

CHAPTER 8: INTERACTING SYSTEMS

8a: General considerations

So far we've treated only systems of particles with no "forces" between them - i.e. the Hamiltonian was the sum of single-particle Hamiltonians, with no interactions:

$$H_{\text{total}} = H(1) + H(2) + \dots + H(N) \quad (8a.1)$$

Treating cases when the Hamiltonian does not split up this way is very difficult in general (the "many-body problem"). We'll briefly discuss four broad classes of attack: 1) perturbation theory; 2) normal modes; 3) Other exact solutions for special cases; 4) mean field theory. Other methods not discussed here include Monte-Carlo methods, the Renormalization Group, high-temperature series, matrix product states, and brute-force numerical simulation.

8b: Perturbation theory.

Recall that all macroscopic properties can be calculated from the partition function $Z_N = \sum_i e^{-\beta E_i(N)}$ or the grand partition function $Q = \sum_N Z_N e^{\beta \mu N}$.

When an inter-particle interaction makes it impossible to find the energy levels exactly we can try treating the interaction calculating the perturbed energies $E_i(N)$ in terms of the unperturbed energies of the $E_{i,0}(N)$ of the noninteracting system: thus from 1st order perturbation theory

$$E_i(N) \approx E_{i,0}(N) + \langle i(0) | \delta H | i(0) \rangle + \dots$$

where $\delta H = H - H_0$ is the interacting term in the Hamiltonian. Since the state $|i(0)\rangle$ has to be a correctly symmetrized many-body product wavefunction, even the first-order calculation of Z or Q this way is messy for quantum systems. The preferred approach uses quantized *fields*, plus Feynman-type diagrams for higher order corrections. This is material for an entire course on its own.

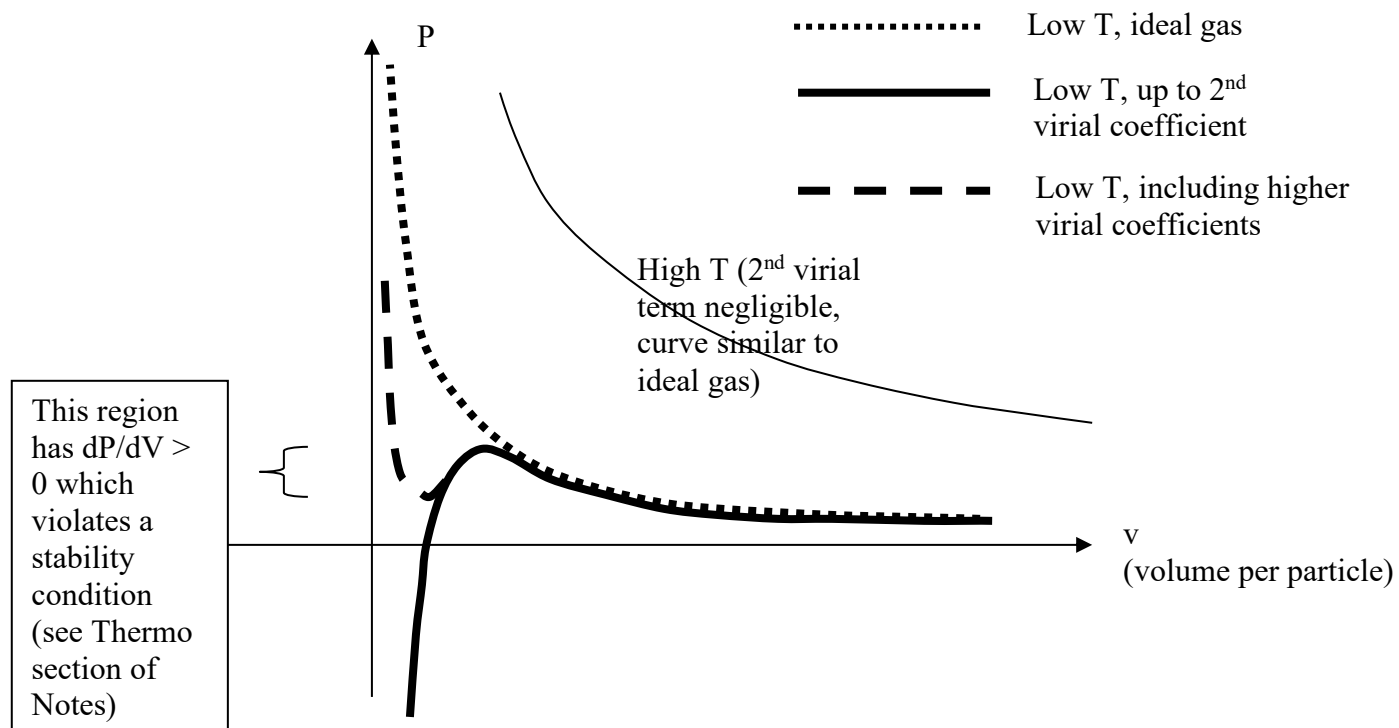
An easier calculation is for a gas in the classical (Maxwell-Boltzmann) regime of temperature. We will not perform that calculation in this course, but if you are interested, refer to the text FSTP by Reif. This method allows one to calculate so-called *virial coefficients* (first defined by Clausius in the 19th century) in terms of the inter-particle potential $U(r)$, assumed to depend only on the distance r between the particles. The first non-trivial virial coefficient was found by Ornstein in his 1908 PhD thesis:

$$B_2(T) = \frac{1}{2} \int_0^\infty 4\pi r^2 [1 - \exp(-\beta U(r))] dr \quad (8b.4)$$

This gives the first-order correction to the ideal gas law, describing a *nonideal* gas, in terms of $v = V/N$:

$$P \approx kT \left(v^{-1} + B_2(T) v^{-2} + \dots \right) \quad (8b.5)$$

This is an expansion in $1/v = N/V$, and so is really a *low-density* expansion. Because the shape of $U(r)$ can be quite complicated (repulsive at short range; attractive at long range), the phases of a non-ideal gas can be complicated. Of course we know this is true because when $1/v$ is larger enough, gases condense into liquids, which don't behave like gases at all. But even before then, complicated phase transitions can occur:



What actually happens in the region predicted to have $dP/dV > 0$ is that **phase separation** occurs. The fluid separates into high-density liquid regions and low-density gas regions. In each of these regions we have $dP/dV < 0$ as required for stability.

8c Normal coordinates (2nd method of dealing with interacting systems)

(see also 3rd year condensed matter Notes)

Consider any array of N distinguishable particles, which are tightly bound to definite sites. Examples of such systems range from the two nuclei in diatomic molecules (the internuclear separation R almost fixed via the Coulomb interaction between the electrons and the nuclei), to ordered solids (crystals) and disordered solids (glasses). Let x_α represent one Cartesian component of the displacement of one "atom" from its site. There are $3N$ such coordinates. We can assign each a mass, M_α , with $M_{3n+1} = M_{3n+2} = M_{3n+3}$ for integer n . Thus the total energy (not really a Hamiltonian yet, but I can't edit the later equations)

$$H = \frac{1}{2} \sum_{\alpha=1}^{3N} M_\alpha \dot{x}_\alpha^2 + V(x_1, \dots, x_{3N}).$$

We can expand the total energy $T+V$ in powers of the small displacements x_α . There is no linear term provided we expand around the mechanical equilibrium configuration, so:

$$H = \frac{1}{2} \sum_{\alpha=1}^{3N} M_\alpha \dot{x}_\alpha^2 + \frac{1}{2} \sum_{\alpha,\beta=1}^{3N} \left. \frac{\partial^2 V}{\partial x_\alpha \partial x_\beta} \right|_{x=0} x_\alpha x_\beta + \text{const} \quad (8c.1)$$

Understanding check: why is there a 1/2 in front of the second sum?

Letting $u_\alpha = \sqrt{M_\alpha} x_\alpha$ we can write this as

$$H = \frac{1}{2} \sum_{\alpha=1}^{3N} \dot{u}_\alpha^2 + \frac{1}{2} \sum_{\alpha,\beta=1}^{3N} u_\alpha G_{\alpha\beta} u_\beta + \text{const} \quad (8c.2)$$

where $G_{\alpha\beta} = \frac{1}{\sqrt{M_\alpha M_\beta}} \left. \frac{\partial^2 V}{\partial x_\alpha \partial x_\beta} \right|_{x=0} = G_{\beta\alpha}$

We can regard the $\{u_\alpha\}$ as components of a $3N$ -dimensional column vector \underline{u} and the $\{G_{\alpha\beta}\}$ as elements of a $3N \times 3N$ real symmetric matrix \underline{G} . Then, removing the constant term,

$$H = \frac{1}{2} \underline{\dot{u}}^T \underline{\dot{u}} + \frac{1}{2} \underline{u}^T \underline{G} \underline{u}. \quad (8c.3)$$

Being real symmetric, G can be diagonalized by an orthogonal matrix:

$$\underline{G} = \underline{P} \underline{\Gamma} \underline{P}^T \quad \text{where} \quad \underline{P} \underline{P}^T = \underline{I}, \quad \text{i.e.} \quad \sum_q P_{\alpha q} P_{\alpha' q} = \delta_{\alpha\alpha'} \quad (8c.4)$$

and also $\det P = \pm 1$ [since $1 = \det I = \det P P^T = (\det P)^2$]. Choose $\det P = +1$ (it is then a "special orthogonal matrix"). Then $\underline{\Gamma}$ is diagonal and its elements are the eigenvalues of G , which we define to be the squares of the normal mode frequencies ω_q , $q=1,2,\dots,3N$:

$$\Gamma_{\approx} = \begin{pmatrix} \omega_1^2 & & & 0 \\ & \omega_2^2 & & \\ & & \ddots & \\ 0 & & & \omega_{3N}^2 \end{pmatrix} \text{ normal mode frequencies} \quad (8c.5)$$

Exercise: check that this is dimensionally correct.

Note that a negative ω^2 indicates mechanical instability, violating our starting assumption.

Thus

$$\begin{aligned} H &= \frac{1}{2} \dot{u}^T \left(P P^T \right) \dot{u} + \frac{1}{2} u^T \left(P \Gamma P^T \right) u \\ &= \frac{1}{2} \dot{Q}^T \dot{Q} + \frac{1}{2} Q^T \Gamma Q \\ &= \sum_q \left(\frac{1}{2} \dot{Q}_q^2 + \frac{1}{2} \Gamma_q Q_q^2 \right) = \sum_q \left(\frac{1}{2} \dot{Q}_q^2 + \frac{1}{2} \omega_q^2 Q_q^2 \right) \end{aligned} \quad (8c.6)$$

where we have introduced normal mode coordinates Q_q ,

$$Q_q = \sum_{\alpha} P_{\alpha q} U_{\alpha}, \text{ ie } Q_{\approx} = P_{\approx}^T u, \quad q = 1, 2, \dots, 3N \quad (8c.7)$$

These are linearly independent combinations of the original displacements $\{x_{\alpha}\}$. This is guaranteed since any real symmetric $D \times D$ matrix G has D linearly independent eigenvectors.

Thus the model reduces to $3N$ uncoupled harmonic oscillators of mass 1 and frequencies ω_q .

Classical limit (ok for high T):

By the equipartition theorem, at thermal equilibrium (canonical ensemble) each harmonic oscillator holds kT of energy (half as kinetic energy and half as potential energy), so the total energy is $E = 3NkT$.

Quantum case (needed for low T).

The quantum case is also straight-forward – it corresponds to the “bosons without number conservation”, section 6n. The q th harmonic oscillator has energy levels $(n + \frac{1}{2})\hbar\omega_q$, and has a mean **phonon** occupation of

$$\bar{n}_q = \frac{1}{e^{\beta\hbar\omega_q} - 1}.$$

Note: the coordinates Q_q are linear combinations of the original x_{α} coordinates, so an energy eigenstate $|\{n_q\}\rangle$, written as a wavefunction of the original $3N$ position coordinates, is a big, complicated thing. That’s why it is easier just to think in terms of phonon excitation.

Debye model of lattice specific heat (low T).

Let the vector $\underline{u}(\underline{R})$ represent the 3 Cartesian components of displacement of "atom" at lattice site \underline{R} . Then in previous notation $x_\alpha = u_\mu(\underline{R})$ i.e. $\alpha = n(\underline{R}) + \mu$, where

μ = the particular cartesian component of displacement (1, 2, or 3) and
 \underline{R} = the unperturbed position of lattice site, defining the n at the start of 8c.

It can be shown that in general, the normal coordinates are sinusoidal combinations of the $\{\underline{u}(\underline{R})\}$ (i.e. they are lattice waves). In terms of previous notation, we "expand" the index q :

$$Q_q = Q_{q,m,j} \text{ so } q = (\underline{q}, m, j):$$

$m = s$ or c labels sin or cos mode; $j = 1$ or 2 or 3 labels polarization

so that we can write

$$\begin{aligned} Q_{q,c,j} &= \sqrt{\frac{2}{N}} \sum_{\underline{R}} \cos(\underline{q} \cdot \underline{R}) \underline{e}(\underline{q}, c, j) \cdot \underline{u}(\underline{R}) \\ Q_{q,s,j} &= \sqrt{\frac{2}{N}} \sum_{\underline{R}} \sin(\underline{q} \cdot \underline{R}) \underline{e}(\underline{q}, s, j) \cdot \underline{u}(\underline{R}) \end{aligned} \quad (8c.14)$$

Here \underline{e} are the polarization vectors (e.g. $j=1$ might be a longitudinal phonon with $\underline{e} \parallel \underline{q}$, while $j=2,3$ might be the 2 transverse polarizations, $\underline{e} \perp \underline{q}$).

To avoid double-counting (because of the symmetry of the sin and cos functions) it is necessary to restrict \underline{q} to one half of its possible values e.g. by saying $q_x > 0$.

Provided that the potential $V(r)$ between atoms (not between phonons!) has a finite range, one can show that $\omega_q \propto |\underline{q}|$ for small $|\underline{q}|$. In anisotropic crystals we can write

$$\omega_q \rightarrow c_{\hat{q},j} |\underline{q}| \text{ as } |\underline{q}| \rightarrow 0.$$

where the c_j are longitudinal ($j=1$) and transverse ($j=2$ or 3) sound velocities that may depend on the direction of the phonon wave, \hat{q} , relative to the crystal lattice. This is known as a *linear dispersion relation*. In an isotropic crystal there is a single scalar sound speed c .

At very low temperatures, the Boltzmann factor implies that only the low-energy modes are excited i.e. modes with small $\omega_q \propto |\underline{q}|$. Thus, it does not hurt to assume a linear dispersion relation for all $3N$ modes, for simplicity. Also assuming an isotropic crystal for simplicity, the energy of mode q is determined entirely by $|\underline{q}|$, and the mean *phonon* occupation number of the mode is equal to $\bar{n}_q = [\exp(\beta c q) - 1]^{-1}$ where we have used q for $|\underline{q}|$, which we can do if we remember [2nd year notes] that the number of modes in an infinitesimal volume element in the wave-vector space is $3V \cdot 4\pi(q/2\pi)^2 d(q/2\pi) = [3V/2\pi^2] q^2 dq$, where V is the volume of the crystal and the 3 comes from the 3 polarisations. (It is only this last factor that differs from the photon case [Chapter 6 of this course / QED notes], as photons have only 2 polarisations.)

Strictly (and unlike electromagnetic waves in the vacuum, for example), there is a finite total number of modes, $3N$, as noted already. Thus there is a maximum possible q , called the Debye q , written q_D , defined by $3N = (V / 2\pi^2) q_D^3$.

Exercise: derive this maximum.

Thus, the mean energy in the Canonical ensemble is

$$\begin{aligned}
 U &= \frac{3V}{(2\pi)^3} \int_0^{q_D} 4\pi q^2 dq (\hbar c q) \left(\frac{1}{\exp(\beta \hbar c q) - 1} + \frac{1}{2} \right) \\
 &= \frac{3V}{(2\pi)^3} \left(\frac{1}{\beta \hbar c} \right)^4 \int_0^{\beta \hbar c q_D} 4\pi x^2 dx (\hbar c x) \frac{1}{\exp x - 1} + (\text{term indep. of } \beta) \\
 &\doteq \frac{3V}{2\pi^2} \frac{(k_B T)^4}{(\hbar c)^3} \int_0^{\beta \hbar c q_D} \frac{x^3 dx}{\exp x - 1}
 \end{aligned}$$

As $T \rightarrow 0$, $\beta \hbar c q_D \rightarrow \infty$ so that

$$U \doteq \frac{3V}{2\pi^2} \frac{(k_B T)^4}{(\hbar c)^3} \int_0^{\infty} \frac{x^3 dx}{\exp x - 1}$$

Therefore $U \propto T^4$ (the same as the photon case).

Thus the specific heat goes as $C_V \propto T^3$ at low T .

[However, since $C_V \propto T$ for electrons (at realistic temperatures at least), the electronic contribution dominates the lattice specific heat at low enough temperature.]

8d: Some exactly soluble models of interacting systems in statistical mechanics

There are not very many exactly solved problems in the statistical mechanics of interacting many-body systems. Two examples are

(i) A 1D system of quantum particles interacting via a point pair potential $V_0 \delta(x_i - x_j)$. This is known as the Bethe model (related also to the Luttinger model) and makes a reasonable simplified description of a surprising number of difficult condensed matter phenomena. We will not study this model here.

(ii) A collection of particles fixed on lattice sites and interacting only via the z -component of their spin variables. This is known as the ISING MODEL and gives a reasonable description of FERROMAGNETISM in some systems. Even this simple model cannot be solved exactly in all cases. A summary of some known exact results is as follows:

1D Ising model, nearest-neighbour interactions only

Solved (and we will do it!), does **not** exhibit ferromagnetism.

1D Ising model, sufficiently long-ranged interactions between spins

Solved for a class of interactions including power law $V = c|R_i - R_j|^{-p}$.

Does exhibit ferromagnetism for $p \leq 2$

2D Ising model with near-neighbour interactions

Solved exactly in zero external B-field, gives ferromagnetism

3D Ising model with near-neighbour interactions

NOT solved in detail, but known to exhibit ferromagnetism.

3D Icing
model



There are a number of generalizations of the Ising model that can be solved in some cases, such as the Baxter models [R. Baxter, Phys. Rev. Lett. **26**, 832 (1971), R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic, London, 1982] and the Potts-type models [F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982)]. These models mimic a number of important problems in physics, and continue to be studied today, but are not on the present syllabus.

We will study the Ising model of ferromagnetism in some detail, in order to gain exposure to an exact and some approximate methods of solution. We first look at some qualitative physics of ferromagnetism, to set the model in context.

8e Ferromagnetism in general

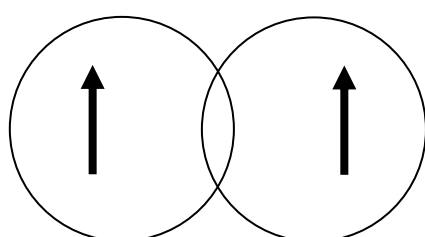
Ferromagnetism is the well-known phenomenon that certain substances (e.g. iron) retain a finite magnetic moment (i.e. they remain "magnetized") even after a polarizing external magnetic field is removed. This phenomenon disappears at temperatures above some maximum value T_c , the *ferromagnetic critical temperature*. As the temperature is raised through T_c the system undergoes a *phase transition* from the magnetized to the unmagnetized state. Various properties are singular functions of T at $T = T_c$.

The physical origin of the persistent magnetization is certainly the magnetic moment of unpaired electron spins: each spin generates a B field that acts on the world outside, and when the spins are partially or completely lined up in the same direction, these B fields add to give the net B field that surrounds a permanent magnet.

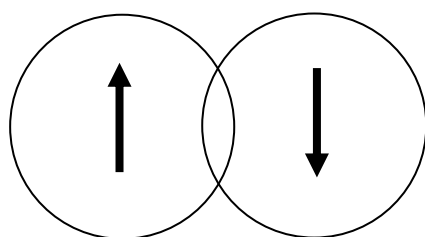
However simple electromagnetic dipole-dipole interactions between these spins are not directly responsible for the coupling that keeps the spins lined up at finite T where thermal effects would tend to randomize the spins. We know this because, for many ferromagnets,

T_c is of order a few thousand Kelvin so that there must be some energy involved in keeping the spins aligned, that is of order $k_B T_c \approx 0.1 \text{ eV}$. Such a large energy cannot possibly arise from the electromagnetic dipole-dipole interaction between either nuclear or electronic spins. Rather the energy is of the order of chemical binding energies, which are essentially Coulombic rather than magnetic.

In fact a coupling $O(eV)$ between spins on electrons in overlapping orbitals results indirectly from Coulomb energies via the Pauli principle. A very simplified argument runs as follows:



Spins parallel, Pauli principle operates, keeping electrons apart. This **lowers** the repulsive (positive) Coulomb interaction energy between them.



Spins **antiparallel**: Pauli principle does not operate, and there is a high repulsive Coulomb energy.

The above argument is actually too simple as it suggests (incorrectly) that the Pauli effect always favours nearby spins to lie parallel. In fact one must also consider the competing effects of electron-nuclear attraction. The combined effect can either have "ferromagnetic sign" as above, or "antiferromagnetic sign" meaning that a lower energy is achieved with antiparallel spins. The argument given does at least show how large electrostatic energies can come into play in determining an apparently magnetic effect involving spins.

8f: Heisenberg spin coupling: Heisenberg model of a ferromagnet

Heisenberg spin-spin coupling

When the above argument is carried out in more detail the resulting interaction energy is often well approximated, for isotropic crystals, by the *Heisenberg interaction energy*

$$\hat{H}_{12} = -J_{12} \hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 \quad (8f.1)$$

where $\hat{\mathbf{S}} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$ represents the set of Pauli spin operators for each spin. The coupling constant J_{12} involves matrix elements over orbitals, and is positive for ferromagnetic coupling (where the two spins tend to line up in the same direction) and negative for antiferromagnetic coupling (where the spins tend to be antiparallel). Note that J_{12} will depend both on the nature of the unpaired electronic orbitals, and on the spatial separation R of the atoms.

Heisenberg model of ferromagnetism

When a large collection of spins in a crystal interact pairwise with their neighbours via the interaction (8f.1), with $J > 0$, we have the *Heisenberg model of ferromagnetism*. Even this relatively simple many-body Hamiltonian cannot be diagonalised exactly for a 3D crystal, and does not give an exactly soluble Statistical Mechanical model. Interesting approximate solutions involve the statistical mechanics of “magnons” or spin waves. We will not cover this topic here, but rather concentrate on a simpler, essentially classical, model of ferromagnetism: the Ising model.

8g: Ising spin-spin coupling, Ising model of ferromagnetism

Ising spin coupling

The Ising model results if we ignore all except the z -component of spin in the Heisenberg pair interaction (8f.1), giving

$$\hat{H}_{12} = -J_{12} \hat{\sigma}_{1z} \hat{\sigma}_{2z} \quad (8g.1)$$

This is known as the *Ising spin-spin coupling Hamiltonian*. (Physically there is some basis for (8g.1) in highly anisotropic substances.) This Hamiltonian commutes with both $\hat{\sigma}_{1z}$ and $\hat{\sigma}_{2z}$, corresponding physically to the fact that spins aligned solely in the z direction are either parallel or antiparallel and so exert no torque on each other. The 4 exact eigenstates of (8g.1) are as follows, where the first arrow represents the state of the first spin, and the second arrow that of the second spin:

$$\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow \text{ and } \downarrow\downarrow$$

That is, the states are *product* (unentangled) states, in which each of the spin projections s_{1z} , s_{2z} can be independently specified. Thus the energies are

$$E(s_{1z}, s_{2z}) = -J_{12} s_{1z} s_{2z}$$

where the spin projections s_{1z} , s_{2z} are now just numbers taking the values ± 1 , not operators.

Ising model of a macroscopic ferromagnet

When many spins interact via (8g.1) and in the presence of an external magnetic field B we have the Ising Hamiltonian

$$\hat{H}^{\text{Ising}} = -B \sum_n \alpha_n \hat{\sigma}_{nz} - \sum_{n>p} J(n, p) \hat{\sigma}_{nz} \hat{\sigma}_{pz} \quad (8g.2)$$

which has exact eigenstates $|i\rangle$ given by independently specifying values ± 1 for the z -spin projections $s_n \equiv s_{nz}$ of all the spins from $n = 1$ to $n = N$:

$$i \equiv (s_1, s_2, \dots, s_N).$$

The exact energy of this eigenstate is

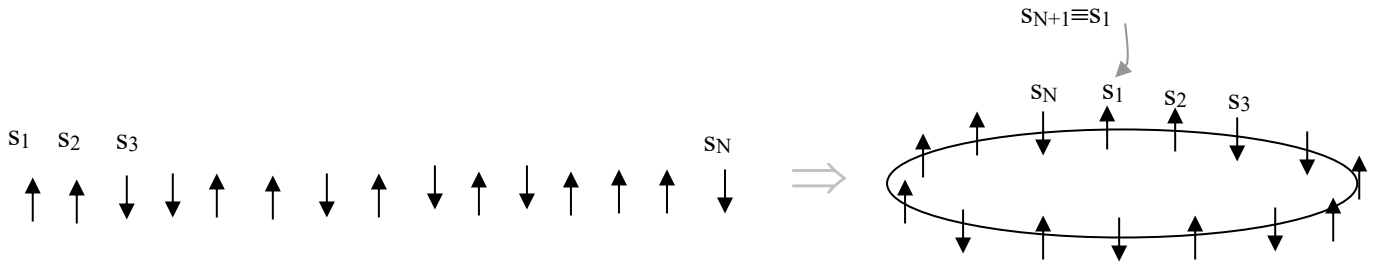
$$E_i^{\text{Ising}} \equiv E^{\text{Ising}}(s_1, \dots, s_N) = -B \sum_n \alpha_n s_n - \sum_{n>p} J(n, p) s_n s_p, \quad s_i = \pm 1 \quad (8g.3)$$

Discussion: why $n>p$ in the second sum?

8h Exact solution of uniform nearest-neighbour 1D Ising model.

The exact solution for the uniform nearest-neighbour 1D Ising model was found by Ising (surprise, surprise) in his 1924 PhD thesis. The proof below, using periodic boundary conditions, is due to Kramers and Wannier 1941. [For a discussion why the model gives the same results without periodic boundary conditions, in the large N limit, see J. Dobson, J. Math. Phys. **9**, 40 (1969).] The periodic boundary conditions means that for spins on a 1D lattice (chain) we regard $s_{N+1} \equiv s_1$. That is, the chain forms a loop. Thus Kramers and Wannier considered the uniform, periodic, nearest-neighbour Ising Hamiltonian

$$H = -\alpha B \sum_{n=1}^N s_n - J \sum_{n=1}^{N-1} s_n s_{n+1} - J s_N s_1 \quad (8h.1)$$



As above, a microstate (energy eigenstate) of the chain is $i \equiv (s_1, s_2, \dots, s_N)$.

The partition function is $\sum_i e^{-\beta E_i} = \sum e^{-\beta H(s_1, \dots, s_N)}$

i.e.
$$Z = \sum_{\substack{s_1 = \pm 1 \\ s_2 = \pm 1, \dots}} e^{-\beta H(s_1, \dots, s_N)} = \sum_{\substack{s_1 = \pm 1 \\ s_2 = \pm 1, \dots}} e^{\beta \left(\alpha B \sum_{n=1}^N s_n + J \sum_{n=1}^N s_n s_{n+1} \right)} \quad (8h.\sqrt{2})$$

Now $\exp(\text{sum}) = \text{product}(\text{exponentials})$, so we can write

$$Z = \sum_{s_1 \dots s_N} \exp(\beta J s_1 s_2 + \beta \alpha B \frac{s_1 + s_2}{2}) \exp(\beta J s_2 s_3 + \beta \alpha B \frac{s_2 + s_3}{2}) \times \dots \times \exp(\beta J s_N s_1 + \beta \alpha B \frac{s_N + s_1}{2})$$

(each term $e^{\beta \alpha B s_n}$ has been split in 2, $e^{\beta \alpha B s_n} = e^{\beta \alpha B s_n / 2} e^{\beta \alpha B s_n / 2}$, to increase symmetry).

Defining a 2×2 matrix T ,

$$T_{ss'} = \exp \left(\beta J s s' + \beta \alpha B \frac{s + s'}{2} \right) \quad (8h.2)$$

we can write Z as

$$\begin{aligned}
Z &= \sum_{s_1 \cdots s_N} T_{s_1 s_2} T_{s_2 s_3} \cdots T_{s_{N-1} s_N} T_{s_N s_1} = \sum_{s_1, \pm 1} (T^N)_{s_1 s_1} \\
&= \text{Trace}(T^N)
\end{aligned} \tag{8h.3}$$

Now since T is real and symmetric, we can diagonalize it using an orthogonal matrix P . That is, $T = P\tau P^T$, where $P^T P = I = P P^T$, and where τ is a diagonal matrix:

$$\tau = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \text{ with diagonals equal to the eigenvalues } \lambda \text{ of } T.$$

To find the eigenvalues, we write out T as a matrix explicitly:

$$T = \begin{pmatrix} e^{\beta J + \beta \alpha B} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta \alpha B} \end{pmatrix} \tag{8h.\pi}$$

Thus the characteristic equation $\det(T - \lambda I) = 0$ is:

$$\begin{aligned}
\lambda^2 - 2[e^{\beta J} \cosh(\beta \alpha B)]\lambda + 2 \sinh 2\beta J &= 0 \\
\therefore \lambda_{\pm} &= e^{\beta J} \cosh(\beta \alpha B) \pm \sqrt{\exp(2\beta J) \cosh^2(\beta \alpha B) - 2 \sinh(2\beta J)}
\end{aligned} \tag{8h.4}$$

Now

$$\begin{aligned}
Z_N &= \text{Tr}(T^N) = \text{Tr}((P\tau P^T)^N) = \text{Tr}(P\tau P^T P\tau P^T \cdots P\tau P^T) \\
&= \text{Tr}(P\tau^N P^T) \text{ using } P P^T = I \\
&= \text{Tr}(P^T P\tau^N) \text{ cyclic property of Trace} \\
&= \text{Tr}(\tau^N) \\
&= \text{Tr} \begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} \\
\therefore Z_N &= \lambda_+^N + \lambda_-^N = \lambda_+^N \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right] \\
\therefore \ln Z_N &= N \ln \lambda_+ + \ln \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right] \\
&\rightarrow N \ln \lambda_+ + o(1) \text{ as } N \rightarrow \infty \text{ provided } \lambda_+ > \lambda_-.
\end{aligned} \tag{8h.5}$$

(Here “o” means of smaller order than its argument, as opposed to “O”, meaning of the same order.) But from (8h.\sqrt{2}), the magnetization per particle is

$$\begin{aligned}\frac{M}{N} &= \frac{\alpha}{N} \left\langle \sum_n s_n \right\rangle = \frac{\alpha}{N} \frac{1}{\alpha\beta} \frac{\partial Z_N}{\partial B} = \frac{1}{N\beta} \frac{\partial}{\partial B} \ln Z_N = \frac{1}{\beta} \frac{\partial}{\partial B} \ln \lambda_+ \\ \therefore \frac{M}{N} &= \frac{1}{\beta} \frac{1}{\lambda_+} \frac{\partial \lambda_+}{\partial B}\end{aligned}\quad (8h.6)$$

From this it can be shown (see problem sheet 2) that

$$\frac{M}{N} = \frac{\alpha e^{\beta J} \sinh(\alpha\beta B)}{\sqrt{[e^{\beta J} \sinh(\alpha\beta B)]^2 + e^{-2\beta J}}}. \quad (8h.7)$$

Clearly this goes to zero as B goes to zero. That is, the model predicts zero average magnetization once the magnetic field is turned off. From this you might be tempted to conclude that there is no ferromagnetism in the 1D Ising model. The conclusion is true (as we'll see), but this reasoning is false. If $B=0$ there is no preferred direction in which the model might have magnetization. If *spontaneous magnetization* were to occur, it would equally likely be positive (i.e. in the $+z$ direction) as negative (in the $-z$ direction). Thus an exact statistical mechanical calculation as we are doing here will *always* predict zero mean spontaneous magnetization, even for models that are ferromagnetic i.e. that do predict spontaneous magnetization.

How then, do we calculate whether spontaneous magnetization will occur. The key idea is to look at *correlations between spins*. If spontaneous magnetization does occur then (roughly) all the spins will be pointing in the same direction, even though that direction is random, equally likely to be up or down. Therefore we should consider the quantity $\langle s_n s_{n+r} \rangle$. This cannot be calculated by a simple derivative of the partition function like $\langle s_n \rangle$ in (8h.6) above. Instead, we have to return to the distribution in the canonical ensemble:

$$\begin{aligned}P(s_1, s_2, \dots, s_N) &= Z^{-1} \exp(\beta J s_1 s_2 + \beta \alpha B \frac{s_1 + s_2}{2}) \times \dots \times \exp(\beta J s_N s_1 + \beta \alpha B \frac{s_N + s_1}{2}) \\ &= Z^{-1} T_{s_1 s_2} T_{s_2 s_3} \dots T_{s_{N-1} s_N} T_{s_N s_1}\end{aligned}$$

Since ultimately we are interested in the macroscopic limit, we will assume $r \ll N$. Also, since the model is translationally invariant (that is, the correlation will depend only on r , not on n), we can take $n=1$ WLOG. Thus we calculate

$$\begin{aligned}\langle s_1 s_{1+r} \rangle &= \sum_{s_1 \dots s_N} P(s_1, s_2, \dots, s_N) s_1 s_{1+r} \\ &= Z^{-1} \sum_{s_1 \dots s_N} s_1 T_{s_1 s_2} T_{s_2 s_3} \dots T_{s_r s_{r+1}} s_{r+1} T_{s_{r+1} s_{r+2}} \dots T_{s_{N-1} s_N} T_{s_N s_1} \\ &= Z^{-1} \sum_{s_1 \dots s_N, z_1, z_{r+1}} \left(z_1 \delta_{z_1 s_1} \right) T_{s_1 s_2} T_{s_2 s_3} \dots T_{s_r s_{r+1}} \left(z_{r+1} \delta_{z_{r+1} s_{r+1}} \right) T_{s_{r+1} s_{r+2}} \dots T_{s_{N-1} s_N} T_{s_N s_1}\end{aligned}$$

where the extra delta-functions introduced here trivially make no difference. But it allows us to rewrite the sum as

$$\begin{aligned}
\langle s_1 s_{1+r} \rangle &= Z^{-1} \sum_{s_1 \dots s_N, z_1, z_{r+1}} (z_1 \delta_{s_1 z_1}) T_{z_1 s_2} T_{s_2 s_3} \dots T_{s_r s_{r+1}} (z_{r+1} \delta_{s_{r+1} z_{r+1}}) T_{z_{r+1} s_{r+2}} \dots T_{s_{N-1} s_N} T_{s_N s_1} \\
&= Z^{-1} \sum_{s_1 \dots s_N, z_1, z_{r+1}} (S_{s_1 z_1}) T_{z_1 s_2} T_{s_2 s_3} \dots T_{s_r s_{r+1}} (S_{s_{r+1} z_{r+1}}) T_{z_{r+1} s_{r+2}} \dots T_{s_{N-1} s_N} T_{s_N s_1} \\
&= Z^{-1} \text{Tr} [S T^r S T^{N-r}],
\end{aligned}$$

where we have defined the matrix S by $S_{ss'} = s \delta_{s,s'}$. That is,

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

using the same basis in which T is expressed as (8h.π). Since there are more T s than S s in the above expression, it again makes sense to change to the basis in which T is diagonal. That is,

$$\begin{aligned}
\langle s_1 s_{1+r} \rangle &= Z^{-1} \text{Tr} [S (P \tau P^T)^r S (P \tau P^T)^{N-r}] \\
&= Z^{-1} \text{Tr} [P^T S P \tau^r P^T S P \tau^{N-r}] \\
&= Z^{-1} \text{Tr} [(P^T S P) \tau^r (P^T S P) \tau^{N-r}]
\end{aligned}$$

Now in the limit of large N , $Z^{-1} \tau^{N-r} \rightarrow (\lambda_+)^{-r} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Evaluating the above in the case of no applied magnetic field, we find (see problem sheet 3),

$$\begin{aligned}
\langle s_1 s_{1+r} \rangle &= \text{Tr} \left[(P^T S P) \begin{pmatrix} 1 & 0 \\ 0 & (\lambda_- / \lambda_+)^r \end{pmatrix} (P^T S P) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right] \\
&\propto (\lambda_- / \lambda_+)^r
\end{aligned} \tag{8h.8}$$

Because $|\lambda_-| < |\lambda_+|$ for any nonzero temperature, the spin-correlation decays to zero as r grows. That is, there is *no long-range order*, which in this case means *no spontaneous magnetization*. Thus, **ID nearest-neighbour Ising model cannot describe ferromagnetism**.

Onsager (1940s) showed that **2D nearest-neighbor Ising model does exhibit ferromagnetism** (K. Huang, “Statistical Physics”, 2nd Ed 1987). One can also show that long-range interactions (e.g. $J_{nm} = c |n - m|^{-p}$, $p \leq 2$) give ID ferromagnetism (Dyson, 1970s).

For 1D problems with beyond-nearest-neighbour interactions, the Kramers-Wannier transfer matrix method can still be applied, but the dimension of the matrix grows rapidly with the range of the interaction. For infinite-range interactions, such as the power-law example above, the transfer matrix also becomes infinite in dimension. In general, from our transfer matrix approach it is clear that the partition function will be a smooth function of temperature and applied magnetic field unless the largest eigenvalue λ_+ becomes degenerate then crosses with another eigenvalue: if that happens then (8h.5) can have a sharp (non-differentiable) features as a function of such controllable parameters. This is known as a *phase transition*, and ferromagnetism is an example (there is spontaneous magnetization below the Curie temperature, but it disappears abruptly at the Curie temperature. It can be shown that in 1D Ising model this only occurs for infinite-range interactions).

The 2D Ising model with nearest-neighbour interactions can be treated using a transfer matrix technique where each matrix represents the interaction between two (infinite) rows of spins: thus the matrices are already infinite even with nearest-neighbour interactions in 2D space.

While it is thus very difficult to derive the Curie phase transition and its temperature from an exact model, it is quite easy using a *mean-field* approximation. We now turn to this.

8i: Mean-field (self-consistent field) approach to Ising model

The following approximate solution to the Ising model is an example of a very general approach known as (heuristic) *mean field theory*.

Assuming a translationally invariant (homogeneous) and directionally invariant (isotropic) 1D model where J_{np} depends only $|n-p|$, the piece of the Hamiltonian containing a particular spin labelled n is

$$H(n) = -\alpha B s_n - \sum_{p: p \neq n} J(|p-n|) s_p s_n = -\alpha B'_n s_n, \quad (8i.1)$$

where

$$B'_n = B + \frac{1}{\alpha} \sum_{p: p \neq n} J(|p-n|) s_p \quad (8i.2)$$

Of course B'_n will fluctuate in time as the $\{s_p\}$ fluctuate; it will also tend to depend on s_n , as neighbouring spins will be correlated – they will more often have same spin for $J > 0$, the ferromagnetic case. **The essence of mean field theory is to ignore these fluctuations and correlations** and replace the dynamical variable s_p in B'_n by its equilibrium average value \bar{s} ; we then find an equation for \bar{s} in terms of itself.

Specifically, for a homogeneous system, we make the replacement $B'_n \rightarrow \bar{B}$ where the equation for the *mean field* is

$$\bar{B} = B + \frac{1}{\alpha} \bar{s} \sum_p J(|p-n|) = B + \frac{1}{\alpha} \bar{s} G \quad \text{where } G = \sum_{R=-\infty}^{\infty} J(R). \quad (8i.3)$$

If B' is replaced by its average \bar{B} , the Hamiltonian (8i.1) becomes that of an ordinary spin in a fixed "external" field \bar{B} (the mean field). We found previously (2nd Year Stat Mech Notes, also Problem Sheet 1) that the average spin of such a paramagnet is

$$\bar{s} = \tanh(\beta \alpha \bar{B}), \quad \beta = \frac{1}{kT} \quad (8i.4)$$

Eqs. (8i.3), (8i.4) are self-consistent equations for \bar{s} and \bar{B} .

It is convenient to introduce a dimensionless parameter $x = \beta \alpha \bar{B}$ so that $\bar{s} = \frac{kT}{J'} x - \frac{\alpha B}{J'}$.

That is, the self consistent condition can be written

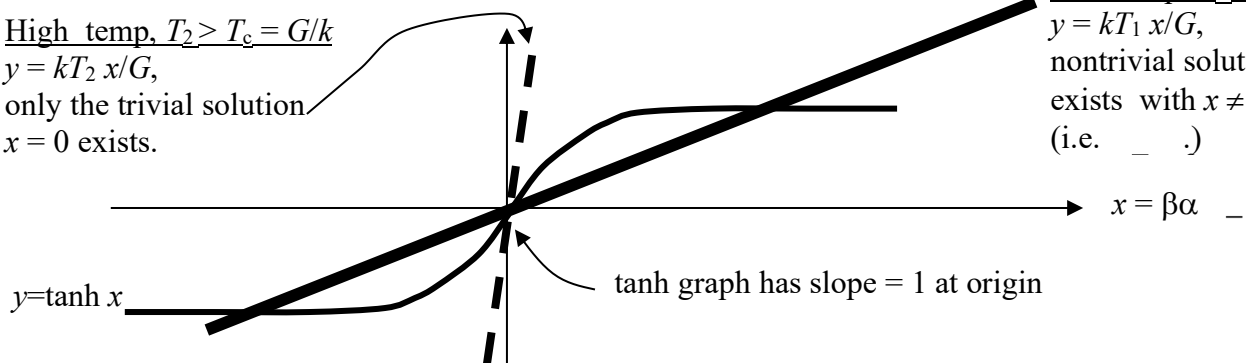
$$\frac{k_B T}{G} x - \frac{\alpha B}{G} = \tanh x \quad (8i.4)$$

This describes the intersection of straight line and tanh graph: We solve (8i.4) graphically

Graphical Solution In Zero External B-field, $B = 0$:

High temp, $T_2 > T_c = G/k$
 $y = kT_2 x/G$,
 only the trivial solution
 $x = 0$ exists.

Low temp, $T_1 < T_c$
 $y = kT_1 x/G$,
 nontrivial solution
 exists with $x \neq 0$.
 (i.e. $\neq 0$.)



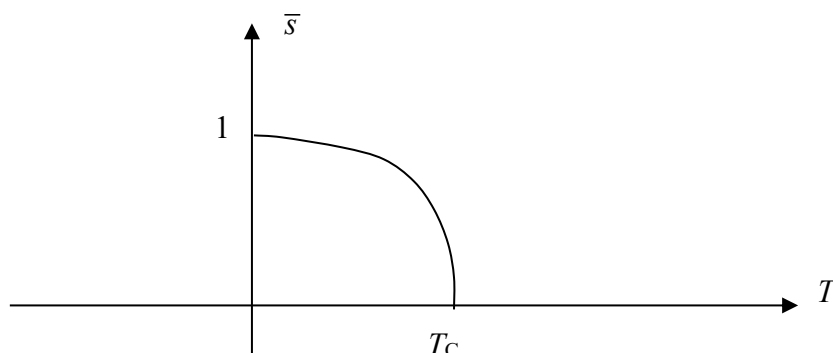
[Only the case $B = 0$ is plotted here, but the same technique can be applied for a finite applied external field B , as this simply moves the straight line up or down.] Note that spontaneous magnetization (ferromagnetism) exists if the straight line cuts the curve more than once. This in turn only happens if the slope of the straight line is less than the maximum slope of $\tanh x$, which is 1 and occurs at $x = 0$. Thus we only get ferromagnetism if $k_B T/G < 1$. That is, we get ferromagnetic order for $T < T_c$ where

$$T_c = \frac{G}{k_B} \text{ where } G = \sum_R J(R) \quad (8i.5)$$

Here we can consider the sum over R to be limited not to the non-zero integers on the line, but rather the positions of all the atoms around a given atom in 1-, 2-, or 3-D. Note that the larger the spin-spin interaction is, or the longer range it is, the higher the Curie temperature will be, as the more energetically favourable it is for the spins to align (for $G > 0$).

Note: This mean field approximation says we get a phase transition to a spontaneously magnetized state regardless of dimensionality. This agrees with known facts for 2D&3D, but *disagrees* with true Ising solution given above, in 1D for finite-range interactions.

The spontaneous magnetization $\alpha \bar{s}$ for $B=0$ is indicated by following graph (qualitative)



We can find how the magnetization dies away when the temperature is just below T_c and rises through T_c . Here x is small and we can Taylor-expand $\tanh x \approx x(1 - x^2/3 + \dots)$. The mean-field condition then becomes (at $B = 0$)

$$x \left(1 - \frac{x^2}{3} \right) = \frac{k_B T}{G} x \Rightarrow x^2 = 3 \left(1 - \frac{k_B T}{G} \right)$$

$$\therefore x = \sqrt{3 \left(1 - \frac{T}{T_C} \right)}$$

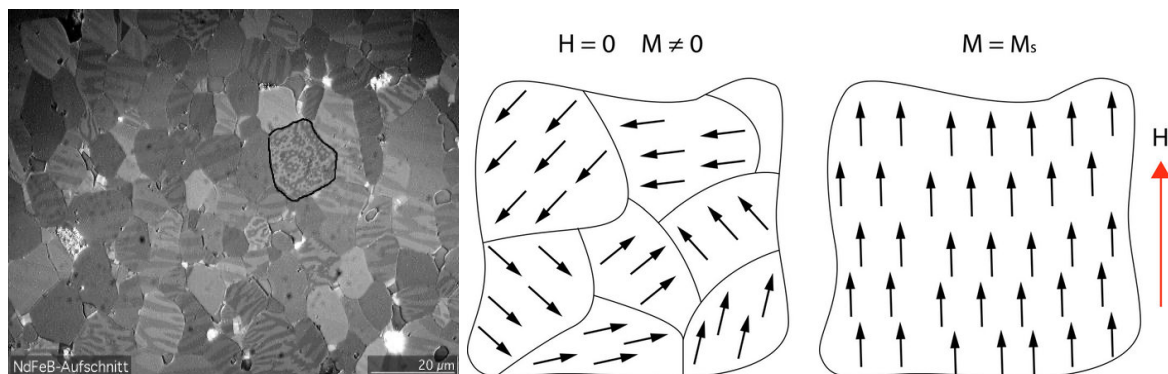
thus the magnetisation, proportional to x , is

$$\bar{s} \propto \left(1 - \frac{T}{T_c} \right)^{1/2} \quad (8i.6)$$

The power $1/2$ appearing in (8i.6) is known as the **critical exponent** for the order parameter in mean-field approximation to the Ising model. The region just below T_C where (8i.6) is valid is called the *critical region*.

The investigation of critical exponents, and the critical region T near T_C generally, has been a prime focus of Statistical Mechanics since the mid 20th century. It is interesting because of its *universality* properties. That is, the same exponents turn up in completely unrelated physical systems. It has given rise to whole new approaches such as the *renormalization group* method, not studied here. Several Nobel prizes have been given in the area of phase transitions, including the three in 2016.

Finally, note that at zero external field, there are in fact always three solutions to (81.4): a positive, a zero, and a negative one. The zero is in fact unstable; the other two have lower free energy (see Problem Sheet). In real materials the two stable solutions typically co-exist. That is, one solution does not extend over the entire material. Rather, there are *magnetic domains* separated by defects, called *domain walls*, in which one or the other stable solution persists for a very long time. The size of these domains is typically of order 10 micrometers. This sounds microscopic, but each domain contains a lot of atoms ($\sim 10^{15}$ in 3D).



Polarisation-sensitive photograph (left) [https://en.wikipedia.org/wiki/Magnetic_domain] and diagram (right) [Ashima Arora, Phd Thesis (2018)] of magnetic domains.