{c}{\gamma}\nabla V = 0$$

$$D\nabla C = -\frac{c}{\gamma}\nabla V$$

But, given our definition of equilibrated c

$$\nabla c = \nabla(c_0 e^{-\beta V}) = -\beta c_0 e^{-\beta V}\nabla V = -\beta c \nabla V$$

and so

$$D\nabla c = -D\beta c \nabla V = -\frac{c}{\gamma}\nabla V$$

$$D = \frac{k_B T}{\gamma} \tag{26}$$

We have therefore determined the diffusion coefficient by making sure it is consistent with our expectations from thermodynamics.

5.1.1 Diffusion Motion

The diffusion motion can be determined by solving the diffusion equation in the absence of potential V. In 1D, we can look at the polymer solutions of diffusion equation to determine that the mean square distance diffused by the particles will be

$$\sqrt{< x^2 > - < x >^2} = \sqrt{2Dt}$$

Therefore, to diffuse to average distance x, the particles need time

$$t = \frac{x^2}{2D}$$

while if the particle would move balistically, they need time

$$t_B = \frac{x}{v}$$

This means that there is a critical time

$$t_C = \frac{x^2}{2D} = \frac{\gamma x^2}{2k_B T}$$

below which the diffusive transport is more effective than the ballistic transport.

5.2 Langevin Equation

In the microscopic description, each particle propagates balistically but is under the effect of the random force due to collisions with other particles and the force due to drag. We can write this down in terms of stochastic (random) variables as the Newton's second law

$$\vec{f} - \gamma\dot{\vec{x}} + \vec{\zeta} = m\ddot{\vec{x}} \tag{27}$$

where $\vec{f}$ is the external (deterministic) force, γ is the drag coefficient and $\vec{\zeta}$ is the stochastic force, which changes with time. This is called the Langevin equation. The question is how to formulate the collision force $\vec{\zeta}$ in terms of probability distributions, so that we can derive probability distributions for $\vec{x}(t)$.

5.2.1 Form of Random Force

Classical choice for $\zeta(t)$ is the Gaussian white noise. This is defined as uncorrelated sampling from the Gaussian distribution for each of the components of $\vec{\zeta}$. The reason behind this is again central limit theorem - if the collisions are frequent, independent of the distribution of each collision, the overall distribution from the collisions will tend towards Gaussian distribution.

We define the properties of ζ as follows

$$\vec{\zeta}(t) = (\zeta_x, \zeta_y, \zeta_z)$$

where each of the numbers are drawn from the Gaussian distribution. Furthermore, we require that

$$<\vec{\zeta}>=0$$

i.e. the Gaussian distributions for the components are all centered on zero and

$$<\vec{\zeta}^2>=A$$

where A is some magnitude of the random force, to be determined by external models. We assume that ζ is uncorrelated in time and in components, i.e.

$$<(\vec{\zeta})_i(t_1)(\vec{\zeta}(t_2))_j>=A\delta_{ij}\delta(t_1-t_2)$$

Also, we assume that the noise produced by ζ is white in the sense that all frequencies are present in variation of ζ - the Fourier transform is a flat line.
The easiest way to formulate some specific example is then to take the case of zero external force. In this case, the Langevin equation becomes

$$m\ddot{\vec{x}}=-\gamma\dot{\vec{x}}+\vec{\zeta}$$

The expectation values of these quantities are then (as ζ is uncorellated variable)

$$m\frac{d^2<\vec{x}>}{dt^2}=-\gamma\frac{d<\vec{x}>}{dt}+<\vec{\zeta}>=-\gamma\frac{d<\vec{x}>}{dt}$$

$$\frac{d^2<\vec{x}>}{dt^2}=-\frac{\gamma}{m}\frac{d<\vec{x}>}{dt}$$

which has a solution

$$\frac{d<\vec{x}>}{dt}\propto e^{-\frac{\gamma}{m}t}$$

Therefore, we predict that the expectation of the displacement of the particles decays over time - the longer we observe the particle, the more probable it becomes that it returns to its starting spot again.

5.2.2 Massless Brownian Particle

Consider now a case when the forces are strong enough that the inertia of the particle does not play a significant role. Then, we have the massless Langevin equation

$$\vec{f}+\vec{\zeta}=\gamma\dot{\vec{x}} \tag{28}$$

In the absence of the external force, this can be integrated as

$$\vec{x}(t)=\vec{x}(t_0)+\int_{t_0}^{t}\frac{\vec{\zeta}}{\gamma}dt$$

Hence, the mean square displacement of the massless particle in time $t-t_0$ will be

$$<|\vec{x}(t)-\vec{x}(t_0)|^2>=\left\langle\int_{t_0}^{t}dt\frac{\vec{\zeta}(t)}{\gamma}\cdot\int_{t_0}^{t}dt'\frac{\vec{\zeta}(t')}{\gamma}\right\rangle=\int_{t_0}^{t}\int_{t_0}^{t}\frac{<\vec{\zeta}(t)\cdot\vec{\zeta}(t')>}{\gamma^2}dtdt'$$

Using the correlation property of ζ as we defined it earlier

$$<\vec{\zeta}(t)\cdot\vec{\zeta}(t')>=\sum_{i=1}^{3}<\zeta_i(t)\zeta_i(t')=3A\delta(t-t')$$

And therefore

$$<|\vec{x}(t)-\vec{x}(t_0)|^2>=\int_{t_0}^{t}\int_{t_0}^{t}\frac{3A\delta(t-t')}{\gamma^2}dtdt'=\int_{t_0}^{t}\frac{3A}{\gamma^2}dt=\frac{3A(t-t_0)}{\gamma^2}$$

Compairing this with the result for diffusion equation (24) in 1D, we have

$$6D(t-t_0)=3A\frac{t-t_0}{\gamma^2}$$

where we have extra factor of 3 because we are working in three dimensions

$$A=2D\gamma^2=2\frac{k_BT}{\gamma}\gamma^2$$

and so we have determined consistency criterion for the Langevin model as

$$A=2k_BT\gamma \tag{29}$$