Why study statistical mechanics?

- it provides a microscopic theory of thermodynamics, and most importantly a physical understanding / basis for entropy.

- "real physics" in the sense that we will ecounter situations that we cannot solve exactly — need to make simplifications guided by physics in order to calculate

- the techniques we will develop have wide application to complex systems
  - many-body quantum systems eg. condensed matter
    astrophysics (white dwarfs and neutron stars)
  - "soft" condensed matter systems
    - biophysics (see colloquia this term!)
    - information theory
  - particle physics (renormalization group)

I: FUNDAMENTALS

[See Chandler 3.1, 3.2, Callen 15, 21-8]

The basic idea of STATISTICAL MECHANICS is to relate <u>micro</u>scopic and <u>macro</u>scopic descriptions of a physical system.

MICROSCOPIC	$\longleftrightarrow$	MACROSCOPIC
$\sim 10^{23}$ variables laws of mechanics		only need a few variables to describe equilibrium systems E, N, V, S
		laws of thermodynamics

In this way it is a microscopic theory of thermodynamics. Perhaps most importantly, it gives a physical understanding / basis for <u>entropy</u>.

In a macroscopic object, there are so many degrees of freedom that it is impossible to keep track of them all, although in principle given initial conditions one could do so. Instead, we adopt a statistical approach – we talk about the probability of a particle in a gas having particular values of $\underline{p}$ or $\underline{x}$ for example. For a body in equilibrium we will see that we can write the probability distribution as a function only of the macroscopic parameters or <u>constraints</u> eg. E, V, N.

For a given set of constraints eg. E, V, N which label the MACROSTATE, there are <u>many</u> different possible MICROSTATES.

eg. a classical gas of non-interacting atoms

$$E = \sum_{i=1}^{N} \frac{p_{x,i}^2 + p_{y,i}^2 + p_{z,i}^2}{2m}$$

For a given energy E, the allowed microstates lie on the surface of a sphere in $3N$-dimensional phase space.

At any moment we can label the microstate of the system as a point on this sphere. Collisions between particles cause the microstate to continuously change, but it always lies on the surface determined by the constraint of fixed total energy E.

The ergodic assumption

"If we wait long enough, the system will eventually flow through all the microstates consistent with the macrostate."

Consequences – 1) consider two systems with the same (E, V, N) but starting in different microstates ν and $\xi\nu'$. According to the ergodic assumption eventually the 2nd system will reach state ν. Its subsequent evolution is identical to the 1st system. $\Rightarrow$ Averaging over a long time must give the same result for both systems — explains why the equilibrium state only depends on the external constraints.

[What do we mean by "long time"? For example, consider a gas at room temperature. Typically, the time between collisions for a given molecule is $\sim 10^{-10}$ s. For $\sim 10^{23}$ particles, this is 10^{33} collisions per second. So on the timescale of a macroscopic measurement, the idea is that the system has plenty of time to explore all the available microstates.]

2) Ensemble averages. The second consequence is that instead of taking the time average of a single system, we can instead

take the average over an ENSEMBLE of systems, each of which is in a different microstate consistent with the macrostate.

eg. observable G

$$G_{obs} = \sum P_\nu G_\nu = \sum P_\nu \langle \nu | G | \nu \rangle$$

$\uparrow$ probability of finding the system in state ν

Principle of equal equilibrium probability

"When a thermally isolated system reaches equilibrium, the mutually-accessible microstates are equally likely."

ie. if $\Omega(E, N, V)$ = # microstates consistent with the constraints

then $P_\nu = \dfrac{1}{\Omega}$ for all ν in the ensemble

[Equal probabilities corresponds to the most random state – ie. equilibrium corresponds to the state of greatest ignorance about the microstate given the constraints.]

[Where does this come from? Quantum mechanics tells us that the transition rates between two states ν and ν' are equal $w_{\nu\nu'} = w_{\nu'\nu}$

(Recall Fermi's Golden Rule $(w_{\nu\nu'}, w_{\nu'\nu}) \propto |M_{\nu\nu'}|^2$)

But in equilibrium we must have $P_\nu w_{\nu\nu'} = P_{\nu'} w_{\nu'\nu}$

$\Rightarrow P_\nu = P_{\nu'}$.]

The Microcanonical Ensemble

is the ensemble of Ω microstates consistent with fixed total energy E [and N, V], with each having uniform probability $P_\nu = \frac{1}{\Omega}$.

A simple physical picture of an isolated system with a well-defined energy.

Let's look at an example, but first a reminder about

Counting

Number of ways of arranging N dissimilar objects is $N!$
If m of them are identical, then the number of ways $= \frac{N!}{m!}$

Permutations	choose m objects from N	${}^{N}P_{m} = \frac{N!}{(N-m)!}$
Combinations	the same, but we don't care about the order	${}^{N}C_{m} = \frac{N!}{(N-m)!\,m!}$

Example: distributing Q quanta among N harmonic oscillators

Each oscillator has $\varepsilon_n = (n+\frac{1}{2})h\nu$. The total $E = Qh\nu$.

What is the number of ways we can do this? To calculate this, visualize the problem as laying down N-1 match sticks and Q marbles (Callen)

eg. 000|0|00|||0|00|0 represents one way to do it. $(3,1,2,0,0,1,2,1)$ $Q=10$ $N=8$

$$\Rightarrow \Omega = \frac{(N-1+Q)!}{Q!\,(N-1)!}$$

the number of ways of putting down $(N-1+Q)$ objects, Q of which are identical and N-1 of which are identical.

Two important points about $\Omega(E)$:

1) The number of states increases exponentially with energy

To see this, use Stirling's approximation for large N

$$\log N! = N(\log N - 1)$$

[an important result to memorize!]

$$\text{then } \log\Omega \approx \log\left(\frac{(N+Q)!}{N!\,Q!}\right)$$

$$\approx (N+Q)\log(N+Q) - N\log N - Q\log Q$$

$$\Rightarrow \log\Omega \approx N\left[\log\left(1+\frac{Q}{N}\right) + \frac{Q}{N}\log\left(\frac{N}{Q}+1\right)\right]$$

For fixed energy per oscillator Q/N, we see that $\Omega \propto e^N$.

It's also instructive to plug in numbers — eg. $N=100$

Q	0	1	2	3	4	5	10	20
Ω	1	100	5050	171700	4.4×10^6	$\sim10^8$	$\sim10^{13.6}$	$\sim10^{22.4}$

2) In a macroscopic system, the spacing between energy levels is extremely small, and we may regard $\Omega(E)$ as a continuous function of E.

[We're glossing over some subtleties here for now, eg. see discussion in Callen 15-5]

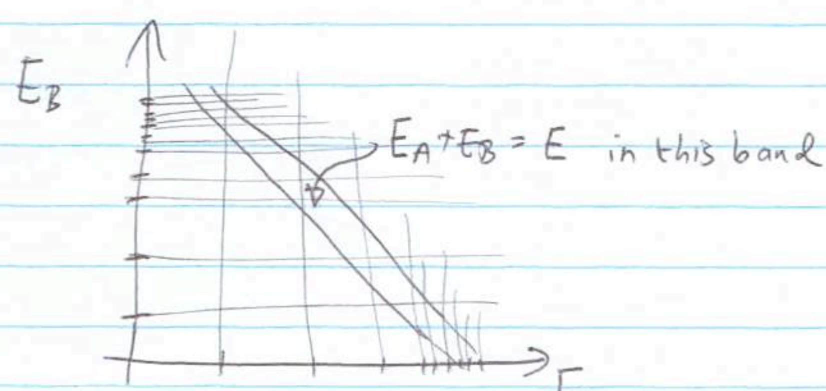
Temperature in statistical mechanics

Now we begin to make the connection to thermodynamics. Consider two systems A and B in thermal contact. The total energy $E = E_A + E_B$ is constant, but the systems may exchange energy with each other.

The number of different arrangements is now

$$\Omega_A(E_A)\,\Omega_B(E_B) = \Omega_A(E_A)\,\Omega_B(E-E_A)$$

Figures 3.1 and 3.2 from Waldram (attached) illustrate the density of states of the joint system in terms of the states of each individual system.



When the two systems are initially placed in contact with some starting values of E_A and E_B, they begin to exchange energy and explore the joint phase space.

We've seen that $\Omega(E)$ increases incredibly rapidly with energy;

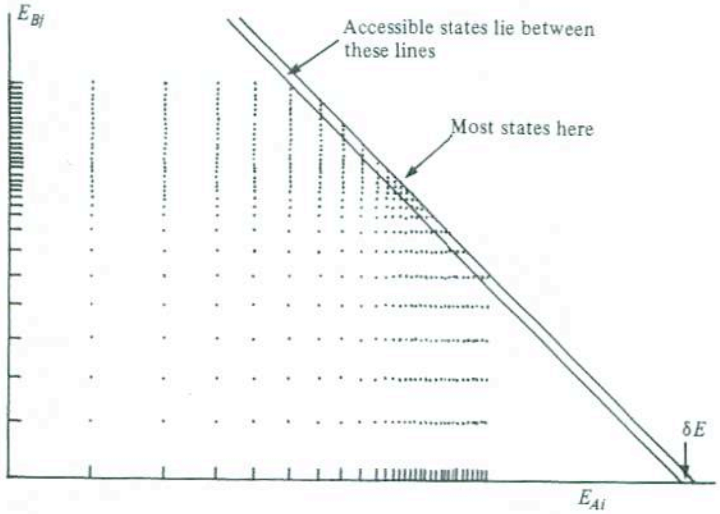
is the total energy in the system, which, for the sake of illustration, I have taken to be 15 000 quanta.

Suppose now that the system starts at some point such as *a*, with a definite energy division between system *A* and system *B*. In general, the joint system will not be in thermal equilibrium, and heat will flow between *A* and *B*, corresponding to a movement parallel to the sloping lines. *What determines the direction of movement?* The answer to this question depends on the values of the product $g_A g_B$ which are marked along the sloping lines. In the direction of state *b* the values of $g_A g_B$ increase *extraordinarily* rapidly (each of the marks between the lines corresponds to an increase in $g_A g_B$ by a factor of 10^{100}). In the direction of state *b′* the number of accessible states *decreases* equally rapidly. This effect is so strong that it seems certain to swamp all other effects, and make the probability of any significant movement towards *b′* totally negligible.

Heat flows so that $g_A g_B$ increases

We can explore this idea in more numerical detail for our model systems. If we assume that *A* and *B* are in relatively weak thermal contact, so that exchange of quanta within each system occurs much more frequently than exchange of quanta between the two systems, then one can show that in the neighbourhood of state *a* the probability that a quantum will jump in the 'right' direction (from *A* to *B*) is twice the probability of the reverse jump. (We omit the details of this calculation.) This makes it clear that *the flow of heat in the 'right' direction is a matter of statistics.* In a time so short that only one or two jumps have occurred the probability that heat will flow in the 'wrong' direction is quite high. However, if we wait long enough for 1500 jumps to have

Fig. 3.1. Joint states of a pair of systems *A* and *B* which are in thermal contact. Each dot represents a joint state. If the joint system is thermally isolated then, using the accessibility convention, only those dots which lie between the sloping lines represent accessible states.



and figure 3.2 of Waldram shows that this is true of $\Omega_A \Omega_B$ at the starting location, such that the number of available states increases when ~~B gives~~ A gives energy to B, and decreases when B gives energy to A. The increase in the number of states is so rapid as energy is transferred from A to B that ~~we will see~~ the system is overwhelmingly likely to move in this direction. Moving the other way (E_A increases) is not forbidden but just very very unlikely — essentially zero probability.

We see then that heat will flow such that

$$d(\Omega_A \Omega_B) > 0$$

$$\text{or} \qquad d\ln(\Omega_A \Omega_B) = d\ln\Omega_A + d\ln\Omega_B > 0$$

Write this in terms of the energy transfer $\delta E = \delta E_A = -\delta E_B$

$$\frac{\partial \ln\Omega_A}{\partial E_A}\,\delta E_A + \frac{\partial \ln\Omega_B}{\partial E_B}\,\delta E_B \geqslant 0$$

$$\delta E \left(\frac{\partial \ln\Omega_A}{\partial E_A} - \frac{\partial \ln\Omega_B}{\partial E_B} \right) > 0$$

$\Rightarrow$ if $\dfrac{\partial \ln\Omega_A}{\partial E_A} > \dfrac{\partial \ln\Omega_B}{\partial E_B}$ then $\delta E > 0$ energy flows from B to A

if $\dfrac{\partial \ln\Omega_A}{\partial E_A} < \dfrac{\partial \ln\Omega_B}{\partial E_B}$ then $\delta E < 0$ energy flows from A to B.

— The quantity "$\dfrac{\partial \ln\Omega}{\partial E}$" appears to be a measure of "coldness".

— We define the temperature of a system as

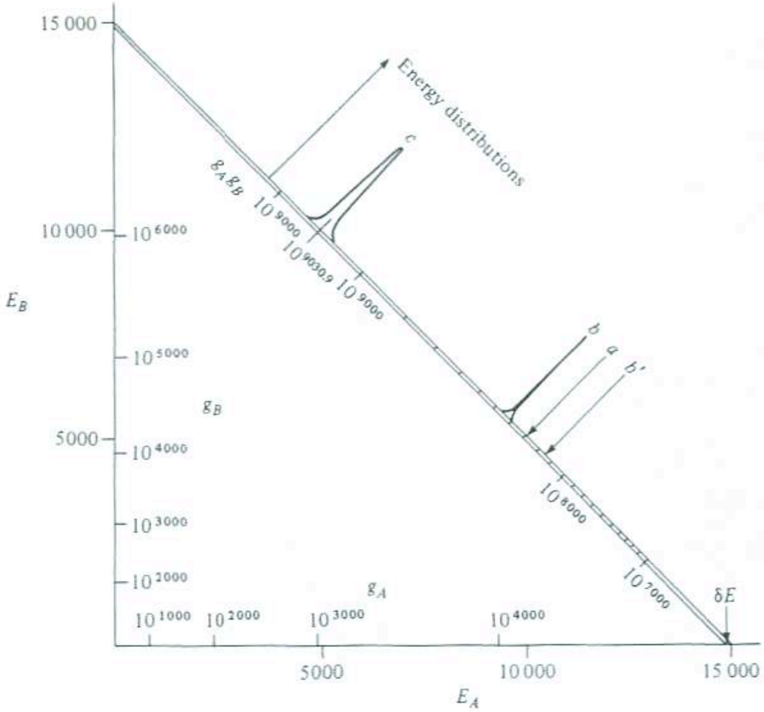
$$\boxed{\frac{1}{k_B T} = \left(\frac{\partial \ln\Omega}{\partial E}\right)_{N,V}}$$

where k_B is Boltzmann's constant $k_B = 1.38 \times 10^{-23}\ \mathrm{JK^{-1}}$

at flows so that
$_B$ increases

occurred, on average we expect 1000 to occur in the 'right' direction and 500 in the 'wrong' direction. We can also work out the standard deviations in these numbers, and we find (again omitting the details) that after 1500 jumps we expect E_A to take the value 9500 ± 36 quanta. The distribution of E_A after 1500 jumps is shown in Fig. 3.2 (curve *b*). Notice how narrow it is. There is not much uncertainty in the quantity of heat which we expect to flow from *A* into *B* in this time. There is, of course, a very small probability that after 1500 jumps the heat will have flowed the 'wrong' way, corresponding to the very small tail of curve *b* on the 'wrong' side of the starting state. For our model system this probability is about 10^{-38}. The probability that a *substantial* amount of heat would have flowed the 'wrong' way, say as far as state *b*′, is about 10^{-150}, a number so small that we can safely say that we should never see this event occur, even if we made the trial every microsecond during the whole history of the universe! Similar considerations apply to all pairs of large systems. *It is therefore safe to assume that when a substantial heat flow occurs between large*

Fig. 3.2. Heat flow between two systems. System *A* consists of 5000 oscillators and system *B* of 10 000 oscillators. The unit of energy is *hν*. The joint system starts in state *a*. After 1500 exchanges of quanta it reaches state *b*: the corresponding energy distribution is plotted normal to the sloping lines. After about 100 000 exchanges the system approaches the equilibrium distribution *c*, which is proportional to $g_A g_B$. The width of curve *c* represents the range of thermal fluctuations of E_A (or E_B) in equilibrium.



We see that the fundamental reason for the Clausius statement of the 2nd law of thermodynamics (heat flows from hotter to colder body) is the extremely rapid increase in the number of available microstates in that direction. <u>The 2nd law is statistical in nature.</u>

<u>Example</u>: let's apply this to our system of harmonic oscillators.

We previously found that

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

where $Q = \frac{E}{h\nu}$ is the number of quanta.

The temperature of the system is given by

$$\frac{1}{k_B T} = \frac{\partial \ln \Omega}{\partial E} = \frac{1}{h\nu}\frac{\partial \ln \Omega}{\partial Q}$$

[treating Q as a continuous variable]

$$= \frac{1}{h\nu}\ln\left(\frac{N}{Q}+1\right)$$

$$\Rightarrow \quad \underline{\frac{k_B T}{h\nu} = \frac{1}{\ln\left(1+\frac{N}{Q}\right)}}$$

or at a given temperature, we can calculate the energy per particle

$$\boxed{\frac{E}{N} = \frac{h\nu}{\exp\left(\frac{h\nu}{kT}\right)-1}}$$

the mean number of quanta per oscillator is

$$\frac{E}{Nh\nu} = \frac{1}{\exp\left(\frac{h\nu}{kT}-1\right)}$$

also known as the occupation number

<u>Two limits</u>

hot $\quad \frac{h\nu}{kT} \ll 1 \qquad E \simeq NkT$ ← each oscillator has kT on average

cold $\quad \frac{h\nu}{kT} \gg 1 \qquad E \simeq Nh\nu\, \underbrace{e^{-h\nu/kT}}_{\text{"Boltzmann factor"}}$

As heat flows from A to B in figure 3.2, the system eventually ends up at the values of E_A and E_B for which $\Omega_A \Omega_B$ is a maximum. The vast majority of the available states are located at this maximum.

The condition for thermal equilibrium is therefore

$$d(\ln \Omega_A \Omega_B) = 0 \Rightarrow \frac{\partial \ln \Omega_A}{\partial E_A} = \frac{\partial \ln \Omega_B}{\partial E_B}$$

or $\boxed{T_A = T_B}$ heat flows to equalize the temperatures.

Entropy

Recall that in classical thermodynamics $\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{N,V}$

which suggests the definition of entropy

$$\boxed{S = k_B \ln \Omega}$$

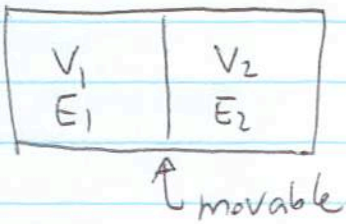
This is Boltzmann's famous formula, and it gives us the physical interpretation of entropy that I promised — the entropy measures the number of available microstates consistent with a given macrostate, or our state of ignorance about the microstate.

The microcanonical ensemble, therefore, gives us a way to calculate all the thermodynamic properties of a system. First determine $\Omega(E, V, N) \longrightarrow S(E, V, N)$

from which all thermodynamic properties can be calculated.

Pressure from the volume dependence of the entropy

Consider two systems joined as before, but able to change their volumes V_1 and V_2 (with the total $V = V_1 + V_2$ constant) as well as exchange energy.



Physically we expect the volumes to change until the pressures of the two systems balance.

How do we understand this in terms of the microcanonical ensemble picture? The number of microstates Ω_1 and Ω_2 depend on the volumes V_1 and V_2. The simplest example is a gas of non-interacting particles, for which we expect $\Omega \propto V^N$. [Since each particle can be put into a number of spatial locations which is $\propto V$]

As before, we expect the system to move in the direction

$$d(\Omega_1 \Omega_2) > 0$$

[as before, note that it is not impossible to move in the other direction, just v. v. v. unlikely!]

$$\Rightarrow \quad \left(\frac{\partial \ln \Omega_1}{\partial E} - \frac{\partial \ln \Omega_2}{\partial E}\right)\delta E + \left(\frac{\partial \ln \Omega_1}{\partial V} - \frac{\partial \ln \Omega_2}{\partial V}\right)\delta V > 0$$

$$\begin{bmatrix} \delta E = \delta E_1 = -\delta E_2 \\ \delta V = \delta V_1 = -\delta V_2 \end{bmatrix}$$

$$\Rightarrow \quad \delta V_1 > 0 \ (\text{and } \delta V_2 < 0) \ \text{ if } \ \frac{\partial \ln \Omega_1}{\partial V} > \frac{\partial \ln \Omega_2}{\partial V}$$

V_1 expands

and vice versa.

The pressure is given by

$$\boxed{p = T \left.\frac{\partial S}{\partial V}\right|_{N,E}}$$

In equilibrium, the system will be at the maximum of $\Omega_1 \Omega_2$

$$\Rightarrow \quad d(\Omega_1 \Omega_2) = 0$$

$$\Rightarrow \quad \boxed{P_1 = P_2 , \qquad T_1 = T_2}$$

Example 2: Ideal gas

How do we understand an ideal gas of N non-interacting particles in the microcanonical ensemble?

Let's calculate Ω

1) spatial part $\quad \Omega \propto \left(\frac{V}{\Delta V}\right)^N$

2) momentum space

$\Omega \propto$ area of a sphere in 3N dimensions with radius $\sqrt{2mE}$

$\propto (\sqrt{E})^{3N-1}$

for large N, $\quad \Omega \propto E^{3N/2}$

$$\Rightarrow \quad \boxed{S = k_B N \ln\left(V E^{3/2}\right) + \text{constant}}$$

The temperature is $\quad \frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{N,V} = \frac{k_B N}{E^{3/2}} \frac{3}{2} E^{1/2}$

$$\Rightarrow \quad \boxed{E = \frac{3}{2} N k_B T}$$

energy per particle $\frac{3}{2}kT$

The pressure is $\quad \frac{P}{T} = \left.\frac{\partial S}{\partial V}\right|_{E,N} \Rightarrow \boxed{P = \frac{N k_B T}{V}}$

the EQUATION OF STATE for the gas

Example 2 As a different example, let's consider a set of N particles which can each be spin up or spin down
↑ ↓

The spin up state has energy ε relative to the spin down state. (This is a "2 state model" see Chandler 3.4). The particles are non-interacting.

Let's calculate the entropy. The number of spin up particles is $m = \frac{E}{\varepsilon}$.

For a total energy E, the number of microstates is

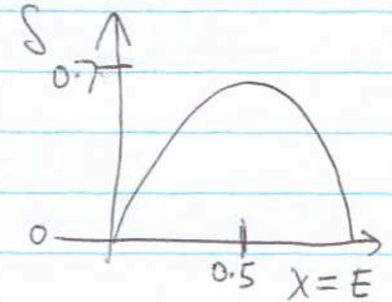
$$\Omega = {}^N C_m = \frac{N!}{m!(N-m)!}$$

$$\Rightarrow \quad S = k_B \ln \frac{N!}{m!(N-m)!}$$

Note that we haven't said anything about the spatial arrangement of the spins, so S is independent of V – there is no pressure in this simple model

use Stirling's approximation

$$S = -k_B N \Big[x \ln x + (1-x)\ln(1-x) \Big] \qquad \text{where } x = \frac{m}{N} = \frac{E}{N\varepsilon}$$



Let's calculate the temperature. We can already see that something strange is happening from a plot of S against E. — for $x > 1/2$ S is a decreasing function of E
⇒ negative temperature!

$$\frac{1}{k_B T} = \left.\frac{\partial S}{\partial E}\right|_{N,V} = \frac{1}{N\varepsilon} \left.\frac{\partial S}{\partial x}\right|_{N,V}$$

$$\Rightarrow \quad \frac{\varepsilon}{k_B T} = -\frac{\partial}{\partial x}\Big(x\ln x + (1-x)\ln(1-x) \Big) = \ln\left(\frac{1-x}{x}\right)$$

$$\Rightarrow \quad \boxed{x = \frac{E}{N\varepsilon} = \frac{1}{1+e^{\varepsilon/kT}}}$$

small T $\quad kT \ll \varepsilon \quad x \simeq e^{-\varepsilon/kT}$ Boltzman factor again

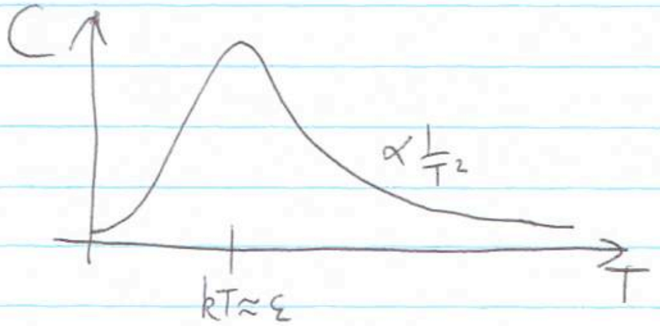
large T $\quad kT \gg \varepsilon \quad x \to \frac{1}{2}$ both states equally occupied

As we noted above if we give the system more energy so that x increases beyond ½ then the temperature is negative — the number of microstates decreases with increasing energy beyond x = ½. [To see that this must be the case consider x = 1 for which there is only one microstate — all particles spin up — and therefore S = 0.) A system with an upper bound on its energy levels is required to be able to achieve negative temperatures.

Note that a negative temperature is NOT colder than T=0 but instead is HOTTER than T=∞ !

ie heat will flow from an object with T<0 to an object with T>0.

The heat capacity of this system is $C = \frac{dE}{dT} = \frac{N\varepsilon^2}{kT^2}\frac{e^{\varepsilon/kT}}{(1+e^{\varepsilon/kT})^2}$



This form of C is referred to as a Schottky heat capacity.

Characteristic of a 2 level system.

The Canonical Ensemble and the Boltzmann Distribution

Now consider a system in equilibrium with a heat bath at constant temperature T. (The heat bath is large enough that its temperature remains constant as it exchanges energy with the system).

What is the probability P_ν of finding the system in state ν with energy E_ν? Since all microstates of the joint system + heat bath are equally likely, then

$$P_\nu \propto \Omega_{bath}(E_{tot} - E_\nu)$$

where $E_{tot} = E_\nu + E_{bath}$ is the total (fixed) energy.

Now, since T is constant, we can expand

$$\ln \Omega_{bath}(E_{tot} - E_\nu) \simeq \ln \Omega_{bath}(E_{tot}) + (-E_\nu) \left.\frac{\partial \ln \Omega}{\partial E_{bath}}\right|_{E_{tot}}$$

$$= \ln \Omega_{bath}(E_{tot}) - \frac{E_\nu}{k_B T}$$

$$\Rightarrow \boxed{P_\nu \propto e^{-E_\nu / k_B T}}$$ This is the Boltzmann distribution

We normalize the probability so that $\sum_\nu P_\nu = 1$, defining the

$$\boxed{\text{PARTITION FUNCTION} \quad Q = \sum_\nu e^{-E_\nu / k_B T}}$$

Then $$P_\nu = \frac{e^{-E_\nu / k_B T}}{Q}$$

[Note that I am using the notation Q for the partition function, which agrees with Chandler. More common is to write the partition function as Z.]

Example: Maxwell-Boltzmann distribution

Distribution of velocities in a gas. Fix u_y, u_z. Then

$$\text{Prob}(u_x)\, du_x \propto e^{-mu_x^2/2k_BT}\, du_x$$

↑ state labelled by u_x

The joint distribution for u_x, u_y, u_z is

$$\text{Prob}(u_x, u_y, u_z)\, du_x\, du_y\, du_z \propto e^{-mu^2/2k_BT}\, du_x\, du_y\, du_z$$

where $u^2 = u_x^2 + u_y^2 + u_z^2$

or since this distribution depends only on the speed $|\underline{u}|$, we can write

$$\text{Prob}(u)\, du \propto e^{-mu^2/2k_BT}\, 4\pi u^2\, du$$

Normalize this such that $\int_0^\infty \text{Prob}(u)\, du = 1$

$$\Rightarrow \boxed{\text{Prob}(u)\, du = \left(\frac{m}{2\pi kT}\right)^{3/2} e^{-mu^2/2kT}\, 4\pi u^2\, du}$$

Maxwell-Boltzmann

The energy of the system is

$$E = \sum E_\nu P_\nu$$

$$= \frac{\sum E_\nu e^{-E_\nu/k_BT}}{Q} = \frac{\sum E_\nu e^{-\beta E_\nu}}{Q}$$

$$= -\frac{\partial Q/\partial\beta}{Q}$$

$\left[\beta = \frac{1}{k_BT}\right.$ use this standard notation $\left.\right]$

$$\Rightarrow \quad \boxed{E = -\frac{\partial \ln Q}{\partial \beta}}$$

We can also calculate the spread in energy

$$\sigma_E^2 = \langle E^2 \rangle - \langle E \rangle^2$$

$$= \sum p_\nu E_\nu^2 - \left(\sum p_\nu E_\nu\right)^2$$

$$= -\frac{\partial}{\partial\beta}\left(\frac{\sum E_\nu e^{-\beta E_\nu}}{\sum e^{-\beta E_\nu}}\right)$$

$$= \frac{\partial^2 \ln Q}{\partial \beta^2} = -\frac{\partial E}{\partial \beta} = kT^2 \frac{\partial E}{\partial T}$$

↑ this is the heat capacity $C_V = \left.\frac{\partial E}{\partial T}\right|_{N,V}$

$$\Rightarrow \quad \boxed{\sigma_E^2 = k_BT^2 C_V}$$

We'll see this kind of result again — the size of statistical fluctuations related to thermodynamic property.

Note that $C_V \propto N$ and $E \propto N$

$$\Rightarrow \boxed{\frac{\sigma_E}{E} \propto \frac{1}{\sqrt{N}}}$$

for macroscopic system, the fluctuations in energy are tiny eg. $N = 10^{23} \Rightarrow \frac{\sigma_E}{E} \sim 10^{-11}$

Relation to Helmholtz free energy

The Helmholtz free energy is $A = E - TS$
(again, follow Chandler – often this is written as F)

The 1st law of thermodynamics $dE = TdS - pdV$

$$\Rightarrow \quad dA = -SdT - PdV$$

$$\Rightarrow \quad \boxed{S = -\frac{\partial A}{\partial T}\bigg|_{N,V}}$$

$$\Rightarrow \quad E = A + TS = A - T\frac{\partial A}{\partial T} = -T^2 \frac{\partial}{\partial T}\left(\frac{A}{T}\right)$$

but $E = -\dfrac{\partial \ln Q}{\partial \beta} = kT^2 \dfrac{\partial \ln Q}{\partial T}$

$$\Rightarrow \boxed{A = -k_B T \ln Q}$$

So in the canonical ensemble, we first compute $Q = \sum_\nu e^{-E_\nu/kT}$

then $A = -kT \ln Q$ and then from A we can obtain all other thermodynamic quantities, eg. $P = -\frac{\partial F}{\partial V}\Big|_T$ $\quad S = -\frac{\partial F}{\partial T}\Big|_V$

Since $A = -k_BT \ln Q$, we can write $Q = e^{-A/k_BT}$

and therefore an alternative way to write P_ν is

$$\boxed{P_\nu = e^{-E_\nu/kT}\, e^{+A/kT}}$$

Why is it that the free energy is a natural quantity to think about when discussing a system coupled to a heat bath — ie. an open system able to transfer energy to/from its surroundings?

We know that the evolution of the system will be such that the total entropy of A and B will increase. If the energy of A changes by ΔE_A, then the entropy of the bath changes by an amount $T\Delta S_B = \Delta E_B = -\Delta E_A$.
Therefore

$$\Delta(S_A + S_B) \geqslant 0 \quad \Rightarrow \quad \Delta(E_A - TS_A) \leqslant 0$$

$$\text{or} \quad \Delta A \leqslant 0$$

The system A evolves in such a way as to minimize its Helmholtz free energy. We can think of this as a simple generalization of the 2nd law for open systems maintained at temperature T by contact with a heat bath. Rather than keeping track of the entropy of the "whole universe" (ie. A + B) we can just follow the free energy of the system that we're interested in.

Uncoupled components: factorization

One of the advantages of the canonical ensemble approach is that if the system consists of uncoupled components, the partition function factorizes.

eg.
$$Q = \sum_{i,j} e^{-\beta(E_i+E_j)} = \left(\sum_i e^{-\beta E_i}\right)\left(\sum_j e^{-\beta E_j}\right)$$
$$= Q_i Q_j$$

Note that this means that the free energy A is additive
$$A = kT \ln Q = kT \ln(Q_i Q_j) = A_i + A_j$$
and internal energy etc.

Example: Let's go back to the N spin up/spin down particles.

The partition function is
$$Q = Q_i^N$$
$$= \left(1 + e^{-\varepsilon/kT}\right)^N$$

The energy is
$$E = -\frac{\partial \ln Q}{\partial \beta} \cancel{= kT^2}$$

$$= -N \frac{\partial}{\partial \beta} \ln\left(1 + e^{-\varepsilon\beta}\right)$$

$$E = \frac{N\varepsilon e^{-\varepsilon/kT}}{1 + e^{-\varepsilon/kT}}$$

$$\Rightarrow \boxed{\frac{E}{N\varepsilon} = \frac{1}{1 + e^{\varepsilon/kT}}}$$

same result as previously

Example: ideal gas

The partition function is

$$Q = \frac{1}{N!} \prod_{i=1}^{3N} \frac{1}{h} \int_0^L dx_i \int_{-\infty}^{\infty} dp_i \; e^{-p_i^2/2m_i kT}$$

indistinguishable particles

Quantum mechanics sets the phase space density

$$= \left(\frac{L}{h} \sqrt{\frac{2\pi m}{\beta}} \right)^{3N} \frac{1}{N!}$$

$$\Rightarrow \quad Q = \left(\frac{L}{\lambda}\right)^{3N} \frac{1}{N!} \quad \text{where} \quad \lambda = \left(\frac{2\pi\hbar^2}{mkT}\right)^{1/2}$$

The energy is

$$E = -\frac{\partial \ln Q}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln \beta^{-3N/2}$$

$$= \frac{3N}{2\beta} = \boxed{\frac{3}{2} N k_B T} \quad ✓$$

The pressure is

$$\frac{\partial A}{\partial V}\Big|_T = k_B T \frac{\partial \ln Q}{\partial V} = k_B T \frac{\partial}{\partial V} \left(\ln V^N\right)$$

$$\Rightarrow \boxed{\frac{N k_B T}{V} = P} \quad ✓$$

The free energy is

$$A = -k_B T \ln Q = -k_B T N \ln\left(\frac{V}{h^3} \left(2\pi m kT\right)^{3/2} \right)$$

$$+ kTN(\ln N - 1) \quad \left[\text{using Stirling}\right]$$

$$\boxed{A = +k_B T N \left[\ln\left(\frac{n}{n_Q}\right) - 1 \right]} \qquad n_Q = \left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2}$$

[n_Q is known as "n-quantum"]

The entropy is $S = -\left.\frac{\partial A}{\partial T}\right|_{V,N}$

$$= -k_B N\left[\ln\left(\frac{n}{n_Q}\right) - 1\right]$$

$$+ k_B T N \frac{\partial \ln \boxed{n_Q}}{\partial T} \quad \propto T^{3/2}$$

$$= -k_B N \ln\left(\frac{n}{n_Q}\right) + \frac{5}{2} k_B N$$

or $$\boxed{S = N k_B\left[\frac{5}{2} - \ln\left(\frac{n}{n_Q}\right)\right]}$$

the famous <u>Sackur-Tetrode formula</u>

Note that the entropy is <u>extensive</u> – if we double N, V then S also doubles. This would not be the case if we hadn't included the N! term in Q to account for indistinguishable particles. This was an early puzzle in the field.

[In fact, Gibbs noted that the entropy increased on mixing two identical samples of gas – the Gibbs' paradox – clearly incorrect since we can reinsert a partition and recover the initial state – the process is reversible. This paradox is resolved once the N! is included.]

Entropy in the canonical ensemble

Let's evaluate the average value of $\ln P_\nu$

ie. $$\sum P_\nu \ln P_\nu = -\frac{\sum P_\nu E_\nu}{kT} - \sum P_\nu \ln Q$$
$$= -\frac{E}{kT} - \ln Q$$
$$= -\frac{E}{kT} + \frac{A}{kT} = -\frac{S}{k_B} \qquad \begin{pmatrix}\text{since} \\ A = E - TS\end{pmatrix}$$

$$\Rightarrow \quad \boxed{S = -k_B \sum P_\nu \ln P_\nu} \; (*)$$ This is known as the GIBBS ENTROPY

This is a very general expression for the entropy of a system.

— For the microcanonical ensemble, where $P_\nu = \frac{1}{\Omega}$ = constant,

it reduces to $$S = -k_B \sum_{n=1}^{\Omega} \frac{1}{\Omega} \ln \frac{1}{\Omega} = k_B \ln \Omega$$ our previous result

— In the case of the microcanonical or canonical ensembles, where the P_ν's refer to the equilibrium state, S is referred to as the "equilibrium entropy". However, equation (*) can be taken as a definition of entropy for non-equilibrium systems.

— In information theory, the Shannon entropy $S = -\sum P_\nu \log_2 P_\nu$ is used as a measure of the <u>disorder</u> or <u>unpredictability</u> in a ~~system~~ signal. The choice of the constant is such that the logarithm is base 2 — the entropy is measured in "bits" of information.

For example, how many shuffles are required to randomize a pack of cards? Initially, the order of the cards is specified — only one configuration is possible $S = 0$. When fully randomized,

[the other example to mention is compression of a data stream
eg. sequence of random 1's and 0's – need all N bits
weather RRSSSSRRRSS .. correlated sequence don't need all N bits]

the number of configurations is 52!, each of which are equally likely, giving $S = -\log_2 52! = -\dfrac{\ln 52!}{\ln 2} = -226$.

Consider a "riffle shuffle" in which the deck is exactly halved and reconstructed by choosing randomly from each pile. The number of configurations increases each time by 2^{52}

$$\Rightarrow \quad \Delta S = -\log_2 2^{52} = -52.$$

$\therefore$ we predict that $\dfrac{226}{52} = 4.3$ shuffles will randomise the deck.

The actual answer is 6 or 7 – see for example Trefethen & Trefethen (2000) Proc R. Soc. London A <u>436</u> 2561.

Probability distributions that maximize entropy

You may recall when we discussed the microcanonical ~~distrib~~ ensemble that I argued that the principle of equal probabilities should be exactly what we expect given our lack of knowledge about the microstates. All we know is the total energy, and so without further knowledge we should expect all microstates compatible with that total energy to be equally likely.

Now we have a way to quantify our lack of knowledge. Let's find the probability distribution P_ν which maximizes $S = -k_B \sum P_\nu \ln P_\nu$ subject to the constraint $\sum P_\nu = 1$.

To do this, use Lagrange multiplier

$$\delta\left(-k_B \sum P_\nu \ln P_\nu + \lambda\left(\sum P_\nu - 1\right)\right) = 0$$

vary P_ν $\Rightarrow$ $-k_B \sum \delta P_\nu \ln P_\nu - k_B \sum \delta P_\nu + \lambda \sum \delta P_\nu = 0$

vary λ $\Rightarrow$ $\delta\lambda\left(\sum P_\nu - 1\right) = 0$ $\rightarrow$ this must be true for all $\delta\lambda$, fixing the constraint $\sum P_\nu = 1$

$$\Rightarrow \quad \sum \delta P_\nu \left[-k_B \ln P_\nu - k_B + \lambda \right] = 0$$

must be true for arbitrary choices of the δP_ν values

$$\Rightarrow \text{ for each } \nu \text{ we must have } \ln P_\nu = \frac{\lambda}{k_B} - 1 = \underline{\text{constant}}.$$

⇒ Equal probabilities P_ν maximize the entropy when the only constraint is $\sum P_\nu = 1$.

What if we have an additional constraint $\sum P_\nu E_\nu = E$?

then $$\delta\left(-k_B \sum P_\nu \ln P_\nu + \lambda\left(\sum P_\nu - 1 \right) + \alpha\left(\sum P_\nu E_\nu - E \right) \right) = 0$$

Same procedure as before

$$\Rightarrow \quad \ln P_\nu + 1 - \frac{\lambda}{k_B} - \frac{\alpha E_\nu}{k_B} = 0$$

$$\Rightarrow \quad \ln P_\nu = \frac{\alpha E_\nu}{k_B} + \text{constant}.$$

Now demand that $\sum P_\nu E_\nu = E$ and calculate $\frac{\partial S}{\partial E} = \frac{1}{T}$

– this fixes the value of $\alpha = -\frac{1}{T}$

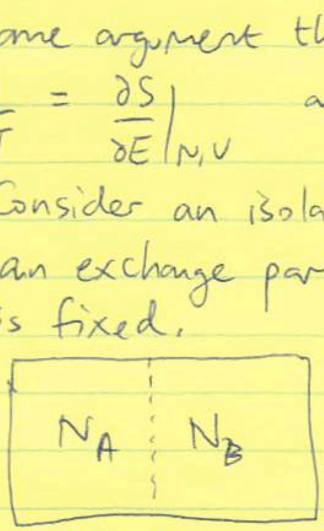
$$\Rightarrow \quad \boxed{\ln P_\nu = -\frac{E_\nu}{k_B T} + \text{constant}}$$

we've recovered the Boltzmann distribution.

Chemical potential

Now let's consider changes in particle number. We'll follow the same argument that led us to the definitions of temperature $\frac{1}{T} = \frac{\partial S}{\partial E}\Big|_{N,V}$ and pressure $\frac{p}{T} = \frac{\partial S}{\partial V}\Big|_{N,E}$.

Consider an isolated system made up of two subsystems that can exchange particles. The total number of particles $N = N_A + N_B$ is fixed.



$N = N_A + N_B$

The system will evolve such that

$$d\,(\Omega_A \Omega_B) \geqslant 0$$

or

$$\left(\frac{\partial \ln \Omega_A}{\partial N} - \frac{\partial \ln \Omega_B}{\partial N}\right)\delta N_A \geqslant 0$$

We define the CHEMICAL POTENTIAL as $\boxed{\mu = -k_B T \frac{\partial \ln \Omega}{\partial N}\Big|_{E,V}}$

or $\boxed{\mu = -T\frac{\partial S}{\partial N}\Big|_{V,E}}$. $\Rightarrow \underline{(\mu_B - \mu_A)\,\delta N_A \geqslant 0}$

if $\mu_A > \mu_B$ then $\delta N_A < 0$ particles flow from A to B
$\mu_A < \mu_B$ " $\delta N_A > 0$ " " B to A

In equilibrium, the chemical potentials balance $\mu_A = \mu_B$.

Other ways to calculate μ:
For a general entropy change

$$T\,dS = T\frac{\partial S}{\partial E}\Big|_{N,V} dE + T\frac{\partial S}{\partial V}\Big|_{N,E} dV + T\frac{\partial S}{\partial N}\Big|_{V,E} dN$$

$$= dE + p\,dV - \mu\,dN$$

(1st law)

$\Rightarrow$ we can also write $\boxed{\mu = \left.\frac{\partial E}{\partial N}\right|_{V,S}}$

In terms of A, recall that $A = E - TS$

$$\Rightarrow \quad dA = -TdS - SdT + dE$$
$$= -SdT - PdV + \mu dN$$

$$\Rightarrow \boxed{\mu = \left.\frac{\partial A}{\partial N}\right|_{T,V}}$$

So for our system in contact with a heat bath, μ is the free energy cost of adding a particle.

For the ideal gas we had $A = k_B T N \left[\ln\left(\frac{N}{V n_Q}\right) - 1 \right]$

$$\Rightarrow \boxed{\mu = k_B T \ln\left(\frac{n}{n_Q}\right)} \quad \text{or} \quad n = n_Q e^{-\mu/kT}$$

Grand Canonical Ensemble

Set up a system in contact with a heat bath as before, but now the heat bath is also a particle reservoir that can supply or accept particles at constant chemical potential.

We make the same argument as for the canonical ensemble. In equilibrium

$$P_\nu \propto \Omega_{bath}(E-E_\nu, N-N_\nu)$$

Expand

$$\ln \Omega_{bath}(E-E_\nu, N-N_\nu) \simeq \ln \Omega_{bath}(E,N) - E_\nu \frac{\partial \ln \Omega_{bath}}{\partial E} - \frac{\partial \ln \Omega_{bath}}{\partial N} N_\nu$$

$$= \ln \Omega_{bath}(E,N) - \frac{E_\nu}{k_B T} + \frac{\mu}{k_B T} N_\nu$$

$$\Rightarrow \boxed{P_\nu \propto \exp\left[-\frac{(E_\nu - \mu N_\nu)}{k_B T} \right]}$$

The "grand partition function" is $\boxed{Z = \sum_\nu e^{-(E_\nu - \mu N_\nu)/k_B T}}$

[this is often written as capital xi, Ξ, but for clarity we'll use Z.]

The associated free energy is the "grand free energy" Φ

$$\boxed{\Phi = -kT \log Z = E - TS - \mu N}$$

As before with the canonical ensemble / Helmholtz free energy, the idea is that the grand free energy keeps track of the entropy of the "rest of the universe" as the system changes its energy and particle number at constant T and μ. In equilibrium, Φ is minimized.

Notes

1) All other thermodynamic quantities can be derived from Φ. Using the 1st law you can show that

$$d\Phi = -p\,dV - S\,dT - N\,d\mu$$

$$\Rightarrow \quad p = -\left.\frac{\partial \Phi}{\partial V}\right|_{T,\mu} \qquad S = -\left.\frac{\partial \Phi}{\partial T}\right|_{\mu,V} \qquad N = -\left.\frac{\partial \phi}{\partial \mu}\right|_{T,V}$$

$$\text{or } N = \left.\frac{\partial \ln Z}{\partial \mu}\right|_{T,V} k_B T$$

The energy is given by $E = \mu N - \frac{1}{Z}\left.\frac{\partial Z}{\partial \beta}\right|_{V,\mu}$.

~~Partial~~

2) We can write the grand partition function as a sum over the canonical partition functions for different particle numbers.

$$Z = \sum_{\text{all microstates}} e^{-E_\nu/kT} e^{\mu N_\nu/kT}$$

$$= \sum_{\substack{\text{particle number} \\ M}} \ \sum_{\substack{\text{energy} \\ \text{eigenstates with} \\ M \text{ particles}}} e^{-E_\nu(M,T,V)/kT} e^{\mu M/kT}$$

$$= \sum_M e^{\mu M/kT} Q(M,T,V)$$

$$= \sum_M e^{-(A(M,T,V)-\mu M)/kT}$$

3) Fluctuations

Evaluate the quantity $\left.\frac{\partial N}{\partial \mu}\right|_{T,V} = -\left.\frac{\partial^2 \Phi}{\partial \mu^2}\right|_{T,V}$

$$= k_B T \frac{\partial}{\partial \mu} \frac{\partial \log Z}{\partial \mu}$$

$$= \cancel{k_B T} \frac{\partial}{\partial \mu} \left[\frac{1}{Z} \sum_\nu N_\nu e^{-E_\nu/kT} e^{\mu N_\nu/kT} \right]$$

$$= \frac{\langle N^2 \rangle}{kT} - \frac{\langle N \rangle^2}{kT}$$

$$\Rightarrow \quad \boxed{\sigma_N^2 = \langle N^2 \rangle - \langle N \rangle^2 = k_B T \left.\frac{\partial N}{\partial \mu}\right|_{T,V}}$$

(compare our previous result $\sigma_E^2 = k_B^2 T^2 C_V$)

4) We could obtain the grand canonical ensemble by maximizing the Gibbs entropy with 3 constraints: $\sum P_\nu = 1$ $\sum P_\nu E_\nu = E$ $\sum P_\nu N_\nu = N$

Different free energies and their ensembles

Depending on the constraints placed on a system, there are several different free energies that could apply. Each has a corresponding ensemble - eg. see Hw2 Q3 Gibbs ensemble.

eg.	$A(T,V,N)$	Helmholtz	$A = E - TS$
	$G(T,P,N)$	Gibbs	$G = E - TS + PV$
	$H(S,P,N)$	Enthalpy	$H = E + PV$
	$\Phi(T,V,\mu)$	Grand	$\Phi = E - TS - \mu N$

II QUANTUM STATISTICAL MECHANICS

(Non-interacting particles)

[Chandler Chapter 4
Sethna Chapter 7]

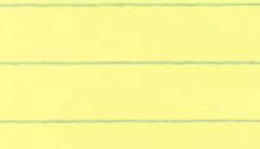
Applications - blackbody radiation
phonons
Bose-Einstein condensation
electrons in metals
white dwarfs and neutron stars

Bose-Einstein and Fermi-Dirac distributions

In a system of non-interacting particles, we can think of the particles as filling the single particle states of the system.

eg. quantum harmonic oscillator

energy levels $\varepsilon_k = (k+\frac{1}{2})\hbar\omega$



[In other words we construct the quantum state of the system from combinations of the single particle states.]

How the states can be filled depends on whether the particles are bosons or fermions.

BOSONS (integer spin) - no limit to the number of particles in each state $n_k = 0, 1, 2, 3 \ldots$

FERMIONS (half-integer spin) - obey the Pauli exclusion principle $n_k = 0$ or 1

What is the set of occupation numbers $\{n_k\}$ for a system in equilibrium with a heat/particle bath at temperature T and chemical potential μ?

The grand canonical partition function is

$$\mathcal{Z} = \sum_\nu e^{-(E_\nu - \mu N_\nu)/kT}$$

where $E_\nu = \sum_k n_k \varepsilon_k$ and $N_\nu = \sum n_k$

We can write this in a different way.
For a given ν (or a set of n_k values)

$$e^{-(E_\nu - \mu N_\nu)/kT} = e^{-\sum n_k(\varepsilon_k-\mu)/kT}$$
$$= \prod_k e^{-n_k(\varepsilon_k-\mu)/kT}$$

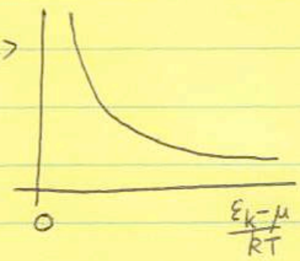
$$\Rightarrow \quad Z = \sum_\nu \prod_k e^{-n_k(\varepsilon_k-\mu)/kT}$$
$$= \prod_k \sum_{\substack{\text{allowed}\\ n_k \text{ values}}} e^{-n_k(\varepsilon_k-\mu)/kT}$$
$$= \prod_k Z_k \qquad \text{where } Z_k = \sum_{n_k} e^{-n_k(\varepsilon_k-\mu)/kT}$$

the grand partition function factorizes.

We can think of each single particle state k being filled independently from the bath.

Two cases –

1) Bosons
$$Z_k = \sum_{n_k=0}^{\infty} e^{-n_k(\varepsilon_k-\mu)/kT}$$
$$= \frac{1}{1-e^{-\beta(\varepsilon_k-\mu)}}$$



$$\Rightarrow \quad \boxed{Z = \prod_k \frac{1}{1-e^{-\beta(\varepsilon_k-\mu)}}} \quad \text{or equivalently } \Phi = \sum \Phi_k$$

The mean occupation number is

$$\boxed{\langle n_k \rangle = -\frac{\partial \Phi_k}{\partial \mu} = \frac{1}{e^{\beta(\varepsilon_k-\mu)}-1}} \quad \text{Bose-Einstein distribution}$$

When the occupancy is low, μ is large and negative $\langle n_k \rangle \simeq e^{\frac{(\mu-\varepsilon_k)}{kT}}$
As μ approaches ε_k from below $\langle n_k \rangle \to \infty$

[therefore μ is always less than the lowest energy eigenvalue]

2) Fermions $$Z_k = 1 + e^{-(\varepsilon_k - \mu)/kT}$$

$$\boxed{Z = \prod_k 1 + e^{-\beta(\varepsilon_k - \mu)}}$$

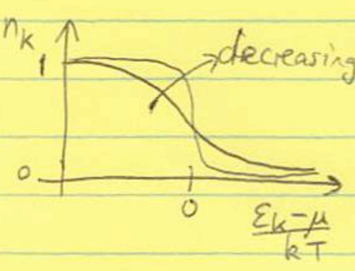
and

$$\boxed{\langle n_k \rangle = \frac{1}{e^{\beta(\varepsilon_k - \mu)} + 1}}$$ Fermi-Dirac distribution

Again, low occupancy corresponds to μ large and negative, and

$\langle n_k \rangle \simeq e^{(\mu - \varepsilon_k)/kT}$

At low T, μ divides the filled and empty states



Counting indistinguishable particles; Maxwell-Boltzmann statistics

The reason that a set of occupation numbers $\{n_k\}$ labels the microstate is that the particles are indistinguishable — we don't need to keep track of which particle goes into which single particle state, we just need to follow how many particles are in each single particle state.

Previously, we dealt with indistinguishable particles by dividing the partition function by N!, ie. $Q = \frac{(Q_{\text{single particle}})^N}{N!}$. We saw that the N! was necessary, for example, to make the entropy of the ideal gas extensive. How does this relate to what we've done here?

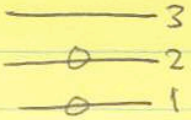
[We'll call this "Maxwell-Boltzmann statistics"]

To see this, consider a simple case: ~~3~~ 2 particles in a system

Use a shorthand:

with 3 single particle states. ("energy levels")

eg.



= we'll write as "12"

Meaning particle 1 into state 1
and " 2 " 2

If the particles are fermions, there can only be one particle in each energy level ⇒ 3 possible microstates 12, 13, 23
[Note that 21 for example is not an additional microstate - it is included when we write 12 because we can't tell in quantum mechanics which particle is in level 1 and which is in level 2 — all we know is that there is one particle in each level].

For bosons, we can have 11, 12, 13, 22, 23, 33 6 states.

Now classically, we can keep track of each particle so 12 and 21 are distinct states for distinguishable particles.
There are 9 states for distinguishable particles:

11 12 13 ~~4~~ 21 22 23 31 32 33

The partition function is a sum over these 9 - alternatively we can multiply

$$Q = Q_1 Q_2 = (1+2+3)(1+2+3) = (Q_{\text{single particle}})^2$$
$$= 11 \quad 12 \quad 13 \quad 21 \quad 22 \quad 23 \quad 31 \quad 32 \quad 33$$

What do we get if we divide by N! ?

$$Q = \frac{(Q_{\text{single particle}})^2}{2!}$$
$$= \underbrace{\frac{(11 + 22 + 33)}{2}} + \underbrace{\frac{21 + 12}{2} + \frac{31 + 13}{2} + \frac{23 + 32}{2}}$$

this is not correct! All of these are unique states of indistinguishable particles - no correction required!

these terms are correct - because if the particles are indistinguishable 21 and 12 are the same so we divide by 2 to correct for overcounting

Notice that the states in which energy levels are singly occupied are correctly included – the problem is with states that are multiply occupied, which are divided by 2. The resulting partition function lies somewhere between the boson and fermion partition functions.

In the classical limit, where $\langle n_k \rangle \ll 1$ (microstates are ~~sparsely sampled~~ sparsely occupied) we get the same answer for fermions, bosons, or with the Maxwell-Boltzmann statistics.

To see this for the MB case, write

$$Z = \sum_M e^{\beta\mu M} \frac{(Q_1)^M}{M!}$$

[using the fact that we can write Z as a sum over number of particles involving $Q(N)$]

$$\Rightarrow \quad Z = \exp\left(e^{\beta\mu} Q_1 \right)$$

$$= \exp\left(\sum_k e^{-\beta(\varepsilon_k - \mu)} \right)$$

$$= \prod_k \exp\left(e^{-\beta(\varepsilon_k - \mu)} \right) = \prod_k Z_k$$

The occupation number is

$$\langle n_k \rangle = \frac{-\partial \Phi_k}{\partial \mu} = e^{-\beta(\varepsilon_k - \mu)}$$

So we have

$$\langle n_k \rangle = \frac{1}{e^{\beta(\varepsilon_k - \mu)} + c} \qquad \text{with } c = \begin{cases} 0 & \text{MB} \\ 1 & \text{FD} \\ -1 & \text{BE} \end{cases}$$

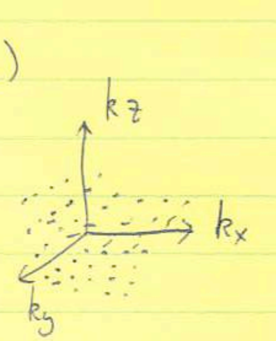
all of which give the same answer for $\langle n_k \rangle \ll 1$ or μ large and negative (classical limit). But when $\langle n_k \rangle$ approaches unity quantum effects become important and we must take into account whether the particles are bosons $(c=-1)$ or fermions $(c=+1)$. [No known particles obey MB statistics $(c=0)$! But it gives the correct classical limit]

The density of states in k-space

We want to calculate the partition function for Fermi and Bose gases, but first we need to know how many states there are with energy E_k? Quantum mechanics tells us that we can describe particles as waves with ~~energy~~ momentum $p = \hbar k = \frac{h}{\lambda}$ where $k = \frac{2\pi}{\lambda}$ is the wavevector and λ the wavelength. In a 1D box of side L, the allowed values of λ are $\lambda = \frac{2L}{n}$ where $n = 1, 2, 3 \ldots$

or $$k = \frac{2\pi n}{2L} = \frac{\pi n}{L}.$$

In 3D, the wavector is $$\underline{k} = \frac{\pi}{L}(n_x, n_y, n_z)$$



Allowed values of k fill a grid in k-space. For macroscopic systems we treat $\underline{k}$ as continuous.

The density of states is $$\left(\frac{1}{\text{Volume per allowed k-value}}\right)$$

$$= \frac{L^3}{\pi^3} = \frac{V}{\pi^3}.$$ in the quadrant $k_x, k_y, k_z > 0$

If we distinguish between left and right going particles then we must allow +ve and −ve k values, and the density of states is $\frac{V}{8\pi^3} = \frac{V}{(2\pi)^3}$.

Another way to arrive at this result is to consider a volume of free space and adopt periodic boundary conditions − then we must have $\lambda = \frac{L}{n}$

$\Rightarrow$ density of states $= \frac{V}{(2\pi)^3}$.

Dividing by V, we get the density of states in momentum-position space

$\#$ states within small volume of $\underline{k}$ and $\underline{x}$ is $\dfrac{d^3\underline{x}\ d^3\underline{k}}{(2\pi)^3}$

or since $p = \hbar k = \dfrac{hk}{2\pi}$

$\Rightarrow$ $\#$ states within small volume of $\underline{p}$ and $\underline{x}$ is $\boxed{\dfrac{d^3\underline{x}\ d^3\underline{p}}{h^3}}$

(A result that we applied earlier to the ideal gas)

If the gas is isotropic then we need only keep track of $|\underline{p}|$

density of states in real space $= g(p)\,dp = \displaystyle\int_{\text{angles}} \frac{d^3\underline{p}}{h^3}$

$$= \frac{4\pi p^2\,dp}{h^3}\ .$$

Using a relation between p and ε eg. $\varepsilon = \dfrac{p^2}{2m}$ for non-relativistic particles

we can write this in terms of energy $g(\varepsilon)\,d\varepsilon$.

<u>Fermi-gas</u>

Let's start with a gas of fermions. The partition function is

$$\ln Z = \sum_k \ln\left(1 + e^{-\beta(\varepsilon_k - \mu)}\right)$$

assume spin $\frac{1}{2}$ particles - can be ↑ or ↓

$$= 2\int_0^\infty \frac{4\pi p^2\,dp}{h^3}\, V\, \ln\left(1 + e^{-\beta(\varepsilon_p - \mu)}\right)$$

now a continuous variable $\varepsilon(p)$

or $$\boxed{\Phi = -2k_BTV\int_0^\infty \frac{dp\,4\pi p^2}{h^3}\,\ln\left(1 + e^{-\beta(\varepsilon_p - \mu)}\right)}$$

The number of particles is $N = -\dfrac{\partial\Phi}{\partial\mu}$

$$= 2\int \frac{4\pi p^2\,dp}{h^3}\, V\cancel{k_BT}\, \frac{e^{-\beta(\varepsilon_p - \mu)}\cancel{\beta}}{1 + e^{-\beta(\varepsilon_p - \mu)}}$$

Fermi gas

We showed previously that the partition function

$$Z = \prod_k Z_k$$

where k labels the single particle states. Equivalently

$$\ln Z = \sum_k \ln Z_k$$

$$= \sum_k \ln \underbrace{\left(1 + e^{-(\varepsilon_k - \mu)/k_BT}\right)}_{\text{only two possibilities } n_k = 0 \text{ or } 1 \text{ for fermions.}}$$

For a gas of particles, we label the states by the momentum p which is a continuous variable; therefore the sum becomes an integral

$$\ln Z = \int_0^\infty dp \underbrace{\left[2 \times \frac{4\pi p^2}{h^3} \times V\right]}_{\substack{\text{density of states} \\ g(p) \\ \text{times the volume } V \\ \text{times a factor of 2} \\ \text{for spin } 1/2 \text{ particles} \\ \text{which can be} \uparrow \text{or} \downarrow}} \ln\left[1 + e^{-(\varepsilon(p)-\mu)/k_BT}\right]$$

therefore

$$\boxed{\Phi = -2k_BTV \int_0^\infty \frac{4\pi p^2 dp}{h^3} \ln\left(1 + e^{-(\varepsilon-\mu)/k_BT}\right)}$$

The number of particles is

$$N = -\left.\frac{\partial \Phi}{\partial \mu}\right|_{T,V} = 2\cancel{k_BT}V \int_0^\infty \frac{4\pi p^2 dp}{h^3} \frac{e^{-(\varepsilon-\mu)/kT}\cancel{\frac{1}{kT}}}{1 + e^{-(\varepsilon-\mu)/kT}}$$

$$\Rightarrow \boxed{N = V\int_0^\infty \frac{8\pi p^2\,dp}{h^3}\,\frac{1}{1+e^{(\varepsilon-\mu)/kT}}}$$

$$= \underbrace{V\int_0^\infty 2g(p)\,dp\,\langle n_k\rangle} = \sum_k \langle n_k\rangle$$

this makes sense- the total number of particles is a sum over the mean occupation of each level k.

This integral gives a relation between $n = \frac{N}{V}$ and μ

Usually we know n and solve to find μ.

Limits 1) classical limit $\langle n_k\rangle \ll 1 \simeq e^{-(\varepsilon-\mu)/kT}$

$$\text{Let } \varepsilon = p^2/2m \Rightarrow d\varepsilon = \frac{p\,dp}{m}$$

$$\Rightarrow n = \int_0^\infty \frac{8\pi p^2\,dp}{h^3}\,e^{-\varepsilon/kT}e^{\mu/kT}$$

$$= \frac{8\pi}{h^3}e^{\mu/kT}\int_0^\infty \sqrt{2m\varepsilon}\,m\,d\varepsilon\,e^{-\varepsilon/kT}$$

$$= \frac{8\pi}{h^3}\sqrt{2m}\,(kT)^{3/2}m\underbrace{\int_0^\infty x^{1/2}\,dx\,e^{-x}}_{\sqrt{\pi}/2}$$

$$= 2\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2}e^{\mu/kT}$$

$$= 2n_Q e^{\beta\mu}$$

$$\Rightarrow \mu = k_B T\ln\left(\frac{n}{2n_Q}\right)$$

this is our previous result except for the factor of 2 which arises from the fact that this time we considered spin $\frac{1}{2}$ particles.

2) degenerate limit $T=0$

then the occupation number is a step function

$$\frac{1}{1+e^{(\varepsilon-\mu)/kT}} = \begin{cases} 1 & \varepsilon<\mu \\ 0 & \varepsilon>\mu \end{cases}$$

All states are filled up to energy $\varepsilon=\mu$. In this limit we refer to μ as the Fermi energy $E_F = \mu(T=0)$

The corresponding Fermi momentum and wavevector are p_F and k_F. (eg. for non-relativistic particles $E_F = p_F^2/2m$).

The integral is now straightforward

$$n = \int_0^{p_F} \frac{8\pi p^2\,dp}{h^3}$$
$$= \frac{8\pi}{3}\frac{p_F^3}{h^3} = \frac{p_F^3}{3\pi^2\hbar^3}$$

or $\boxed{p_F = \hbar(3\pi^2 n)^{1/3}}$ $\boxed{k_F = (3\pi^2 n)^{1/3}}$

For non-relativistic particles $E_F = \frac{\hbar^2}{2m}k_F^2 = \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3}$.

The energy density is $E = \frac{3}{5}nE_F = \int_0^{p_F}\frac{8\pi p^2\,dp}{h^3}\,\varepsilon$

The pressure is $P = \frac{2}{5}nE_F = \frac{2}{3}E$ (we'll show this soon)

[the corresponding results for relativistic particles $E_F = p_F c$, $\varepsilon = pc$ are $E = \frac{3}{4}nE_F$ and $P = \frac{1}{4}nE_F = \frac{1}{3}E$]

You should contrast these expressions with the ideal gas results

$$P = nk_BT \qquad E = \frac{3}{2}nk_BT$$

In a degenerate gas, it is E_F and not k_BT that sets the energy scale of the particles (and therefore internal energy and pressure). The condition for the $T=0$ results to hold is therefore $\frac{k_B T}{E_F} \ll 1$. When $k_B T \ll E_F$, we are dealing with a degenerate gas.

3) Finite temperature effects in a degenerate gas may be calculated using Sommerfeld's expansion

$$\int_0^\infty \phi(\varepsilon)\, d\varepsilon \frac{1}{e^{(\varepsilon-\mu)/k_BT}+1} \simeq \int_0^\mu \phi(\varepsilon)\, d\varepsilon + \frac{\pi^2}{6}(k_BT)^2 \phi'(\mu) + \frac{7\pi^4}{360}(k_BT)^4 \phi'''(\mu) + \dots$$

valid for $k_BT \ll \mu$. Note that only odd derivatives appear $\phi'(\varepsilon) = \frac{d\phi}{d\varepsilon}$, $\phi''' = \frac{d^3\phi}{d\varepsilon^3}$ etc.

Examples

— chemical potential is determined by

$$n = \int_0^\infty \frac{8\pi p^2 dp}{h^3} \frac{1}{e^{(\varepsilon-\mu)/kT}+1}$$

For NR particles, $\varepsilon = p^2/2m$, we find

$$n = \frac{8\pi m^{3/2}\sqrt{2}}{h^3} \int \frac{\varepsilon^{1/2}\, d\varepsilon}{e^{(\varepsilon-\mu)/k_BT}+1}$$

Now apply the expansion

$$\underbrace{\frac{h^3 n}{8\pi m^{3/2}\sqrt{2}}}_{\substack{\text{this is} \\ \frac{2}{3}E_F^{3/2}}} = \underbrace{\int_0^\mu \varepsilon^{1/2}\, d\varepsilon}_{\frac{2}{3}\mu^{3/2}} + \frac{\pi^2}{6}(k_BT)^2 \frac{1}{2}\mu^{-1/2}$$

$$\Rightarrow \quad E_F^{3/2} = \mu^{3/2} + \frac{\pi^2}{8}\frac{(k_BT)^2}{\mu^{1/2}}$$

$$\Rightarrow \quad \mu = E_F\left(1 - \frac{\pi^2}{8}\left(\frac{kT}{\mu}\right)^2\right)^{2/3}$$

$$\boxed{\mu \approx E_F\left(1 - \frac{\pi^2}{12}\left(\frac{k_BT}{E_F}\right)^2\right)}$$

to the same order in $\frac{kT}{E_F}$ we can replace μ by E_F here

μ decreases with T

— heat capacity

We can similarly use the Sommerfeld expansion to calculate the internal energy density $\int \frac{8\pi p^2\,dp}{h^3}\frac{1}{e^{(\varepsilon-\mu)/kT}+1}\varepsilon$

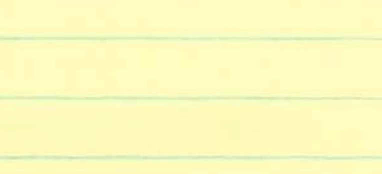
Differentiating with respect to T gives

$$\boxed{C_V = \frac{\pi^2}{2}k_B\left(\frac{k_BT}{E_F}\right)}$$

heat capacity per particle

(see HW 3 Q3)

This is much smaller than the ideal gas heat capacity of $\simeq k_B$ per particle. What's happening is that only a small fraction of the particles ($\simeq k_BT/E_F$) near the Fermi energy are available to accept thermal energy.



Pressure of a fermi gas

We can derive the pressure in two ways:

1. "Kinetic theory" approach

eg. Look at the flux of x-momentum across unit area in the y-z plane.

$$\text{momentum flux} = \int \underset{\substack{\uparrow \\ (\text{spin } \frac{1}{2})}}{2} \frac{p^2\, dp\, d\phi \sin\theta\, d\theta}{h^3} \underset{\substack{\uparrow \\ \text{x-momentum}}}{(p\cos\theta)} \underset{\substack{\uparrow \\ \text{x-velocity}}}{(v\cos\theta)} \underset{\substack{\uparrow \\ \#\text{ particles} \\ \text{with momentum } p}}{\langle n \rangle}$$

$$= \int \frac{p^2\, dp}{h^3}\, 4\pi \int_0^1 d\mu \;\; p v \langle n \rangle \mu^2$$

$$= \frac{4\pi}{3h^3} \int p^2\, dp\; v p \frac{1}{e^{(\varepsilon-\mu)/kT}+1}$$

If there was a wall at that location, the particles would reflect and transmit 2x their momentum to the wall.

$$\Rightarrow \quad \text{pressure} = 2 \times (\text{momentum flux})$$

$$\Rightarrow \quad \boxed{P = \frac{1}{3} \int \frac{8\pi p^2\, dp}{h^3} \frac{p v}{1+e^{(\varepsilon-\mu)/k_BT}}} \quad \text{—}(*)$$

2. from the free energy

$$\Phi = -k_BTV \int \frac{8\pi p^2\, dp}{h^3} \ln\left(1+e^{-(\varepsilon-\mu)/k_BT}\right)$$

$$\Rightarrow \quad P = -\left.\frac{\partial \Phi}{\partial V}\right|_{T,\mu} = +k_BT \int \frac{8\pi p^2\, dp}{h^3} \ln\left(1+e^{-\beta(\varepsilon-\mu)}\right)$$

Now integrate by parts

$$\frac{1}{3}\frac{d}{dp}\left(p^3\ln\left(1+e^{-\beta(\varepsilon-\mu)}\right)\right) = p^2\ln\left(1+e^{-\beta(\varepsilon-\mu)}\right) + \frac{p^3}{3}\frac{1}{1+e^{-\beta(\varepsilon-\mu)}}\,e^{-\beta(\varepsilon-\mu)}(-\beta)\frac{d\varepsilon}{dp}$$

The surface term vanishes

$$\Rightarrow \quad P = \cancel{k_BT}\int\frac{8\pi p^3\,dp}{3h^3}\,\frac{1}{1+e^{\beta(\varepsilon-\mu)}}\,\cancel{\beta}\,\frac{d\varepsilon}{dp}$$

A useful result is $\frac{d\varepsilon}{dp} = v$ (true for arbitrary $\frac{v}{c}$)

[to see this recall that $\varepsilon^2 = p^2c^2+(mc^2)^2$ $\qquad \left(\gamma^2=\frac{1}{1-\beta^2}\right)$

$$\Rightarrow \quad \frac{d\varepsilon}{dp} = \frac{pc^2}{\varepsilon} \qquad (\beta = v/c)$$

but $\left.\begin{matrix} p=\gamma m v \\ \varepsilon=\gamma mc^2\end{matrix}\right\}$ $v = \frac{pc^2}{\varepsilon} = \frac{d\varepsilon}{dp}$.]

$\Rightarrow$ same result as before (eq. *).

<u>Two limits</u>: 1) $k_BT \gg \mu$

$$P = \frac{1}{3}\int\frac{8\pi p^2\,dp}{h^3}\,v\,p\,e^{-(\varepsilon-\mu)/kT}$$

$$= \frac{8\pi}{3h^3}(2m)^{3/2}\underbrace{\int_0^\infty \varepsilon^{3/2}e^{-\beta\varepsilon}\,d\varepsilon}_{\frac{3}{4}\sqrt{\pi}\,(k_BT)^{5/2}}\;e^{\mu/kT}$$

$$= k_BT\,\underbrace{2\left(\frac{2\pi m k_BT}{h^2}\right)^{3/2}e^{\mu/kT}}_{2n_Q e^{\mu/kT} = n} = n k_BT \checkmark$$

2) $k_B T \ll \mu$ degenerate

$$p = \frac{1}{3}\int_0^{p_F} \frac{8\pi p^2 \, dp}{h^3} \, p v = \frac{8\pi}{3h^3} (2m)^{3/2} \underbrace{\int_0^{E_F} \varepsilon^{3/2} \, d\varepsilon}_{\frac{2}{5} E_F^{5/2}}$$

$$= \frac{2}{5} E_F \left(\frac{(2m E_F)^{3/2}}{3\pi^2 \hbar^3} \right)$$

$$= \underline{\underline{\frac{2}{5} n E_F}} \, .$$

Simple estimate of C_V for a Fermi gas

First, calculate the density of states at the Fermi energy.

We know $$g(p)\,dp = \frac{2\times 4\pi p^2\,dp}{h^3}$$

$$= \frac{3}{2}\times\frac{p^2\,dp}{\hbar^3 3\pi^2}\times 2$$

but $p_F^3 = 3\pi^2\hbar^3 n \Rightarrow dp\, g(p) = \frac{3}{2}\, n\, \frac{2p^2}{p_F^3}\,dp$

Now change variables to $\varepsilon = p^2/2m$: $\quad p = (2m\varepsilon)^{1/2}$

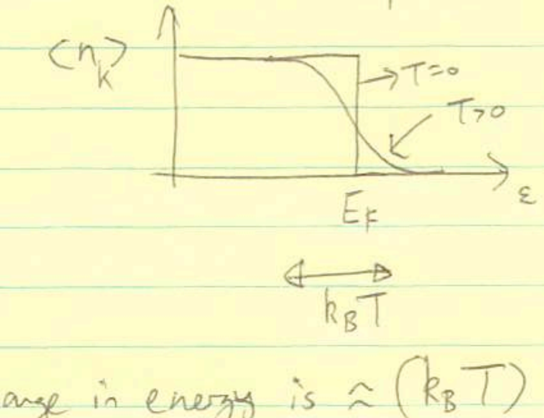
$$dp = \frac{1}{2}\frac{(2m)^{1/2}}{\varepsilon^{1/2}}\,d\varepsilon$$

$$\Rightarrow g(\varepsilon)\,d\varepsilon = \frac{3}{2}\, n\, \frac{1}{p_F^3}\,\frac{(2m)^{1/2}}{\varepsilon^{1/2}}\,d\varepsilon\; 2m\varepsilon$$

$$= \frac{3}{2}\, n\, \frac{\varepsilon^{1/2}\,d\varepsilon}{E_F^{3/2}}$$

$$\Rightarrow \quad \underline{g(E_F) = \frac{3}{2}\frac{n}{E_F}}$$

Now consider adding energy to the Fermi gas to take it from $T=0$ to finite T. The occupation number looks like



the Fermi surface is broadened it has a width $\simeq kT$.

$$\text{The change in energy is } \simeq (k_BT)\times\Big(\overbrace{g(E_F)\,k_BT}^{\text{\# of particles that can accept } kT \text{ worth of energy}}\Big)$$

$$\simeq \frac{3}{2}\, n\, k_BT\left(\frac{k_BT}{E_F}\right) \Rightarrow C_V \simeq 3\, n\, k_B\left(\frac{k_BT}{E_F}\right)$$

We get the correct scalings, but the prefactor is not quite right. We must do the integral over the FD distribution (HW3 Q3) to get the correct prefactor.

But this brings out the physics — only a fraction $\left(\frac{kT}{E_F}\right)$ of particles are able to accept thermal energy.

Example — Electrons in metals

In a metal, the overlap of the wavefunctions of the outermost electrons in each atom leads to states for these conduction or valence electrons that extend throughout the crystal lattice. Ignoring interactions between the electrons, we think of the electrons as independently moving in the background potential of the ions (nucleus + core electrons). For a constant background potential, the states are plane wave states $\psi \propto e^{i\mathbf{k}\cdot\mathbf{r}}$ and we can apply our results for the Fermi gas.

To see that the electrons are ~~g~~ degenerate, calculate

$$\frac{kT}{E_F} = \frac{kT}{\hbar^2(3\pi^2 n_e)^{2/3}/2m_e}$$

Let's estimate n_e for eg. Copper

$$n_e = \underset{\substack{\uparrow\\ \text{number of}\\ \text{valence}\\ \text{electrons}}}{Z} \; \frac{\rho}{\underset{\substack{\uparrow\\ \text{atomic}\\ \text{mass}}}{A} m_p}$$

for Copper $\rho \simeq 9\,\mathrm{g/cm^3}$
$Z = 1$
$A = 63$

$$\frac{kT}{E_F} = \frac{2m_e kT}{\hbar^2(3\pi^2)^{2/3}}\left(\frac{Am_p}{Z\rho}\right)^{2/3} = 4\times10^{-3}\left(\frac{T}{300\,\mathrm{K}}\right)\left(\frac{A/Z}{63}\right)^{2/3}\left(\frac{9\,\mathrm{g/cm^3}}{\rho}\right)^{2/3}$$

So we see that $kT \ll E_F$: the electrons in the metal are indeed degenerate.

For $T = 300\,\mathrm{K}$, $k_B T = 25\,\mathrm{meV}$

$$\Rightarrow \underline{E_F \simeq 6\,\mathrm{eV}}$$

Let's check how the wavelength of the electrons at the Fermi-level compares with the lattice spacing.

$$\lambda_F = \frac{2\pi}{k_F}$$

lattice spacing $\frac{4\pi a^3}{3} \cdot n_i = 1$ $\left(\begin{array}{l}\text{Sphere of radius } a\\ \text{contains one ion}\end{array}\right)$

$$\Rightarrow \quad \frac{\lambda_F}{a} = \frac{2\pi}{(3\pi^2 n_e)^{1/3}} \cdot \left(\frac{4\pi n_i}{3}\right)^{1/3} = \left(\frac{32\pi^2}{9Z}\right)^{1/3} = \underline{\underline{\frac{3.3}{Z^{1/3}}}}$$

But $\frac{n_i}{n_e} = \frac{1}{Z}$

So we see that for the most energetic electrons $\lambda_F \sim a$ so the assumption of a constant potential is probably not such a good one.

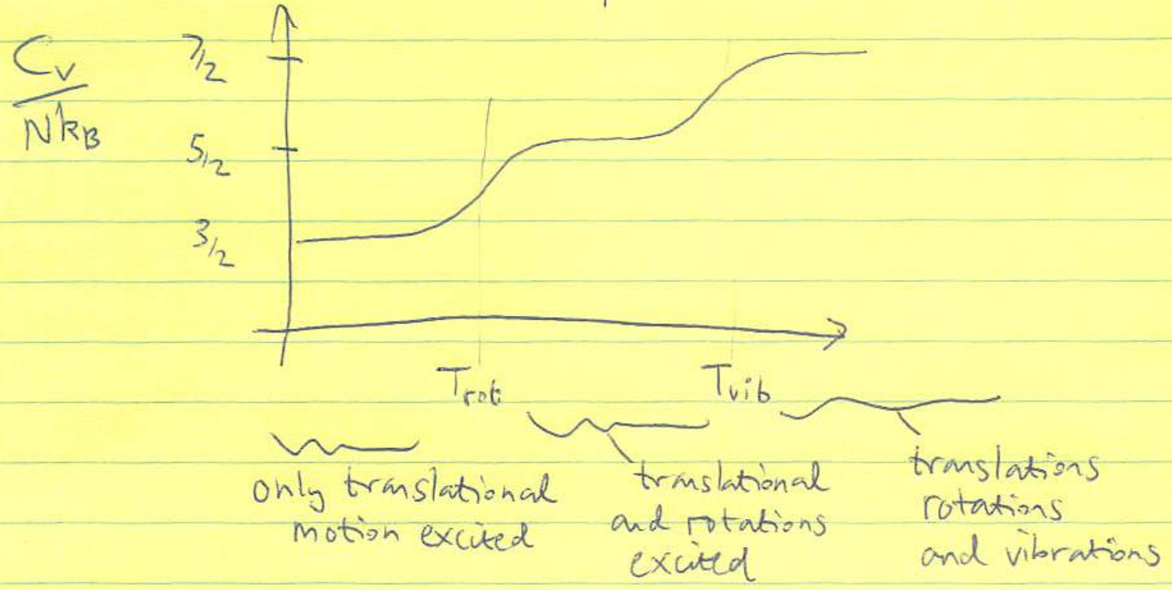
Even so, the free electron model of metals successfully explains a range of metallic properties eg. thermal and electrical conductivities, and the <u>electronic contribution to the heat capacity</u>. To see this, we must go to very low temperatures, where the ion heat capacity becomes very small ($\propto T^3$ we'll see this very soon). Below a few degrees K, a linear dependence $C_V = \gamma T$ is observed as expected from our result

$$C_V = \frac{\pi^2}{2} n k_B \left(\frac{k_B T}{\epsilon_F}\right) \propto T$$

Historically, the Drude model (1900) treated the free electrons classically and was successful in predicting conductivities, but also predicted a large electronic heat capacity ($\frac{3}{2} k_B$ per particle) that was not seen in experiments! Sommerfeld applied the Pauli exclusion principle to the electrons, and obtained the (kT/E_F) suppression of heat capacity nicely explaining the experimental data.

Classical Equipartition Theorem

We saw in HW3 Q1 that the specific heat of a diatomic molecule varies with temperature as

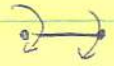


The plateaus at $3/2$, $5/2$, $7/2$ k_B are examples of the classical equipartition theorem, which states that in a classical system each degree of freedom has an energy $\frac{1}{2}k_BT$ in thermal equilibrium.

By "degree of freedom," we mean each quadratic term in the energy of a particle

eg. translational 3 d.o.f. $\quad \varepsilon = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}$

rotational $\quad \varepsilon = \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2$

(only two d.o.f. here because the third moment of inertia vanishes for a linear molecule )

vibration 2 d.o.f. $\quad \varepsilon = \frac{1}{2}kx^2 + \frac{1}{2}mv^2$ 

Quantization of energy levels results in "freezing out" of d.o.f. at low temperature, when $kT \ll \hbar\omega$ the motion cannot be excited.

The theorem is straightforward to prove with the canonical ensemble. For a system with n d.o.f. we can write

$$\varepsilon_\nu = \sum_{i=1}^{n} \alpha_i x_i^2$$

where x_i could represent momentum, position, angular velocity etc.

Then the single particle partition function is

$$Q_1 = \int dx_1\, dx_2 \dots dx_n\, e^{-\sum_{i=1}^{n} \alpha_i x_i^2/kT}$$

$$= \prod_{i=1}^{n} \int dx_i\, e^{-\alpha_i x_i^2/kT}$$

$$\propto (kT)^{n/2} \propto \beta^{-n/2}$$

$$\Rightarrow E_1 = -\frac{\partial \ln Q_1}{\partial \beta} = \frac{n}{2} k_B T.$$

Each quadratic contribution to the particle energy gives an additional $\frac{1}{2} k_B T$ of energy in thermal equilibrium.

Zero-temperature stars : white dwarfs and neutron stars

A star is a ball of gas supported against gravity by gas pressure. Its structure is determined by the equation of hydrostatic balance

$$\underbrace{\frac{dP}{dr}}_{\substack{\text{outwards}\\ \text{pressure}\\ \text{gradient}}} = \underbrace{-\rho\frac{Gm}{r^2}}_{\substack{\text{inwards}\\ \text{pull of gravity}}}$$

m= mass within radius r
$= \int 4\pi r^2 \rho \, dr$

The pressure and density decrease smoothly from their central values to zero at the stellar surface.

Therefore, we can estimate $\frac{dP}{dr} \simeq \frac{P_c}{R}$ ← central pressure

↖ radius of star

$$\Rightarrow \quad \frac{P_c}{R} \simeq \left(\frac{M}{R^3}\right)\left(\frac{GM}{R^2}\right)$$

$$\underline{\underline{P_c \simeq \frac{GM^2}{R^4}}}$$

This is the "back of the envelope" estimate for the central pressure of a star with mass M and radius R.

For the Sun, the equation of state is ideal gas

$$P = n k_B T$$
$$= \frac{\rho k_B T}{m_p} \quad \text{for pure hydrogen}$$

(the Sun is actually 70% H 30% He)

Therefore $\quad P_c \simeq \frac{M}{R^3} \cdot \frac{k_B T_c}{m_p} = \frac{GM^2}{R^4}$

$$\Rightarrow \quad \boxed{\frac{k_B T_c}{m_p} \simeq \frac{GM}{R}}$$

Input numbers for the Sun:

$M = 2 \times 10^{33}\,g$

$R = 7 \times 10^{10}\,cm$

$\Rightarrow \underline{T_c = 2 \times 10^{7}\,K}$

(detailed model gives $1.5 \times 10^{7}\,K$).

The temperature is maintained by nuclear burning

$4H \rightarrow {}^{4}He$

over the $\sim 10^9$ yrs of the Sun's life.

Later, shorter-lived stages involve nuclear burning to heavier elements. $3He \rightarrow C$ etc.

But eventually the <u>fuel runs out</u>! The star collapses to form a compact object.

Initial mass	Remnant	Pressure support
$\lesssim 8\,M_\odot$	White dwarf	electron degeneracy
8-$25\,M_\odot$	Neutron star	neutron degeneracy.
$\gtrsim 25\,M_\odot$	Black hole	none!

Neutron stars and white dwarfs are a "zero-temperature stars" in the sense that $k_B T \ll E_F$

the particles are degenerate.
providing the pressure

First let's calculate the radius of a <u>white dwarf</u> star, which is held up by degenerate electrons.

Then $$P_c = \frac{2}{5} n_e E_F = \frac{2}{5} n_e \frac{\hbar^2}{2m_e}(3\pi^2 n_e)^{2/3}$$
$$= \frac{\hbar^2}{5m_e}(3\pi^2)^{2/3} n_e^{5/3}$$

A white dwarf is typically made up of carbon and oxygen, fully ionized. Therefore

$$n_e \simeq \underbrace{\frac{\rho}{12m_p}}_{\substack{\text{number density of}\\ \text{nuclei for pure } ^{12}\text{C}}} \times 6 \;\; \nwarrow \text{6 electrons per nucleus} \qquad \simeq \rho/2m_p.$$

$$\Rightarrow \boxed{P_c = \frac{\hbar^2}{5m_e}\frac{(3\pi^2)^{2/3}}{(2m_p)^{5/3}} \rho_c^{5/3}} \quad \propto \quad \frac{M^{5/3}}{R^5}$$

But we saw earlier that $P_c \simeq \frac{GM^2}{R^4}$ for hydrostatic balance

$$\Rightarrow \quad \frac{M^2}{R^4} \propto \frac{M^{5/3}}{R^5}$$
$$\Rightarrow \quad \boxed{R \propto M^{-1/3}}$$

The prefactor is
$$R = \underbrace{8.7 \times 10^8 \text{ cm}}_{8700\text{ km}} \left(\frac{M}{M_\odot}\right)^{-1/3}$$

Recall that the radius of the Earth is 6400 km so this is a star with the mass of the Sun (300,000 $M_\oplus$) but the radius of the Earth!

The mean density is $\simeq \frac{10^{33}\text{ g}}{(10^9\text{ cm})^3} \simeq 10^6 \text{ g/cm}^3$

(Compare to water $\rho_{water} = 1\text{ g/cm}^3$)

Another way to write the hydrostatic balance is

$$\frac{P_c}{R} = \frac{2}{5}\frac{n_e E_F}{R} \sim \frac{\rho E_F}{m_p R} \sim \rho \frac{GM}{R^2}$$

$$\Rightarrow \quad \boxed{E_F \sim \frac{GMm_p}{R}}$$

Since $R \propto M^{-1/3}$ we see that $E_F \propto M^{4/3}$
grows with mass.

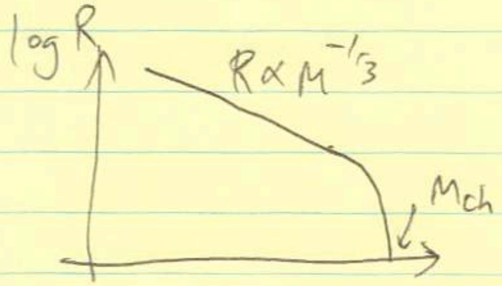
Eventually there is a mass where the electrons become relativistic.
Then $E_F = p_F c \propto \rho^{1/3}$

$$\Rightarrow \quad P \propto \rho^{4/3}$$

$$\Rightarrow \quad \frac{GM^2}{R^4} \propto \left(\frac{M}{R^3}\right)^{4/3}$$

the radius drops out of the equation!
This defines a characteristic mass, the Chandrasekhar mass

$$M_{ch} = 1.45 M_\odot$$



As $M \to M_{ch}$, the central density exceeds $\rho \simeq 10^9 \, g/cm^3$

This leads to thermonuclear ignition of carbon $^{12}C + {}^{12}C \to (\;\;)$

The energy release is explosive
→ Type Ia Supernova.

Neutron stars

Massive stars burn their fuel all the way to heavy elements — this produces an iron core with $M \simeq M_{Ch}$. The core collapses to form a neutron star (and the outer layers are thrown off in a supernova explosion).

The final object must have $E_F \approx \frac{GMm_p}{R}$

and $E_F = \frac{p_F^2}{2m} \propto \frac{1}{m}$

↖ mass of the degenerate particle

So if neutrons and protons provide the degeneracy pressure we expect the radius to be $\frac{m_p}{m_e} \simeq 2000$ times smaller than a white dwarf.

$$\Rightarrow \quad R_{NS} \simeq \frac{10^4\ \text{km}}{2000} = 5\,\text{km}!$$

This is about right — detailed models give $R \simeq 10$-15 km.
These are <u>extreme objects</u>!

1) The density $\frac{M}{\frac{4\pi}{3}R^3} \gtrsim 10^{14}\ \text{g/cm}^3$

and the central density is $> 10^{15}\ \text{g/cm}^3$.

— this exceeds the density of an atomic nucleus!

$$\left(\rho_{\text{nucleus}} \sim \frac{m_p}{(1\,\text{fm})^3} \sim \frac{10^{-27}\,\text{kg}}{(10^{-15}\,\text{m})^3} \sim 10^{18}\,\frac{\text{kg}}{\text{m}^3} \sim 10^{15}\,\frac{\text{g}}{\text{cm}^3}\right)$$

2) The Fermi energy is 100's of MeV $\Rightarrow$ produce exotic particles eg. pions

3) General relativity is important also $\frac{GM}{Rc^2} \simeq 0{\cdot}3$.

Photon gas

Electromagnetic waves in a box

density of states $\frac{d^3k}{(2\pi)^3}$

isotropic gas $\frac{4\pi k^2 dk}{(2\pi)^3} \times \underset{\substack{\uparrow \\ \text{2 polarizations}}}{2} = \frac{k^2 dk}{\pi^2}$

in terms of frequency $\omega = ck$,

density of states is $g(\omega)\, d\omega = \frac{\omega^2\, d\omega}{\pi^2 c^3}$

(per unit spatial volume)

For a classical gas in equilibrium at temperature T, we expect each mode to have energy $k_B T$.

$$\Rightarrow \text{energy density} \quad \mathcal{E}(\omega)\, d\omega = \frac{k_B T\, \omega^2\, d\omega}{\pi^2 c^3}$$

This is the <u>Rayleigh-Jeans</u> formula. It fits well at low frequency, but at high frequencies it predicts that the energy density continues to grow — leading to infinite total energy

the "UV catastrophe"

To resolve this, Planck proposed that the energy in each mode was quantized with energy $n\hbar\omega$.

Just like the quantum harmonic oscillator $Q = \sum_0^\infty e^{-n\frac{\hbar\omega}{kT}}$

$$= \frac{1}{1 - e^{-\hbar\omega\beta}}$$

$$E = -\frac{\partial \ln Q}{\partial \beta} = \frac{\hbar\omega}{e^{\hbar\omega/kT}-1}.$$ average energy per mode.

For frequencies $\omega \gg \frac{kT}{\hbar}$ the energy is $\hbar\omega \underbrace{e^{-\hbar\omega/kT}}_{\text{exponentially suppressed}}$

Solving the UV catastrophe.

The energy density is $$\boxed{\varepsilon(\omega)\,d\omega = \frac{\omega^2}{\pi^2 c^3}\frac{\hbar\omega}{e^{\hbar\omega/kT}-1}\,d\omega}$$

the "<u>Planck distribution</u>"

Limits: $\hbar\omega/kT \ll 1$ $\quad \varepsilon(\omega) = kT\,\frac{\omega^2}{\pi^2 c^3}$ <u>Rayleigh-Jeans</u>

$\hbar\omega/kT \gg 1$ $\quad \varepsilon(\omega) = \frac{\omega^2}{\pi^2 c^3}\,\hbar\omega\, e^{-\hbar\omega/kT}$ <u>Wien's law</u>

The total energy density is
$$U = \frac{E}{V} = \int_0^\infty \varepsilon(\omega)\,d\omega = \frac{(kT)^4}{\hbar^3}\frac{1}{\pi^2 c^3}\underbrace{\int_0^\infty \frac{dx\,x^3}{e^x-1}}_{\frac{\pi^4}{15}}$$

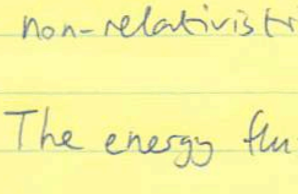
$$\Rightarrow \qquad = \underbrace{\left(\frac{\pi^2 k_B^4}{15(\hbar c)^3}\right)}_{\equiv\text{"a" radiation constant}} T^4 = \boxed{aT^4 = \frac{E}{V}}$$

Recall the Bose-Einstein distribution $\frac{1}{e^{(\varepsilon-\mu)/kT}-1}$. We see that the radiation can be thought of as a gas of bosons (photons) with $\mu = 0$. Physically, this comes about because the photon number is not conserved, for example as they are absorbed and emitted by the walls of the container. Contrast this with a gas of particles, eg. ^{4}He, which has a fixed number of particles in it.

One can show that the pressure of the photon gas is $P = \frac{1}{3}\left(\frac{E}{V}\right) = \frac{u}{3}$

[Standard result for relativistic particles; recall that a gas of non-relativistic particles has $P = {}^2\!/_3\,(E/V)$.]

The energy flux across unit area is



$$\int \underbrace{uc\cos\theta}_{\text{energy per second per unit area in x direction}} \underbrace{\left(\frac{\sin\theta\, d\theta\, 2\pi}{4\pi}\right)}_{\text{fraction of photons travelling at angle }\theta}$$

$\mu = \cos\theta$

$$= uc\,\frac{1}{2}\int_0^1 \mu\, d\mu = \frac{1}{4}cu = \frac{1}{4}acT^4$$

$$\Rightarrow \boxed{\text{flux} = \frac{1}{4}acT^4 = \sigma_{SB}T^4}$$

σ_{SB} Stefan-Boltzmann constant

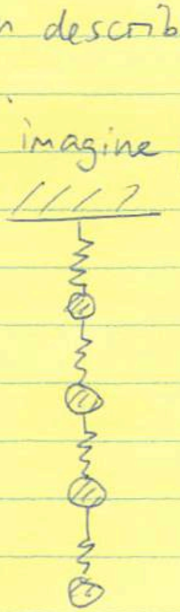
A "blackbody" is an object that absorbs all radiation falling on it. In equilibrium with a photon gas at temperature T, it must emit radiation to replace the radiation absorbed

$\Rightarrow$ flux of radiation from the surface of a blackbody $= \sigma T^4$

Phonons and the lattice heat capacity.

Now think about the ions making up the crystal lattice. We can describe their motions in terms of 3N oscillation modes, or "normal modes"

eg.
imagine playing with a set of masses connected by springs



low frequency modes — long range correlations — adjacent masses moving together

↓

high frequency modes - alternate masses move oppositely.

In quantum mechanics, these oscillation modes have quantized energy levels with $\Delta E = \hbar\omega$

$$\Rightarrow \text{ each one has } \quad \langle \varepsilon \rangle = \frac{\hbar\omega}{e^{\hbar\omega/kT}-1}$$

when in thermal eqm at temperature T.

The idea is to calculate the thermodynamic properties of the lattice, we need to sum over the modes.

$$\text{eg. energy } \quad E = \underset{\nearrow}{\int g(\omega)\, d\omega}\, \frac{\hbar\omega}{e^{\hbar\omega/kT}-1}$$

we need the <u>density of states</u> $g(\omega)$

In general, $g(\omega)$ is a complicated function depending on the details of the lattice - arrangement of the ions and the potential. We'll discuss two simple approximations.

1) Einstein solid assume all modes have the same frequency ω_E

$$\Rightarrow \quad E = \frac{3N\hbar\omega_E}{e^{\hbar\omega_E/kT} - 1}$$

At high temperature $kT \gg \hbar\omega_E$ $E \simeq 3Nk_BT$

and $\boxed{C_V = 3Nk_B}$ ✓

"Dulong-Petit" law"

But At low T, predict exponential suppression of the heat capacity

$$C_V \to 3Nk_B\left(\frac{\hbar\omega}{kT}\right)^2 e^{-\hbar\omega/kT}$$

which disagrees with experiments. The problem is that all modes are suppressed at low T, but there are always some long wavelength modes available at low frequency. We need to include those.

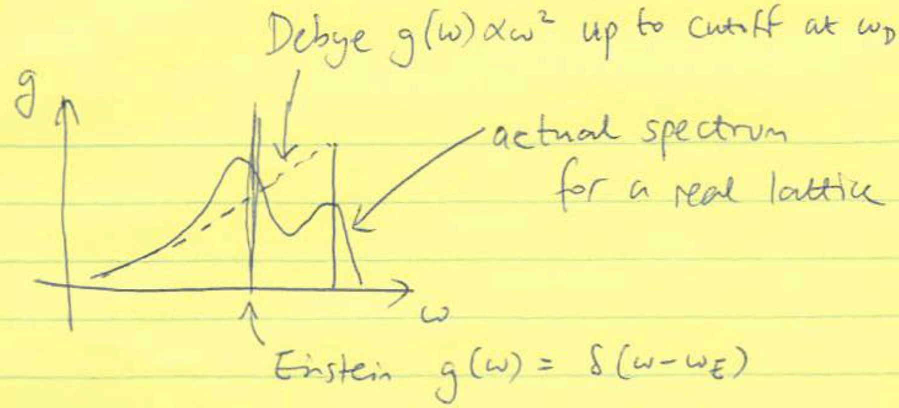
2) Debye solid

An elastic solid supports transverse (2 polarizations) and longitudinal waves, with velocities c_T, c_L. At low frequency / long wavelengths we can treat the modes as elastic waves without worrying about details of the lattice structure. (The lattice appears as a continuum to the elastic mode).

$$\text{Then } d\omega\, g(\omega) = \frac{4\pi k^2 dk}{(2\pi)^3} = \frac{\omega^2 d\omega}{2\pi^2}\left(\frac{1}{c_L^3} + \frac{2}{c_T^3}\right)$$

$$= \underline{\alpha\,\omega^2\, d\omega} \qquad (\alpha = \text{constant})$$

The idea is to assume $g(\omega)$ continues to have this form at higher frequencies.



But we must cutoff the Debye spectrum to ensure the correct number of modes:

$$\int_0^{\omega_D} \alpha \omega^2 \, d\omega = 3N \quad \Rightarrow \quad \omega_D = \left(\frac{9N}{\alpha}\right)^{1/3}$$

"Debye frequency"

The energy is

$$E = \int_0^{\omega_D} \alpha \omega^2 \, d\omega \, \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}$$

$$= 3NkT \left[3\left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} \frac{x^3 \, dx}{e^x - 1} \right]$$

where Θ_D = "Debye temperature" defined such that $k\Theta_D = \hbar\omega_D$.

For $T \gg \Theta_D$ $\quad C_V = 3Nk_B$ as expected

$T \ll \Theta_D$

$$E = \frac{3\pi^4}{5} \frac{NkT^4}{\Theta_D^3}$$

$$\Rightarrow C_V = \frac{12\pi^4}{5} Nk_B \left(\frac{T}{\Theta_D}\right)^3 \quad \underline{\underline{\propto T^3}}$$

agrees with experiment

(23.31)

ɪre, while Fermi tem-
ıe temperature T_0 is
degrees Kelvin. This
ɔnly observed at low

L DENSITY)

cific heat (23.15), are

(23.32)

grals, by introducing
) that $g(\omega)\ d\omega$ is the
ange between ω and
ŋ, the sum or integral

e extent that results
of $\Theta_D(T)$ rather than
tion back to specific
is is shown in Figure
le 23.2. In Table 23.3
were determined by
5) at the point where

ebye approximation (in
(From J. de Launay,

integrand in (23.26)
his is to be expected,
$_D$.) Thus the Debye
ıtions as the Fermi
re a measure of the
ıtum statistics must
istical mechanics is
ways well below T_F,
ɔ both classical and

---

are known as van Hove singularities.[18] A typical density of levels displaying these singularities is shown in Figure 23.6, and a concrete illustration of how the singularities arise in the linear chain is given in Problem 3.

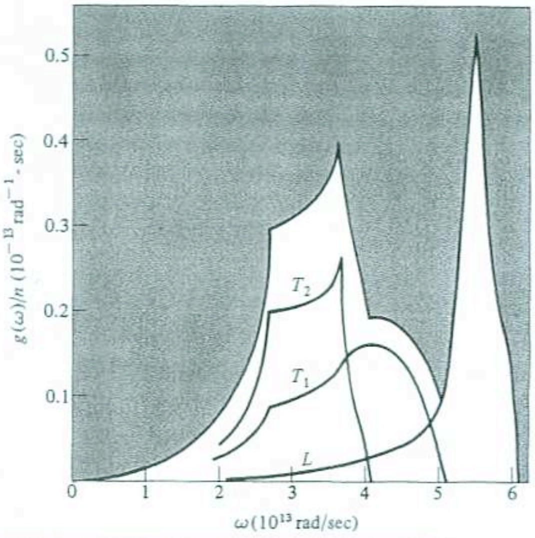


Figure 23.6
Phonon density of levels in aluminum, as deduced from neutron scattering data (Chapter 24). The highest curve is the full density of levels. Separate level densities for the three branches are also shown. (After R. Stedman, L. Almqvist, and G. Nilsson, *Phys. Rev.* **162**, 549 (1967).)

Table 23.3
DEBYE TEMPERATURES FOR SELECTED ELEMENTS[a]

ELEMENT	Θ_D (K)	ELEMENT	Θ_D (K)
Li	400	A	85
Na	150	Ne	63
K	100		
		Cu	315
Be	1000	Ag	215
Mg	318	Au	170
Ca	230		
		Zn	234
B	1250	Cd	120
Al	394	Hg	100
Ga	240		
In	129	Cr	460
Tl	96	Mo	380
		W	310
C (diamond)	1860	Mn	400
Si	625	Fe	420
Ge	360	Co	385
Sn (grey)	260	Ni	375
Sn (white)	170	Pd	275
Pb	88	Pt	230
As	285	La	132
Sb	200	Gd	152
Bi	120	Pr	74

[a] The temperatures were determined by fitting the observed specific heats c_v to the Debye formula (23.26) at the point where $c_v = 3nk_B/2$. Source: J. de Launay, *Solid State Physics*, vol. 2, F. Seitz and D. Turnbull, eds., Academic Press, New York, 1956.

from Ashcroft & Mermin "Solid State Physics"

Bose-Einstein condensation

Now consider a gas of N bosons, where N is fixed. Following the same approach as for fermions, we write

$$N = \int_0^\infty \frac{g(\varepsilon)\, d\varepsilon}{e^{(\varepsilon-\mu)/kT}-1} \qquad \text{which determines } \mu(N,T) \qquad (*)$$

For a classical gas, we know that μ is large and negative, $\mu = kT \ln\left(\frac{n}{n_Q}\right)$.

As density increases, μ approaches zero from below. It can never exceed zero (the energy of the lowest energy state) because as it approaches zero, the occupation number of the $\varepsilon = 0$ state diverges

$$\langle n(\varepsilon=0)\rangle = \frac{1}{e^{-\mu/kT}-1} \simeq \frac{kT}{-\mu} \to \infty \qquad \text{"Bose-Einstein Condensation"}$$

We soon see that there is a problem with our integral $(*)$ because it converges to a finite value when $\mu = 0$

$$N = \int_0^\infty \frac{4\pi p^2\, dp}{(2\pi)^3 \hbar^3}\, \underset{\overset{\nearrow}{\text{spin}}}{g_s} V \frac{1}{e^{(\varepsilon-\mu)/kT}-1}$$

$$= \frac{g_s V m^{3/2}}{\sqrt{2}\,\pi^2 \hbar^3} \int_0^\infty \frac{d\varepsilon\, \varepsilon^{1/2}}{e^{(\varepsilon-\mu)/kT}-1}$$

$$= \underbrace{g_s V \left(\frac{2\pi m k T}{h^2}\right)^{3/2}}_{\text{this is } g_s V n_Q} \underbrace{\frac{2}{\sqrt{\pi}} \int_0^\infty \frac{dx\, \sqrt{x}}{e^{x-\alpha}-1}}_{\substack{\text{when } \mu = 0 \text{ this is} \\ \zeta(3/2) = 2.612}} \qquad \left(\alpha = \frac{\mu}{kT}\right)$$

$$\Rightarrow \quad \underline{N(\mu=0) = 2.612\, g_s V n_Q}$$

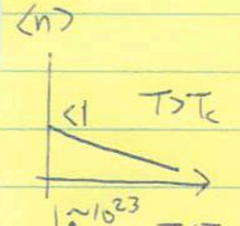
But what if I have more particles than this value? They go into the ground-state as $\mu \to 0$ as we discussed. The problem is that the integral does not correctly count the ground state! The continuous approximation for the density of states breaks down when the ground state starts to become occupied by a large number of particles. In the integral, the ground state gets zero weight because of the $\sqrt{\varepsilon}$ factor.

To fix this, we can include the ground state separately:

$$N = \underbrace{\frac{g_s}{e^{-\alpha}-1}}_{\substack{\#\text{particles in}\\ \text{grond state}\\ N_0}} + \underbrace{g_s V n_Q F_{3/2}(\alpha)}_{\substack{\#\text{particles in the}\\ \text{excited states}\\ N_e}}$$

$$\left[\text{where } F_{3/2}(\alpha) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{x}\,dx}{e^{x-\alpha}-1}\right]$$

Bose-Einstein condensation occurs for $T < T_c$



where T_c is given by $N = g_s V n_Q F_{3/2}(0)$

$$\Rightarrow \boxed{k_B T_c = \frac{\hbar^2}{2\pi m}\left[\frac{N}{g_s V \zeta(3/2)}\right]^{2/3}} \quad \text{Critical temperature}$$

For $T > T_c$ $N_0 \ll N_e$ and we can write $N = g_s V n_Q F_{3/2}(\alpha)$

For $T < T_c$, N_0 becomes significant and $N \to N_0$ as $T \to 0$.

macroscopically large!

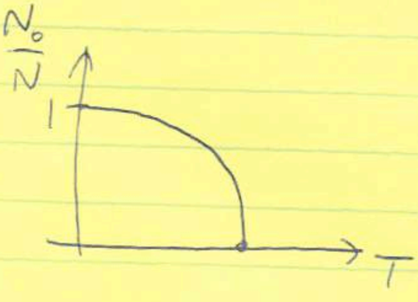
To see how N_0 grows as $T \to 0$ note that for $T < T_c$

$$N_e = 2.612\, g_s V n_Q \propto T^{3/2}$$

and at $T = T_c$ $N_e = N$

$$\Rightarrow \quad \frac{N_e}{N} = \left(\frac{T}{T_c}\right)^{3/2}$$

$$\Rightarrow \quad \boxed{\frac{N_0}{N} = 1 - \left(\frac{T}{T_c}\right)^{3/2}}$$



What about the energy for $T < T_c$?
Only the excited states contribute

$$E = \frac{g_s V m^{3/2}}{\sqrt{2}\pi^2 \hbar^3} \int_0^\infty \frac{\varepsilon^{3/2}\, d\varepsilon}{e^{(\varepsilon-\mu)/kT} - 1}$$

~~(Also for $T < T_c$)~~

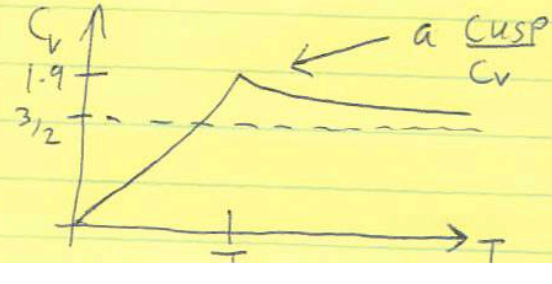
$$= \frac{3}{2} kT \; V g_s n_Q \; \underbrace{\frac{4}{3\sqrt{\pi}} \int_0^\infty \frac{x^{3/2}\, dx}{e^{x-\alpha} - 1}}_{\zeta(5/2) = 1.341 \text{ for } \alpha = 0} \qquad \alpha = \frac{\mu}{kT}$$

otherwise write $F_{5/2}(\alpha)$

$$\Rightarrow \quad E = \frac{3}{2} N kT \left(\frac{N_e}{N}\right) \left[\frac{F_{5/2}(\alpha)}{F_{3/2}(\alpha)}\right]$$

For $T < T_c$ $\alpha = 0$, $\frac{N_e}{N} = \left(\frac{T}{T_c}\right)^{3/2} \Rightarrow \boxed{E = 0.76\, N k_B T_c \left(\frac{T}{T_c}\right)^{5/2}}$

$$\Rightarrow \quad \boxed{C_V = 1.9\, N k_B \left(\frac{T}{T_c}\right)^{3/2}}$$



a <u>cusp</u> in C_V

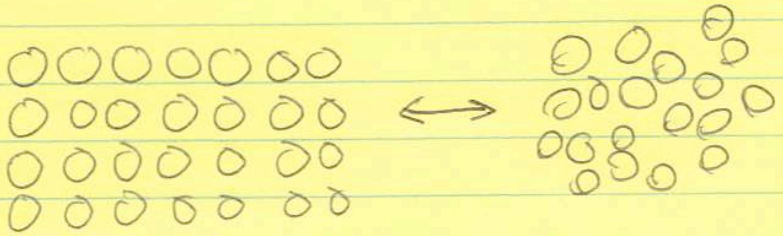
Applications

- Superfluidity in ^4He
  if we use our formula for T_c $\left(\frac{V}{N} = 27.6 \text{ cm}^3/\text{mol}\right.$ for liquid He$\left.\right)$
  we get 3.13 K, close to the superfluid transition at 2.176 K.
- Superconductivity — electrons pair.
- BEC was achieved experimentally in 1995 with a dilute gas of Rb-87 atoms.

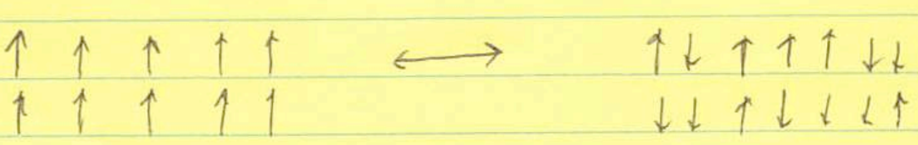
III PHASE TRANSITIONS (Chandler Chp 5)

A discontinuity or singularity in a thermodynamic function.

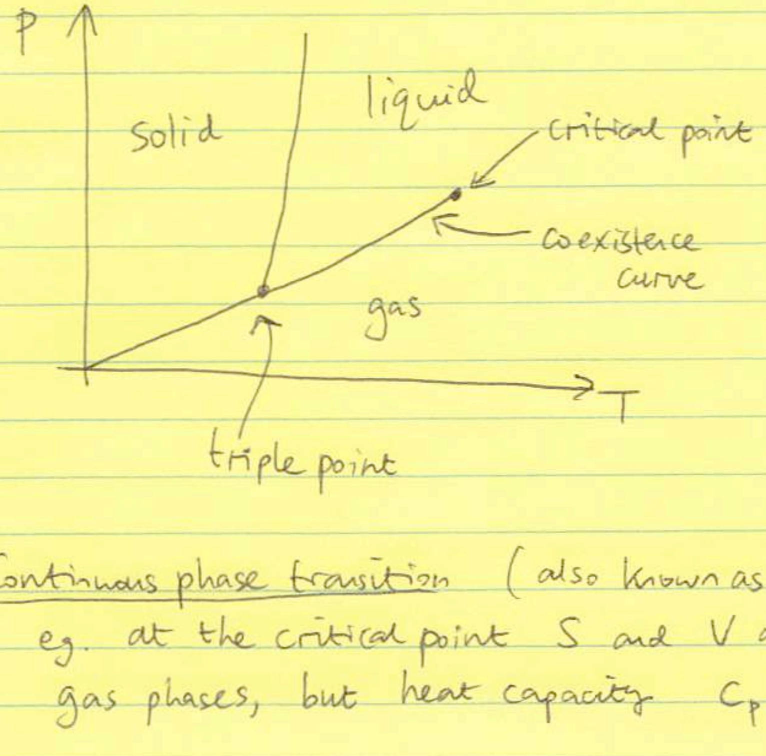
eg. Solid – liquid transition



eg. Ferromagnet at the Curie temperature



Phase diagram



1st order phase transition

discontinuous 1st derivatives of G

$$V = \frac{\partial G}{\partial P}\bigg|_T \qquad \underbrace{S = -\frac{\partial G}{\partial T}\bigg|_p}_{\text{gives latent heat } T\Delta S}$$

Continuous phase transition (also known as 2nd order)

eg. at the critical point S and V are the same in the liquid and gas phases, but heat capacity $C_p \propto |t|^{-\alpha}$ ← critical exponent

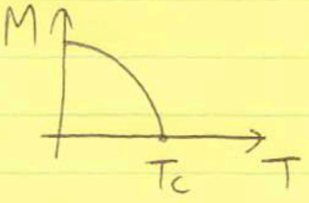
$$t = \frac{T}{T_c} - 1$$

There is hysteresis in these transitions – eg. superheated water is in a metastable state at a T above its boiling point.

We want to understand these properties from a microscopic point of view. Interactions between particles are key. We'll start with a specific example – the Ising model of interacting spins, but first list some general features:

- The phase at lower T is generally more ordered (see sketches of solid-liquid or magnetization above) – introduce an order parameter, eg. M magnetization of the system

find that $M \propto |t|^{\beta}$ ← another example of a critical exponent



[Compare plot of $\frac{N_0}{N}$ for BEC]

- Long-range order is associated with a correlation length that diverges $\xi \propto |t|^{-\nu}$

- Symmetry breaking M chooses a direction.

- Universality different physical systems often have similar critical exponents

- Large fluctuations occur near a critical point

[recall that for example $\sigma_E^2 = k_B T^2 C_V$
if C_V shows divergent behavior it can lead to large fluctuations.
"Critical opalescence" – density fluctuations near critical point scatter light]

(we'll follow Chandler quite closely here)

Ising model – Overview

System of spins with nearest-neighbor interactions.

The energy is $$E_\nu = -\sum_{i=1}^{N} \mu H s_i - J \underset{\substack{\uparrow \\ \text{nearest neighbor} \\ \text{pairs}}}{\sum} s_i s_j$$

In this model we assume the spins point up or down $s_i = \pm 1$ (more general model Heisenberg model has interaction term of the form $\underline{s}_i \cdot \underline{s}_j$ but same qualitative features).

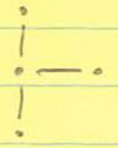
μ = magnetic moment
H = applied magnetic field
The energy of a spin is $\mp \mu H$ if it points in the same/opposite direction as the field H.

$J > 0$	spins want to align with each other	ferromagnetic
$J < 0$	" anti-align	anti-ferromagnetic

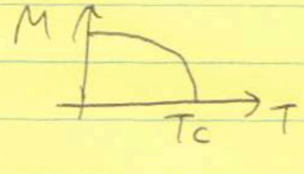
Order parameter $\langle M \rangle = \sum_{i=1}^{N} \mu s_i$

eg. $H = 0$ cubic lattice

lowest energy state $E_0 = -\underset{\substack{\uparrow \\ \text{dimensionality}}}{D} N J$



The fact that neighboring spins prefer to be aligned suggests that there could be a <u>spontaneous magnetization</u> (even with no applied field) at low enough T.



The partition function is

$$Q(T, N, H) = \sum_{\nu} e^{-\beta E_\nu}$$

$$= \sum_{s_1}\sum_{s_2}\cdots\sum_{s_N=\pm 1} \exp\left[\beta\mu H \sum_{i=1}^{n} s_i + \beta J \sum_{ij} s_i s_j\right]$$

In 1D, the interaction term can be written as a single sum,

$$-J\sum_{i=1}^{N} s_i s_{i+1} \qquad \left(\begin{array}{l}\text{assuming periodic boundary}\\ \text{conditions } s_{N+1} = s_1\end{array}\right)$$

the partition function can then be summed eg. For $H=0$ the result is

$$Q(T, N, 0) = \left[2\cosh(\beta J)\right]^N$$

More on this later, but for now note that there is no spontaneous magnetization in 1D. To understand this consider the energy cost of creating a domain wall

$$\begin{array}{cccccccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & & E = -JN \\ 1 & \cdots & & & N/2 & \cdots & & N & & \end{array}$$

$$\begin{array}{ccccccccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & E = -J(N-4) \\ 1 & - & - & - & N/2 & - & - & - & - & N & \end{array}$$

the energy difference is of order $\frac{1}{N}$

In 2D eg.

$$\begin{array}{cccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{array}$$

energy difference from the ground state is $\sim \frac{1}{\sqrt{N}}$

this larger energy difference can stabilize the ground state —

there is spontaneous magnetization in 2D.

Onsager found an exact solution in 2D

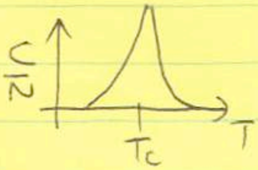
$$Q(T, N, H=0) = \left[2\cosh(\beta J) e^{I} \right]^N$$

$$I = \frac{1}{2\pi}\int_0^{\pi} d\phi \ \ln\left\{ \frac{1}{2}\left[1 + (1-k^2\sin^2\phi)^{1/2} \right] \right\}$$

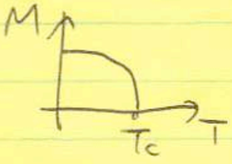
$$k = \frac{2\sinh(2\beta J)}{\cosh^2(2\beta J)}$$

The critical temperature is $k_B T_c = 2.269 J$.

Critical behavior: $\frac{C}{N} \sim \frac{8k_B}{\pi}(\beta J)^2 \ln\frac{1}{|T-T_c|}$



$T<T_c$ $\frac{M}{N} \propto (T_c - T)^{\beta}$ $\beta = 1/8$



No exact solution has been found in 3D. Numerical simulations show that in 3D

$$T_c \simeq (2 \times \text{the 2D critical temperature})$$

$$\frac{C}{N} \propto |T-T_c|^{-\alpha} \qquad \alpha \simeq 0.125$$

$$\frac{M}{N} \propto (T_c - T)^{\beta} \qquad \beta \simeq 0.313$$

A Monte Carlo simulation of the 2D Ising model is available at
http://physics.ucsc.edu/~peter/ising/ising.html
[website of Peter Young, UCSC]
which runs in the browser window.

Mean field theory

Now discuss an approximate solution to the Ising model in which we treat each spin as independently interacting with the thermodynamic average value of the surrounding spins.

The energy of the Ising model is

$$E = -\mu H \sum s_i - J \sum_{i,j} s_i s_j$$

sum over unique pairs of nearest neighbors (n.n.)

The idea is to write

$$s_i s_j = (s_i - \langle s\rangle + \langle s\rangle)(s_j - \langle s\rangle + \langle s\rangle)$$

$$= \langle s\rangle^2 + (s_i - \langle s\rangle)\langle s\rangle + (s_j - \langle s\rangle)\langle s\rangle$$

$$+ \underbrace{\cancel{(s_i - \langle s\rangle)(s_j - \langle s\rangle)}}$$

we throw away this term – ie. neglect correlations in neighbouring spins

$$\Rightarrow \quad \sum s_i s_j = -\langle s\rangle^2 \underbrace{\frac{Nz}{2}} + \sum_{i,j} \langle s\rangle (s_i + s_j)$$

the total number of unique n.n. pairs

z = number of nearest neighbors per lattice site.

We can rewrite the last term

sum over all pairs

$$\frac{1}{2} \sum_{i,j} \langle s\rangle (s_i + s_j)$$

$$= \frac{1}{2}\, 2 \sum_{i,j} s_i \langle s\rangle$$

$$= z \langle s\rangle \sum_i s_i$$

here do the sum over nearest neighbors

we're left with a sum over lattice sites

$$\Rightarrow \quad E = -\mu H \sum s_i + \frac{JNz}{2} \langle s \rangle^2 - Jz \langle s \rangle \sum s_i$$

$\Rightarrow$ we can write $\boxed{E = \sum_{i=1}^{N} E_i}$

where $\boxed{E_i = -\mu H s_i - Jz \langle s \rangle s_i + \frac{Jz}{2} \langle s \rangle^2}$

We can write the energy as a sum over energies of individual spins. Each spin has an energy E_i set by the external field H and the mean spin $\langle s \rangle$.

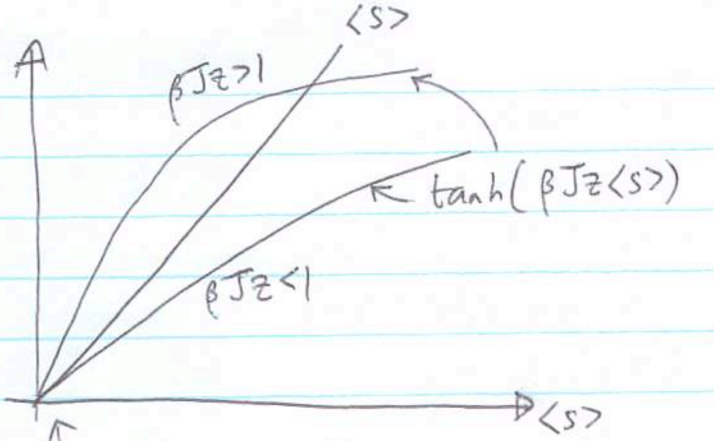
To make the theory self-consistent, we must have

$$\frac{\sum_{s_i = \pm 1} s_i e^{-\beta E_i}}{\sum_{s_i = \pm 1} e^{-\beta E_i}} = \langle s \rangle$$

$$= \frac{\exp\left(\beta\mu H + \beta J z \langle s \rangle\right) - \exp\left(-\beta\mu H - \beta J z \langle s \rangle\right)}{\exp\left(\beta\mu H + \beta J z \langle s \rangle\right) + \exp\left(-\beta\mu H - \beta J z \langle s \rangle\right)}$$

or $\boxed{\langle s \rangle = \tanh\left(\beta\mu H + \beta J z \langle s \rangle\right)}$ — (*)

This equation determines $\langle s \rangle$ for particular values of T, H.
eg. if $H = 0$ we can solve to get $\langle s \rangle (T)$.
Do this graphically by plotting the LHS and RHS and looking for the intersection point.



The magnetization of the system is

$$M = N\mu\langle s\rangle$$

the slope of $\tanh(\beta J z\langle s\rangle)$ for $\langle s\rangle$ close to zero is just $\beta J z$ because $\tanh(\alpha x) \approx \alpha x$ for $x \ll 1$

When $\beta J z > 1$ then there is a solution at finite $\langle s\rangle$

– ie. <u>spontaneous magnetization</u> occurs in this model for $T < T_c$

where $\frac{Jz}{k_B T_c} = 1$

or $\boxed{T_c = \frac{zJ}{k_B}}$

How does this compare with the exact solution?

	MFT	Exact
1D	$T_c = \frac{2J}{k_B}$	no phase transition!
2D	$T_c = \frac{4J}{k_B}$	$2.269\, J/k_B$
3D	$\frac{6J}{k_B}$	$\approx 4.5\, J/k_B$

the MFT provides a better estimate as the number of n.n. increases

this is a general result that MFT becomes more accurate in higher dimensions → so then the net fluctuations are smaller

What do we predict for the critical exponent near $T = T_c$?
First, rewrite equation (*) as

$$\frac{2\langle s\rangle T_c}{T} = 2\beta J z \langle s\rangle = \ln\left(\frac{1+\langle s\rangle}{1-\langle s\rangle}\right)$$

and expand near $\langle s\rangle = 0$

$$\ln\left(\frac{1+\langle s\rangle}{1-\langle s\rangle}\right) \simeq 2\langle s\rangle + \frac{2\langle s\rangle^3}{3} \cdots$$

$$\Rightarrow \quad \frac{T_c}{T} \simeq 1 + \frac{\langle s\rangle^3}{3}$$

$$\Rightarrow \quad \langle s\rangle \simeq \sqrt{3}\left(\frac{T_c}{T} - 1\right)^{1/2}$$

or defining $t = 1 - \frac{T}{T_c}$

so that $\frac{T_c}{T} - 1 = \frac{t}{1-t} \simeq t$

$$\Rightarrow \quad \boxed{\langle s\rangle \propto t^{1/2}}$$

The critical exponent is $\frac{1}{2}$ independent of the number of dimensions.
[This compares to $\beta = 0{\cdot}313$ in 3D $\quad \beta = 0{\cdot}125$ in 2D
$M \propto t^{\beta}$]

These differences at the critical point arise because the mean field theory ignores fluctuations in all except the one spin that is acted on by the mean field. At the critical point, fluctuations on large lengthscales are important, and the mean field theories fail to capture this aspect, even when extended to include fluctuations in the nearest neighbors etc.

Landau theory

This is a phenomenological theory of phase transitions in which the symmetries of the system are used to write down an expression for the free energy as a function of the order parameter.

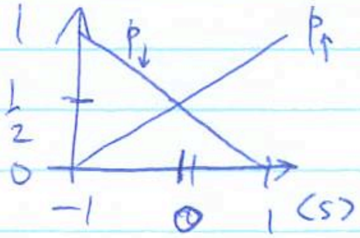
To see what the free energy looks like in MFT for the Ising model as a function of $\langle s \rangle$, we assume $\langle s \rangle$ is given and calculate $A = E - TS$.

The energy is $E = -\mu H N \langle s \rangle - \frac{1}{2} z J N \langle s \rangle^2$.

For the entropy, we can evaluate the Gibbs entropy

$$S = -k_B \sum p_i \ln p_i$$

with $p_\uparrow = \dfrac{1 + \langle s \rangle}{2}$ $\qquad p_\downarrow = \dfrac{1 - \langle s \rangle}{2}$



[these probabilities satisfy $p_\uparrow + p_\downarrow = 1$
$p_\uparrow - p_\downarrow = \langle s \rangle$]

$$\Rightarrow \quad S = -k_B N \left[\frac{1}{2} (1 + \langle s \rangle) \ln\left(\frac{1 + \langle s \rangle}{2}\right) + \left(\frac{1 - \langle s \rangle}{2}\right) \ln\left(\frac{1 - \langle s \rangle}{2}\right) \right].$$

The equilibrium value of $\langle s \rangle$ minimizes $A = E - TS$:

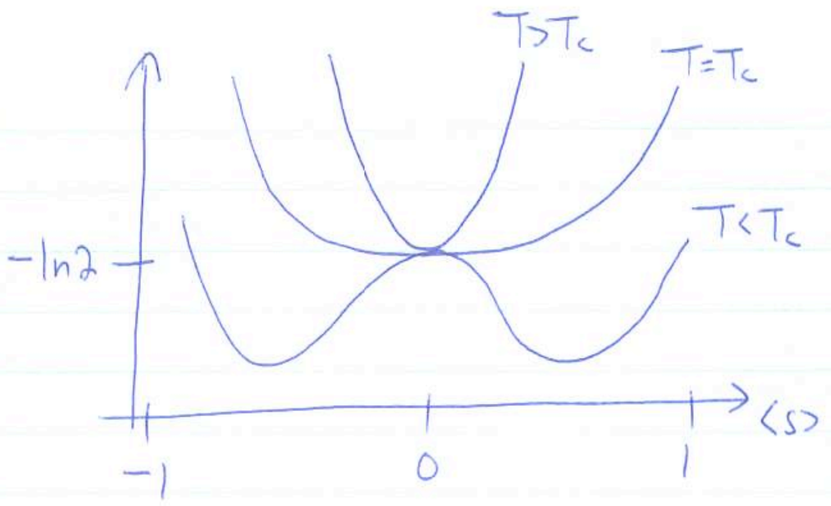
$$\frac{\partial A}{\partial \langle s \rangle} = 0 \quad \Rightarrow \quad -\frac{\mu H N}{N k_B T} - \frac{z J N \langle s \rangle_{eqm}}{N k_B T} + \frac{1}{2} \ln\left(\frac{1 + \langle s \rangle_{eqm}}{1 - \langle s \rangle_{eqm}}\right) = 0$$

The last term is $\tanh^{-1} \langle s \rangle_{eqm}$

$$\Rightarrow \quad \boxed{\langle s \rangle_{eqm} = \tanh\left(\mu H + z J \langle s \rangle_{eqm}\right)}$$

which is what we found last time.

Plot $A(\langle s\rangle)$:



In equilibrium, the system finds the value of $\langle s\rangle$ that minimizes A. The phase transition is associated with a new minimum appearing in A at finite $\langle s\rangle$ for $T<T_c$.

For small $\langle s\rangle$, we can expand

$$\frac{A}{Nk_BT} \approx -\log 2 + \frac{1}{2}\langle s\rangle^2\left(1-\frac{T_c}{T}\right) + \frac{\langle s\rangle^4}{12} - \frac{\mu H}{k_BT}\langle s\rangle + O(\langle s\rangle^6)$$

Landau wrote down a free energy of this form using symmetry arguments. For example, for $H=0$ the system has up-down symmetry ie. we expect $A(\langle s\rangle, T) = A(-\langle s\rangle, T)$

$\Rightarrow$ only even powers of $\langle s\rangle$ should appear for $H=0$.

The change in sign of the $\langle s\rangle^2$ term below T_c gives rise to new minima of A at non-zero $\langle s\rangle$. Symmetry breaking occurs when the system chooses a value of $\langle s\rangle$ (>0 or <0).
At $T=T_c$ the $\langle s\rangle^2$ term vanishes which gives the flat-bottomed curve in the plot at $T=T_c$. ($A = \text{constant} + \langle s\rangle^4$)
Introducing non-zero H lowers one of the two minima with respect to the other

The susceptibility χ and fluctuations in M

The partition function for the Ising model is

$$Q = \sum_{\{s_k\}} \exp\left(\beta \mu H \sum s_i + \beta J \sum s_i s_j \right)$$

$\uparrow$ sum over all possible configurations of spins

$$\Rightarrow \quad \frac{\partial \ln Q}{\partial H} = \beta\mu \left\langle \sum s_i \right\rangle = \frac{M}{k_B T}$$

or $\boxed{M = -\left.\frac{\partial A}{\partial H}\right|_T}$

Take a second derivative:

$$\frac{\partial^2 \ln Q}{\partial H^2} = -\beta^2 \langle M \rangle^2 + \beta^2 \langle M^2 \rangle$$

$$\Rightarrow \quad \frac{\partial M}{\partial H} = \chi = \beta \sigma_M^2$$

or $\boxed{\sigma_M^2 = k_B T \chi}$

(compare our previous result for energy fluctuations
$\sigma_E^2 = k_B^2 T C_V$)

Two more critical exponents in mean field theory

Let's evaluate χ in MFT:

$$\chi = \frac{\partial M}{\partial H} = N\mu \frac{\partial \langle s \rangle}{\partial H}$$

but

$$\langle s \rangle = \tanh\left(\beta\mu H + \langle s \rangle T_c/T \right)$$

$$\Rightarrow \quad \frac{\partial \langle s \rangle}{\partial H} = \operatorname{sech}^2\left(\beta\mu H + \frac{\langle s \rangle T_c}{T} \right) \left(\beta\mu + \frac{T_c}{T} \frac{\partial \langle s \rangle}{\partial H} \right)$$

$$\Rightarrow \quad \frac{\partial \langle s \rangle}{\partial H} = \frac{\frac{\mu}{k_B T_c} \frac{T_c}{T}}{\cosh^2(\;) - \frac{T_c}{T}}$$

but $\cosh^2(\;) = \frac{1}{1-\tanh^2(\;)} = \frac{1}{1-\langle s \rangle^2}$

$$\Rightarrow \quad \boxed{\chi = \frac{\left(\frac{N\mu^2}{k_B T_c}\right)\left(\frac{T_c}{T}\right)}{\frac{1}{1-\langle s \rangle^2} - \frac{T_c}{T}}}$$

the susceptibility in MFT

What is the behavior close to the critical point?
We know that $\langle s \rangle \propto t^{\beta}$ with $t = 1 - \frac{T}{T_c}$ and $\beta = \frac{1}{2}$.

Writing χ in terms of t we find $\chi \propto t^{-1}$ for small t.

This is another critical exponent $\boxed{\chi \propto \left(1 - \frac{T}{T_c}\right)^{-\gamma}}$

MFT gives $\boxed{\gamma = 1}$

Another critical exponent is δ which describes how M depends on the applied field at $T = T_c$: $M(T=T_c) \propto H^{1/\delta}$.
To derive this, write

$$\langle s \rangle \frac{T_c}{T} + \frac{\mu H}{k_B T} = \tanh^{-1}\langle s \rangle \simeq \langle s \rangle + \frac{\langle s \rangle^3}{3}$$

When $T = T_c$ $\quad \langle s \rangle^3 \propto H \quad$ or $\quad M \propto H^{1/3}$

$$\Rightarrow \quad \boxed{\delta = 3}$$

The values for the exact solution in 3D are $\gamma = 1.24$, $\delta = 4.8$.

Renormalization group theory

For a summary of this approach and historical background, see Ken Wilson's Nobel lecture in RMP vol. 55, p583 (1983)

In the mean field theory approach, we averaged over all degrees of freedom except for a single spin. The idea behind the renormalization group approach is to average a finite number of degrees of freedom at a time. We average out the short wavelength behavior, finding a new system with the same behavior at long wavelengths but new values (renormalized) of parameters.

We follow Chandler and start with the 1D Ising model as an example.

$$Q = \sum_{s_1, s_2, \dots s_N = \pm 1} \exp\left[K \left(\cdots s_1 s_2 + s_2 s_3 + s_4 s_5 + \cdots \right)\right]$$

where we define $K = \frac{J}{k_B T}$ the "coupling constant" (reciprocal temperature).

The idea is to first sum over the even numbered spins

$$Q = \sum_{s_1, s_2 \dots} \exp\left[K(s_1 s_2 + s_2 s_3)\right] \exp\left[K(s_3 s_4 + s_4 s_5)\right] \cdots$$

Do the sum over even spins:

$$Q = \sum_{s_1, s_3, s_5 \dots} \left\{ \exp(K(s_1+s_3)) + \exp(-K(s_1+s_3)) \right\}$$
$$\times \left\{ \exp\left[K(s_3+s_5)\right] + \exp\left[-K(s_3+s_5)\right] \right\} \cdots$$

Now try to cast this new partition function in the form of an Ising model with only N/2 spins and a different coupling K'.

So, we look for a function $f(K)$ and a K' such that

$$e^{K(s+s')} + e^{-K(s+s')} = f(K)\, e^{K's's}$$

for all choices of $s, s' = \pm 1$.

Then $$Q(K,N) = \sum_{s_1, s_3, s_5 \ldots} f(K)^{N/2} \exp(K's_1 s_3)\exp(K's_3 s_5)\cdots$$

or $$\boxed{Q(K,N) = f(K)^{N/2}\, Q\left(K', \tfrac{N}{2}\right)} \quad (*)$$

<u>a Kadanoff transformation</u>

To find K' and $f(K)$:

1) $s = s' = \pm 1$ $\qquad e^{2K} + e^{-2K} = f(K)\, e^{K'}$

2) $s = -s' = \pm 1$ $\qquad 2 = f(K)\, e^{-K'}$

$$\Rightarrow \boxed{\begin{aligned} K' &= \tfrac{1}{2} \ln \cosh(2K) \\ f(K) &= 2\sqrt{\cosh(2K)} \end{aligned}}$$

the "Renormalization group" equations

Next, write the free energy as $-kT \ln Q = -kT\, N g(K)$

the recursion relation $(*)$ gives

$$\boxed{g(K') = 2g(K) - \ln\left[2\sqrt{\cosh(2K)}\right]}$$

The idea is that if we can calculate g for some value of K then we can iterate using the recursion relation and obtain g for other values of K.

We can invert these relations to go in the other direction

$$\boxed{\begin{aligned} K &= \frac{1}{2}\cosh^{-1}\left(e^{2K'}\right) \\ \text{and} \qquad & \\ g(K) &= \frac{1}{2}\left[g(K') + \ln 2 + K'\right] \end{aligned}} \;\text{—}\; (*)$$

these equations always give $K' < K$
ie. the new value of K is larger than the previous value
(or temperature is lower)
[the previous set of equations move in the other direction)

We can test these equations for the 1D Ising model. At high temperature / low values of K $\quad Q \simeq 2^N$

$$\Rightarrow \quad g(K \approx 0) = \ln 2$$

Choose a starting value $K = 0{\cdot}01$, set $g(0{\cdot}01) \simeq \ln 2$ and iterate. Chandler gives a table* which shows that the values of $g(K)$ obtained in this way are close to the exact answer. [Recall the exact solution for $H=0$ gives

$$g(K) = \frac{\ln Q}{N} = \ln\left[2\cosh(K)\right] \quad \rfloor$$

(iterate using equations (*) to update the values of K and $g(K)$ at each step.)

Notice what we did here – we took an initial value of $K \ll 1$ for which we can trivially calculate $g(K)$. Then we applied the recurrence relations to move to a value of K where it is difficult to calculate $g(K)$ directly.

(* Chandler's table is attached)

temperature) K. If this
recursion relation from
g from a system with another
we look for a function of K,
such that

$f(K)e^{K'ss'}$

quantities, then

$s_3)f(K)\exp(K's_3s_5)\cdots$

relation. A transformation like

and $f(K)$, we note that if

$(K)e^{K'}$.

±1 from which we have

$-K'$

K'.

unknowns. The solution is

osh $(2K)$, (a)

(K).

s a free energy, and since free
$g(K)$ to be intensive—that is,
rom the recursion relation,
/2), we have $g(K)=(1/2)\ln f(K)$
$^{1/2}(2K)$,

$[2\sqrt{\cosh(2K)}]$. (b)

normalization group (RG) equa-
s which obey the group property,
scheme.) If the partition function
can generate $\ln Q = Ng(K)$ for
"renormalization." Notice that in
(a) and (b), the new coupling
is always less than K.

---

from Chandler

An alternative set of RG equations would be

$$K = (1/2)\cosh^{-1}(e^{2K'}), \tag{c}$$

which is the inverse of (a), and

$$g(K) = (1/2)g(K') + (1/2)\ln 2 + K'/2, \tag{d}$$

which is obtained by noting $f(K) = 2\exp(K')$.

Exercise 5.12 Derive these RG equations, and show that $K > K'$.

To see how these equations work, we will apply (c) and (d) starting with a small value of the coupling constant. Repeated application will generate $g(K)$ at successively larger values of the coupling constant. Let's start with $K' = 0.01$. For such a small coupling constant, interactions between spins are nearly negligible. Thus, $Q(0.01, N) \approx Q(0, N) = 2^N$. As a result,

$$g(0.01) \approx \ln 2.$$

We now start the iteration. From (c) and (d) we find

$$K = 0.100\,334,$$
$$g(K) = 0.698\,147.$$

We now use these numbers as the new K primed quantities and obtain

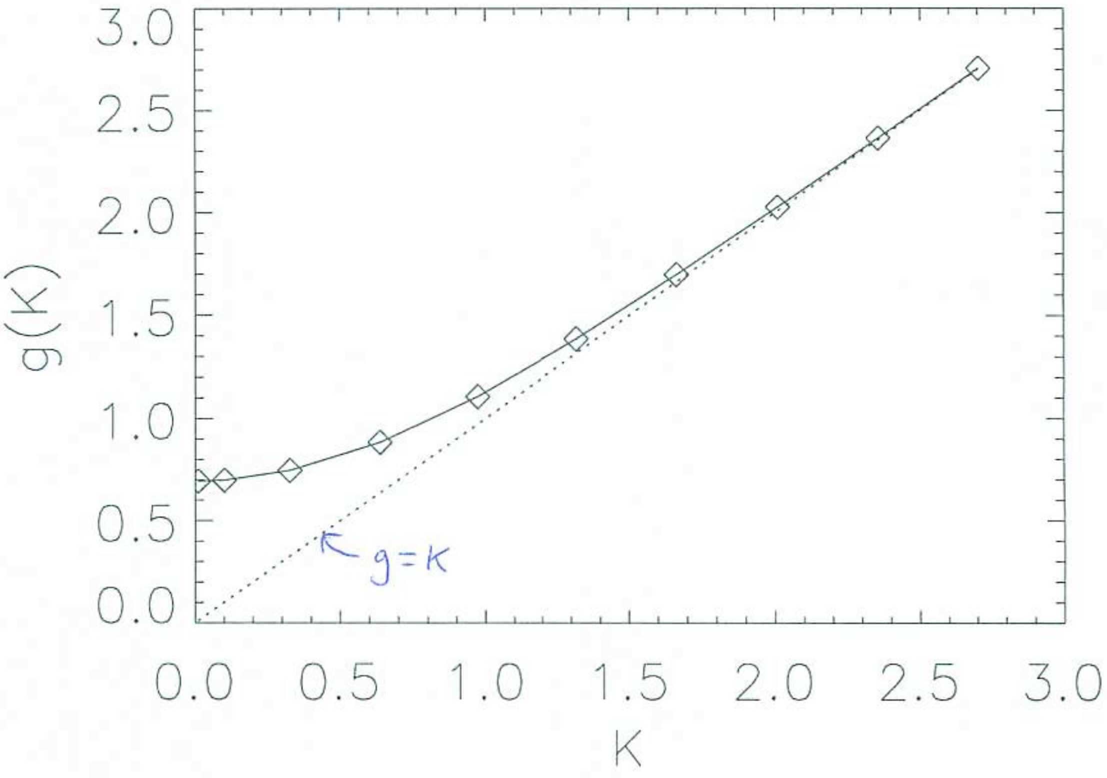
$$K = 0.327\,447,$$
$$g(K) = 0.745\,814,$$

and so on.

exact Solution
$\ln[2\cosh K]$

	K	Renormalization group	Exact
Successive application of RG equations (c) and (d) ↓	0.01	ln 2	0.693 197
	0.100 334	0.698 147	0.698 172
	0.327 447	0.745 814	0.745 827
	0.636 247	0.883 204	0.883 210
	0.972 710	1.106 299	1.106 302
	1.316 710	1.386 078	1.386 080
	1.662 637	1.697 968	1.697 968
	2.009 049	2.026 876	2.026 877
	2.355 582	2.364 536	2.364 537
	2.702 146	2.706 633	2.706 634

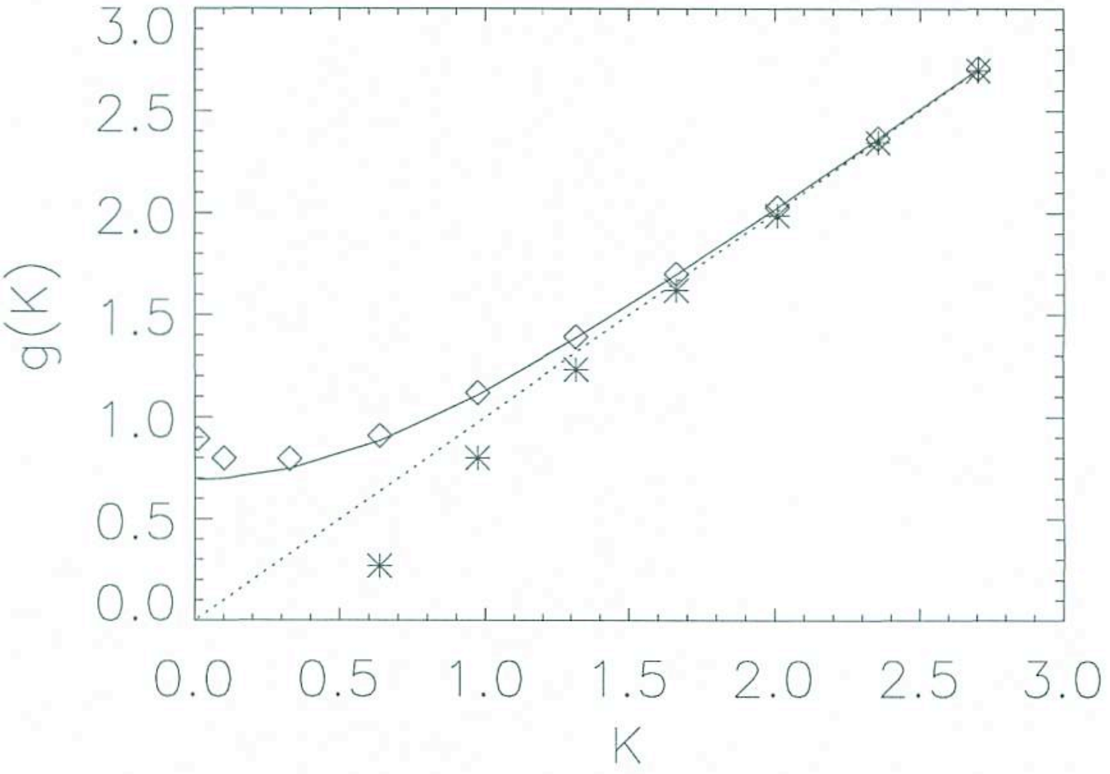
Solution to 1D RG equations



Solid line exact solution $g = \ln[2 \cosh K]$

points are using RG equations starting at

$g = \ln 2$
$K = 0.01$



diamonds:

as before, but now starting at $g = \ln 2 + 0.2$
Note the solution quickly converges to the exact solution.

stars: now I take the final point at $K \simeq 2.7$
subtract 0.01 from g and now evolve backwards to smaller K.
Note the solution diverges away from the exact solution in this case.

Program in IDL that evaluates the RG equations for 1D Ising model

```
pro rg
; Evaluate RG equations for 1D Ising model

; number of iterations
n=10
g=dblarr(n)
K=dblarr(n)

; initial values for K and g
K[0]=0.01d
g[0]=alog(2d0) +0.2
; here I've included an offset in the initial value of g
; so that we can see the convergence to the correct solution

for i=1,n-1 do begin

   ; calculate the new value of g
   g[i]=0.5*(g[i-1]+alog(2.0)+K[i-1])

   ; calculate the new K
   K[i]=0.5d*alog(exp(2.0*K[i-1])+sqrt(exp(4.0*K[i-1])-1))
```

here I'm using the identity $\cosh^{-1} x = \ln\left[x+\sqrt{x^2-1}\right]$

```
endfor

; the exact solution
gexact=alog(2.0*cosh(K))

; plot
plot, K, g, xtitle='K', ytitle='g(K)', charsize=2, linestyle=0, psym=4,
symsize=1.5
oplot, K, gexact, linestyle=0
oplot, K,K,linestyle=1


; now take our final values of (K,g) and try to work backwards
; we should see divergence from the exact solution

; introduce an offset
g[n-1]-=0.01

for i=n-2,0,-1 do begin

   ; calculate the new value of g
   g[i]=2.0*g[i+1]-alog(2.0*sqrt(cosh(2.0*K[i+1])))
```

```
    ; calculate the new K
    K[i]=0.5*alog(cosh(2.0*K[i+1]))

endfor
oplot, K, g, symsize=1.5,psym=2

end
```

More on the 1D RG solution : convergence to the exact solution

1) First check that the exact solution satisfies the RG equations

$$g(K') = 2g(K) - \ln\left[2\sqrt{\cosh 2K}\right]$$

$$\text{where } K' = \tfrac{1}{2}\ln\left[\cosh 2K\right]$$

The exact solution $\Rightarrow$ $g(K') = \ln[2\cosh K']$

$$= \ln\left[2\cosh\left(\tfrac{1}{2}\ln\cosh 2K\right)\right]$$

but $2\cosh(\ln x) = e^{\ln x} + e^{-\ln x}$
$= x + \frac{1}{x}$

$$\Rightarrow g(K') = \ln\left[\sqrt{\cosh 2K} + \frac{1}{\sqrt{\cosh 2K}}\right]$$

$$= \ln[\underbrace{\cosh 2K + 1}_{2\cosh^2 K}] - \ln\sqrt{\cosh 2K}$$

$$= \ln 2 - \ln\sqrt{\cosh 2K} + 2\ln(\cosh K)$$

$$= 2\ln[2\cosh K] - \ln(2\sqrt{\cosh 2K})$$

$$= 2g(K) - \ln(2\sqrt{\cosh 2K}) \quad ✓$$

2) Now write $g(K) = \ln[2\cosh K] + \delta g$

(agrees with my numerical solution)

and $g(K') = \ln[\cancel{2\cosh K'}] + \delta g'$

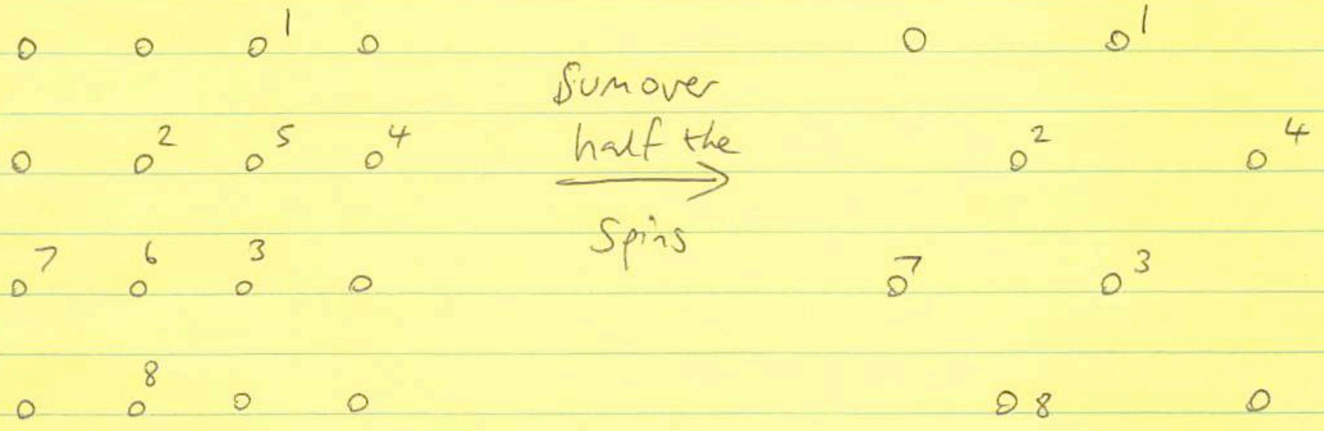
↑

$$\underset{\text{(use recurrence relation)}}{=} 2\delta g + \cancel{2\ln[2\cosh K]} - \cancel{\ln[2\sqrt{\cosh 2K}]}$$

$$\Rightarrow \boxed{\delta g' = 2\delta g} \quad \text{or} \quad \boxed{\delta g = \frac{\delta g'}{2}}$$

when using the inverse transforms the error decreases by factor 2 each time.

Now apply these ideas to 2D.



To do the sum, gather together the terms in $s_5, s_6 \dots$

$$Q = \sum \cdots \exp\left[K(s_1+s_2+s_3+s_4)s_5\right]$$

$$\times \exp\left[K(s_2+s_3+s_7+s_8)s_6\right] \cdots$$

do the sum $\Rightarrow$

$$Q = \sum \cdots \left\{\exp\left[K(s_1+s_2+s_3+s_4)\right] + \exp\left[-K(s_1+s_2+s_3+s_4)\right]\right\}$$

$$\times \left\{\exp\left[K(s_2+s_3+s_7+s_8) + \exp\left[-K(s_2+s_3+s_7+s_8)\right]\right\}$$

$$\cdots\cdots$$

Following the same approach as in 1D, we'd like to write

$$\cancel{Q}\ \exp\left[K(s_1+s_2+s_3+s_4)\right] + \exp\left[-K(s_1+s_2+s_3+s_4)\right)$$

$$= f(K)\exp\left[K'(s_1s_2+s_2s_3+s_3s_4+s_1s_4)\right]$$

But in this expression there are only two degrees of freedom $f(K)$ and K' not enough because there are >2 choices for (s_1, s_2, s_3, s_4) [whereas before we had only two choices $s=s'=\pm 1$ or $s=-s'=\pm 1$).

There are four unique choices for the spins s_1, s_2, s_3, s_4

$$\left.\begin{array}{ll} s_1 = s_2 = s_3 = s_4 = \pm 1 & \text{all the same} \\ s_1 = s_2 = s_3 = -s_4 = \pm 1 & \text{three the same} \\ s_1 = s_2 = -s_3 = -s_4 = \pm 1 & \text{opposite sides} \\ s_1 = -s_2 = s_3 = -s_4 = \pm 1 & \text{opposite corners} \end{array}\right\} (*)$$

A rewriting that does work is

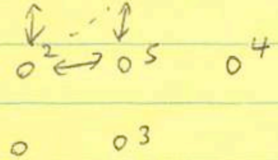
nearest neighbors

$$\boxed{\begin{aligned} & e^{K(s_1+s_2+s_3+s_4)} + e^{-K(s_1+s_2+s_3+s_4)} \\ & \qquad = f(K) \exp\left[\frac{K_1}{2}(s_1s_2 + s_2s_3 + s_3s_4 + s_4s_1) + K_2(s_1s_3 + s_2s_4) \right. \\ & \qquad\qquad \left. + K_3(s_1s_2s_3s_4) \right] \end{aligned}} \quad (**)$$

Inserting the set of possibilities (*) and solving for $K_1, K_2, K_3, f(K)$ one finds:

$$\boxed{\begin{aligned} K_1 &= \tfrac{1}{4} \ln \cosh(4K) \\ K_2 &= \tfrac{1}{8} \ln \cosh(4K) \\ K_3 &= \tfrac{1}{8} \ln \cosh(4K) - \tfrac{1}{2} \ln \cosh(2K) \\ f(k) &= 2\left[\cosh(2K)\right]^{1/2}\left[\cosh(4K)\right]^{1/8} \end{aligned}}$$

Now we insert the rewrite (**) into the partition function sum. The nearest neighbour terms (K_1) appear twice in the sum – eg. s_1 and s_2 are coupled through the sums over spins 5 and 9:



whereas the next nearest neighbors eg. $s_1 s_3$ are coupled only by s_5 and so that term appears only once.

This is why we wrote $K_1/_2$ in equation (**).

The final result is

$$Q(K,N) = f(K)^{N/2} \sum_{N/2\text{ spins}} \exp\Big[K_1 \overset{\substack{\text{nearest}\\\text{neighbors}\\\downarrow}}{\sum_{ij}{}'} s_i s_j + K_2 \overset{\substack{\text{next - nearest}\\\text{neighbors}\\\downarrow}}{\sum_{\ell m}{}''} s_\ell s_m + K_3 \underset{\substack{\uparrow\\\text{sets of 4 spins in}\\\text{a square.}}}{\sum_{pqrt}{}'''} s_p s_q s_r s_t \Big]$$

---

* To obtain a useful iterative solution as in the 1D case, we must * approximate the RHS, to try to get it in the same form as the original Q.
[in 1D this happens automatically because of the simple topology]

First try – what if we keep only the nearest neighbor interactions?
Set $K_2 = K_3 = 0$

$$\text{Then} \quad Q(K,N) = [f(K)]^{N/2} \, Q(K_1, N/2)$$

$$\text{where} \quad K_1 = \tfrac{1}{4} \ln \cosh(4K)$$
$$f(K) = 2[\cosh(2K)]^{1/2} [\cosh(4K)]^{1/8}$$

Almost the same as the 1D case → no phase transition!

Second try Let's discard $K_3 = 0$ but keep K_2, ie keep the next-nearest neighbor interactions.

One way to do this is to write

$$K_1 \sum_{ij}{}' s_i s_j + K_2 \sum_{\ell m}{}'' s_\ell s_m = K' \sum_{ij}{}' s_i s_j$$

in other words, drop K_2 term, but increase K_1 to include the aligning tendency of the K_2 term.

When all the spins are aligned, if we choose $K' = K_1 + K_2$ we get the same energy either way, so let's choose

$$K' = K_1 + K_2 = \frac{3}{8} \ln \cosh(4K) \qquad \text{— } (***)$$

and

$$Q(K,N) = f(K)^{N/2}\, Q(K', N/2)$$

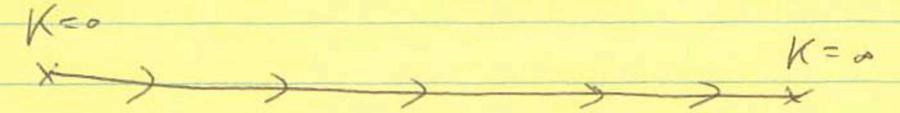
in terms of g, $$\underline{\underline{g(K') = 2g(K) - \ln\left\{ 2\sqrt{\cosh(2K)}\,[\cosh(4K)]^{1/8} \right\}}}$$

A new feature appears in equation $(***)$ — there is a non-trivial fixed point, ie. a value of K that gives $K' = K$ when inserted into $(***)$. Solving $K_c = \frac{3}{8} \ln \cosh(4K_c)$

we find $K_c = 0.506981$

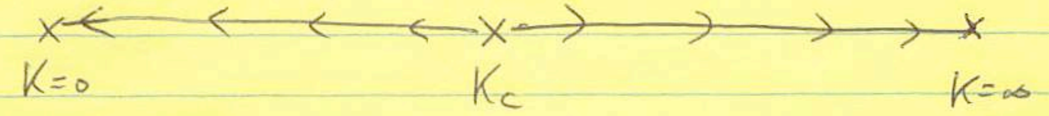
[Onsager exact 2D solution gives $K_c = 0.440723$]

The 1D solution we had previously had fixed points at $K=0$ and $K=\infty$.



(stable flow is to increasing values of K – moving in the opposite direction leads to diverging errors)

Here in 2D we have



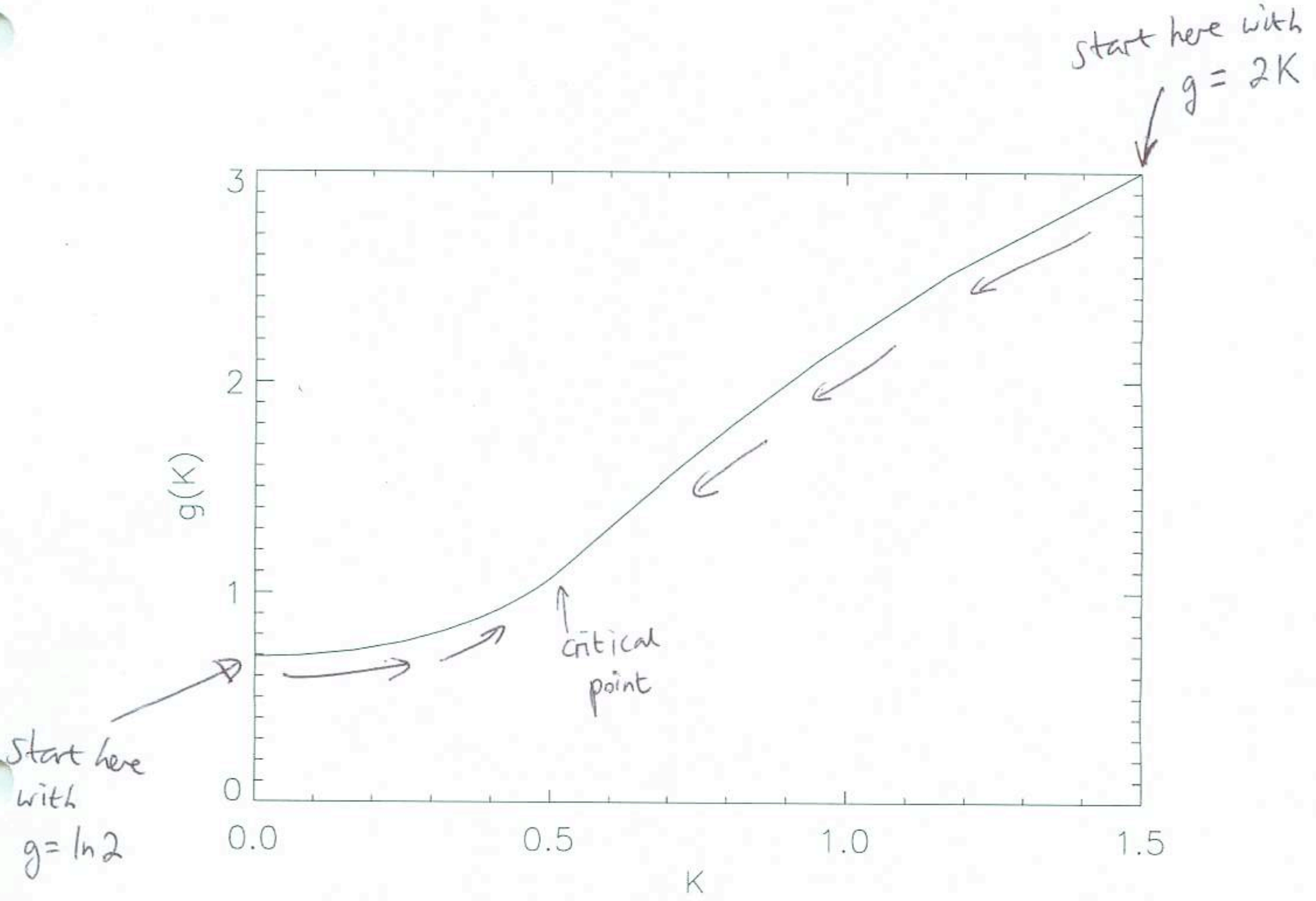
[eq. $(***)$ gives $K' < K$ for $K < K_c$ and $K' > K$ for $K > K_c$]

The inverse equations are

$$\boxed{\begin{aligned} K &= \frac{1}{4}\cosh^{-1}\left[e^{8K'/3}\right] \\ g(K) &= \frac{g(K')}{2} + \frac{1}{2}\ln\left\{ 2e^{2K'/3}\left[\cosh\left(\frac{4K'}{3}\right)\right]^{1/4} \right\} \end{aligned}}$$

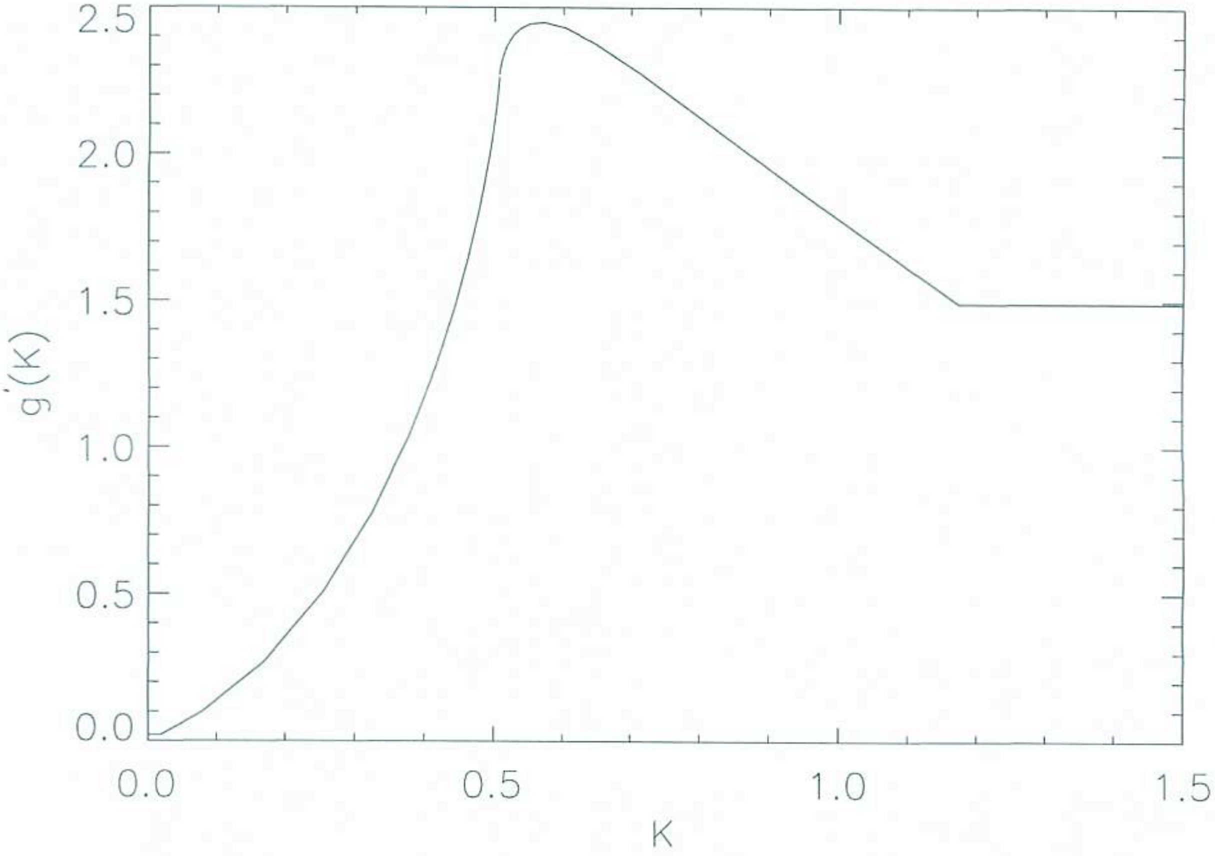
or $g(K) = \frac{1}{2}\left[g(K') + \ln 2 + \frac{2K'}{3} + \frac{1}{4}\ln\cosh\left(\frac{4K'}{3}\right) \right\}$

Use these inverse equations to evaluate $g(K)$:


start here with $g = 2K$

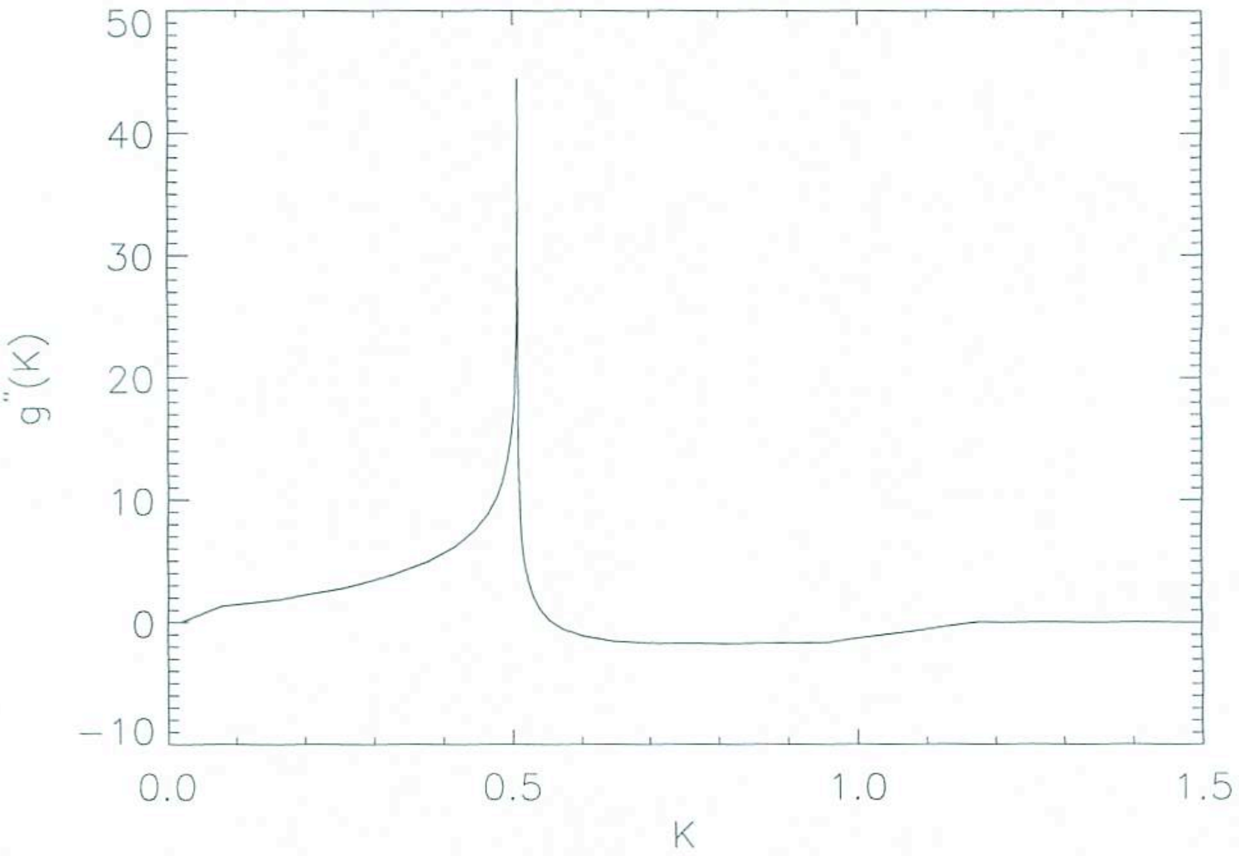
critical point

Start here with $g = \ln 2$



$\frac{dg}{dK}$ calculated using $\frac{dg}{dK} = \frac{g[i] - g[i-1]}{K[i] - K[i-1]}$

$$E = -NJ \frac{\partial g}{\partial K}$$



$\frac{d^2g}{dK^2}$ calculated using $\frac{d^2g}{dK^2} = \frac{g'[i] - g'[i-1]}{K[i] - K[i-1]}$

$C_V = Nk_B \frac{\partial^2 g}{\partial K^2} K^2$

RG solution of 2D Ising model – behavior of C_v near the critical point.

The numerical solution shows a divergence of the heat capacity near the critical point.

[Note that: $E = -\dfrac{\partial \ln Q}{\partial \beta} = -NJ\dfrac{\partial g}{\partial K}$

$$\Rightarrow \quad \frac{C_v}{Nk_B} = K^2 \frac{\partial^2 g}{\partial K^2}$$
]

Assume that the free energy has a non-analytic term

$$g(K) \sim a|K-K_c|^{2-\alpha} \quad \Rightarrow \quad C_v \propto \frac{1}{|K-K_c|^{\alpha}}$$

Near the critical point, expand $K' = K_c + (K-K_c)\left.\dfrac{dK'}{dK}\right|_{K=K_c}$

Then

$$\underbrace{g(K')}_{} = \underbrace{2g(K) - \ln[2\cosh^{1/2} 2K \cosh^{1/8} 4K]}_{}$$

$$a\left|(K-K_c)\left.\frac{dK'}{dK}\right|_{K=K_c}\right|^{2-\alpha} = 2a|K-K_c|^{2-\alpha}$$

$$\Rightarrow \quad 2 = \left(\left.\frac{dK'}{dK}\right|_{K=K_c}\right)^{2-\alpha}$$

$$\Rightarrow \quad \boxed{\alpha = 2 - \frac{\ln 2}{\ln \left.\frac{dK'}{dK}\right|_{K=K_c}}}$$

Using the relation $K' = \frac{1}{4}\ln(\cosh 4K) \Rightarrow \left.\dfrac{dK'}{dK}\right|_{K_c} = \dfrac{3}{2}\tanh 4K_c$

and $K_c = 0.506981 \quad \Rightarrow \quad \boxed{\alpha = 0.131} \longrightarrow$ the exact 2D solution gives $C_v \propto -\ln|T-T_c|$

Monte Carlo numerical solution of the Ising model

The idea is to construct a sample of microstates from the Ising model that can be used to calculate observables. The total number of microstates is a huge number 2^N – eg. even for $N = 100$ (10×10 array of spins) then $\Omega = 2^{100} \approx 10^{30}$. Instead, we focus on those states that are statistically most important by choosing microstates with a probability given by their Boltzmann factor

$$\underset{\substack{\nearrow \\ \text{probability of} \\ \text{choosing state } i \text{ in our} \\ \text{sample of } N_{sample} \text{ microstates}}}{p_i} = \frac{N_i}{N_{samples}} = e^{-\beta E_i}$$

Then observable M for example is calculated using

$$\langle M \rangle = \frac{\sum_\nu M_\nu e^{-\beta E_\nu}}{\sum_\nu e^{-\beta E_\nu}} \simeq \frac{1}{N_{samples}} \sum_{j=1}^{N_{samples}} M_j .$$

A commonly used method for doing this is the Metropolis algorithm (Metropolis et al. 1953). Start with some initial configuration of spins, then generate a new microstate by:

1) Choose a spin at random
2) calculate the change in energy on flipping that spin ΔE

3) if $\Delta E < 0$ then flip the spin
if $\Delta E > 0$ flip the spin with probability $e^{-\beta \Delta E}$
(otherwise the new microstate is the same as the old one.)

We repeat this procedure to generate a sequence of microstates. The occurrence rate of a given microstate will be proportional to its Boltzmann factor $e^{-\beta E_i}$.

To see why this works, write down the transition probability $W_{\mu\nu}$ ← probability of going from state μ to ν.

The values of $W_{\mu\nu}$ satisfy $\sum_\nu W_{\mu\nu} = 1$.

They do not depend on the previous states visited – the probability of jumping to state ν depends only on the current state μ (a "Markov" process).

The probability of choosing microstate μ is p_μ, and satisfies

$$\frac{dp_\mu}{dt} = -\sum_\nu p_\mu W_{\mu\nu} + \sum_\nu p_\nu W_{\nu\mu} \qquad \left(\text{"Master equation"}\right)$$

In equilibrium, we have <u>detailed balance</u> $\boxed{p_\mu W_{\mu\nu} = p_\nu W_{\nu\mu}}$

Therefore given that in equilibrium $\frac{p_\nu}{p_\mu} = \exp\left(-\beta(E_\nu - E_\mu)\right)$

we must choose $\boxed{\frac{W_{\mu\nu}}{W_{\nu\mu}} = \exp\left(-\beta(E_\nu - E_\mu)\right)}$

which is satisfied by the Metropolis algorithm.

[Early in the course, we used a similar argument to motivate the principle of apriori equal probabilities – the argument there is that quantum mechanics gives $W_{\mu\nu} = W_{\nu\mu}$ and therefore $p_\mu = p_\nu$ in equilibrium.]

Given a chain of microstates generated by the Metropolis algorithm, the mean magnetization M and energy E can be calculated. The susceptibility χ and heat capacity C_V can be obtained by also calculating $\langle M^2 \rangle$ and $\langle E^2 \rangle$, and using

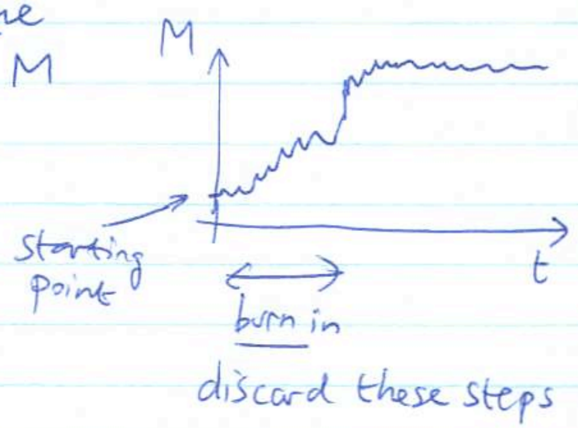
$$C_V = \frac{\beta^2}{N} \left(\langle E^2 \rangle - \langle E \rangle^2 \right)$$

$$\chi = \beta \left(\langle M^2 \rangle - \langle M \rangle^2 \right) .$$

You will get a chance to implement the Metropolis algorithm in the homework: Some points to note are

1) in general the starting microstate will not be from the equilibrium sample and so the algorithm has an equilibriation or "burn-in" time

eg. plot M



2) Some tricks to make the implementation more efficient:

- the Boltzmann factors $e^{-\beta \Delta E}$ can be calculated in advance since ΔE only takes a few values

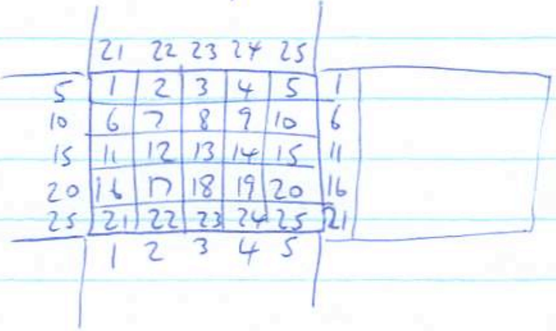
$$\Delta E = 2 s_{ij} \left(s_{i-1,j} + s_{i+1,j} + s_{i,j-1} + s_{i,j+1} \right).$$

$$= 0, \pm 4, \pm 8$$

- rather than calculate M and E for each microstate by summing over the lattice each time, you can just update $E \to E + \Delta E$, $M \to M + \Delta M$.

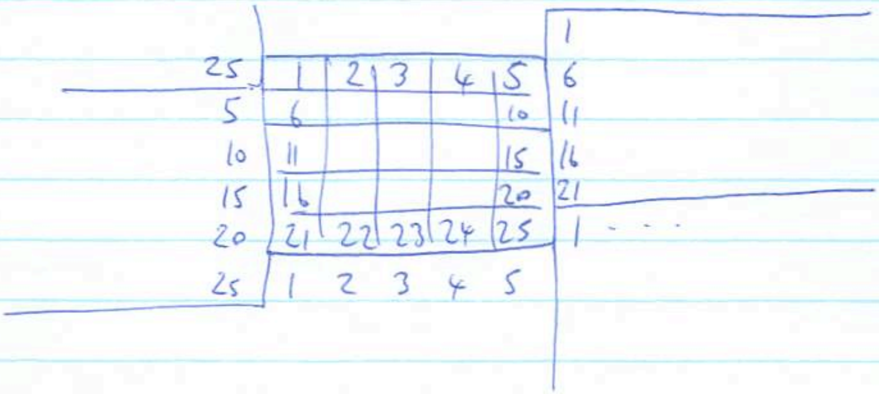
- the burn-in time can be reduced by using the final microstate from the previous temperature as the initial microstate for a neighboring value of temperature.

3) boundary conditions: obvious boundary conditions to use are periodic which we can represent like this:



	21	22	23	24	25	
5	1	2	3	4	5	1
10	6	7	8	9	10	6
15	11	12	13	14	15	11
20	16	17	18	19	20	16
25	21	22	23	24	25	21
	1	2	3	4	5	

An alternative is helical b.c.'s:



						1
25	1	2	3	4	5	6
5	6				10	11
10	11				15	16
15	16				20	21
20	21	22	23	24	25	1 . . .
25	1	2	3	4	5	

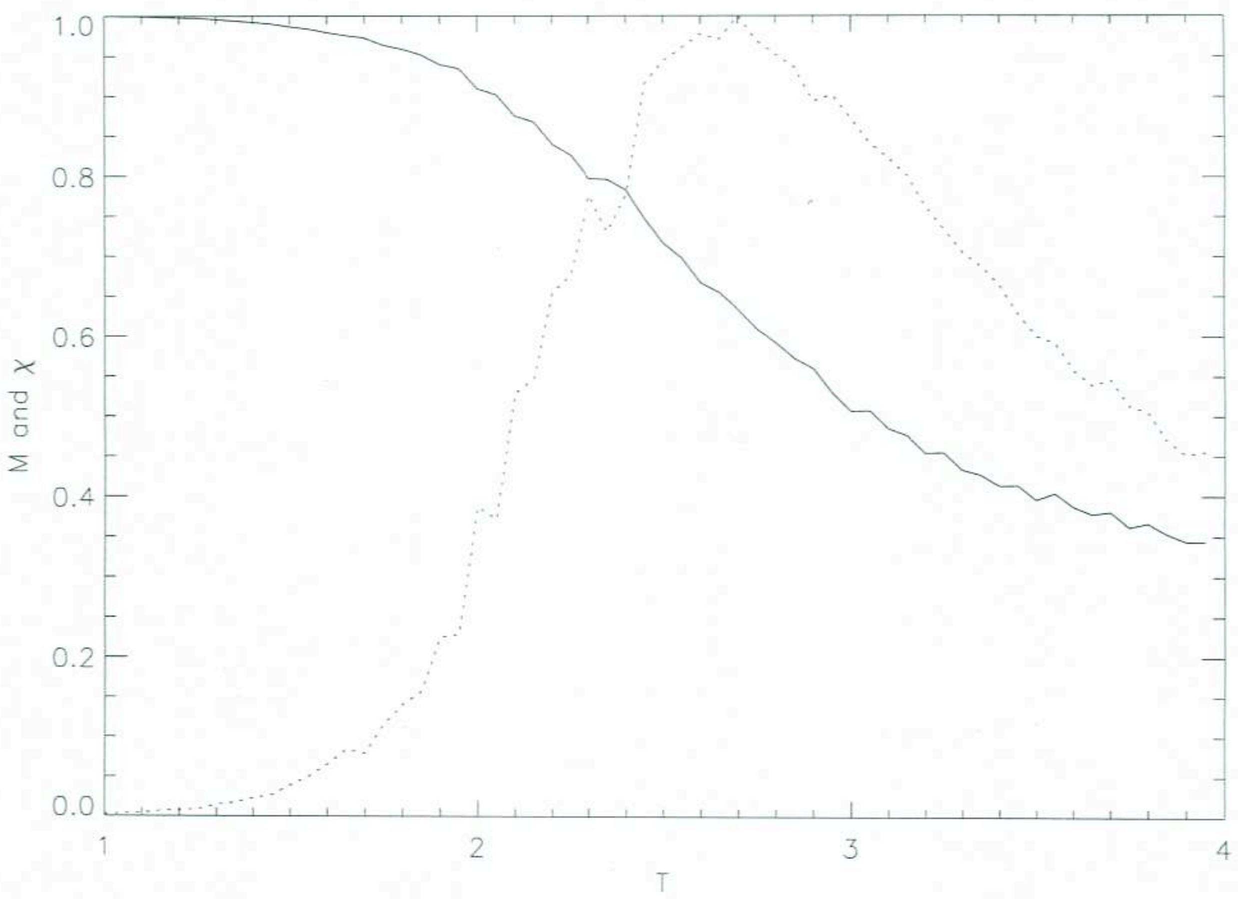
This makes indexing the nearest neighbors simpler.

4) to implement step 3 "flip the spin with probability $e^{-\beta \Delta E}$", you can simply choose a uniformly-distributed random number x (between 0 and 1) and flip the spin if $x < e^{-\beta \Delta E}$.

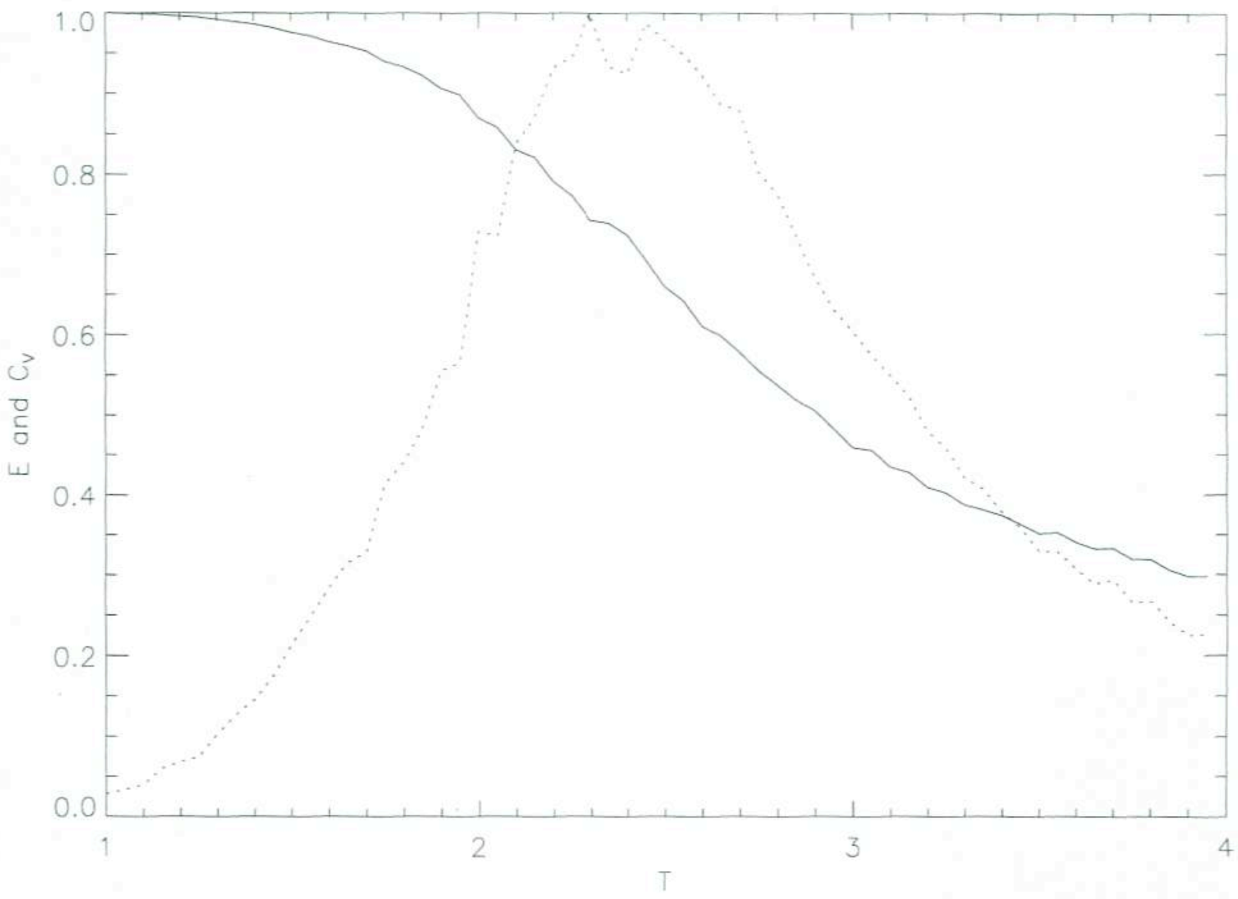
More on point 1 (equilibriation) – it helps here that we know what solution we're looking for. In more complex problems it is not always clear whether the chain has reached equilibrium. ~~One technique is~~ or for example is stuck in a local minimum. One technique is to run multiple chains from different starting points and compare the answers.

Metropolis algorithm results

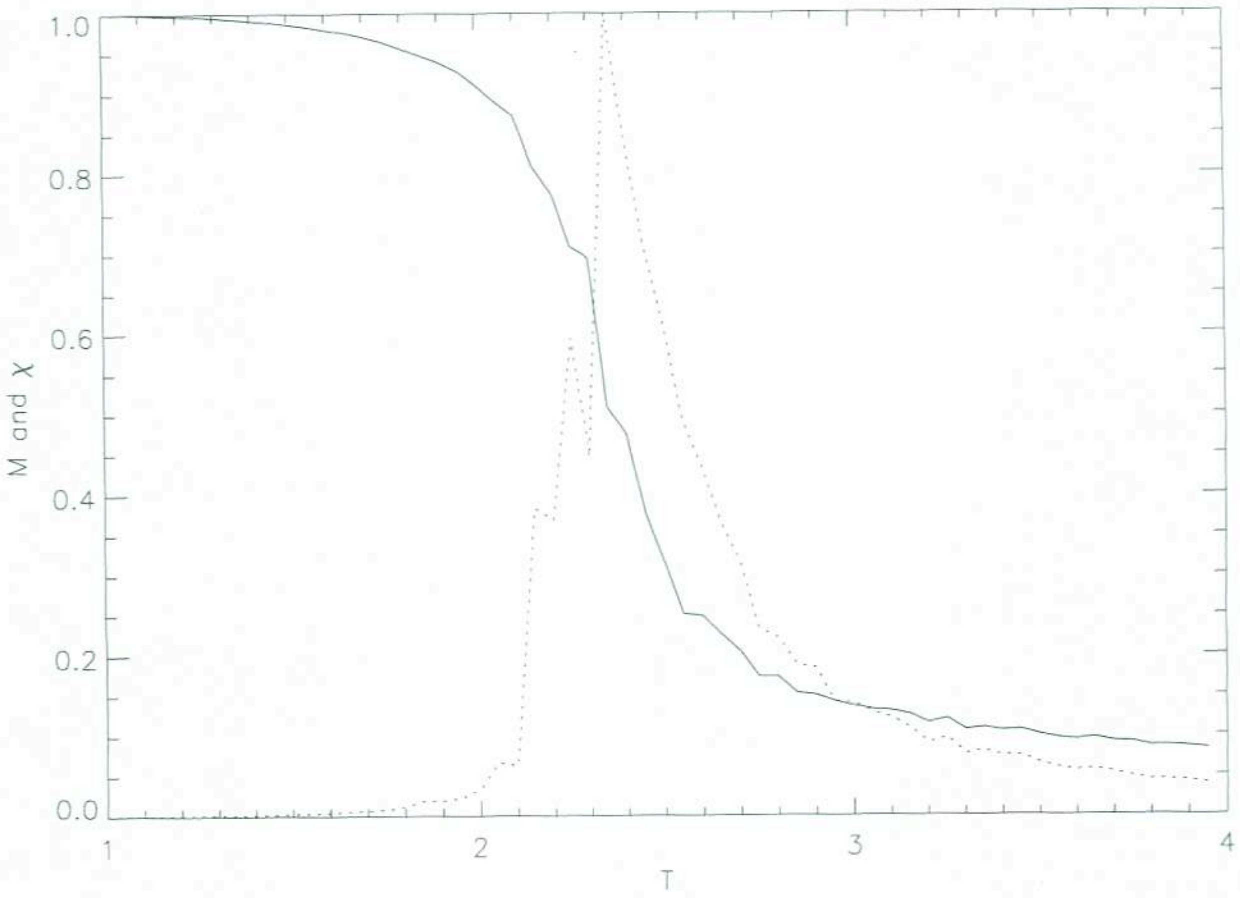
5×5 grid



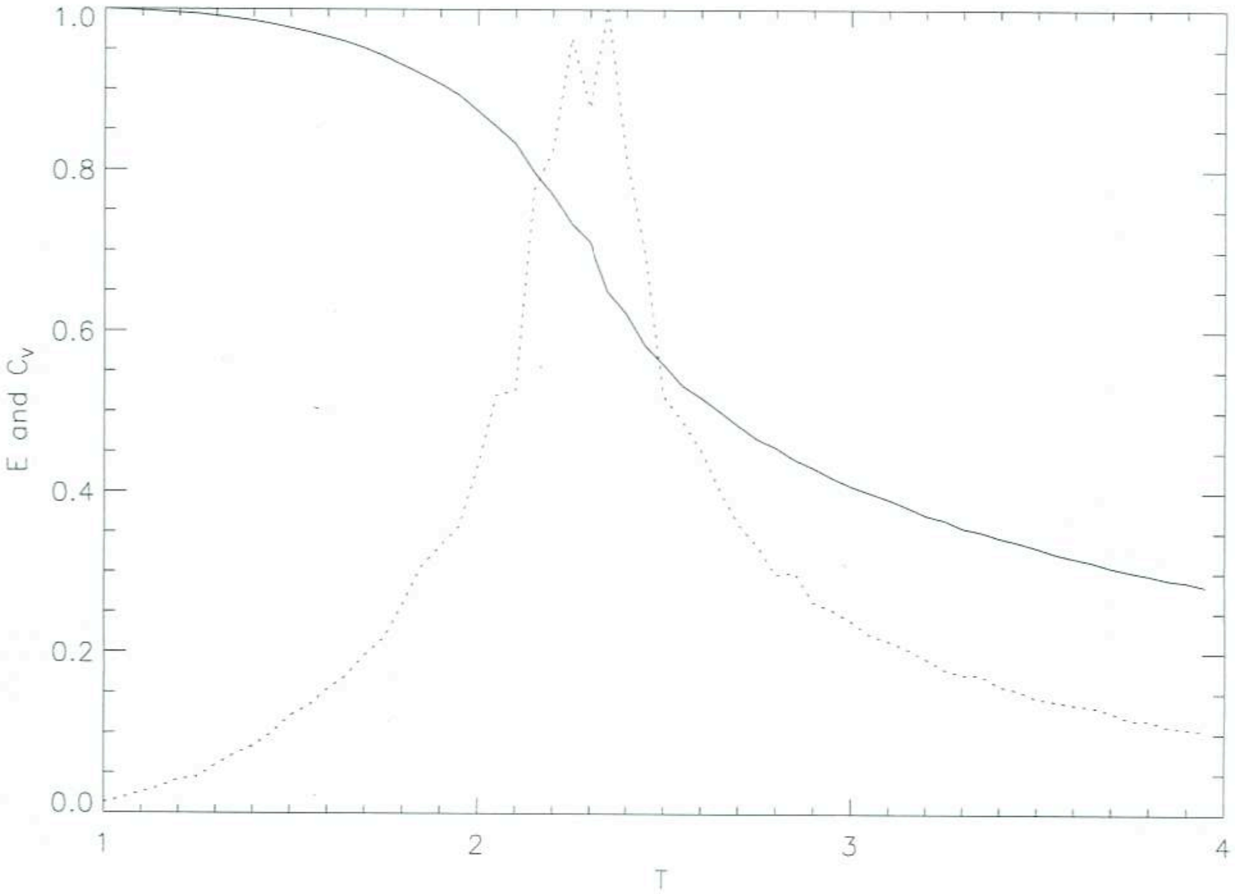
5x5 grid



20×20 grid



20 x 20 grid



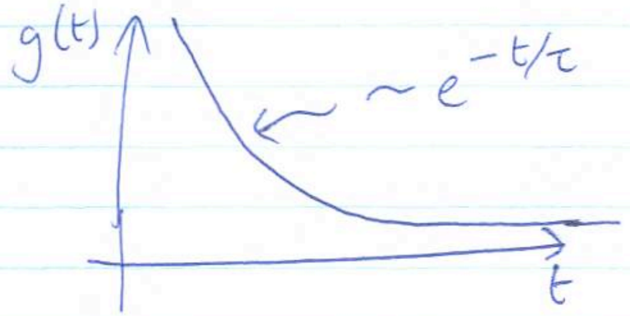
Autocorrelation function

You will notice that there are correlations in for example the values of M_i between different states in the chain. To quantify this, we can calculate the autocorrelation function

$$g(t) = \int dt' \left(M(t') - \langle M \rangle\right)\left(M(t'+t) - \langle M \rangle\right)$$

$$= \int dt' \left(M(t') M(t'+t) - \langle M \rangle^2 \right)$$

typically this will look like



τ = correlation time
(ie number of steps)

The number of independent samples is $\sim \frac{N_{samples}}{\tau}$

which is the number that should be used to calculate the errors in derived quantities, eg. $\langle M \rangle \pm \frac{\sigma_M}{\sqrt{N_{samples}/\tau}}$.

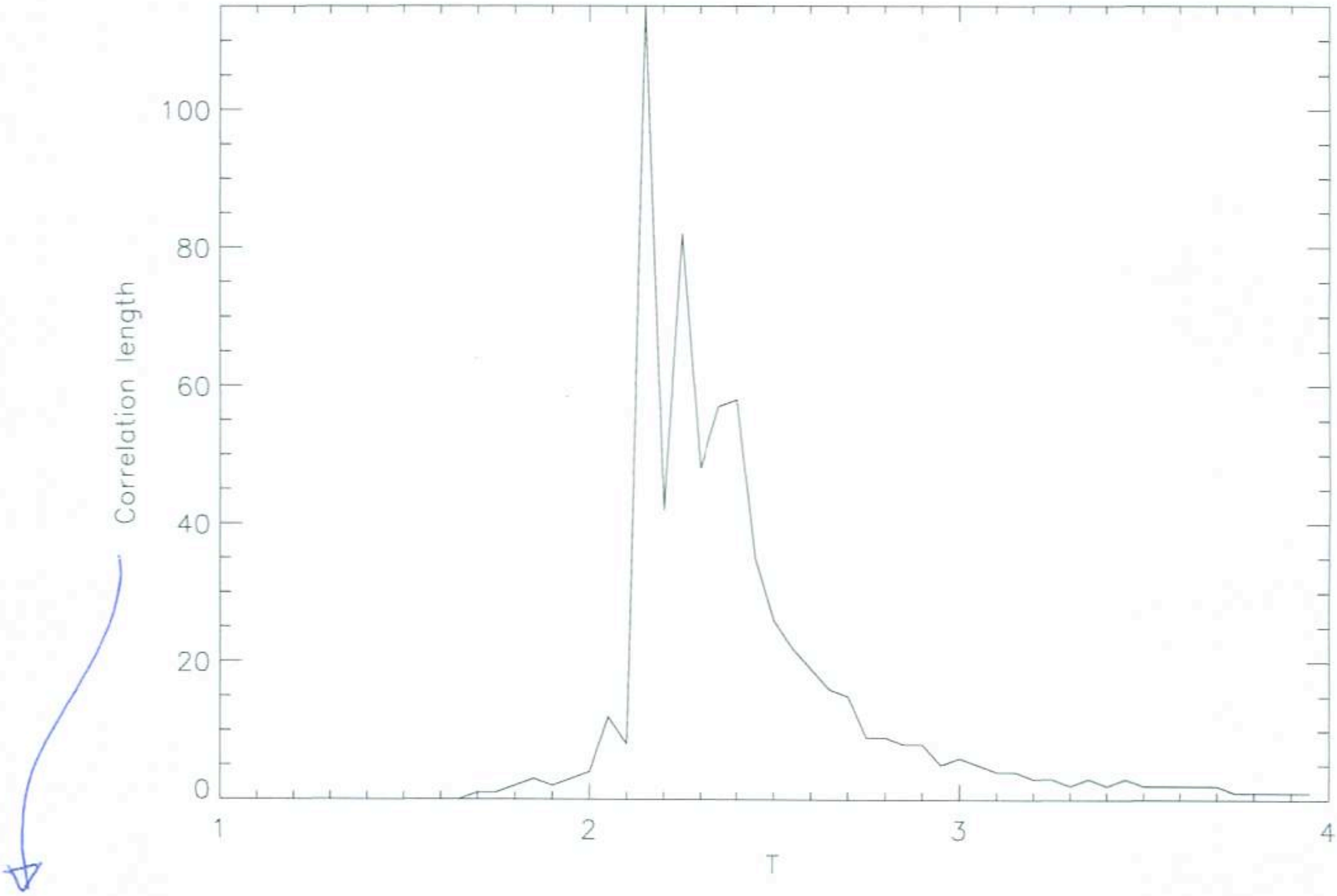
where $\sigma_M^2 = \langle M^2 \rangle - \langle M \rangle^2$.

I've included a plot of ~~g &~~ τ against T for a 20×20 Ising model, where τ is calculated as the value for which $g(\tau) = g(0)/e$. We see that τ increases sharply near $T = T_c$. This is because there are large domains in which all spins are aligned, and it is hard to flip these one spin at a time. For example, the acceptance probability for a flip

$$\begin{array}{ccc} & \uparrow & \\ \uparrow & \uparrow & \uparrow \\ & \uparrow & \end{array} \longrightarrow \begin{array}{ccc} & \uparrow & \\ \uparrow & \downarrow & \uparrow \\ & \uparrow & \end{array}$$

is $e^{-8J/T_c} = 0.03$.

20 × 20 grid



the value of τ where $\frac{g(\tau)}{g(0)} = \frac{1}{e}$

τ is measured in units of n^2

10,000 × (20 × 20) trials at each T

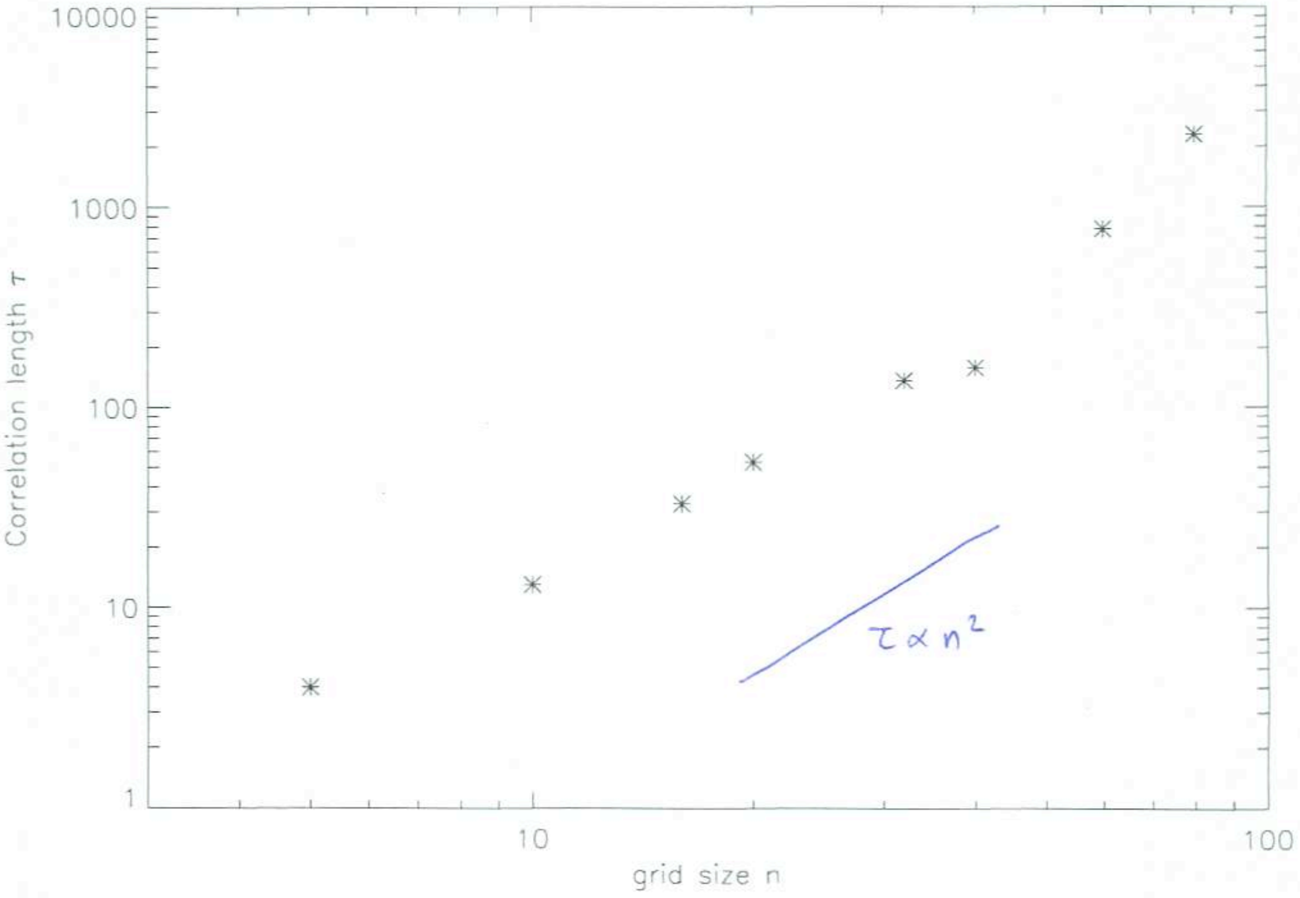
the first 1000 × 20 × 20 are discarded.

The next plot shows how τ increases with grid size, roughly $\tau \propto n^2$ for the Metropolis algorithm (here we take $T = 2.3\, J/k_B$).

Note that we are measuring time in time per spin, so the number of iterations is actually scaling (# iterations between independent measurements) $\propto n^4$!

There are much better algorithms for investigating the behavior near the critical point, eg. the Wolff algorithm in which clusters of aligned spins are flipped all in one go. For this algorithm the scaling is $\tau \underset{\sim}{\propto} n^{1/4}$.

A good place to look for more information is the book "Monte Carlo Methods in Statistical Physics" by Newman and Barkema.



τ is the value at which the autocorrelation function $g(\tau)$ has dropped to $\frac{1}{e}$ of $g(0)$.

Roughly, the scaling is $\tau \propto n^2$

For each grid size, we run 10,000 iterations $(\times n^2)$
Temperature is fixed at 2.3

For example at $n = 10$ $\tau \sim 10 \Rightarrow 1000$ independent measurements $\Rightarrow$ we can get a few % measurement of $\langle M \rangle$. But at $n = 60$, $\tau \approx 80$ which implies only 12 independent measurements $\Rightarrow$ 30% error in $\langle M \rangle$.

(ie. error in $\langle M \rangle$ $= 30\% \times \sigma_M$)

Spatial correlation function

We've already discussed the fact that long-range spatial correlations arise close to the critical point.

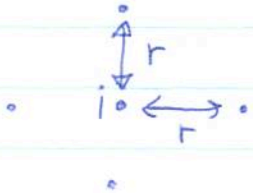
We can measure this using the correlation function

$$G(r) = \langle (s_i - \langle s_i \rangle)(s_j - \langle s_j \rangle) \rangle$$
$$= \langle s_i s_j \rangle - \langle s \rangle^2$$

since all spins are equivalent, the correlation function depends only on the distance between spins.

Typically $G(r) \sim e^{-r/\xi}$ where we define the <u>correlation length</u> ξ

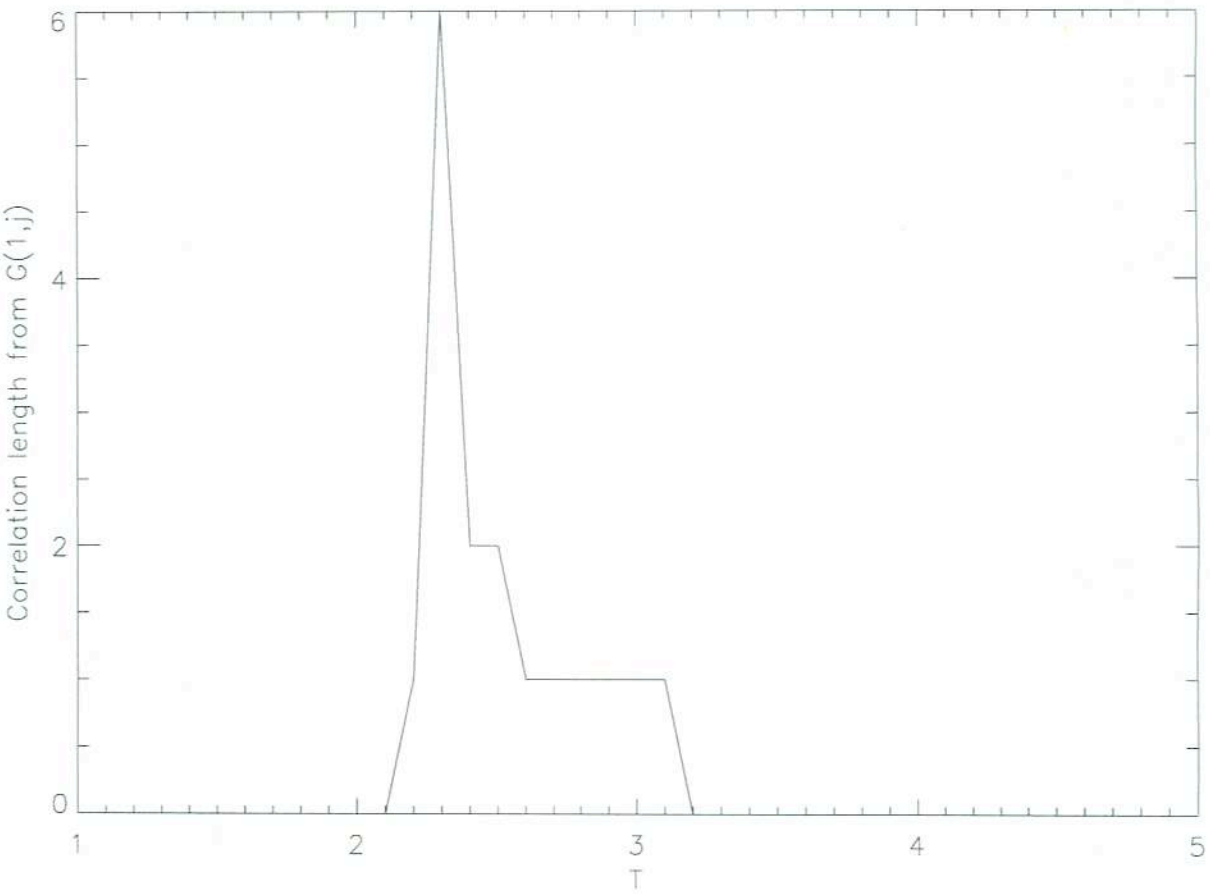
In my code, I calculate $G(r)$ by setting $i=1$ (1st spin) and averaging over the 4 spins j a distance r away



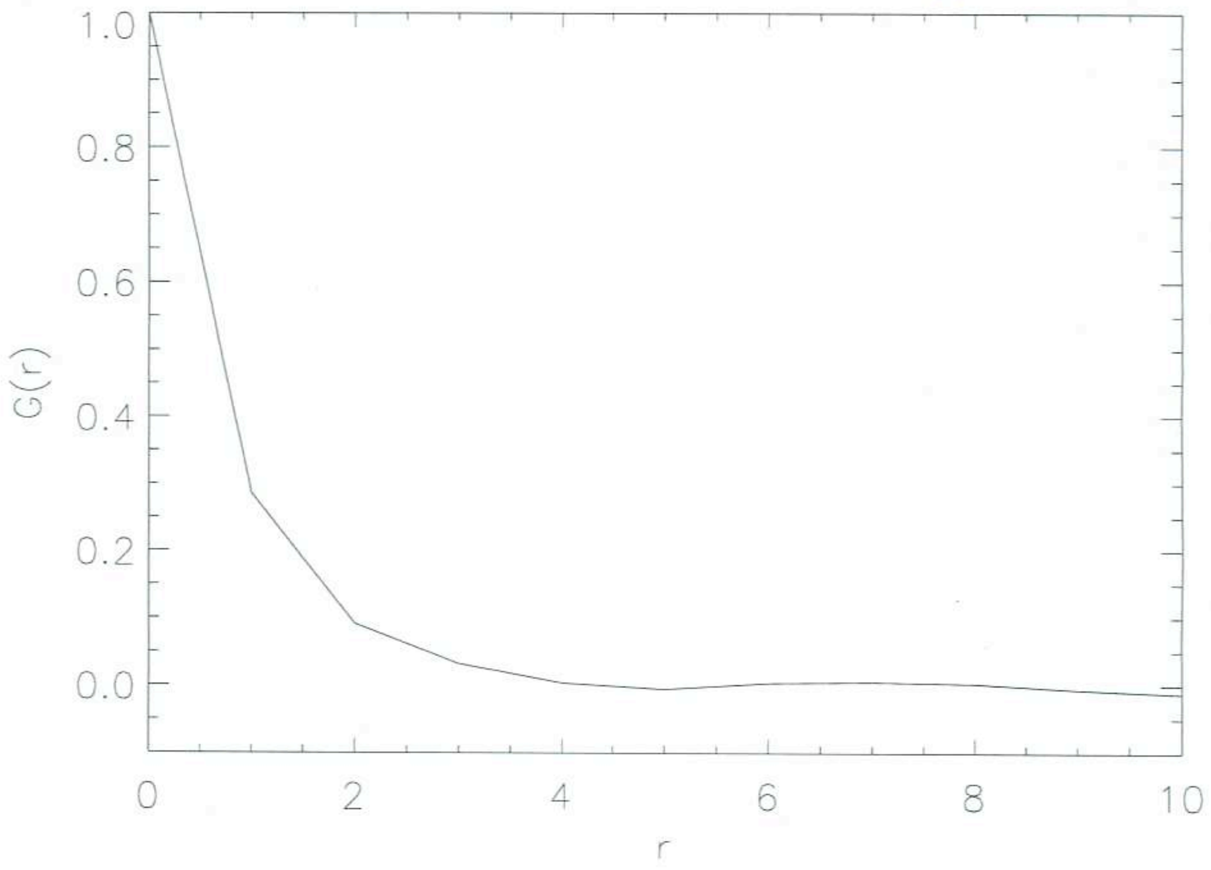
A crude measure of ξ is the value of r where $G(r)/G(0) = \frac{1}{e}$.

I show this as a function of T in the following plot – note that ξ increases in integer steps, (because of the way I've defined it.) You can see the dramatic increase in ξ for $T \simeq T_c$.

20×20 grid

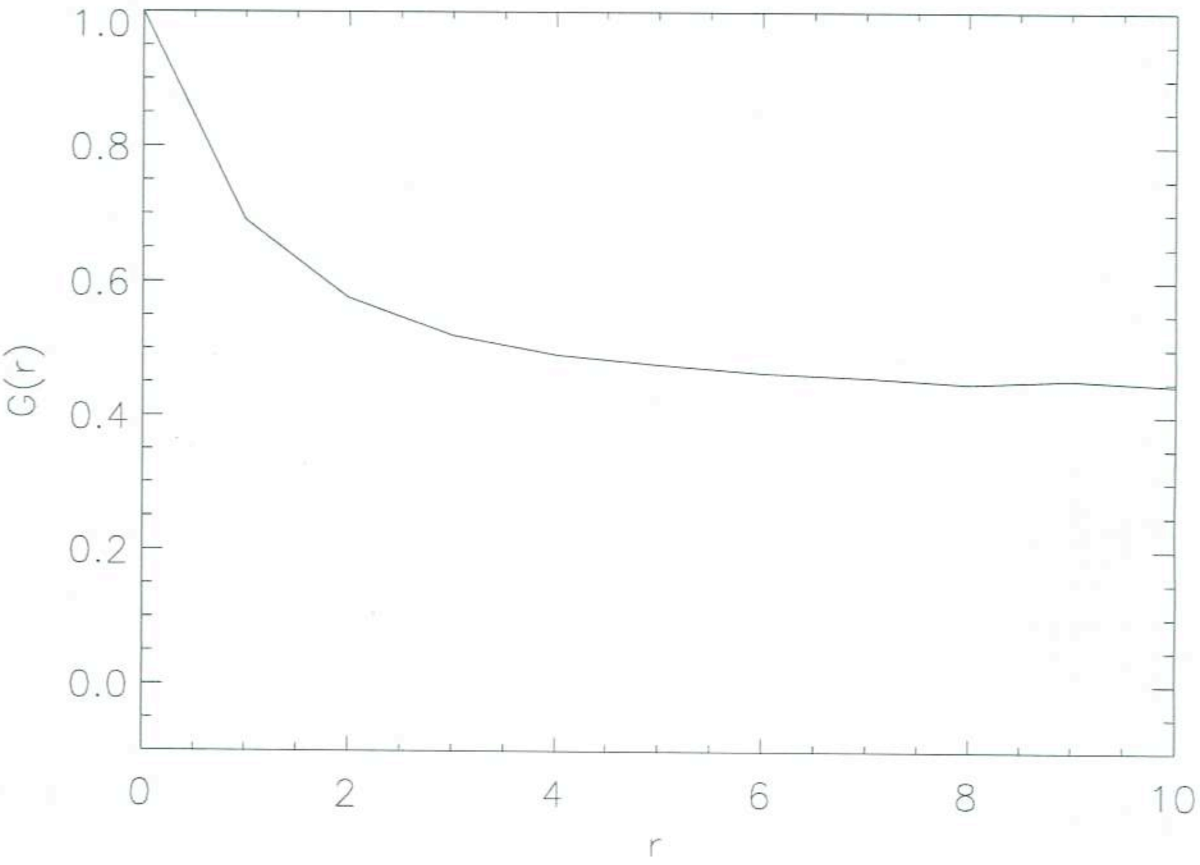


$T = 4.0\ J/k_B$



20x20

$T = 2.3\ J/k_B$



20 x 20

We already saw that there is a relation between χ and the fluctuations in M:

$$\chi = \beta \sigma_M^2 = \beta\left(\langle M^2 \rangle - \langle M \rangle^2 \right).$$

There is also a connection to $G(r)$:

$$\chi = \beta\mu^2 \left\langle \sum_{i=1}^{N} (s_i - \langle s \rangle) \sum_{j=1}^{N} (s_j - \langle s \rangle) \right\rangle$$

$$= \beta\mu^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \underbrace{\langle (s_i - \langle s \rangle)(s_j - \langle s \rangle) \rangle}_{\text{this is the two point correlation function.}}$$

Since all spins are equivalent we can write

$$\chi = \beta\mu^2 N \underbrace{\sum_{j=1}^{N} \langle (s_1 - \langle s \rangle)(s_j - \langle s \rangle) \rangle}_{\sim \text{ sum over correlations with spin 1}}$$

So we see that

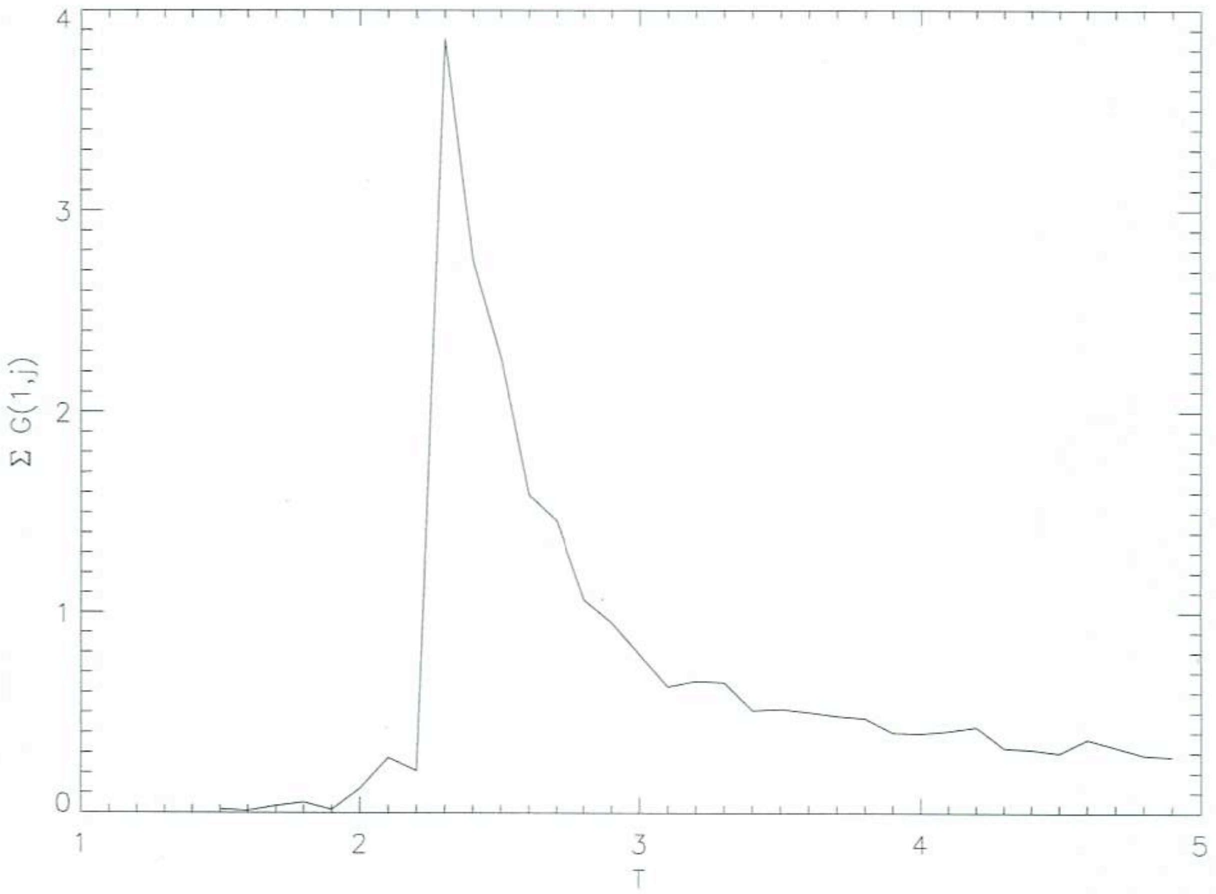
fluctuations ⟷ susceptibility ⟷ ~~long range order~~ spatial correlations

at $T \approx T_c$

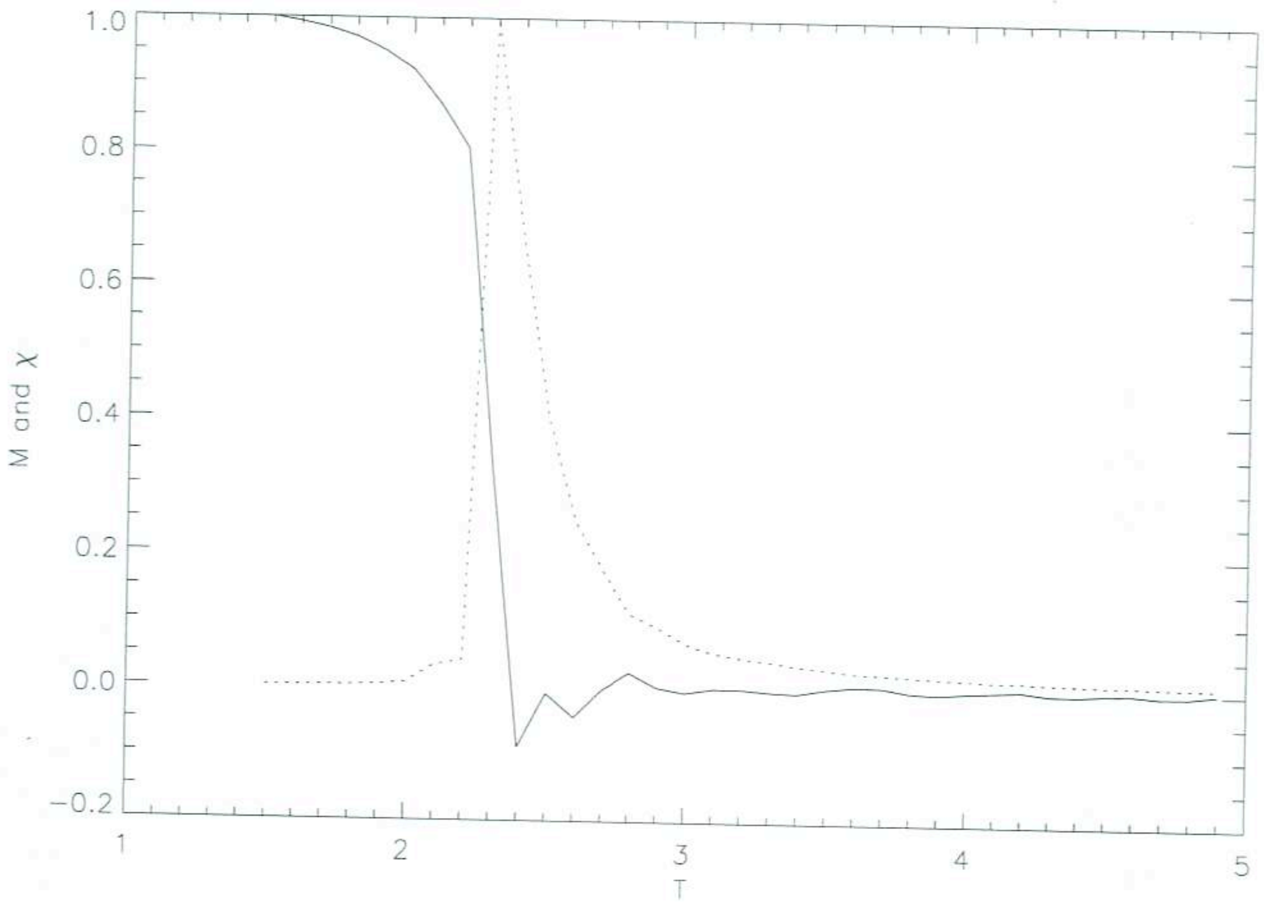
large fluctuations ⟷ diverging χ ⟷ long range order

The plot overleaf shows $G(r)$ summed over separations r as a function of temperature. Note the similarity to the following plot which shows χ as a function of T.

20 x 20 grid



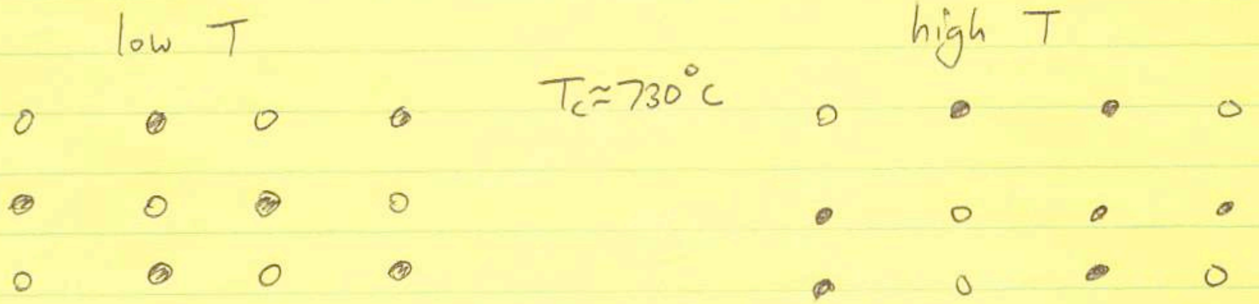
20 x 20 grid



This time I haven't taken the absolute value of M.

Binary alloys

eg. β-brass is a roughly 50/50 mixture of Cu and Zn. It has a phase transition between an ordered phase at low T in which Cu and Zn atoms occupy alternate sites on a cubic lattice, and a high temperature, disordered phase in which the sites are randomly occupied by Cu or Zn.



We model this by writing different energies for the 3 different kinds of nearest neighbour pairs

$$\boxed{E = -E_{AA}\underset{\text{eg. number of A-A nearest neighbor pairs}}{\underbrace{N_{AA}}} - E_{BB}N_{BB} - E_{AB}N_{AB}} \qquad -(*)$$

It turns out that this is the Ising model in disguise, but with an extra (constant) term in the energy to give us 3 degrees of freedom.

To show this, we'll start with the Ising model and show that we can rewrite it in the same form as equation (*).

$$\boxed{E = -J\sum s_i s_j - H\sum s_i - CN}$$

where $s_i = \begin{cases} 1 & \text{if atom A occupies site } i \\ -1 & \text{if atom B occupies site } j \end{cases}$

or in terms of the different N_{AA} values etc,

$$E = -J(N_{AA} + N_{BB} - N_{AB}) - H(N_A - N_B) - CN$$

We can rewrite the N_A and N_B terms using the relations

$$zN_A = 2N_{AA} + N_{AB} \quad \text{(count the links to A atoms)}$$
$$zN_B = 2N_{BB} + N_{AB} \quad \text{('' '' B atoms)}$$

$$\Rightarrow \quad E = -N_{AA}\left(J + \frac{2}{z}(C+H)\right) - N_{BB}\left(J + \frac{2}{z}(C-H)\right) - N_{AB}\left(-J + \frac{2}{z}C\right)$$

Compare this with (*) → we have

$$\left.\begin{aligned} E_{AA} &= J + \frac{2(C+H)}{z} \\ E_{BB} &= J + \frac{2(C-H)}{z} \\ E_{AB} &= -J + \frac{2}{z}C \end{aligned}\right\} \text{solve these} \Rightarrow \boxed{\begin{aligned} H &= \left(\frac{z}{4}\right)(E_{AA} - E_{BB}) \\ C &= \left(\frac{z}{4}\right)\left(E_{AB} + \frac{E_{AA} + E_{BB}}{2}\right) \\ J &= \frac{1}{2}\left(\frac{E_{AA} + E_{BB}}{2} - E_{AB}\right) \end{aligned}}$$

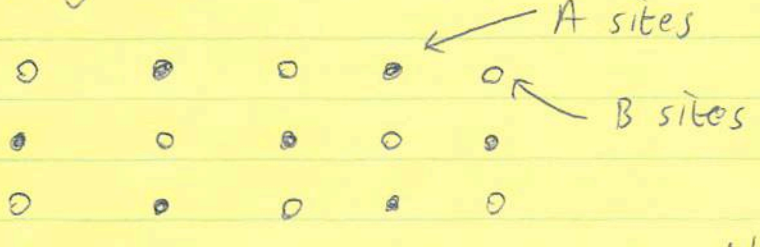
these are the constants in the Ising model that is equivalent to (*).

For β-brass, $J < 0$ or $E_{AB} > \frac{E_{AA} + E_{BB}}{2}$

equivalent to an antiferromagnetic Ising model. [ground state:
↑↓↑↓↑
↓↑↓↑↓]

We can prove a useful result relating antiferromagnetic $(J<0)$ and ferromagnetic $(J>0)$ cases, using the symmetry of the Ising Hamiltonian for zero applied field $(H=0)$.

We divide the lattice into two sub-lattices such that alternate sites belong to each sublattice



then $$E = -J \sum_{i,j} S_i^A \, s_j^B$$

the nearest neighbors to an atom on site A are all from lattice B.

then $$E(-J, \{s^A\}, \{s^B\}) = E(J, \{-s^A\}, \{s^B\}) \quad -(1)$$
$$= E(J, \{s^A\}, \{-s^B\})$$

(ie. reversing the sign of J is equivalent to reversing all the spins on one of the sublattices)

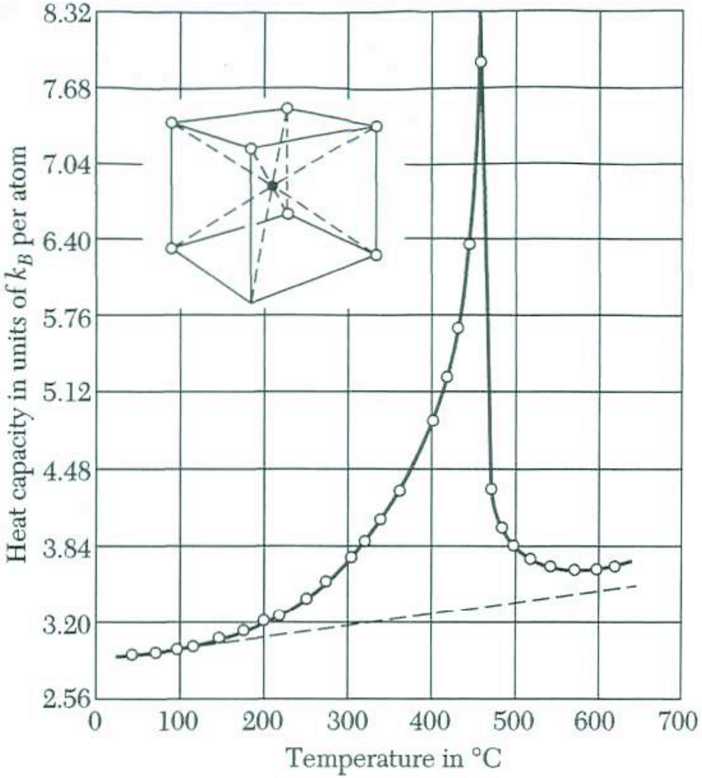
For any function $f(\{s\})$, $$\sum_{\{s=\pm 1\}} f(\{s\}) = \sum_{\{s=\pm 1\}} f(\{-s\}) \quad -(2)$$

$\Rightarrow$
$$Q(-J,T) = \sum_{\{s^A\}} \sum_{\{s^B\}} \exp(-\beta E(-J, \{s^A\} \{s^B\}))$$
$$= \sum \sum \exp(-\beta E(J, \{-s^A\}, \{s^B\})) \quad \text{using (1)}$$
$$= \sum \sum \exp(-\beta E(J, \{s^A\}, \{s^B\})) \quad \text{using (2)}$$
$$= Q(J,T).$$

See attached plots of C_v vs. T for 1) ferromagnet 2) antiferromagnet 3) binary alloy

$\Rightarrow$ the ferromagnetic and antiferromagnetic Ising models ~~should~~ have the same thermal properties at zero field.

Chp 22



9 Heat capacity versus tem-
of CuZn alloy (β-brass).

ange order parameter P is defined so that the number of A's on the a is equal to $\frac{1}{2}(1 + P)N$. The number of A's on lattice b is equal to)N. When $P = \pm 1$, the order is perfect and each lattice contains only pe of atom. When $P = 0$, each lattice contains equal numbers of A and B and there is no long-range order.

e consider that part of the internal energy associated with the bond en- of AA, AB, and BB nearest-neighbor pairs. The total bond energy is

The energy (3) becomes

$$E =$$

where

$$E_0 = 2N(U_{AA} + U_{BB} + 2U$$

We now calculate the entrop $\frac{1}{2}(1 + P)N$ atoms A and $\frac{1}{2}(1 - P)N$ atoms A and $\frac{1}{2}(1 + P)N$ atoms B on these atoms is

$$G = \left[\frac{}{[\frac{1}{2}(1 +} \right.$$

From the definition of the entrop approximation,

$$S = 2Nk_B \ln 2 - Nk_B[(1 +$$

This defines the **entropy of mixing**

The equilibrium order is det energy $F = E - TS$ be a minimum differentiating F with respect to P,

$$4NPU +$$

The transcendental equation for smoothly decreasing curve show expand (9) to find $4NPU + 2Nk_BT$ so that

Bleaney & Bleaney

properties should be greatest at this point, followed by a sharp
above the Curie point where M becomes zero.

An experimental curve showing the variation of C with T for
is given in Fig. 21.8. The anomalous specific heat is appreciable
near the Curie point, but the drop above the Curie point spreads
range of temperature, instead of appearing as a sharp discontinuit

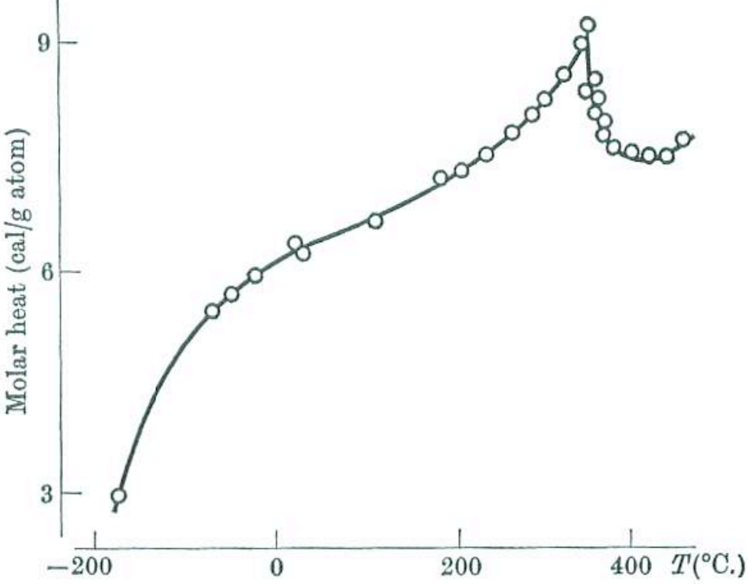


Fig. 21.8. The molar heat of nickel, from the measurements of Grew, 1934.

order to obtain a value for λ from the specific heat anomaly, C_M mu
estimated and subtracted from the measured specific heat, so that
the magnetic contribution remains. Measurements are made at con
pressure, so that we can write $C_M = C_v+(C_p-C_v)+C_e$. C_v is obt
by extrapolation, using the Debye formula, from measurements a
temperatures; (C_p-C_v) may be found from the expansion coefficien
compressibility using a standard thermodynamical formula; C_e i
electronic specific heat. This is abnormally large in a ferromag
metal and difficult to estimate since it is associated with a high ele
density in the 3d band (see § 18.4). dM^2/dT must be found by plo
M^2 as a function of temperature, and then λ is obtained. This is
very accurate method of finding λ and the value does not agree too
with the value obtained from the magnetization curve, probably be
of errors in C_e. However, the general form of the specific heat cu
not incompatible with theory and this also applies to iron and co

Bleaney & Bleaney

22.2. The molecular field—two sub-lattice model

Let the two sub-lattices be denoted by A and B. Then a dipole in lattice A is subject to an external field B_0 and an internal field proportional to the magnetization of sub-lattice B which we may write as $-\lambda M_B$, where the minus sign appears because of the reversed sign of the exchange integral. The effective field acting on a dipole in A is therefore

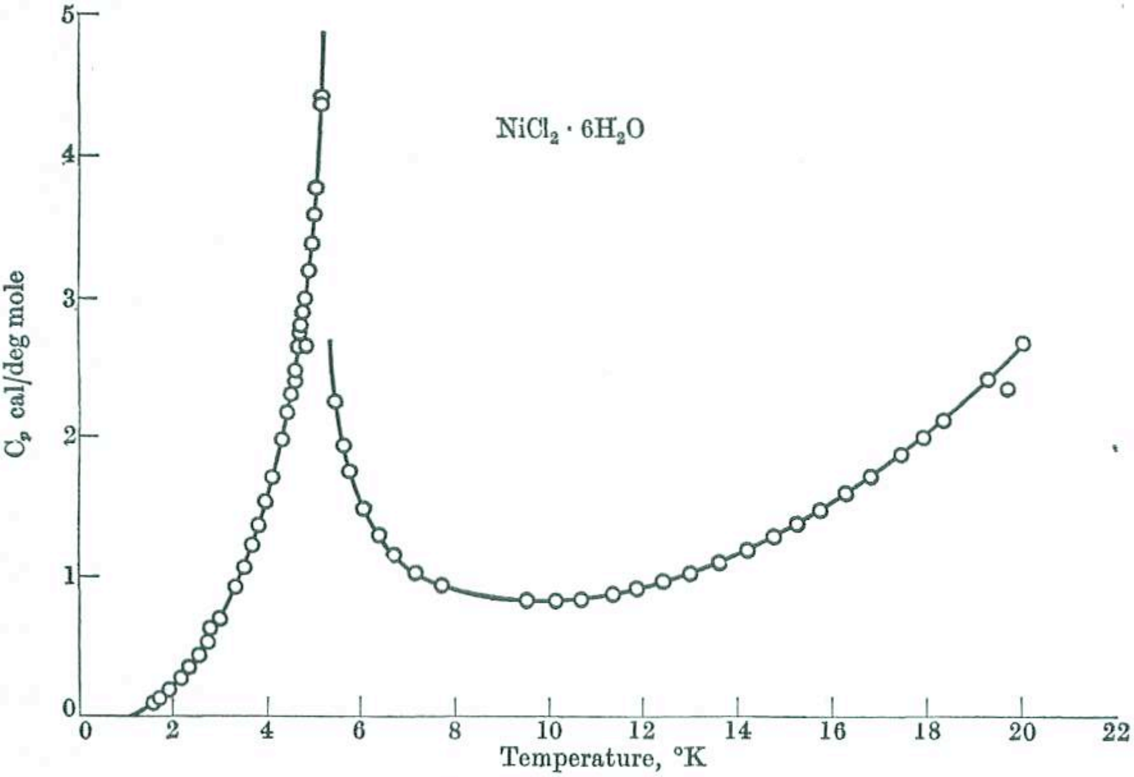


FIG. 22.2. Specific heat of NiCl$_2$,6H$_2$O at low temperatures, showing the λ-type anomaly at the Néel temperature, 5·34° K (after Robinson and Friedberg, 1960). The entropy in the anomaly is $R \ln 3$, corresponding to the threefold degeneracy of the $S = 1$ ground state of the Ni^{++} ion. The rise at high temperatures is due to the lattice specific heat. Note the different shape of the co-operative anomaly from that due to a simple level splitting in another nickel salt (Fig. 20.15).

$$\left.\begin{aligned} B_A &= B_0-\lambda M_B \\ B_B &= B_0-\lambda M_A \end{aligned}\right\}. \qquad (22.1)$$

Similarly

At high temperatures where the dipoles are randomly oriented the magnetization of each sub-lattice should obey Curie's law if we take the effective field instead of the external field. Thus we have

$$M_A = \tfrac{1}{2}CB_A/\mu_0 T, \qquad M_B = \tfrac{1}{2}CB_B/\mu_0 T, \qquad (22.2)$$

Mean field theory for the antiferromagnetic Ising model

We've seen that the binary alloy model we wrote down for β-brass is equivalent to the anti-ferromagnetic $(J<0)$ Ising model. Let's develop a mean field for this. In the context of binary alloys, this is known as the Bragg-Williams approximation after work by these authors in the 1930's.

The total magnetization M is not a good order parameter to use in this case, because $M=0$ in both ordered and disordered states for an antiferromagnet.

↑ ↓ ↑ ↓ ↑   ↑ ↑ ↓ ↑
↓ ↑ ↓ ↑ ↓   ↑ ↓ ↑ ↑

$T<T_c$ $T>T_c$

(the critical temperature T_c is known as the Néel temperature in this case)

To solve this problem, we divide the lattice into two sub-lattices A and B such that atoms on the A lattice only have B atoms as nearest neighbors and vice-versa.

Then $$E = -J\sum_{\text{pairs}} s_i^A s_j^B - H\sum_{\text{A lattice}} s_i^A - H\sum_{\text{B lattice}} s_j^B - \underbrace{CN}_{\text{we need this term for the binary alloy}}$$

Now try a MFT approach:

$$\sum s_i^A s_j^B = \sum \left(s_i^A - \langle s^A\rangle + \langle s^A\rangle\right)\left(s_j^B - \langle s^B\rangle + \langle s^B\rangle\right)$$

$$= \cancelto{0}{\sum (s_i^A - \langle s_A\rangle)(s_j^B - \langle s^B\rangle)} \text{ throw away this term} + \langle s^A\rangle (s_j^B - \langle s_B\rangle)$$

$$\ldots + \langle s^{B}_{\cancel{B}} \rangle \left(s_i^A - \langle s^A \rangle \right) + \langle s^A \rangle \langle s^B \rangle$$

$$= \sum_{i,j} \langle s^A \rangle s_j^B + \sum_{i,j} \langle s^B \rangle s_i^A - \sum_{i,j} \langle s^A \rangle \langle s^B \rangle$$

$$= z \langle s^A \rangle \sum_{\text{B lattice}} s_j^B + \langle s^B \rangle z \sum_{\text{A lattice}} s_i^A - \frac{Nz}{2} \langle s^A \rangle \langle s^B \rangle$$

⇒ the energy is

$$E = - \sum_{\text{A lattice}} s_i^A \left(Jz \langle s^B \rangle + H \right) - \sum_{\text{B lattice}} s_i^B \left(Jz \langle s^A \rangle + H \right)$$
$$+ \frac{zNJ}{2} \langle s^A \rangle \langle s^B \rangle - CN$$

The A and B lattices contribute separate terms to the energy in this approximation ⇒ the partition function factorizes

$$Q = \left(e^{-\beta \left[\frac{zJ}{2} \langle s^A \rangle \langle s^B \rangle - C \right]} \right)^N \left(2 \cosh \left(\beta H + \beta z J \langle s^A \rangle \right) \right)^{N/2}$$
$$\left(2 \cosh \left(\beta H + \beta z J \langle s^B \rangle \right) \right)^{N/2}$$

⇒ $A = -kT \ln Q$ is

$$\frac{A}{NkT} = \frac{1}{kT} \left[\frac{zJ}{2} \langle s_i^A \rangle \langle s^B \rangle - C \right] - \frac{1}{2} \ln 2 \cosh \left(\beta H + \beta J z \langle s^A \rangle \right)$$
$$- \frac{1}{2} \ln 2 \cosh \left(\beta H + \beta J z \langle s^B \rangle \right)$$

To find $\langle s^A \rangle$ and $\langle s^B \rangle$, minimize A with respect to each of these.

eg. $$\frac{\partial A}{\partial \langle s_i^A \rangle} = 0 = - \frac{\beta J z}{2} \tanh \left(\beta H + \beta z J \langle s^A \rangle \right) + \frac{zJ}{2kT} \langle s^B \rangle$$

$$\Rightarrow \quad \langle s^B \rangle = \tanh \left(\beta H + \beta z J \langle s^A \rangle \right)$$

and vice vessa, so we find that $\langle S^A \rangle$ and $\langle S^B \rangle$ are given by

$$\langle S^A \rangle = \tanh\left(\beta H + \beta z J \langle S^B \rangle \right)$$

$$\langle S^B \rangle = \tanh\left(\beta H + \beta z J \langle S^A \rangle \right)$$

For $J > 0$ (ferromagnetic) the solution is $\langle S^A \rangle = \langle S^B \rangle = \langle S \rangle$

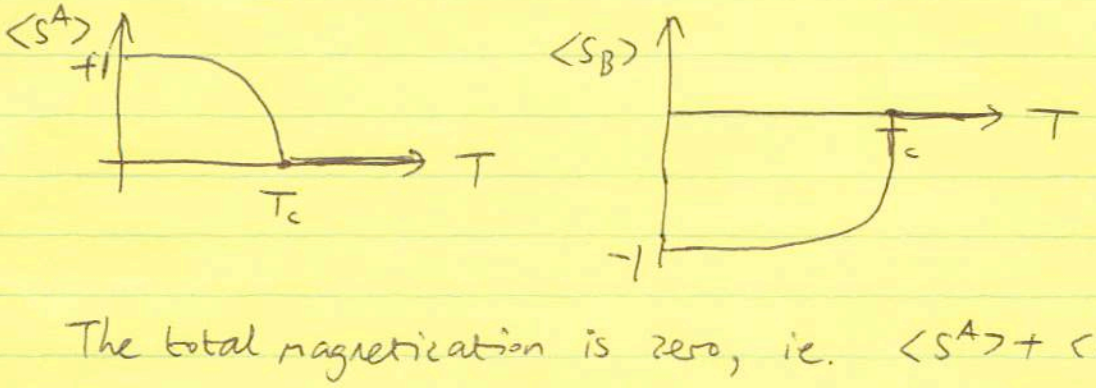
where $\langle S \rangle = \tanh(\beta H + \beta z J \langle S \rangle)$
as before

For $J < 0$ (antiferromagnetic) the solution is for $H = 0$

$$\langle S_A \rangle = - \langle S_B \rangle = \langle S \rangle$$

where

$$\langle S \rangle = \tanh\left(\beta z |J| \langle S \rangle \right)$$



The total magnetization is zero, ie. $\langle S^A \rangle + \langle S^B \rangle = 0$

What is the solution for finite H? Lets calculate the magnetization $M = \langle S^A \rangle + \langle S^B \rangle$ for small H. Since $\chi = \frac{\partial M}{\partial H}$

then the magnetization for small H will give us the susceptibility χ, ie. $M = \chi H$ for small H and M.

$h = \beta H$
$j = \beta z J$

When $H=0$ $\langle s_A \rangle = s$, $\langle s_B \rangle = -s$

then $$s + \delta s_A = \tanh\left(h + j(-s + \delta s_B)\right)$$

$$-s + \delta s_B = \tanh\left(h + j(s + \delta s_A)\right)$$

we want $$\chi = \left(\frac{\delta s_A + \delta s_B}{h}\right) \times (\beta) \times \left(\frac{N}{2}\right)$$

expand: $$\cancel{s} + \delta s_A \simeq \underbrace{\tanh(-js)}_{\cancel{s}} + (h + \delta s_B j)\underbrace{(1 - \tanh^2(-js))}_{1-s^2}$$

$$\Rightarrow \boxed{\delta s_A = (1-s^2)(h + j\delta s_B)}$$

Similarly $$-\cancel{s} + \delta s_B = \underbrace{\tanh(js)}_{-\cancel{s}} + (h + j\delta s_A)\underbrace{(1 - \tanh^2 js)}_{1-s^2}$$

$$\Rightarrow \boxed{\delta s_B = (1-s^2)(h + j\delta s_A)}$$

Solve these by adding them

$$\delta s = \delta s_A + \delta s_B = (1-s^2)(2h + j\,\delta s)$$

$$\delta s\left[1 - j(1-s^2)\right] = (1-s^2)\,2h$$

Subtracting we find that $\delta s_A - \delta s_B = 0$
$\Rightarrow \delta s_A = \delta s_B = \frac{\delta s}{2}$
So each magnetization shifts by the same amount.

$$\Rightarrow \chi = \frac{2\beta(1-s^2)}{1 - j(1-s^2)}\,\frac{N}{2} = \frac{\cancel{2}\beta N}{\frac{1}{1-s^2} - j}$$

or $$\boxed{\chi = \frac{N \cancel{2\beta} \frac{1}{kT}}{\frac{1}{1-s^2} - \frac{zJ}{kT}}}$$

~~for $J<0$ the denominator is $\frac{1}{1-s^2} + \frac{z|J|}{kT}$~~

This is the same as we found in the HW for $J > 0$

$$\boxed{\chi = \frac{N/kT}{\frac{1}{1-s^2} - \frac{T_c}{T}}} \qquad T_c = zJ/k_B$$

for $J < 0$ we can write $T_c = \frac{z|J|}{k_B}$

and
$$\boxed{\chi = \frac{N/kT}{\frac{1}{1-s^2} + \frac{T_c}{T}}} \text{ — } (*)$$

(we've dropped factors of μ needed to get the units right ie. $M = N\mu\langle s\rangle$)

Plot it:

This agrees well with observations — see over for data for MnF_2, an antiferromagnetic material. (Fig 22.6 from Bleaney & Bleaney).

You'll see in this plot that the $\chi_{\parallel}$ looks like eq. (*)
But $\chi_{\perp} \simeq$ constant below T_c — in real materials we have to worry about directions.
The $\parallel$ and $\perp$ refer to the direction between the magnetization and the applied field. In our model the H and M are automatically aligned.

Bleaney & Bleaney Electricity & Magnetism

The two sub-lattice model is valid for many anti-ferromagnetic substances, but in some cases there are more (in a face-centred cubic lattice there are generally four). As in ferromagnetism, the exchange interaction itself gives no preference to any particular orientation of the spins relative to the crystal axes; this arises from the anisotropy energy. In a simple tetragonal crystal such as MnF_2, the spins are aligned along the tetragonal axis in a simple two sub-lattice anti-parallel arrangement, but much more complicated arrangements are possible in which the vector sum of the dipole moments is zero—the distinctive feature of an anti-ferromagnetic.

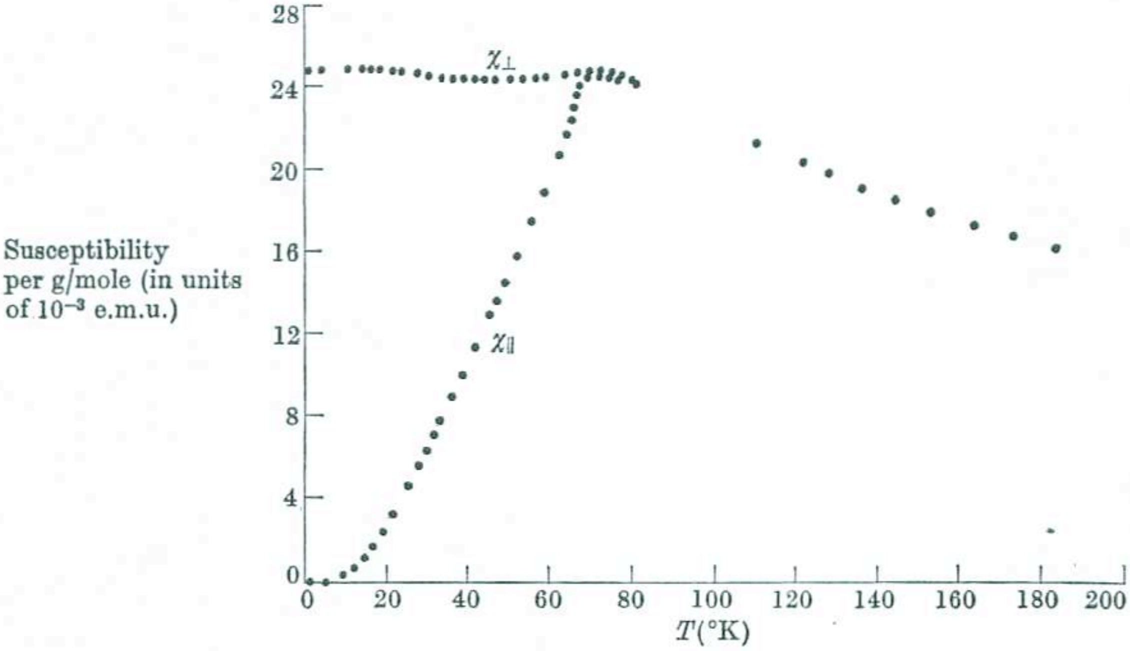


FIG. 22.6. Principal susceptibilities of a single crystal of MnF_2 (after C. Trapp and J. W. Stout, 1963).

22.3. Ferrimagnetism

The technical importance of magnetic materials in electrical industry has increased continuously, the ideal substance being one with a large magnetic moment at room temperature, which is also an electrical insulator. Ferromagnetic metals and alloys have been widely exploited, but their high electrical conductivity is a serious handicap in radio-frequency applications because of the eddy current losses. For this reason a number of magnetic oxides ('ferrites', of which magnetite, Fe_3O_4, is the most famous as the original 'lodestone') became of great technical interest because of their low electrical conductivity. They show spontaneous magnetization, remanence, and other properties

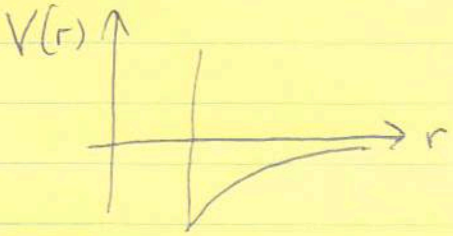
Lattice gas

Consider a container with volume V divided into cells of size v. Then $N_{cells} = \frac{V}{v}$. Only one atom can occupy each cell so the occupancy $n_i = 0$ or 1.

We introduce an attractive interaction between nearest neighbors so that the energy of microstate $\{n_i\}$ is $E = -\varepsilon \sum_{n.n.} n_i n_j$

where the sum is over unique nearest-neighbor pairs.

The idea underlying this model is to approximate an intermolecular potential which has a close-range attraction and a repulsive core



Let's develop a mean-field theory for the lattice gas.
The mean occupancy is $\bar{n} = \frac{N}{N_{cells}} = \frac{Nv}{V}$.

Following the same argument as for the Ising model (leave as an exercise)

$$E = \sum_{i=1}^{N_{cells}} E_i \quad \text{where } E_i = 3\varepsilon\bar{n}(\bar{n} - 2n_i).$$

Now use the grand canonical ensemble.

$$\mathcal{Z} = \mathcal{Z}_i^{N_{cells}} \qquad \mathcal{Z}_i = \left(1 + e^{6\beta\varepsilon\bar{n} + \beta\mu}\right) e^{-3\beta\varepsilon\bar{n}^2}$$

and grand ~~free energy~~

$$\boxed{\Phi = 3N_{cells}\varepsilon\bar{n}^2 - k_B T N_{cells} \ln\left(1 + e^{6\beta\bar{n}\varepsilon + \mu\beta}\right)}$$

Now we can derive the properties of the system.

i) <u>Number of particles</u>

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

$$= -\frac{6\varepsilon N}{N_{cells}}\frac{\partial N}{\partial \mu} + kT N_{cells}\left(\frac{e^{6\beta\bar{n}\varepsilon+\mu\beta}}{1+e^{6\beta\bar{n}\varepsilon+\mu\beta}}\right)\left(\beta + \frac{6\beta v\varepsilon}{V}\frac{\partial N}{\partial \mu}\right)$$

$$= -\frac{6\varepsilon N}{N_{cells}}\frac{\partial N}{\partial \mu} + \cancel{kT}\,\cancel{\beta}\, N_{cells}\,\frac{e^{6\beta\bar{n}\varepsilon+\mu\beta}}{1+e^{6\beta\bar{n}\varepsilon+\mu\beta}}$$

$$+ \frac{6\varepsilon}{\cancel{N_{cells}}}\frac{\partial N}{\partial \mu}\left(\frac{e^{6\beta\bar{n}\varepsilon+\mu\beta}}{1+e^{6\beta\bar{n}\varepsilon+\mu\beta}}\right)$$

$$\Rightarrow \frac{N}{N_{cells}} - \frac{e^{x}}{1+e^{x}} = \left(6\varepsilon\frac{\partial N}{\partial \mu}\right)\left[-\frac{N}{N_{cells}} + \frac{e^{x}}{1+e^{x}}\right]$$

$$\Rightarrow \boxed{\bar{n} = \frac{N}{N_{cells}} = \frac{e^{x}}{1+e^{x}}} \qquad x = 6\beta\bar{n}\varepsilon + \beta\mu$$

We can invert this to get μ in terms of N

$$e^{x} = \frac{\bar{n}}{1-\bar{n}} \Rightarrow \boxed{\mu = -6\bar{n}\varepsilon + kT\ln\left(\frac{\bar{n}}{1-\bar{n}}\right)}$$

when $\bar{n} \ll 1$

$$\mu \approx \bar{n}(kT - 6\varepsilon) + \underbrace{kT\ln\bar{n}}_{\text{this term dominates as } \bar{n}\to 0}$$

$\Rightarrow$ as $\bar{n} \to 0$ $\quad \mu \to kT \ln\left(\frac{N}{V} v\right)$

This looks like the ideal gas result $kT\ln(n/n_Q)$ with v^{-1} playing the role of n_Q.

2) Pressure

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

$$= -\frac{\partial}{\partial V}\left[\frac{3\varepsilon N^2 v}{V} - \frac{kT\cancel{V}}{\cancel{V}v}\ln\left(1+e^x\right)\right]$$

$$= \frac{3\varepsilon N^2 v}{V^2} - \frac{6\varepsilon N v}{V}\frac{\partial N}{\partial V} + \frac{kT}{v}\ln(1+e^x) + \frac{kTV}{v}\underbrace{\frac{e^x}{1+e^x}}_{\bar{n}}\cdot\frac{\partial x}{\partial V}$$

$$= \frac{3\varepsilon N^2 v}{V^2} - \cancel{\frac{6\varepsilon N v}{V}\frac{\partial N}{\partial V}} + \frac{kT}{v}\ln(1+e^x) + \cancel{kT}N\left[6\cancel{\beta}\varepsilon\right]\left(\cancel{\frac{\partial N}{\partial V}\frac{1}{N_{sites}}} - \cancel{2}\frac{Nv}{V^2}\right)$$

$$\boxed{P = -3\varepsilon\left(\frac{N}{V}\right)^2 v + \frac{kT}{v}\ln(1+e^x)}$$

Again check the ideal gas limit $\bar{n} \ll 1$

first write $\ln(1+e^x) = -\ln(1-\bar{n}) \approx \bar{n}$

$$\Rightarrow P \approx -3\varepsilon\cancel{v}\,\bar{n}^2 + \frac{kT\bar{n}}{v}$$

$$= \frac{NkT}{V} + O(\bar{n}^2)$$

= ideal gas

We looked at a gas model with close-range repulsion before (HW2 Q1) — the "hard sphere" model.
There we found that

$$P = \frac{NkT}{V-b}$$

where $b = 4 \times$ (volume of a single sphere)

Here, We have $$P = \underbrace{\frac{NkT}{V}\left[\frac{1}{\bar{n}}\ln\left(\frac{1}{1-\bar{n}}\right)\right]} - 3\varepsilon v\left(\frac{N}{V}\right)^2$$

this term has the same qualitative behavior as the hard sphere gas — the repulsion leads to extra pressure when the volume of the molecules becomes comparable to the volume of the container.

The second term is due to the attractive part of the potential and leads to a reduction in the pressure. This kind of term appears in the van der Waals equation of state

$$P = \frac{NkT}{V-b} - a\left(\frac{N}{V}\right)^2$$

often written as $$\left[P + a\left(\frac{N}{V}\right)^2\right]\left[V-b\right] = NkT$$

attice gas in the canonical ensemble

How can we calculate the lattice gas using the canonical ensemble? Here is an approach which is equivalent to MFT. We imagine laying down N particles into the N_{cells} grid cells. The number of ways to do this is

$$\frac{N_{cells}!}{(N_{cells}-N)!\,N!} = W$$

The number of pairs is

$$N_{pairs} \simeq N_{cells}\left(\frac{N}{N_{cells}}\right)\left(\frac{6N}{N_{cells}}\right)\frac{1}{2}$$

[writing the number of pairs this way is the Bragg-Williams approximation] →

$$= \frac{3N^2}{N_{cells}}$$

$\Rightarrow$ each of the W configurations have energy $E = -\dfrac{3\varepsilon N^2}{N_{cells}}$

$$\Rightarrow \quad Q \simeq \frac{N_{cells}!}{(N_{cells}-N)!\,N!}\exp\left(+\frac{\beta 3\varepsilon N^2}{N_{cells}}\right)$$

$$\Rightarrow \quad A = -kT\ln Q$$

$$= -\frac{3\varepsilon N^2}{N_{cells}} - kT\Big[\underbrace{N_{cells}\ln N_{cells} - N\ln N - (N_{cells}-N)\ln(N_{cells}-N)}_{N_{cells}\left[-\bar{n}\ln\bar{n} - (1-\bar{n})\ln(1-\bar{n})\right]}\Big]$$

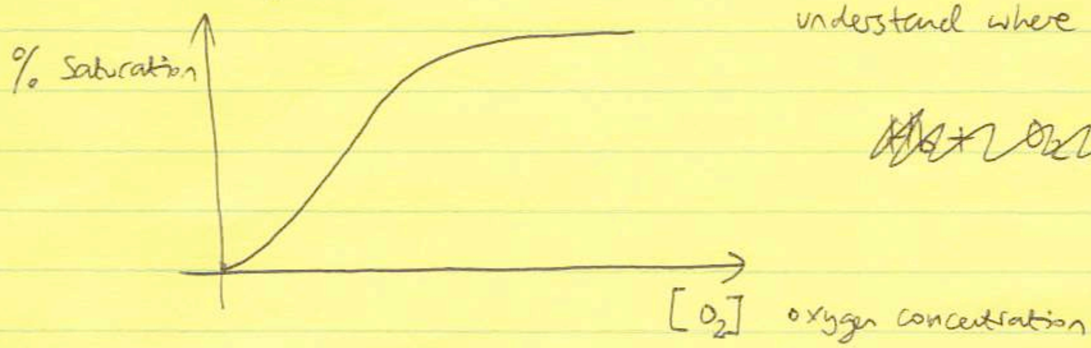
$$= \underbrace{-\frac{3\varepsilon N^2}{N_{cells}}}_{E} + \underbrace{Tk_B\Big\{N_{cells}\left[\bar{n}\ln\bar{n} + (1-\bar{n})\ln(1-\bar{n})\right]}_{-TS}$$

which is the same result as previously.

Cooperative binding in hemoglobin

Hemoglobin is the protein in red blood cells responsible for transporting oxygen. We'd like to understand how does the oxygen bind to the protein.

The oxygen-binding curve is sigmoidal or S-shaped. Can we understand where this comes from?



~~$Hb + 2O_2 \rightleftharpoons HbO_2$~~

The simplest model is to consider the protein as having two states, bound and unbound, with the bound state being lowest in energy.

$\sigma = 0$ $\quad\quad \varepsilon = 0$ $\quad\quad \left[Hb + O_2 \rightleftharpoons HbO_2 \right]$

$\sigma = 1$ $\quad\quad \varepsilon = \varepsilon_b < 0$ $\quad\quad$ lower energy state

Treat the surrounding solution of oxygen as a heat and particle bath — grand canonical ensemble.

$$\mathcal{Z} = \sum_{\nu} e^{-\beta(E_\nu - \mu N_\nu)}$$

$$= 1 + e^{-\beta(\varepsilon_b - \mu)}$$

the mean occupation number is $\langle n \rangle = \dfrac{e^{-\beta(\varepsilon_b - \mu)}}{1 + e^{-\beta(\varepsilon_b - \mu)}}$

$$\Rightarrow \quad \langle n \rangle = \frac{1}{1 + e^{\beta(\varepsilon_b - \mu)}} .$$

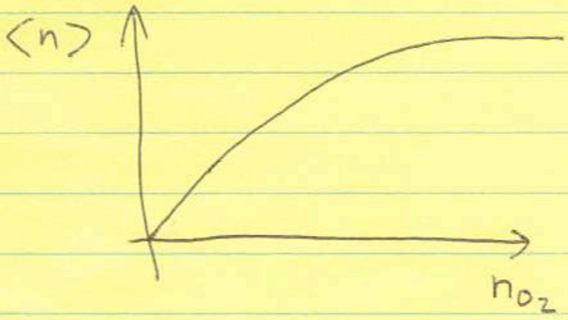
The chemical potential of the O_2 molecules in solution is

$$\mu = kT \ln\left(\frac{n_{O_2}}{n_Q}\right) \qquad n_Q = \left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2}$$

$$\Rightarrow \quad e^{-\beta\mu} = \frac{n_Q}{n_{O_2}}$$

$$\Rightarrow \quad \boxed{\langle n \rangle = \frac{n_{O_2}}{n_{O_2} + n_Q e^{\varepsilon_b/kT}}}$$

Plot this:



it has the wrong shape!
This curve does match the oxygen binding curve for myoglobin – so this model seems to work there.

The answer is that there are multiple binding sites (4) on the protein ~~what if there are m~~ and the binding is cooperative – binding one oxygen molecule makes it easier for the next to bind, and so on.

To see the effect on the binding curve, let's make the simplest possible "all or nothing" model. Consider again two states, but now these are

~~[illegible] and~~ no molecules bound $(\varepsilon = 0)$
~~$\varepsilon = 0$~~ all molecules bound $(\varepsilon = \varepsilon_b < 0)$

if there are m binding sites then

$$\langle n \rangle = \frac{1}{1 + e^{+m\beta(\varepsilon_b - \mu)}} = \frac{1}{1 + \left(\frac{n_Q}{n_{O_2}}\right)^m e^{+\beta m \varepsilon_b}}$$

$Hb + mO_2 \rightleftharpoons Hb(O_2)_m$ (for $m>1$ this has an S-shape)

or

$$\boxed{\langle n \rangle = \frac{(n_{O_2})^m}{(n_{O_2})^m + \left(n_Q e^{\beta \varepsilon_b}\right)^m}}$$

This is the famous <u>Hill equation</u>

Fitting to the measured hemoglobin curve gives $m \simeq 3$ even though from the structure of the protein we know that there are 4 binding sites. We need to improve our model further.

The <u>Pauling model</u> is to introduce an interaction between pairs of binding sites

$$E = \varepsilon_b \sum_{i=1}^{4} \sigma_i + \frac{J}{2} \sum_{i,j} \sigma_i \sigma_j$$

where $\sigma_i = \begin{cases} 1 & \text{if site } i \text{ is occupied} \\ 0 & \text{if site } i \text{ is unoccupied} \end{cases}$

The partition function is $Z = \sum_{\{\sigma_i\}} e^{-\beta(E - N\mu)}$

Count the different states:

0 bound	$E = 0$	1 state
1 bound	$E = \varepsilon$	4 states
4 bound	$E = 4\varepsilon + 6J$	1 state
2 bound	$E = 2\varepsilon + J$	6 states
3 bound	$E = 3\varepsilon + 3J$	4 states

$$\Rightarrow Z = 1 + 4e^{-\beta(\varepsilon-\mu)} + 6e^{-2\beta(\varepsilon-\mu) - \beta J}$$
$$+ 4e^{-3\beta(\varepsilon-\mu) - 3\beta J} + e^{-4\beta(\varepsilon-\mu) - 6\beta J}$$

Then $\langle n \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z$

This gives $$\langle n \rangle = \frac{4x + 12x^2 j + 12x^3 j^3 + 4x^4 j^6}{1 + 4x + 6x^2 j + 4x^3 j^3 + x^4 j^6}$$

where $x = e^{-\beta(\varepsilon - \mu)}$

$j = e^{-\beta J}$

Check the limits: $J = 0$ no interactions

$\Rightarrow j = 1$

$$\langle n \rangle = \frac{4x + 12x^2 + 12x^3 + 4x^4}{1 + 4x + 6x^2 + 4x^3 + x^4}$$

$$= \frac{4x \left(1 + 3x + 3x^2 + x^3\right)}{(1+x)^4} \longrightarrow (1+x)^3$$

$$\langle n \rangle = \frac{4x}{1+x}$$

$$= \frac{4}{1 + e^{\beta(\varepsilon - \mu)}}$$

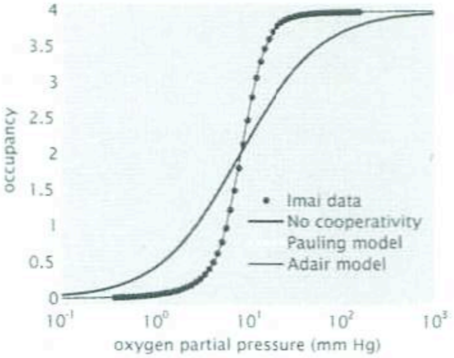
Same as 4 independent sites ✓

and as T drops, x and j increase and $\langle n \rangle$ moves from $0 \to 1 \to 2 \to 3 \to 4$ ✓

Can get a good fit to the data by varying J and ε.

Other models go further, eg. Adair model has 3 and 4 site interactions.

Phillips, Kondev & Theriot
"Physical Biology of the Cell"



Adair model has 3 and 4 site interactions

Figure 7.18: Hemoglobin binding. Comparison of the oxygen binding isotherms for different models of hemoglobin using the two-level system description. (Data from K. Imai, *Biophys. Chem.*, 37:1, 1990.)

the next, rather, is our choice of energy function. In the case of the Pauling model, the physical content of the cooperativity arises because it is assumed that there is a pairwise interaction between oxygens on different sites. If we think of the four binding sites as the vertices of a tetrahedron, there are six interactions corresponding to the six edges of the tetrahedron. If we label the four vertices 1, 2, 3 and 4, these pairwise interactions are between 1 and 2, 1 and 3, etc. and there are a total of six distinct such interactions.

Within this model, the energy of the system is written in the form

$$E = \varepsilon \sum_{\alpha=1}^{4} \sigma_\alpha + \frac{J}{2} \sum_{(\alpha,\gamma)}{}' \sigma_\alpha \sigma_\gamma, \tag{7.42}$$

where the sums over α and γ run from 1 to 4 and the prime $\sum'$ instructs us *not* to include terms in the sum when $\alpha = \gamma$ and J is divided by 2 to account for the presence of terms like $\sigma_1\sigma_2$ and $\sigma_2\sigma_1$ which both occur in the sum. Whenever two different sites are occupied, there is a corresponding term in the energy with a contribution J. The partition function corresponding to this energy is given by

$$\mathcal{Z} = \sum_{\sigma_1=0}^{1} \sum_{\sigma_2=0}^{1} \sum_{\sigma_3=0}^{1} \sum_{\sigma_4=0}^{1} e^{-\beta(\varepsilon-\mu)\sum_{\alpha=1}^{4}\sigma_\alpha - \beta\frac{J}{2}\sum'_{\alpha,\gamma}\sigma_\alpha\sigma_\gamma}, \tag{7.43}$$

which once again corresponds to summing over all eight states of occupancy of the hemoglobin molecule by its partner oxygens. As before, the partition function can be evaluated analytically and is given by

$$\mathcal{Z} = \underbrace{1}_{0\text{ bound}} + \underbrace{4e^{-\beta(\varepsilon-\mu)}}_{1\text{ bound}} + \underbrace{6e^{-2\beta(\varepsilon-\mu)-\beta J}}_{2\text{ bound}} + \underbrace{4e^{-3\beta(\varepsilon-\mu)-3\beta J}}_{3\text{ bound}} + \underbrace{e^{-4\beta(\varepsilon-\mu)-6\beta J}}_{4\text{ bound}}. \tag{7.44}$$

The law of mass action (Some chemistry)

We could have derived the expression for $\langle n \rangle$ for the simple single site model ($Hb + O_2 \rightleftarrows HbO_2$) or the "all or nothing" (Hill equation $Hb + mO_2 \rightleftharpoons Hb(O_2)_m$) model using the law of mass action

which states that if the reaction

$$aA + bB \rightleftharpoons cC + dD$$

is in equilibrium then $$\frac{[C]^c [D]^d}{[A]^a [B]^b} = \underset{\uparrow \text{"equilibrium constant" or "rate constant"}}{K}$$

eg. Hill equation $$\frac{[Hb(O_2)_m]}{[Hb][O_2]^m} = K$$

define $Y = \dfrac{[Hb(O_2)_m]}{[Hb] + [Hb(O_2)_m]}$ which is equivalent to our $\langle n \rangle$.

$$= \frac{1}{1 + \frac{1}{K[O_2]^m}} = \underline{\underline{\frac{[O_2]^m}{[O_2]^m + K^{-1}}}} \quad — (*)$$

Same as before if we identify $\underline{\underline{K^{-1} = \left(n_Q e^{\beta \varepsilon_b}\right)^m}}$

Let's understand in more detail where the law of mass action comes from. What drives chemical reactions? The system is trying to reduce its free energy. So, in equilibrium it must be the case that the free energy of the reactants and products are equal. The appropriate free energy here is the Gibbs free energy G (constant T and P).

$$\Rightarrow \quad G(\text{reactants}) = G(\text{products})$$

or since $G = \mu N$ → [if you're not sure why $G = \mu N$ see the end of this section]

then eg for the reaction $aA + bB \rightleftharpoons cC + dD$ we must have

$$\underline{a\mu_a + b\mu_b = c\mu_c + d\mu_d}$$

When we write down the chemical potential for each species, we must be careful to include the chemical energy of the molecule, ε

$$\mu = \varepsilon + kT\ln\left(\frac{n}{n_Q}\right)$$

[To see this, note that if we include an extra constant energy per molecule ε, the partition function is multiplied by $\left(e^{-\beta\varepsilon}\right)^N$

or the free energy has an extra term $N\varepsilon$ $\Rightarrow$ an extra term ε in the chemical potential because $\mu = \left.\frac{\partial A}{\partial N}\right|_{T,V}$ or $\left.\frac{\partial G}{\partial N}\right|_{T,P}$]

In chemistry, it is standard practice to write

$$\mu = kT\ln\left(\frac{c}{c_0}\right) + \mu^\circ$$

μ° ← standard chemical potential

c_0 = standard concentration (often $\frac{1\,\text{mole}}{L}$)

We see that $\mu^\circ = \varepsilon + kT \ln c_0 - kT \ln n_Q$

$$= \underline{\varepsilon + kT \ln} \left(\frac{c_0}{n_Q}\right)$$

[The "availability" is then

$$\frac{c}{c_0} = \exp\left(\frac{\mu - \mu^\circ}{kT}\right)$$]

So, for our reaction $Hb + m(O_2)^* \rightleftharpoons Hb(O_2)_m$

we have

$$-\mu_{Hb} - m\,\mu_{O_2} + \mu_{Hb(O_2)_n} = 0$$

or $$0 = \underbrace{\frac{-\mu^\circ_{Hb} - m\mu^\circ_{O_2} + \mu^\circ_{Hb(O_2)_n}}{kT}}_{\text{write this as } -\ln K'} + \ln\left[\left(\frac{c_{Hb(O_2)_m}}{c_0}\right)\left(\frac{c_{Hb}}{c_0}\right)^{-1}\left(\frac{c_{O_2}}{c_0}\right)^{-m}\right]$$

then $$\boxed{\frac{c_{Hb(O_2)_m}}{c_{Hb}\left(c_{O_2}\right)^m} = \frac{K'}{c_0^{\,m}}}$$

We've derived the law of mass action.

we see that $$K' = e^{-\beta\left[\Sigma\mu^\circ(\text{products}) - \Sigma\mu^\circ(\text{reactants})\right]}$$

$$= e^{-\Delta G^\circ/kT}$$

where ΔG° is the standard free energy change.

Statistical mechanics gives a microscopic basis for the law of mass action!

In our case $\mu^{0}_{Hb(O_2)_m} - m\mu^{0}_{O_2} - \mu^{0}_{Hb}$

$$= m\varepsilon_b + kT\ln\left(\frac{c_0}{\underset{n_{Q,O_2}}{\cancel{n_Q}}}\right) \quad \cancel{(1\,4 - m - 1)}$$

this is the n_Q for O_2: we assume $n_Q(Hb(O_2)_m) \simeq n_Q(H_b)$ or in other words mass of Hb >>> mass of O_2

$$\Rightarrow \quad K' = \exp\left(-\beta\varepsilon_b m\right) \times \left(\frac{c_0}{n_Q}\right)^m$$

$$\text{or} \quad \left(\frac{K'}{c_0^{\,m}}\right)^{-1} = n_{Q,O_2}^{\,m}\, e^{m\varepsilon_b/kT}$$

put this into (*)

$$\Rightarrow \quad Y = \frac{[O_2]^m}{[O_2]^m + \left(n_{Q_{O_2}} e^{\varepsilon_b/kT}\right)^m}$$

How to show that $G = \mu N$:

The 1st law $\quad dE = TdS - pdV + \mu dN$

but E, S, V are all <u>extensive variables</u>

$\Rightarrow$ we can integrate $\quad E = TS - PV + \mu N$

$$\Rightarrow \quad G = E - TS + PV = \mu N$$

(this step uses Euler's theorem for homogeneous functions)

Some notes on Euler's theorem of homogeneous functions

The theorem is that if a function satisfies

$$f(\lambda x, \lambda y) = \lambda^n f(x,y)$$

then $$x \left.\frac{\partial f}{\partial x}\right|_y + y \left.\frac{\partial f}{\partial y}\right|_x = nf.$$

The application in thermodynamics is to obtain expressions such as $G = \mu N$, $\Phi = -PV$ etc.

eg. Gibb's free energy $G = E - TS + PV$

$$dG = -S\,dT + V\,dP + \mu\,dN$$

(using $dE = T\,dS - p\,dV + \mu\,dN$)

G is a function of $\underbrace{T, P}_{\text{intensive}}, \overset{\nwarrow \text{extensive}}{N}$

$$\Rightarrow \quad G = N g(T, P)$$

$$\Rightarrow \quad \underbrace{N \left.\frac{\partial G}{\partial N}\right|_{T,P}}_{\mu} = G$$

$$\Rightarrow \quad \underline{G = \mu N}$$

eg. Grand free energy $\Phi = E - TS - \mu N$

function of T, μ, V

$$\Phi = V \phi(T, \mu)$$

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

But $d\Phi = -S\,dT - p\,dV - N\,d\mu \Rightarrow \left.\frac{\partial \Phi}{\partial V}\right|_{T,\mu} = -P$

$\Rightarrow \underline{\Phi = -PV}$

eg. Helmholtz $A = E - TS$

function of T, N, V

obeys

$$A(\lambda N, \lambda V) = \lambda A \quad \leftarrow$$

(ie. double the volume and # particles ⇒ double the free energy)

$$\Rightarrow A = N \left.\frac{\partial A}{\partial N}\right|_{V,T} + V \left.\frac{\partial A}{\partial V}\right|_{N,T}$$

But $dA = -S\,dT - p\,dV + \mu\,dN$

$$\Rightarrow \underline{A = \mu N - PV}$$

[check for ideal gas $A = N k_B T \ln\left(\frac{n}{n_Q}\right) - \frac{N}{V} V k_B T$

$= N k_B T \left[\ln\left(\frac{n}{n_Q}\right) - 1\right]$ ✓]

Universality

Experimental data on several different critical exponents is given in the table on the next page, taken from Pathria's book.

The definitions are

			MFT	Observed
β	$m \propto (T_c - T)^{\beta}$	$T \lesssim T_c \quad H \to 0$	$\frac{1}{2}$	$0 \cdot 3 - 0 \cdot 36$
γ	$\chi \propto (T - T_c)^{-\gamma}$,	$T \gtrsim T_c \quad H \to 0$	1	$1 \cdot 2 - 1 \cdot 4$
	$(T_c - T)^{-\gamma'}$	$T \lesssim T_c \quad H \to 0$		$1 \cdot 0 - 1 \cdot 2$
δ	$m\big\vert_{T=T_c} \propto H^{1/\delta}$	$T = T_c,\ H \to 0$	3	$4 \cdot 2 - 4 \cdot 8$
α	$C_V \propto (T - T_c)^{-\alpha}$	$T \gtrsim T_c$	0	$0 \cdot 0 - 0 \cdot 2$
	$(T_c - T)^{-\alpha'}$	$T \lesssim T_c$		

For a gas, $\rho - \rho_c$ and $P - P_c$ play the roles of H and M.

Remarkably, very different systems are observed to have the same critical exponents. This is the idea of universality. Systems which show the same critical exponents are said to belong to the same universality class. Which universality class applies in a given case depends on only a small number of parameters – including the dimensionality d, the number of components of the order parameter, and the range of the interactions (eg. nearest neighbors vs. infinite).

The critical exponents are not independent, but satisfy relations. For example, Rushbrooke (1963) showed that

$$(\alpha' + 2\beta + \gamma') \geq 2 \qquad \left(\begin{array}{l}\text{in MFT} \\ \alpha + 2\beta + \gamma = 2\end{array}\right)$$

[The argument is a thermodynamic one:

Pathria p 336.

TABLE 11.1. EXPERIMENTAL DATA ON CRITICAL EXPONENTS

Critical exponents	Magnetic systems$^{(a)}$	Gas–liquid systems$^{(b)}$	Binary fluid mixtures$^{(c)}$	Binary alloys$^{(d)}$	Ferroelectric systems$^{(e)}$	Superfluid He$^{4(f)}$	Mean field results
α,α'	0.0–0.2	0.1–0.2	0.05–0.15	– – –	– – –	−0.026	0
β	0.30–0.36	0.32–0.35	0.30–0.34	0.305 ± 0.005	0.33–0.34	– – –	1/2
γ	1.2–1.4	1.2–1.3	1.2–1.4	1.24 ± 0.015	1.0 ± 0.2	inaccessible	1
γ'	1.0–1.2	1.1–1.2	– – –	1.23 ± 0.025	1.23 ± 0.02	inaccessible	1
δ	4.2–4.8	4.6–5.0	4.0–5.0	– – –	– – –	inaccessible	3
ν	0.62–0.68	– – –	– – –	0.65 ± 0.02	0.5–0.8	0.675	1/2
η	0.03–0.15	– – –	– – –	0.03–0.06	– – –	– – –	0

(a) Stierstadt *et al.* (1990).
(b) Voronel (1976); Rowlinson and Swinton (1982).
(c) Rowlinson and Swinton (1982).
(d) J. Als-Nielsen (1976); data pertain to beta-brass only.
(e) Kadanoff *et al.* (1967); Lines and Glass (1977).
(f) Ahlers (1980).

start with

$$C_H - C_M = -T\left(\frac{\partial H}{\partial T}\right)_M \left(\frac{\partial M}{\partial T}\right)_H = \frac{T}{\chi}\left(\frac{\partial M}{\partial T}\right)_H^2$$

but $C_M > 0 \Rightarrow C_H \geqslant \frac{T}{\chi}\left(\frac{\partial M}{\partial T}\right)_H^2$

Now let $H \to 0$ $T \to T_c$ from below

$$\Rightarrow \quad k_1 \left(T_c - T\right)^{-\alpha'} \geqslant k_2 \left(T_c - T\right)^{\gamma' + 2(\beta - 1)} \quad - (*)$$

where k_1~~&~~, k_2 are positive constants.

Eq. (*) only holds for all $T_c - T$ if the inequality $\alpha' + \gamma' + 2(\beta - 1) \geqslant 0$ holds.

┘

Sethna

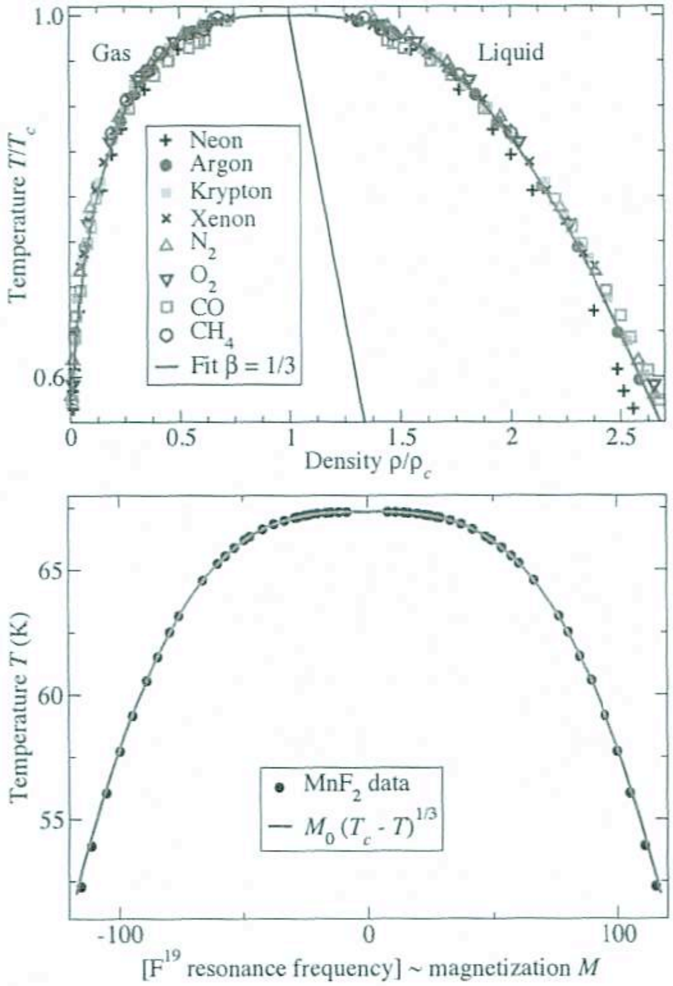


Fig. 12.6 Universality. (a) Universality at the liquid–gas critical point. The liquid–gas coexistence lines ($\rho(T)/\rho_c$ versus T/T_c) for a variety of atoms and small molecules, near their critical points (T_c, ρ_c) [54]. The curve is a fit to the argon data, $\rho/\rho_c = 1 + s(1 - T/T_c) \pm \rho_0(1 - T/T_c)^\beta$ with $s = 0.75$, $\rho_0 = 1.75$, and $\beta = 1/3$ [54]. (b) Universality: ferromagnetic–paramagnetic critical point. Magnetization versus temperature for a uniaxial antiferromagnet MnF_2 [56]. We have shown both branches $\pm M(T)$ and swapped the axes so as to make the analogy with the liquid–gas critical point (above) apparent. Notice that both the magnet and the liquid–gas critical point have order parameters that vary as $(1 - T/T_c)^\beta$ with $\beta \approx 1/3$. (The liquid–gas coexistence curves are tilted; the two theory curves would align if we defined an effective magnetization for the liquid–gas critical point $\rho_{\text{eff}} = \rho - 0.75\rho_c(1 - T/T_c)$ (thin midline, above). This is not an accident; both are in the same universality class, along with the three-dimensional Ising model, with the current estimate for $\beta = 0.325 \pm 0.005$ [148, chapter 28].

of falling apart, they become similar to one another! In particular, all signs of the original lattice structure and microscopic rules have disappeared.[4]

Thus we observe in these cases that different microscopic systems look the same near critical points, if we ignore the microscopic details and confine our attention to long length scales. To study this systematically, we need a method to take a kind of continuum limit, but in systems which remain inhomogeneous and fluctuating even on the largest scales. This systematic method is called the *renormalization group*.[5]

The renormalization group starts with a remarkable abstraction: it

[4] Notice in particular the *emergent symmetries* in the problem. The large percolation clusters at p_c are statistically both translation invariant and rotation invariant, independent of the grids that underly them. In addition, we will see that there is an emergent *scale invariance*—a kind of symmetry connecting different length scales (as we also saw for random walks, Fig. 2.2).

[5] The word renormalization grew out of quantum electrodynamics, where the effective charge on the electron changes size (norm) as a function of length scale. The word group is usually thought to refer to the family of coarse-graining operations that underly the method (with the group product being repeated coarse-graining). However, there is no inverse operation to coarse-graining, so the renormalization group does not satisfy the definition of a mathematical group.

More about fluctuations

We've seen expressions for fluctuations in several quantities in this course, for example

$$\sigma_E^2 = k_B T^2 C_V = k_B T^2 \left.\frac{\partial E}{\partial T}\right|_V$$

$$\sigma_N^2 = k_B T \left.\frac{\partial N}{\partial \mu}\right|_{T,V}$$

$$\sigma_V^2 = -k_B T \left.\frac{\partial V}{\partial P}\right|_T$$

$$\sigma_M^2 = k_B T \chi = k_B T \left.\frac{\partial M}{\partial H}\right|_T$$

They all take the form

$$\text{fluctuation in quantity } X = (k_B T) \overbrace{\left(\begin{array}{l}\text{rate of change of } X \\ \text{as the controlling} \\ \text{parameter } Y \text{ is varied}\end{array}\right)}^{\substack{\text{this is a thermodynamic} \\ \text{response function} \\ \text{("susceptibility")}}}$$

Which variables are held constant in the derivatives on the RHS depends on which ensemble is being used – ie. what the constraints are. For example, we derived σ_E^2 in the canonical ensemble and so $\sigma_E^2 = k_B T^2 \left.\frac{\partial E}{\partial T}\right|_{V,N}$. In the grand canonical ensemble both μ and T are held constant so we have $\sigma_E^2 = k_B T^2 \left.\frac{\partial E}{\partial T}\right|_{V,\mu}$.

Thermodynamic relations can be used to get these relations into a more useful form. For example

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

grand free energy

but since $\Phi = E - TS - \mu N = -pV$

then $-SdT - Nd\mu = -VdP$

and $\left.\frac{\partial \mu}{\partial P}\right|_T = \frac{1}{n}$

use $dE = TdS - pdV + \mu dN$

$$\Rightarrow \quad \left.\frac{\partial N}{\partial \mu}\right|_{T,V} = V \left.\frac{\partial n}{\partial P}\right|_{T,V} n$$

$$\Rightarrow \quad \sigma_N^2 = \frac{k_B T N}{V^{\cancel{2}}} \left.\frac{\partial N}{\partial P}\right|_{T,V}$$

but $\left.\frac{\partial N}{\partial P}\right|_V \left.\frac{\partial P}{\partial V}\right|_N \left.\frac{\partial V}{\partial N}\right|_P = -1 \Rightarrow \left.\frac{\partial N}{\partial P}\right|_{T,V} = -\left.\frac{\partial V}{\partial P}\right|_{T,N} \left.\frac{\partial N}{\partial V}\right|_{P,T}$

$= N/V \; \cancel{2}$

(extensive)

$$\Rightarrow \quad \frac{\sigma_N^2}{N^2} = -\frac{k_B T}{\cancel{N} V^{\cancel{2}}} \left.\frac{\partial V}{\partial P}\right|_{T,N} \frac{\cancel{N}}{V}$$

$$= -\frac{k_B T}{V} \left.\frac{\partial \ln V}{\partial P}\right|_{T,N} = \underline{-\frac{k_B T \kappa_T}{V}}$$

(isothermal compressibility)

We also have seen the relation between χ (and therefore σ_M^2) and the correlation function $\langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$

$$\sigma_M^2 = k_B T \chi = \mu^2 N \underbrace{\sum_{j=1}^{N} \left(\langle s_1 s_j \rangle - \langle s \rangle^2 \right)}_{\sum_{\underline{r}} G(\underline{r})}$$

A similar argument relates fluctuations in number of particles to the spatial correlation function $\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle$ (lattice gas $n_i = 0$ or 1 occupation # of site i).

To see this note that the number of particles is $N = \sum_{i=1}^{N_{cells}} n_i$

$$\Rightarrow \quad \sigma_N^2 = \langle N^2 \rangle - \langle N \rangle^2$$
$$= \sum_{i=1}^{N_{cells}} \sum_{j=1}^{N_{cells}} \left(\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle \right)$$
$$= N_{cells} \sum_{j=1}^{N_{cells}} C_{1j} \quad .$$

We can use this to check the case of an ideal gas. Then the positions of particles should be uncorrelated

$$\langle n_i n_j \rangle = \langle n_i \rangle \langle n_j \rangle$$
$$\left(\text{ie. } \langle (n_i - \langle n \rangle)(n_j - \langle n \rangle) \rangle = 0 \right)$$

Also the cells are sparsely occupied $\langle n \rangle \ll 1$ and $n_i = 0$ or 1 $\Rightarrow$ $\langle n_i \rangle = \langle n_i^2 \rangle$

$$\Rightarrow \quad \frac{\sigma_N^2}{N_{cells}} = \sum_{j=1}^{N_{cells}} \left(\langle n_1 n_j \rangle - \langle n_1 \rangle \langle n_j \rangle \right)$$
$$= \langle n_1^2 \rangle - \langle n_1 \rangle^2 = \langle n \rangle (1 - \langle n \rangle)$$
$$\simeq \langle n \rangle$$
$$= \frac{N}{N_{cells}}$$

$$\Rightarrow \quad \sigma_N^2 = N \, .$$

But $\frac{\sigma_N^2}{N^2} = -\frac{k_B T}{V} \left.\frac{\partial \ln V}{\partial P}\right|_{T,N} = \frac{N}{N^2} \Rightarrow \left.\frac{\partial P}{\partial V}\right|_{T,N} = -\frac{N k_B T}{V^2}$

$$\Rightarrow \boxed{P = \frac{N}{V} k_B T} \checkmark$$

[An aside: I'm not sure if we ever plugged numbers into n_Q.

$$n_Q = \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \approx 10^{28}\ \mathrm{cm}^{-3} \left(\frac{T}{300\mathrm{K}}\right)^{3/2} \left(\frac{m}{m_p}\right)^{3/2}.$$

Compared to air in the room $\rho = 10^{-3}\ \mathrm{g/cm^{+3}}$ $\left(\frac{1}{1000}\text{ of water}\right)$

$$\Rightarrow n \simeq \frac{10^{-3}\ \mathrm{g/cm^3}}{30\times(10^{-24}\mathrm{g})} \simeq 3\times 10^{19}\ \mathrm{cm}^{-3}.$$

$$\Rightarrow \frac{n}{n_Q} \sim 10^{-10}$$

very sparsely occupied.]

When interactions between particles are included, then particle positions become correlated. An important quantity is the two point correlation function $g(r)$

$\uparrow$ (only depends on distance in isotropic system)

Given a particle at the origin, the probability of finding a second particle at $\underline{r}$ is $g(r)\, d^3\underline{r}$.

[ideal gas has uniform $g(r) = 1$ for all r so $\sigma_N^2 = N$ as before.]

You can show that

$$\frac{\sigma_N^2}{N} = 1 + n\int (g(r)-1)\, d^3\underline{r}.$$

So for example in a Monte Carlo simulation of interacting particles one can measure $g(r)$ and from there derive thermodynamic properties – eg. equation of state $P(\rho, T)$

Brownian motion

Random motion of pollen grains observed by botanist Robert Brown in 1828. Due to random collisions with molecules. The pollen grains follow a random walk. This was analyzed by Einstein (1905) and Smoluchowski (1906).

Let's model this in 1D – assume each collision occurs after a time t_c on average, and causes the molecule to move a distance ℓ left or right, with equal probability.

What is the probability that we find the particle at position x after n jumps? (ie. after $t = nt_c$ on average).

To get to x, it must have made $m = \frac{x}{\ell}$ more jumps in the +ve x-direction than in the -ve x-direction.

Jumps to the right $= \frac{n+m}{2}$

jumps to the left $= \frac{n-m}{2}$

$$\text{Prob}_n(m) = \frac{n!}{\left(\frac{1}{2}(n+m)\right)!\,\left(\frac{1}{2}(n-m)\right)!}\left(\frac{1}{2}\right)^n$$

which gives
$$\overline{m} = 0$$
$$\overline{m^2} = n$$

or in terms of position (go to large enough time that $x = \ell m$ is a good approximation)

$$\overline{x(t)} = 0$$
$$\overline{x^2(t)} = \ell^2 n^2 = \ell^2 \frac{t}{t_c} = \left(\frac{\ell^2}{t_c}\right) t \quad \propto t.$$

or $\underline{\sqrt{\overline{x^2}} \propto t^{1/2}}$.

rms displacement increases as $t^{1/2}$.

Use Stirling's approximation for large n and assume $m \ll n$

$$n! \simeq (2\pi n)^{1/2}\left(\frac{n}{e}\right)^n$$

(good approx. since $m \sim n^{1/2}$)

$$\log \mathrm{Prob}_n(m) = -n\log 2 + \left(n+\tfrac{1}{2}\right)\log n$$
$$- \left(\frac{n+m+1}{2}\right)\log\left(\frac{n+m}{2}\right) - \left(\frac{n-m+1}{2}\right)\log\left(\frac{n-m}{2}\right)$$
$$- \tfrac{1}{2}\ln 2\pi$$

which gives $\quad \mathrm{Prob}_n(m) \simeq \dfrac{2}{\sqrt{2\pi n}}\, e^{-m^2/2n}$

$\Rightarrow$ we can write

$$p(x)\,dx = \frac{dx}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right)$$

where $D = \dfrac{\ell^2}{2t_c}$.

Instead of following only one particle, we could ~~also~~ follow the collection of particles. Their number density $n(\underline{x})$ satisfies a diffusion equation

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

If the particles are concentrated initially at the origin, the solution is

$$n(r,t) = \frac{N}{(4\pi D t)^{1/2}} \exp\left(-\frac{\underline{x}^2}{4Dt}\right).$$

This is normalized

$$\int_{-\infty}^{\infty} n \, dx = N.$$

We can calculate $\quad \langle \underline{x}(t)\rangle = 0$

and

$$\langle x^2(t) \rangle = \frac{1}{N} \int_{\infty}^{\infty} n \, dx \, x^2 = 2 D t$$

in agreement with our earlier result

$$\overline{x^2} = \frac{\ell^2}{t_c} t \quad \text{if we identify } D = \frac{\ell^2}{2t_c}.$$

The ensemble of particles diffuses outwards.

Langevin equation

Given some initial velocity, a particle such as a pollen grain will slow and stop its net motion due to collisions with the surrounding molecules. If the particle is large compared to the mean free path of the molecules, so that they may be treated as a viscous fluid, the drag force is given by Stokes' law

$$F_{drag} = -6\pi\mu R v$$

particle velocity

dynamic viscosity

radius of spherical particle

we shall write this as $F_{drag} = -\dfrac{v}{B}$

where B is the "mobility".

Langevin wrote an equation of motion for a particle

$$M\frac{dv}{dt} = \underbrace{F(t)}_{\substack{\text{fluctuating}\\ \text{part of the force}}} - \underbrace{\frac{v}{B}}_{\substack{\text{average part}\\ = \text{viscous drag}}}$$

$$\overline{F(t)} = 0$$

If we take an ensemble average $\quad M\dfrac{d\langle v\rangle}{dt} = -\dfrac{\langle v\rangle}{B}$

$$\Rightarrow \quad \langle v\rangle = v(0)\, e^{-t/\tau}$$

where $\tau = MB$ is the relaxation time

$\Rightarrow$ the mean velocity decays to zero from its initial value.

The instantaneous acceleration is $\quad \dfrac{d\underline{v}}{dt} = -\dfrac{\underline{v}}{\tau} + \underline{A}(t) \quad -(*)$

where $\underline{A}(t) = \underline{F}(t)/M$ and $\overline{\underline{A}(t)} = 0$.

Take $\langle \underline{r} . (*) \rangle$:

LHS $$\left\langle \underline{r} . \frac{d\underline{v}}{dt} \right\rangle = \left\langle \frac{d}{dt}(\underline{r}.\underline{v}) \right\rangle - \langle v^2 \rangle$$

$$= \left\langle \frac{1}{2}\frac{d^2 r^2}{dt^2} \right\rangle - \langle v^2 \rangle \qquad \text{using } \underline{r}.\underline{v} = \frac{1}{2}\frac{dr^2}{dt}$$

$$= \frac{1}{2}\frac{d^2\langle r^2\rangle}{dt^2} - \langle v^2 \rangle$$

RHS $$\underbrace{-\frac{\langle \underline{r}.\underline{v}\rangle}{\tau}}_{-\frac{1}{\tau}\frac{1}{2}\frac{d\langle r^2\rangle}{dt}} + \underbrace{\cancelto{0}{\langle \underline{r}.\underline{A}\rangle}}_{\substack{\text{no correlation between} \\ \underline{A} \text{ and } \underline{r}}}$$

$$\Rightarrow \quad \boxed{\frac{d^2\langle r^2\rangle}{dt^2} + \frac{1}{\tau}\frac{d\langle r^2\rangle}{dt} = 2\langle v^2\rangle} \quad \text{— } (\dagger)$$

To solve this, let's assume the particles has come into equilibrium with the surrounding molecules $\langle v^2 \rangle = \frac{3k_BT}{M}$. Then the RHS is constant, and $(\dagger)$ can be integrated:

$$\boxed{\langle r^2 \rangle = \frac{6k_BT}{M}\tau^2\left[\frac{t}{\tau} - (1 - e^{-t/\tau})\right]}$$

where we take $\langle r^2 \rangle = 0$ at $t = 0$

$$\frac{d\langle r^2\rangle}{dt} = 2\langle \underline{r}.\underline{v}\rangle = 0 \quad \text{at} \quad t = 0$$

The short time behavior $(t \ll \tau)$ is $\langle r^2 \rangle = \frac{3 k_B T}{M} t^2$

$$= \langle v^2 \rangle t^2$$

"ballistic motion"

the long time behavior $(t \gg \tau)$ is

$$\langle r^2 \rangle = \frac{6 k_B T \tau}{M} t$$

$$= (6 k_B T B) t \quad \propto t$$

"diffusion"

Comparing to our results from last time, we see that

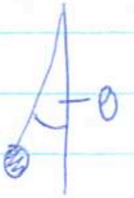
$$\boxed{D = B k_B T}$$

This is an example of the FLUCTUATION - DISSIPATION THEOREM

relates fluctuations of the equilibrium state

dissipative process related to how a system comes into equilibrium.

As another example, consider a pendulum suspended in a surrounding medium.



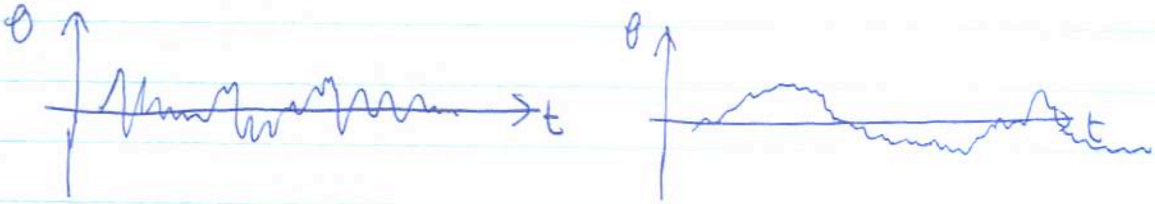
In thermal equilibrium, $P.E. = \frac{1}{2} mg\ell \langle \theta^2 \rangle = \frac{1}{2} k_B T$

$$\Rightarrow \langle \theta^2 \rangle = \frac{k_B T}{mg\ell} \qquad \left(\langle \theta \rangle = 0 \right)$$

This depends only on T. So for example if the density of the surrounding medium changes, we must get the same answer. This implies a direct connection between the fluctuations and the damping.

eg. if the damping is smaller $\Rightarrow$ each impact has more effect $\Rightarrow$

to get the same amplitude there must be fewer fluctuations.



See this in the time series - $\langle\theta^2\rangle$ is the same, but different time structure.

So we see that the kind of relations we've been finding between fluctuations and susceptibilities are extremely powerful because by looking at fluctuations around the equilibrium state we can learn about how a non-equilibrium system (not too far from eqm.) relaxes back towards equilibrium.

Exact solution of 1D Ising model (Transfer matrix method)

The partition function is

$$Q = \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \exp\left[\beta \mu H \sum_{i=1}^{N} s_i + \beta J \sum s_i s_j \right]$$

in 1D with periodic b.c.'s, we can write

$$Q = \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \exp\left[\beta \sum_{i=1}^{N} \left(\frac{1}{2} \mu H (s_i + s_{i+1}) + J s_i s_{i+1} \right)\right]$$

Rewrite this in a matrix formalism $s_i = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

define $$\underline{\underline{q}} = \begin{pmatrix} e^{\beta(J+\mu H)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu H)} \end{pmatrix}$$

then $$Q = \sum_{s_1} \cdots \sum_{s_N} \langle s_1 | q | s_2 \rangle \langle s_2 | q | s_3 \rangle \langle s_3 | q | s_4 \rangle \cdots \langle s_N | q | s_1 \rangle$$

$$= \sum_{s_1=\pm 1} \langle \sigma_1 | q^N | \sigma_1 \rangle$$

$$= \mathrm{Trace}(q^N)$$

$$= \lambda_1^N + \lambda_2^N$$ where λ_1 and λ_2 are the eigenvalues of matrix q.

To find λ_1, λ_2: $$\begin{vmatrix} e^{\beta(J+\mu H)} - \lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu H)} - \lambda \end{vmatrix} = 0$$

$$\Rightarrow \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = e^{\beta J} \cosh(\beta \mu H) \pm \left[e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta \mu H) \right]^{1/2}$$

Now, since $\lambda_2 < \lambda_1$, $\left(\frac{\lambda_2}{\lambda_1}\right)^N \to 0$ for $N \gg 1$ and so we need only take the 1st eigenvalue.

$$\Rightarrow \frac{\ln Q}{N} = \ln\left[e^{\beta J}\cosh(\beta\mu H) + \left\{ e^{-2\beta J} + e^{2\beta J}\sinh^2(2\mu H)\right\}^{1/2}\right]$$

$$A = -NJ - NkT\ln\left[\cosh(\beta\mu H) + \left\{e^{-4\beta J} + \sinh^2(\beta\mu H)\right\}^{1/2}\right]$$

The magnetization is

$$M = -\left(\frac{\partial A}{\partial H}\right)_T$$

$$\boxed{M = \frac{N\mu\sinh(\beta\mu H)}{\left\{e^{-4\beta J} + \sinh^2(\beta\mu H)\right\}^{1/2}}}$$

$\Rightarrow$ as $H \to 0$ then $M \to 0$ <u>no spontaneous magnetization</u>.

For zero field, $Q = \left[2\cosh(\beta J)\right]^N$

the energy is $E = -\frac{\partial \ln Q}{\partial \beta} = -\frac{\partial}{\partial\beta} N\ln(2\cosh(\beta J))$

$$= -\frac{N}{2\cosh(\beta J)} J\, 2\sinh(\beta J)$$

$$\boxed{E = -NJ\tanh(\beta J)}$$

PHYS 362 Problem Set 1

This problem set is due Monday January 19, 2009, either in class or by 5pm in ERP 219.

1. **Collisions of air molecules**

In class I mentioned that air molecules undergo collisions every $\sim 10^{-9}$ s on average. Do you believe me? Estimate this number for yourself. Don't worry about factors of two - you just want to check the order of magnitude. Your starting points should be the density and temperature of air in the room, and the typical size and mass of an air molecule.

2. **Molecules in a room**

The lecture room contains approximately 10^{28} air molecules. How many times more likely is it that the air molecules are split evenly between the front and back of the room than that there is a 49-51% split? What is the biggest front-to-back asymmetry that you expect to occur during a one-hour lecture on average?

3. **A simple model of a rubber band**

A polymer molecule is a long chain molecule made of subunits that can change their relative orientations. Fixing one end of the molecule at the origin, you can think of the subunits making a random walk in space.

(a) Make a simple 1D model of such a molecule, by assuming the molecule lies along the x-axis and that individual subunits point either forwards or backwards along the x-axis (don't worry about the fact that the molecule can double back on itself, we'll allow that to happen in our simple model). Each subunit has a length d, and therefore the total length of the molecule is $L = d(n_+ - n_-)$, where n_+ and n_- are the number of subunits that point in the forward and backward directions, respectively. Derive an expression for the number of configurations Ω and therefore the entropy as a function of the length of the molecule L and the total number of subunits N. You may assume N is large and use Stirling's approximation.

(b) Argue that there must be a tension in the molecule, and calculate that tension.

(c) This model can be taken as a simple model of a rubber band, which is made up of many long polymer molecules. What do you predict should happen to the length of a rubber band when it is heated?

(d) Where does this tension come from physically? Give two different qualitative arguments, one based on how entropy changes as the length of the molecule changes, and one based on kinetic theory ideas.

[You may find it helpful for this question to think about the analogy with volume changes and the pressure of a gas.]

PHYS 362 Homework 1 Solutions

1. Collisions of Air Molecules

The speed of air molecules is roughly given by setting

$$\frac{3}{2} kT = \frac{1}{2} m v^2$$

$$\Rightarrow \quad v = \left(\frac{3kT}{m}\right)^{1/2}$$

Take $T = 300\,\mathrm{K}$ $m = 28\, m_p$ (N_2 molecule)

$$\Rightarrow \quad \underline{v = 515\ \mathrm{m/s}} \simeq 500\ \mathrm{m/s}.$$

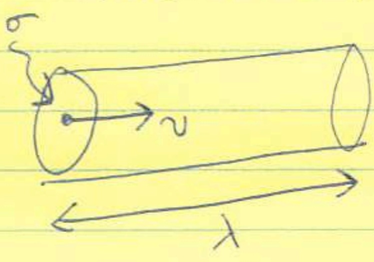
(this sounds about right because remember the speed of sound in air is 330 m/s)

The number density is $n = \dfrac{\rho}{28\, m_p} = \dfrac{1\ \mathrm{kg/m^3}}{28\, m_p}$

$$\Rightarrow \underline{n \approx 2 \times 10^{25}\ \mathrm{m^{-3}}}$$

The size of a molecule should be $\sim 10^{-10}$ m or perhaps slightly larger, let's say 3×10^{-10} m.
Then the cross-section is $\underline{\sigma \sim 10^{-19}\ \mathrm{m^2}}$.

Put these together: the mean free path is given by



$$n \sigma \lambda = 1 \quad \Rightarrow \quad \lambda = \frac{1}{n\sigma} = 0.5\,\mu\mathrm{m}.$$

The time between collisions is $\dfrac{\lambda}{v} = \underline{10^{-9}\,\mathrm{s}}$

2. Molecules in a room

Since the volume of each half of the room is the same, we can just do this by counting. The number of ways of selecting N_1 molecules out of N to go into the front half is $\binom{N}{N_1} = \frac{N!}{N_1!(N-N_1)!}$.

$$\Rightarrow \frac{\Omega_{50/50}}{\Omega_{51/49}} = \left[\frac{\left(\frac{N}{2}\right)!\left(\frac{N}{2}\right)!}{(Nf)!\,(N(1-f))!} \right]^{-1} \qquad \text{where } f = 0.49$$

Use Stirling's approximation $\ln N! \simeq N\ln N - N$

$$\Rightarrow \boxed{\ln\left(\frac{\Omega_{50/50}}{\Omega_{51/49}}\right) = N\left[f\ln f + (1-f)\ln(1-f) + \ln 2 \right]}$$

For $f = 0.49$, the term in brackets is 0.0002

$\Rightarrow$ the 50/50 arrangement is more likely by a factor $e^{0.0002N}$

or $\underline{\underline{e^{2\times10^{24}}}}$!

[If the volumes of the two partitions were different, we would have to include factors of V^N for each partition. But here, since the V's are the same, these volume factors cancel.]

An alternative way to write this is $\underline{\underline{(1.0002)^{10^{28}}}}$

(since $(1+x)^N \simeq e^{Nx}$ for $\substack{N \gg 1 \\ (x \ll 1)}$)

For the last part,

expand $\ln\left(\frac{\Omega_{50/50}}{\Omega(f)}\right)$ around $f=\frac{1}{2}$

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

then $f\ln f + (1-f)\ln(1-f)$

$$= \left(x+\tfrac{1}{2}\right)\ln\left(x+\tfrac{1}{2}\right) + \left(\tfrac{1}{2}-x\right)\ln\left(\tfrac{1}{2}-x\right)$$

$$= x\ln\left(\frac{1+2x}{1-2x}\right) + \underbrace{\frac{1}{2}\ln\left[\left(\tfrac{1}{2}+x\right)\left(\tfrac{1}{2}-x\right)\right]}_{}$$

$$+\underbrace{\frac{1}{2}\ln\left(\frac{1}{4}\right)}_{-\ln 2} + \frac{1}{2}\ln\left((1+2x)(1-2x)\right)$$

$$\Rightarrow \quad \ln\left(\frac{\Omega_{50/50}}{\Omega(f)}\right) = N\left[x\ln\left(\frac{1+2x}{1-2x}\right) + \frac{1}{2}\ln\left((1+2x)(1-2x)\right)\right]$$

for small x, $\ln\left(\frac{\Omega_{50/50}}{\Omega(f)}\right) = N\left[4x^2 \cancel{+} -2x^2\right]$

$$\simeq 2Nx^2$$

if during 1 hour there are N_{states} visited we expect on average the maximum fluctuation has

$$2Nx^2 = \ln N_{\text{states}}$$

or $\boxed{x = \frac{1}{\sqrt{2N}}\sqrt{\ln N_{\text{states}}}}$

A typical molecule has $v \simeq 300\,\text{m/s}$ ($\simeq$ sound speed)

and room size $= 10\,\text{m}$

$$\Rightarrow \quad N_{\text{states}} \simeq \cancel{\frac{80}{300}} \cancel{\times}\, 3600 \times \frac{300}{10}$$

$$\simeq \cancel{100}\ \underline{\underline{10^5}} \text{ per molecule}$$

$$\Rightarrow \quad N_{\text{states}} \simeq N\,10^5$$

$$\Rightarrow \quad x = \frac{1}{\sqrt{2N}} \sqrt{\ln 10^5 N}$$

for $N = 10^{28}$, I get $x = \left(\sqrt{\frac{33 \times 2.303}{2}}\right) 10^{-14}$

$$\simeq \underline{\underline{10^{-13}}}.$$

$\Rightarrow$ <u>The maximum fluctuation during a 1 hour lecture is about 1 part in 10^{13}.</u>

3. A simple model of a rubber band

(a) The number of microstates is $\binom{N}{n_+} = \dfrac{N!}{n_+!\,(N-n_+)!}$

where $N = n_+ + n_-$.
In terms of the length L, $n_+ = \dfrac{N}{2} + \dfrac{L}{2d}$

The entropy is $S = k_B \ln \Omega$

use Stirling's approximation

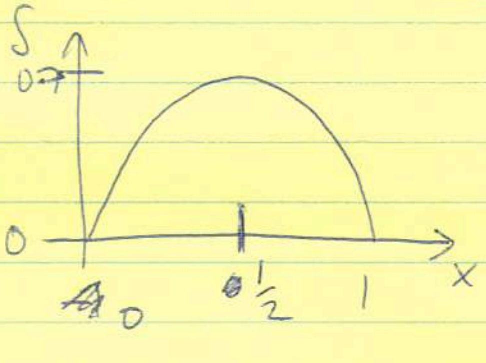
$$\ln\left(\frac{N!}{n_+!(N-n_+)!}\right) = N\ln N - n_+ \ln n_+ - (N-n_+)\ln(N-n_+$$

$$\Rightarrow \quad S = k_B N\left[\ln N - \frac{n_+}{N}\ln n_+ - \left(1-\frac{n_+}{N}\right)\ln\left(N-n_+\right)\right]$$

$$S = -k_B N\left[\frac{n_+}{N}\ln\left(\frac{n_+}{N}\right) + \left(1-\frac{n_+}{N}\right)\ln\left(1-\frac{n_+}{N}\right)\right]$$

define $x = \dfrac{n_+}{N} = \dfrac{1 + L/Nd}{2}$ $\qquad 0<|x|<1$

then $$\boxed{S = -k_B N\left[x\ln x + (1-x)\ln(1-x)\right]}$$



— when fully stretched ($x = 0$ or $x = 1$)
the entropy is 0 (only 1 configuration)
— maximum entropy for $x =$ ~~0~~ $\frac{1}{2}$

(b) The entropy depends on length – therefore there must be a tension.

The 1st law for this system is $dE = TdS + FdL$

$$\Rightarrow \quad F = -T\frac{\partial S}{\partial L}\bigg|_E$$

$$= -T\left(\frac{\partial S}{\partial x}\right)\frac{1}{2Nd}$$

$$\Rightarrow \boxed{F = \frac{\cancel{N}k_BT}{2\cancel{N}d}\ln\left(\frac{x}{1-x}\right)}$$

In practise, a polymer molecule will have $L \ll Nd$

or $x = \frac{1}{2} + \varepsilon \qquad \varepsilon \ll 1 \qquad \varepsilon = \frac{L}{2Nd}$

expand: $$F = \frac{k_BT}{2d}\ln\left[\frac{\frac{1}{2}+\varepsilon}{\frac{1}{2}-\varepsilon}\right]$$

$$\approx \underline{\underline{\frac{k_BTL}{Nd^2}}}$$ like a spring force ∝ length

(c)

① the entropy increases as the length decreases

⇒ the molecule "wants to" become shorter

② the molecules making up the polymer are continuously undergoing thermal motions as the molecule moves between microstates. These motions give rise to a force on the end of the molecule.

(d) If the rubber band is held under constant tension, then as temperature goes up the rubber band will *shrink*.

Here, we repeat question 2 using the full Stirling's approximation. This allows us to get the total number of states right.

Stirling's approximation is $\boxed{N! \simeq \sqrt{2\pi N}\left(\frac{N}{e}\right)^N}$

or $\ln N! = N\ln N - N + \ln\sqrt{2\pi N}$

We have $\Omega(m) = \frac{N!}{(N-m)!\,m!}$

$$\ln\Omega = N\ln N + \ln\sqrt{2\pi N} - (N-m)\ln(N-m)$$
$$- \ln\sqrt{2\pi(N-m)} - m\ln m - \ln\sqrt{2\pi m}$$

$$= -N\left[(1-f)\ln(1-f) + f\ln f\right] + \frac{1}{2}\ln\left(\underbrace{\frac{\cancel{N}\,1}{(1-f)\,fN}}_{\frac{1}{Nf(1-f)}}\frac{1}{\sqrt{2\pi}}\right)$$

$$\boxed{\ln\Omega(f) = -N\left[(1-f)\ln(1-f) + f\ln f\right] + \frac{1}{2}\ln\left(\frac{1}{Nf(1-f)\sqrt{2\pi}}\right)}$$

Let's check $\Omega(50/50) = \exp\left[+N\ln 2\right]\frac{1}{\sqrt{N/4\,2\pi}}$

$$= \frac{\cancel{4}\,2^N}{\sqrt{N/4}} = \underline{\underline{\frac{2^{N+1}}{\sqrt{N2\pi}}}} \quad ✓$$

which matches $\lim_{N\to\infty}\left(\frac{N!}{\left(\frac{N}{2}\right)!\left(\frac{N}{2}\right)!}\right)$

Now check the total number of states.

$$\Omega_{tot} = \sum_{m=0}^{N} \frac{N!}{(N-m)!\,m!} = 2^N$$

as expected
since each particle
can go left or right.

In terms of ~~an~~ integral

$$\Omega_{tot} = N\int_0^1 df\ \Omega(f)$$

$$= N\int_0^1 df\ \frac{\exp\left[-N\left(f\ln f + (1-f)\ln(1-f)\right)\right]}{\sqrt{2\pi N}\ \sqrt{f(1-f)}}$$

Now write $x = f - \tfrac{1}{2}$

$\Rightarrow f = x + \tfrac{1}{2}$

$1 - f = \tfrac{1}{2} - x$

$$N\left(f\ln f + (1-f)\ln(1-f)\right)$$

$$= Nx\ln\left(\frac{\frac{1}{2}+x}{\frac{1}{2}-x}\right) + \underbrace{\frac{N}{2}\ln\left[f(1-f)\right]}$$

~~As cancels the denominator!~~

then

$$\boxed{\Omega_{tot} = \frac{N}{\sqrt{2\pi N}} \int_{-1/2}^{1/2} dx\ \frac{\exp\left[-Nx\ln\left(\frac{1+2x}{1-2x}\right)\right]}{\left(\sqrt{f(1-f)}\right)^{N+1}}} \quad \text{—(*)}$$

$$\cancel{\neq \sqrt{\frac{N}{2\pi}} \int_{-1/2}^{1/2} dx \left(\frac{1+2x}{1-2x}\right)^{-Nx}}$$

$y = Nx$

$$\cancel{= \sqrt{\frac{1}{2\pi N}} \int_{-N/2}^{N/2} dy \left(\frac{1+2y/N}{1-2y/N}\right)^{-y}}$$

$$\Rightarrow \Omega_{tot} = \sqrt{\frac{N}{2\pi}} \int_{-1/2}^{1/2} dx \; \frac{\exp\left(-Nx \ln\left(\frac{1+2x}{1-2x}\right)\right)}{\underset{(1+2x)(1-2x)}{\left(\cancel{(x+1/2)(\tfrac{1}{2}-x)}\right)^{\frac{N+1}{2}}}} \quad 2^{N+1}$$

$$= 2^N \sqrt{\frac{2N}{\pi}} \int_{-1/2}^{1/2} dx \; \underbrace{\left(\frac{1+2x}{1-2x}\right)^{-Nx} \frac{1}{\left[(1+2x)(1-2x)\right]^{\frac{N+1}{2}}}}_{(1+2x)^{-Nx-\frac{N+1}{2}} \; (1-2x)^{Nx-\frac{N+1}{2}}}$$

$$\frac{\Omega_{tot}}{2^N} = \left(\frac{2N}{\pi}\right)^{1/2} \int_{-1/2}^{1/2} dx \; \exp\left[-Nx \ln\left(\frac{1+2x}{1-2x}\right) - \left(\frac{N+1}{2}\right) \ln\left((1+2x)(1-2x)\right)\right]$$

Now assume the integrand is strongly peaked near $x = 0$.

$$\int_{-1/2}^{1/2} dx \; \exp\Big[\underbrace{-N4x^2 + \boxed{\left(\frac{N+1}{2}\right)} \cancel{4}x^2}_{4x^2\left(\frac{1-N}{2}\right)}\Big]$$

$$\simeq -2Nx^2$$

$$\int_{-1/2}^{1/2} dx \; \exp(-2Nx^2) \simeq \int_{-N/2}^{N/2} \frac{dy}{\sqrt{2N}} \cdot e^{-y^2} \simeq \underline{\underline{\sqrt{\frac{\pi}{2N}}}} \; \checkmark$$

$$\Rightarrow \quad \underline{\underline{\Omega_{tot} = 2^N}} \quad \checkmark$$

PHYS 362 Problem Set 2

This problem set is due Wednesday January 28th 2009, either in class or by 5pm in ERP 219.

1. **Hard sphere gas**

Consider a gas of N indistinguishable hard spheres each with radius r occupying a volume V. Write down the number of spatial microstates Ω available to the gas, taking into account the reduced volume that each new molecule added to the box has available to it because of the space taken up by the other spheres. Assume that the total volume of the spheres is much smaller than the volume of the box. Ignore any effects arising from the edges of the box. The result

$$\sum_{n=1}^{N} \log\left(1-(n-1)\epsilon\right) = N \log\left(1-\frac{(N-1)\epsilon}{2}\right) + O\left(\epsilon^2\right)$$

will be useful. Show that the equation of state for the gas is $P = Nk_BT/(V-b)$ where b is four times the total volume of the spheres.

2. **DNA zipper**

A DNA molecule consists of two polymer chains joined by N links. A link can be opened only if a neighbouring link is already open (or if the link is at the end of the molecule). In this way, it behaves like a zipper. Opening a link takes an energy ϵ.

(a) First consider the case where the link at one end of the molecule is held closed, so that the molecule can unzip only from the other end. Write down the partition function, and calculate the mean number of open links as a function of temperature. What is the mean number of open links when $k_BT \ll \epsilon$ and when $k_BT \gg \epsilon$?

(b) Repeat part (a) for the case where the molecule can unzip from either end.

3. **The Gibbs ensemble**

A system is in equilibrium with a bath with which it is able to exchange energy and volume. Write down the probability of finding the system in a particular microstate, and the corresponding partition function. Show that the partition function can be expressed in terms of the Gibbs free energy $G = E - TS + PV$.

PHYS 362 Problem Set 2 Solutions

1. Hard sphere gas

A sphere of radius r excludes a volume $8 \times \frac{4\pi r^3}{3}$

Now add the spheres one by one:

$$\Omega = \underset{\text{indistinguishable particles}}{\frac{1}{N!}} \prod_{n=0}^{N-1} \left[V - n\frac{4\pi}{3}(2r)^3 \right]$$

$$\Rightarrow \log \Omega = \sum_{n=0}^{N-1} \log\left[V - n\frac{32\pi r^3}{3} \right] - N\log N + N$$

$\frac{\pi r^3}{V} \ll 1$

$$= N\log V + \underbrace{\sum_{n=1}^{N} \log\left(1 - \frac{(n-1)}{V}\frac{32\pi r^3}{3} \right)}_{\text{apply identity}} + N - N\log N$$

$$= N\log\left(1 - \frac{(N-1)}{2} \cdot \frac{32\pi r^3}{3V} \right)$$

$$\Rightarrow \boxed{\log\Omega = N\left[1 + \log\left(\frac{V}{N} - \frac{(N-1)}{N}\frac{16\pi r^3}{3} \right) \right]}$$

The pressure is $\boxed{P = T\frac{\partial S}{\partial V}\Big|_{N,E} = \frac{k_B T N}{V-b}}$ ✓ $\boxed{b = (N-1)\,4\times\frac{4\pi r^3}{3}}$ ✓

2. DNA zipper

(a) The microstates are

$n=0$ $E=0$

$n=1$ $E=\varepsilon$

$n=2$ $E=2\varepsilon$

etc.

$\Rightarrow$ partition function
$$Q = \sum_{n=0}^{N-1} e^{-n\varepsilon/kT}$$

This is a geometric series, which we can sum

$$\Rightarrow \boxed{Q = \frac{1-e^{N\varepsilon/kT}}{1-e^{-\varepsilon/kT}}}$$

The mean number of open links is

$$\langle n \rangle = \sum_{n=0}^{N-1} P_n n$$
$$= \frac{\sum_{n=0}^{N-1} n e^{-n\varepsilon/kT}}{Q}$$

define $x = e^{-\varepsilon/kT}$ $\Rightarrow$ $\langle n \rangle = \dfrac{\sum_{n=0}^{N-1} n x^n}{(1-x^N)/(1-x)}$

again we can sum this

$$\sum_{n=0}^{N-1} n x^n = \frac{x - Nx^N - x^{1+N} + Nx^{1+N}}{(x-1)^2}$$
$$= \left[-N(1-x)x^N - (x^N-1)x\right] / (1-x)^2$$

$$\Rightarrow \quad \langle n \rangle = \frac{-Nx^N(1-x) + (1-x^N)x}{(1-x)(1-x^N)}$$

$$\boxed{\langle n \rangle = \frac{x}{1-x} - \frac{Nx^N}{1-x^N}}$$

1) $k_B T \ll \varepsilon$ then $x \ll 1$ and $\boxed{\langle n \rangle \simeq x = e^{-\varepsilon/kT}}$ Cold

2) $k_B T \gg \varepsilon$ or $x \simeq 1$

write $\begin{pmatrix} y = 1-x \\ x = 1-y \end{pmatrix} \Rightarrow y \ll 1$

$$\langle n \rangle = \frac{1-y}{y} - \frac{N(1-y)^N}{1-(1-y)^N}$$

$$\Rightarrow \quad \boxed{\langle n \rangle \simeq \frac{N-1}{2} + O(y)} \quad \text{hot}$$

Approximately half the links are open when $k_B T \gg \varepsilon$. This makes sense - recall that in the two level system as $T \to \infty$ the 2 levels become equally populated.

[Note a faster way to get the expression for $\langle n \rangle$ is to find the energy $E = -\frac{\partial \ln Q}{\partial \beta} = -\frac{\partial}{\partial \beta}\left[\ln(1-e^{-\beta N\varepsilon}) - \ln(1-e^{-\beta\varepsilon})\right]$

$$= -\frac{N\varepsilon e^{-\beta N \varepsilon}}{1-e^{-\beta N\varepsilon}} + \frac{\varepsilon e^{-\beta\varepsilon}}{1-e^{-\beta\varepsilon}}$$

$$= \varepsilon\left[\frac{x}{1-x} - \frac{Nx^N}{1-x^N}\right]$$

$$= \varepsilon \langle n \rangle.$$

]

(b) all links can open. If n links are open, there are $n+1$ ways to arrange the links (eg. 0, 1, 2 ... n on the left, the remainder on the right).

$$\Rightarrow \quad Q = \sum_{n=0}^{N-1} (n+1) e^{-n\varepsilon\beta} + \underbrace{e^{-N\varepsilon\beta}}_{\text{all links open}}$$

$$\boxed{Q = \sum_{n=0}^{N-1} (n+1) x^n + x^N}$$

To get $\langle n \rangle$, we can write

$$\langle n \rangle = \frac{E}{\varepsilon} = -\frac{1}{\varepsilon}\frac{\partial \ln Q}{\partial \beta}$$

$$= -\frac{1}{\varepsilon}\frac{\partial x}{\partial \beta}\frac{\partial \ln Q}{\partial x}$$

$$\frac{\partial x}{\partial \beta} = -\varepsilon x$$

$$= x\frac{\partial \ln Q}{\partial x}$$

$$\Rightarrow \quad \boxed{\langle n \rangle = \frac{\partial \ln Q}{\partial \ln x}}$$

$$\Rightarrow \quad \langle n \rangle = x\frac{\partial}{\partial x}\left[\ln\left\{\sum_{n=0}^{N-1}(n+1)x^n + x^N\right\}\right]$$

$$\boxed{\langle n \rangle = \frac{\sum_{n=0}^{N-1}(n+1)n x^n + N x^N}{\sum_{n=0}^{N-1}(n+1)x^n + x^N}}$$

because in the sums we take the largest terms $\sum (n+1) n x^n \simeq 2x$

$\sum (1+n) x^n \simeq 1$

for $x \ll 1$ (cold) $\langle n \rangle \simeq 2x = 2e^{-\varepsilon/kT}$

for $x \simeq 1$ (hot) $\langle n \rangle \simeq \dfrac{\sum_0^{N-1} n(n+1) + N}{\sum_0^{N-1} (n+1)}$

$$\langle n \rangle = \frac{\frac{1}{3}N(N^2-1) + N}{N + \frac{N(N-1)}{2}}$$

$$= \frac{\frac{1}{3}(N^2-1) + 1}{\frac{1}{2}(N-1) + 1} = \frac{\frac{1}{3}N^2 + \frac{2}{3}}{\frac{1}{2}N + \frac{1}{2}}$$

$$= \frac{2N^2+4}{3N+3} \approx \underline{\underline{\frac{2N}{3}}} \quad \varepsilon.$$

$\frac{2}{3}$ of links are open on average at the highest temperatures.

3. The Gibbs Ensemble

Follow the same method as in class:

$$P_\nu \propto \Omega_{bath}(E-E_\nu, V-V_\nu)$$

Expand

$$\ln \Omega_{bath} \approx \ln \Omega_{bath}(E,V) \underbrace{- E_\nu \frac{\partial \ln \Omega_{bath}}{\partial E}}_{-\frac{E_\nu}{k_B T}} \underbrace{- V_\nu \frac{\partial \ln \Omega_{bath}}{\partial V}}_{\cancel{+} - \frac{PV_\nu}{k_B T}}$$

$$\Rightarrow \boxed{P_\nu \propto \exp\left(-\frac{(E_\nu + PV_\nu)}{k_B T}\right)}$$

Normalize by the partition function

$$\boxed{Z = \sum_\nu \exp\left(-\frac{(E_\nu + PV_\nu)}{k_B T}\right)}$$

We want to write this in terms of $G = E - TS + PV$

First, calculate $\frac{1}{Z}\frac{\partial Z}{\partial \beta} = \frac{-1}{Z}\sum_\nu (E_\nu + PV_\nu)\exp\left(-\frac{(E_\nu + PV_\nu)}{kT}\right)$

$$= -E \cancel{+} - PV$$

actually we don't need this!

Now calculate $TS = -k_B T \sum P_\nu \ln P_\nu$

$$= -k_B T \sum P_\nu \left(-\frac{(E_\nu + PV_\nu)}{kT} - \ln Z \right)$$

$$TS = E + PV + k_B T \ln Z \Rightarrow \boxed{G = -k_B T \ln Z}$$

PHYS 362 Problem Set 3

Due Friday February 6th 2009, either in class or by 5pm in ERP 219.

1. **Diatomic gas.**

A diatomic molecule has vibrational energy levels $\epsilon_n = (n + 1/2)\hbar\omega$, $n = 0, 1, 2...$, and rotational energy levels $\epsilon_l = l(l+1)\hbar^2/2I$ with degeneracy $2l + 1$, where I is the moment of inertia of the molecule and $l = 0, 1, 2....$

(a) Show that the contribution to the heat capacity per particle from the vibrational modes is

$$C_V = k_B x \left(\frac{\ln x}{1-x}\right)^2$$

where $x = \exp(-\hbar\omega/k_BT)$. What are the low and high temperature limits?

(b) Derive the contribution to the heat capacity from the rotational modes in the low and high temperature limits by either (1) taking only the first two terms in the partition function sum, or (2) approximating the sum as an integral, explaining which of these methods is appropriate for which limit and why.

(c) Estimate the vibrational and rotational energy splittings $\hbar\omega$ and $\hbar^2/2I$ for a hydrogen molecule H_2. Use these estimates to sketch a graph of heat capacity against temperature for H_2 gas, labelling each axis as quantitatively as you can. (Assume that the translational degrees of freedom always contribute $(3/2)k_B$ per particle to the heat capacity.)

2. **Two-state RNA**

Read the attached pages which are taken from the book "Biological Physics: Energy, Information, Life" by Philip Nelson. They describe an experiment which investigates an RNA molecule that exhibits "two-state" behavior, in which the hairpin turn of the molecule is either open or closed.

(a) In fact, the molecule has many more than two microstates available to it, but the idea is that we can categorize the microstates as either being associated with the hairpin being "open" or associated with it being "closed". If the molecule is in contact with a heat bath at temperature T, show that the probability of the hairpin being open is

$$\frac{1}{e^{\Delta F/k_BT} + 1}$$

where ΔF is the free energy difference between the open and closed states. (Hint: divide the partition function into two pieces).

(b) Now the ends of the molecule are pulled with a force f that is kept constant during the experiment. By using an appropriate ensemble (hint: something similar to the Gibbs ensemble from problem set 2), derive an expression for the probability that the molecule is in the "closed" state. You should find the same formula that is plotted in Figure 6.10b of Nelson's book. [The good agreement between this formula and the experimental data is used to argue that indeed we can divide the microstates of this molecule into "open" and "closed" states - the molecule acts as a two-state system.]

3. **Heat capacity of a Fermi gas**

Show that the heat capacity of a degenerate gas of non-relativistic electrons is

$$c_V = \frac{\pi^2}{2} n k_B \left(\frac{k_B T}{E_F} \right).$$

6.7 EXCURSION: "RNA FOLDING AS A TWO-STATE SYSTEM" BY J. LIPHARDT, I. TINOCO, JR., AND C. BUSTAMANTE

Recently, we set out to explore the mechanical properties of RNA, an important biopolymer. In cells, RNA molecules store and transfer information, and catalyze biochemical reactions. We knew that numerous biological processes like cell division and protein synthesis depend on the ability of the cell to unfold RNA (as well as to unfold proteins and DNA) and that such unfolding involves mechanical forces, which one might be able to reproduce by using biophysical techniques. To investigate how RNA might respond to mechanical forces, we needed to find a way to grab the ends of individual molecules of RNA. Then we wanted to pull on them and watch them buckle, twist, and unfold under the effect of the applied external force.

We used an **optical tweezer** apparatus, which allows small objects, like polystyrene beads with a diameter of $\approx 3\,\mu\mathrm{m}$, to be manipulated by using light (Figure 6.9). Although the beads are transparent, they do bend incoming light rays, transferring some of the light's momentum to each bead, which accordingly experiences a force. A pair of opposed lasers, aimed at a common focus, can thus be used to hold the beads in prescribed locations. Because the RNA is too small to be trapped by itself, we attached it to molecular "handles" made of DNA, which were chemically modified to stick to specially prepared polystyrene beads (Figure 6.9, inset). As sketched in the inset, the RNA sequence we studied has the ability to fold back on itself, thereby forming a "hairpin" structure (see Figure 2.16 on page 52).

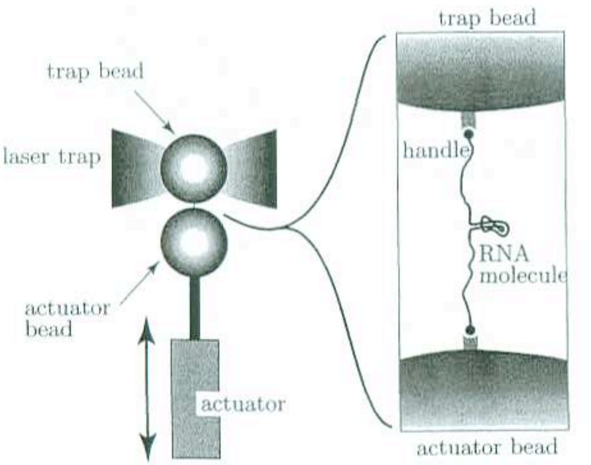


Figure 6.9: (Schematic.) Optical tweezer apparatus. A piezoelectric actuator controls the position of the bottom bead. The top bead is captured in an optical trap formed by two opposing lasers, and the force exerted on the polymer connecting the two beads is measured from the change in momentum of light that exits the optical trap. Molecules are stretched by moving the bottom bead vertically. The end-to-end length of the molecule is obtained as the difference of the position of the bottom bead and the top bead. *Inset:* The RNA molecule of interest is coupled to the two beads via DNA "handles." The handles end in chemical groups that stick to complementary groups on the bead. The drawing is not to scale: Relative to the diameter of the beads ($\approx 3000\,\mathrm{nm}$), the RNA is tiny ($\approx 20\,\mathrm{nm}$). [Figure kindly supplied by J. Liphardt.]

When we pulled on the RNA via the handles, we saw the force initially increase smoothly with extension (Figure 6.10a, black curve), just as it did when we pulled on the handles alone: The DNA handles behaved much like a spring (a phenomenon to be discussed in Chapter 9). Then, suddenly, at $f = 14.5\,\mathrm{pN}$, there was a small discontinuity in the force-extension curve (points labeled *a* and *b*). The change in length ($\Delta z \approx 20\,\mathrm{nm}$) of that event was consistent with the known length of the part of the RNA that could form a hairpin. When we reduced the force, the hairpin refolded and the handles contracted. Different samples gave slightly different values for the critical force, but in every case it was sharply defined.

To our surprise, the observed properties of the hairpin were entirely consistent with those of a two-state system. Even though the detailed energetics of RNA folding are known to be rather complex, involving hydration effects, Watson–Crick base-pairing and charge shielding by ions, the overall behavior of the RNA hairpin under external force was that of a system with just two allowed states, folded and unfolded. We stretched and relaxed the RNA hairpin many times and then plotted the fraction of folded hairpins versus force (Figure 6.10b). As the force increased, the fraction folded decreased, and that decrease could be fit to a model used to describe two-state systems (Equation 6.34 and Figure 6.10b, inset). Just as an external magnetic field can be used to change the probability of an atomic magnet to point up or down,[7] the work done by the external force ($f\,\Delta z$) was apparently changing the free energy difference $\Delta F = F_{\mathrm{open}} - F_{\mathrm{closed}}$ between the two states and thus controlling the probability $P(f)$ of the hairpin being folded. But if the ΔF could be so easily manipulated by changing the external force, it might be possible to watch a hairpin "hop" between the two states if we tuned the strength of the external force to the right critical value (such that $P(f) \approx \frac{1}{2}$) and held it there by force-feedback.

Indeed, about one year after starting our RNA unfolding project, we were able to observe this predicted behavior (Figure 6.10c). After showing RNA hopping to everyone who happened to be in the Berkeley physics building that night, we began to investigate this process more closely to see how the application of increasing force tilts the equilibrium of the system toward the longer, unfolded form of the molecule. At forces slightly below the critical force, the molecule stayed mostly in the short folded state except for brief excursions into the longer unfolded state (Figure 6.10c, lower curves). When the force was held at 14.1 pN, the molecule spent roughly equal times in either state ($\approx 1\,\mathrm{s}$). Finally, at 14.6 pN, the effect was reversed: The hairpin spent more time in the extended, unfolded form and less time in the short, folded form. Thus, it is possible to control the thermodynamics and kinetics of the folding reaction in real time, simply by changing the external force. The only remaining question had to do with the statistics of the hopping reaction. Was RNA hopping a simple process characterized by a constant probability of hopping per unit time at a given force? It appears so: Histograms of the dwell times can be fit to simple exponentials (see Figure 6.10d and Equation 6.31).

---

[7] See Problem 6.5.

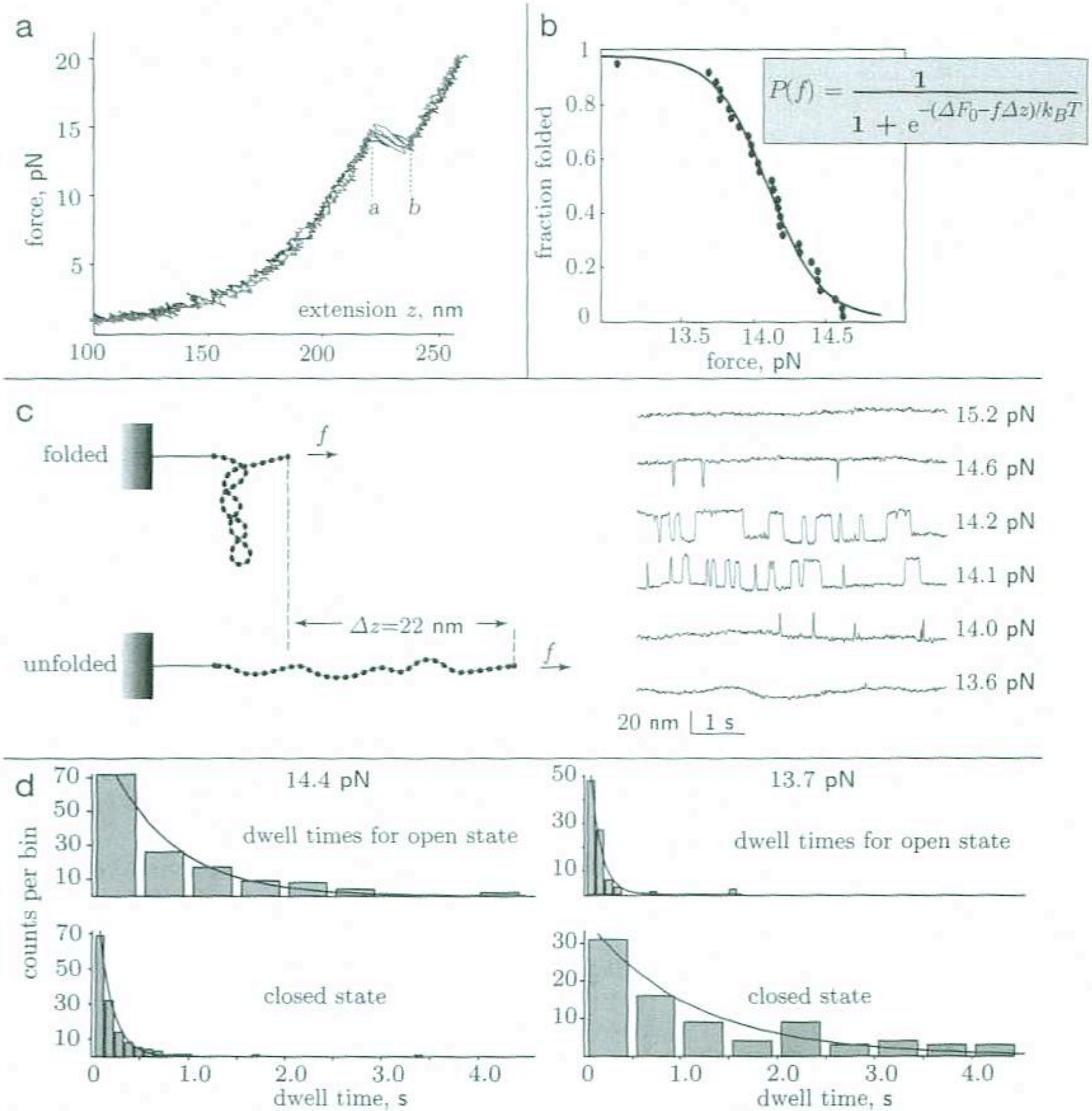


Figure 6.10: (Experimental data) (a) Force-extension curves of an RNA hairpin with handles. Stretching (*black*) and relaxing (*gray*) curves are superimposed. Hairpin unfolding occurs at about 14.5 pN (labeled a). (b) Fraction $P(f)$ of hairpins folded versus force. Data (*filled circles*) are from 36 consecutive pulls of a single RNA hairpin. *Solid line*, probability versus force for a two-state system (see Equation 6.34 on page 225). Best-fit values, $\Delta F_0 = 79 k_B T_r$, $\Delta z = 22$ nm, consistent with the observed Δz seen in panel (a). (c) Effect of mechanical force on the rate of RNA folding. Length versus time traces of the RNA hairpin at various constant forces. Increasing the external force increases the rate of unfolding and decreases the rate of folding. (d) Histograms of the dwell times in the open and closed states of the RNA hairpin at two different forces ($f = 14.4$ and 13.7 pN). The solid lines are exponential functions fit to the data (see Equation 6.31), giving rate constants for folding and unfolding. At 13.7 pN, the molecule is mostly folded, with $k_{\mathrm{open}} = 0.9\,\mathrm{s}^{-1}$, and $k_{\mathrm{fold}} = 8.5\,\mathrm{s}^{-1}$. At 14.4 pN, the unfolded state predominates, with $k_{\mathrm{open}} = 7\,\mathrm{s}^{-1}$ and $k_{\mathrm{fold}} = 1.5\,\mathrm{s}^{-1}$. [Figure kindly supplied by J. Liphardt.]

PHYS 362 Problem Set 3 Solutions

1. Diatomic gas

(a) vibrational modes $\varepsilon_n = (n+\tfrac{1}{2})\hbar\omega$

$$Q = \sum_{n=0}^{\infty} e^{-\varepsilon_n/kT} = \sum_{n=0}^{\infty} e^{-\hbar\omega(n+\frac{1}{2})/kT}$$

$$= \sum_{n=0}^{\infty} e^{-\hbar\omega/2kT}\left(e^{-\hbar\omega/kT}\right)^n$$

$$= \cancel{e}\; x^{1/2}\,\frac{1}{1-x} \qquad x = e^{-\beta\hbar\omega}$$

$$\frac{\partial x}{\partial\beta} = -\hbar\omega x$$

$$E = -\frac{\partial \ln Q}{\partial\beta} = -\frac{\partial}{\partial\beta}\left[\tfrac{1}{2}\ln x - \ln(1-x)\right]$$

$$= +\frac{1}{2x}\hbar\omega x + \frac{1}{1-x}\hbar\omega x$$

$$= \frac{\hbar\omega}{2} + \frac{\hbar\omega x}{1-x}$$

$$C_v = \frac{\partial E}{\partial T} = -\frac{1}{kT^2}\frac{\partial E}{\partial\beta} = \frac{\hbar\omega}{kT^2}\, x\, \frac{\partial E}{\partial x}$$

$$\frac{\partial E}{\partial x} \longrightarrow \underbrace{\frac{\hbar\omega}{1-x} + \frac{\hbar\omega x}{(1-x)^2}}_{\frac{\hbar\omega}{(1-x)^2}}$$

$$= \underbrace{\left(\frac{\hbar\omega}{kT}\right)^2}_{\uparrow\,(\ln x)^2} k_B \frac{x}{(1-x)^2}$$

$$\Rightarrow \boxed{C_V = k_B\, x\left(\frac{\ln x}{1-x}\right)^2}$$

✓

4b) Rotational modes $\varepsilon_\ell = \frac{\ell(\ell+1)\hbar^2}{2I}$

$$Q = \sum_{\ell=0}^{\infty} (2\ell+1)\, e^{-\varepsilon_\ell/kT}$$

① low temperature $kT \ll \varepsilon_\ell$ we need only take the first terms in the sum:

$$Q = 1 + 3e^{-\frac{\hbar^2}{IkT}}$$

then $$E = -\frac{\partial \ln Q}{\partial \beta} = +\frac{1}{1+3e^{-\beta\hbar^2/I}}\, 3e^{-\beta\hbar^2/I}\, \hbar^2/I$$

$$= \frac{3\hbar^2/I}{e^{\beta\hbar^2/I}+3}$$

$$\Rightarrow \quad C_v = +\frac{1}{kT^2}\,\frac{3\hbar^2}{I}\,\frac{1}{(e^{\beta\hbar^2/I}+3)^2}\cdot\frac{\hbar^2}{I}\,e^{\beta\hbar^2/I}$$

$y = e^{\hbar^2\beta/I}$

$$= 3k_B\,\frac{(\ln y)^2}{(3+y)^2}\,y$$

② high T many modes occupied – reasonable to approximate as continuous spectrum

$$Q = \int_0^\infty d\ell\,(2\ell+1)\,e^{-\frac{\hbar}{2IkT}\ell(\ell+1)}$$

$$\simeq \frac{2IkT}{\hbar}\cdot\int_0^\infty dx\, e^{-x} = \frac{2IkT}{\hbar} = \frac{2I}{\beta\hbar}$$

$$\Rightarrow E = -\frac{\beta\hbar}{2I}\cdot\left(-\frac{2I}{\beta^2\hbar}\right) = kT \quad \Rightarrow \quad C_v = k_B.$$

c) $\frac{\hbar^2}{2I}$ $\quad I \simeq mr^2$ $\quad$ take $m = m_p = 1.67 \times 10^{-27}$ kg

$r = 10^{-10}$ m

$$\Rightarrow I \simeq 10^{-47}$$

$$\Rightarrow \frac{\hbar^2}{2I} \simeq \frac{10^{-68}}{10^{-47}} = \underline{10^{-21}\ \text{J}} \quad \text{or} \simeq 0.01\text{eV}$$

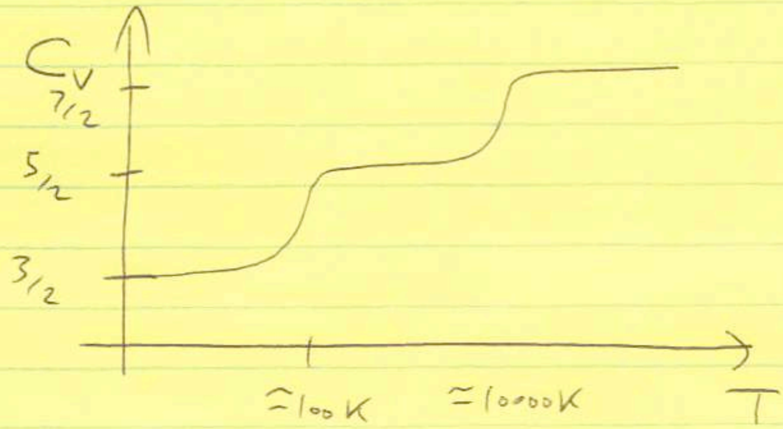
if we define a temperature $T = \frac{\hbar^2}{2Ik_B} = \frac{10^{-21}\text{J}}{10^{-23}} = \underline{100\text{K}}$.

$\hbar\omega$ what is the vibrational frequency? for a very rough estimate, write the energy $m\omega^2 r^2 = \frac{e^2}{4\pi\varepsilon_0 r}$

$$\omega = \left(\frac{e^2}{4\pi\varepsilon_0 r^3 m}\right)^{1/2}$$

$$\simeq 10^{15}\ \text{s}^{-1}$$

$k_B T = \hbar\omega \Rightarrow \underline{T = 10000\text{K}}$ a little high



2. Two-state RNA

(a) The partition function sum $Q = \sum e^{-E_\nu/kT}$

divides into two pieces $Q = \underset{\text{"open states"}}{\sum} e^{-E_\nu/kT} + \underset{\text{"closed states"}}{\sum} e^{-E_\nu/kT}$

$$= Q_{open} + Q_{closed}$$

in terms of free energies $Q = e^{-A_{open}/kT} + e^{-A_{closed}/kT}$

where $A_{open} = -kT \ln Q_{open}$, $A_{closed} = -kT \ln Q_{closed}$

$\Rightarrow$ Probability of being open

$$= \frac{e^{-A_{open}/kT}}{Q}$$

$$= \frac{1}{1 + e^{+\Delta A/kT}} \quad \text{where } \Delta A = A_{open} - A_{closed}$$

(b) Now the molecule is kept at constant force. Use an ensemble analogous to Gibb's ensemble but in 1D.

Whereas $P = \cancel{\tfrac{1}{T}}\, T \frac{\partial S}{\partial V}$ here we have $f = -T \frac{\partial S}{\partial x}$

The probability of being in microstate ν is

$$P_\nu \propto \exp\left(- \frac{(E_\nu - f x_\nu)}{kT} \right)$$

with a corresponding partition function $Z = \sum_\nu e^{-(E_\nu - f x_\nu)/kT}$

and free energy $G = -kT\ln Z$.

Again, divide the microstates into two groups

$\Rightarrow$

$$\boxed{\begin{array}{c} \text{Prob(open)} = \dfrac{1}{1 + e^{\Delta G/kT}} \\ \text{where } \Delta G = \Delta A - f\Delta x \end{array}}$$

Same as before but with ΔG.

(because I can write down a free energy for each group of microstates

$$G_{open} = E_{open} - T S_{open} - f x_{open}$$

$$G_{closed} = E_{closed} - T S_{closed} - f x_{closed}$$

)

4. Heat capacity of a Fermi gas

The energy is $$E = \int \underbrace{\frac{4\pi p^2 dp \times 2}{h^3}} \frac{\varepsilon}{e^{(\varepsilon-\mu)/kT}+1}$$

$p^2 = 2m\varepsilon$ $$\frac{8\pi}{h^3} m\,d\varepsilon\,(2m\varepsilon)^{1/2}$$

$$= \underbrace{\int_0^\infty \frac{\varepsilon^{3/2}\,d\varepsilon}{e^{(\varepsilon-\mu)/kT}-1}} \cdot \frac{8\pi}{h^3}\sqrt{2}\,m^{3/2}$$

Use Sommerfeld expansion on this term

$$\underbrace{\int_0^\mu \varepsilon^{3/2}\,d\varepsilon}_{\frac{2}{5}\mu^{5/2}} + \frac{\pi^2}{6}(kT)^2\,\frac{3}{2}\,\mu^{1/2}$$

We want $c_V = \frac{dE}{dT}$ but we must be careful! μ is the chemical potential at finite temperature, not E_F

From class, $\mu = E_F\left(1 - \frac{\pi^2}{12}\left(\frac{kT}{E_F}\right)^2\right)$

$$\Rightarrow \frac{d\mu}{dT} = -E_F\,\frac{\pi^2}{12}\,\frac{k_B^2}{E_F^2}\,2T$$

$$= -\frac{\pi^2}{6}\frac{k_B^2 T}{E_F}$$

$$\Rightarrow \frac{1}{[illegible]\,m^{3/2}}\frac{h^3}{8\pi}\frac{dE}{dT} = \mu^{3/2}\frac{d\mu}{dT} + \frac{\pi^2}{6}k_B^2 T\,\frac{3}{4}\,\mu^{1/2} + \frac{\pi^2}{6}(kT)^2\,\frac{3}{2}\,\frac{1}{2}\,\mu^{-1/2}\frac{d\mu}{dT}$$

$$= -\mu^{3/2}\,\frac{\pi^2}{6}\frac{k_B^2 T}{E_F} + \frac{\pi^2}{6}k_B^2 T\,\frac{3}{4}\,\mu^{1/2} + O\left(\frac{kT}{E_F}\right)^3$$

$$\Rightarrow C_v = \frac{8\pi}{h^3}\sqrt{2}\, m^{3/2} \cdot \frac{\pi^2}{\cancel{6}3} k_B^2 T E_F^{1/2}$$

~~$\sqrt{2m}E_F^{1/2} = p_F = \hbar(3\pi^2 n)^{1/3}$~~

$$= \frac{\pi^2}{2}\left(\frac{k_B^2 T}{E_F}\right) \cdot \underbrace{\frac{\cancel{8\pi}^{3} E_F^{3/2} m^{3/2}\sqrt{2}}{\hbar^3}\, \frac{2}{3\pi^2}}$$

$$\frac{(2mE_F)^{3/2}}{3\pi^2\hbar^3} = n$$

$$\boxed{C_v = \frac{\pi^2}{2} n k_B \left(\frac{k_B T}{E_F}\right)}$$ ✓

PHYS 362 Problem Set 4

Due Friday 20th February 2009, **in class**.

1. **Non-relativistic and relativistic Fermi gases**

Show that the pressure P of a non-relativistic Fermi gas is related to its internal energy density E by $P = (2/3)E$, for any degree of degeneracy (for any value of μ/k_BT).

Show that a relativistic gas has $P = (1/3)E$ for any degree of degeneracy.

(By non-relativistic gas I mean that the particle energy is $\epsilon = p^2/2m$ whereas for a relativistic gas $\epsilon = pc$).

2. **Thermodynamic properties of a photon gas**

Calculate the entropy, free energy, and pressure of a photon gas, using two methods:

(a) by first calculating the partition function and grand free energy of the gas (remember that a photon gas has $\mu = 0$). The entropy and pressure are derivatives of the free energy.

(b) starting with the internal energy of the gas $E = aT^4V$ which we know from class, calculate C_V, and then $S = \int C_V dT/T$. Once you know E and S, you can write down the free energy and get the pressure by differentiation.

362 Problem Set 4 Solutions

1.) The internal energy is $\int_0^\infty g(p)\,dp\ \varepsilon\ \frac{1}{e^{(\varepsilon-\mu)/kT}+1}$

whereas from class the pressure is

$$\frac{1}{3}\int_0^\infty g(p)\,dp\ pv\ \frac{1}{e^{(\varepsilon-\mu)/kT}+1}$$

Now for a NR gas, $\frac{p^2}{2m}=\varepsilon$

$$\text{and}\quad pv = \frac{p^2}{m} = 2\varepsilon$$

$$\Rightarrow \boxed{P = \frac{2}{3}E}\ ✓$$

For a relativistic gas

$$\left.\begin{aligned} \varepsilon &= pc \\ \text{and}\quad pv &= pc \end{aligned}\right\} \Rightarrow \boxed{P=\frac{1}{3}E}$$

2) (a) Start with $Z = \prod_k \frac{1}{1-e^{-\beta(\varepsilon_k-\mu)}}$ for bosons.

Photons have $\mu=0$.

$$\Rightarrow \quad \Phi = kT\sum_k \ln\left(1-e^{-\beta\varepsilon_k}\right)$$

Replace the sum with an integral over the density of states

$$\Phi = kTV\int \frac{4\pi\omega^2\,d\omega}{(2\pi)^3c^3}\cdot 2\cdot \ln\left(1-e^{-\beta\hbar\omega}\right)$$

$$\Rightarrow \quad \underline{\Phi = kT\,V \int_0^\infty \frac{\omega^2\, d\omega}{\pi^2 c^3} \ln\left(1 - e^{-\beta\hbar\omega}\right).}$$

Rewrite this by defining $x = \frac{\hbar\omega}{kT} = \beta\hbar\omega$

$$\Rightarrow \quad \Phi = \frac{V(kT)^4}{\hbar^3 \pi^2 c^3} \underbrace{\int_0^\infty x^2\, dx\, \ln\left(1 - e^{-x}\right)}_{\text{integrate by parts}}$$

$$\frac{1}{3}\frac{d}{dx}\left[x^3 \ln\left(1 - e^{-x}\right)\right] = x^2 \ln\left(1 - e^{-x}\right) + \frac{1}{3} x^3 \frac{1}{1 - e^{-x}} e^{-x}$$

$$= \frac{-V(kT)^4}{\hbar^3 \pi^2 c^3} \frac{1}{3} \int_0^\infty \frac{x^3}{e^x - 1}\, dx$$

or $\boxed{\Phi = -\frac{1}{3} a T^4 V}$ using the definition of a from class

Now everything else follows: $S = -\frac{\partial \Phi}{\partial T}\Big|_V = \underline{\frac{4}{3} a T^3 V}$ ✓

$$p = -\frac{\partial \Phi}{\partial V}\Big|_T = \underline{\frac{1}{3} a T^4} \quad ✓$$

(b) $E = aT^4 V$

$\Rightarrow \quad C_V = 4aT^3 V$

$\Rightarrow \quad S = \int \frac{C_V\, dT}{T} = \underline{\frac{4}{3} a T^3 V}$

$A = E - TS = -\frac{1}{3} a T^4 V$

$\Rightarrow \quad p = -\frac{\partial A}{\partial V}\Big|_T$

$= \underline{\frac{1}{3} a T^4}$

PHYS 362 Problem Set 5

Due on Friday March 20th 2009, either in class or by 5pm in ERP 219.

1. **Thermal properties of the mean field Ising model**

(a) Starting by writing down the partition function for the Ising model in mean field theory, show that the energy, heat capacity, and entropy are

$$E = -\frac{1}{2} N k_B T_c \langle s \rangle^2$$

$$\frac{C_V}{N k_B} = \left(\frac{T_c}{T}\right)^2 \langle s \rangle^2 \left[\frac{1}{1-\langle s \rangle^2} - \left(\frac{T_c}{T}\right)\right]^{-1}$$

$$\frac{S}{N k_B} = \ln\left(\frac{2}{\sqrt{1-\langle s \rangle^2}}\right) - \frac{T_c}{T}\langle s \rangle^2 .$$

(b) Show that as temperature approaches the critical point from below,

$$\langle s \rangle \approx \sqrt{3\left(1-\frac{T}{T_c}\right)} \qquad C_V \rightarrow \frac{3}{2} N k_B \qquad S \rightarrow N k_B \ln 2.$$

(c) Now show that as $T \rightarrow 0$ both C_V and S go to zero.

(d) What are E, C_V, and S for $T > T_c$? Give a physical interpretation for this value of S.

(e) Sketch E, C_V, and S as a function of temperature. Is this a first order or continuous phase transition? How does the exact 2D or 3D solution differ from these mean field results?

2. **More on the DNA zipper**

This question is an extension of the zipper model for DNA that we looked at in homework 2. A DNA molecule consists of two polymer chains joined by N links. The link at one end of the molecule is held closed, but the molecule can "unzip" from the other end, with an energy cost ϵ for each link that opens. The difference from the previous homework problem is that we now include the fact that each of the open links can rotate around the axis of the polymer chain, with g different orientations available per open link. Therefore there are g^n different states of the molecule with n open links.

(a) Show that the energy and heat capacity are

$$E = \langle n \rangle \epsilon = \epsilon \left[\frac{x}{1-x} - \frac{N x^N}{1-x^N} \right]$$

$$C_V = k_B \left(\frac{\epsilon}{k_B T} \right)^2 \left[\frac{x}{(1-x)^2} - \frac{N^2 x^N}{(1-x^N)^2} \right]$$

where $x = g \exp(-\beta \epsilon)$ and $\langle n \rangle$ is the mean number of open links.

(b) Assuming $N \gg 1$, show that at the critical temperature $T_c = \epsilon / k_B \ln g$ (where $x = 1$ and the denominators in the expressions for E and C_V diverge),

$$\langle n \rangle = N/2$$

and

$$\frac{C_V}{N k_B} = \frac{N (\ln g)^2}{12}.$$

(c) Sketch $\langle n \rangle$ and C_V as a function of x and indicate how the curves change as N increases.

(d) For large N, use the results you've derived so far to argue that there is a first order phase transition between the zipped and unzipped states of the molecule. Explain physically why the critical temperature T_c at which this transition occurs is smaller for larger values of g.

(e) In the laboratory, DNA is observed to unzip at temperatures above 90°C. If the energy required to open a link is $\epsilon \approx 0.3$eV, estimate the value of g.

Useful results for geometric series:

$$\Sigma_{n=0}^{N-1} x^n = \frac{1-x^N}{1-x} \qquad \Sigma_{n=0}^{N-1} n x^n = \left(\frac{1-x^N}{1-x} \right) \left[\frac{x}{1-x} - \frac{N x^N}{1-x^N} \right]$$

PHYS 362 Problem Set 5 Solutions

1. Mean field theory

(a) The energy of a single spin is $-Jz\langle s_i\rangle^2$

$$= -k_B T_c \langle s_i\rangle^2$$

For N spins, including a factor of ½ to avoid overcounting

$$\boxed{E = -\frac{1}{2} N k_B T_c \langle s_i\rangle^2}$$

Then $$C_V = \frac{\partial E}{\partial T} = -\frac{1}{k_B T^2}\frac{\partial E}{\partial \beta} = \underline{N k_B \frac{T_c}{T}\langle s_i\rangle \, \beta \frac{\partial\langle s_i\rangle}{\partial \beta}} \quad (*)$$

Use $\langle s_i\rangle = \tanh\left(\frac{T_c}{T}\langle s_i\rangle\right)$

$$\Rightarrow \frac{\partial\langle s_i\rangle}{\partial\beta} = \operatorname{sech}^2\left(\frac{T_c}{T}\langle s_i\rangle\right)\left[k_B T_c \langle s_i\rangle + \frac{T_c}{T}\frac{\partial\langle s_i\rangle}{\partial\beta}\right]$$

$$\Rightarrow \beta\frac{\partial\langle s_i\rangle}{\partial\beta} = \frac{\frac{T_c}{T}\langle s_i\rangle}{-\frac{T_c}{T} + \cosh^2\left(\frac{T_c}{T}\langle s_i\rangle\right)} = \frac{\frac{T_c}{T}\langle s_i\rangle}{\frac{1}{1-\langle s_i\rangle^2} - \frac{T_c}{T}}$$

$$\left(\text{Since } \cosh^2 x = \frac{1}{\operatorname{sech}^2 x} = \frac{1}{1-\tanh^2 x}\right)$$

$$\Rightarrow \boxed{\frac{C_V}{N k_B} = \frac{\left(\frac{T_c}{T}\right)^2\langle s_i\rangle^2}{\frac{1}{1-\langle s_i\rangle^2} - \frac{T_c}{T}}}$$

The entropy is $$\frac{S}{N k_B} = \int_0^T \frac{C_V\, dT}{N k_B T} \underset{\text{use }(*)}{=} -\int \frac{\cancel{dT}}{T}\frac{T_c}{\cancel{T}}\langle s_i\rangle \cancel{T}\frac{\partial\langle s_i\rangle}{\cancel{\partial T}}$$

$$= -\int_1^{\langle s_i\rangle}\frac{T_c}{T}\langle s_i'\rangle\, d\langle s_i'\rangle$$

Solution for
1. (a) using the partition function to get E and S.

In the mean field theory approximation, each spin makes an independent contribution to the energy

$$E_i = -Jz\langle s\rangle s_i + \frac{Jz}{2}\langle s\rangle^2$$

⇒ the partition function factorizes $\quad Q = (Q_1)^N$

$$Q_1 = e^{-\beta Jz\langle s\rangle^2/2}\; 2\cosh(\beta Jz\langle s\rangle)$$

Then
$$\begin{aligned} E &= -\frac{\partial \ln Q}{\partial\beta} \\ &= -N\frac{\partial}{\partial\beta}\left[-\frac{\beta Jz\langle s^2\rangle}{2} + \ln 2\cosh(\beta Jz\langle s\rangle)\right] \\ &= \frac{NJz\langle s^2\rangle}{2} - N\underbrace{\tanh(\beta Jz\langle s\rangle)}_{\langle s\rangle}\, Jz\langle s\rangle \\ &= \underline{-\frac{1}{2}NJz\langle s\rangle^2} \quad ✓ \end{aligned}$$

The free energy is
$$\begin{aligned} A &= -kT\ln Q \\ &= \frac{Jz\langle s\rangle^2 N}{2} - Nk_BT\ln 2\cosh(\beta Jz\langle s\rangle) \end{aligned}$$

$$\Rightarrow \quad S = \frac{-A+E}{T}$$

$$\Rightarrow \quad \frac{S}{Nk_B} = \underbrace{-\frac{Jz\langle s\rangle^2}{2k_BT}}_{-\frac{T_c}{T}\frac{\langle s\rangle^2}{2}} + \ln\underbrace{2\cosh(\beta Jz\langle s\rangle)}_{\frac{2}{\sqrt{1-\tanh^2()}} = \frac{2}{\sqrt{1-\langle s\rangle^2}}} \overbrace{-\frac{1}{2}\frac{Jz\langle s\rangle^2}{k_BT}}^{E}$$

$$= -\frac{T_c}{T}\langle s\rangle^2 + \ln\left(\frac{2}{\sqrt{1-\langle s\rangle^2}}\right) \quad ✓$$

but $\frac{T_c}{T}\langle s_i\rangle = \tanh^{-1}\langle s_i\rangle$

$$\Rightarrow \quad \frac{S}{Nk_B} = \int_{\langle s_i\rangle}^{1} d\langle s_i\rangle \tanh^{-1}\langle s_i\rangle$$

(use Mathematica) $= \left[\log 2 - \frac{1}{2}\log(1-\langle s_i\rangle^2) - \langle s_i\rangle\tanh^{-1}\langle s_i\rangle\right]$

$$\Rightarrow \quad \boxed{S = Nk_B\left[\log\left(\frac{2}{\sqrt{1-\langle s_i\rangle^2}}\right) - \langle s_i\rangle^2\frac{T_c}{T}\right]}$$

(b) for $T \simeq T_c$, $\langle s_i\rangle \ll 1$ and so we can expand

$$\langle s_i\rangle = \tanh\left(\langle s_i\rangle\frac{T_c}{T}\right) \simeq \langle s_i\rangle\frac{T_c}{T} - \frac{\langle s_i\rangle^3}{3}\left(\frac{T_c}{T}\right)^3 \cdots$$

$$\Rightarrow \quad \boxed{\langle s_i\rangle = \left(\frac{T}{T_c}\right)\sqrt{3\left(1-\frac{T}{T_c}\right)}} \quad ✓$$

Write $x = 1-\frac{T}{T_c}$ or $\frac{T}{T_c} = 1-x$

$$\Rightarrow \quad \frac{C_v}{Nk_B} = \frac{3x}{\frac{1}{1-(1-x)^2\,3x} - \frac{1}{1-x}} \approx \frac{3x(1-x)}{\underbrace{(1-x)(1+3x)-1}_{2x}}$$

$$\simeq \frac{3}{2} + O(x)$$

$$\Rightarrow \quad \boxed{C_v \rightarrow \frac{3}{2}Nk_B} \text{ at } T = T_c^- \quad ✓$$

As $\langle s_i\rangle \rightarrow 0$ $\boxed{S = Nk_B\log 2}$ ✓

(c) We need to work out how $\langle s_i\rangle$ behaves as $T \rightarrow 0$.

$$\underbrace{\tanh^{-1}\langle s_i\rangle}_{} = s_i \frac{T_c}{T}$$

write as $\tanh^{-1}(1-x)$ where $x \to 0$ as $T \to 0$

$$\| $$

$$\frac{1}{2}\ln\left(\frac{2-x}{x}\right) \simeq \frac{1}{2}\left[\ln 2 - \frac{x}{2} - \ln x\right]$$

$$\Rightarrow \quad 2(1-x)\frac{T_c}{T} \simeq \ln\left(\frac{2}{x}\right) - \frac{x}{2}$$

as T and x go to zero, we must have $\ln\left(\frac{2}{x}\right) = 2\frac{T_c}{T}$

$$\Rightarrow \boxed{x = 2e^{-2T_c/T}}$$

$$\text{or} \quad \boxed{\langle s_i\rangle = 1 - 2e^{-2T_c/T}}.$$

Now
$$\frac{C_V}{Nk_B} = \frac{\left(\frac{T_c}{T}\right)^2 (1-x)^2}{\frac{1}{\underbrace{1-(1-x)^2}_{\simeq 2x}} - \frac{T_c}{T}} \simeq \frac{(1-x)^2\left(\frac{T_c}{T}\right)^2}{\frac{1}{2x} - \frac{T_c}{T}}$$

$$\simeq 2x\left(\frac{T_c}{T}\right)^2$$

$$\boxed{\frac{C_V}{Nk_B} \simeq 4e^{-2T_c/T}\left(\frac{T_c}{T}\right)^2} \quad \text{as } T \to 0$$

$$\to 0 \text{ as } T \to 0. \checkmark$$

And the entropy as $T \to 0$

$$\frac{S}{Nk_B} \to \underbrace{\ln\left(\frac{2}{\sqrt{2x}}\right)}_{\frac{T_c}{T}} - \cancel{(1-2x)^2}\langle s_i\rangle^2 \frac{T_c}{T}$$

$$= \frac{T_c}{T}\left(1 - \langle s_i\rangle^2\right) = (2x)\left(\frac{T_c}{T}\right) \to \underline{\underline{0}} \checkmark$$

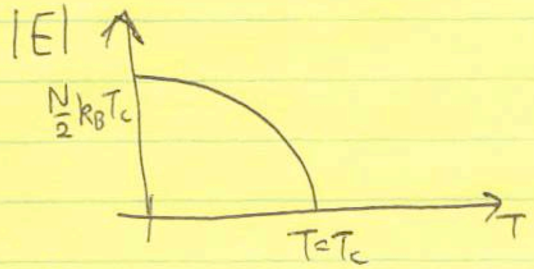
(d) For $T > T_c$, $\langle s_i \rangle = 0$, giving $E = 0$

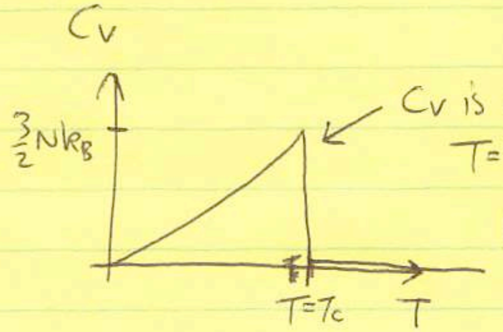
$C_V = 0$

$$\frac{S}{k_B} = N \ln 2 = \ln(2^N)$$

N spins, each can be up or down

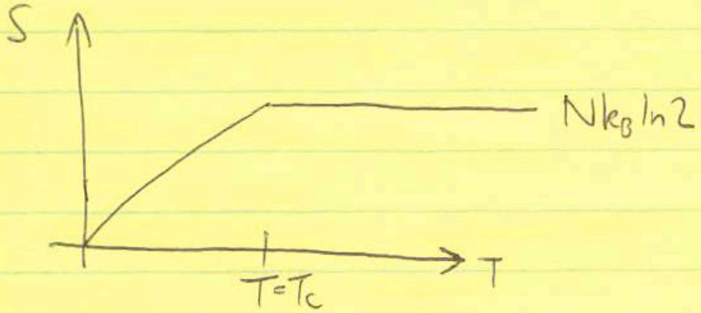
(e) Plots:





C_V is discontinuous at $T = T_c$

($\Rightarrow$ continuous phase transition)



2. DNA zipper

(a)
$$\frac{E}{\varepsilon} = \langle n \rangle = \frac{\sum_{n=0}^{N-1} n x^n}{\sum_{n=0}^{N-1} x^n} = \underline{\underline{\frac{x}{1-x} - \frac{N x^N}{1-x^N}}}$$

where $x = g e^{-\beta\varepsilon}$

$$C_v = \frac{\partial E}{\partial T} = -\frac{1}{kT^2}\frac{\partial E}{\partial \beta} = k_B\left(\frac{\varepsilon}{k_B T}\right)^2 x \frac{\partial \langle n \rangle}{\partial x}$$

$$= \underline{\underline{k_B\left(\frac{\varepsilon}{k_B T}\right)^2\left[\frac{x}{(1-x)^2} - \frac{N^2 x^N}{(1-x^N)^2}\right]}}$$

(b) Analyse the behavior around $x=1$. Let $\eta = x-1$ $x = 1+\eta$ $\eta \ll 1$

$$\langle n \rangle = -\frac{1+\eta}{\eta} - \frac{N(1+\eta)^N}{1-(1+\eta)^N}$$

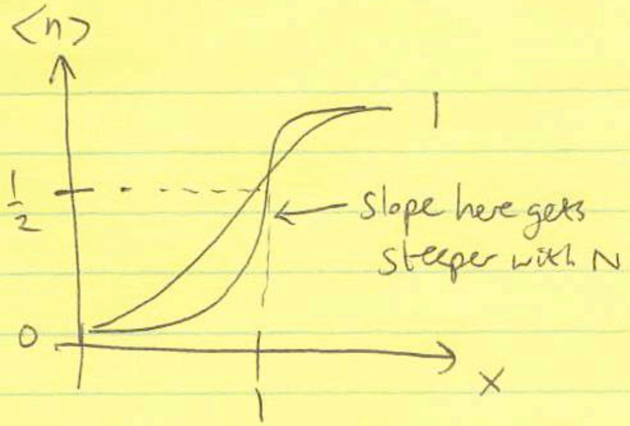
expanding gives $\langle n \rangle = \frac{N-1}{2} + \left(\frac{N^2-1}{12}\right)(x-1) \cdots$

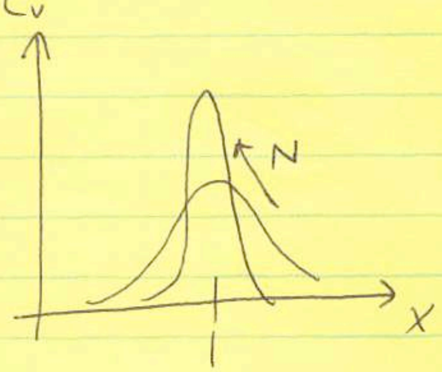
$$\Rightarrow \boxed{\langle n \rangle \simeq \frac{N}{2} \quad \text{at } x=1}$$

and $\frac{d\langle n \rangle}{dx} = \frac{N^2}{12}$ giving $\boxed{\frac{C_v}{N k_B} = \frac{N(\ln g)^2}{12}}$

at $x=1$.

(c)





(d) As $N \to \infty$ $\quad \frac{C_V}{N k_B} \to \infty$ or in other words the entropy is discontinuous at $T = T_c$ ⇒ 1st order transition (first derivative of free energy has a discontinuity.)

For larger values of g, the entropy of the unzipped molecule is greater, allowing phase transition at lower temperature $(A = E - TS)$

[Zipped molecule has $E = S = A = 0$
unzipped has $E = N\varepsilon$, $S = k_B \cancel{g} N \ln g$

$$\Rightarrow A = 0 \text{ when } T = \frac{E}{S} = \frac{\varepsilon}{k_B \ln g} \checkmark$$
]

(e) $T_c = 90°C \simeq 350\,K$

$$\Rightarrow \ln g \simeq \frac{\varepsilon}{k_B T_c} \simeq \frac{0.3 \times 10^{-19}}{10^{-23} \times 300} \simeq 10$$

$$\Rightarrow g \simeq e^{10} = 10^{10/2.303} \simeq \underline{\underline{10^4}}.$$

PHYS 362 Problem Set 6

Due on Wednesday April 1st 2009, either in class or by 5pm in ERP 219.

Monte Carlo simulation of the 2D Ising model

Implement the Metropolis algorithm discussed in class for a 5×5 spin Ising model. Plot the magnetization M, susceptibility χ, energy E, and heat capacity C_V as a function of T.

[You should hand in a brief description of how you implemented the algorithm, your code, and the plots.]

PHYS 362 Problem Set 7

Due on Tuesday 14th April by 5pm in ERP 219.

1. **Lattice gas.**

Calculate the entropy of the lattice gas in mean field theory by differentiating the free energy and by writing down the Gibbs entropy, and check that the two methods give the same answer.

2. **Mean field Ising model at non-zero applied magnetic field**

For the mean-field Ising model, plot M against T for different values of H and M against H for different values of T.

3. **Reaction equilibria**

(a) Consider the process

$$A \rightleftharpoons B$$

where A and B refer to different isomer states of a molecule. Assume the process takes place in a dilute gas and the energy difference between state A and B is $\Delta\epsilon$. In thermal equilibrium, calculate the expected ratio of number densities n_A and n_B using two methods: from the Boltzmann distribution, and by setting $\mu_A = \mu_B$. Make sure that they give the same answer.

(b) Derive the equilibrium number density ratio of reactants and products for the three-body reaction

$$A + B \rightleftharpoons C.$$

Discuss physically the factors that determine whether $A + B$ or C dominate in equilibrium.

4. **Adsorption**.

This question is from last year's final.

(a) A metal surface is exposed to a gas of helium atoms at temperature T and pressure P. The metal surface has binding sites for helium atoms. The energy required to remove a bound helium atom from the surface is ϵ. In equilibrium, what fraction of binding sites are occupied? Write your answer in terms of T and P. What are the high and low temperature limits of your expression? [Treat the gas as a heat and particle bath. You may use the result that the chemical potential of an ideal gas is

$$\mu = k_B T \ln\left(\frac{n}{n_Q}\right)$$

where $n_Q = (m k_B T/2\pi\hbar^2)^{3/2}$.]

(b) Repeat the calculation for the case where the bound helium atoms can move freely on the surface, so that they form a two-dimensional gas of non-interacting particles. Discuss how your answer differs from part (a).

PHYS 362 Problem Set 7 Solutions

1. Lattice gas

In class we found

$$\Phi = 3N_{\text{cells}}\,\varepsilon\,\bar{n}^2 - k_BT\,N_{\text{cells}}\ln\left(1+e^{6\beta\bar{n}\varepsilon+\mu\beta}\right)$$

The entropy is $S = -\left.\frac{\partial\Phi}{\partial T}\right|_{\mu,V}$

$$= -\left[\cancel{\frac{6\varepsilon N}{N_{\text{cells}}}\frac{\partial N}{\partial T}} - k_BN_{\text{cells}}\ln\left(1+e^{6\beta\bar{n}\varepsilon+\mu\beta}\right) - k_BT\underbrace{\boxed{N_{\text{cells}}\frac{e^x}{1+e^x}}}_{N}\left(-\frac{(6\bar{n}\varepsilon+\mu)}{k_BT^2} + \cancel{\frac{6\varepsilon}{k_BT}\frac{1}{N_{\text{cells}}}\frac{\partial N}{\partial T}}\right)\right]$$

Cancels first term

$$= k_BN_{\text{cells}}\ln\underbrace{\left(1+e^x\right)}_{1+\frac{\bar{n}}{1-\bar{n}}\,=\,\frac{1}{1-\bar{n}}}\ \underset{-}{\cancel{+}}\ \frac{N}{T}\underbrace{\left(6\bar{n}\varepsilon+\mu\right)}_{\text{from class this is } k_BT\ln\left(\frac{\bar{n}}{1-\bar{n}}\right)}$$

$$= -k_BN_{\text{cells}}\ln\left(1-\bar{n}\right) - N\,k_B\ln\left(\frac{\bar{n}}{1-\bar{n}}\right)$$

$$\frac{S}{k_BN_{\text{cells}}} = -(1-\bar{n})\ln(1-\bar{n}) - \bar{n}\ln\bar{n}$$

which is the Gibbs entropy $S = -k_B\sum p_i\ln p_i$

(Not asked for in the Q, but for information:)
One additional check we can do is to calculate

$$E = \Phi + TS + \mu N$$

$$= 3N_{\text{cells}}\varepsilon\bar{n}^2 \underset{+}{\cancel{-}} k_BT N_{\text{cells}} \ln(\overset{1-\bar{n}}{\cancel{1+e^x}}) \cancel{+} - k_BT N_{\text{cells}}\begin{pmatrix}\bar{n}\ln\bar{n} \\ +(1-\bar{n})\ln(1-\bar{n})\end{pmatrix}$$

$$+ N_{\text{cells}}\bar{n}\left(-6\bar{n}\varepsilon + k_BT\ln\left(\frac{\bar{n}}{1-\bar{n}}\right)\right)$$

$$= N_{\text{cells}}\bar{n}^2\varepsilon(3-6)$$

$$= -3N_{\text{cells}}\bar{n}^2\varepsilon \quad \text{as expected ✓}$$

$$= -3N_{g}\bar{n}\varepsilon$$

2. MFT at non-zero applied field

See attached plots.

3. Reaction equilibria

(a) $A \rightleftharpoons B$

In thermal eqm, we expect $\frac{n_A}{n_B} = e^{-\frac{\Delta\varepsilon}{k_BT}}$ from the Boltzmann distribution, where A is the more energetic state.

Now use chemical potentials: in eqm

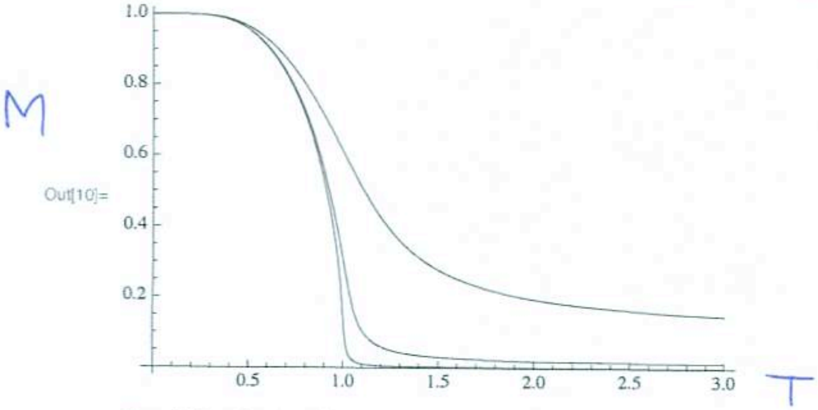
$$\mu_A = \mu_B \Rightarrow \quad k_BT\ln\left(\frac{n_A}{n_{Q,A}}\right) = k_BT\ln\left(\frac{n_B}{n_{Q,B}}\right) - \Delta\varepsilon$$

2a

```
In[5]:= f[T_, H_] := FindRoot[x == Tanh[H + x / T], {x, Sign[H]}]
```

```
In[10]:= Show[Plot[x /. f[y, 0.001], {y, 0, 3}],
  Plot[x /. f[y, 0.01], {y, 0, 3}], Plot[x /. f[y, 0.1], {y, 0, 3}]]
```

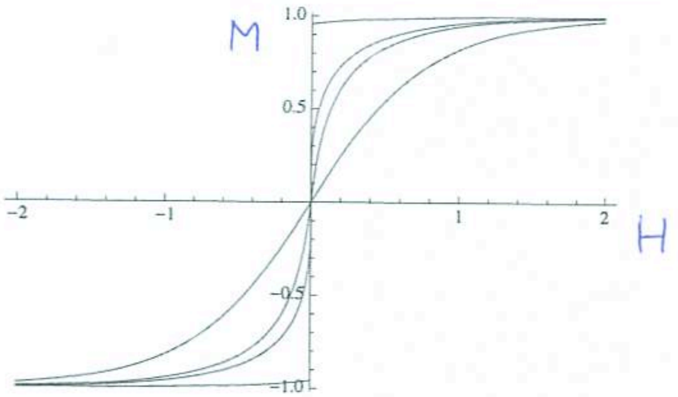
Out[10]=



$H = 10^{-3}, 10^{-2}, 10^{-1}$

```
In[18]:= Show[Plot[x /. f[5.0, y], {y, -2, 2}], Plot[x /. f[1.2, y], {y, -2, 2}],
  Plot[x /. f[1.0, y], {y, -2, 2}], Plot[x /. f[0.5, y], {y, -2, 2}]]
```

Out[18]=



$\frac{T}{T_c} = 0{\cdot}5, 1{\cdot}0, 1{\cdot}2, 5{\cdot}0$

An isomer is a different structural form of the same molecule so that the particle masses are the same ⇒ $n_{Q,A} = n_{Q,B}$

Therefore $\frac{n_A}{n_B} = e^{-\Delta\varepsilon/k_BT}$ ✓

$$n_Q = \left(\frac{2\pi mkT}{h^2}\right)^{3/2}$$

(b) $A + B \rightleftharpoons C$

the molecule C is energetically preferred

$$\mu_A + \mu_B = \mu_C$$

$$k_BT\ln\left(\frac{n_A}{n_{Q,A}}\right) + k_BT\ln\left(\frac{n_B}{n_{Q,B}}\right) = k_BT\ln\left(\frac{n_C}{n_{Q,C}}\right) - \Delta\varepsilon$$

$$\Rightarrow \quad \frac{n_A n_B}{n_C} = \left(\frac{n_{A,Q}\, n_{B,Q}}{n_{C,Q}}\right) e^{-\Delta\varepsilon/k_BT}$$

$$= \left(\frac{m_A m_B}{m_C}\right)^{3/2} e^{-\Delta\varepsilon/k_BT} \left(\frac{2\pi k_BT}{h^2}\right)^{3/2}$$

---

part(a)

The point I'm trying to get across here is that in the first case the position of the equilibrium is set by k_BT vs. $\Delta\varepsilon$. At low temperatures, the lower energy state B dominates ($n_A \ll n_B$) whereas at high temperatures $k_BT > \Delta\varepsilon$ then the energy difference no longer matters and $n_A \simeq n_B$.

In the second case – part(b) – this is no longer the only factor. There is a prefactor $\left(\frac{2\pi mk_BT}{h^2}\right)^{3/2} = n_Q$ which takes into account the fact that even though C is energetically preferred, there is an entropy increase when C breaks up into two molecules A and B – So the transition from C to (A+B) takes place at a lower temperature than $k_B T = \Delta \varepsilon$.

4. Adsorption

See solutions for last year's final exam for this one.