

2 Equilibrium statistical mechanics: Curie–Weiss model

Equilibrium statistical mechanics imagines a system with N degrees of freedom described by a Hamiltonian H and weakly coupled to a bath of temperature $T = \beta^{-1}$. The probability distribution over states of the system is then described by the Boltzmann distribution

$$p(\mathbf{s}) = \frac{1}{Z} e^{-\beta H(\mathbf{s})} \quad (1)$$

where the partition function Z is a normalization constant defined by

$$Z = \int_{\mathbb{R}^N} d\mathbf{s} e^{-\beta H(\mathbf{s})} \quad \text{or} \quad Z = \sum_{\mathbf{s}} e^{-\beta H(\mathbf{s})} \quad (2)$$

depending on whether \mathbf{s} is continuous or discrete. This simple framework contains within the entire complexity of equilibrium statistical mechanics, from phase transitions to fluctuation relations. In particular, many thermodynamic quantities of interest can be extracted from Z , or rather the free energy

$$F = -\beta^{-1} \log Z \quad (3)$$

The average energy is given by

$$\begin{aligned} E = \langle H \rangle &= \frac{1}{Z} \int_{\mathbb{R}^N} d\mathbf{s} e^{-\beta H(\mathbf{s})} H(\mathbf{s}) = -\frac{1}{Z} \frac{\partial}{\partial \beta} \int_{\mathbb{R}^N} d\mathbf{s} e^{-\beta H(\mathbf{s})} \\ &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta} \log Z = \frac{\partial}{\partial \beta} (\beta F) \end{aligned} \quad (4)$$

whereas the entropy is given by

$$TS = E - F = \beta \frac{\partial F}{\partial \beta} \quad (5)$$

A prototypical model of equilibrium statistical mechanics is the Ising model, defined for a set of spins that take values ± 1 on each site of a lattice by the Hamiltonian

$$H(\mathbf{s}) = -\frac{1}{2} \sum_{\langle ij \rangle} s_i s_j \quad (6)$$

where the notation $\langle ij \rangle$ means the set of i, j for which the corresponding lattice sites are neighbors. The Ising model is a model for ferromagnetic materials, with the direction of the spins representing the magnetic moment of atoms in a material and their interaction representing the effective exchange interaction. As simple as it is, it is not solvable under most circumstances, so for illustrative purposes it is more convenient to use the Curie–Weiss model

$$H(\mathbf{s}) = -\frac{1}{2} \frac{1}{N} \sum_{i,j=1}^N s_i s_j \quad (7)$$

The only differences between the Curie–Weiss model and the Ising model are (1) the sum now runs over all pairs of spins and (2) the interaction strength is rescaled by N to make H typically order- N .

For our purposes, we will simplify the model yet further, and instead take $\mathbf{s} \in \mathbb{R}^N$ as a real N -dimensional configuration confined to the sphere in N dimensions $\|\mathbf{s}\|^2 = N$. This is convenient because it will have a more direct connections with the problems we will study later in the course, and because it makes analyzing the dynamics of the model easier. The partition function of this model is given by

$$Z = \int_{\mathbb{R}^N} d\mathbf{s} \sqrt{N} \delta\left(\frac{1}{2}(N - \|\mathbf{s}\|^2)\right) e^{\frac{1}{2}\beta \frac{1}{N} \sum_{i,j=1}^N s_i s_j} \quad (8)$$

Define an order parameter $m = \frac{1}{N} \sum_{i=1}^N s_i = \frac{1}{N} \mathbf{s} \cdot \mathbf{1}$, the magnetization. We can write the partition function in terms of this order parameter as

$$Z = \sqrt{N} \int_{\mathbb{R}} dm e^{\frac{1}{2}\beta N m^2} \int_{\mathbb{R}^N} d\mathbf{s} \delta\left(\frac{1}{2}(N - \|\mathbf{s}\|^2)\right) N \delta(Nm - \mathbf{s} \cdot \mathbf{1}) \quad (9)$$

where we have added another δ function to enforce the definition of the order parameter. The integral over \mathbf{s} as a function of fixed m gives the entropy associated with a given fixed magnetization. Converting the δ functions to their Fourier representation, we have

$$Z = \sqrt{N^3} \int_{\mathbb{R}} dm e^{\frac{1}{2}\beta N m^2} \int_{\mathbb{R}} d\hat{m} d\lambda \int_{\mathbb{R}^N} d\mathbf{s} e^{i\lambda \frac{1}{2}(N - \|\mathbf{s}\|^2) + i\hat{m}(Nm - \mathbf{s} \cdot \mathbf{1})} \quad (10)$$

The integral in \mathbf{s} is Gaussian, and we can perform it explicitly, giving

$$Z = \sqrt{N^3} \int_{\mathbb{R}} dm e^{\frac{1}{2}\beta N m^2} \int_{\mathbb{R}} d\hat{m} d\lambda e^{i\frac{1}{2}N\lambda + iN\hat{m}m} \sqrt{\frac{(2\pi)^N}{\det(i\lambda I_N)}} e^{-\frac{1}{2} \frac{\hat{m}^2}{i\lambda} \mathbf{1} \cdot \mathbf{1}} \quad (11)$$

$$= \sqrt{N^3} \sqrt{(2\pi)^N} \int_{\mathbb{R}} dm e^{\frac{1}{2}\beta N m^2} \int_{\mathbb{R}} d\hat{m} d\lambda e^{i\frac{1}{2}N\lambda + iN\hat{m}m - \frac{1}{2}N \frac{\hat{m}^2}{i\lambda} - \frac{1}{2}N \log(i\lambda)}$$

The integral over \hat{m} is also Gaussian, and we can perform it to give

$$Z = \sqrt{N^3} \sqrt{(2\pi)^N} \int_{\mathbb{R}} dm e^{\frac{1}{2}\beta N m^2} \int_{\mathbb{R}} d\lambda \sqrt{\frac{2\pi}{N/i\lambda}} e^{i\frac{1}{2}N\lambda - \frac{1}{2}N \log(i\lambda) - \frac{1}{2}iN\lambda m^2} \quad (12)$$

$$= N \sqrt{(2\pi)^{N+1}} \int_{\mathbb{R}} dm e^{\frac{1}{2}\beta N m^2} \int_{\mathbb{R}} d\lambda e^{i\frac{1}{2}N(1-m^2)\lambda - \frac{1}{2}(N-1) \log(i\lambda)}$$

The integral over λ now has the form that is amenable to a saddle point treatment, with

$$\mathcal{S}_m(\lambda) = \frac{1}{2}i\lambda(1 - m^2) - \frac{1}{2} \log(i\lambda) \quad (13)$$

The integral should be evaluated by finding a saddle point of S_m , whose conditions are

$$0 = \frac{\partial S_m}{\partial \lambda} = \frac{1}{2}i(1 - m^2) - \frac{1}{2\lambda} \quad (14)$$

which is solved by $\lambda^* = \frac{-i}{1-m^2}$. Since

$$\frac{\partial^2 S_m(\lambda^*)}{\partial \lambda^2} = \frac{1}{2} \frac{1}{(\lambda^*)^2} = -\frac{1}{2}(1 - m^2)^2 \quad (15)$$

we have

$$\begin{aligned} Z &\simeq N \sqrt{(2\pi)^{N+1}} \int_{\mathbb{R}} dm e^{\frac{1}{2}Nm^2} \sqrt{\frac{2\pi/N}{\frac{1}{2}(1-m^2)^2}} e^{\frac{1}{2}N - \frac{1}{2}N \log \frac{1}{1-m^2}} \\ &= \sqrt{\frac{N(2\pi)^2}{\frac{1}{2}(1-m^2)^2}} \int_{\mathbb{R}} dm e^{\frac{1}{2}\beta Nm^2 + \frac{1}{2}N(1 + \log(2\pi)) + \frac{1}{2}N \log(1-m^2)} \end{aligned} \quad (16)$$

The integrand is now a function of the magnetization alone, and is in the form for applying the saddle point approximation, with effective action

$$S(m) = \frac{1}{2}(1 + \log 2\pi) + \frac{1}{2}\beta m^2 + \frac{1}{2} \log(1 - m^2) \quad (17)$$

The saddle points of this action are given by

$$0 = \frac{\partial S}{\partial m} = \beta m - \frac{m}{1 - m^2} \quad (18)$$

or $0 = m(\beta - 1 - \beta m^2)$. This is solved for either $m^* = 0$, or

$$m^* = \pm \sqrt{\frac{\beta - 1}{\beta}} \quad (19)$$

These nonzero solutions are only sensible for $\beta > 1$, which marks the phase transition in this model. Substituting these solutions into $S(m)$, we can write the free energy per site as

$$\begin{aligned} f &= \frac{1}{N} F = -\frac{1}{N} \beta^{-1} \log Z = -\beta^{-1} S(m^*) \\ &= \begin{cases} -\frac{1}{2\beta} (1 + \log 2\pi) & \beta < 1 \\ -\frac{1}{2\beta} (1 + \log 2\pi) - \frac{1}{2} \frac{\beta-1}{\beta} + \frac{1}{2\beta} \log \beta & \beta > 1 \end{cases} \end{aligned} \quad (20)$$

Note that $m^* = \pm \sqrt{T - T_c}$: we have a square root singularity in the emergence of the order parameter at the phase transition in mean field.

Equilibrium dynamics can be studied using the Langevin equation, with

$$\dot{\mathbf{s}}(t) = -\nabla H(\mathbf{s}(t)) + \boldsymbol{\xi}(t) \quad (21)$$

where $\xi(t)$ is a random Gaussian function with zero mean and covariance $\langle \xi_i(t)\xi_j(t') \rangle = 2T\delta_{ij}\delta(t-t')$. For our spherical Curie–Weiss model, we need to add another term to keep the dynamics confined to the sphere. This is

$$\dot{\mathbf{s}}(t) = -H'(\mathbf{s}(t)) - \mu(t)\mathbf{s}(t) + \xi(t) \quad (22)$$

where μ is a time-dependent force that cancels any dynamics that would remove the configuration from the sphere. For the spherical Curie–Weiss model, the gradient of the energy gives

$$\dot{\mathbf{s}}(t) = \frac{1}{N}(1 \cdot \mathbf{s}(t))1 - \mu(t)\mathbf{s}(t) + \xi(t) \quad (23)$$

An equation for μ can be determined by requiring that $\|\mathbf{s}\|^2$ is a constant of the motion, or

$$\begin{aligned} 0 &= \frac{1}{2} \frac{\partial}{\partial t} \|\mathbf{s}(t)\|^2 = \mathbf{s}(t) \cdot \dot{\mathbf{s}}(t) = \frac{1}{N}(1 \cdot \mathbf{s}(t))^2 - \mu(t)\|\mathbf{s}(t)\|^2 + \mathbf{s}(t) \cdot \xi(t) \\ &= \frac{1}{N}(1 \cdot \mathbf{s}(t))^2 - \mu(t)N + \mathbf{s}(t) \cdot \xi(t) \end{aligned}$$

which implies

$$\mu(t) = \frac{1}{N^2}(1 \cdot \mathbf{s}(t))^2 + \frac{1}{N}\xi(t) \cdot \mathbf{s}(t) \quad (24)$$

averaged over instantiations of the noise. We want to understand the behavior of the typical correlation and response functions

$$C(t, t') = \frac{1}{N} \langle \mathbf{s}(t) \cdot \mathbf{s}(t') \rangle \quad R(t, t') = \frac{1}{N} \sum_{i=1}^N \left\langle \frac{\partial s_i(t)}{\partial \xi_i(t')} \right\rangle \quad (25)$$

We can determine equations for these self-consistently by manipulating the equation of motion. A differential equation for the correlation function can be written by taking the scalar product of the equation of motion with $\mathbf{s}(t')$, or

$$\dot{\mathbf{s}}(t) \cdot \mathbf{s}(t') = \frac{1}{N}(1 \cdot \mathbf{s}(t))(1 \cdot \mathbf{s}(t')) - \mu(t)(\mathbf{s}(t) \cdot \mathbf{s}(t')) + \xi(t) \cdot \mathbf{s}(t') \quad (26)$$

Taking the average of both sides over noise, we find

$$\frac{\partial}{\partial t} C(t, t') = m(t)m(t') - \langle \mu(t) \rangle C(t, t') + \frac{1}{N} \langle \xi(t) \cdot \mathbf{s}(t') \rangle \quad (27)$$

where using the equation for μ we have

$$\langle \mu(t) \rangle = m(t)^2 + \frac{1}{N} \langle \xi(t) \cdot \mathbf{s}(t) \rangle \quad (28)$$

We can treat the average over the scalar product between the noise and the spin using Wick's theorem, which gives

$$\begin{aligned}
\frac{1}{N} \langle \boldsymbol{\xi}(t) \cdot \mathbf{s}(t') \rangle &= \frac{1}{N} \sum_{ij=1}^N \int dt'' \langle \xi_i(t) \xi_j(t'') \rangle \left\langle \frac{\delta s_i(t')}{\delta \xi_j(t'')} \right\rangle \\
&= 2T \frac{1}{N} \sum_{ij=1}^N \delta_{ij} \int dt'' \delta(t-t'') \left\langle \frac{\delta s_i(t')}{\delta \xi_j(t'')} \right\rangle \\
&= 2T \frac{1}{N} \sum_{i=1}^N \left\langle \frac{\delta s_i(t')}{\delta \xi_i(t)} \right\rangle = 2TR(t', t)
\end{aligned} \tag{29}$$

where we have identified the definition of the response function. We therefore have

$$\partial_t C(t, t') = m(t)m(t') - \langle \mu(t) \rangle C(t, t') + 2TR(t', t) \tag{30}$$

We can build an equation for the response function by taking the variation of the equation of motion with respect to the noise, which gives

$$\frac{1}{N} \sum_i \frac{\delta \dot{s}_i(t)}{\delta \xi_i(t')} = \frac{1}{N^2} \sum_{ij} \frac{\delta s_j(t)}{\delta \xi_i(t')} - \frac{1}{N} \sum_i \frac{\delta \mu(t)}{\delta \xi_i(t')} s_i(t) - \mu(t) \frac{1}{N} \sum_i \frac{\delta s_i(t)}{\delta \xi_i(t')} + \delta(t-t')$$

The first two terms are smaller by a factor of N than the others. For the first, this is true because off-diagonal parts of the response ($i \neq j$) are smaller than diagonal ones by a factor of N . Then

$$\frac{1}{N^2} \sum_{ij} \left\langle \frac{\delta s_j(t)}{\delta \xi_i(t')} \right\rangle \simeq \frac{1}{N^2} \sum_i \left\langle \frac{\delta s_i(t)}{\delta \xi_i(t')} \right\rangle = \frac{1}{N} R(t, t') \tag{31}$$

To treat the second term, we note that

$$\begin{aligned}
\frac{1}{N} \sum_i \left\langle \frac{\delta \mu(t)}{\delta \xi_i(t')} s_i(t) \right\rangle &= \frac{2}{N^3} \left\langle (1 \cdot \mathbf{s}(t)) \sum_{ij} \frac{\delta s_j(t)}{\delta \xi_i(t')} s_i(t) \right\rangle + \frac{1}{N^2} \left\langle \sum_i s_i(t)^2 \right\rangle \delta(t-t') \\
&\simeq \frac{2}{N^3} \langle 1 \cdot \mathbf{s}(t) \rangle \sum_i \left\langle \frac{\delta s_i(t)}{\delta \xi_i(t')} \right\rangle \langle s_i(t) \rangle + \frac{1}{N} \delta(t-t') \\
&= \frac{2}{N} m(t)^2 R(t, t') + \frac{1}{N} \delta(t-t')
\end{aligned}$$

Neglecting these two terms, taking the average of the equation above gives

$$\frac{\partial}{\partial t} R(t, t') = \frac{1}{N} \sum_i \left\langle \frac{\delta \dot{s}_i(t)}{\delta \xi_i(t')} \right\rangle \simeq -\mu(t) R(t, t') + \delta(t-t')$$

We can write an equation for the magnetization by taking the product of the equation of motion with 1, which gives

$$\begin{aligned}\dot{\mathbf{m}}(t) &= \frac{1}{N} \langle \mathbf{1} \cdot \dot{\mathbf{s}}(t) \rangle \\ &= \frac{1}{N^2} \langle \mathbf{1} \cdot \mathbf{1} \rangle \langle \mathbf{1} \cdot \mathbf{s}(t) \rangle - \frac{1}{N} \langle \boldsymbol{\mu}(t) \rangle \langle \mathbf{1} \cdot \mathbf{s}(t) \rangle + \frac{1}{N} \langle \boldsymbol{\xi}(t) \cdot \mathbf{1} \rangle \\ &= \mathbf{m}(t)(1 - \boldsymbol{\mu}(t))\end{aligned}$$

since $\langle \boldsymbol{\xi}(t) \cdot \mathbf{1} \rangle = 0$.

In equilibrium, the equations will reach a steady state, which means that the correlation and response will be invariant under translations in time, and therefore that they will depend only on time differences $\tau = t - t'$. Likewise, $\mathbf{m}(t) = \mathbf{m}$ and $\langle \boldsymbol{\mu}(t) \rangle = \boldsymbol{\mu}$ will be constants. It follows that

$$\begin{aligned}C'(\tau) &= \mathbf{m}^2 - \boldsymbol{\mu}C(\tau) + 2\text{TR}(-\tau) \\ R'(\tau) &= -\boldsymbol{\mu}R(\tau) + \delta(\tau) \\ \boldsymbol{\mu} &= \mathbf{m}^2 + 2\text{TR}(0) \\ 0 &= \mathbf{m}(1 - \boldsymbol{\mu})\end{aligned}$$

The equation for R is solved by

$$R(\tau) = \Theta(\tau)e^{-\boldsymbol{\mu}\tau} \quad (32)$$

where Θ is the Heaviside function that is zero for negative argument and one for positive argument. Its presence, caused by the Dirac δ function in the equation, preserves causality: responses cannot precede perturbations.

It is difficult to draw conclusions about the value of $\boldsymbol{\mu}$ from the above equation, since $R(0)$ is point of discontinuity of the response function. However, there is another way to fix its value. Since C is symmetric under exchange of t and t' , when expressed in a single variable it must be symmetric under $\tau \mapsto -\tau$. It therefore follows that $C'(0_+) = -C'(0_-)$. Since $R(0_-) = 0$ and $R(0_+) = 1$ and $C(0) = 1$, we have

$$C'(0_+) = \mathbf{m}^2 - \boldsymbol{\mu} \quad C'(0_-) = \mathbf{m}^2 - \boldsymbol{\mu} + 2\mathbf{T} \quad (33)$$

$$C'(0_+) = -C'(0_-) \mapsto \boldsymbol{\mu} = \mathbf{m}^2 + \mathbf{T} \quad (34)$$

The equation for \mathbf{m} , which is satisfied for $\mathbf{m} = 0$ or $\boldsymbol{\mu} = 1$, is therefore solved by $\mathbf{m} = 0$ or $\mathbf{m}^2 = 1 - \mathbf{T} = \frac{\beta-1}{\beta}$, exactly as in equilibrium. Finally, the equation for C is solved by

$$C(\tau) = e^{-\boldsymbol{\mu}|\tau|} + (1 - e^{-\boldsymbol{\mu}|\tau|}) \frac{\mathbf{m}^2}{\boldsymbol{\mu}} \quad (35)$$

The characteristic relaxation timescale of the system is

$$\tau_0 = \frac{1}{\mu} = \begin{cases} \beta & \beta < 1 \\ 1 & \beta > 1 \end{cases} \quad (36)$$

Here, we see the traditional correspondence between dynamics and statics in equilibrium: the phase transition and its ergodicity breaking is reflected in the correlation function no longer decaying to zero. One can check also that at all temperatures, the dynamics satisfies the fluctuation relation

$$\mathcal{R}(\tau) = -\beta\Theta(\tau)C'(\tau) \quad (37)$$