

2 The Curie–Weiss Model

In statistical mechanics, a *mean-field approximation* is often used to approximate a model by a simpler one, whose global behavior can be studied with the help of *explicit* computations. The information thus extracted can then be used as an indication of the kind of properties that can be expected from the original model. In addition, this approximation turns out to provide quantitatively correct results in sufficiently high dimensions.

The Ising model, which will guide us throughout the book, is a classical example of a model with a rich behavior but with no explicit solution in general (the exceptions being the one-dimensional model, see Section 3.3, and the two-dimensional model when $h = 0$). In this chapter, we consider its mean-field approximation, in the form of the *Curie–Weiss model*. Although it is an over-simplification of the Ising model, the Curie–Weiss model still displays a phase transition, with distinct behaviors at high and low temperature. It will also serve as an illustration of various techniques and show how the probabilistic behavior is intimately related to the analytic properties of the thermodynamic potentials (free energy and pressure) of the model.

2.1 The mean-field approximation

Consider a system of Ising spins living on \mathbb{Z}^d , described by the Ising Hamiltonian defined in (1.44). In that model, the spin ω_i located at vertex i interacts with the rest of the system via its neighbors. The contribution to the total energy coming from the interaction of ω_i with its $2d$ neighbors can be written as

$$-\beta \sum_{j:j \sim i} \omega_i \omega_j = -2d\beta \omega_i \cdot \frac{1}{2d} \sum_{j:j \sim i} \omega_j. \quad (2.1)$$

Written this way, one can interpret the contribution of ω_i to the total energy as an interaction of ω_i with a local magnetization density, produced by the average of its $2d$ nearest neighbors:

$$\frac{1}{2d} \sum_{j:j \sim i} \omega_j.$$

Of course, this magnetization density is *local* and varies from one point to the other. The **mean-field approximation** consists in assuming that each local magnetization

density can be approximated by the *global* magnetization density,

$$\frac{1}{N} \sum_{j=1}^N \omega_j,$$

where N is the number of spins in the system. The mean-field approximation of the Ising model thus amounts to do the following transformation on the Hamiltonian (1.44) (up to a multiplicative constant that will be absorbed in β):

$$\text{Replace } -\beta \sum_{i \sim j} \omega_i \omega_j \text{ by } -\frac{d\beta}{N} \sum_{i,j} \omega_i \omega_j.$$

The term involving the magnetic field, on the other hand, remains unchanged. This leads to the following definition.

Definition 2.1. The *Curie–Weiss Hamiltonian* for a collection of spins $\omega = (\omega_1, \dots, \omega_N)$ at inverse temperature β and with an external magnetic field h is defined by

$$\mathcal{H}_{N;\beta,h}^{\text{CW}}(\omega) \stackrel{\text{def}}{=} -\frac{d\beta}{N} \sum_{i,j=1}^N \omega_i \omega_j - h \sum_{i=1}^N \omega_i. \quad (2.2)$$

In contrast to those of the Ising model, the interactions of the Curie–Weiss model are global: each spin interacts with all other spins in the same way, and the relative positions of the spins can therefore be ignored. Actually, due to this lack of geometry, one may think of this model as defined on the **complete graph** with N vertices, which has an edge between any pair of distinct vertices:

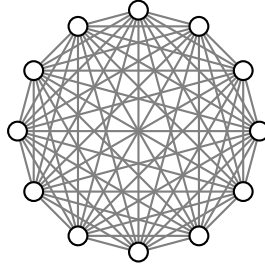


Figure 2.1: The complete graph with 12 vertices. In the Curie–Weiss model, all spins interact: a pair of spins living at vertices i, j contributes to the total energy by an amount $-\frac{d\beta}{N} \omega_i \omega_j$.

We denote by $\Omega_N \stackrel{\text{def}}{=} \{\pm 1\}^N$ the set of all possible configurations of the Curie–Weiss model. The Gibbs distribution on Ω_N is written

$$\mu_{N;\beta,h}^{\text{CW}}(\omega) \stackrel{\text{def}}{=} \frac{e^{-\mathcal{H}_{N;\beta,h}^{\text{CW}}(\omega)}}{Z_{N;\beta,h}^{\text{CW}}}, \quad \text{where} \quad Z_{N;\beta,h}^{\text{CW}} \stackrel{\text{def}}{=} \sum_{\omega \in \Omega_N} e^{-\mathcal{H}_{N;\beta,h}^{\text{CW}}(\omega)}.$$

As mentioned in the introduction, the expectation (or average) of an observable $f : \Omega_N \rightarrow \mathbb{R}$ under $\mu_{N;\beta,h}^{\text{CW}}$ will be denoted by $\langle f \rangle_{N;\beta,h}^{\text{CW}}$.

Our aim, in the rest of the chapter, is to show that the Curie–Weiss model exhibits paramagnetic behavior at high temperature and ferromagnetic behavior at low temperature.

2.2 The behavior for large N when $h = 0$

We will first study the model in the absence of a magnetic field. The same heuristic arguments given in Section 1.4.3 for the Ising model also apply here. For instance, when $h = 0$, the Hamiltonian is invariant under the global spin flip $\omega \mapsto -\omega$ (which changes each ω_i into $-\omega_i$), which implies that the magnetization density

$$m_N \stackrel{\text{def}}{=} \frac{M_N}{N}, \quad \text{where} \quad M_N \stackrel{\text{def}}{=} \sum_{i=1}^N \omega_i,$$

has a symmetric distribution: $\mu_{N;\beta,0}^{\text{CW}}(m_N = -m) = \mu_{N;\beta,0}^{\text{CW}}(m_N = +m)$. In particular,

$$\langle m_N \rangle_{N;\beta,0}^{\text{CW}} = 0. \quad (2.3)$$

As discussed in Section 1.4.3, we expect that the spins should be essentially independent when β is small, but that, when β is large, the most probable configurations should have most spins equal and thus be close to one of the two ground states, in which all spins are equal. The following theorem confirms these predictions.

Theorem 2.2. ($h = 0$) Let $\beta_c = \beta_c(d) \stackrel{\text{def}}{=} \frac{1}{2d}$. Then, the following holds.

1. When $\beta \leq \beta_c$, the magnetization concentrates at zero: for all $\epsilon > 0$, there exists $c = c(\beta, \epsilon) > 0$ such that, for large enough N ,

$$\mu_{N;\beta,0}^{\text{CW}}(m_N \in (-\epsilon, \epsilon)) \geq 1 - 2e^{-cN}.$$

2. When $\beta > \beta_c$, the magnetization is bounded away from zero. More precisely, there exists $m^{*,\text{CW}}(\beta) > 0$, called the **spontaneous magnetization**, such that, for all small enough $\epsilon > 0$, there exists $b = b(\beta, \epsilon) > 0$ such that if

$$J_*(\epsilon) \stackrel{\text{def}}{=} (-m^{*,\text{CW}}(\beta) - \epsilon, -m^{*,\text{CW}}(\beta) + \epsilon) \cup (m^{*,\text{CW}}(\beta) - \epsilon, m^{*,\text{CW}}(\beta) + \epsilon),$$

then, for large enough N ,

$$\mu_{N;\beta,0}^{\text{CW}}(m_N \in J_*(\epsilon)) \geq 1 - 2e^{-bN}.$$

β_c is called the **inverse critical temperature** or **inverse Curie temperature**

In other words, when N is large,

$$\forall \beta \leq \beta_c, \quad m_N \simeq 0 \quad \text{with high probability,}$$

whereas

$$\forall \beta > \beta_c, \quad m_N \simeq \begin{cases} +m^{*,\text{CW}}(\beta) & \text{with probability close to } \frac{1}{2}, \\ -m^{*,\text{CW}}(\beta) & \text{with probability close to } \frac{1}{2}. \end{cases}$$

This behavior is understood easily by simply plotting the distribution of m_N :

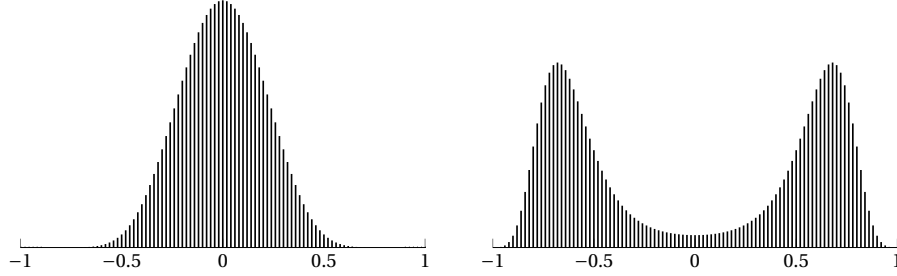


Figure 2.2: The distribution of the magnetization of the Curie–Weiss model, $\mu_{N;\beta,0}^{\text{CW}}(m_N = \cdot)$, with $N = 100$ spins, when $h = 0$, plotted using (2.9) below.

At high temperature (on the left, $2d\beta = 0.8$), m_N concentrates around zero. At low temperature (on the right, $2d\beta = 1.2$), the distribution of m_N becomes bimodal, with two peaks near $\pm m^{*,\text{CW}}(\beta)$. In both cases, $\langle m_N \rangle_{N;\beta,0}^{\text{CW}} = 0$. The width of the peaks in the above pictures tends to 0 when $N \rightarrow \infty$, which means that

$$\lim_{N \rightarrow \infty} \mu_{N;\beta,0}^{\text{CW}}(m_N \in \cdot) = \begin{cases} \delta_0(\cdot) & \text{if } \beta \leq \beta_c, \\ \frac{1}{2}(\delta_{+m^{*,\text{CW}}(\beta)}(\cdot) + \delta_{-m^{*,\text{CW}}(\beta)}(\cdot)) & \text{if } \beta > \beta_c, \end{cases}$$

where δ_m is the Dirac mass at m (that is, the probability measure on $[-1, 1]$ such that $\delta_m(A) = 1$ or 0 , depending on whether A contains m or not.)

Remark 2.3. We emphasize that, when $\beta > \beta_c$ and N is large, the above results say that the typical values of the magnetization observed when sampling a configuration are close to either $+m^{*,\text{CW}}(\beta)$ or $-m^{*,\text{CW}}(\beta)$. Of course, this does not contradict the fact that it is always zero *on average*: $\langle m_N \rangle_{N;\beta,0}^{\text{CW}} = 0$. The proper interpretation of the latter average comes from the Law of Large Numbers. Namely, let us fix N and sample an infinite sequence of independent realizations of the magnetization density: $m_N^{(1)}, m_N^{(2)}, \dots$, each distributed according to $\mu_{N;\beta,0}^{\text{CW}}$. Then, by the Strong Law of Large Numbers, the empirical average over the n first samples converges almost surely to zero as $n \rightarrow \infty$:

$$\frac{m_N^{(1)} + \dots + m_N^{(n)}}{n} \longrightarrow \langle m_N \rangle_{N;\beta,0}^{\text{CW}} = 0.$$

There is another natural Law of Large numbers that one might be interested in this context. When $\beta = 0$, the random variables

$$\sigma_i(\omega) \stackrel{\text{def}}{=} \omega_i$$

are independent Bernoulli random variables of mean 0, which also satisfy a Law of Large Numbers: their empirical average $\frac{1}{N} \sum_{i=1}^N \sigma_i$ converges to $\langle \sigma_1 \rangle_{N;0,0}^{\text{CW}} = 0$ in probability. One might thus wonder whether this property survives the introduction of an interaction between the spins: $\beta > 0$. Since $\frac{1}{N} \sum_{i=1}^N \sigma_i = m_N$, Theorem 2.2 shows that this is the case if and only if $\beta \leq \beta_c$. \diamond

To the inverse critical temperature β_c corresponds the critical temperature $T_c = \frac{1}{\beta_c}$. The range $T > T_c$ (that is, $\beta < \beta_c$) is called the **supercritical regime**, while the range $T < T_c$ (that is, $\beta > \beta_c$) is the **subcritical regime**, also called the regime of **phase coexistence**. The value $T = T_c$ corresponds to the **critical regime**. The result above shows that the Curie–Weiss model is not ordered in this regime; it can be shown, however, that the magnetization possesses peculiar properties at T_c , such as non-Gaussian fluctuations.

The Curie–Weiss model possesses a remarkable feature, which makes its analysis much easier than that of the Ising model on \mathbb{Z}^d : since

$$\sum_{i,j=1}^N \omega_i \omega_j = \left(\sum_{i=1}^N \omega_i \right)^2 \equiv M_N^2,$$

the Hamiltonian $\mathcal{H}_{N;\beta,0}^{\text{CW}}$ is entirely determined by the magnetization density:

$$\mathcal{H}_{N;\beta,0}^{\text{CW}} = -d\beta m_N^2 N. \quad (2.4)$$

This property will make it possible to compute explicitly the thermodynamic potentials (and other quantities) associated to the Curie–Weiss model.

The thermodynamic potential that plays the central role in the study of the Curie–Weiss model is the free energy:

Definition 2.4. Let $e(m) \stackrel{\text{def}}{=} -dm^2$ and

$$s(m) \stackrel{\text{def}}{=} -\frac{1-m}{2} \log \frac{1-m}{2} - \frac{1+m}{2} \log \frac{1+m}{2}.$$

Then

$$f_\beta^{\text{CW}}(m) \stackrel{\text{def}}{=} \beta e(m) - s(m) \quad (2.5)$$

is called the **free energy** of the Curie–Weiss model.

The claims of Theorem 2.2 will be a direct consequence of the following proposition, which shows the role played by the free energy in the asymptotic distribution of the magnetization.

Proposition 2.5. For any β ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Z_{N;\beta,0}^{\text{CW}} = - \min_{m \in [-1,1]} f_\beta^{\text{CW}}(m). \quad (2.6)$$

Moreover, for any interval $J \subset [-1, 1]$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mu_{N;\beta,0}^{\text{CW}}(m_N \in J) = - \inf_{m \in J} I_\beta^{\text{CW}}(m), \quad (2.7)$$

where

$$I_\beta^{\text{CW}}(m) \stackrel{\text{def}}{=} f_\beta^{\text{CW}}(m) - \min_{\tilde{m} \in [-1,1]} f_\beta^{\text{CW}}(\tilde{m}). \quad (2.8)$$



One can write (2.7) roughly as follows:

$$\text{For large } N, \quad \mu_{N;\beta,0}^{\text{CW}}(m_N \in J) \simeq \exp\left(-\left\{\inf_{m \in J} I_\beta^{\text{CW}}(m)\right\}N\right).$$

(In the language of large deviations theory, I_β^{CW} is called a **rate function**.) Notice that $I_\beta^{\text{CW}} \geq 0$ and

$$\min_{m \in [-1, 1]} I_\beta^{\text{CW}}(m) = 0.$$

Thus, if $J \subset [-1, 1]$ is such that I_β^{CW} is uniformly strictly positive on J ,

$$\inf_{m \in J} I_\beta^{\text{CW}}(m) > 0,$$

then $\mu_{N;\beta,0}^{\text{CW}}(m_N \in J)$ converges to zero exponentially fast when $N \rightarrow \infty$, meaning that the magnetization is very likely to take values outside J . This shows that the typical values of the magnetization correspond to the regions where I_β^{CW} vanishes. \diamond

Proof of Proposition 2.5: Observe that for a fixed N , m_N is a random variable taking values in the set

$$\mathcal{A}_N \stackrel{\text{def}}{=} \left\{ -1 + \frac{2k}{N} : k = 0, \dots, N \right\} \subset [-1, 1].$$

Let $J \subset [-1, 1]$ be an interval. Then,

$$\mu_{N;\beta,0}^{\text{CW}}(m_N \in J) = \sum_{m \in J \cap \mathcal{A}_N} \mu_{N;\beta,0}^{\text{CW}}(m_N = m).$$

Since there are exactly $\binom{N}{\frac{1+m}{2}N}$ configurations $\omega \in \Omega_N$ which have $m_N(\omega) = m$, one can express explicitly the distribution of m_N using (2.4):

$$\mu_{N;\beta,0}^{\text{CW}}(m_N = m) = \sum_{\substack{\omega \in \Omega_N: \\ m_N(\omega) = m}} \frac{e^{-\mathcal{H}_{N;\beta,0}^{\text{CW}}(\omega)}}{Z_{N;\beta,0}^{\text{CW}}} = \frac{1}{Z_{N;\beta,0}^{\text{CW}}} \binom{N}{\frac{1+m}{2}N} e^{d\beta m^2 N}. \quad (2.9)$$

In the same way,

$$Z_{N;\beta,0}^{\text{CW}} = \sum_{m \in \mathcal{A}_N} \binom{N}{\frac{1+m}{2}N} e^{d\beta m^2 N}. \quad (2.10)$$

Since we are interested in its behavior on the exponential scale and since it is a sum of only $|\mathcal{A}_N| = N + 1$ positive terms, $Z_{N;\beta,0}^{\text{CW}}$ can be estimated by keeping only its *dominant term*:

$$\max_{m \in \mathcal{A}_N} \binom{N}{\frac{1+m}{2}N} e^{d\beta m^2 N} \leq Z_{N;\beta,0}^{\text{CW}} \leq (N + 1) \max_{m \in \mathcal{A}_N} \binom{N}{\frac{1+m}{2}N} e^{d\beta m^2 N}.$$

To study the large N behavior of the binomial factors, we use Stirling's Formula. The latter implies the existence of two constants $c_-, c_+ > 0$ such that, for all $m \in \mathcal{A}_N$,

$$c_- N^{-1/2} e^{Ns(m)} \leq \binom{N}{\frac{1+m}{2}N} \leq c_+ e^{Ns(m)}. \quad (2.11)$$

Exercise 2.1. Verify (2.11).

We can thus compute an upper bound as follows:

$$\begin{aligned} Z_{N;\beta,0}^{\text{CW}} &\leq c_+ (N + 1) \exp\left(N \max_{m \in \mathcal{A}_N} \{d\beta m^2 + s(m)\}\right) \\ &\leq c_+ (N + 1) \exp\left(-N \min_{m \in [-1, 1]} I_\beta^{\text{CW}}(m)\right), \end{aligned}$$

which yields

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log Z_{N;\beta,0}^{\text{CW}} \leq - \min_{m \in [-1,1]} f_{\beta}^{\text{CW}}(m).$$

For the lower bound, we first use the continuity of $m \mapsto \{d\beta m^2 + s(m)\}$ on $[-1, 1]$ and consider some $m' \in [-1, 1]$ for which $f_{\beta}^{\text{CW}}(m') = \min_m f_{\beta}^{\text{CW}}(m)$. Fix $\epsilon > 0$, and choose some $m \in \mathcal{A}_N$ such that $|f_{\beta}^{\text{CW}}(m) - f_{\beta}^{\text{CW}}(m')| \leq \epsilon$, which is always possible once N is large enough. We then have

$$Z_{N;\beta,0}^{\text{CW}} \geq c_- N^{-1/2} \exp\{-N(f_{\beta}^{\text{CW}}(m') + \epsilon)\}.$$

This yields

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \log Z_{N;\beta,0}^{\text{CW}} \geq - \min_{m \in [-1,1]} f_{\beta}^{\text{CW}}(m) - \epsilon.$$

Since ϵ was arbitrary, (2.6) follows.

A similar computation can be done for the sum over $m \in J \cap \mathcal{A}_N$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \sum_{m \in J \cap \mathcal{A}_N} \left(\frac{N}{\frac{1+m}{2} N} \right) e^{d\beta m^2 N} = - \inf_{m \in J} f_{\beta}^{\text{CW}}(m),$$

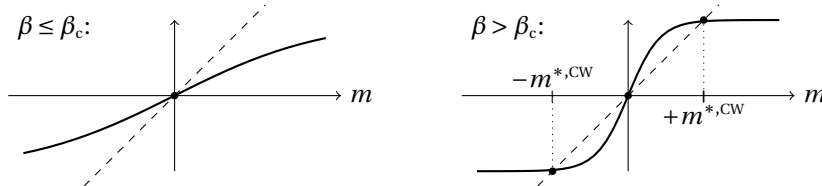
and we get (2.7). \square

Proof of Theorem 2.2. As discussed after the statement of Proposition 2.5, we must locate the zeros of I_{β}^{CW} . Since the latter is smooth and $I_{\beta}^{\text{CW}} \geq 0$, the zeros correspond to the solutions of $\frac{\partial I_{\beta}^{\text{CW}}}{\partial m} = 0$. After a straightforward computation, we easily see that this condition is equivalent to the **mean-field equation**:

$$\tanh(2d\beta m) = m. \quad (2.12)$$

Since $\lim_{m \rightarrow \pm\infty} \tanh(\beta m) = \pm 1$, there always exists at least one solution and, as an analysis of the graph of $m \mapsto \tanh(\beta m)$ shows (see below), the number of solutions of (2.12) depends on whether $2d\beta$ is larger or smaller than 1, that is, whether β is larger or smaller than β_c .

On the one hand, when $\beta \leq \beta_c$, (2.12) has a unique solution, given by $m = 0$. On the other hand, when $\beta > \beta_c$, there are two additional non-trivial solutions, $+m^{*,\text{CW}}(\beta)$ and $-m^{*,\text{CW}}(\beta)$ (which depend on β):



The trivial solution $m = 0$ is a local maximum of I_{β}^{CW} , whereas $+m^{*,\text{CW}}(\beta)$ and $-m_{\beta}^{*,\text{CW}}$ are global minima (see Figure 2.3).

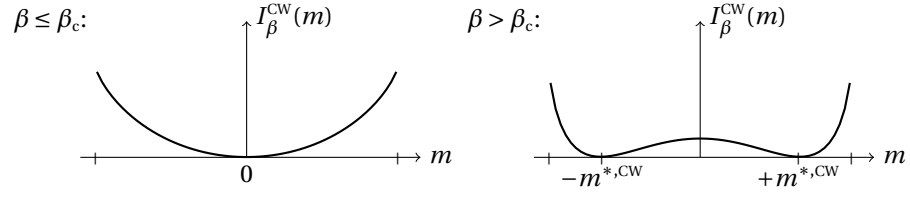


Figure 2.3: The rate function of the Curie–Weiss model. The values taken by the magnetization density of a very large system lie in a neighborhood of the points m at which I_β^{CW} vanishes, with probability very close to 1, as seen in Figure 2.2. In the supercritical and critical phases ($\beta \leq \beta_c$), there exists a unique global minimum $m = 0$. In the subcritical phase ($\beta > \beta_c$), there exist two non-zero typical values $\pm m^{*,\text{CW}}(\beta)$: there is a *phase transition* at β_c .

Combined with (2.7), this analysis proves the theorem. \square

Remark 2.6. One clearly sees from the graphical characterization of $m^{*,\text{CW}}(\beta)$ that

$$m^{*,\text{CW}}(\beta) \downarrow 0 \quad \text{as} \quad \beta \downarrow \beta_c^{\text{CW}}. \quad (2.13)$$

A more quantitative analysis is provided in Section 2.5.3. \diamond

The above analysis revealed that the typical values of the magnetization of the model are those near which the function I_β^{CW} vanishes. Since I_β^{CW} differs from f_β^{CW} only by a constant, this means that *the typical values of the magnetization are those that minimize the free energy*, a property typical of the thermodynamic behavior studied in Section 1.1.5 (when letting a system exchange energy with a heat reservoir), or as was already derived non-rigorously in Section 1.3.1.

The bifurcation of the typical values taken by the magnetization in the Curie–Weiss model at low temperature originated in the appearance of two global minima in the free energy. On the one hand, $e(m) = -dm^2$ is the **energy density** associated to configurations of magnetization density m ; it is minimal when $m = +1$ or -1 (all spins equal). On the other hand, $s(m)$ is the **entropy density**, which measures the number of configurations with a magnetization density m ; it is maximal at $m = 0$ (equal proportions of $+$ and $-$ spins). Since $\beta e(m)$ and $s(m)$ are both concave, the convexity/concavity properties of their difference depend on the temperature. When β is small, entropy dominates and $f_\beta^{\text{CW}}(m)$ is strictly convex. When β is large, energy starts to play a major role by favoring configurations with small energy: $f_\beta^{\text{CW}}(m)$ is not convex and has two global minima.

As already mentioned in Chapter 1, this interplay between energy and entropy is fundamental in the mechanism leading to phase transition.

Remark 2.7. The non-convex free energy observed at low temperature in the Curie–Weiss model is a consequence of the lack of geometry in the model. As will be seen later, the free energy of more realistic systems (such as the Ising model on \mathbb{Z}^d , or the lattice gas of Chapter 4) is always convex. \diamond

Exercise 2.2. Let $\zeta : \mathbb{N} \rightarrow \mathbb{R}_{>0}$. Consider the following modification of the Curie–Weiss Hamiltonian (with $h = 0$):

$$\widetilde{\mathcal{H}}_{N;\beta,0}(\omega) \stackrel{\text{def}}{=} -\frac{\beta}{\zeta(N)} \sum_{i,j=1}^N \omega_i \omega_j.$$

Denote by $\tilde{\mu}_{N;\beta}$ the corresponding Gibbs distribution. Show that the following holds.

1. If $\lim_{N \rightarrow \infty} \frac{\zeta(N)}{N} = \infty$, then m_N tends to 0 in probability for all $\beta \geq 0$.
2. If $\lim_{N \rightarrow \infty} \frac{\zeta(N)}{N} = 0$, then $|m_N|$ tends to 1 in probability for all $\beta > 0$.

This shows that the only scaling leading to a nontrivial dependence in β is when $\zeta(N)$ is of the order of N .

2.3 The behavior for large N when $h \neq 0$

In the presence of an external magnetic field h , the analysis is similar. The relevant thermodynamic potential associated to the magnetic field is the *pressure*.

Theorem 2.8. *The pressure*

$$\psi_\beta^{\text{CW}}(h) \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} \frac{1}{N} \log Z_{N;\beta,h}^{\text{CW}}.$$

exists and is convex in h . Moreover, it equals the **Legendre transform of the free energy**:

$$\psi_\beta^{\text{CW}}(h) = \max_{m \in [-1,1]} \{hm - f_\beta^{\text{CW}}(m)\}. \quad (2.14)$$

Proof. We start by decomposing the partition function as in (2.10):

$$Z_{N;\beta,h}^{\text{CW}} = \sum_{m \in \mathcal{A}_N} \sum_{\substack{\omega \in \Omega_N: \\ m_N(\omega) = m}} e^{-\mathcal{H}_{N;\beta,h}^{\text{CW}}(\omega)} = \sum_{m \in \mathcal{A}_N} \binom{N}{\frac{1+m}{2}N} e^{(hm + d\beta m^2)N}.$$

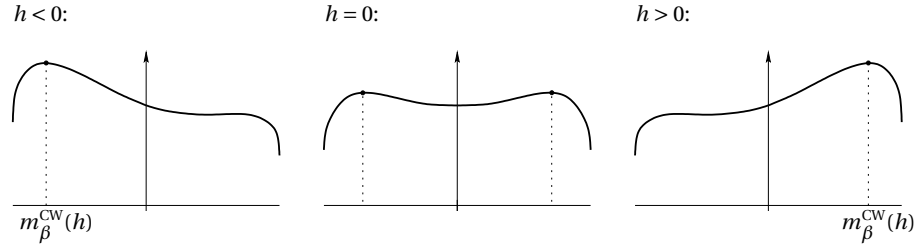
We can then proceed as in the proof of Theorem 2.2. For example,

$$Z_{N;\beta,h}^{\text{CW}} \leq \frac{c_+(N+1)}{\sqrt{N}} \exp \left\{ N \max_{m \in [-1,1]} \{hm - f_\beta^{\text{CW}}(m)\} \right\}.$$

A lower bound of the same type is not difficult to establish, yielding (2.14) in the limit $N \rightarrow \infty$. As shown in Appendix B.2.3, a Legendre transform is always convex. \square

We first investigate the behavior of the pressure as a function of the magnetic field and later apply it to the study of the typical values of the magnetization density.

Again, since $hm - f_\beta^{\text{CW}}(m)$ is smooth (analytic, in fact) in m , we can find the maximum in (2.14) by explicit differentiation. Before that, let us plot the graph of $m \mapsto hm - f_\beta^{\text{CW}}(m)$ for different values of h (here, at low temperature):



When $h \neq 0$, the supremum of $hm - f_\beta^{\text{CW}}(m)$ is attained at a unique point which we denote by $m_\beta^{\text{CW}}(h)$. This point can be computed by solving $\frac{\partial}{\partial m} \{hm - f_\beta^{\text{CW}}(m)\} = 0$, which is equivalent to $\frac{\partial f_\beta^{\text{CW}}}{\partial m} = h$, and can be written as the modified mean-field equation:

$$\tanh(2d\beta m + h) = m. \quad (2.15)$$

Again, this equation always has at least one solution. Let $\beta_c (= 1/2d)$ denote the inverse critical temperature introduced before. When $\beta < \beta_c$, the solution to (2.15) is unique. When $\beta > \beta_c$, there can be more than one solution, depending on h ; in every case, $m_\beta^{\text{CW}}(h)$ is the largest (resp. smallest) one if $h > 0$ (resp. $h < 0$).

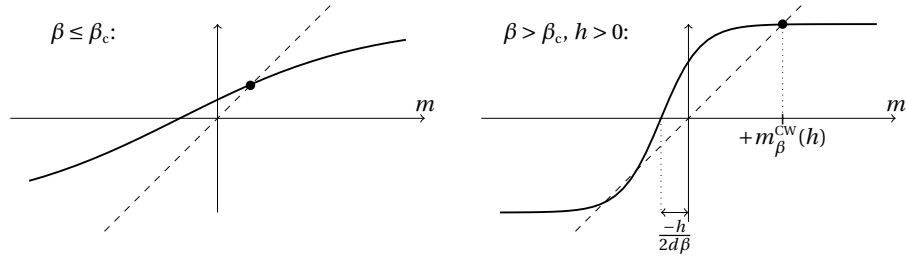


Figure 2.4: Equation (2.15) has a unique solution when $\beta \leq \beta_c$ (left), but up to 3 different solutions when $\beta > \beta_c$ (right) and one must choose the largest (resp. smallest) one when $h > 0$ (resp. $h < 0$).

On the one hand, a glance at the above graph shows that, when $\beta \leq \beta_c$,

$$\lim_{h \downarrow 0} m_\beta^{\text{CW}}(h) = \lim_{h \downarrow 0} m_\beta^{\text{CW}}(h) = 0. \quad (2.16)$$

On the other hand, when $\beta > \beta_c$,

$$\lim_{h \downarrow 0} m_\beta^{\text{CW}}(h) = -m^{*,\text{CW}}(\beta) < +m^{*,\text{CW}}(\beta) = \lim_{h \downarrow 0} m_\beta^{\text{CW}}(h). \quad (2.17)$$

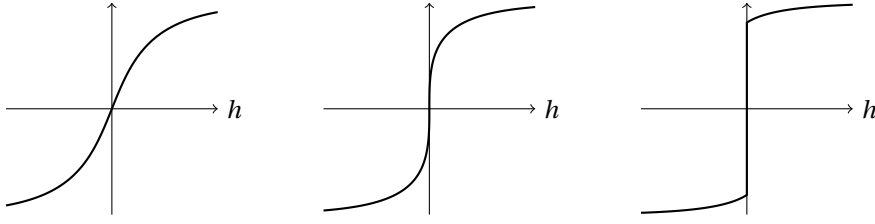


Figure 2.5: The magnetization $h \mapsto m_\beta^{\text{CW}}(h)$ for $\beta < \beta_c$ (on the left), $\beta = \beta_c$ (center) and $\beta > \beta_c$ (on the right). These pictures were made by a numerical study of the solutions of (2.15).

Exercise 2.3. Show that $h \mapsto m_\beta^{\text{CW}}(h)$ is analytic on $(-\infty, 0)$ and on $(0, \infty)$.

Exercise 2.4. Using (2.14), show that the pressure can be written explicitly as

$$\psi_\beta^{\text{CW}}(h) = -d\beta m_\beta^{\text{CW}}(h)^2 + \log \cosh(2d\beta m_\beta^{\text{CW}}(h) + h) + \log 2.$$

Conclude, in particular, that it is analytic on $(-\infty, 0)$ and $(0, +\infty)$.

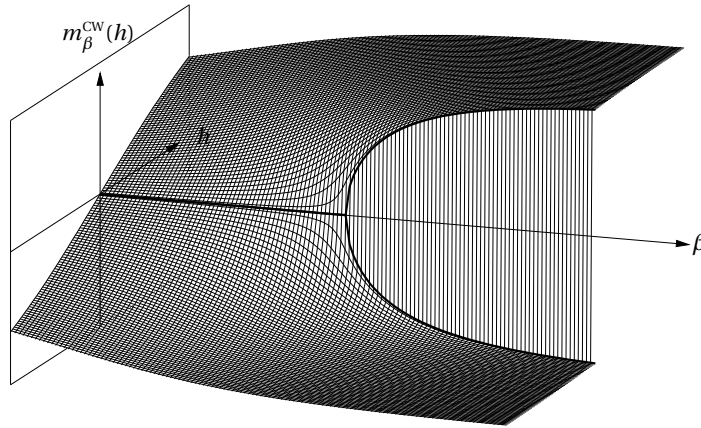


Figure 2.6: The graph of $(\beta, h) \mapsto m_\beta^{\text{CW}}(h)$. At fixed $\beta > 0$, one observes the curves $h \mapsto m_\beta^{\text{CW}}(h)$ of Figure 2.5; in particular these are discontinuous when $\beta > \beta_c$.

Using the terminology of Section 1.4.1, we thus see that the Curie–Weiss model provides a case in which *paramagnetism* is observed at high temperature and *ferromagnetism* at low temperature.

We then move on to the study of the pressure, by first considering non-zero magnetic fields: $h \neq 0$. In this case, we can express $\psi_\beta^{\text{CW}}(h)$ using the Legendre transform:

$$\psi_\beta^{\text{CW}}(h) = h \cdot m_\beta^{\text{CW}}(h) - f_\beta^{\text{CW}}(m_\beta^{\text{CW}}(h)),$$

from which we deduce, using Exercise 2.3 and the analyticity of $m \mapsto f_\beta^{\text{CW}}(m)$ on $(-1, 1)$, that $h \mapsto \psi_\beta^{\text{CW}}(h)$ is analytic on $(-\infty, 0) \cup (0, \infty)$. Differentiating with respect

to h yields, when $h \neq 0$,

$$\frac{\partial \psi_\beta^{\text{CW}}}{\partial h}(h) = m_\beta^{\text{CW}}(h). \quad (2.18)$$

To study the behavior at $h = 0$, we first notice that since ψ_β^{CW} is convex, Theorem B.28 guarantees that its one-sided derivative, $\frac{\partial \psi_\beta}{\partial h^+}|_{h=0}$ (resp. $\frac{\partial \psi_\beta}{\partial h^-}|_{h=0}$) exists and is right-continuous (resp. left-continuous). If $\beta \leq \beta_c$, (2.16) gives

$$\begin{aligned} \frac{\partial \psi_\beta^{\text{CW}}}{\partial h^-}|_{h=0} &= \lim_{h \rightarrow 0^-} \frac{\partial \psi_\beta^{\text{CW}}}{\partial h} = \lim_{h \rightarrow 0^-} m_\beta^{\text{CW}}(h) \\ &= 0 \\ &= \lim_{h \rightarrow 0^+} m_\beta^{\text{CW}}(h) = \lim_{h \rightarrow 0^+} \frac{\partial \psi_\beta^{\text{CW}}}{\partial h} = \frac{\partial \psi_\beta^{\text{CW}}}{\partial h^+}|_{h=0}. \end{aligned}$$

As a consequence, ψ_β^{CW} is differentiable at $h = 0$. Assume then that $\beta > \beta_c$. By (2.18) and (2.17), the same argument yields

$$\frac{\partial \psi_\beta^{\text{CW}}}{\partial h^-}|_{h=0} = -m^{*,\text{CW}}(\beta) < 0 < m^{*,\text{CW}}(\beta) = \frac{\partial \psi_\beta^{\text{CW}}}{\partial h^+}|_{h=0},$$

and so ψ_β^{CW} is not differentiable at $h = 0$.

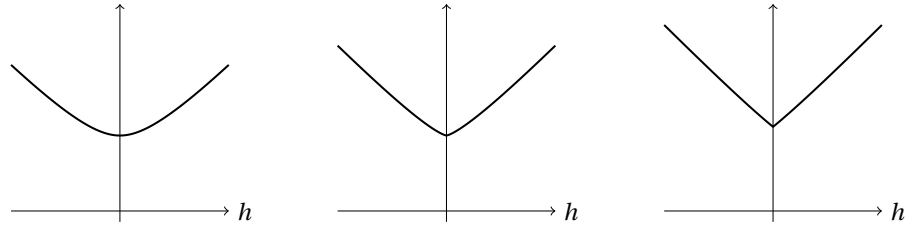


Figure 2.7: The pressure $\psi_\beta^{\text{CW}}(h)$ of the Curie–Weiss model with the same values of β as in Figure 2.5.

Finally, we let the reader verify that, when $h \neq 0$, the magnetization density m_N concentrates exponentially fast on $m_\beta^{\text{CW}}(h)$.

Exercise 2.5. Adapting the analysis of the case $h = 0$, show that an expression of the type (2.7) holds:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mu_{N;\beta,h}^{\text{CW}}(m_N \in J) = - \inf_{m \in J} I_{\beta,h}^{\text{CW}}(m), \quad (2.19)$$

with

$$I_{\beta,h}^{\text{CW}}(m) \stackrel{\text{def}}{=} f_\beta^{\text{CW}}(m) - hm - \min_{\tilde{m} \in [-1,1]} (f_\beta^{\text{CW}}(\tilde{m}) - h\tilde{m}).$$

Show that, when $h \neq 0$, the rate function $I_{\beta,h}^{\text{CW}}(m)$ has a unique global minimum at $m_\beta^{\text{CW}}(h)$, for all $\beta > 0$ (see the figure below).

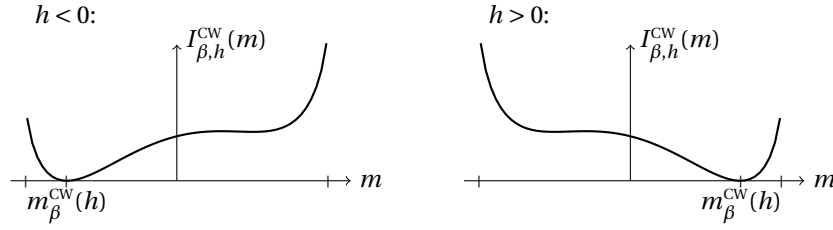


Figure 2.8: The rate function of the Curie–Weiss model with a magnetic field $h \neq 0$ has a unique global minimum at $m_{\beta}^{CW}(h)$.

2.4 Bibliographical references

The Curie–Weiss model, as it is described in this chapter, has been introduced independently by many people, including Temperley [328], Husimi [167] and Kac [183]. There exist numerous mathematical treatments where the interested reader can get much more information, such as Ellis’ book [100].

In our study of the van der Waals model of a gas in Section 4.9, we will reinterpret the Curie–Weiss model as a model of a lattice gas. There, we will derive, in a slightly different language, additional information on the free energy, the pressure and the relationship between these two quantities.

2.5 Complements and further reading

2.5.1 The “naive” mean-field approximation.

This approximation made its first appearance in the early 20th century work of Pierre-Ernest Weiss, based on earlier ideas of Pierre Curie, in which the method now known as *mean-field theory* was developed. The latter is somewhat different from what is done in this chapter, but leads to the same results.

Namely, consider the nearest-neighbor Ising model on $\Lambda \subseteq \mathbb{Z}^d$. The distribution of the spin at the origin, conditionally on the values taken by its neighbors, is given by

$$\mu_{\Lambda;\beta,h}(\sigma_0 = \pm 1 \mid \sigma_j = \omega_j, j \neq 0) = \frac{1}{Z} \exp\{\pm(\beta \sum_{j \sim 0} \omega_j + h)\},$$

where $Z \stackrel{\text{def}}{=} 2 \cosh(\beta \sum_{j \sim 0} \omega_j + h)$ is a normalization factor. The naive mean-field approximation corresponds to assuming that each of the neighboring spins ω_j can be replaced by its mean value m . This yields the following distribution

$$\nu(\sigma_0 = \pm 1) \stackrel{\text{def}}{=} \frac{1}{Z'} \exp\{\pm(2d\beta m + h)\},$$

with the normalization $Z' \stackrel{\text{def}}{=} 2 \cosh(2d\beta m + h)$. The expected value of σ_0 under ν is equal to $\tanh(2d\beta m + h)$. However, for this approximation to be self-consistent, this expected value should also be equal to m . This yields the following consistency condition:

$$m = \tanh(2d\beta m + h),$$

which is precisely (2.15). Notice that the approximation made above, replacing each ω_j by m , seems reasonable in large dimensions, where the average of the $2d$ nearest-neighbor spins is expected to already have a value close to the expected magnetization m .

2.5.2 Alternative approaches to analyze the Curie–Weiss model.

Our analysis of the Curie–Weiss model was essentially combinatorial. We briefly describe two other alternative approaches, whose advantage is to be more readily generalizable to more complex models.

The Hubbard–Stratonovich transformation

The first alternative approach has a more analytic flavor; it relies on the Hubbard–Stratonovich transformation [322, 166].

Observe first that, for any $\alpha > 0$, a simple integration yields

$$\exp\{\alpha x^2\} = \frac{1}{\sqrt{\pi\alpha}} \int_{-\infty}^{\infty} \exp\left\{-\frac{y^2}{\alpha} + 2yx\right\} dy. \quad (2.20)$$

This can be used to express the interactions among spins in the Boltzmann weight as

$$\exp\left\{\frac{d\beta}{N} \left(\sum_{i=1}^N \omega_i\right)^2\right\} = \sqrt{\frac{N}{\pi d\beta}} \int_{-\infty}^{\infty} \exp\left\{-\frac{N}{d\beta} y^2 + 2y \sum_{i=1}^N \omega_i\right\} dy.$$

The advantage of this reformulation is that the quadratic term in the spin variables has been replaced by a linear one. As a consequence, the sum over configurations, in the partition function, can now be performed as in the previous subsection:

$$\begin{aligned} Z_{N;\beta,h}^{\text{CW}} &= \sum_{\omega \in \Omega_N} e^{-\mathcal{H}_{N;\beta,h}^{\text{CW}}(\omega)} = \sqrt{\frac{N}{\pi d\beta}} \int_{-\infty}^{\infty} e^{-Ny^2/d\beta} \prod_{i=1}^N \sum_{\omega_i=\pm 1} \exp\{(2y+h)\omega_i\} dy \\ &= \sqrt{\frac{N}{\pi d\beta}} \int_{-\infty}^{\infty} e^{-N\varphi_{\beta,h}(y)} dy, \end{aligned}$$

where

$$\varphi_{\beta,h}(y) \stackrel{\text{def}}{=} y^2/d\beta - \log(2 \cosh(2y+h)).$$

Exercise 2.6. Show that

$$\lim_{N \rightarrow \infty} \sqrt{N} \int_{-\infty}^{\infty} e^{-N(\varphi_{\beta,h}(y) - \min_y \varphi_{\beta,h}(y))} dy > 0.$$

Hint: Use second-order Taylor expansions of $\varphi_{\beta,h}$ around its minima.

We then obtain

$$\psi_{\beta}^{\text{CW}}(h) = \lim_{N \rightarrow \infty} \frac{1}{N} \log Z_{N;\beta,h}^{\text{CW}} = -\min_y \varphi_{\beta,h}(y),$$

and leave it as an exercise to check that this expression coincides with the one given in Exercise 2.4 (it might help to minimize over $m \stackrel{\text{def}}{=} y/d\beta$).

Stein's methods for exchangeable pairs

The second alternative approach we mention, which is more probabilistic, relies on *Stein's method for exchangeable pairs*. We only describe how it applies to the Curie–Weiss model and refer the reader to Chatterjee's paper [66] for more information.

We start by defining a probability measure P on $\Omega_N \times \Omega_N$ by sampling (ω, ω') as follows: (i) ω is sampled according to the Gibbs distribution $\mu_{N;\beta,h}^{\text{CW}}$; (ii) an index $I \in \{1, \dots, N\}$ is sampled uniformly (with probability $\frac{1}{N}$); (iii) we set $\omega'_j = \omega_j$, for all $j \neq I$, and then let ω'_I be distributed according to $\mu_{N;\beta,h}^{\text{CW}}$, conditionally on the other spins ω_j , $j \neq I$. That is, $\omega'_I = +1$ with probability

$$\frac{\exp(2d\beta\check{m}_I + h)}{\exp(2d\beta\check{m}_I + h) + \exp(-2d\beta\check{m}_I - h)},$$

where

$$\check{m}_i = \check{m}_i(\omega) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{j \neq i} \omega_j.$$

The reader can easily check that the pair (ω, ω') is **exchangeable**:

$$P((\omega, \omega')) = P((\omega', \omega)) \quad \text{for all } (\omega, \omega') \in \Omega_N \times \Omega_N.$$

Let $F(\omega, \omega') \stackrel{\text{def}}{=} \sum_{i=1}^N (\omega_i - \omega'_i)$. The pairs (ω, ω') with $P((\omega, \omega')) > 0$ differ on at most one vertex, and so $|F(\omega, \omega')| \leq 2$. Denoting by E the expectation with respect to P , let

$$f(\omega) \stackrel{\text{def}}{=} E[F(\omega, \omega') \mid \omega] = \frac{1}{N} \sum_{i=1}^N \{\omega_i - \tanh(2d\beta\check{m}_i(\omega) + h)\}.$$

Again, for a pair (ω, ω') with non-zero probability,

$$|f(\omega) - f(\omega')| \leq \frac{2 + 4d\beta}{N}.$$

(We used $|\tanh x - \tanh y| \leq |x - y|$.) The next crucial observation is the following: for any function g on Ω_N ,

$$E(f(\omega)g(\omega)) = E(F(\omega, \omega')g(\omega)) = E(F(\omega', \omega)g(\omega')) = -E(F(\omega, \omega')g(\omega')).$$

We used the tower property of conditional expectation in the first identity and exchangeability in the second. Combining the first and last identities,

$$E[f(\omega)g(\omega)] = \frac{1}{2} E[F(\omega, \omega')(g(\omega) - g(\omega'))].$$

In particular, since each element of the pair (ω, ω') has distribution $\mu_{N;\beta,h}^{\text{CW}}$, and since $E[f] = 0$,

$$\text{Var}_{N;\beta,h}^{\text{CW}}(f) = E[f(\omega)^2] = \frac{1}{2} E[F(\omega, \omega')(f(\omega) - f(\omega'))] \leq \frac{2 + 4d\beta}{N}.$$

Therefore, by Chebyshev's inequality (B.18), for all $\epsilon > 0$,

$$\mu_{N;\beta,h}^{\text{CW}}(|f| > \epsilon) \leq \frac{2 + 4d\beta}{N\epsilon^2}.$$

Finally, since $|\tilde{m}_i(\omega) - m_N(\omega)| \leq 1/N$ for all i ,

$$|f(\omega) - \{m_N(\omega) - \tanh(2d\beta m_N(\omega) + h)\}| \leq \frac{2d\beta}{N},$$

from which we conclude that, for all N large enough,

$$\mu_{N;\beta,h}^{\text{CW}}(|m_N(\omega) - \tanh(2d\beta m_N(\omega) + h)| > 2\epsilon) \leq \mu_{N;\beta,h}^{\text{CW}}(|f| > \epsilon) \leq \frac{2+4d\beta}{N\epsilon^2}.$$

This implies that the magnetization density m_N concentrates, as $N \rightarrow \infty$, on the solution of (2.15). Further refinements can be found in [66], such as much stronger concentration bounds and the computation of the distribution of the fluctuations of the magnetization density in the limit $N \rightarrow \infty$.

2.5.3 Critical exponents

As we have seen in this chapter, the Curie–Weiss model exhibits two types of phase transitions:

- When $\beta > \beta_c = 1/2d$, the magnetization, as a function of h , is discontinuous at $h = 0$: there is a *first-order phase transition*.
- When $h = 0$, the magnetization, as a function of β is continuous, but not analytic at β_c : there is a *continuous phase transition*.

It turns out that the behavior of statistical mechanical systems at continuous phase transitions displays remarkable properties, which will be briefly described in Section 3.10.11. In particular, the different models of statistical mechanics fall into broad *universality classes*, in which all models share the same type of *critical behavior*, characterized by their *critical exponents*.

In this section, we will take a closer look at the critical behavior of the Curie–Weiss model in the neighborhood of the point $\beta = \beta_c, h = 0$ at which a continuous phase transition takes place. This will be done by defining certain critical exponents associated to the model. Having the graph of Figure 2.6 in mind might help the reader understand the definitions of these exponents.

To start, let us approach the transition point by varying the temperature from high to low. We know that $h \mapsto m_\beta^{\text{CW}}(h)$ is continuous at $h = 0$ when $\beta \leq \beta_c$, but discontinuous when $\beta > \beta_c$. We can consider this phenomenon from different points of view, each associated to a way of fixing one variable and varying the other. First, one can see how the derivative of $m_\beta^{\text{CW}}(h)$ with respect to h at $h = 0$ diverges as $\beta \uparrow \beta_c$. Let us thus consider the **magnetic susceptibility** , ,

$$\chi(\beta) \stackrel{\text{def}}{=} \left. \frac{\partial m_\beta^{\text{CW}}(h)}{\partial h} \right|_{h=0},$$

which is well defined for all $\beta < \beta_c$. Since $\chi(\beta)$ must diverge when $\beta \uparrow \beta_c$, one might expect a singular behavior of the form

$$\chi(\beta) \sim \frac{1}{(\beta_c - \beta)^\gamma}, \quad \text{as } \beta \uparrow \beta_c, \quad (2.21)$$

for some constant $\gamma > 0$. More precisely, the last display should be understood in terms of the following limit:

$$\gamma \stackrel{\text{def}}{=} - \lim_{\beta \downarrow \beta_c} \frac{\log \chi(\beta)}{\log(\beta_c - \beta)}.$$

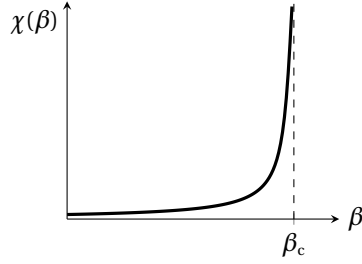


Figure 2.9: Magnetic susceptibility of the Curie–Weiss model.

On the other hand, one can fix $\beta = \beta_c$ and consider the fast variation of the magnetization at $h = 0$:

$$m_{\beta_c}^{\text{CW}}(h) \sim h^{1/\delta}, \quad \text{as } h \downarrow 0, \quad (2.22)$$

with δ defined by

$$\delta^{-1} \stackrel{\text{def}}{=} \lim_{h \downarrow 0} \frac{\log m_{\beta_c}^{\text{CW}}(h)}{\log h}.$$

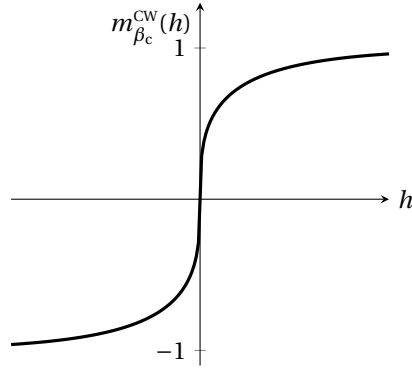


Figure 2.10: Magnetization of the Curie–Weiss model as a function of h at β_c .

But one can also approach the transition by varying the temperature from low to high. So, for $\beta > \beta_c$, consider the magnetization $m^{*,\text{CW}}(\beta)$, in the vicinity of β_c . We have already seen in Remark 2.6 that $m^{*,\text{CW}}(\beta)$ vanishes as $\beta \downarrow \beta_c$, and one is naturally led to expect some behavior of the type:

$$m^{*,\text{CW}}(\beta) \sim (\beta - \beta_c)^b, \quad (2.23)$$

with b defined by

$$b \stackrel{\text{def}}{=} \lim_{\beta \downarrow \beta_c} \frac{\log m^{*,\text{CW}}(\beta)}{\log(\beta - \beta_c)}.$$

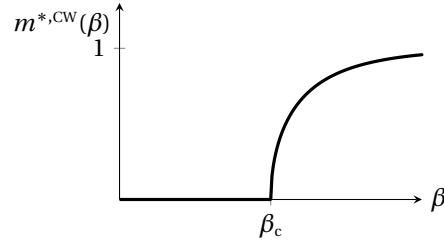


Figure 2.11: The spontaneous magnetization of the Curie–Weiss model at $h = 0$, as a function of β .

(Usually, the letter used for b is β , but we rather use b for obvious reasons.)

Let us introduce one last pair of exponents. This time, in order for our definition to match the standard one in physics, we consider the dependence on the temperature $T = \beta^{-1}$ and on H , defined by $h \equiv \beta H$. Let us define the **internal average energy density** by

$$u(T, H) \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} \frac{1}{\beta N} \langle \mathcal{H}_{N; \beta, \beta H}^{\text{CW}} \rangle_{N; \beta, \beta H}^{\text{CW}},$$

and define the **heat capacity**

$$c_H(\beta) \stackrel{\text{def}}{=} \frac{\partial u}{\partial T}.$$

The exponents α and α' are defined through

$$c_{H=0}(\beta) \sim \begin{cases} (\beta_c - \beta)^{-\alpha} & \text{as } \beta \uparrow \beta_c, \\ (\beta - \beta_c)^{-\alpha'} & \text{as } \beta \downarrow \beta_c, \end{cases}$$

or, more precisely, by

$$\alpha \stackrel{\text{def}}{=} - \lim_{\beta \uparrow \beta_c} \frac{\log c_{H=0}(\beta)}{\log(\beta_c - \beta)},$$

and similarly for α' .

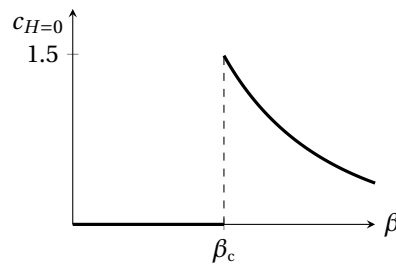


Figure 2.12: Heat capacity of the Curie–Weiss model at $H = 0$.

The numbers $\alpha, \alpha', b, \gamma, \delta$ are examples of **critical exponents**. Similar exponents can be defined for any model at a continuous phase transition, but are usually difficult to compute. These exponents can vary from one model to the other, but coincide for models belonging to the same universality class. We have seen, for instance, in Exercise 1.14, that $b = 1/8$ for the two-dimensional Ising model (more information on this topic can be found in Section 3.10.11).

Theorem 2.9. *For the Curie–Weiss model,*

$$\alpha = \alpha' = 0, \quad b = \frac{1}{2}, \quad \gamma = 1, \quad \delta = 3.$$

Proof. We start with b . Since $m^{*,\text{CW}}(\beta) > 0$ is the largest solution of (2.12) and since $\beta m^{*,\text{CW}}(\beta)$ is small when β is sufficiently close to β_c , we can use a Taylor expansion for $\tanh(\cdot)$:

$$\begin{aligned} m^{*,\text{CW}}(\beta) &= \tanh(2d\beta m^{*,\text{CW}}(\beta)) \\ &= 2d\beta m^{*,\text{CW}}(\beta) - \frac{1}{3}(2d\beta m^{*,\text{CW}}(\beta))^3 + O((\beta m^{*,\text{CW}}(\beta))^5) \\ &= 2d\beta m^{*,\text{CW}}(\beta) - (1 + o(1)) \frac{(2d\beta)^3}{3} (m^{*,\text{CW}}(\beta))^3, \end{aligned}$$

where $o(1)$ tends to zero when $\beta \downarrow \beta_c$. We thus get

$$m^{*,\text{CW}}(\beta) = (1 + o(1)) \left(\frac{3(\beta - \beta_c)}{4d^2\beta^3} \right)^{1/2}, \quad (2.24)$$

(using the fact that $\beta_c = 1/2d$) which shows that $b = 1/2$.

To study $\chi(\beta)$ with $\beta < \beta_c$, we start by the definition of $m_\beta^{\text{CW}} = m_\beta^{\text{CW}}(h)$, as the unique solution to the mean-field equation (2.15), which we differentiate implicitly with respect to h , to obtain

$$\chi(\beta) = \frac{\partial m_\beta^{\text{CW}}}{\partial h} \Big|_{h=0} = \frac{1 - \tanh^2(2d\beta m_\beta^{\text{CW}})}{1 - 2d\beta(1 - \tanh^2(2d\beta m_\beta^{\text{CW}}))} \Big|_{h=0} = \frac{\beta_c}{\beta_c - \beta},$$

which shows that $\gamma = 1$.

Let us now turn to the exponents α, α' . The internal energy density of the Curie–Weiss model at $H = 0$ is given by

$$u = -d \lim_{N \rightarrow \infty} \langle m_N^2 \rangle_{N;\beta,0}^{\text{CW}}.$$

Now, by Theorem 2.2,

$$\lim_{N \rightarrow \infty} \langle m_N^2 \rangle_{N;\beta,0}^{\text{CW}} = \begin{cases} 0 & \text{if } \beta < \beta_c, \\ m^{*,\text{CW}}(\beta)^2 & \text{if } \beta > \beta_c. \end{cases}$$

We immediately deduce that $u = 0$ when $\beta < \beta_c$ and, thus, $\alpha = 0$. When $\beta > \beta_c$,

$$\frac{\partial u}{\partial T} = 2d\beta^2 m^{*,\text{CW}}(\beta) \frac{\partial m^{*,\text{CW}}(\beta)}{\partial \beta}.$$

Differentiating (2.12) with respect to β , we get

$$\frac{\partial m^{*,\text{CW}}(\beta)}{\partial \beta} = (1 - \tanh^2(2d\beta m^{*,\text{CW}}(\beta))) (2d m^{*,\text{CW}}(\beta) + 2d\beta \frac{\partial m^{*,\text{CW}}(\beta)}{\partial \beta}).$$

Rearranging, expanding the hyperbolic tangents to leading order and using (2.24), we obtain that, as $\beta \downarrow \beta_c$,

$$\frac{\partial m^{*,\text{CW}}(\beta)}{\partial \beta} = \frac{1 - \tanh^2(2d\beta m^{*,\text{CW}}(\beta))}{1 - 2d\beta + 2d\beta \tanh^2(2d\beta m^{*,\text{CW}}(\beta))} 2d m^{*,\text{CW}}(\beta) = \frac{(1 + o(1))}{2(\beta - \beta_c)} m^{*,\text{CW}}(\beta).$$

Using once more (2.24), we conclude that, as $\beta \downarrow \beta_c$,

$$c_{H=0}(\beta) = \frac{d\beta^2}{\beta - \beta_c} (m^{*,\text{CW}}(\beta))^2 (1 + o(1)) = \frac{3}{2} \frac{\beta_c}{\beta} (1 + o(1)),$$

so that $\lim_{\beta \downarrow \beta_c} c_{H=0}(\beta) = 3/2$ and $\alpha' = 0$.

We leave the proof that $\delta = 3$ to the reader. \square

2.5.4 Links with other models on \mathbb{Z}^d .

One of the main reasons for the interest in mean-field models is that the results obtained often shed light on the type of behavior that might be expected in more realistic lattice models on \mathbb{Z}^d , of the type discussed in the rest of this book. In view of the approximation involved, one might expect the agreement between a lattice spin system on \mathbb{Z}^d and its mean-field version to improve as the number of spins with which one spin interacts increases, which happens when either the range of the interaction becomes large, or the dimension of the lattice increases. It turns out that, in many cases, this can in fact be quantified rather precisely. We will not discuss these issues in much detail, but will rather provide some references. Much more information can be found in Sections II.13–II.15 and V.3–V.5 of Simon's book [308] and in Section 4 of Biskup's review [22].

Rigorous bounds. A first type of comparison between models on \mathbb{Z}^d and their mean-field counterpart is provided by various bounds on some quantities associated to the former in terms of the corresponding quantities associated to the latter.

First, the mean-field pressure is known to provide a rigorous lower bound on the pressure in very general settings. For example, we will show in Theorem 3.53 that

$$\psi_{\beta}^{\text{Ising on } \mathbb{Z}^d}(h) \geq \psi_{\beta}^{\text{CW}}(h).$$

(Remember that the dimensional parameter d also appears in the definition of the Curie–Weiss model in the right-hand side.) See also Exercise 6.28 for a closely related result.

Second, the mean-field critical temperature is known to provide rigorous upper bound on the critical temperature of models on \mathbb{Z}^d . Again, this is done for the Ising model, via a comparison of the magnetizations of these two models, in Theorem 3.53. More information and references on this topic can be found in [308, Sections V.3 and V.5].

Convergence. The bounds mentioned above enable a general comparison between models in arbitrary dimensions and their mean-field approximation, but do not yield quantitative information about the discrepancy. Here, we consider various limiting procedures, in which actual convergence to the mean-field limits can be established.

First, one can consider spin systems on \mathbb{Z}^d with spread-out interactions (for example, such that any pair of spins located at a distance less than some large value interact). A prototypical example are models with *Kac interactions*. An example of the latter is discussed in detail in Section 4.10 and a proof that the corresponding pressure converges to the corresponding mean-field pressure when the range of the interactions diverges is provided in Theorem 4.31. Additional information and references on this topic can be found at the end of Chapter 4.

Another approach is to consider models on \mathbb{Z}^d and prove convergence of the pressure or the magnetization as $d \rightarrow \infty$. A general result can be found in [308, Theorem II.14.1] with the relevant bibliography. Alternatively, one might try to provide quantitative bounds for the difference between the magnetization of a model on \mathbb{Z}^d and the magnetization of its mean-field counterpart. This is the approach developed in [23, 24, 68]; see also the lecture notes by Biskup [22]. This approach

is particularly interesting for models in which the mean-field magnetization is discontinuous. Indeed, once the dimension is large enough (or the interaction is sufficiently spread-out), the error term becomes small enough that the magnetization of the corresponding model on \mathbb{Z}^d must necessarily be also discontinuous. This provides a powerful technique to prove the existence of first-order phase transitions in some models.

Critical exponents. Finally, as mentioned in Section 2.5.3 and as will be discussed in more detail in Section 3.10.11, when a continuous phase transition occurs, describing qualitatively quantities exhibiting singular behavior is of great interest. In particular, a challenging problem is to determine the corresponding critical exponents, as we did for the Curie–Weiss model in Section 2.5.3. It is expected that the critical exponents of models on \mathbb{Z}^d *coincide* with those of their mean-field counterpart *for all large enough dimensions* (and not only in the limit!). Namely, there exists a critical dimension d_u , known as the **upper critical dimension**, such that the critical exponents take their mean-field values for all $d > d_u$. This has been proved in several cases, such as the Ising model, for which $d_u = 4$. A thorough discussion can be found in the book by Fernández, Fröhlich and Sokal [102].

