Average total magnetization $\vec{M}$ is oriented $\parallel$ to $\vec{h}$. If we choose $\vec{h} = h\hat{z}$ along $\hat{z}$, then

$$M_z = N\langle \mu\cos\theta \rangle = N \frac{\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta\, e^{\beta\mu h\cos\theta}\, \mu\cos\theta}{\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta\, e^{\beta\mu h\cos\theta}}$$

$$= \frac{N}{\beta} \frac{\frac{\partial}{\partial h}\left[2\pi \int_0^{\pi} d\theta \sin\theta\, e^{\beta\mu h\cos\theta}\right]}{2\pi \int_0^{\pi} d\theta \sin\theta\, e^{\beta\mu h\cos\theta}}$$

$$= \frac{N}{\beta} \frac{\frac{\partial}{\partial h}(Q_1)}{Q_1} = \frac{N}{\beta} \frac{\partial}{\partial h}(\ln Q_1)$$

$$= \frac{\partial}{\partial h}\left[k_B T \ln Q_1^N\right] = -\frac{\partial}{\partial h} A(T, h)$$

when we apply $\vec{h}$, the magnetic field $\vec{h}$ is a new thermodynamic variable. The above

$$M_z = -\frac{\partial A(T,h)}{\partial h}$$

says that total magnetization is the thermodynamic conjugate variable to magnetic field.

Using our result for Q_1, we get

$$\frac{M_z}{N} = \frac{1}{\beta}\frac{\partial}{\partial h}(\ln Q_1) = \frac{1}{\beta}\frac{1}{Q_1}\frac{\partial Q_1}{\partial h}$$

$$= \frac{4\pi}{\beta}\left[\frac{\cosh(\beta\mu h)}{h} - \frac{\sinh(\beta\mu h)}{\beta\mu h^2}\right]$$

$$\frac{4\pi \sinh(\beta\mu h)}{\beta\mu h}$$

$$= \mu h\left[\frac{\coth}{h}(\beta\mu h) - \frac{1}{\beta\mu h^2}\right]$$

$$\boxed{\frac{M_z}{N} = \mu\left[\coth(\beta\mu h) - \frac{1}{\beta\mu h}\right]}$$

↑ hyperbolic cotan

$\frac{M_z}{N} = \langle \mu_z \rangle$ average spin along $\hat{z}$

$L(x) = \coth x - \frac{1}{x}$ Langevin function

for large x, $L(x) \to 1$

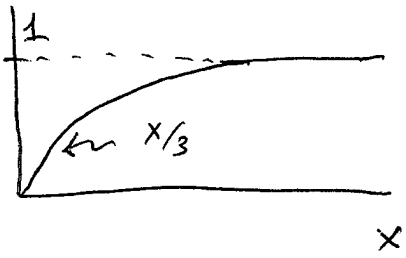
for small x, $L(x) = \frac{\cosh x}{\sinh x} - \frac{1}{x}$

$$\simeq \frac{1+\frac{x^2}{2}}{x+\frac{x^3}{6}} - \frac{1}{x} = \frac{1+\frac{x^2}{2}}{x\left(1+\frac{x^2}{6}\right)} - \frac{1}{x}$$

$$\simeq \frac{\left(1+\frac{x^2}{2}\right)\left(1-\frac{x^2}{6}\right)}{x} - \frac{1}{x} \simeq \frac{1+\frac{x^2}{2}-\frac{x^2}{6}}{x} - \frac{1}{x}$$

$$\simeq \frac{x}{3}$$

So $L(x)$



$x = \beta\mu h$

$\Rightarrow$ at small h or at large T (small β)

$$\langle \mu_z \rangle = \frac{\mu^2 \beta h}{3} = \frac{\mu^2 h}{3 k_B T}$$

$$M_z = \frac{N\mu^2 h}{3 k_B T}$$

magnetic susceptibility $\chi = \lim_{h \to 0} \frac{\partial M_z}{\partial h} = \frac{N\mu^2}{3k_B T} \propto \frac{1}{T}$

Curie law of paramagnetism

$$\chi \propto \frac{1}{T}$$

A note about coordinates

We can write the partition function as

Microcanonical $\Omega(E, V, N) = \int_{\text{all degrees of freedom}} \delta(E - \mathcal{H})$

Canonical $Q_N(T, V) = \int_{\text{all degrees of freedom}} e^{-\beta \mathcal{H}}$

where by $\int_{\text{all degrees of freedom}}$ I mean a sum or integral over all the degrees of freedom that characterize the system. If degrees of freedom are continuous, then the integration should include an appropriate factor (like $\frac{1}{h^{3N}}$ for the gas of N classical particles) so that the partition function is dimensionless.

For classical systems with continuous degrees of freedom, it is ESSENTIAL that the degrees of freedom one integrates over be a set of Hamiltonian canonically conjugate coordinate-momentum pairs (q_i, p_i).

The reason for this has to do with Liouville's Theorem. To describe equilibrium we needed that the probability density for the system to be at a particular point in phase space should NOT change in time, i.e. $\frac{\partial \rho}{\partial t} = 0$. Liouville's theorem told us that this will be the case whenever all states of a given total energy E are equally likely. But Liouville's theorem only applies if we are labeling our states by a set of Hamiltonian canonically conjugate coordinate - momentum pairs. - recall we had to use Hamilton's equations of motion in deriving Liouville's Theorem.

If (q_i, p_i) are such a set of canonically conjugate degrees of freedom, then we can write the canonical partition function, for example, as

$$Q_N(T,V) = \int dq_1, \ldots, dq_N \int dp_1 \cdots dp_N \; e^{-\beta H(q_1, \ldots q_N, p_1, \ldots p_N)}$$

Now sometimes it might be convenient to label states by some other set of coordinates, for example $(q_i, \dot{q}_i)$ instead of (q_i, p_i). In that case one can compute the partition function in terms of the convenient coordinates, provided one makes the correct transformation of the variables of integration.

$$Q_N(T,V) = \int d q_i \, dp_i \; e^{-\beta H(q_i, p_i)}$$

$$= \int dq_i \, d\dot{q}_i \; J \, e^{-\beta H(q_i, \dot{q}_i)}$$

Where J is the Jacobian of the transformation from (q_i, p_i) to $(q_i, \dot{q}_i)$. It is crucial to include this Jacobian factor to get the correct result for the partition function.

Example A classical gas of N particles in an external magnetic field.

$$H = \sum_{i=1}^{N} \frac{1}{2m} \left| \vec{p}_i - \frac{e}{c} \vec{A}(\vec{r}_i) \right|^2$$

$\vec{A}(\vec{r})$ is the magnetic vector potential.
The particles velocity, in terms of the canonical momentum $\vec{p}$, is

$$\vec{v}_i = \frac{\vec{p}_i - \frac{e}{c}\vec{A}(\vec{r}_i)}{m}$$

At first glance ~~it seems~~ it seems that we could not compute the average total energy $E = \langle H \rangle$ using the equipartition theorem, since H couples $\vec{p}_i$ to $\vec{r}_i$ via $\vec{A}(\vec{r}_i)$.

However, if we directly compute Q_1

$$Q_1 = \int d^3r \int \frac{d^3p}{h^3} \, e^{-\beta \frac{|\vec{p} - \frac{e}{c}A(\vec{r})|^2}{2m}}$$

we can do the following trick. Since $\vec{r}$ is fixed when we do the $\vec{p}$ integration, then as far as the $\vec{p}$ integration is concerned $\frac{e}{c}\vec{A}(\vec{r})$ is just some constant $\vec{p}_0$.

We can then make a transformation of variables in the integration $\vec{p}' = \vec{p} - \frac{e}{c}\vec{A}(\vec{r})$ to write

$$Q_1 = \int \frac{d^3r \int d^3p'}{h^3} e^{-\beta |\vec{p}'|^2/2m}$$

which gives exactly the same Q_1 as for a gas that is NOT in a magnetic field!

Hence we conclude that the total energy of the gas in the magnetic field is the same as in zero field, and so $E = \frac{3}{2} N k_B T$.

We can view this result in terms of coordinates. Instead of writing the partition function above in terms of the canonical coordinates $\vec{r}_i, \vec{p}_i$, we could use instead the coordinates $\vec{r}_i, \vec{v}_i$. We then have $\vec{v} = \frac{\vec{p} - \frac{e}{c}\vec{A}(\vec{r})}{m}$

$$Q_1 = \int \frac{d^3r \int d^3v}{h^2} J e^{-\beta \frac{mv^2}{2}}$$

where $J = m^3$ is the Jacobian of the transformation from $\vec{r}, \vec{p}$ to $\vec{r}, \vec{v}$.

We can now do the integration over $\vec{v}$ and get the same result for Q_1 as for the case with no magnetic field. The crucial

thing that made things work simply in this case was that the Jacobian J was a constant and not some function of $\bar{r}, \bar{v}$.

Entropy & Information

In canonical ensemble we had

prob to be in energy E $\quad P(E) = \dfrac{\Omega(E)\, e^{-E/k_B T}}{Q_N}$

or if we label microstates by an index i then the prob to be in state i is

$$P_i = \frac{e^{-E_i/k_B T}}{Q_N} \quad \text{where} \quad Q_N = \sum_i e^{-E_i/k_B T}$$

Consider the average value of $\ln P_i$

$$\langle \ln P_i \rangle = \sum_i P_i \ln P_i \quad \text{by definition of average}$$

But also $\langle \ln P_i \rangle = \left\langle \ln \left[\dfrac{e^{-E_i/k_B T}}{Q_N} \right] \right\rangle$

$$= -\frac{\langle E \rangle}{k_B T} - \ln Q_N$$

$$\Rightarrow k_B T \langle \ln P_i \rangle = -\langle E \rangle + A = -T\langle S \rangle$$

(Entropy as computed in canonical ensemble)

as $A = E - TS$

$$\Rightarrow \boxed{\langle S \rangle = -k_B \sum_i P_i \ln P_i}$$

where P_i is the prob to be in state i

Note: above was derived for canonical ensemble. But it also holds for the microcanonical ensemble. In microcanonical, the prob to be in state i is $1/\Omega(E)$ for a state with $E_i = E$, and zero otherwise. equally likely to be in any state with energy E

$$\Rightarrow \quad -k_B \sum_i p_i \ln p_i = -k_B \sum_i \left(\frac{1}{\Omega}\right) \ln\left(\frac{1}{\Omega}\right)$$

↑ sum over only states in energy shell about E.

But the terms in the sum are all equal, and the number of terms is just the number of states at energy E, i.e. Ω.

$$\Rightarrow -k_B \sum_i \left(\frac{1}{\Omega}\right) \ln\left(\frac{1}{\Omega}\right) = -k_B \left(\frac{1}{\Omega}\right) \ln\left(\frac{1}{\Omega}\right) \sum_i$$

$$= -k_B \left(\frac{1}{\Omega}\right) \ln\left(\frac{1}{\Omega}\right) (\Omega) = -k_B \ln\left(\frac{1}{\Omega}\right)$$

$$= k_B \ln \Omega$$

So $\quad -k_B \sum_i p_i \ln p_i = k_B \ln \Omega = S(E)$ entropy in microcanonical ensemble

So $\quad \boxed{S = -k_B \sum_i p_i \ln p_i}$ works for both microcanonical and canonical ensembles!

Shannon (1948) turned this relation backwards, in developing a close relation between entropy and information theory.

Consider a system with states labeled by i, and P_i is the probability for the system to be in state i.

We want to define a measure of how disordered the distribution P_i is. Call this disorder measure S (it will turn out to be the entropy). The bigger (smaller) S is, the more (less) disordered the system is, the less (more) information we have about the probable state of the system.

We want S to satisfy the following properties

1) If $p_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$ then the state of the system is exactly known to be i. This should have $S=0$ as there is no uncertainty, no disorder

2) For equally likely p_i, ie all $p_i = 1/N$ for N states, the system is maximally disordered, ie S is max possible value for all possible N state distributions.

3) S should be <u>additive</u> over independently random systems.

To explain what we mean by (3),
suppose we have one system with N equally likely states labeled by $n = 1, \cdots, N$, and a second system with M equally likely states labeled by $m = 1, \cdots, M$.

The combined system ~~has~~ has $N \times M$ equally states labeled by the pairs (n, m). We want

$$S(N \times M) = S(N) + S(M)$$

The function with this property is the logarithm. We use the natural log, although any base would do.

$\Rightarrow$ For a system of N equally likely states,

$$S = k \ln N$$ where k is an arbitrary proportionality constant.

(Note: if we take $k = k_B$ then above is same as the definition of entropy in the microcanonical ensemble!)

Suppose that all states are not equally likely.
What is S in such a case?

Consider a system which has two possible states 1 and 2. The prob to be in 1 is p_1. The prob to be in 2 is $p_2 = 1 - p_1$. In general $p_1 \neq p_2$, ie the states need not be equally likely.