

Ising Model

(Lesson 19 of
22/04/20)
Compiled: May 17,
2020

Statistical Mechanics, at its core, allows us to understand and quantify macroscopic phenomena starting from microscopic dynamics. In particular, it provides a window in how surprisingly *complex emergent* behaviours arise from the interaction of many relatively *simple* components.

One such example is given by **phase transitions**, i.e. abrupt changes of a system's properties when surpassing a well-defined threshold. For instance, water becomes ice when its temperature dips below $T = 0^\circ\text{C}$, or steam when $T = 100^\circ\text{C}$ is reached.

Phase transitions

Perhaps one of the simplest models describing a phase-transition is the **Ising Model**. At its origin, it was meant as an explanation of **ferromagnetism**.

Certain materials possess no net magnetic moment above a certain temperature T_c (the **Curie** temperature), and develop a temporary induced magnetization only in the presence of an external magnetic field (**paramagnetic** phase). However, when T dips below T_c , they exhibit **spontaneous magnetization**, even in the absence of any external field, and behave like *permanent magnets* (**ferromagnetic** phase).

*Ferromagnetism
and paramagnetism*

Classical Statistical Mechanics, by itself, cannot explain this kind of behaviour. In fact, if we suppose that magnetization arises from *tiny current loops*, its (canonical) thermal average is predicted to be always 0, regardless of temperature¹. This is because paramagnetism and ferromagnetism are inherently **quantum** phenomena, arising from the alignment of intrinsic magnetic dipoles of atoms, i.e. **spins**.

*Quantum nature of
ferromagnetism*

In 1920, Wilhelm Lenz proposed a model of interacting spins on a lattice to his student Ernst Ising, who then found an analytic solution for the one-dimensional case in 1924. Underwhelmingly, the model did not exhibit any kind of phase transition - but it was still able to capture the attention of many researcher.

An analytic solution for the $d = 2$ generalization was found by Lars Onsager in 1944, requiring a long and sophisticated mathematical derivation. In this case, however, the model was complex enough to capture a **phase transition** - a very important result for Statistical Mechanics.

The Ising Model a lot of research and applications. Nowadays, the Ising Model

Applications

¹^This is consequence of the Bohr-van Leeuwen theorem. See [ferromagnets]

is relevant for simulating the behaviour of gases on a discretized grid (**lattice gases**), more complex **spin glasses**, and also the activity of **neural networks** (e.g. Hopfield networks). Also higher dimensional cases are of interest - for example the $d = 4$ model is relevant for modelling *spacetime* (3 dimensions for space plus 1 for time). While no analytic solution for $d > 2$ is known, efficient numerical methods are available, and will be examined in a later chapter.

1.1 Lattice gas

The Ising Model can be introduced as a purely classical model when applied to the behaviour of gas particles in a discretized grid (lattice), completely bypassing the need to deal with quantum effects which are often difficult to interpret.

Consider a system \mathcal{S} of N particles of mass m enclosed in a volume V . Its state is completely specified by $3N$ positions and $3N$ momenta:

$$\begin{aligned}\mathbf{Q} &= (q_{1x}, q_{1y}, q_{1z}, \dots, q_{Nx}, q_{Ny}, q_{Nz}) = (\mathbf{q}_1, \dots, \mathbf{q}_N) \in \mathbb{R}^{3N} \\ \mathbf{P} &= (p_{1x}, p_{1y}, p_{1z}, \dots, p_{Nx}, p_{Ny}, p_{Nz}) = (\mathbf{p}_1, \dots, \mathbf{p}_N) \in \mathbb{R}^{3N}\end{aligned}$$

Suppose that the particles interact through a potential $V_N(\mathbf{Q})$ depending only on the spatial coordinates. The Hamiltonian is then given by:

$$\mathcal{H}_N(\mathbf{Q}, \mathbf{P}) = \sum_{i=1}^N \frac{\|\mathbf{p}_i\|^2}{2m} + V_N(\mathbf{Q}) \quad (1.1)$$

The system is at equilibrium with a much larger environment, with which it exchanges both energy and particles. Denoting with P the grand-canonical pressure, the grand-canonical partition function is given by:

$$e^{\beta PV} = \sum_{N=0}^{+\infty} \frac{1}{N!} \int_{\Gamma_N} \prod_{i=1}^N \frac{d^3 \mathbf{q}_i d^3 \mathbf{p}_i}{h^3} \exp(-\beta[\mathcal{H}_N(\mathbf{Q}, \mathbf{P}) - \mu N]) \quad (1.2)$$

where μ is the system's chemical potential, representing the energy cost of adiabatically adding one particle to \mathcal{S} such that the resulting $N + 1$ system is still at equilibrium. Physically, the value of μ *fixes* the average number of particles $\langle N \rangle$ in \mathcal{S} : if we “forcefully empty \mathcal{S} ”, making $N = 0$, particles will flow in \mathcal{S} from the environment until the “cost” of adding a new particle reaches μ , and then N will only slightly oscillate.

Since in (??) momenta and positions are independent, the integral over \mathbf{P} in (??) is just a gaussian integral, resulting in:

$$\int_{\mathbb{R}^{3N}} \frac{d^{3N} \mathbf{P}}{h^{3N}} \exp\left(-\beta \sum_{i=1}^N \frac{\|\mathbf{p}_i\|^2}{2m}\right) = \left(\frac{2\pi m}{\beta h^2}\right)^{\frac{3N}{2}} \equiv \lambda^{-3N} \quad (1.3)$$

where:

$$\lambda = \sqrt{\frac{\beta h^2}{2\pi m}} \underset{\substack{\beta=1/k_B T \\ \hbar=h/2\pi}}{=} \sqrt{\frac{2\pi \hbar}{m k_B T}}$$

is called the **thermal wavelength**², and has dimensions of a length (in fact $[d^3 \mathbf{p} / h^3] = 1/\text{m}^3$).

Substituting (??) back in (??) leads to:

$$\begin{aligned} e^{\beta PV} &= \sum_{N=0}^{+\infty} \frac{1}{N!} \int_{V^N} d^{3N} \mathbf{Q} e^{-\beta V_N(\mathbf{Q})} \left(\underbrace{\frac{e^{\beta \mu}}{\lambda^3}}_z \right)^N = \\ &= \sum_{N=0}^{+\infty} \frac{z^N}{N!} \int_{V^N} d^{3N} \mathbf{Q} e^{-\beta V_N(\mathbf{Q})} \end{aligned} \quad (1.4)$$

We now **discretize** the system's volume V in small cubic *sites*, each with edges of size a (fig. ??). Space is thus divided in cells, each labelled by its 3-dimensional integer indices $\mathbf{x}_i \in \mathbb{Z}^3$.

1. Lattice discretization

More in general, we may consider a d -dimensional system, of d -volume³ V . Then, each cell i will be labelled by d indices: $\mathbf{x}_i \in \mathbb{Z}^d$. In fact, many physical systems can be modelled with $d \neq 3$ dimensions. For example, oxygen interacting with a graphite plane is intrinsically a $d = 2$ system.

Also, cells may be of different shapes: it suffices that they are all equal and that they *tessellate* the entire space. For instance, in $d = 2$ we may subdivide a plane in triangles instead of squares, or with tetrahedra instead of cubes in $d = 3$. In our case we will focus on the simplest choice, the cubic one.

Suppose now that $V_N(\mathbf{Q})$ may be written as sum of two-body interactions:

2. Pair-wise close-range potential

$$V_N(\mathbf{Q}) = \sum_{i < j}^N v(q_{ij}) \quad q_{ij} \equiv \|\mathbf{q}_i - \mathbf{q}_j\| \quad (1.5)$$

where $v(q_{ij})$ is the potential of two particles i and j separated by a (relative) distance q_{ij} . If the gas is made of neutral particles, then $v(q)$ is a close-range attractive interaction, with the shape of a **Lennard-Jones potential** (fig. ??).

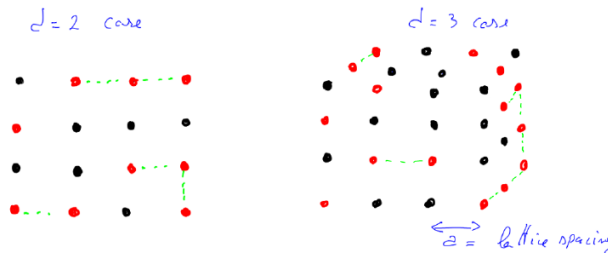


Figure (1.1) – Examples of cubic lattices in $d = 2$ and $d = 3$.

²It is the average de Broglie wavelength of particles of an ideal gas at temperature T , i.e. of particles with energy of order $k_B T$.

³If $d = 1$, then V is a length, if $d = 2$ it is an area, and if $d = 3$ it is the usual volume. For higher dimensions it is a *hyper*-volume.

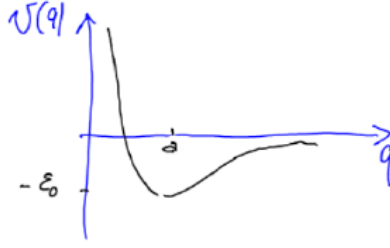


Figure (1.2) – Shape of the Lennard-Jones potential. Two particles separated by a sufficiently small distance $q \sim a$ are weakly attracted to each other, and repelled if $q \rightarrow 0$ (as if they were *hard spheres*). As $v(q)$ quickly vanishes for $q \rightarrow \infty$, particles that are too far from each other *do not interact*. Physically, such $v(q)$ is due to interactions of the electronic clouds of neutral atoms. For q too small, the (Pauli) repulsion of electrons dominates. Due to quantum fluctuations, the electronic clouds are not completely uniform: sometimes an electron “spends more time on one side”, producing a short-lived polarization. If $q \sim a$, one such “instantaneous dipole” may induce a temporary dipole in the other atom. It is this correlation between temporary polarizations that leads to a weak attractive force between the atoms (van der Waals force).

If we choose the lattice step length a as the position of the minimum of $v(q)$, then on average at most one particle will occupy a given cell at any time - because two particles in the same site would be separated by $q < a$, for which $v(q)$ is strongly repulsive. Moreover, only particles in neighbouring cells are sufficiently close to interact with each other.

So, in **approximation**, we consider a model in which cells can contain **at most one particle** at a time (at its centre), and interactions between non-neighbouring cells are neglected. Thus the distance q between two particles must be a multiple of a , and the interaction potential is given by:

3. *Discretized potential*

$$v(q) = \begin{cases} +\infty & q = 0 \\ -\epsilon_0 & q = a \\ 0 & \text{otherwise} \end{cases}$$

We then define the **occupancy** n_i of cell i as a binary variable:

$$n_i = \begin{cases} 1 & \text{Site } i \text{ contains a particle} \\ 0 & \text{Site } i \text{ is empty} \end{cases}$$

If we know all $\{n_i\}$, then the system (in this discretized approximation) is completely determined. Note that dealing with occupancies *automatically* takes into account the indistinguishability of particles: $n_i = 1$ regardless if cell i is occupied by particle #3 or #42.

Shortcomings of the lattice model. In summary, the lattice model is a way to deal with the complex integration in ??, making the following approximations:

1. **Discretization:** space is divided in small cubic units, each containing **at most one particle**.

2. **Close-range pair-wise interaction:** the only possible interactions are between pairs of neighbouring cells.

As a consequence of these approximation, there is a **maximum density** in the system, corresponding to the case when all cells are occupied - meaning that the system cannot be compressed over a certain threshold.

Moreover, there is no way to understand if, in the densest case, particles are still moving from one cell to the other, “exchanging places” with each other, or if they just stay forever in their original site: in both cases, all the n_i will be equal to 1, and remain constant. In other words, there is no difference between the **liquid** phase and the **solid** one, meaning that the model **cannot appreciate the liquid-solid phase transition**, and so it is not very good for explaining the usual three phases of matter. However, it can describe accurately the liquid-vapour transition, and - more importantly - it leads to the Ising Model, which is very important in physical mechanics.

There are of course more complex methods that relax the lattice approximation, dealing directly with the grand partition function in the continuum. In general they apply perturbation theory to the pair-wise potential, and through quite involved expansions (e.g. Virial expansion, or Mayer cluster expansion) they lead to equations of state capturing both the solid-liquid transition and the liquid-vapour transition. These are all well-known techniques (Mayer worked in the 1940s), with not much conceptual difficulty, apart of some very long and uninspiring computations. For more information, see [huang].

We can now write $V_N(\mathbf{Q})$ as function of the $\{n_i\}$. We consider two particles occupying neighbouring cells as being separated by a distance of a , and experiencing a potential $V(a) = -\epsilon_0$. Then the total potential experienced at cell i is given by summing a contribution of $V(a)$ for each occupied neighbouring cell:

$$V(\mathbf{x}_i) = -\epsilon_0 \sum_{\langle j,i \rangle} n_j$$

The notation $\langle j,i \rangle$ represents a sum over all cells j that are the nearest neighbours of cell i . Then $\sum_{\langle j,i \rangle} n_j$ is exactly to the number of cells *around* i that are occupied. In particular, the total number N of particles in the system is the number of occupied cells:

$$N = \sum_x n_x \tag{1.6}$$

Then, the total potential $V_N(\mathbf{Q})$ can be approximated as the sum of $V(\mathbf{x}_i)$ terms over all cells i that are occupied, leading to:

$$V_N(\{\sigma_i\}) = -\epsilon_0 \sum_{\langle x,y \rangle} n_x n_y \tag{1.7}$$

Note that the product $n_x n_y$ is 1 if and only if both cells are occupied.

So, in other words, V_N is obtained by multiplying the average potential of an interaction ($-\epsilon_0$) by the number of such pairwise interactions, i.e. the number of pairs of neighbouring cells $\langle x,y \rangle$ that are both occupied.

Note that not all sites are treated equally: the ones at the **boundaries** of V have a lower number of neighbours than the ones in the **bulk**. If we consider this asymmetry, letting cells at the margins interact only with their neighbours, the system is said to have **open boundaries**. Alternatively, boundaries may be removed by “connecting” sites at one margin with the ones from the opposite side (**periodic boundaries**). In this way, all cells have exactly the same number of neighbours, and can then be treated the same, achieving **translational invariance**.

Open and periodic boundary conditions

Periodic boundary conditions alter the system’s **topology**. For example, in $d = 1$, the *open* case can be represented as a segment, while the *periodic* one as a circle (fig. ??). In $d = 2$, open boundaries result in a *planar* topology, while periodic conditions produce a *toroidal* surface (fig. ??).

Periodic boundaries and topology

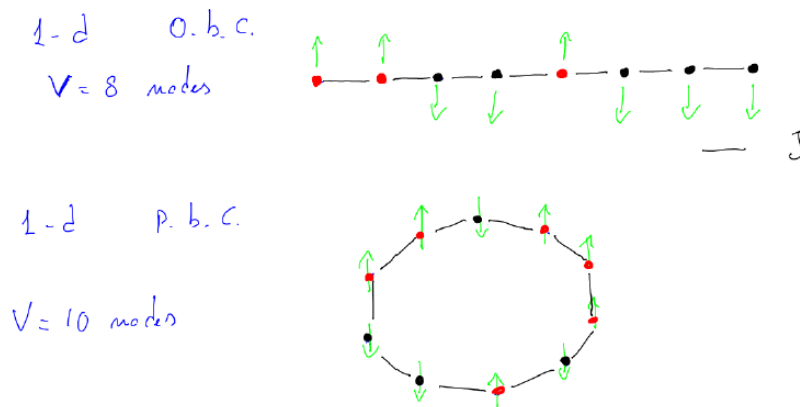


Figure (1.3) – Ising model in $d = 1$ with **open** boundaries (top), or **periodic** boundaries (bottom). Periodic conditions are obtained by *deforming* the initial line to “attach” its two ends.

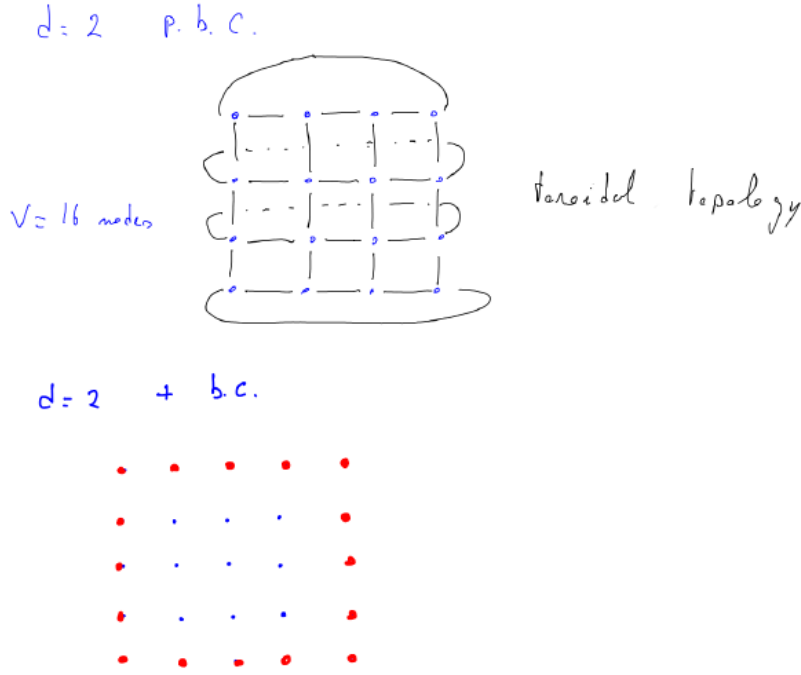


Figure (1.4) – Ising model in $d = 2$, with **periodic** boundaries (top) or **open** boundaries (bottom). Starting from a square, two opposite edges are attached together, forming a cylinder. Then the two circles at the boundaries are attached, deforming the cylinder into a torus.

Boundaries may also be **fixed** to a particular *state*: for example to be always empty or full of particles.

It is now useful to *shift* the occupancies n_i so that they assume symmetrical values. So, we define a “spin-like” occupancy σ_i as follows:

$$\sigma_i = \begin{cases} +1 & \text{Cell } i \text{ is occupied} \\ -1 & \text{Cell } i \text{ is empty} \end{cases} \quad (1.8)$$

Clearly $\sigma_i = 2n_i - 1$, and so:

$$n_i = \frac{1 + \sigma_i}{2}$$

And so we may rewrite (??) and (??) as follows:

$$N = \sum_x \frac{1 + \sigma_x}{2} \quad (1.9)$$

$$V_N(\{\sigma_i\}) = -\epsilon_0 \sum_{\langle x,y \rangle} n_x n_y \quad (1.10)$$

We can then use (??) and (??) to approximate (??) on the lattice:

$$\begin{aligned} Z_{\text{g.c.}} &\equiv e^{\beta PV} \stackrel{(\text{??})}{=} \sum_{\{\sigma\}} \exp(-\beta V_N(\sigma) + N(\sigma) \ln z) = \\ &\stackrel{(\text{??})}{=} \sum_{\{\sigma\}} \exp \left(\beta \epsilon_0 \sum_{\langle x,y \rangle} \frac{1 + \sigma_x}{2} \frac{1 + \sigma_y}{2} + \ln z \sum_x \frac{1 + \sigma_x}{2} \right) \end{aligned} \quad (1.11)$$

In fact the sum over all possible unique⁴ configurations of N particles over all values of N is equivalent to the sum over all possible occupancies $\sigma = \{\sigma_i\}$:

$$\sum_{N=0}^{+\infty} \int_{V^N} \frac{d^{d \cdot N} \mathbf{Q}}{N!} \xrightarrow{\text{discretization}} \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_{\text{last}}=\pm 1} = \sum_{\{\sigma\}}$$

The choice of using spin-like variables (??) has lead to the appearance of constant terms, independent of the system's state, in (??), that we now extract. Consider the argument of the exponential, and expand the sums:

$$\frac{\beta \epsilon_0}{4} \left[\sum_{\langle x,y \rangle} 1 + \sum_{\langle x,y \rangle} (\sigma_x + \sigma_y) + \sum_{\langle x,y \rangle} \sigma_x \sigma_y \right] + \frac{\ln z}{2} \left(\sum_x 1 + \sum_x \sigma_x \right) \quad (1.12)$$

Summing ones over all cells is just the total number of cells in the lattice, each with volume a^d :

$$\sum_x 1 = N_{\text{cells}} = \frac{V}{a^d}$$

For simplicity, let's choose our units so that $a = 1$, and thus:

$$\sum_x 1 = V \quad (1.13)$$

Similarly $\sum_{\langle x,y \rangle} 1$ counts the number of *pairs* of neighbours. Graphically, if we represent cells as nodes in a graph, each connected to its neighbours by edges, then $\sum_{\langle x,y \rangle} 1$ is just the number of edges (fig. ??). Let's consider, for simplicity, a system with periodic boundary conditions - so that all cells have the same number q of neighbours (this is also true for a system with open boundaries, in the limit of a lattice of infinite size). Then, the number of pairs (edges) will be:

$$\sum_{\langle x,y \rangle} 1 = \text{N. of pairs} = \frac{N_{\text{cells}} q}{2} = \frac{V q}{2}$$

where the division by 2 accounts for the fact that edges are *undirected* - i.e. if a is connected to b , then b is connected to a by the same edge. Without this division, we would be counting every edge *twice*.

For a cubic lattice (with no boundaries), each cell has exactly 2 neighbours in each *direction*. So the total number of neighbours q will be twice the number of dimensions d :

$$\sum_{\langle x,y \rangle} 1 = \frac{2Vd}{2} = Vd \quad (1.14)$$

Finally, consider the second sum:

$$\sum_{\langle x,y \rangle} \sigma_x + \sigma_y$$

⁴Due to the division by $N!$, configurations where only N cells are occupied are considered distinguishable only if the occupied cells are different. In other words, permuting the position of two particles is not counted.

Practically, to compute it we may inspect each edge (x, y) , and sum the values σ_x and σ_y of the *spins* at its extrema. When we are done, each value σ_x at a node will have been considered exactly q times - one for every neighbour of x :

$$\sum_{\langle x, y \rangle} \sigma_x + \sigma_y = \sum_x \sigma_x \text{dg}(x) \quad (1.15)$$

where $\text{dg}(x)$ is the **degree** of node x , i.e. the number of connections (edges) involving cell x . In our case, $\text{dg}(x) = q = 2d$:

$$\sum_{\langle x, y \rangle} \sigma_x + \sigma_y = 2d \sum_x \sigma_x$$

This relation will allow to highlight the common factor $\sum_x \sigma_x$.

Substituting (??), (??) and (??) back in (??) leads to:

$$\begin{aligned} & \beta \frac{\epsilon_0}{4} \left[Vd + 2d \sum_x \sigma_x + \sum_{\langle x, y \rangle} \sigma_x \sigma_y \right] + \frac{\ln z}{2} \left(V + \sum_x \sigma_x \right) = \\ & = \left[\underbrace{\beta \frac{\epsilon_0}{4} \sum_{\langle x, y \rangle} \sigma_x \sigma_y}_J + \underbrace{\left[\frac{\beta \epsilon_0 d}{2} + \frac{\ln z}{2} \right] \sum_x \sigma_x}_{\beta b} \right] + V \left(\beta \frac{\epsilon_0}{4} d + \frac{\ln z}{2} \right) \end{aligned} \quad (1.16)$$

The term in the first set of square parentheses is the only one depending on σ , and thus the one capturing the *essence* of the Ising Model. The last one is just a scaling factor given by the current application (lattice gas).

For simplicity, let's define:

$$J \equiv \frac{\epsilon_0}{4}; \quad \beta b \equiv \frac{\beta \epsilon_0 d}{2} + \frac{\ln z}{2} \quad (1.17)$$

In this way we may collect a β in (??):

$$\beta \left[J \sum_{\langle x, y \rangle} \sigma_x \sigma_y + b \sum_x \sigma_x \right] + \beta V \left(Jd + \frac{\ln z}{2\beta} \right)$$

Substituting back in (??):

$$Z_{\text{g.c.}} \equiv e^{\beta PV} = \underbrace{\left(\sum_{\{\sigma\}} \exp \left[\beta \left(J \sum_{\langle x, y \rangle} \sigma_x \sigma_y + b \sum_x \sigma_x \right) \right] \right)}_Z \exp \left(\beta V \left[Jd + \frac{\ln z}{2\beta} \right] \right) \quad (1.18)$$

We define the Ising Model partition function Z to contain the only relevant factor:

$$Z_{\text{Ising}} \equiv \sum_{\{\sigma\}} \exp \left[\beta \left(J \sum_{\langle x, y \rangle} \sigma_x \sigma_y + b \sum_x \sigma_x \right) \right] \equiv e^{-\beta V f(T, b)} \quad (1.19)$$

where $f(T, b)$ (defined by the above relation) is the model's **free energy density**. Then, taking the logarithm of both terms in (??):

$$\beta P \mathcal{V} \stackrel{(\text{??})}{=} -\beta \mathcal{V} f(T, b) + \beta \mathcal{V} J d + \beta \mathcal{V} \frac{\ln z}{2\beta} \quad (1.20)$$

Rearranging the definition of b (??):

$$b = \frac{\epsilon_0 d}{2} + \frac{\ln z}{2\beta} \Rightarrow \frac{\ln z}{2\beta} = b - \frac{\epsilon_0 d}{2}$$

and substituting in (??) we get the grand-canonical pressure of the lattice gas:

$$\begin{aligned} P_{\text{g.c.}} &= -f(T, b) + \frac{\epsilon_0}{4} d + b - \frac{\epsilon_0 d}{2} = -f(T, b) - \frac{\epsilon_0}{4} d + b = \\ &= -f(T, b) + b - Jd \end{aligned} \quad (1.21)$$

The physical interpretation of J and b depends on which application we are considering. In the lattice gas, J is clearly proportional to the strength of interaction between gas molecules occupying neighbouring cells, while b depends on the chemical potential μ (through z), the temperature and the system's dimensionality.

A clearer meaning for b is found when the Ising Model is applied to ferromagnetism. In this case, the $\{\sigma_i\}$ represent the direction (up or down) of the particle's *spins*, i.e. their *intrinsic* magnetic momenta (of pure quantum origin). Then $b \sum_x \sigma_x$ measures the correlation between the *spin directions* and b . In particular, configurations for which the majority of σ_i have the same sign of b (i.e. are “parallel” to b) have a higher probability (as each term in the sum over states in Z is the probability associated with a particular configuration σ). So, b can be interpreted as an **external magnetic field**, pushing each particle to align its spin to it.

On the other hand, for a quantum system (such as a ferromagnet), J is called the **exchange energy**, and measures the overlap of electronic clouds of neighbouring atoms.

From (??) we can extract the Hamiltonian for the Ising Model:

$$Z_{\text{Ising}} \equiv \sum_{\{\sigma\}} e^{-\beta \mathcal{H}(\sigma)} \Rightarrow \mathcal{H}(\sigma) = -J \sum_{\langle x, y \rangle} \sigma_x \sigma_y - b \sum_x \sigma_x \quad (1.22)$$

$\beta \mathcal{H}(\sigma)$ is called the **reduced Hamiltonian**:

$$\begin{aligned} \beta \mathcal{H}(\sigma) &= - \underbrace{\beta J}_{K} \sum_{\langle x, y \rangle} \sigma_x \sigma_y - \underbrace{\beta b}_h \sum_x \sigma_x \\ &= -K \sum_{\langle x, y \rangle} \sigma_x \sigma_y - h \sum_x \sigma_x \end{aligned}$$

Intuitively, $\mathcal{H}(\sigma)$ is the sum of two interactions:

- **Spin-spin interactions:** $-J \sum_{\langle x, y \rangle} \sigma_x \sigma_y$. If $J > 0$, the term is minimized if $\sigma_x \sigma_y > 0$, i.e. if neighbouring spins are all *parallel* to each other (**ferromagnetic order**). If $J < 0$, spins tend instead to be *anti-parallel* to each other (**anti-ferromagnetic order**), as can be seen in fig. ??.

- **Spin-field interaction:** $-b \sum_x \sigma_x$, which is minimized if spins are aligned (parallel) to the magnetic field b (when applying the model to a ferromagnet).

In particular, if $h = 0$ (and so $b = 0$), then there is no “preferred direction” for the alignment of spins. This results in two equivalent *ground states*, as $H(\boldsymbol{\sigma}) = H(-\boldsymbol{\sigma})$. For instance, if $J > 0$, the two possibilities are all $\sigma_i = +1$ (up), or all $\sigma_i = -1$ (down).

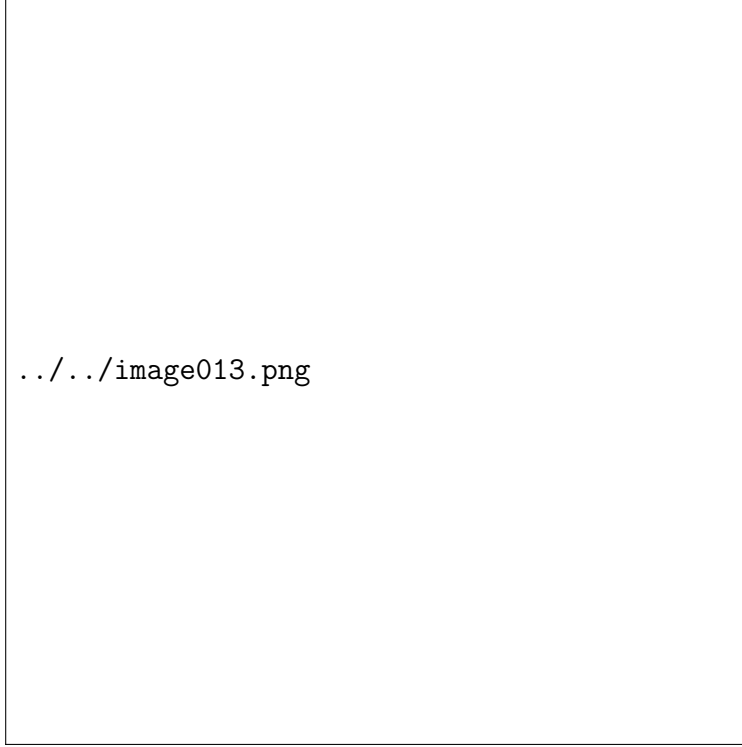


Figure (1.5) – The spin-spin interaction energy when all neighbouring spins are ordered in a (anti)parallel way, depending on the sign of J .

1.2 Ising Partition Function

Our next goal is to compute explicitly the Ising partition function (??):

$$Z(K, h) = \sum_{\{\boldsymbol{\sigma}\}} \exp \left(K \sum_{\langle x, y \rangle} \sigma_x \sigma_y + h \sum_x \sigma_x \right) = e^{-\beta V f(K, h)} \quad \begin{matrix} K = \beta J \\ h = \beta b \end{matrix} \quad (1.23)$$

Note that the sum is over $2^{N_{\text{cells}}} = 2^V$ (as the lattice step a is set to 1) possible states (spin configurations).

We define the **magnetization** m as the average *alignment* of spins:

Magnetization

$$\begin{aligned} m &\equiv \frac{1}{V} \langle \sum_x \sigma_x \rangle = \frac{1}{V} \frac{1}{Z} \sum_{\{\boldsymbol{\sigma}\}} \exp \left(K \sum_{\langle x, y \rangle} \sigma_x \sigma_y + h \sum_x \sigma_x \right) \sum_x \sigma_x = \\ &= \frac{1}{V} \frac{\partial}{\partial h} \ln Z(K, h) \underset{(\text{??})}{=} -\beta \frac{\partial}{\partial h} f(K, h) \end{aligned} \quad (1.24)$$

The average number of particles $\langle N \rangle$ in a grand-canonical ensemble is given by:

Average number of particles $\langle N \rangle$

$$\langle N \rangle_{\text{g.c.}} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z_{\text{g.c.}} = \frac{1}{\beta} \frac{\partial \cancel{z}}{\partial \mu} \frac{\partial}{\partial \cancel{z}} \ln Z_{\text{g.c.}} \quad (1.25)$$

with:

$$\frac{\partial z}{\partial \mu} = \frac{\partial}{\partial \mu} \frac{e^{\beta \mu}}{\lambda^3} = \beta \underbrace{\frac{e^{\beta \mu}}{\lambda^3}}_z = \beta z$$

and so:

$$\langle N \rangle_{\text{g.c.}} = \frac{1}{\beta} \beta z \frac{\partial}{\partial z} \ln Z_{\text{g.c.}} = z \frac{\partial}{\partial z} \ln Z_{\text{g.c.}} \quad (1.26)$$

Let v be the average volume per particle, i.e. $\langle N \rangle / V$. Then the average **density** of particles is v^{-1} :

Particle density

$$\begin{aligned} v^{-1} &\equiv \frac{\langle N \rangle}{V} \stackrel{(\text{??})}{=} \frac{z}{V} \frac{\partial}{\partial z} \ln Z_{\text{g.c.}} \stackrel{(\text{??})}{=} \frac{z}{V} \frac{\partial}{\partial z} \beta P V = \\ &\stackrel{(\text{??})}{=} \frac{z}{V} \frac{\partial}{\partial z} [-\beta V f(T, b(z)) + \beta V b(z) - \beta V J d] = \\ &\stackrel{(\text{??})}{=} \frac{z}{V} \frac{\partial}{\partial z} [\ln Z_{\text{Ising}} + \beta V b] \end{aligned} \quad (1.27)$$

as the third term does not depend on z . It is useful to change variables from $z \rightarrow \ln z$, as b is a function of $\ln z$:

$$z \frac{\partial}{\partial z} = z \frac{\partial(\ln z)}{\partial z} \frac{\partial}{\partial(\ln z)} = \cancel{z} \frac{1}{\cancel{z}} \frac{\partial}{\partial(\ln z)} = \frac{\partial}{\partial(\ln z)}$$

So that:

$$\frac{z}{V} \frac{\partial}{\partial z} \beta V b = \beta \frac{\partial b}{\partial(\ln z)} \stackrel{(\text{??})}{=} \beta \frac{1}{2\beta} = \frac{1}{2} \quad (1.28)$$

With another change of variable $z \rightarrow h$ we can rewrite $z \partial_z \ln Z_{\text{Ising}}$ as a function of the magnetization m :

$$\frac{z}{V} \frac{\partial}{\partial z} \ln Z_{\text{Ising}} = \frac{\partial \cancel{h}}{\partial(\ln z)} \underbrace{\frac{1}{V} \frac{\partial}{\partial \cancel{h}} \ln Z_{\text{Ising}}}_m = \left(\frac{\partial(\ln z)}{\partial(h)} \right)^{-1} m \stackrel{(a)}{=} \frac{m}{2} \quad (1.29)$$

where the derivative in (a) can be computed by isolating $\ln z$ in the definition of b (??):

$$b = \frac{\epsilon_0 d}{2} + \frac{\ln z}{2\beta} \Rightarrow 2\beta b = \underbrace{\epsilon_0}_{4J} d\beta + \ln z \Rightarrow \ln z = 2\beta(b - 2Jd) = 2(h - 2Kd) \Rightarrow \frac{\partial(\ln z)}{\partial h} = 2$$

Substituting (??) and (??) back in (??) leads to:

Particle density and magnetization

$$v^{-1} = \frac{m}{2} + \frac{1}{2} = \frac{m+1}{2} \quad (1.30)$$

So a higher m corresponds (in the lattice gas model) to a higher particle density. Intuitively, a high m means that, *on average*, cells are more occupied - meaning that particles will be, in general, closer together. Note that $m \in [-1, +1]$ (which happens, respectively, when all spins point *up* or *down*, i.e. if all cells are occupied or empty), and so $v^{-1} \in [0, 1]$ as expected.

Assuming periodic boundary conditions, the system is **translational invariant**: there is no “preferred” position in the lattice - all cells are exactly equal to each other. This means that $\langle \sigma_x \rangle \equiv \bar{\sigma}$ must be independent of x , and so it is equal to the magnetization m :

Consequences of translational invariance

$$m = \frac{1}{V} \langle \sum_x \sigma_x \rangle = \frac{1}{V} \sum_x \langle \sigma_x \rangle = \frac{1}{V} V \bar{\sigma} = \bar{\sigma} = \langle \sigma_x \rangle$$

On the other hand, the 2-point correlation function $\langle \sigma_x \sigma_y \rangle$ must depend only on the distance $\|\mathbf{r}_x - \mathbf{r}_y\|$ between the two cells, because of translational invariance.

Note that for any finite V , the sum in (??) is a sum over a finite number 2^V of states. As each term is an analytic function, Z is also an analytic function, and so it is $\ln Z = -\beta V f(K, h)$, and in particular the free energy $f(K, h)$. However, we expect phase-transitions to correspond to points at which the free energy is non-analytic - which would explain the *sudden* changes in the system’s properties that are experimentally observed during such a transition.

So, in any finite lattice we won’t be able to see any phase-transition. Conversely, to observe a phase-transition, an **infinite lattice** is required, which is obtained when $V \rightarrow \infty$, i.e. in the **thermodynamic limit**.

Example 1 (Non-interacting spins):

Let’s examine the simplest possible case in the Ising Model, occurring when $J = 0$ (or, equivalently, $K = 0$). The Hamiltonian becomes:

$$\mathcal{H}(\boldsymbol{\sigma}) = -h \sum_x \sigma_x \quad (1.31)$$

meaning that spins (or cells) are completely independent from each other (**decoupled**). In fact, the partition function factorizes:

$$Z = \sum_{\{\boldsymbol{\sigma}\}} \exp\left(h \sum_x \sigma_x\right) = \sum_{\sigma_1=\pm 1} e^{h\sigma_1} \dots \sum_{\sigma_V=\pm 1} e^{h\sigma_V}$$

Noting that:

$$\sum_{\sigma_i=\pm 1} e^{h\sigma_i} = 2 \frac{e^h + e^{-h}}{2} = 2 \cosh h$$

the partition function becomes:

$$Z = (2 \cosh h)^V \quad (1.32)$$

Then the **free energy** f is:

$$Z = e^{-\beta V f(h)} \Rightarrow f(h) = -\frac{\ln Z}{\beta V} \stackrel{??}{=} -\frac{\ln(2 \cosh h)}{\beta} \quad (1.33)$$

The **magnetization**:

$$m(h) \stackrel{(\text{??})}{=} -\beta \frac{\partial}{\partial h} f(h) = \frac{\partial}{\partial h} \ln(\cosh h) = \frac{\sinh h}{\cosh h} = \frac{e^h - e^{-h}}{e^h + e^{-h}} = \tanh h \quad (1.34)$$

A plot of $m(h)$ can be seen in fig. ??.

Inverting (??) allows to express h as a function of m :

$$h = \beta b = \tanh^{-1} m$$

This can be solved by letting $t = e^h$, and so:

$$\begin{aligned} m = \tanh h &= \frac{t + 1/t}{t - 1/t} = \frac{t^2 - 1}{t^2 + 1} \Rightarrow m(t^2 + 1) = t^2 - 1 \\ \Rightarrow t^2(m + 1) &= -(m - 1) \Rightarrow t = \pm \sqrt{\frac{1 + m}{1 - m}} \end{aligned}$$

As $t = e^h > 0$, only the positive solution is acceptable, leading to:

$$e^h = \sqrt{\frac{1 + m}{1 - m}} \Rightarrow h = \ln \sqrt{\frac{1 + m}{1 - m}} = \frac{1}{2} \ln \frac{1 + m}{1 - m} \quad (1.35)$$

Substituting back:

$$h = \beta b = \tanh^{-1} m = \frac{1}{2} \ln \frac{1 + m}{1 - m} \quad -1 < m < +1 \quad (1.36)$$

We can also express the free energy (??) as function of m . First note that:

$$\begin{aligned} \ln \cosh h &= \frac{1}{2} \ln(\cosh h)^2 = \frac{1}{2} \ln \underbrace{\frac{\cosh^2 h}{\cosh^2 h - \sinh^2 h}}_1 = \\ &= -\frac{1}{2} \ln \frac{\cosh^2 h - \sinh^2 h}{\cosh^2 h} = -\frac{1}{2} \ln(1 - \tanh^2 h) = \\ &\stackrel{(\text{??})}{=} -\frac{1}{2} \ln(1 - m^2) \end{aligned} \quad (1.37)$$

and so:

$$\begin{aligned} f &\stackrel{(\text{??})}{=} -\frac{\ln(2 \cosh h)}{\beta} \Rightarrow \beta f = -\ln(\cosh h) - \ln 2 = \\ &\stackrel{(\text{??})}{=} \frac{1}{2} \ln(1 - m^2) - \frac{1}{2} \ln 2^2 = \frac{1}{2} \ln \frac{1 - m^2(h)}{4} \end{aligned} \quad (1.38)$$

Note that the free energy is an **even** function of the magnetization, whereas m is an **odd** function of h .

Finally, from (??) note that the derivative of f with respect to b is $-m$:

$$\frac{\partial f}{\partial b} = \frac{\partial^2 f}{\partial h \partial \frac{\partial h}{\partial b}} \stackrel{h=\beta b}{=} -\frac{m}{\beta} \stackrel{(\text{??})}{=} -m$$

So, the **(non-standard) Legendre transform** of $f(b)$ with respect to b is $\gamma(m)$ defined by the following relation:

$$-\gamma(m) + f(b(m)) = \frac{\partial f}{\partial b} \cdot b(m) = -mb(m) \Rightarrow \gamma(m) = f + bm \quad (1.39)$$

$b(m)$ is obtained rearranging (??):

$$\frac{h}{\beta} = b = \frac{1}{2\beta} \ln \frac{1+m}{1-m} \quad (1.40)$$

Then substituting (??) and (??) in (??) leads to:

$$\begin{aligned} \gamma(m) &= \frac{1}{2\beta} \ln \frac{1-m^2}{4} + \frac{m}{2\beta} \ln \frac{1+m}{1-m} = \\ &= \frac{1}{2\beta} \left[\ln \frac{1-m}{2} \ln \frac{1+m}{2} + m \left(\ln \frac{1+m}{2} + \cancel{\ln 2} - \ln \frac{1-m}{2} - \cancel{\ln 2} \right) \right] = \\ &= \frac{1}{2\beta} \left[\ln \frac{1-m}{2} + \ln \frac{1+m}{2} + m \left(\ln \frac{1+m}{2} - \ln \frac{1-m}{2} \right) \right] = \\ &= \frac{1}{2\beta} \left[(1+m) \ln \frac{1+m}{2} + (1-m) \ln \frac{1-m}{2} \right] = \\ &= \left[\frac{1-m}{2} \ln \frac{1-m}{2} + \frac{1+m}{2} \ln \frac{1+m}{2} \right] \beta^{-1} \end{aligned}$$

Note that, due to the properties of the Legendre transform, we have:

$$\frac{\partial \gamma(m)}{\partial m} = b$$

Then, the **entropy density** is obtained by differentiating f (or equivalently its Legendre transform γ , as T is not involved in the transformation):

$$s = -\frac{\partial f}{\partial T} = -\frac{\partial \gamma}{\partial T} = -k_B \left[\frac{1-m}{2} \ln \frac{1-m}{2} + \frac{1+m}{2} \ln \frac{1+m}{2} \right] \quad (1.41)$$

This can be derived also from the definition of the Shannon Entropy. Recall that the probability of a certain spin configuration σ is:

$$\rho(\sigma) = \frac{1}{Z} e^{-\beta \mathcal{H}(\sigma)}$$

As spins are decoupled, $\rho(\sigma)$ factorizes:

$$\rho(\sigma) \stackrel{(\text{??})}{=} \prod_x \frac{e^{h\sigma_x}}{2 \cosh h} \equiv \prod_x \rho_1(\sigma_x) \quad \rho_1(\sigma_x) \equiv \frac{e^{h\sigma_x}}{2 \cosh h} \quad (1.42)$$

Any generic function $g(\sigma)$ of a binary variable $\sigma \in \{\pm 1\}$ can be written as:

$$g(\sigma) = \frac{g(+1) + g(-1)}{2} + \sigma \frac{g(+1) - g(-1)}{2}$$

In fact:

$$g(+1) = \frac{g(+1) + \cancel{g(-1)}}{2} + \frac{g(+1) - \cancel{g(-1)}}{2} = g(+1)$$

$$g(-1) = \frac{g(+1) + g(-1)}{2} - \frac{g(+1) - g(-1)}{2} = g(-1)$$

In particular, if $g(\sigma) = \exp(h\sigma)$:

$$e^{h\sigma} = \frac{e^h + e^{-h}}{2} + \sigma \frac{e^h - e^{-h}}{2} = \cosh h + \sigma \sinh h \quad (1.43)$$

And so:

$$\rho_1(\sigma) \stackrel{(\text{??})}{=} \frac{e^{h\sigma}}{2 \cosh h} \stackrel{(\text{??})}{=} \frac{\cosh h + \sigma \sinh h}{2 \cosh h} = \frac{1 + \sigma \tanh h}{2} \stackrel{(\text{??})}{=} \frac{1 + \sigma m}{2} \quad (1.44)$$

The ρ_1 so defined is already normalized:

$$\sum_{\sigma=\pm 1} \rho_1(\sigma) = \frac{1+m}{2} + \frac{1-m}{2} = 1 \quad (1.45)$$

We are now ready to compute the **information entropy**:

$$\begin{aligned} S_I[\rho] &= -k_B \sum_{\{\sigma\}} \rho(\sigma) \ln \rho(\sigma) \stackrel{(\text{??})}{=} -k_B \sum_{\{\sigma\}} \left(\prod_{x=1}^V \rho_1(\sigma_x) \right) \ln \left(\prod_{y=1}^V \rho_1(\sigma_y) \right) = \\ &= -k_B \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_V=\pm 1} \prod_{x=1}^V \rho_1(\sigma_x) \sum_{y=1}^V \ln \rho_1(\sigma_y) = \\ &\stackrel{(a)}{=} -k_B \sum_{y=1}^V \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_V=\pm 1} \left(\prod_{x \neq y} \rho_1(\sigma_x) \right) \rho_1(\sigma_y) \ln \rho_1(\sigma_y) = \\ &\stackrel{(b)}{=} -k_B \sum_{y=1}^V \sum_{\sigma_y=\pm 1} \rho_1(\sigma_y) \ln \rho_1(\sigma_y) \underbrace{\prod_{x \neq y} \sum_{\sigma_x=\pm 1} \rho_1(\sigma_x)}_1 = \\ &\stackrel{(c)}{=} -k_B \sum_{y=1}^V \sum_{\sigma_y=\pm 1} \rho_1(\sigma_y) \ln \rho_1(\sigma_y) = -k_B V \sum_{\sigma_y=\pm 1} \rho_1(\sigma_y) \ln \rho_1(\sigma_y) \end{aligned} \quad (1.46)$$

In (a) we exchange the sum over cells y with the one over states σ . Then, we split the product over x (highlighted in blue) in two factors: one with $x \neq y$, and the other with $x = y$. Then, in (b) we exchange the order of the sums over cell states, bringing the one over σ_y first, and factoring out everything that depends only on σ_y . The remaining term is a product of ρ_1 , and thanks to the spin-independence we can bring all the other sum over states inside it, and then apply normalization (??) to reach the result in (c). Due to translational invariance, the inner sum evaluates to a constant, and so $\sum_{y=1}^V$ amounts merely to multiplying by the number of cells V .

Finally, substituting (??) in the last step (??) and dividing by V leads back to (??):

$$s = \frac{S_I[\rho]}{V} = -k_B \left[\frac{1+m}{2} \ln \frac{1+m}{2} + \frac{1-m}{2} \ln \frac{1-m}{2} \right] \quad (1.47)$$

The grand-canonical pressure is given by:

$$\begin{aligned}
P &\stackrel{(\text{??})}{=} [-f(T, b) + b - Jd] \Big|_{J=0} \stackrel{(\text{??})}{=} -\frac{1}{2\beta} \ln \frac{1-m^2}{4} + \frac{1}{2\beta} \ln \frac{1+m}{1-m} = \\
&= \frac{1}{2\beta} \left[-\ln \frac{1-m}{2} - \ln \frac{1+m}{2} + \ln \frac{1+m}{2} - \ln \frac{1-m}{2} \right] = -\beta^{-1} \ln \frac{1-m}{2} = \\
&= -\beta^{-1} \ln \frac{2-1-m}{2} = -\beta^{-1} \ln \left(1 - \frac{m+1}{2} \right) \stackrel{(\text{??})}{=} -\beta^{-1} \ln(1-v^{-1}) = \\
&= -\beta^{-1} \ln \left(1 - \frac{\langle N \rangle}{V} \right) \xrightarrow{v^{-1} \rightarrow 0} -\beta^{-1} \left(-\frac{\langle N \rangle}{V} \right) = k_B T \frac{\langle N \rangle}{V}
\end{aligned}$$

So, in the limit of *low density* $v^{-1} \rightarrow 0$, i.e. of a rarefied gas, we get the equation of state for an ideal gas in the grand-canonical ensemble.

However, if we consider the full equation, we note a singularity for $v^{-1} \rightarrow 1$ (densest case, where all cells are filled), corresponding to the *liquid phase*. A plot of $P(v)$ is shown in fig. ??, and compared with the one of a real gas. In particular, no *phase-transition* is captured by such *free* Ising Model: as we will see, spin-spin interactions are fundamental. However, even with $J = 0$, the model is able to describe two phases: that of a rarefied gas and of a liquid.

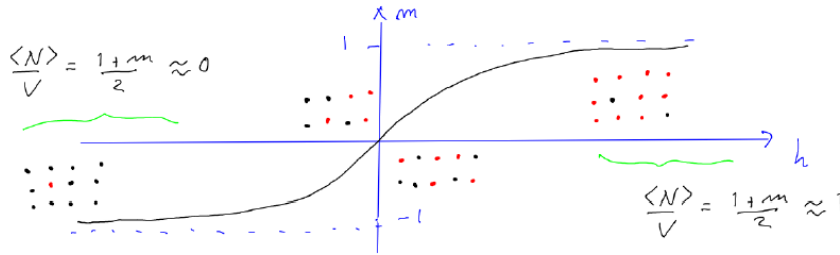


Figure (1.6) – Magnetization m as function of $h = \beta b$ for the Ising Model with decoupled spins ($J = 0$). For $h \rightarrow \pm\infty$, $m = \tanh h \rightarrow \pm 1$. For the negative magnetization, the lattice is almost “empty”, while for $m \rightarrow +1$ it is almost full.

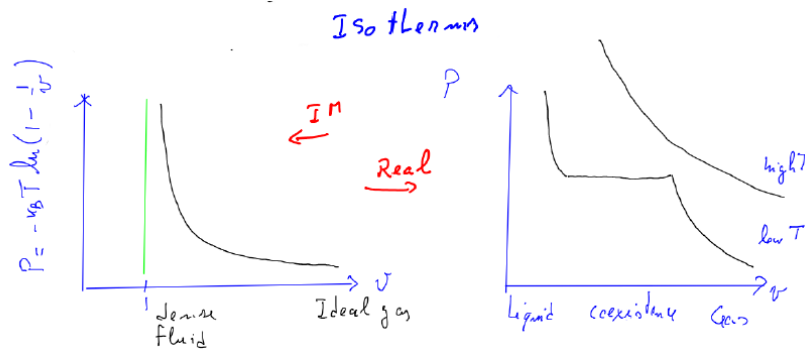


Figure (1.7) – Isothermal curves $P(v)$ for the lattice gas (left) and a real gas (right), where v is the average volume per particle (the reciprocal of the density v^{-1}). In the Ising Model pressure *diverges* when density approaches its maximum $v^{-1} \rightarrow 1$. However, in the real case there is a range of temperatures at which, for a range of values of v , the liquid and gas phases are coexisting (which is where the phase transition is happening). So the model is able to capture *some* of the behaviour (the dilute “ideal gas” state and the *liquid* phase), but not all.

Exercise 1.2.1 (2-point correlation):

Show that:

$$\langle \sigma_x \sigma_y \rangle = (\tanh h)^2 = \langle \sigma_x \rangle \langle \sigma_y \rangle$$