# Lecture-17: Interacting particle systems

## 1 Ferromagnets and Ising models

Magnetic materials contain molecules with individual magnetic moments that tend to align with the external magnetic field. Magnetic moments of different molecules interact with each other. In many materials, the energy is lower when the moments align. A simple mathematical model for considering a number of particles with interacting moments is the Ising model, which describes the magnetic moments by Ising spins localized at the vertices of a certain region of a $d$-dimensional cubic lattice $\mathbb{L}$. The cubic lattice $\mathbb{L} = (V, F)$ is determined by the set of vertices $V \triangleq [L]^d$ and the edges between the nearest neighbors defined as

$$F \triangleq (\{i,j\} \in [L]^d \times [L]^d : \sum_{k=1}^{d} |i_k - j_k| = 1).$$

### 1.1 Energy function

Ising spins of particles at lattice points is denoted by $\sigma \in \mathcal{X} \triangleq \mathcal{Z}^{[L]^d}$ where the configuration of a particle at each coordinate $i \in [L]^d$ is an Ising spin $\sigma_i \in \mathcal{Z} \triangleq \{-1, 1\}$. We denote the number of particles $N \triangleq L^d$ and $\mathcal{X}$ is the space of configuration. The energy of an $N$ particle configuration $\sigma$ is given by

$$E(\sigma) = -\sum_{\{i,j\} \in F} \sigma_i \sigma_j - B \sum_{i \in [L]^d} \sigma_i, \tag{1}$$

where the sum over $(i, j)$ runs over all the unordered pairs of sites $i, j \in [L]^d$ which are nearest neighbors and $B$ is the applied external magnetic field. Determining the free energy density $f(\beta)$ in the thermodynamic limit for this model is a non-trivial task. In 1924, Ernst Ising solved the $d = 1$ case and showed the absence of phase transitions. In 1948, Lars Onsager solved the $d = 2$ case, exhibiting the first soluble *finite-dimensional* model with a second-order phase transition. The problem is unsolved in higher dimensions, although many important features of the solution are well understood.

### 1.2 Temperature limits

The two limiting cases that can be considered are at high and low temperatures.

**High temperature limit,** $\beta \to 0$. The energy no longer matters and the Boltzmann distribution is uniform over all configurations $\sigma \in \mathcal{Z}^N$. That is, $\mu_\beta(\sigma) = \frac{1}{2^N}$ for all $\sigma \in \mathcal{Z}^N$.

**Low temperature limit,** $\beta \to \infty$**.** The Boltzmann distribution concentrates onto the ground state(s). In the absence of external magnetic field, i.e $B = 0$, the two degenerate ground states are given by,

$$\sigma^+ = (\sigma_i = 1 : i \in [L]^d), \qquad \sigma^- = (\sigma_i = -1 : i \in [L]^d).$$

If the magnetic field is set to some non-zero value, one of the two configuration dominates. The ground state is $\sigma^+$ if $B > 0$ and the ground state is $\sigma^-$ if $B < 0$.

### 1.3 Important observables for ferromagnets

**Definition 1.1. Expected spin** with respect to the Boltzmann distribution for any particle $i \in [L]^d$, is defined as $\langle \sigma_i \rangle \triangleq \sum_{\sigma \in \mathcal{X}} \mu_\beta(\sigma) \sigma_i$.

**Definition 1.2.** The extent of alignment in a region due to an external magnetic field $B$ is given by **average magnetization**, defined as $M_N(\beta, B) \triangleq \frac{1}{N}\sum_{i\in[L]^d} \langle\sigma_i\rangle$.

**Lemma 1.3.** *Average magnetization $M_N(\beta, B)$ is an odd function of $B$, and hence $M_N(\beta, 0) = 0$.*

*Proof.* Recall that $\mu_N(\beta, B) = \frac{1}{Z_N(\beta,B)} e^{-\beta E(\sigma,B)}$ where $-\beta E(\sigma, B) = \beta \sum_{i,j\in F} \sigma_i\sigma_j + \beta B \sum_{i\in[L]^d} \sigma_i$. We observe that $-\beta E(\sigma, B) = -\beta E(-\sigma, -B)$. Thus, we can write

$$Z_N(\beta, B) = \sum_\sigma e^{-\beta E(\sigma,B)} = \sum_\sigma e^{-\beta E(-\sigma,-B)} = \sum_{-\sigma} e^{-\beta E(\sigma,-B)} = Z_N(\beta, -B).$$

Further, we can write the numerator as $\sum_\sigma \sigma_i e^{-\beta E(\sigma,B)} = \sum_\sigma \sigma_i e^{-\beta E(-\sigma,-B)} = -\sum_{-\sigma} \sigma_i e^{-\beta E(\sigma,-B)}$. Combining the two results, we obtain

$$M_N(\beta, B) = \frac{1}{Z_N(\beta, B)} \sum_\sigma \sigma_i e^{-\beta E(\sigma,B)} = -\frac{1}{Z_N(\beta, -B)} \sum_{-\sigma} \sigma_i e^{-\beta E(\sigma,-B)} = M_N(\beta, -B).$$

□

**Definition 1.4.** The **spontaneous magnetization** is defined as the large particle and low positive magnetic field limit of average magnetization, i.e. $M_+(\beta) \triangleq \lim_{B\downarrow 0} \lim_{N\to\infty} M_N(\beta, B)$.

*Remark* 1. We observe the following properties hold for spontaneous magnetization.

1. At high temperatures when $\beta \to 0$, the alignment of spins are random, and hence $M_+(0) = 0$. We observe that this is true for all $d$ and is referred to as the *paramagnetic phase*.
2. At low temperatures when $\beta \to \infty$ and any positive magnetic field $B \downarrow 0$, the alignment of spins is $\sigma^+$, i.e. $\langle\sigma_i\rangle = 1$ for all $i \in [L]^d$ and hence $M_+(\infty) = 1$.
3. Phase transition for the system occurs at the critical temperature $T_c \triangleq 1/\beta_c$ which depends on the number of dimensions $d$. We have $M_+(\beta) = 0$ for all $\beta < \beta_c$ and $M_+(\beta) > 0$ for all $\beta > \beta_c(d)$.
4. In one-dimensional systems ($d = 1$), a phase transition occurs at $T_c = 0$.
5. For the number of dimensions $d \geqslant 2$, the critical temperature is non-zero.

## 1.4 Rescaled magnetic field

To analyze the limiting behavior upon the application of a magnetic field, we define a *rescaled* magnetic field $x = \beta B$, with inverse temperature $\beta \to 0$ or $\beta \to \infty$, keeping $x$ fixed. With this, we will subsequently study some of the qualitative properties of the resultant model. We can write the scaled energy function for the ferromagnetic Ising model as

$$-\beta E(\sigma) = \beta \sum_{\{i,j\}\in F} \sigma_i\sigma_j + x \sum_{i\in[L]^d} \sigma_i.$$

**High temperature limit** $\beta \to 0$. We have $\lim_{\beta\to 0} -\beta E(\sigma) = x \sum_i \sigma_i$, corresponding to non-interacting systems, and hence

$$\mu_\beta(\sigma) = \prod_i \mu_\beta(\sigma_i), \text{ where } \mu_\beta(\sigma_i) = \frac{e^{x\sigma_i}}{e^x + e^{-x}}.$$

Therefore, we can write $\langle\sigma_i\rangle = \sum_{\sigma_i=\pm 1} \sigma_i \mu_\beta(\sigma_i) = \tanh(x)$.

**Low temperature limit** $\beta \to \infty$. The ground state is given by $\sigma_i = \text{sign}(x)$ for all $i \in [L]^d$. Therefore, we can write the scaled energy function as $-\beta E(x) = \beta |F| + xN$. Since $|F| \approx Nd$, we get $-\beta E(x) \approx N(\beta d + |x|)$. We can write the Boltzmann distribution as $\mu_\beta(\sigma) = \frac{e^{N|x|}}{e^{-Nx}+e^{Nx}}$. It follows that

$$\langle\sigma_i\rangle = \frac{e^{Nx} - e^{-Nx}}{e^{Nx} + e^{-Nx}} = \tanh(Nx). \tag{2}$$

## 1.5 The one-dimensional case

Consider a one-dimensional system where $d = 1$, and hence has $L$ spins with energy $E(\sigma) \triangleq -\sum_{i=1}^{L-1} \sigma_i \sigma_{i+1} - B \sum_{i=1}^{L} \sigma_i$.

**Definition 1.5.** The **partial partition function** where the configurations of all spins $(\sigma_1, \ldots, \sigma_p)$ have been summed over and $\sigma_{p+1}$ is fixed, is defined as

$$z_p(\beta, B, \sigma_{p+1}) \triangleq \sum_{\sigma_1,\ldots,\sigma_p} \exp\left[\beta \sum_{i=1}^{p} \sigma_i \sigma_{i+1} + \beta B \sum_{i=1}^{p} \sigma_i\right].$$

We define row vector $z_p(\beta, B) \triangleq \begin{bmatrix} z_p(\beta, B, -1) & z_p(\beta, B, 1) \end{bmatrix}$ and $2 \times 2$ transfer matrix $T(\sigma_p, \sigma_{p+1}) \triangleq e^{\beta \sigma_p \sigma_{p+1} + \beta B \sigma_p}$ for all $\sigma_p, \sigma_{p+1} \in \mathcal{Z} = \{-1, 1\}$, such that

$$T = \begin{bmatrix} e^{\beta - \beta B} & e^{-\beta - \beta B} \\ e^{-\beta + \beta B} & e^{\beta + \beta B} \end{bmatrix}.$$

*Remark* 2. The eigenvalues of the transfer matrix $T$ are $\lambda_{1,2} \triangleq e^{\beta} \cosh(\beta B) \pm \sqrt{e^{2\beta} \sinh^2(\beta B) + e^{-2\beta}}$. For an eigenvalue $\lambda_i$ the associated left and right eigenvectors are denoted by $\psi_i$ and $\phi_i^T$ respectively, where $i \in \{1, 2\}$. We observe that

$$\psi_{1,2} e^{-\beta(1-B)} = -\psi_{1,1}(e^{\beta(1-B)} - \lambda_1), \qquad \phi_{1,2} e^{-\beta(1+B)} = -\phi_{1,1}(e^{\beta(1-B)} - \lambda_1). \tag{3}$$

**Proposition 1.6.** *The partition function for one-dimensional ferromagnetic model with $L$ spins is given by*

$$Z_L(\beta, B) = u_1 \lambda_1^{L-1} \langle \psi_1, \psi_R \rangle + u_2 \lambda_2^{L-1} \langle \psi_2, \psi_R \rangle. \tag{4}$$

*Proof.* In terms of matrix $T$ and row vector $z_p(\beta, B)$, we write the following recursive relation for all $p \in [L-1]$ and $\sigma_{p+1} \in \mathcal{Z}$ as

$$z_p(\beta, B, \sigma_{p+1}) = \sum_{\sigma_p \in \mathcal{Z}} z_{p-1}(\beta, B, \sigma_p) T(\sigma_p, \sigma_{p+1}) = (z_{p-1}(\beta, B) T)_{\sigma_{p+1}}.$$

We can rewrite this recursion as $z_p(\beta, B) = z_{p-1}(\beta, B) T$ for $p \in [L-1]$. Using the definition of $z_{L-1}(\beta, B)$ and $Z_L(\beta, B)$, we can write the partition function as

$$Z_L(\beta, B) = \sum_{\sigma \in \mathcal{Z}^L} e^{\beta \sum_{i=1}^{L-1} \sigma_i \sigma_{i+1} + \beta B \sum_{i=1}^{L} \sigma_i} = \sum_{\sigma_L \in \mathcal{Z}} z_{L-1}(\beta, B, \sigma_L) e^{\beta B \sigma_L}.$$

We define two row vectors $\psi_L \triangleq \begin{bmatrix} 1 & 1 \end{bmatrix} = z_0(\beta, B)$ and $\psi_R \triangleq \begin{bmatrix} e^{-\beta B} & e^{\beta B} \end{bmatrix}$, we obtain

$$z_{L-1}(\beta, B) = \psi_L T^{L-1}, \qquad Z_L(\beta, B) = \left\langle \psi_L T^{L-1}, \psi_R \right\rangle.$$

Let $u_1, u_2 \in \mathbb{R}$ such that $\psi_L = u_1 \psi_1 + u_2 \psi_2$. It follows that $\psi_L T^{L-1} = u_1 \lambda_1^{L-1} \psi_1 + u_2 \lambda_2^{L-1} \psi_2$, and the result follows. $\square$

**Lemma 1.7.** *Free entropy density for a one-dimensional ferromagnetic system with finite $\beta$ is given by $\phi(\beta, B) = \log \lambda_1$.*

*Proof.* The result is immediate from the definition of free entropy density $\phi(\beta, B) = \lim_{N \to \infty} \frac{1}{N} \log Z_N(\beta, B)$, the eigenvalue decomposition of partition function in (4), and the fact that $\lambda_1 > \lambda_2$. $\square$

**Lemma 1.8.** *Defining $2 \times 2$ diagonal matrix $\hat{\sigma} \triangleq \operatorname{diag}(-1, 1)$, the expected spin under the Boltzmann distribution for a one-dimensional ferromagnetic system with finite $\beta$ is*

$$\langle \sigma_i \rangle = \frac{1}{Z_L(\beta, B)} \left\langle \psi_L T^{i-1} \hat{\sigma} T^{L-i}, \psi_R \right\rangle. \tag{5}$$

*Proof.* From the definition of expected spin under the Boltzmann distribution for a one-dimensional ferromagnetic system with finite $\beta$, we write

$$\langle \sigma_i \rangle = \frac{1}{Z_L(\beta, B)} \sum_{\sigma \in \mathcal{Z}^L} \sigma_i e^{\beta \sum_{j=1}^{L-1} \sigma_j \sigma_{j+1} + \beta B \sum_{j=1}^{L} \sigma_j} = \frac{1}{Z_L(\beta, B)} \sum_{\sigma_i \in \mathcal{Z}} z_{i-1}(\beta, B, \sigma_i) \sigma_i (T^{L-i} \psi_R^T)_{\sigma_i}.$$

The result is immediate from the fact that $z_{i-1}(\beta, B, \sigma_i) = (\psi_L T^{i-1})_{\sigma_i}$ and $\langle a, b \rangle = ab^T$. □

**Lemma 1.9.** *Average magnetization for a one-dimensional ferromagnetic system with finite $\beta$ is*

$$M_N(\beta, B) = \frac{u_1 v_1 \lambda_1^{N-1} \langle \psi_1 \hat{\sigma}, \phi_1 \rangle + u_2 v_2 \lambda_2^{N-1} \langle \psi_2 \hat{\sigma}, \phi_2 \rangle + \frac{\lambda_1^N - \lambda_2^N}{N(\lambda_1 - \lambda_2)} (u_1 v_2 \langle \psi_1 \hat{\sigma}, \phi_2 \rangle + u_2 v_1 \langle \psi_2 \hat{\sigma}, \phi_1 \rangle)}{u_1 \lambda_1^{N-1} \langle \psi_1, \psi_R \rangle + u_2 \lambda_2^{N-1} \langle \psi_2, \psi_R \rangle}.$$

*Proof.* From the definition of average magnetization $M_N(\beta, B) = \frac{1}{N} \sum_{i=1}^{N} \langle \sigma_i \rangle$. We substitute the expression for partition function $Z_N(\beta)$ for one-dimensional ferromagnetic system from (4) in the denominator of the expression for $\langle \sigma_i \rangle$ in (5). For the numerator, we recall that $\psi_L = u_1 \psi_1 + u_2 \psi_2$ where $\psi_i$ is a left eigenvector of $T$ with eigenvalues $\lambda_i$ for $i \in \{1, 2\}$. It follows that

$$\left\langle \psi_L T^{i-1} \hat{\sigma} T^{L-i}, \psi_R \right\rangle = \left\langle (u_1 \lambda_1^{i-1} \psi_1 + u_2 \lambda_2^{i-1} \psi_2) \hat{\sigma} T^{L-i}, \psi_R \right\rangle.$$

Let $v_1, v_2 \in \mathbb{R}$ be such that $\psi_R^T = v_1 \phi_1^T + v_2 \phi_2^T$ where $\phi_i^T$ is a right eigenvector of $T$ with eigenvalues $\lambda_i$ for $i \in \{1, 2\}$. Then, $T^{L-i} \psi_R^T = v_1 \lambda_1^{L-i} \phi_1^T + v_2 \lambda_2^{L-i} \phi_2^T$, and hence $\left\langle \psi_L T^{i-1} \hat{\sigma} T^{L-i}, \psi_R \right\rangle$ equals

$$u_1 v_1 \lambda_1^{L-1} \langle \psi_1 \hat{\sigma}, \phi_1 \rangle + u_2 v_2 \lambda_2^{L-1} \langle \psi_2 \hat{\sigma}, \phi_2 \rangle + u_1 v_2 \lambda_1^{i-1} \lambda_2^{L-i} \langle \psi_1 \hat{\sigma}, \phi_2 \rangle + u_2 v_1 \lambda_2^{i-1} \lambda_1^{L-i} \langle \psi_2 \hat{\sigma}, \phi_1 \rangle.$$

Since $\frac{1}{N} \sum_{i=1}^{N} \lambda_1^{i-1} \lambda_2^{N-i} = \frac{\lambda_1^N - \lambda_2^N}{N(\lambda_1 - \lambda_2)}$, we obtain the result by evaluating $\frac{1}{N} \sum_{i=1}^{N} \left\langle \psi_L T^{i-1} \hat{\sigma} T^{N-i}, \psi_R \right\rangle$. □

**Lemma 1.10.** *Spontaneous magnetization $M_+(\beta) = 0$ for a one-dimensional ferromagnetic system with finite $\beta$.*

*Proof.* Since $\lambda_1 > \lambda_2$ and $\langle \psi_i, \phi_j \rangle = I_{ij}$, we can obtain the thermodynamic limit of average magnetization, as

$$\lim_{N \to \infty} M_N(\beta, B) = \frac{\langle \psi_1 \hat{\sigma}, \phi_1 \rangle}{\langle \psi_1, \phi_1 \rangle} = \frac{-\psi_{1,1} \phi_{1,1} + \psi_{1,2} \phi_{1,2}}{\psi_{1,1} \phi_{1,1} + \psi_{1,2} \phi_{1,2}}.$$

Substituting (3) and the identity $e^{\beta(1-B)} - \lambda_1 = -e^{\beta} \sinh(\beta B) - \sqrt{e^{2\beta} \sinh^2(\beta B) + e^{-2\beta}}$ in the above equation, we obtain

$$\lim_{N \to \infty} M_N(\beta, B) = \frac{-e^{-2\beta} + (e^{\beta(1-B)} - \lambda_1)^2}{e^{-2\beta} + (e^{\beta(1-B)} - \lambda_1)^2} = \frac{\sinh(\beta B)}{\sqrt{\sinh^2(\beta B) + e^{-4\beta}}}.$$

For $\beta < \infty$, the average magnetization is an analytic function of $\beta$ and $B$. In particular, at any non-zero temperature, the spontaneous magnetization is zero, i.e. $M_+(\beta) = 0$ for all $\beta < \infty$. □

*Remark* 3. For one-dimensional configurations $\sigma \in \mathcal{X} \triangleq \{-1, 1\}^L$ with $L$ spins, we recall that

$$E(\sigma, B) = -\sum_{i=1}^{L-1} \sigma_i \sigma_{i+1} - B \sum_{i=1}^{L} \sigma_i,$$

and hence $-\frac{\partial E(\sigma, B)}{\partial B} = \sum_{i=1}^{L} \sigma_i$. Thus, we can write the average magnetization as

$$\frac{1}{L} \left\langle \sum_{i=1}^{L} \sigma_i \right\rangle_{\beta, B} = -\frac{1}{L} \left\langle \frac{\partial E(\sigma, B)}{\partial B} \right\rangle_{\beta, B}.$$

Hence, we could have computed the thermodynamic limit directly by observing that $\Phi_L(\beta, B) = \ln Z_L(\beta, B)$, and noticing that

$$\frac{\partial \Phi_L(\beta, B)}{\partial B} = -\beta \sum_{x \in \mathcal{X}} \mu_{L, \beta, B}(x) \frac{\partial E(x, B)}{\partial B} = -\beta \left\langle \frac{\partial E(x, B)}{\partial B} \right\rangle_{\beta, B}.$$

Thus, dividing both sides by $L$, taking limit $L \to \infty$, interchanging limits on the left hand side, and rearranging the above equation, we obtain

$$\frac{1}{\beta}\frac{\partial \phi(\beta, B)}{\partial B} = \lim_{L\to\infty} \frac{1}{L}\left\langle \sum_{i=1}^{L} \sigma_i \right\rangle_{\beta,B}.$$

For one-dimensional ferromagnet model, the free entropy density is $\phi(\beta, B) = \lim_{L\to\infty} \frac{1}{L}\Phi_L(\beta, B) = \ln \lambda_1$. Hence, we obtain

$$\frac{1}{\beta}\frac{\partial \phi(\beta, B)}{\partial B} = \frac{1}{\beta\lambda_1}\frac{\partial \lambda_1}{\partial B} = \frac{1}{\beta\lambda_1}\left(\beta e^{\beta}\sinh(\beta B) + \frac{\beta e^{2\beta}\sinh(\beta B)\cosh(\beta B)}{\sqrt{e^{2\beta}\sinh^2(\beta B) + e^{-2\beta}}}\right) = \frac{\sinh(\beta B)}{\sqrt{\sinh^2(\beta B) + e^{-4\beta}}}.$$

**Definition 1.11 (Susceptibility).** The susceptibility associated with limiting average magnetization $M(\beta, B) \triangleq \lim_{N\to\infty} M_N(\beta, B)$ is defined as

$$\chi_M(\beta) \triangleq \frac{\partial M(\beta, 0)}{\partial B}. \tag{6}$$

*Remark* 4. Intuitively, susceptibility is the tendency of a site in a region to have the same alignment as its neighbors. For one-dimensional Ferromagnets, we have

$$\chi_M(\beta) = \frac{\partial M(\beta, B)}{\partial B}\Big|_{B=0} = \frac{\beta e^{-4\beta}\cosh(\beta B)}{(\sinh^2(\beta B) + e^{-4\beta})^{\frac{3}{2}}}\Big|_{B=0} = \beta e^{2\beta}.$$

*Remark* 5. A single spin in a field has a susceptibility $\chi_M(\beta) = \beta$. If we consider $N$ spins constrained to take the same value, the corresponding susceptibility will be $N\beta$, as in (2). For one-dimensional Ferromagnets with $N$ spins, the system behaves like the spins were blocked into groups of $\frac{\chi(\beta)}{\beta}$ spins each. The spins in each group are restricted to a value, while spins in different groups are independent.

**Example 1.12.** Consider one-dimensional Ferromagnetic model with $N$ spins and zero magnetic field $B = 0$ and $\delta N < i < j < (1-\delta)N$. Defining correlation length of the model $\xi(\beta) \triangleq -\frac{1}{\log\tanh\beta}$ as the distance below which two spins are well correlated, one finds the correlation function at large $N$, as

$$\langle \sigma_i \sigma_j \rangle = e^{-\frac{|i-j|}{\xi(\beta)}} + \Theta(e^{-\alpha N}).$$

The correlation length $\xi(B)$ increases with decrease in temperature, that is, spins become more correlated at lower temperatures. The relation between correlation length and susceptibility is given by,

$$\chi_M(\beta) = \beta \sum_{i=-\infty}^{\infty} e^{\frac{-|i|}{\xi(\beta)}} + \Theta(e^{-\alpha N}). \tag{7}$$

This makes it evident that a large susceptibility must correspond to a large correlation length.

# 2 Curie-Weiss Model

The exact solution of the one-dimensional model, lead Ising to think that there couldn't be a phase transition in any dimension. This was debunked by a qualitative theory of ferromagnetism which was put forward by Pierre Curie. It assumed the existence of a phase transition at non-zero temperature $T_c$ (Curie point) and a non-vanishing spontaneous magnetization for $T < T_c$. The dilemma was eventually solved by Onsager solution of the two-dimensional model.

**Definition 2.1 (Curie-Weiss model).** Consider $N$ Ising spins $\sigma_i \in \mathcal{Z} \triangleq \{-1, 1\}$ and a configuration $\sigma \in \mathcal{X} \triangleq \mathcal{Z}^N$. The energy function in the presence of a magnetic field $B$, is defined as

$$E(\sigma) \triangleq -\frac{1}{2N}\sum_{i\neq j\in[N]} \sigma_i\sigma_j - B\sum_{i=1}^{N} \sigma_i.$$

*Remark* 6. Unlike the Ising model, the spins are not a part of a $d$-dimensional lattice, instead, they all interact in pairs. The absence of any finite-dimensional geometrical structure makes the Curie-Weiss model one among the mean-field models.

*Remark* 7. It needs to be mentioned that the summation over $i, j \in [N]$ involves $O(N^2)$ terms of order $O(1)$. Therefore, the energy function is scaled by $\frac{1}{N}$ to obtain a non-trivial free-energy density in the thermodynamic limit.

*Remark* 8. Curie-Weiss model is an example of the more general mean-field model.

**Definition 2.2.** For a configuration $\sigma \in \mathcal{X}^N$, the **empirical or instantaneous magnetization** is given by $m(\sigma) \triangleq \frac{1}{N}\sum_{i=1}^N \sigma_i$.

*Remark* 9. We can write the energy function $E(\sigma, B)$ in the presence of a magnetic field $B$ in terms of empirical magnetization $m(\sigma)$, as

$$E(\sigma, B) = -\frac{1}{2N}\sum_{i=1}^N \sigma_i(Nm(\sigma) - \sigma_i) - BNm(\sigma) = -\frac{N}{2}m(\sigma)^2 + \frac{1}{2} - BNm(\sigma). \tag{8}$$

*Remark* 10. An immediate observation is that $\langle m(\sigma)\rangle = \frac{1}{N}\sum_{i=1}^N \langle \sigma_i \rangle = M_N(\beta, B)$. That is, the instantaneous magnetization is a function of the empirical magnetization $m(\sigma)$.

**Proposition 2.3 (Free energy density for Curie-Weiss model).** *The free energy density for Curie-Weiss model is*

$$\phi(\beta, B) = \sup_{m \in [-1,1]} \phi_{\text{mf}}(m; \beta, B),$$

*where the map* $\phi_{\text{mf}} : [-1, 1] \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ *is defined for all* $(m, \beta, B) \in [-1, 1] \times \mathbb{R}_+ \times \mathbb{R}$ *as*

$$\phi_{\text{mf}}(m; \beta, B) \triangleq \frac{\beta}{2}m^2 + \beta B m + H\Big(\frac{m+1}{2}\Big). \tag{9}$$

*Further, the free energy density is maximized at* $m_*$ *that solves the following implicit equation*

$$m_* = \tanh(\beta m_* + \beta B). \tag{10}$$

*Proof.* From the expression (8) for energy in terms of empirical magnetization $m$, we can write the partition function as

$$Z_N(\beta, B) = e^{-\frac{N\beta}{2}} \sum_{\sigma \in \mathcal{X}} e^{\beta N m(\sigma)(\frac{m(\sigma)}{2} + B)} \overset{(a)}{=} e^{-\frac{N\beta}{2}} \sum_m \binom{N}{\frac{N(m+1)}{2}} e^{\beta N m(\frac{m}{2}+B)} \doteq_N e^{-\frac{N\beta}{2}} \sum_m e^{NH(\frac{(m+1)}{2})} e^{\beta N m(\frac{m}{2}+B)},$$

where equation (a) follows from the observation that the number of positive spins in a configurations $\sigma \in \mathcal{X}$ is $N(m(\sigma)+1)/2$. Moreover, we use $H$ to represent the binary entropy function expressed in nats. In terms of $\phi_{\text{mf}}$ defined in (9), we can approximate the partition function for large $N$, as

$$Z_N(\beta, B) \doteq_N \int_{-1}^{1} e^{N\phi_{\text{mf}}(m;\beta,B)} dm. \tag{11}$$

The largest contribution to the integral in the above equation comes from the largest exponent, and hence the result follows for the free energy density. To maximize $\phi_{\text{mf}}$, we take its derivative with respect to $m$ and equate it to zero. Since $\frac{dH(p)}{dp} = \ln(\frac{1}{p} - 1)$, we obtain

$$0 = \frac{\partial \phi_{\text{mf}}(m; \beta, B)}{\partial m}\Big|_{m=m_*} = \beta m_* + \beta B + \frac{1}{2}\ln\left(\frac{2}{m_*+1} - 1\right).$$

The solution of this equation is the solution to the implicit equation (10). $\square$

**Proposition 2.4 (Phase transition).** *For* $B = 0$*, the free energy density has a unique maximum for* $\beta < 1$*, and has two maxima for* $\beta > 1$*. That is, there is a phase transition at* ***Curie temperature*** $T_c \triangleq \frac{1}{\beta_c} = 1$.

*Proof.* We observe that $\phi_{\rm mf}(m;\beta,0) = \phi_{\rm mf}(-m;\beta,0)$ is an even function of $m$. Further, we observe from (10) that

$$\frac{\partial \phi_{\rm mf}(m;\beta,0)}{\partial m} = \beta m + \frac{1}{2}\ln\Big(\frac{2}{m+1}-1\Big), \qquad \frac{\partial^2 \phi_{\rm mf}(m;\beta,0)}{\partial^2 m} = \beta - \frac{1}{1-m^2}.$$

We observe that $\frac{\partial \phi_{\rm mf}(m;\beta,0)}{\partial m} = 0$ at $m = 0$.

$\beta < 1$. Since $\frac{1}{1-m^2} > 1$, it follows that $\phi_{\rm mf}(m;\beta,0)$ is concave decreasing in $[0,1]$. Thus, $\phi_{\rm mf}(m;\beta,0)$ is maximized uniquely at $m_* = 0$.

$\beta > 1$. For small $m > 0$, we have $\beta > \frac{1}{1-m^2}$ and $\beta < \frac{1}{1-m^2}$ for large $m \to 1$. That is, $\phi_{\rm mf}(m;\beta,0)$ is convex for $m \in [-\sqrt{1-\frac{1}{\beta}}, \sqrt{1-\frac{1}{\beta}})$, and concave in $[-1, -\sqrt{1-\frac{1}{\beta}}) \cup (\sqrt{1-\frac{1}{\beta}}, 1]$. It follows that $\phi_{\rm mf}(m;\beta,0)$ is maximized at $m_+ = -m_- \in (\sqrt{1-\frac{1}{\beta}}, 1]$.

□

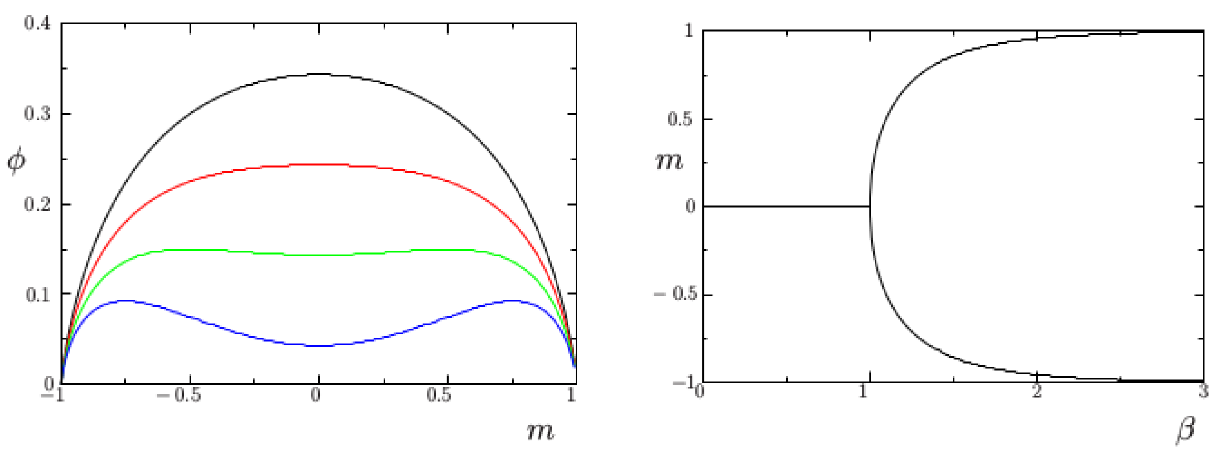


Figure 1: The plot on the left shows the variation of $\phi_{\rm mf}(m;\beta,0)$ with $m$, for different values of $\beta$. For $\beta < 1$, there is a unique maximum, and for $\beta > 1$, there are two maxima—indicating a phase transition at $\beta = \beta_c = 1$. On the right, the plot shows the variation of the values of $m$ that maximize $\phi_{\rm mf}(m;\beta,0)$, with $\beta$. The phase transition at $\beta = 1$ is indicated by a bifurcation.

# 3 The Ising spin-glass (or Edwards-Anderson) model

**Definition 3.1 (Edwards-Anderson).** Consider a configuration $\sigma \in \mathcal{X}^N = \{-1,+1\}^{\mathbb{L}}$ of the $N$-particle system with $\mathbb{L} = \{1,\ldots,L\}^d$ representing a $d$-dimensional lattice. In the Ising spin-glass model under magnetic field $B$, the energy is defined as

$$E(\sigma) \triangleq -\sum_{(ij)} J_{i,j}\sigma_i\sigma_j - B\sum_{i\in\mathbb{L}} \sigma_i.$$

*Remark* 11. Here, the first summation runs over each edge of the lattice, and the multiplying factor, $J_{i,j} \in \mathbb{R}$, for $i,j \in \mathbb{L}$. Note the difference in the energy function from the Ising model, in that now, each 2-particle interaction is multiplied by a (possibly) different factor.

We state here that it is not straightforward to arrive at a low energy configuration in this model (by satisfying each local constraint), as elucidated in the example below.

**Example 3.2.** Consider the Ising spin-glass model, for an $L = 2,\ d = 2$ system, with $B = 0$. The lattice is hence a 2-dimensional square, with vertices $V = \{(1,1),(1,2),(2,1),(2,2)\}$. Let $J_{(1,1)\sim(1,2)} = 1$, $J_{(1,1)\sim(2,1)} = 1$, $J_{(2,1)\sim(2,2)} = 1$ and $J_{(2,2)\sim(1,2)} = -1$, where the notation $(a,b) \sim (c,d)$ is used to represent the edge between $(a,b)$ and $(c,d)$, for $(a,b),(c,d) \in \mathbb{L}$. We observe that the two configurations,

$$\sigma^1 = \big(\sigma_{(1,1)} = 1, \sigma_{(1,2)} = 1, \sigma_{(2,1)} = 1, \sigma_{(2,2)} = 1\big)$$
$$\text{and } \sigma^2 = \big(\sigma_{(1,1)} = 1, \sigma_{(1,2)} = -1, \sigma_{(2,1)} = 1, \sigma_{(2,2)} = 1\big)$$

are degenerate, with $E(\sigma^1) = E(\sigma^2) = 2$. This is, however, a **frustrated** system, since it is impossible to satisfy each local constraint induced by the individual $J_{i,j}$s, $i,j \in \mathbb{L}$.

# 4 Optimization and statistical physics

Combinatorial optimization problems present inherent difficulties owing to the "discreteness" (or lack of smoothness) of the space. In general, in such problems, given a configuration space $\mathcal{X}$, we wish to find a configuration $x \in \mathcal{X}$ with the smallest cost. It is possible to introduce a Boltzmann probability distribution $\mu_\beta \in \mathcal{M}(\mathcal{X})$ on the space of configurations $\mathcal{X}$, such that for any configuration $x \in \mathcal{X}$

$$\mu_\beta(x) \triangleq \frac{1}{Z(\beta)} e^{-\beta E(x)}, \qquad Z(\beta) \triangleq \sum_{x \in \mathcal{X}} e^{-\beta E(x)}.$$

In the limit as $\beta \to \infty$, the probability distribution concentrates on the ground states—which is the case when all optimization constraints are satisfied.

**Example 4.1 (Min-cuts on graphs).** We consider again, the Ising spin-glass model, with $B = 0$, and $E(\sigma) = \sum_{(ij)} J_{i,j}\sigma_i\sigma_j$. Each configuration $\sigma$ partitions the set $[N]$ into two subsets, defined as

$$V_+ \triangleq \{i \in [N] : \sigma_i = +1\}, \qquad V_- \triangleq \{i : \sigma_i = -1\}.$$

Defining $C \triangleq \sum_{(ij)} J_{i,j}$, and $\gamma(V_+) \triangleq \{(i,j) : i \in V_+, j \in V_-\}$, we obtain $E(\sigma) = -C + 2\sum_{(ij)\in\gamma(V_+)} J_{i,j}$. Solving for the lowest energy configuration, is hence, exactly equivalent to finding the min-cut in the graph $G = (V_+ \cup V_-, E)$, with $E$ being the set of edges induced by the particle interactions.

**Example 4.2 (Error-correcting codes).** In this example, we illustrate the potential hardness of the decoding problem, for binary codes. Let the binary alphabet be $\mathcal{Z} \triangleq 0, 1$, and set of $N$-length binary vectors $\mathcal{X} \triangleq \mathcal{Z}^N$. For $M$ information bits, we denote the set of messages as $\mathcal{M} \triangleq \{0, \ldots, 2^M - 1\}$. Recall the setting of a communication system, which consists of an encoder $e : \mathcal{M} \mapsto \mathcal{X}$ that maps the output of an *i.i.d.* uniformly distributed source $m$ to a codeword $x = e(m) \in \mathcal{X}$. Let $y \in \mathcal{X}$ be the output of the channel described by the conditional distribution $Q(y|x)$ for all $y, x \in \mathcal{X}$. The decoder, $d : y \in \mathcal{X} \mapsto \mathcal{X}$ outputs an estimate, $\hat{x} \triangleq d(y)$ of the codeword $x$. The average probability of error

$$P_B^{avg} = \frac{1}{2^M} \sum_m \sum_{y: d(y) \neq e(m)} Q(y|e(m)) = 1 - \frac{1}{2^M} \sum_y \sum_m Q(y|e(m)) \mathbb{1}_{\{d(y)=e(m)\}}.$$

It is, therefore, obvious after interchanging summations that the optimal decoder that minimizes $P_B^{avg}$ must map the received word $y$ to that codeword, $\hat{\mathbf{x}}$, that maximizes $Q(y|\hat{\mathbf{x}})$. However, one can note that this procedure of finding the "most likely" codeword, involves searching over all possible $2^M$ codewords, leading to exponential time complexity. In fact, the general problem of decoding codes that admit a concise specification (polynomial in the block-length) is NP-hard.