

Chapter 3

Ferromagnetism *alla* Ising

3.1 Introduction

In certain alloys, particularly those containing Fe, Co or Ni, the electrons have a tendency to align their spins in a common direction. This phenomenon is called **ferromagnetism** and is characterized by the existence of a finite magnetization even in the absence of a magnetic field. It has its origin in a genuinely quantum mechanical effect known as the **exchange interaction** and related to the overlap between the wave functions of neighboring electrons.

The spin ordering diminishes as the temperature of the sample increases and, above a certain temperature T_c , the magnetization vanishes entirely. This is a **phase transition** and T_c is called the **critical temperature** or the **Curie temperature**. A typical behavior of the magnetization $M(T)$ as a function of temperature is shown in Fig. 3.1(a). Below T_c the magnetization is finite and the material is termed *ferromagnetic*. At T_c the magnetization becomes identically zero and for $T > T_c$ the material behaves like the paramagnets we have studied before. The values of T_c for some selected materials are shown in Table 3.1. Magnetite (Fe_2O_3) was historically the first magnetic material found in nature.¹ A funny example is Gd. It has $T_c = 292 \text{ K} = 19^\circ \text{ C}$ and therefore is ferromagnetic in the winter, but paramagnetic in the summer. Another important example is $\text{Nd}_2\text{Fe}_{14}\text{B}$. It was developed in 1982 and nowadays has many applications, from electric cars to fusion reactors. But its popularity has nothing to do with its T_c value. Rather, it is because it is a **hard magnet**.

In terms of applications, what matters the most is the behavior of M as a function of the magnetic field H , below T_c (above T_c the material behaves like a paramagnet so there is nothing very interesting to see). A typical curve is shown in Fig. 3.1(b). This type of curve is called a **hysteresis loop**.² The first part of Fig. 3.1(b), marked as 1, is what happens when you take a sample

¹ Fe_2O_3 is actually *ferrimagnetic*. But the idea is the same.

² The word hysteresis means *lag* because the magnetization is always lagging behind the magnetic field, in contrast to paramagnetism where a given field determines a given magnetization.

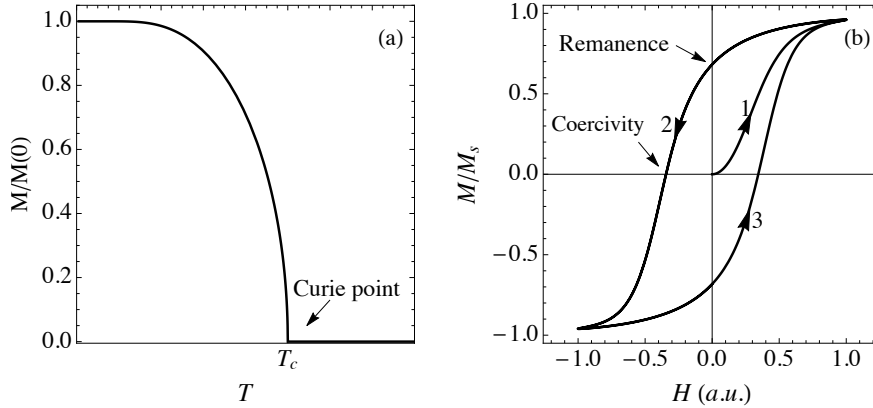


Figure 3.1: Typical behavior of a ferromagnetic material. (a) M vs. T . (b) Hysteresis loop, M vs. H .

Table 3.1: Curie temperature and coercive field for some selected magnetic materials. Data taken from completely untrustworthy sources around the internet. The values of the coercive field H_c are rather arbitrary since, as we will learn in the end of the chapter, this value depends sensibly on many details, such as the shape of the sample and so on.

	Fe ₂ O ₃	Fe	Co	Ni	EuO	Nd ₂ Fe ₁₄ B	Gd
T_c (K)	858	1043	1400	627	69	600-700	292
μ (μ_B /atom)		2.22	1.72	0.6			
H_c (kA/m)		0.16	0.8	0.056		900	

straight out of the oven and apply a magnetic field. The magnetization simply grows (usually much more quickly than in a paramagnetic material) towards the saturation value. The novel effect appears when we begin to reduce the field from saturation toward zero, as in the curve marked 2 in Fig. 3.1(b). What we see is that M does not go to zero when $H \rightarrow 0$, as in a paramagnet. But rather, it stops at a finite value called the **remanence**. The magnetization at remanence is precisely the function $M(T)$ illustrated in Fig. 3.1(a). If we want to take M back to zero we must apply a field in the opposite direction. The field at which $M(H) = 0$ is called the **coercive field** H_c . The coercive field measures how “strong” can the magnet resist to external fields. Values for selected materials are shown in Table 3.1. From these values it becomes quite clear why Nd₂Fe₁₄B is a “hard” magnet. It has a coercive field which is more than 10000 larger than raw iron, corresponding to a magnetic field $B = \mu_0 H_c$ of more than 1 T. That means that a very large field must be applied in order to demagnetize Nd₂Fe₁₄B.

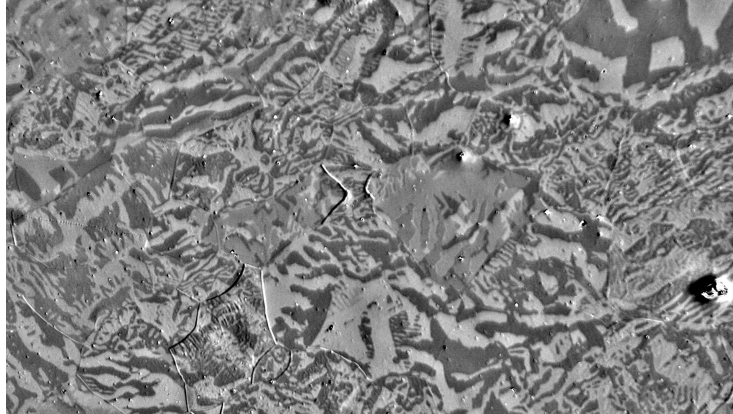


Figure 3.2: Magnetic domains in an FeSi sample observed with a Kerr polarized optical microscope. Courtesy of A. D. Santos.

The dipolar interaction

The exchange interaction is the microscopic mechanism behind the ferromagnetic order. However, this effect is *isotropic*. It causes the spins to align in a common direction, irrespective of which direction that is. Notwithstanding, it is well known that the magnetic properties of a material are strongly influenced by the **shape** of the sample. For instance, when the sample is carved in the form of a needle, as in a compass, there is a tendency for the magnetization to remain in the direction of the needle. Or, you may remember those old U-shaped magnets. It turns out that making the sample in the shape of a U strongly forces the magnetization to lie along the U.

This shape dependence is a consequence of the **dipolar interaction**. Each spin is a magnetic dipole and therefore produces a magnetic field which influences all other spins. The dipolar interaction is not an *essential* ingredient for the existence of a phase transition; the essential ingredient is the exchange interaction. However, once in the ferromagnetic phase, the dipolar interaction becomes important and starts to compete with the exchange interaction. The outcome is the formation of **magnetic domains**. Over very small distances, the exchange energy is dominant and microscopic chunks of the material are always magnetized in the same direction. However, having the entire material pointing in the same direction becomes energetically costly for the dipolar interaction and therefore the material divides itself into domains, each domain possessing a different magnetization. What we measure experimentally is the sum of the magnetization of each domain. The coercive point $M(H) = 0$ is therefore precisely the point where the net effect of all domains add up to zero.

Suppose now that you have a piece of iron and you apply a magnetic field and then remove it. The sample will become magnetized, meaning that the field will make all magnetic domains point in the same direction. In fact, iron has the highest magnetic moment in nature, so your sample will have a very

large magnetization. However, iron is *soft*. It has a low coercivity and therefore tiny perturbations, such as hitting it with a hammer, or even thermal fluctuations, will tend to reintroduce a domain structure which you will observe as a demagnetization of your sample.

The trick to obtain a hard magnet is therefore to find a way to make the formation of domains energetically costly. Historically, this was done by exploiting the shape of the sample, like in the U-shaped magnets. The big breakthrough of Nd₂Fe₁₄B magnets was to exploit instead the crystal structure. The Nd and B atoms are not magnetic, only Fe is. But they are introduced so as to create a very specific crystal structure which forces the Fe spins to prefer to be aligned in a given direction. This is known as *magnetocrystalline anisotropy*. Of course, from a technological standpoint, it may also be interesting to produce extremely soft magnets. One way to do that is by producing amorphous materials (ie, materials which are not crystalline). That way you completely eliminate any magnetocrystalline anisotropy.

Battle plan

We will come back to magnetism a million times again. This chapter is just a first passage through this problem. We will therefore start by investigating the simplest model possible containing the necessary symmetries of a ferromagnetic system. This is the **Ising model**, which is the simplest model to describe ferromagnetism. It was proposed by Wilhelm Lenz in 1920 as a PhD project to his student Ernst Ising. We will learn that exact solutions of the Ising model are difficult to come by. Here we will give the solution in 1D. It can also be solved in 2D, but we will leave that to another chapter since the solution is quite complicated. In 3D no one has been able to solve it so far.

Then we will discuss an approximation technique called the **mean-field theory**. This type of theory appear in many other areas of physics and therefore it is a good idea to have a firm understanding of what they do. For instance, the Hartree-Fock approximation used in electronic structure calculations is a mean-field. They are also the basis for the BCS theory of superconductivity.

3.2 The 1D Ising model: zero magnetic field

The one-dimension Ising model, which was the one actually studied by Ising in his PhD, is defined by a one-dimensional lattice with N sites, each being represented by a Pauli matrix σ_i^z (see Fig. 3.3). The Hamiltonian is taken to be

$$H = -J \sum_{i=1}^{N-1} \sigma_i^z \sigma_{i+1}^z \quad (3.1)$$

where, in this section, we will assume there is no magnetic field. For simplicity, we will assume $J > 0$, corresponding to a ferromagnetic case. Usually these models are solved with periodic boundary conditions (which means we include

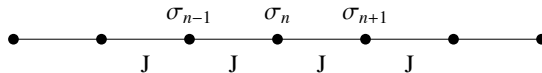


Figure 3.3: The Ising model in 1D. Each site is described by a variable $\sigma_n = \pm 1$ and interacts with nearest neighbors with exchange constant J .

an extra term $\sigma_N^z \sigma_1^z$). But in this section only we will assume open boundary conditions, for reasons that will become clear as we move along.

This Hamiltonian is already diagonal in the $|\sigma_1, \dots, \sigma_N\rangle$ basis, with energy levels

$$E = -J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1} \quad (3.2)$$

In total there are N spins but only $N-1$ bonds. The ground-state (ie the lowest energy state) is obtained when we set all spins up:

$$E_{\text{gs}} = -(N-1)J \quad (3.3)$$

The ground-state is therefore ferromagnetic (all spins aligned). The relevant question is whether that remains true at finite temperatures. To do that we must compute the partition function

$$Z = \text{tr}(e^{-\beta H}) = \sum_{\sigma_1, \dots, \sigma_N} e^{-\beta E} \quad (3.4)$$

The partition function is a sum over all possible configurations of the spin variables $\sigma_1, \dots, \sigma_N$. Since each σ_n takes on the values ± 1 , there is a total of 2^N distinct terms in this sum. Moreover, this is not like the systems we have worked with before, where Z factored into a product of partition functions for each individual particle. In this case the particles interact so we must face the full Z .

Lucky for us, there is a trick to compute Eq. (3.4). As far as I know, it only works in this particular situation. But it is really cute so will discuss it anyway. A more general method will be given in the next section. The idea is to compute first the sum over σ_N :

$$Z = \sum_{\sigma_1, \dots, \sigma_{N-1}} e^{\beta J(\sigma_1 \sigma_2 + \dots + \sigma_{N-2} \sigma_{N-1})} \left(\sum_{\sigma_N} e^{\beta J \sigma_{N-1} \sigma_N} \right)$$

This sum gives

$$\sum_{\sigma_N} e^{\beta J \sigma_{N-1} \sigma_N} = e^{\beta J \sigma_{N-1}} + e^{-\beta J \sigma_{N-1}} = 2 \cosh(\beta J \sigma_{N-1})$$

But now comes the fun part: \cosh is an even function and $\sigma_{N-1} = \pm 1$. Whence, $\cosh(\beta J \sigma_{N-1}) = \cosh(\beta J)$. This part of the sum therefore factors out of the

partition function and we are left with

$$Z = (2 \cosh \beta J) \sum_{\sigma_1, \dots, \sigma_{N-1}} e^{\beta J(\sigma_1 \sigma_2 + \dots + \sigma_{N-2} \sigma_{N-1})}$$

What remains is exactly the partition function for a system with $N - 1$ spins. If we continue to repeat this procedure, we get a factor of $2 \cosh \beta J$ each time. The only different term will be the last one (the sum over σ_1). It will have no argument so we will get simply $\sum_{\sigma_1} 1 = 2$. Hence, the partition function is

$$Z = 2(2 \cosh \beta J)^{N-1} \quad (3.5)$$

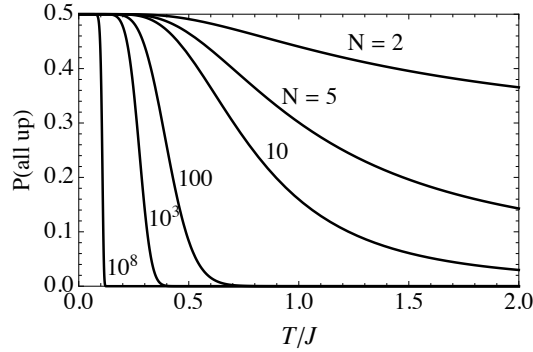


Figure 3.4: Probability that the system is fully magnetized as a function of temperature and the number of particles, Eq. (3.6).

Let us compute the probability that all spins are up. The probability for any state $\sigma = (\sigma_1, \dots, \sigma_N)$ is, of course,

$$P_\sigma = \frac{e^{-\beta E_\sigma}}{Z}$$

When all spins are up, $\sigma_1 \sigma_2 + \dots + \sigma_{N-1} \sigma_n = (N - 1)$, so

$$P(\text{all up}) = \frac{e^{\beta J(N-1)}}{2(\cosh \beta J)^{N-1}} = \frac{1}{2} \left(\frac{e^{\beta J}}{\cosh \beta J} \right)^{N-1} \quad (3.6)$$

By symmetry, this is the same as the probability that all spins are down. This result is plotted as a function of T in Fig. 3.4 for several values of N . As can be seen, when N is moderately small, $P(\text{all up})$ changes gradually from a small value at high T to $1/2$ at $T = 0$. However, we see that as we increase N the probability of finding all spins up tends to a step function, being zero above a certain temperature and exactly $1/2$ below. Moreover, this temperature where

the jump occurs is also pushed to smaller and smaller values as N increases. Hence, as $N \rightarrow \infty$, $P(\text{all up})$ becomes discontinuous, being exactly zero for all finite temperatures but then jumping to $1/2$ exactly at $T = 0$. This is a strong indication that finite temperatures destroy the ferromagnetic order (ie, that the model is ferromagnetic only at $T = 0$). But before we can conclude this for sure, we need to investigate thermodynamic quantities.

Thermal properties

I will assume N is sufficiently large and therefore replace $N - 1 \simeq N$. The main thermodynamic quantities are then

$$\begin{aligned}
 F &= -T \ln Z = -T \ln 2 - NT \ln[2 \cosh(J/T)] \\
 U &= -\frac{\partial}{\partial \beta} \ln Z = -NJ \tanh(J/T) \\
 C &= \frac{\partial U}{\partial T} = N \left(\frac{J}{T}\right)^2 \operatorname{sech}^2(J/T) \\
 S &= \frac{U - F}{T} = \ln 2 + N \left\{ -\frac{J}{T} \tanh(J/T) + \ln[2 \cosh(J/T)] \right\}
 \end{aligned}$$

These results are presented in Fig. 3.5.

In the entropy we also see a lonely term $\ln 2$. This term is negligible in comparison with the N -dependent term. But it has nonetheless an interesting physical interpretation. At zero field the ground-state is two-fold degenerate since the spins can be all up or all down. At $T = 0$ we know the entropy should equal the logarithm of the degeneracy of the ground-state [cf. Eq. (2.116)]. And this is exactly what we have here: at $T = 0$ the N -dependent term in S vanishes and we are left with $S(T = 0) = \ln 2$. On the other hand, when $T \rightarrow \infty$, the entropy tends to $N \ln 2$, which is the logarithm of the total dimension of the Hilbert space, 2^N (at infinite temperatures all states become equally likely).

What about the magnetization? It turns out that for $h = 0$ it is a bit hard to find it. We learned that

$$M = -\frac{\partial F}{\partial h}$$

But since $h = 0$ there is no h in F to differentiate. We could try to compute the magnetization from its definition as

$$M = \frac{1}{Z} \sum_{\sigma_1, \dots, \sigma_N} \mathcal{M} e^{-\beta E}, \quad \mathcal{M} = \sum_{n=1}^N \sigma_n$$

However, this result will certainly be zero since the system has up-down symmetry: for each state with a given positive \mathcal{M} , there will be a corresponding state with negative \mathcal{M} and the same probability.

In the case of the 1D Ising model this difficulty is only temporary: in the next section we will discuss how to solve the problem with $h \neq 0$ [We will learn

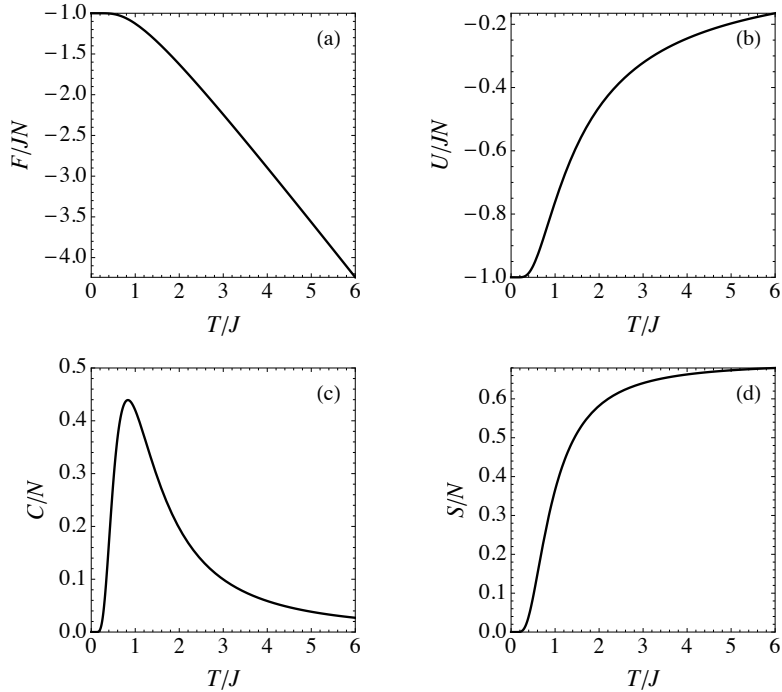


Figure 3.5: Thermodynamic quantities for the Ising model under zero field

that indeed there is no ferromagnetic order, with M tending to zero as $h \rightarrow 0$]. However, for the Ising model in 2D this has been, historically, a big problem because we only know the solution for $h = 0$. It turns out that the way to compute the magnetization is look for two-point correlation functions which are infinitely far apart. In a ferromagnetic phase these turn out to be proportional to the magnetization (although demonstrating that analytically is not an easy task).

The σ - τ transformation

Eq. (3.5) looks a lot like the partition function for non-interacting spins in a magnetic field, Eq. (2.45):

$$Z = [2 \cosh(\beta h)]^N$$

This similarity is no coincidence. Consider the following transformation of variables:

$$\begin{aligned}\tau_1 &= \sigma_1 \\ \tau_2 &= \sigma_2 \sigma_1 \\ \tau_3 &= \sigma_3 \sigma_2 \\ &\vdots \\ \tau_N &= \sigma_N \sigma_{N-1}\end{aligned}$$

These τ_n also take on the values ± 1 . To find the inverse formulas we use the fact that $\sigma_n = \pm 1$, so $\sigma_n^2 = 1$. Moreover, $1/\sigma_n = \sigma_n$ (I think this property quite entertaining). We then get

$$\begin{aligned}\sigma_1 &= \tau_1 \\ \sigma_2 &= \tau_2 / \sigma_1 = \tau_2 \tau_1 \\ \sigma_3 &= \tau_3 / \sigma_2 = \tau_3 \tau_2 \tau_1 \\ &\vdots \\ \sigma_N &= \tau_N \tau_{N-1} \dots \tau_2 \tau_1\end{aligned}$$

With this transformation,

$$J(\sigma_1 \sigma_2 + \dots + \sigma_{N-1} \sigma_N) = J(\tau_2 + \dots + \tau_N)$$

Hence, with this mapping we can transform **bonds** into **sites**. This is an example of a **duality transformation**. It means that the energy (3.2) is physically equivalent to $N-1$ spins under a “magnetic field” J . Using the τ ’s, the partition function becomes

$$\begin{aligned}Z &= \sum_{\tau_1, \dots, \tau_N} e^{\beta J (\tau_2 + \dots + \tau_N)} \\ &= \left(\sum_{\tau_1} 1 \right) \left(\sum_{\tau_2} e^{\beta J \tau_2} \right) \dots \left(\sum_{\tau_N} e^{\beta J \tau_N} \right) \\ &= 2(2 \cosh \beta J)^{N-1}\end{aligned}$$

which is the same as Eq. (3.5).

Correlation function

Another interesting quantity to compute is the correlation function $\langle \sigma_n \sigma_{n+r} \rangle$. It represents the degree of statistical correlation between two spins at different

positions. It is defined as

$$\langle \sigma_n \sigma_{n+r} \rangle = \frac{1}{Z} \sum_{\sigma_1, \dots, \sigma_N} \sigma_n \sigma_{n+r} e^{\beta J (\sigma_1 \sigma_2 + \dots + \sigma_{N-1} \sigma_N)}$$

To compute the correlation function it is easier to use the τ 's. Since $\sigma_n^2 = 1$, we may write

$$\begin{aligned} \sigma_n \sigma_{n+r} &= (\sigma_n \sigma_{n+1})(\sigma_{n+1} \sigma_{n+2}) \dots (\sigma_{n+r-1} \sigma_{n+r}) \\ &= \tau_{n+1} \dots \tau_{n+r} \end{aligned}$$

The correlation function then becomes

$$\langle \sigma_n \sigma_{n+r} \rangle = \frac{1}{2(2 \cosh J)^{N-1}} \sum_{\tau_1, \dots, \tau_N} (\tau_{n+1} \dots \tau_{n+r+1}) e^{\beta J (\tau_2 + \dots + \tau_N)}$$

The sum is now completely factored. The sum over τ_1 cancels the 2 downstairs and the sums that do not have τ_n multiplying the exponential will cancel the cosh downstairs. What survives are r sums of the form

$$\frac{1}{2 \cosh \beta J} \sum_{\tau} \tau e^{\beta J \tau} = \frac{1}{2 \cosh \beta J} (e^{\beta J} - e^{-\beta J}) = \tanh \beta J$$

Hence, we find

$$\boxed{\langle \sigma_n \sigma_{n+r} \rangle = \tanh^r(J/T)} \quad (3.7)$$

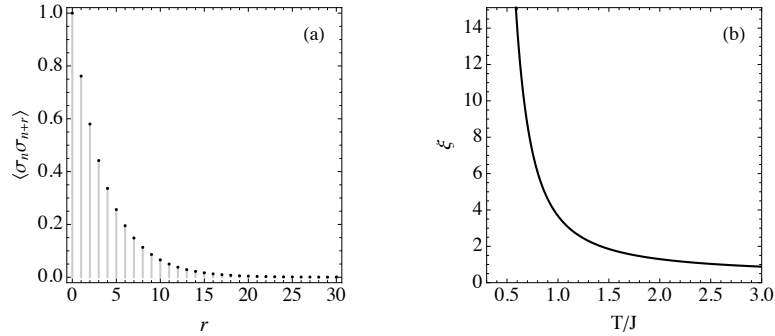


Figure 3.6: (a) The two-point correlation function versus the lattice spacing r , for the 1D Ising model, Eq. (3.7) with $J/T = 1$. (b) The corresponding correlation length, Eq. (3.8) as a function of T/J .

This result is presented in Fig. 3.6(a). As can be seen, the correlation between two sites decays quickly with the sites separation r . Usually, these correlation functions can be written as

$$\langle \sigma_n \sigma_{n+r} \rangle = e^{-r/\xi}$$

where ξ is called the **correlation length**. In our case we may write $\tanh^r(J/T) = e^{r \ln \tanh(J/T)}$. We then see that the correlation length is

$$\xi = -\frac{1}{\ln \tanh(J/T)} \quad (3.8)$$

When $T \rightarrow 0$, this correlation length diverges [see Fig. 3.6(b)]. This means that two spins infinitely far apart will be perfectly correlated. This type of long-range correlation is a signature of ferromagnetic order, which agrees with our previous discussion that the system is ferromagnetic at $T = 0$.

Susceptibility

Lastly, we turn to the susceptibility:

$$\chi = -\frac{\partial^2 F}{\partial h^2}$$

In a previous occasion, we showed that χ was related to the fluctuations of the magnetization,

$$\chi = \frac{1}{T} \left[\langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2 \right]$$

The second term is zero because, as we have already seen, there is no spontaneous magnetization. Moreover, expanding $\mathcal{M} = \sum_n \sigma_n$, we get

$$\chi = \frac{1}{T} \sum_{n,m} \langle \sigma_n \sigma_m \rangle$$

This formula shows the relationship between the correlation function and the susceptibility. We can also manipulate it further, as follows:

$$\begin{aligned} \chi &= \frac{1}{T} \left[\sum_n \langle \sigma_n \sigma_n \rangle + \sum_{n,m \neq n} \langle \sigma_n \sigma_m \rangle \right] \\ &= \frac{1}{T} \left[\sum_n (1) + 2 \sum_{n,m > n} \langle \sigma_n \sigma_m \rangle \right] \\ &= \frac{1}{T} \left[N + 2N \sum_{r=1}^N \langle \sigma_1 \sigma_{1+r} \rangle \right] \end{aligned}$$

where, in the last line we used the fact that $\langle \sigma_n \sigma_m \rangle$ depends only on the distance between n and m . We then get

$$\chi = \frac{N}{T} \left[1 + 2 \sum_{r=1}^N \tanh^r \beta J \right]$$

We may compute this resulting sum if we assume that $N \rightarrow \infty$. In this case it resembles the geometric series, which converges because $\tanh J < 1$. We just need to be a bit careful because the geometric series starts at $r = 0$. So we write

$$\chi = \frac{N}{T} \left[1 - 2 + 2 \sum_{r=0}^{\infty} \tanh^r \beta J \right] = \frac{1}{T} \left[-1 + \frac{2}{1 - \tanh \beta J} \right]$$

Simplifying we get

$$\chi = \frac{N}{T} \left(\frac{1 + \tanh(J/T)}{1 - \tanh(J/T)} \right) \quad (3.9)$$

This result is shown in Fig. 3.7. As can be seen, it diverges as we approach $T \rightarrow 0$.

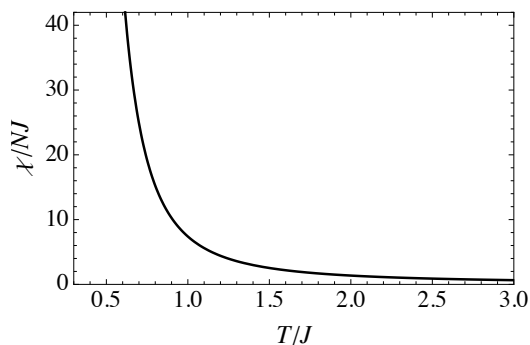


Figure 3.7: Susceptibility for the 1D Ising model, Eq. (3.9), as a function of T/J .

3.3 Non-zero field and the transfer matrix

Next we turn to the 1D Ising model when $h \neq 0$. In this case it is convenient to use periodic boundary conditions so we will write the energy as

$$E = -J \sum_{n=1}^N \sigma_n \sigma_{n+1} - h \sum_{n=1}^N \sigma_n \quad (3.10)$$

and the partition function is

$$Z = \sum_{\sigma_1, \dots, \sigma_N} e^{-\beta E} = \sum_{\sigma_1, \dots, \sigma_N} \exp \left\{ \beta J \sum_{n=1}^N \sigma_n \sigma_{n+1} + \beta h \sum_{n=1}^N \sigma_n \right\} \quad (3.11)$$

This sum may be computed using a technique known as the **transfer matrix**, which can also be used in other models, in particular in the 2D Ising model.

We begin by introducing the function

$$V(\sigma_1, \sigma_2) = \exp \left\{ \beta J \sigma_1 \sigma_2 + \frac{\beta h}{2} (\sigma_1 + \sigma_2) \right\} \quad (3.12)$$

This is almost a typical term in Eq. (3.11), except that I slightly altered the magnetic field term to make things look more symmetric. The partition function (3.11) may now be written as (please take your time to check that this is indeed correct)

$$Z = \sum_{\sigma_1, \dots, \sigma_N} V(\sigma_1, \sigma_2) V(\sigma_2, \sigma_3) \dots V(\sigma_{N-1}, \sigma_N) V(\sigma_N, \sigma_1) \quad (3.13)$$

The key to the **transfer matrix technique** is to recognize the similarity between this expression and the trace of a matrix.

Given some matrix A , its trace is the sum of the diagonal entries:

$$\text{tr}(A) = \sum_i A_{i,i}$$

The trace of A^2 , on the other hand, is

$$\text{tr}(A^2) = \sum_i (A^2)_{i,i} = \sum_{i,j} A_{i,j} A_{j,i}$$

And the trace of A^3 is

$$\text{tr}(A^3) = \sum_{i,j,k} A_{i,j} A_{j,k} A_{k,i}$$

Note how this is exactly what appears in Eq. (3.13), provided we interpret $V(\sigma_i, \sigma_j)$ as the matrix elements of a matrix V . This is acceptable since each σ_i can take on only two values so that V will be 2×2 .

The matrix elements of V can be computed directly from the definition in Eq. (3.12):³

$$V = \begin{pmatrix} V(1,1) & V(1,-1) \\ V(-1,1) & V(-1,-1) \end{pmatrix} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix} \quad (3.14)$$

The partition function in Eq. (3.13) may now be written as

$$\boxed{Z = \text{tr}(V^N)} \quad (3.15)$$

This is the idea of the transfer matrix: the matrix V transfers the solution from one site to the other and the partition function is simply the trace of a product of V s.

We now use two general theorems of linear algebra:

³ The reason why we symmetrize Eq. (3.12) is precisely because it produces a symmetric matrix V , which is much easier to work with.

1. The trace of a matrix is the sum of its eigenvalues.
2. If λ_i are the eigenvalues of a matrix A , then the eigenvalues of A^N are λ_i^N .

Let us denote the eigenvalues of V by λ_1 and λ_2 . They may be easily computed since V is 2×2 . The characteristic equation is

$$\det(V - \lambda) = \lambda^2 - \lambda(2e^{\beta J} \cosh \beta h) + (e^{2\beta J} - e^{-2\beta J}) = 0$$

which gives for the two eigenvalues

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta h) \pm \sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta h)} \quad (3.16)$$

Using our results from linear algebra now allow us to write the partition function as

$$Z = \lambda_+^N + \lambda_-^N \quad (3.17)$$

From Eq. (3.16) we see that $\lambda_+ > \lambda_-$. We may therefore write

$$Z = \lambda_+^N \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right]$$

If we are interested in the thermodynamic limit, $N \rightarrow \infty$, the term inside parenthesis will tend to zero since $\lambda_-/\lambda_+ < 1$. Hence, we conclude that

$$Z = \lambda_+^N = \lambda_{\max}^N \quad (3.18)$$

This is actually a general result, valid for any problem that can be solved using transfer matrices: *the partition function is the largest eigenvalue of the transfer matrix, to the power N* . For non-interacting problems we had $Z = Z_1^N$. Now we don't have that because the particles interact. Notwithstanding, we recover something quite similar.

To finish the problem with style, let us write the result explicitly:

$$Z = \left[e^{\beta J} \cosh(\beta h) + \sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta h)} \right]^N \quad (3.19)$$

In the limit that $h \rightarrow 0$ this reduces to Eq. (3.5). The free energy is

$$F = -NT \ln \lambda_+ = -NT \left\{ e^{\beta J} \cosh(\beta h) + \sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta h)} \right\} \quad (3.20)$$

Whence, the magnetization per particle is

$$m = \frac{M}{N} = -\frac{1}{N} \frac{\partial F}{\partial h} = \frac{e^{\beta J} \sinh(\beta h)}{\sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta h)}} \quad (3.21)$$

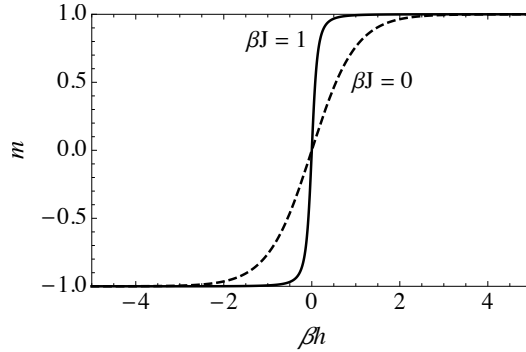


Figure 3.8: Magnetization of the 1D Ising model, Eq. (3.21), as a function of βh for two different values of βJ .

This result is plotted in Fig. 3.8 as a function of βh for two different values of βJ . The dashed line correspond to the limit $\beta J \rightarrow 0$, where we obtain

$$m = \tanh(\beta h)$$

corresponding to independent spins [cf. Eq (2.49)]. As can be seen, the effect of changing J is to make the curve grow up faster. However, as $h \rightarrow 0$ we always obtain a zero magnetization. This is the definitive proof that the system is not ferromagnetic at any finite temperature.

Some properties of Pauli matrices

Computing the eigenvalues of V in Eq. (3.14) is an easy task since the matrix is 2×2 . However, to compute correlation functions we also need the eigenvectors, which are usually more cumbersome to deal with, even for 2×2 matrices. Whenever you are faced with a problem like this, I always recommend you use Pauli matrices. The idea is to use the fact that the most general 2×2 matrix may be written as

$$A = a_0 + \mathbf{a} \cdot \boldsymbol{\sigma} \quad (3.22)$$

for four numbers a_0, a_x, a_y and a_z . Here $\mathbf{a} \cdot \boldsymbol{\sigma}$ is an abbreviation for $a_x \sigma_x + a_y \sigma_y + a_z \sigma_z$. We may also let $a = |\mathbf{a}|$ and introduce the unit vector $\mathbf{n} = \mathbf{a}/a$. Then

$$A = a_0 + a(\mathbf{n} \cdot \boldsymbol{\sigma}) \quad (3.23)$$

The idea of using Pauli matrices is to exploit the following properties:

1. The eigenvectors of A are the same as those of $\mathbf{n} \cdot \boldsymbol{\sigma}$.
2. If λ is an eigenvalue of $\mathbf{n} \cdot \boldsymbol{\sigma}$, then $a_0 + a\lambda$ is an eigenvalue of A .

Let us look more closely at the matrix

$$\mathbf{n} \cdot \boldsymbol{\sigma} = n_x \sigma_x + n_y \sigma_y + n_z \sigma_z = \begin{pmatrix} n_z & n_x - i n_y \\ n_x + i n_y & -n_z \end{pmatrix}$$

It is straightforward to verify that the eigenvalues of this matrix are ± 1 . Hence, the eigenvalues of A in Eq. (3.23) are

$$a_0 \pm a \quad (3.24)$$

As for the eigenvectors, it is convenient to parametrize \mathbf{n} in spherical coordinates,

$$\begin{aligned} n_x &= \sin \theta \cos \phi \\ n_y &= \sin \theta \sin \phi \\ n_z &= \cos \theta \end{aligned}$$

Then the eigenvectors may be written as

$$|v_+\rangle = \begin{pmatrix} \cos \theta/2 \\ e^{i\phi} \sin \theta/2 \end{pmatrix} \quad |v_-\rangle = \begin{pmatrix} -e^{-i\phi} \sin \theta/2 \\ \cos \theta/2 \end{pmatrix} \quad (3.25)$$

This procedure gives a systematic way to find eigenvalues and eigenvectors of 2×2 matrices. And the resulting formulas are much more convenient to work with. Every time I have to deal with 2×2 matrices, I always use this procedure.

Back to the transfer matrix

Now let us go back to the transfer matrix in Eq. (3.14). It can be written as

$$\begin{aligned} V &= e^{\beta J} \cosh(\beta h) + e^{-\beta J} \sigma_x + e^{\beta J} \sinh(\beta h) \sigma_z \\ &= e^{\beta J} \cosh(\beta h) + \sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta h)} \left\{ \sin \theta \sigma_x + \cos \theta \sigma_z \right\} \end{aligned} \quad (3.26)$$

where

$$\tan \theta = \frac{e^{-2\beta J}}{\sinh(\beta h)} \quad (3.27)$$

This is precisely in the form of Eq. (3.22). The eigenvalues of V are therefore $a_0 \pm a$, or

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta h) \pm \sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta h)}$$

which is the same as Eq. (3.16). Moreover, the eigenvectors are

$$\boxed{|v_+\rangle = \begin{pmatrix} \cos \theta/2 \\ \sin \theta/2 \end{pmatrix} \quad |v_-\rangle = \begin{pmatrix} -\sin \theta/2 \\ \cos \theta/2 \end{pmatrix}, \quad \tan \theta = \frac{e^{-2\beta J}}{\sinh(\beta h)}} \quad (3.28)$$

We shall use these formulas below to exploit more sophisticated properties of the transfer matrix.

Magnetization

Let us look at

$$\langle \sigma_1 \rangle = \frac{1}{Z} \sum_{\sigma_1, \dots, \sigma_N} \sigma_1 V(\sigma_1, \sigma_2) V(\sigma_2, \sigma_3) \dots V(\sigma_N, \sigma_1)$$

The quantity $\sigma_1 V(\sigma_1, \sigma_2)$ may be interpreted as another matrix, V' . It has values [cf. Eq. (3.14)]

$$V' = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ -e^{-\beta J} & -e^{\beta(J-h)} \end{pmatrix}$$

So we may also write

$$\langle \sigma_1 \rangle = \frac{1}{Z} \text{tr}(V' V^{N-1})$$

But you may verify that

$$V' = \sigma_z V$$

Hence, we have

$$\langle \sigma_1 \rangle = \frac{1}{Z} \text{tr}(\sigma_z V^N)$$

This logic also extends to the magnetization at any other site. In general,

$$\langle \sigma_n \rangle = \frac{1}{Z} \text{tr}(V^{n-1} \sigma_z V^{N-n+1}) \quad (3.29)$$

But using the cyclic property of the trace we find that

$$\langle \sigma_n \rangle = \frac{1}{Z} \text{tr}(\sigma_z V^N) \quad (3.30)$$

The cyclic property of the trace thence reflects the **translational invariance** of the model. We may compute this trace in the $|v_{\pm}\rangle$ basis which diagonalizes V . We then get

$$\langle \sigma_n \rangle = \frac{1}{Z} \left\{ \langle v_+ | \sigma_z | v_+ \rangle \lambda_+^N + \langle v_- | \sigma_z | v_- \rangle \lambda_-^N \right\}$$

In the thermodynamic limit, however, $Z = \lambda_+^N$ so the term containing λ_- vanishes. We are then left with

$$\langle \sigma_n \rangle = \langle v_+ | \sigma_z | v_+ \rangle \quad (3.31)$$

This result has a very important physical interpretation. As can be seen, we have reduced the problem to a **quantum mechanical** expectation value. We

can envisage $|v_+\rangle$ as being the *ground-state of a 2-level system*. In this sense the transfer matrix has reduced a 1D thermal average into a zero-dimensional (ie, a single particle) ground-state average.

A similar thing will happen in the 2D Ising model. There the transfer matrix will transfer one row to another and is therefore $N \times N$, not 2×2 . We will then show how to use the transfer matrix to transform the 2D Ising model into a 1D quantum mechanical model at zero temperature, which will turn out to be the Ising model in a transverse field. In the literature, this type of transformation goes by the name of **Hamiltonian limit**.

For sanity purposes, let us compute Eq. (3.31) and see if it matches Eq. (3.21). Using Eq. (3.28) we find

$$\langle \sigma_n \rangle = \cos^2(\theta/2) - \sin^2(\theta/2) = \cos \theta = \frac{e^{\beta J} \sinh(\beta h)}{\sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta h)}}$$

which is indeed Eq. (3.21), so everything works. This result also gives a physical interpretation to the angle θ .

Correlation function

Finally, let us study correlation functions $\langle \sigma_n \sigma_{n+r} \rangle$. Using the same argument that led us to Eq. (3.29), we may write

$$\langle \sigma_n \sigma_{n+r} \rangle = \frac{1}{Z} \text{tr}(V^{n-1} \sigma_z V^r \sigma_z V^{N-n-r+1}) = \frac{1}{Z} \text{tr}(\sigma_z V^r \sigma_z V^{N-r})$$

We now compute this trace in the $|v_\pm\rangle$ basis, already assuming that $Z = \lambda_+^N$:

$$\langle \sigma_n \sigma_{n+r} \rangle = \frac{1}{\lambda_+^N} \left\{ \langle v_+ | \sigma_z V^r \sigma_z | v_+ \rangle \lambda_+^{N-r} + \langle v_- | \sigma_z V^r \sigma_z | v_- \rangle \lambda_-^{N-r} \right\}$$

Again, the second term vanishes and we are left with

$$\boxed{\langle \sigma_n \sigma_{n+r} \rangle = \frac{\langle v_+ | \sigma_z V^r \sigma_z | v_+ \rangle}{\lambda_+^r}} \quad (3.32)$$

To make things more symmetric, you can also write this as

$$\langle \sigma_n \sigma_{n+r} \rangle = \frac{\langle v_+ | \sigma_z V^r \sigma_z | v_+ \rangle}{\langle v_+ | V^r | v_+ \rangle}$$

Carrying out the computation finally gives

$$\langle \sigma_n \sigma_{n+r} \rangle = \cos^2 \theta + \left(\frac{\lambda_-}{\lambda_+} \right)^r \sin^2 \theta \quad (3.33)$$

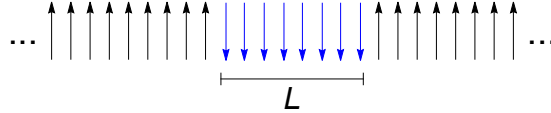


Figure 3.9: A single magnetic domain in the 1D Ising model.

The correlation between two spins very far apart therefore tends to $\cos^2 \theta$. It is convenient to subtract this part and define the **connected correlation**

$$G(r) = \langle \sigma_n \sigma_{n+r} \rangle - \langle \sigma_n^z \rangle \langle \sigma_{n+r}^z \rangle = \left(\frac{\lambda_-}{\lambda_+} \right)^r \sin^2 \theta \quad (3.34)$$

I wrote $G(r)$ and not $G(n, n+r)$ to emphasize that, due to translation symmetry, this result depends only on the distance between the spins. When $h \rightarrow 0$ this result reduces to Eq. (3.7).

3.4 The Peierls arguments

Ising solved the 1D Ising model (around 1925) in a way similar to how we did it and found no phase transition, except at $T = 0$. He then concluded that the Ising model in any dimension would not have a phase transition. But that conclusion is wrong: dimensionality *matters*; it matters *a lot*.

The best way to understand why dimensionality is so important is through an argument put forth by R. Peierls in 1936 on a paper entitled *On the Model of Ising for Ferromagnetism*. Consider a one-dimensional Ising chain. We already know that the ground-state is the fully magnetized state, with all spins up (or down). Now we want to test how robust this state is, in comparison with other states. To do that, suppose we produce a single magnetic domain of size L in the middle of the chain, as illustrated in Fig. 3.9. Recall that the Ising energies are

$$E = -J \sum_n \sigma_n \sigma_{n+1}$$

Every time we flip a spin, we increase the energy by $2J$. For the domain configuration in Fig. 3.9 there are two flips, so the energy will be $E = E_0 + 4J$, where E_0 is the ground state energy.

Recall, however that at finite temperatures we should look not for the state of smallest energy, but for the state of smallest *free* energy $F = U - TS$. The entropy of a given microscopic configuration is $S = \ln \Omega$, where Ω is the number of ways to produce an equivalent configuration. Since the domain length is L , we see that there are roughly L different ways of making a domain of length L . Thus, $S \sim \ln L$, which will give a free energy

$$F = U - TS = (E_0 + 4J) - T \ln L$$

Now comes Peierls' argument: the first term is independent of size, so for any $T \neq 0$, it is always possible to find a domain size L for which the second term will dominate. That is, we can always find a size for which the free energy of having a domain is smaller than the free energy of *not* having a domain. Consequently, there can be no magnetic order.

The situation is dramatically different in 2D. Consider a single domain in the 2D Ising model, as in 3.10, and assume that the perimeter of the domain is L . This means that there are L up-down bonds and therefore the energy of this domain will be $E = E_0 + 2JL$. Unlike the 1D case, this energy now depends on L . The number of microscopic configurations giving a domain of perimeter L now also increase dramatically. In fact, it can be shown that this number is of the order $\Omega \sim c^L$ where c is a number between 2 and 3. Therefore, the free energy will now be

$$F = U - TS \simeq E_0 + 2JL - TL \ln c$$

If

$$T > \frac{2J}{\ln c}$$

it is more advantageous (from a free energy point of view) to have the system dividing up into an infinite number of domains, which would destroy the ferromagnetic order. However, if T is smaller than this quantity the fully magnetized state is the state which continues to give the smallest energy. Consequently, at very low temperatures there should be an ordered ferromagnetic phase.

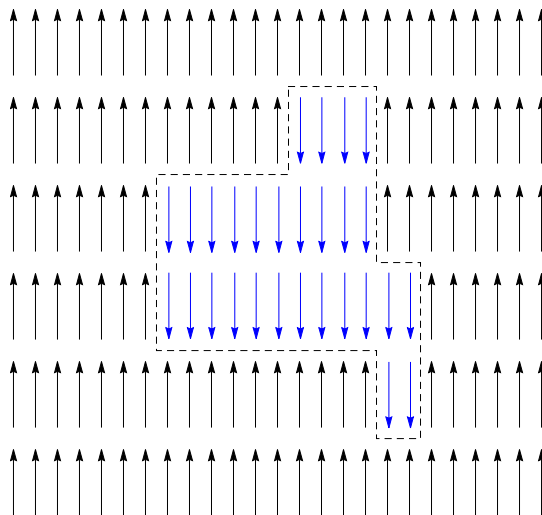


Figure 3.10: A single magnetic domain in the 2D Ising model with perimeter L

Peierls' argument therefore shows that *dimensionality matters*. Every statistical model has something called a **lower critical dimension**, which is the

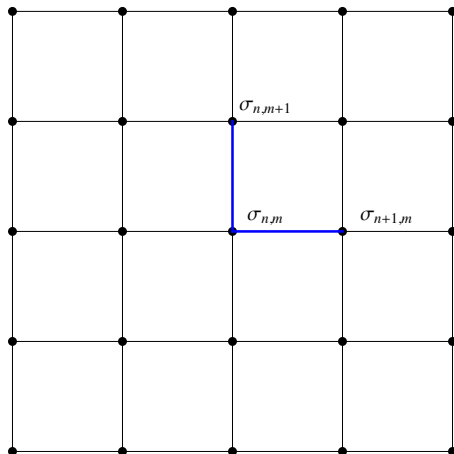


Figure 3.11: The Ising model in 2D The blue lines denote the terms to include in the sum when writing the energy explicitly.

smallest dimension at which there is no phase transition. For the Ising model the lower critical dimension is 1.

3.5 The mean-field approximation

The mean-field approximation is *the* most widely used approximation to deal with interacting models. We will discuss this approximation now in the context of the Ising model, but it is also applicable to an enormous class of problems. For instance, it is the basis of the Hartree-Fock approximation in electronic structure calculations and also of the BCS theory of superconductivity.

Consider the Ising model in a d -dimensional hypercubic lattice with only nearest-neighbor interactions. The energy is taken to be

$$E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (3.35)$$

Here the notation $\langle i,j \rangle$ means a sum over nearest neighbors. For instance, if we are in a 2D lattice (fig. 3.11) there will be N sites, each having 2 bonds. In general, the number of bonds is Nd , assuming periodic boundary conditions. Our goal, as usual, is to compute the partition function

$$Z = \sum_{\{\sigma\}} e^{-\beta E} \quad (3.36)$$

where $\{\sigma\}$ is a short-hand notation to mean the sum over all spin configurations.

The basic idea of the mean-field approximation is to separate the mean value $\langle \sigma_i \rangle$ from its *fluctuations*

$$\delta \sigma_i = \sigma_i - \langle \sigma_i \rangle \quad (3.37)$$

Our model has translation invariance so we expect that all $\langle \sigma_i \rangle$ will have the same value:

$$\langle \sigma_i \rangle := m \quad (\text{same for all sites } i) \quad (3.38)$$

The idea is not to substitute $\sigma_i = m + \delta\sigma_i$ into the first term of Eq. (3.35). We then get

$$\begin{aligned} \sigma_i \sigma_j &= (m + \delta\sigma_i)(m + \delta\sigma_j) \\ &= m^2 + m(\delta\sigma_i + \delta\sigma_j) + \delta\sigma_i \delta\sigma_j \end{aligned}$$

So far this is all exact. Now comes the actual **mean-field approximation**: **we neglect the term $\delta\sigma_i \delta\sigma_j$** . Reason: $\delta\sigma_i$ represent *fluctuations* and we assume (hope) that fluctuations will be small. So $\delta\sigma_i \delta\sigma_j$ is quadratic in the fluctuations and hence negligible. We will therefore approximate

$$\begin{aligned} \sigma_i \sigma_j &\simeq m^2 + m(\delta\sigma_i + \delta\sigma_j) \\ &= m^2 + m[(\sigma_i - m) + (\sigma_j - m)] \\ &= -m^2 + m(\sigma_i + \sigma_j) \end{aligned}$$

The energy (3.35) then becomes

$$E \simeq J \sum_{\langle i,j \rangle} m^2 - Jm \sum_{\langle i,j \rangle} (\sigma_i + \sigma_j) - h \sum_i \sigma_i$$

Now: the sum over $\langle i, j \rangle$ contains Nd terms so the first term becomes $JNdm^2$. Moreover,

$$\sum_{\langle i,j \rangle} \sigma_i = d \sum_i \sigma_i$$

and

$$\sum_{\langle i,j \rangle} \sigma_j = d \sum_j \sigma_j = d \sum_i \sigma_i$$

Hence, the energy finally becomes

$$E = JNdm^2 - (h + 2Jdm) \sum_i \sigma_i \quad (3.39)$$

This energy now depends on $m = \langle \sigma_i \rangle$, which is a little bit weird. But we will learn how to deal with it in a second.

The important aspect of this result is that the energy is now *linear* in the σ_i . It is essentially the energy of N independent spins in an “effective magnetic field” $\tilde{h} = h + 2Jmd$. Therefore the partition function will be simply

$$Z = e^{-\beta Jm^2 Nd} \left\{ 2 \cosh[\beta(h + 2Jmd)] \right\}^N \quad (3.40)$$

We know that for independent spins $\langle \sigma_i \rangle = \tanh(\beta h)$. In our case the result must be the same, with h replaced by \tilde{h} . But we also have that $\langle \sigma_i \rangle = m$ so we conclude that

$$m = \tanh(\beta h + \beta 2Jmd) \quad (3.41)$$

This is called the **Curie-Weiss equation**. It is a self-consistent equation that needs to be solved for m .

Curie and Weiss actually arrived at this result following a different logic. For isolated spins $m = \tanh(\beta h)$. They then argued that in a ferromagnetic material the spins also feel an effective field due to the presence of all other spins around it. This field should itself depend on the magnetization so they proposed to replace

$$h \rightarrow h + \lambda m$$

where λ was a constant which they called the **molecular field constant**. The word “molecular” here is to be interpreted as “microscopic” because this is the microscopic field created by all other particles. Comparing with Eq. (3.41) we then see that $\lambda = 2Jd$. The idea of having a molecular field is actually wrong: the field would have to be unrealistically large. Instead, around 30 years after Curie and Weiss’ original model, Heisenberg showed that this type of interaction is actually electrostatic (it is the **exchange interaction**). Notwithstanding, I think that the idea of Curie and Weiss is very interesting and important to understand: it represents a *feedback mechanism* where the surroundings of a spin influence each other.

Analysis for $h = 0$

When $h = 0$ Eq. (3.41) becomes

$$m = \tanh(\beta 2Jmd) \quad (3.42)$$

This equation can be solved graphically. We have two functions, m and $\tanh(\beta 2Jmd)$ and we must find the point where they coincide. The idea is illustrated in Fig. 3.12 for three possible values of $2\beta Jd$.

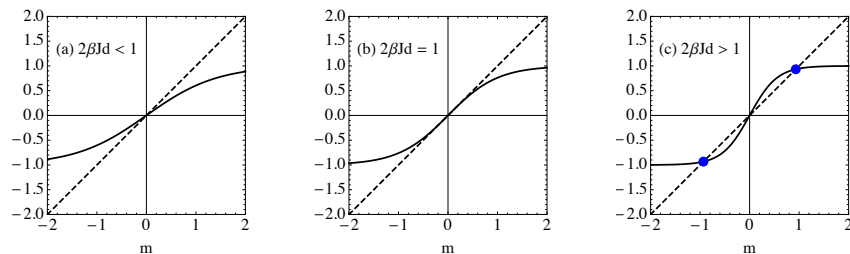


Figure 3.12: Graphical solution of Eq. (3.42) for three different values of $2\beta Jd$.

The key aspect that you need to have in mind here is the *slope* of $\tanh(2\beta Jdm)$ close to $m = 0$. When m is small we may expand

$$\tanh(x) \simeq x - \frac{x^3}{3} \quad (3.43)$$

Eq. (3.42) then becomes, approximately,

$$m \simeq 2\beta Jdm - (2\beta Jd)^3 \frac{m^3}{3} \quad (3.44)$$

One solution is always $m = 0$. However, we see that we may also have additional solutions. In particular, if we write this as

$$\frac{(2\beta Jd)^3}{3} m^2 = 2\beta Jd - 1 \quad (3.45)$$

we see that there will be other real solutions provided $2\beta Jd \geq 1$. This defines the **critical temperature**

$$\boxed{T_c = 2Jd} \quad (3.46)$$

The critical temperature increases with J , which makes sense since a stronger spin-spin interaction should lead to a higher resistance to thermal fluctuations. Moreover, it increases with d since, in higher dimensions, the number of bonds is higher.

Graphically, you should try to understand the critical temperature as being the temperature where the slope of $\tanh(2\beta Jdm)$ at $m = 0$ becomes identically 1, as in Fig. 3.12(b). Above T_c the slope will be smaller than one and there will only be a trivial solution $m = 0$ [Fig. 3.12(c)] and below T_c the slope will be greater than 1 and two new solutions will appear [Fig. 3.12(a)].

Our solution is summarized in Fig. 3.13. For $T > T_c$ we have only one possibility, which is $m = 0$. But below T_c there are three possible solutions. Later we will show that the solution with $m = 0$ becomes unstable below T_c (we don't have the tools to talk about stability just yet). Moreover, the two solutions with $m \neq 0$ have equal magnitude and just differ in sign. This is the up-down symmetry of magnetic materials. In any case, what is experimentally observed in the end is $|m|$, so instead of Fig. 3.13 we observe in practice Fig. 3.1(a).

Returning to Eq. (3.45) we may write it as

$$\frac{(T_c/T)^3}{3} m^2 = \frac{T_c}{T} - 1$$

so that the magnetization, close to the critical point, becomes

$$m^2 = \frac{3T^2}{T_c^3} (T_c - T) \quad (3.47)$$

This is valid for $T < T_c$. Above the critical temperature $m = 0$.

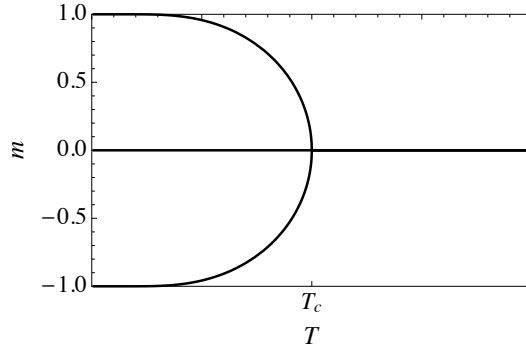


Figure 3.13: All possible solutions of Eq. (??).

The two-point correlation function is

$$\langle \sigma_n \sigma_{n+r} \rangle = \langle (m + \delta\sigma_n)(m + \delta\sigma_{n+r}) \rangle \quad (3.48)$$

But, by definition, $\langle \delta\sigma_n \rangle = 0$. Moreover, since we are discarding quadratic terms in the fluctuations, we are left only with

$$\langle \sigma_n \sigma_{n+r} \rangle = m^2 \quad (3.49)$$

The correlation function is independent of distance and is always the magnetization squared.

Analysis for $h \neq 0$

Now let us return to Eq. (3.41). Inverting Eq. (??) and writing $2\beta J = T_c$ we obtain

$$h = -T_c m + T \tanh^{-1}(m) \quad (3.50)$$

This is the **equation of state** for the system, just like for the ideal gas, where we have $pV = NT$. Here h plays the role of $-p$; ie, it is something which externally alters the state of the system, whereas m plays the role of V/N . A typical m vs. T curve for $h \neq 0$ is shown in Fig. 3.14. Close to the critical point we may expand $\tanh^{-1}(m) \simeq m + m^3/3$ and therefore write

$$h \simeq (T - T_c)m + \frac{Tm^3}{3} \quad (3.51)$$

Exactly at $T = T_c$ we see that

$$m = \left(\frac{3h}{T_c} \right)^{1/3} \quad (3.52)$$

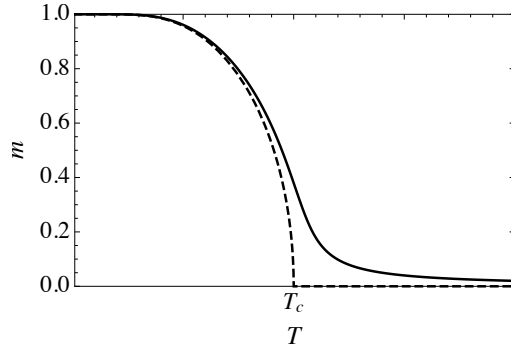


Figure 3.14: m vs. T for $h = 0$ (dashed line) and $h \neq 0$ (solid line).

Next we turn to the susceptibility:

$$\chi = \left. \frac{\partial m}{\partial h} \right|_{h=0} \quad (3.53)$$

We may obtain a formula for χ close to the critical point by differentiating both sides of Eq. (3.51). This gives

$$1 = -(T_c - T) \frac{\partial m}{\partial h} + T m^2 \frac{\partial m}{\partial h}$$

or,

$$\chi(T) = \frac{1}{(T - T_c) + T m^2} \quad (3.54)$$

This formula is illustrated in Fig. 3.15. As can be seen, the susceptibility diverges around the critical point. Above the critical point $m = 0$ and Eq. (3.54) simplifies to

$$\chi = \frac{1}{T - T_c} \quad (3.55)$$

which is known as the Curie-Weiss law. Recall that for paramagnets we used to have $\chi = C/T$. In the present case, even though at $T > T_c$ the material behaves like a paramagnet, we still see some reflections of the interactions, which modify Curie's law to $C/(T - \theta)$. In our present model $\theta = T_c$. In real materials we find that both are usually different. For $T < T_c$ we substitute Eq. (3.47) to find

$$\chi = \frac{1}{2(T_c - T)} \quad (3.56)$$

It also diverges with the same exponent and a slightly different pre-factor.

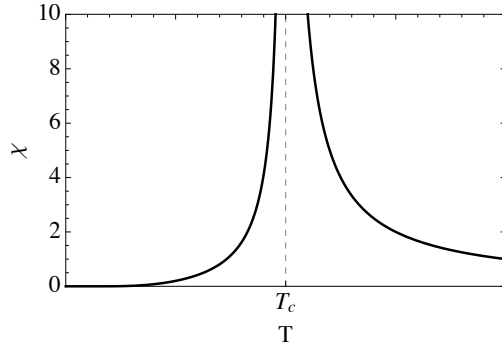


Figure 3.15: χ vs. T , computed from Eq. (3.54).

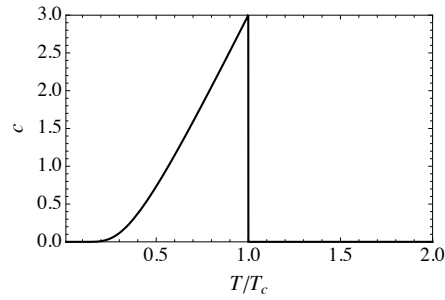


Figure 3.16: The specific heat for the mean-field model, computed using Eq. (??) together with the numerical solution of Eq. (??).

Internal energy and specific heat

The energy of the system, in the mean-field approximation, is given by Eq. (3.39). Thus, the internal energy will be

$$U = \langle E \rangle = Jm^2Nd - (h + 2Jmd)(Nm)$$

or,

$$U = -Nhm - NJdm^2 \quad (3.57)$$

The heat capacity at $h = 0$ will then be

$$C = \frac{\partial U}{\partial T} = -NJd \frac{\partial}{\partial T} m^2 \quad (3.58)$$

Above the critical temperature $m = 0$ so the specific heat will be zero. Below the critical temperature it increases with T since m decreases monotonically. This result is shown in Fig. 3.16

Critical exponents and universality

Now let me introduce to you some of the jargons of critical phenomena. The magnetization is an **order parameter**. This is the name given to any quantity which is zero in one phase but non-zero in the other. The order parameter therefore characterizes which phase you are in. The important part of critical systems is *how* the order parameter tends to zero as we approach the critical point. By “how”, I mean that we don’t care about pre-factors or silly numerical values. All we care about is the functional form of this decay. In the literature you will often find this written as

$$m \sim (T_c - T)^\beta \quad (3.59)$$

The important part is the exponent β (which is not $1/T$, sorry!). It is called a **critical exponent**. Other thermodynamic quantities also have their own critical exponents. For instance, the susceptibility diverges at the critical point from either side as

$$\chi \sim \frac{1}{|T - T_c|^\gamma} \quad (3.60)$$

and similarly for the specific heat:

$$c \sim \frac{1}{|T - T_c|^\alpha} \quad (3.61)$$

Exactly at $T = T_c$, the magnetization usually becomes algebraic in the magnetic field, changing as

$$m \sim |h|^{1/\delta} \quad (3.62)$$

Finally, the correlation length ξ diverges as

$$\xi \sim \frac{1}{|T - T_c|^\nu} \quad (3.63)$$

The use of the letters β , α , γ , δ and ν is usually standard in the literature.

Now comes the surprising part: the set of critical exponents depends *only* on the symmetries of your model and the dimensionality. This is called **Universality**. The pre-factors and the value of T_c are completely different, but the exponents are the same. For instance, the Ising model in a 2D square lattice and in a 2D triangular lattice will have the same critical exponents because both have the same symmetries (up-down in this case) and dimensionality.

This introduces the idea of **Universality classes**. We say, for instance, that the mean-field approximation defines the *mean-field universality class*. Any mean-field approximation that you do in your life, irrespective of what kind of physical system you are considering, *always* has the same set of critical exponents as the mean-field model. It doesn’t even have to be a model in magnetism. For the mean-field the exponents can be read off from Eqs. (3.47), (3.49), (3.52), (3.56) and (3.58). They are summarized in Table (3.2) together with the exponents for the 2D and 3D Ising models. The former was computed exactly and the latter by Monte Carlo simulations. This is why the 2D exponents are given as rational numbers and the 3D are not.

Table 3.2: Critical exponents, Eqs. (3.59)-(3.63). Copied without permission from Nishimori and Ortiz, *Elements of Phase Transitions and Critical Phenomena*.

Exponent	Mean-field	Ising 2D	Ising 3D
α	0	0	0.110
β	1/2	1/8	0.3265
γ	1	7/4	1.2372
δ	3	15	4.789
ν	1/2	1	0.63

3.6 The Landau free energy and hysteresis

The partition function in the mean-field approximation is given in Eq. (3.40). The free energy will then be $F = -T \ln Z$ or

$$F = \frac{NT_c}{2}m^2 - NT \ln 2 - NT \ln \left[\cosh \left(\frac{h + mT_c}{T} \right) \right] \quad (3.64)$$

where I already used the fact that $2Jd = T_c$. The free energy per spin will be

$$f = \frac{F}{N} = \frac{T_c}{2}m^2 - T \ln 2 - T \ln \left[\cosh \left(\frac{h + mT_c}{T} \right) \right] \quad (3.65)$$

The free energy is now a function of m . This only happens when we are in the mean-field approximation (usually, F is only a function of h and T). We already know that thermodynamic equilibrium corresponds to the state which minimizes the free energy. Hence, in thermal equilibrium the value of m and f will be determined from

$$\frac{\partial f}{\partial m} = 0 \quad (3.66)$$

Taking the derivative of Eq. (3.65) we then find

$$\frac{\partial f}{\partial m} = T_c m - T_c \tanh \left(\frac{h + mT_c}{T} \right) = 0 \quad (3.67)$$

which is precisely Eq. (3.41). This shows that the solutions plotted in Fig. 3.13 are nothing but the extrema of the free energy.

Now let us assume that we are close to the critical point so that m is small. Moreover, we will also assume that h is small. This allows us to expand the last term as

$$\ln[\cosh(x)] \simeq \frac{x^2}{2} - \frac{x^4}{12} \quad (3.68)$$

When we do that we will get a power series expansion in both m and h . For the terms which contain m alone, we will keep the expansion up to order m^4 (for reasons that will become clear in a second). But for terms which contain

h , we will only retain the term which is linear in h . If you have a little patience to work out the expansion, you will find that this leaves us with

$$f \simeq -T \ln 2 - \frac{T_c}{T} m h + \frac{T_c}{2T} (T - T_c) m^2 + \frac{T_c^4}{12T^3} m^4$$

Since we are talking about an expansion close to criticality, we may also approximate $T_c/T \simeq 1$. Finally, we get rid of the first term, $-T \ln 2$, since it is independent of m . We then finally arrive at

$$f \simeq -m h + \frac{(T - T_c)}{2} m^2 + \frac{T_c}{12} m^4 \quad (3.69)$$

This is called the **Landau free energy**. When $h = 0$ the free energy will preserve up-down symmetry since it only contains even terms in the magnetization. The magnetic field term is odd in m and therefore breaks this symmetry, as we would expect.

Eq. (3.69) allows for a very interesting interpretation of ferromagnetism, due to Landau. It is based on the *sign* of the quadratic term. In Fig. 3.17 we plot f as a function of m for $T > T_c$ (paramagnetic phase) and $T < T_c$ (ferromagnetic phase), for the case $h = 0$. In the paramagnetic phase $T - T_c > 0$ so the only minimum of the Landau free energy is at $m = 0$. But in the ferromagnetic phase this coefficient becomes negative, causing the appearance of two minima at non-zero values of m . These new minima represent the *spontaneous* ferromagnetic phases. They are precisely the solutions in Fig. 3.13. When we discussed Fig. 3.13 I said that the solution with $m = 0$ was unstable in the ferromagnetic phase. Now you can see why: it corresponds to a maximum, not a minimum, of the free energy.

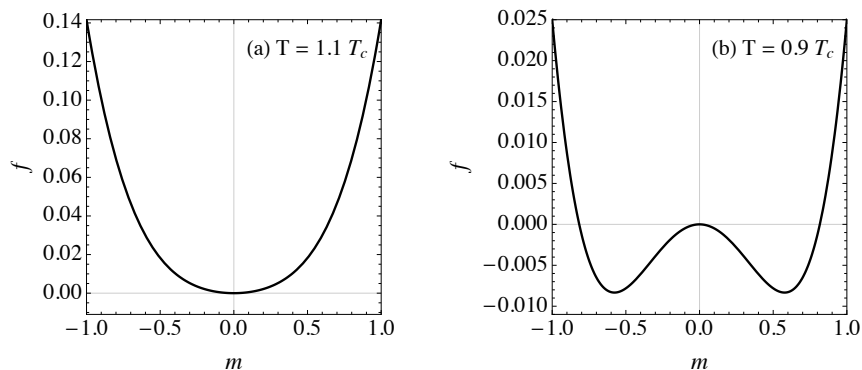


Figure 3.17: The Landau free energy f , Eq. (3.69) as a function of m for (a) $T > T_c$ (paramagnetic phase) and (b) $T < T_c$ (ferromagnetic phase).

Landau arrived at Eq. (3.69) using only symmetry arguments. His basic idea is:

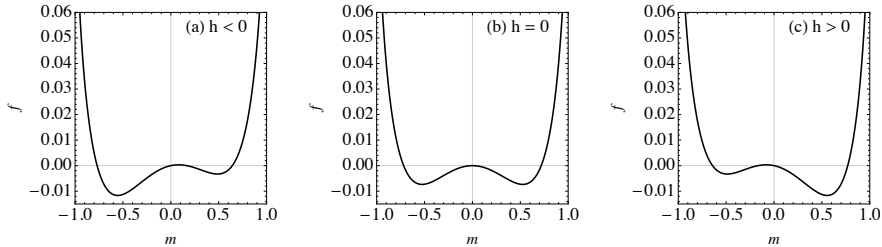


Figure 3.18: The Landau free energy (3.69) for non-zero magnetic fields in the ferromagnetic phase ($T < T_c$).

1. The free energy should be an analytic function of the order parameter and may therefore be expanded in a power series close to the critical point.
2. The free energy must have the same symmetries of the original Hamiltonian.

In our case the system must be invariant under $m \rightarrow -m$. Thus, an expansion of $f(m)$ for $h = 0$ may only contain *even* terms of m . Something like:

$$f = a_0 + a_2 m^2 + a_4 m^4 + a_6 m^6 + \dots \quad (3.70)$$

The coefficients a_i will be functions of T . In order for us to have a phase transition at $T = T_c$, the coefficient a_2 must change sign. Hence, it must be proportional to $T - T_c$. The coefficient a_4 , on the other hand, must always remain positive to ensure the stability of the free energy. This type of reasoning can be applied to all phase transitions. It is called the **Landau theory of phase transitions**. All we need to do is identify the order parameter and then expand the free energy while preserving the symmetries of the original Hamiltonian.

Now I want to discuss a very interesting concept called **spontaneous symmetry breaking**. Look again at Fig. 3.17. There seems to be a contradiction between this figure and experiment. The free energy has up-down symmetry (it is invariant under exchange of $m \rightarrow -m$). However, in practice we either find the system magnetized one way or another. Thus, even though the states with $m > 0$ and $m < 0$ should be equally likely, the system “chooses” one of these two states to stay on. The up-down symmetry has been *spontaneously broken*.

To understand why this happens, consider the effect of adding a non-zero field (Fig. 3.18). The field breaks up-down symmetry, making one of the minima deeper than the other. Now imagine that we start with a big and positive field, pushing the magnetization to the positive side. Then we start to slowly reduce the field until we reach $h = 0$. We will then have restored up-down symmetry but the magnetization will continue to be on the positive side. Thus, by applying a field, we have forced the system to choose one of two solutions.

Let $\pm m^*$ denote the two minima of f [ie, the two solutions we found in Eq. (3.47)]. If the system is in $+m^*$, there is a certain probability that jumps

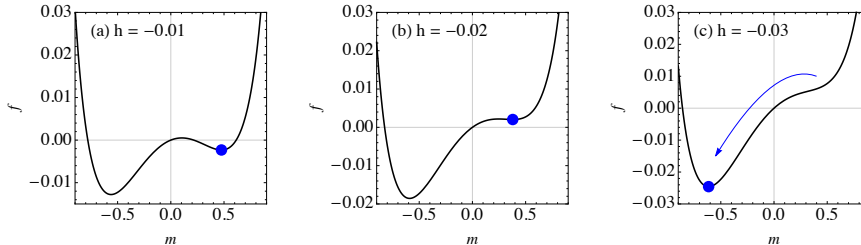


Figure 3.19: The Landau free energy f , Eq. (3.69) as a function of m for $T = 0.9T_c$ and several values of h (in units of T_c).

toward the $-m^*$ state. But to do that, it must overcome an *energy barrier*. The energy barrier is the difference $\Delta F = F(+m^*) - F(0)$ [cf. Fig. 3.17]. Using Eq. (3.47) we find

$$\Delta F = \frac{9N(T - T_c)^2}{4T_c} \quad (3.71)$$

If $T = T_c$ the energy barrier is zero. But for any $T < T_c$, the energy barrier will be proportional to the number of particles N . Hence, even though the probability exists, in the thermodynamic limit it will become vanishingly small: the barrier becomes insurmountable. In other words, the time the system will take to jump from $+m^*$ to $-m^*$ becomes infinite. Thus, in practice, we have gone from a situation where the two states are equally likely, to a situation where the system only chooses one of the two states: the up-down symmetry has been spontaneously broken.

Hysteresis

Suppose that the system is in the $+m^*$ state and we start to apply a small *negative* field. This will not instantaneously push the magnetization to $-m^*$ since there will still be an energy barrier to surmount. In 1948, Stoner and Wolfarth proposed that the jump should occur only when the energy barrier separating the two minima becomes a saddle point,⁴ i.e., when the shallowest of the two minima disappear. This is illustrated in Fig. 3.19. The magnetization is initially on the right well, and it will stay there until the field is strong enough to completely destroy this minimum. Then it will have no energy barrier to surmount and will simply slide down to the now unique global minimum. Notice how this introduces a *memory* effect. The state of m is no longer uniquely determined by h . But rather, it depends on the previous history of the sample.

This is the idea behind hysteresis. If we continue to follow this procedure we will obtain a **hysteresis loop**, like the one shown in Fig. 3.20. I computed this loop using only Eq. (3.69), by forcing the magnetization to switch from one minima to another only when there was no barrier between them. The field where the switch occurs is the **coercive field** h_c .

⁴See E. C. Stoner and E. P. Wolfarth, *Phil. Trans. Roy. Soc.* **A240** (1948) 599

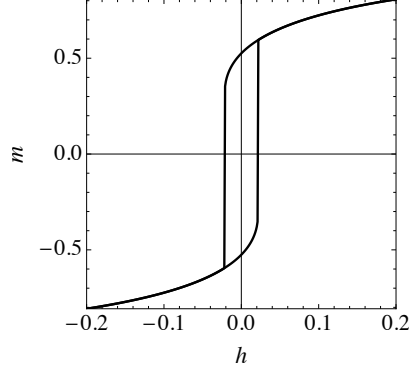


Figure 3.20: A hysteresis loop computed using the Stoner and Wolfarth model, for $T_c = 0.9$.

Relaxation to equilibrium

If we put the magnetization of the system in a non-equilibrium state, it will slowly relax toward the thermal equilibrium solutions $\pm m^*$. We may obtain a rough description of *how* this relaxation takes place using the Landau free energy. The condition for equilibrium is that $\partial f / \partial m = 0$. Therefore, close to equilibrium we may assume that the dynamics of $m(t)$ may be described by

$$\tau \frac{dm}{dt} = -\frac{\partial f}{\partial m} \quad (3.72)$$

where τ is a constant measuring the time scale of the process (this is usually difficult to calculate analytically, but it can always be measured experimentally). This equation describes a sort of viscous relaxation toward equilibrium. It is certainly phenomenological, but it does give fairly good results.

Let us take $h = 0$. Close to the critical value, we may use Eq. (3.67) to write

$$\tau \frac{dm}{dt} = -(T - T_c)m - \frac{T_c m^3}{3}$$

Multiplying on both sides by m and writing $m \frac{dm}{dt} = \frac{1}{2} \frac{dm^2}{dt}$ we get

$$\frac{\tau}{2} \frac{dm^2}{dt} = -(T - T_c)m^2 - \frac{T_c m^4}{3}$$

The solution of this equation is

$$m(t)^2 = \frac{T - T_c}{\alpha e^{2(T - T_c)t/\tau} - T_c/3} \quad (3.73)$$

where α is a constant determined from the initial conditions. If $T > T_c$ the exponential is positive and we find that m relaxes exponentially to zero as $m(t) \sim e^{-(T-T_c)t/\tau}$, in agreement with what we expect from a paramagnetic phase. Conversely, if $T < T_c$ we see that $m(t)$ tends to its equilibrium value in Eq. (3.47).

Special care must be taken to the solution exactly at the critical point. We then have instead

$$\tau \frac{dm}{dt} = -\frac{T_c m^3}{3}$$

The general solution of this equation is

$$m(t) = \frac{1}{2\frac{T_c}{3}\frac{t}{\tau} + \lambda}$$

where λ is another constant. Hence, at the critical point the magnetization also relaxes to zero, but it does so *algebraically*. This is very common of critical phenomena: outside criticality the relaxation is exponential, but at the critical point it becomes algebraic.