

Phase Transitions and the Ising Model

Magnet, n. Something acted upon by magnetism.

Magnetism, n. Something acting upon a magnet.

The two definitions immediately foregoing are condensed from the works of one thousand eminent scientists, who have illuminated the subject with a great white light, to the inexpressible advancement of human knowledge.

Ambrose Bierce, in *The Devil's Dictionary*

The Ising model is deceptively simple. It can be defined in a few words, but it displays astonishingly rich behavior. It originated as a model of ferromagnetism in which the magnetic moments were localized on lattice sites and had only two allowed values. This corresponds, of course, to a spin-one-half model, but since it is tiresome to continually write factors of $\hbar/2$, it is traditional to take the values of a ‘spin’ σ_j , located on lattice site j to be

$$\sigma_j = +1 \text{ or } -1. \quad (31.1)$$

The energy of interaction with a magnetic field also requires a factor of the magnetic moment of the spin, but to make the notation more compact, everything will be combined into a single ‘magnetic field’ variable h , which has units of energy,

$$H_h = -h\sigma_j. \quad (31.2)$$

Eq. (31.2) implies that the low-energy state is for $\sigma_j = +1$ when $h > 0$.

There should not be any confusion of the ‘ h ’ in eq. (31.2) with Planck’s constant, because the latter does not appear in this chapter.

The energy of interaction between spins on neighboring lattice sites, j and k , is given by the product of the spins,

$$H_{\mathcal{J}} = -\mathcal{J}\sigma_j\sigma_k. \quad (31.3)$$

The constant \mathcal{J} is taken to have units of energy. The low-energy states of eq. (31.3) are $\sigma_j = \sigma_k = +1$ and $\sigma_j = \sigma_k = -1$ for $\mathcal{J} > 0$. This form of interaction is also called an ‘exchange’ interaction, due to the role of exchanging electrons in its derivation from the quantum properties of neighboring atoms. That derivation is, of course, beyond the scope of this book.

If we put the two kinds of interactions together on a lattice, we can write the Hamiltonian as

$$H = -\mathcal{J} \sum_{\langle j,k \rangle} \sigma_j \sigma_k - h \sum_{j=1}^N \sigma_j \quad (31.4)$$

where the notation $\langle j,k \rangle$ denotes that the sum is over nearest-neighbor pairs of sites. In the rest of this chapter we will discuss the remarkable properties of the model described by eq. (31.4).

31.1 The Ising Chain

The first calculations on the Ising model were carried out by Ernst Ising (German physicist, 1900–1998), who was given the problem by his adviser, Wilhelm Lenz (German physicist, 1888–1957). In 1924 Ising solved a one-dimensional model with N spins, in which each spin interacted with its nearest neighbors and a magnetic field,

$$H = -\mathcal{J} \sum_j \sigma_j \sigma_{j+1} - h \sum_{j=1}^N \sigma_j. \quad (31.5)$$

This model is also referred to as an ‘Ising chain’. The limits on the first sum have not been specified explicitly in eq. (31.5) to allow it to describe two kinds of boundary conditions.

1. Open boundary conditions

The first sum goes from $j = 1$ to $j = N - 1$. This corresponds to a linear chain for which the first spin only interacts with the second spin, and the spin at $j = N$ spin only interacts with the spin at $j = N - 1$.

2. Periodic boundary conditions

A spin $\sigma_{N+1} = \sigma_1$ is defined, and the first sum goes from $j = 1$ to $j = N$. This corresponds to a chain of interactions around a closed loop.

In both cases, the second sum runs from $j = 1$ to $j = N$ to include all spins.

We will investigate various aspects of this model in the following sections before going on to discuss the (qualitatively different) behavior in more than one dimension. We begin by ignoring the interactions and setting $\mathcal{J} = 0$ in the next section.

31.2 The Ising Chain in a Magnetic Field ($J = 0$)

Begin by setting $\mathcal{J} = 0$ in eq. (31.5),

$$H = -h \sum_{j=1}^N \sigma_j. \quad (31.6)$$

The canonical partition function is then given by

$$Z = \sum_{\{\sigma\}} \exp[-\beta(-h \sum_{j=1}^N \sigma_j)], \quad (31.7)$$

where we have used the notation that $\{\sigma\} = \{\sigma_j | j = 1, \dots, N\}$ is the set of all spins.

Eq. (31.7) can be simplified by using the best trick in statistical mechanics (factorization of the partition function: Sections 19.12 and 24.7),

$$Z = \sum_{\{\sigma\}} \prod_{j=1}^N \exp[\beta h \sigma_j] = \prod_{j=1}^N \sum_{\sigma_j} \exp[\beta h \sigma_j] = \prod_{j=1}^N Z_j, \quad (31.8)$$

where

$$Z_j = \sum_{\sigma_j} \exp[\beta h \sigma_j]. \quad (31.9)$$

This reduces the N -spin problem to N -identical single-spin problems, each of which involves only a sum over two states.

Since each $\sigma_j = \pm 1$, the sums can be carried out explicitly:

$$Z_j = \exp(\beta h) + \exp(-\beta h) = 2 \cosh(\beta h). \quad (31.10)$$

It is easy to show that the average magnetization is given by

$$m = m_j = \langle \sigma_j \rangle = \frac{\exp(\beta h) - \exp(-\beta h)}{\exp(\beta h) + \exp(-\beta h)} = \tanh(\beta h). \quad (31.11)$$

Since it will be important later to be familiar with the graphical form of eq. (31.11), it is shown in Fig. 31.1. The shape of the curve makes sense in that it shows that $m = 0$ when $h = 0$, while $m \rightarrow \pm 1$ when $h \rightarrow \pm \infty$.

The average energy of a single spin is then easily found,

$$U_j = \langle -h \sigma_j \rangle = -h m_j = -h m = -h \tanh(\beta h). \quad (31.12)$$

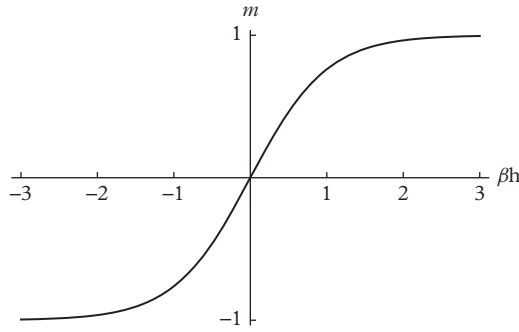


Fig. 31.1 Plot of the magnetization of an isolated spin against the magnetic field, $m = \tanh(\beta h)$.

The energy of the entire Ising chain is given by simply multiplying by N ,

$$U = \langle H \rangle = -h \sum_{j=1}^N \langle \sigma_j \rangle = -hNm = -hN \tanh(\beta h). \quad (31.13)$$

The specific heat is found by differentiating with respect to the temperature,

$$c = \frac{1}{N} \frac{\partial U}{\partial T} = \frac{-1}{Nk_B T^2} \frac{\partial U}{\partial \beta} = k_B \beta^2 h^2 \operatorname{sech}^2(\beta h). \quad (31.14)$$

The specific heat goes to zero at both $T = 0$ and $T = \infty$.

31.3 The Ising Chain with $h = 0$, but $J \neq 0$

Next, we will put the interactions between spins back into the Ising chain, but remove the magnetic field. To simplify the problem we will use open boundary conditions

$$H' = -J \sum_{j=1}^{N-1} \sigma_j \sigma_{j+1}. \quad (31.15)$$

At first sight, finding the properties in this case looks considerably more difficult than what we did in the previous section. Each spin is linked to its neighbor, so the Hamiltonian cannot be split up in the same way to factorize the partition function. However, we can factorize the partition function if we introduce new variables.

Let $\tau_1 = \sigma_1$, and $\tau_j = \sigma_{j-1} \sigma_j$ for $2 \leq j \leq N$. Clearly, each τ_j takes on the values ± 1 . Given the set of N values $\{\tau_j | j = 1, \dots, N\}$, we can uniquely find the original values of the σ_j 's.

The Hamiltonian in eq. (31.15) can now be rewritten as

$$H = -\mathfrak{J} \sum_{j=2}^N \tau_j. \quad (31.16)$$

This makes the Hamiltonian look just like eq. (31.6), which we derived in the previous section for the case of independent spins in a magnetic field. Although the $N-1$ τ -variables all appear in the Hamiltonian, σ_1 does not. Summing over $\sigma_1 = \pm 1$ in the partition function is responsible for an overall factor of two in the equations below.

Because eq. (31.16) is a sum of independent terms, the partition function can again be calculated by factorization.

$$Z = \sum_{\{\tau\}} \prod_{j=1}^N \exp[\beta \mathfrak{J} \tau_j] = 2 \prod_{j=2}^N \sum_{\tau_j} \exp[\beta \mathfrak{J} \tau_j] = 2 \prod_{j=2}^N Z'_j \quad (31.17)$$

where

$$Z'_j = \sum_{\tau_j=\pm 1} \exp[\beta \mathfrak{J} \tau_j] = \exp[\beta \mathfrak{J}] + \exp[-\beta \mathfrak{J}] = 2 \cosh(\beta \mathfrak{J}) \quad (31.18)$$

and

$$Z = 2(Z'_j)^{N-1} = 2(2 \cosh(\beta \mathfrak{J}))^{N-1}. \quad (31.19)$$

The Massieu function, $\tilde{S}[\beta]$, for the Ising chain with $h = 0$ is given by the logarithm of the partition function in eq. (31.19)

$$\tilde{S}[\beta] = \ln Z = \ln 2 + (N-1) \ln Z'_j = \ln 2 + (N-1) \ln (2 \cosh(\beta \mathfrak{J})). \quad (31.20)$$

Relating the partition function to a Massieu function is preferable here because this model exhibits negative temperatures.

The energy of the system is easily found to be

$$U = -(N-1)\mathfrak{J} \tanh(\beta \mathfrak{J}), \quad (31.21)$$

and the specific heat per spin is

$$c = k_B \beta^2 \mathfrak{J}^2 \operatorname{sech}^2(\beta \mathfrak{J}) (1 - 1/N) \approx k_B \beta^2 \mathfrak{J}^2 \operatorname{sech}^2(\beta \mathfrak{J}). \quad (31.22)$$

Note that the form of eq. (31.22) is essentially the same as that of eq. (31.14), which we found in the previous section. The specific heat in this case also goes to zero at both $T = 0$ and $T = \infty$.

31.4 The Ising Chain with both $J \neq 0$ and $h \neq 0$

The full Hamiltonian in eq. (31.5) with both exchange interactions and a magnetic field is more challenging. We cannot factorize the partition function in the same way that we have done with other problems, so we will have to use a different approach, known as the Transfer Matrix method. The basis of the method is that the partition function can be expressed as a product of relatively small matrices—in this case, just 2×2 matrices—which can then be diagonalized and their determinants evaluated.

31.4.1 Transfer Matrix

We will consider the case of periodic boundary conditions, so that $\sigma_{N+1} = \sigma_1$, as discussed earlier. To make the matrices more symmetric, we will write the Hamiltonian, eq. (31.5), in the form

$$H = -J \sum_{j=1}^N \sigma_j \sigma_{j+1} - \frac{h}{2} \sum_{j=1}^N (\sigma_j + \sigma_{j+1}). \quad (31.23)$$

The upper limits on the both summations is N , so that the last terms in the sums are $-J\sigma_N\sigma_1$ and $(h/2)(\sigma_N + \sigma_1)$, due to the periodic boundary conditions.

Using eq. (31.23), we can transform the partition function into a product of 2×2 matrices,

$$\begin{aligned} Z &= \sum_{\{\sigma\}} \exp \left(-\beta \left(-J \sum_j \sigma_j \sigma_{j+1} - \frac{1}{2} h \sum_j (\sigma_j + \sigma_{j+1}) \right) \right) \\ &= \sum_{\{\sigma\}} \prod_j \exp \left(\beta J \sigma_j \sigma_{j+1} + \frac{1}{2} \beta h (\sigma_j + \sigma_{j+1}) \right) \\ &= \sum_{\{\sigma\}} \prod_j T(\sigma_j, \sigma_{j+1}). \end{aligned} \quad (31.24)$$

In this equation,

$$T(\sigma_j, \sigma_{j+1}) = \exp \left(\beta J \sigma_j \sigma_{j+1} + \frac{1}{2} \beta h (\sigma_j + \sigma_{j+1}) \right), \quad (31.25)$$

is defined as the transfer matrix. Note that we have not exchanged the sums and products, because we have not eliminated the coupling between variables. The sums remain and can be interpreted as matrix multiplication. As the name implies, we can also write the transfer matrix explicitly as a matrix,

$$T(\sigma_j, \sigma_{j+1}) = \begin{pmatrix} \exp(\beta\mathcal{J} + \beta h) & \exp(-\beta\mathcal{J}) \\ \exp(-\beta\mathcal{J}) & \exp(\beta\mathcal{J} - \beta h) \end{pmatrix}. \quad (31.26)$$

Eq. (31.24) is equivalent to representing the partition function as a product of matrices. To see this more clearly, consider the factors involving only three spins, somewhere in the middle of the chain, at locations $j-1, j$, and $j+1$,

$$\sum_{\sigma_j=\pm 1} T(\sigma_{j-1}, \sigma_j) T(\sigma_j, \sigma_{j+1}) = T^2(\sigma_{j-1}, \sigma_{j+1}) \quad (31.27)$$

The expression on the left of eq. (31.27) can also be written in terms of the explicit matrices from eq. (31.26),

$$\begin{pmatrix} \exp(\beta\mathcal{J} + \beta h) & \exp(-\beta\mathcal{J}) \\ \exp(-\beta\mathcal{J}) & \exp(\beta\mathcal{J} - \beta h) \end{pmatrix} \begin{pmatrix} \exp(\beta\mathcal{J} + \beta h) & \exp(-\beta\mathcal{J}) \\ \exp(-\beta\mathcal{J}) & \exp(\beta\mathcal{J} - \beta h) \end{pmatrix}. \quad (31.28)$$

Carrying out the matrix multiplication explicitly, we find

$$T^2(\sigma_{j-1}, \sigma_{j+1}) = \begin{pmatrix} \exp(2\beta\mathcal{J} + 2\beta h) + \exp(-2\beta\mathcal{J}) & \exp(\beta h) + \exp(-\beta h) \\ \exp(\beta h) + \exp(-\beta h) & \exp(2\beta\mathcal{J} - 2\beta h) + \exp(-2\beta\mathcal{J}) \end{pmatrix}. \quad (31.29)$$

If we carry out the sums over σ_2 through σ_N , we find the matrix $T^N(\sigma_1, \sigma_{N+1})$. Since $\sigma_{N+1} = \sigma_1$, the final sum over σ_1 yields the partition function

$$Z = \sum_{\sigma_1=\pm 1} T^N(\sigma_1, \sigma_1). \quad (31.30)$$

Fortunately, we do not have to carry out all the matrix multiplications explicitly. That is the power of the Transfer Matrix method—as we will show in the next section.

31.4.2 Diagonalizing $T(\sigma_j, \sigma_{j+1})$

The next step is the one that makes this calculation relatively easy. The trace is invariant under a similarity transformation of the form

$$\tilde{T} = RTR^{-1}. \quad (31.31)$$

In this equation, R and its inverse R^{-1} are both 2×2 matrices. If we write out the product of T 's we can see how it reduces to the original product of T 's. The series of products brings the R and R^{-1} matrices together, so that their product gives unity. Using the compact notation, $T(\sigma_j, \sigma_{j+1}) \rightarrow T_{jj+1}$, we see that

$$\tilde{T}_{1,2}\tilde{T}_{2,3}\tilde{T}_{3,4} = RT_{1,2}R^{-1}RT_{2,3}R^{-1}RT_{3,4}R^{-1} = RT_{1,2}T_{2,3}T_{3,4}R^{-1}. \quad (31.32)$$

This procedure can be extended to arbitrarily many T matrices. Since we have periodic boundary conditions, after applying the transformation eq. (31.31) N times and using $R^{-1}R = 1$ between the T -matrices, eq. (31.30) becomes

$$Z = \text{Tr} R \tilde{T}^N R^{-1} = \text{Tr} \tilde{T}^N. \quad (31.33)$$

For the last equality, we have used the fact that the trace of a product of matrices is invariant under a cyclic permutation.

Finding the trace is much easier if we capitalize on the freedom of making a similarity transformation to diagonalize the T matrix. The diagonal elements of T' will then be given by the eigenvalues λ_+ and λ_- ,

$$\tilde{T}(\sigma_j, \sigma_{j+1}) = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}. \quad (31.34)$$

The product of N diagonal matrices is trivial

$$(T')^N = \begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix}, \quad (31.35)$$

and the partition function becomes

$$Z = \lambda_+^N + \lambda_-^N. \quad (31.36)$$

The only remaining task is to find the eigenvalues of the T matrix in eq. (31.25). This involves solving the eigenvalue equation

$$T\chi = \lambda\chi, \quad (31.37)$$

or

$$\begin{pmatrix} \exp(\beta\mathcal{J} + \beta h) & \exp(-\beta\mathcal{J}) \\ \exp(-\beta\mathcal{J}) & \exp(\beta\mathcal{J} - \beta h) \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \lambda \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \quad (31.38)$$

This matrix equation is equivalent to two simultaneous equations for λ , and can be solved by setting the determinant equal to zero,

$$\begin{vmatrix} \exp(\beta\mathcal{J} + \beta h) - \lambda & \exp(-\beta\mathcal{J}) \\ \exp(-\beta\mathcal{J}) & \exp(\beta\mathcal{J} - \beta h) - \lambda \end{vmatrix} = 0. \quad (31.39)$$

Solving the resulting quadratic equation—and leaving the algebra to the reader—we find the required eigenvalues,

$$\lambda_{\pm} = e^{\beta\mathcal{J}} \cosh(\beta h) \pm \sqrt{e^{2\beta\mathcal{J}} \cosh^2(\beta h) - 2 \sinh(2\beta\mathcal{J})}. \quad (31.40)$$

The partition function of a system of N spins is given by inserting eq. (31.40) into eq. (31.36).

It is interesting to see how this form of the partition function behaves in the limit that $N \rightarrow \infty$ (the thermodynamic limit). Since $\lambda_+ > \lambda_-$, the ratio $\lambda_-/\lambda_+ < 1$. We can use this fact to rewrite eq. (31.36) in a convenient form

$$Z = \lambda_+^N \left(1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right). \quad (31.41)$$

The Massieu function then takes on a simple form

$$\begin{aligned} \tilde{S}[\beta] &= \ln Z \\ &= N \ln \lambda_+ + \ln \left(1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right) \\ &\approx N \ln \lambda_+ + \left(\frac{\lambda_-}{\lambda_+} \right)^N. \end{aligned} \quad (31.42)$$

The Massieu function per spin in the thermodynamic limit is just

$$\frac{\tilde{S}[\beta]}{N} \approx \ln \lambda_+ + \frac{1}{N} \left(\frac{\lambda_-}{\lambda_+} \right)^N \rightarrow \ln \lambda_+,$$

or, inserting λ_+ from eq. (31.40),

$$\lim_{N \rightarrow \infty} \frac{\tilde{S}[\beta]}{N} = \ln \left(e^{\beta\mathcal{J}} \cosh(\beta h) + \sqrt{e^{2\beta\mathcal{J}} \cosh^2(\beta h) - 2 \sinh(2\beta\mathcal{J})} \right). \quad (31.43)$$

Perhaps the most important thing to notice about eq. (31.43) is that the Massieu function is an analytic function of the temperature everywhere except at $\beta = \pm\infty$. Since a phase transition can be defined as a point at which the free energy is not analytic in the limit of an infinite system (thermodynamic limit), this means that the one-dimensional Ising model does not exhibit a phase transition at any non-zero temperature.

This lack of a phase transition was one of the main results that Ising found (by a method different from that which we have used). Based on this result, he speculated that the model would not show a phase transition in higher dimensions, either. In this speculation, he was wrong. However, it took from 1924, when Ising did his thesis work,

until 1944, when Lars Onsager (Norwegian chemist and physicist, 1903–1976, who became an American citizen in 1945) derived an exact solution for the two-dimensional Ising model (without a magnetic field) that showed a phase transition.

Although Onsager's exact solution of the Ising model is beyond the scope of this book, we will discuss some approximate methods for calculating the behavior of the Ising model at its phase transition, beginning with the Mean-Field Approximation in the next section.

31.5 Mean-Field Approximation

The history of the Mean-Field Approximation (MFA) is interesting in that it precedes the Ising model by seventeen years. In 1907, Pierre-Ernest Weiss (French physicist, 1865–1940) formulated the idea of describing ferromagnetism with a model of small magnetic moments arranged on a crystal lattice. He assumed that each magnetic moment should be aligned by an average 'molecular field' due to the other magnetic moments, so that the idea was originally known as the 'Molecular Field Theory'. The justification of the model was rather vague, but comparison with experiment showed it to be remarkably good.

In the following subsections we will derive the Mean-Field Approximation for the Ising model in an arbitrary number of dimensions,

$$H = -J \sum_{\langle j,k \rangle} \sigma_j \sigma_k - h \sum_{j=1}^N \sigma_j. \quad (31.44)$$

Recall that the notation $\langle j, k \rangle$ indicates that the summation in the first term is taken over the set of all nearest-neighbor pairs of sites in the lattice.

31.5.1 MFA Hamiltonian

The basic idea of MFA is to approximate the effects of interactions on a spin by its neighbors as an average (or 'mean') magnetic field. Specifically, we would like to calculate an approximation to the average magnetization of a spin at site j ,

$$m_j = \langle \sigma_j \rangle. \quad (31.45)$$

For simplicity, we will assume periodic boundary conditions, so that the properties of the model are translationally invariant. Since all spins see the same average environment, they should all have the same average magnetization, m ,

$$m = \frac{1}{N} \sum_k m_k = m_j. \quad (31.46)$$

For any given spin, σ_j , we can express the sum over the nearest-neighbor sites as a sum over the neighbors δ , and extract all terms in eq. (31.44) that contain σ_j ,

$$H_j = -\mathcal{J} \sum_{\delta} \sigma_j \sigma_{j+\delta} - h \sigma_j = - \left(\mathcal{J} \sum_{j+\delta} \sigma_{j+\delta} + h \right) \sigma_j. \quad (31.47)$$

Although eq. (31.47) can be regarded as a reduced Hamiltonian for spin σ_j , we cannot interpret its average value as representing the average energy per spin, since $H \neq \sum_j H_j$ because of double counting the exchange interactions. This is a mental hazard that has caught many students and even a few textbook authors.

Note that eq. (31.47) has the same form as the Hamiltonian for a single spin in a magnetic field, $-h\sigma$, from eq. (31.2), except that the external magnetic field is replaced by an effective magnetic field,

$$h_{j,\text{eff}} = \mathcal{J} \sum_{\delta} \sigma_{j+\delta} + h. \quad (31.48)$$

The difficulty with the effective field found in eq. (31.48) is that it depends on the instantaneous values of the neighboring spins, which are subject to fluctuations. MFA consists of replacing the fluctuating spins in eq. (31.48) by their average (or mean) values, which must all be equal to the magnetization m , due to translational invariance,

$$h_o^{MFA} = \mathcal{J} \sum_{\delta} \langle \sigma_{j+\delta} \rangle + h = \mathcal{J} z m + h. \quad (31.49)$$

In the second equality in eq. (31.49) we have introduced the parameter z to denote the number of nearest-neighbor sites in the lattice of interest ($z = 4$ for a two-dimensional square lattice). The MFA Hamiltonian,

$$H_o^{MFA} = -h_o^{MFA} \sigma_o, \quad (31.50)$$

has exactly the same form as eq. (31.2). An equation for the magnetization can be written down immediately in analogy to eq. (31.11),

$$m = \tanh(\beta h_o^{MFA}) = \tanh(\beta \mathcal{J} z m + \beta h). \quad (31.51)$$

Unfortunately, there is no closed-form solution for eq. (31.51). However, there is a nice graphical solution that makes it relatively easy to understand the essential behavior of the solution, as well as guiding us to good approximations.

31.5.2 Graphical Solution of eq. (31.51) for $h = 0$

First consider the case of a vanishing magnetic field, $h = 0$, for which eq. (31.51) simplifies,

$$m = \tanh(\beta \mathcal{J} z m). \quad (31.52)$$

It might seem strange, but we can further simplify the analysis of eq. (31.52) by expressing it as two equations,

$$m = \tanh(x) \quad (31.53)$$

$$x = \beta \mathcal{J} z m. \quad (31.54)$$

The second equation can then be rewritten as

$$m = \left(\frac{k_B T}{z \mathcal{J}} \right) x. \quad (31.55)$$

Since eqs. (31.53) and (31.55) are both equations for m as a function of x , it is natural to plot them on the same graph, as shown in Fig. 31.2. Intersections of the two functions then represent solutions of eq. (31.52).

When the slope of the line described by eq. (31.55) equals 1, it provides a dividing line between two qualitatively different kinds of behavior. The temperature at which this happens is known as the critical temperature, T_c , or the Curie temperature, in honor of Pierre Curie (French physicist, 1859–1906, awarded the 1903 Nobel Prize, together with his wife, Maria Sklodowska-Curie, 1867–1934)

$$k_B T_c = z \mathcal{J}. \quad (31.56)$$

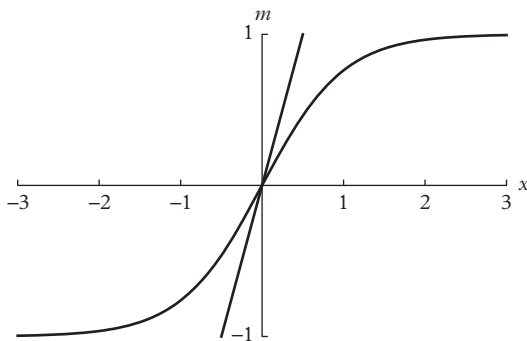


Fig. 31.2 Plot of $m = \tanh x$ and $m = (T/T_c)x$ for $T > T_c$.

This expression for the mean-field value of T_c means that eq. (31.55) can be rewritten as

$$m = \left(\frac{T}{T_c} \right) x. \quad (31.57)$$

If $T > T_c$, the slope of the straight line described by eq. (31.57) is greater than 1, and the two curves in Fig. 31.2 will only intersect at $m = 0$ and $x = 0$. This gives the reasonable result that the magnetization above T_c is zero.

For temperatures below T_c there are three intersections of the two curves, as shown in Fig. 31.3. The intersection at $m = 0$ is not stable, which will be left to the reader to show. The other two intersections provide symmetric solutions $m = \pm m(T)$, which are both stable.

The two solutions found in Fig. 31.3 are expected by the symmetry of the Ising Hamiltonian under the transformation $\sigma_j \rightarrow -\sigma_j$ for all spins.

31.5.3 Graphical Solution of eq. (31.51) for $h \neq 0$

The situation is qualitatively different when the magnetic field, h , is non-zero. The Ising Hamiltonian, eq. (31.4), is no longer invariant under the transformation $\sigma_j \rightarrow -\sigma_j$ for all spins, because the second term changes sign. Eq. (31.51) can again be simplified by writing it in terms of two equations,

$$m = \tanh(x) \quad (31.58)$$

$$x = \beta \mathcal{J} z m + \beta h. \quad (31.59)$$

The second equation can then be rewritten as

$$m = \left(\frac{k_B T}{z \mathcal{J}} \right) x - \frac{h}{z \mathcal{J}}. \quad (31.60)$$

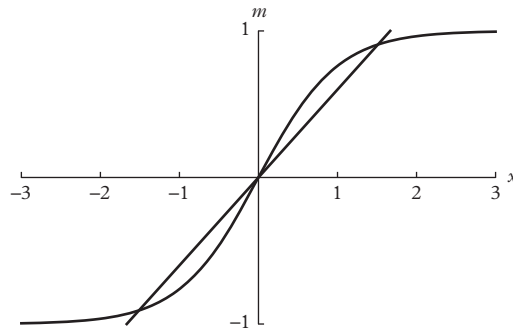


Fig. 31.3 Plot of $m = \tanh x$ and $m = (T_c/T)x$ for $T < T_c$.

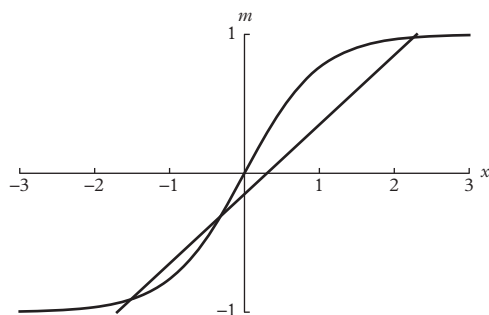


Fig. 31.4 Plot of $m = x$ and $m = (T_c/T)x - h/k_B T$ for $T < T_c$.

Fig. 31.4 shows a plot of eqs. (31.58) and (31.59) on the same graph. In analogy to the previous section, the intersections of the two functions represent solutions to eq. (31.11).

In Fig. 31.4, $T < T_c$ and $h > 0$. Of the three points of intersection that mark solutions to eq. (31.11), only the one at the far right, with $m > 0$, is stable. The intersection to the far left is metastable, and the intersection nearest the origin is unstable. (Proofs will be left to the reader.)

It can be seen by experimenting with various parameters that there is always a stable solution with $m > 0$ when $h > 0$. Similarly, when $h < 0$ there is always a stable solution with $m < 0$.

As long as $h \neq 0$, the MFA prediction for the magnetization is an analytic function of the temperature, and there is no phase transition. This property is also true for the exact solution; a phase transition appears only in the absence of an applied magnetic field.

31.6 Critical Exponents

A curious aspect of phase transitions is that many properties turn out to have some sort of power-law behavior. For example, the magnetization near the critical temperature can be approximated by the expression

$$m \approx A|T - T_c|^\beta, \quad (31.61)$$

where A is a constant.

There is a serious source of potential confusion in eq. (31.61). The use of the symbol β as a dimensionless critical exponent is in direct conflict with the use of the same symbol as $\beta = 1/k_B T$. Unfortunately, both uses of β are so firmly rooted in tradition that an alternative notation is impossible. I can only council patience and a heightened degree of alertness when you encounter the symbol β anywhere in statistical mechanics or thermodynamics.

There is a full Greek alphabet soup of critical exponents defined in a similar manner. For example, for the specific heat,

$$c \propto |T - T_c|^{-\alpha}, \quad (31.62)$$

and for the magnetic susceptibility χ ,

$$\chi = \left(\frac{\partial m}{\partial h} \right)_{h=0} \propto |T - T_c|^{-\gamma}, \quad (31.63)$$

$$\chi = \left(\frac{\partial m}{\partial h} \right)_{T=T_c} \propto |h|^{1/\delta}. \quad (31.64)$$

Correlations between spins can be described by a correlation function, which gives rise to additional critical exponents. The correlation function has the following asymptotic form for large separation $r \equiv |\vec{r}_j - \vec{r}_k|$ between spins,

$$f(r) \equiv \langle \sigma_j \sigma_k \rangle - \langle \sigma_j \rangle \langle \sigma_k \rangle \propto r^{-d-2+\eta} \exp[-r/\xi]. \quad (31.65)$$

An exponent η describes the power-law fall-off of the correlation function. The correlation length ξ , in the exponent in eq. (31.65), is the characteristic size of fluctuations near T_c . ξ diverges with an exponent ν ,

$$\xi \propto |T - T_c|^{-\nu}. \quad (31.66)$$

All of these critical exponents can be evaluated in MFA, as illustrated in the following subsection.

31.7 Mean-Field Exponents

We can derive the value of β *within the mean-field approximation* by making use of the expansion of the hyperbolic tangent in eq. (31.67)

$$\tanh x = x - \frac{1}{3}x^3 + \dots. \quad (31.67)$$

Near the critical point, the magnetization m is small, so that the argument of the hyperbolic tangent in eq. (31.52) is also small. We can therefore use eq. (31.67) to approximate eq. (31.52)

$$m = \tanh(\beta \mathcal{J} z m) \approx \beta \mathcal{J} z m - \frac{1}{3}(\beta \mathcal{J} z m)^3 + \dots. \quad (31.68)$$

One solution of eq. (31.68) is obviously $m = 0$, which is the only solution above T_c . Below T_c , the non-zero solutions for the magnetization are stable, and to find them we can divide eq. (31.68) by $m \neq 0$. Solving the resulting equation for the magnetization, we find

$$m^2 = 3 \left(\frac{T}{T_c} \right)^2 \left(1 - \frac{T}{T_c} \right) + \dots \quad (31.69)$$

Close to T_c , the magnetization behaves as

$$m = \sqrt{3} \left(\frac{T}{T_c} \right) \left(1 - \frac{T}{T_c} \right)^{1/2} + \dots, \quad (31.70)$$

so that we find

$$\beta_{MFA} = \frac{1}{2}. \quad (31.71)$$

The calculation of each of the critical exponents can be carried out within the mean-field approximation. However, it is only fair to let the reader have some of the fun, so we will leave the evaluation of the other critical exponents as exercises.

31.8 Analogy with the van der Waals Approximation

In Chapter 17 we discussed the van der Waals fluid and showed that it predicted a transition between liquid and gas phases. Although we did not emphasize it there, the van der Waals fluid is also an example of a mean-field theory. The interactions between particles are approximated by an average (mean) attraction strength (represented by the parameter a) and an average repulsion length (represented by the parameter b).

In the van der Waals model, the difference in density between the liquid and gas phases was found to be given by the expression

$$\rho_{liq} - \rho_{gas} \propto |T - T_c|^{\beta_{vdW}} \quad (31.72)$$

close to the critical point.

Eq. (31.72) for the discontinuity in density in the vdW fluid is directly analogous to eq. (31.61) for the discontinuity in magnetization in the Ising model from $+|m|$ to $-|m|$ as the applied magnetic field changes sign. Indeed, the analogy between these two models can be extended to all critical properties. Just as the density discontinuity is analogous to the magnetization, the compressibility is analogous to the magnetic susceptibility.

The analogy between the vdW fluid and the MFA Ising model even extends to both models having the same value of the critical exponent with a value of $\beta_{vdW} = 1/2 = \beta_{MFA}$. Indeed, all critical exponents are the same in both models, which will be discussed in the next section.

31.9 Landau Theory

The great Russian physicist, Lev Davidovich Landau (1908–1968) made a remarkable generalization of the mean-field approximation. His theory included not only the van der Waals and Ising models that we have discussed, but any extension of them with more distant or more complicated interactions.

Landau first introduced the concept of an ‘order parameter’; that is, an observable quantity that describes a particular phase transition. For the Ising model, the order parameter is the magnetization, which measures the degree of ordering of the spins. For a fluid, the order parameter is given by the discontinuity of the density. In both cases, the order parameter vanishes above the critical temperature, T_c .

To indicate the generality of the theory, we will denote the order parameter by ψ . Landau postulated that the free energy F should be an analytic function of ψ , as indeed it is for both the vdW fluid and the MFA model of ferromagnetism. Since we are interested in the behavior near the critical temperature, we expect ψ to be small, so Landau expanded the free energy in powers of ψ

$$F = h\psi + r\psi^2 + u\psi^4 + \dots \quad (31.73)$$

In eq. (31.73), the first term indicates the interaction with an external magnetic field, which is why h multiplies the lowest odd power of ψ . There will, of course, also be higher-order odd terms, but we will ignore them in this introduction to the theory. For a magnetic model, all odd terms vanish when $h = 0$ because of the symmetry of the Hamiltonian.

The thermodynamic equilibrium state will be given by the value of the order parameter that corresponds to the minimum of the free energy. This leads us to assume that $u > 0$, because otherwise the minimum free energy would be found at $\psi = \pm\infty$. With this assumption we can find the minimum by the usual procedure of setting the first derivative equal to zero. Ignoring higher-order terms, we find

$$\frac{\partial F}{\partial \psi} = h + 2r\psi + 4u\psi^3 = 0. \quad (31.74)$$

First look at the solutions of eq. (31.74) for $h = 0$. There are two solutions. The simplest is $\psi = 0$, which is what we expect for $T > T_c$. The other solution is

$$\psi = \pm \sqrt{\frac{-2r}{4u}}. \quad (31.75)$$

Clearly, there is only a real solution to eq. (31.74) for $r < 0$. This led Landau to assume that r was related to the temperature, and to lowest order,

$$r = r_o(T - T_c). \quad (31.76)$$

Eq. (31.75) now becomes

$$\psi = \pm \sqrt{\frac{2r_o(T_c - T)}{3u}}. \quad (31.77)$$

Since the definition of the critical exponent β in Landau theory is that for very small ψ ,

$$\psi \propto (T_c - T)^\beta, \quad (31.78)$$

we see that quite generally, $\beta = 1/2$.

All other critical exponents can also be calculated within Landau theory, and their values are identical to those obtain from MFA or the vdW fluid. The derivations will be left to the reader.

The calculation of the critical exponents is remarkable for two reasons. The first is that the same set of values are obtained from Landau theory for *all* models. This result is known as ‘universality’. It is a very powerful statement. It says that the particular interactions in any model of interest are completely unimportant in determining the power laws governing critical behavior. The range, strength, or functional form of the interactions make no difference at all.

With some important limitations, universality is even true!

The second reason for the importance of Landau theory is that the particular values he found for the critical exponents are wrong! They disagree with both experiment and some exact solutions of special models.

The discovery that the Landau theory is wrong was a disaster. A great strength of Landau theory is its generality; tweaking parameters has no effect on the values of the critical exponents. Unfortunately, that means that if those values are wrong, there is no easy way to modify the theory to obtain the right answer.

31.10 Beyond Landau Theory

The resolution of the crisis created by the disagreement of Landau theory with experiment was not achieved until 1971, when Kenneth G. Wilson (American physicist, 1936–2013) applied renormalization-group theory to the problem. The key to the solution turned out to be that at a critical point, fluctuations at all wavelengths interact with each other; approximating the behavior of the system without including all interactions between different length scales always leads to the (incorrect) Landau values of the critical exponents.

The story of how physicists came to understand phase transitions is fascinating, but unfortunately would take us beyond the scope of this book. On the other hand, the intention of this book is only to be a starting point; it is by no means a complete survey.

We come to the end of our introduction to thermal physics, which provides the foundations for the broad field of condensed matter physics. It is my hope that you are

now well prepared for further study of this field, which is both intellectually challenging and essential to modern technology.

31.11 Problems

PROBLEM 31.1

An Ising chain

We have solved the one-dimensional Ising chain *with free boundary conditions* for either $\mathcal{J} = 0$ or $h = 0$. We also solved the one-dimensional Ising chain *with periodic boundary conditions* with both $\mathcal{J} \neq 0$ and $h \neq 0$.

For this assignment we will return to the one-dimensional Ising chain with $\mathcal{J} \neq 0$, but $h = 0$.

1. Use the same matrix methods we used to solve the general one-dimensional chain to find the free energy with $\mathcal{J} \neq 0$, but $h = 0$. (Do not simply copy the solution and set h equal to zero! Repeat the derivation.)
2. Find the free energy per site F_N/N in the limit of an infinite system for both free boundary conditions and periodic boundary conditions. Compare the results.
3. The following problem is rather challenging, but it is worth the effort.

A curious problem related to these calculations caused some controversy in the late 1990s. Consider the following facts.

- (a) Free boundary conditions.

The lowest energy level above the ground state corresponds to having all spins equal to $+1$ to the left of some point and all spins equal to -1 to the right. Or vice versa. The energy of these states is clearly $2\mathcal{J}$, so that the thermal probability of the lowest state should be proportional to $\exp(-2\beta\mathcal{J})$, and the lowest-order term in an expansion of the free energy should be proportional to this factor.

- (b) Periodic boundary conditions.

The ground state still has all spins equal to 1 (or -1), but the deviation from perfect ordering must now occur in *two* places. One of them will look like

$$\cdots + + + + - - - \cdots$$

and the other will look like

$$\cdots - - - - + + + \cdots$$

The energy of these states is clearly $4\mathcal{J}$, so the thermal probability of the lowest state will be $\exp(-4\beta\mathcal{J})$ for all N . In this case, the lowest-order term in an expansion of the free energy should be proportional to $\exp(-4\beta\mathcal{J})$.

In the limit of an infinite system, this difference is not seen in the free energies per site. Can you find the source of this apparent contradiction?

PROBLEM 31.2

Mean-Field Approximation (MFA) for the Ising model

We derived the basic MFA equation for the magnetization in class,

$$m = \langle \sigma_j \rangle = \tanh(\beta z \mathfrak{J} m + \beta h)$$

z is the number of nearest neighbors. For a two-dimensional square lattice, $z = 4$.

We found the critical temperature to be given by

$$k_B T_c = z \mathfrak{J}.$$

1. Consider the case of zero magnetic field ($h = 0$). For temperatures below the critical temperature, the magnetization is non-zero. As $T \rightarrow T_c$, the magnetization goes to zero as a power law

$$m \propto (T_c - T)^\beta.$$

Find the value of β .

2. For temperatures above the critical temperature, find an exact solution for the magnetic susceptibility

$$\chi \equiv \frac{\partial m}{\partial h}$$

at zero field.

The magnetic susceptibility diverges at the critical temperature with an exponent called γ .

$$\chi \propto (T - T_c)^{-\gamma}$$

What is the mean-field value of γ ?

Suggestions for writing programs:

The next few problems deal with Monte Carlo simulations of the two-dimensional Ising model. In writing the programs, use a sequential selection of spins for the MC steps. Even though it is not exactly correct, the differences are very small and a sequential selection is much faster. You can implement the algorithm any way you like, but I recommend using a loop of the form:

```
while n < N_MC:
    nxm = L - 2
    nx  = L - 1
    nxp = 0
```

```

while nxp < L:
    nym = L - 2
    ny  = L - 1
    nyp = 0
    while nyp < L:
        STATEMENTS INSERTED HERE FOR MC UPDATE
        nym = ny
        ny  = nyp
        nyp += 1
    nxm = nx
    nx  = nxp
    nxp += 1
STATEMENTS TO RECORD MC DATA INSERTED HERE
n += 1

```

PROBLEM 31.3

Monte Carlo simulation of a $d = 2$ Ising model

Write a computer program to perform a Monte Carlo computer simulation of the two-dimensional Ising model in zero magnetic field,

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

where the sum is over nearest-neighbor pairs of sites on a square lattice.

1. Write a program to perform a Monte Carlo simulation of the two-dimensional Ising model. You should be able to simulate any size $L \times L$ lattice, with L determined at the beginning of the program. You should be able to set the temperature T at the beginning of the program. You can also write the program for zero magnetic field. (We will use non-zero fields in later problems.)
2. Beginning with a random configuration on a fairly large lattice ($L = 32$), perform a single run of 20 MC sweeps at $T = 5.0$. and look at the printout of the configuration.
3. Beginning with a random configuration on a fairly large lattice ($L = 32$), perform a single run of 20 MC sweeps at $T = 3.0$. and look at the print-out of the configuration.
4. Beginning with a random configuration on a fairly large lattice ($L = 32$), perform a single run of 20 MC sweeps at $T = 2.0$, and look at the print-out of the configuration.
5. Compare the pictures of the configurations that you have generated at $T = 5.0, 2.0, 2.0$. What do you see?
6. Beginning with a random configuration on a fairly large lattice ($L = 16$ or $L = 32$), perform several runs with about 50 or 100 MC sweeps at $T = 0.5$. Keep the length of the runs short enough so that you can do it several times without wasting your time. Look at the print-outs of several configurations. If you see something that seems unusual, print it out. Explain why you think it is unusual.

PROBLEM 31.4

Monte Carlo simulation of a $d = 2$ Ising model (revisited)

For this problem we will modify the MC program to make a series of runs with temperatures separated by a value ΔT .

1. Use the Python function definition feature to create a function that does the entire MC simulation. Then put the function call inside a loop that first equilibrates the spins for some number of sweeps and then generates the data you want from a longer number of sweeps. The loop should change the temperature by an amount ΔT for every iteration of equilibration and data taking.
2. Calculate the average energy, specific heat, magnetization, and magnetic susceptibility for temperatures of $T = 2.0, 4.0$, and 6.0 for an 8×8 lattice. Use a large enough numbers of MC updates to obtain reasonable results, but not so many that you waste your time sitting in front of a blank computer screen.
3. What do you notice about the data from your MC simulation?

This is a fairly open-ended question. The main purpose is to think about the results of the simulation.

PROBLEM 31.5

More Monte Carlo simulations of a two-dimensional Ising model

The Hamiltonian of a two-dimensional Ising model is: use periodic boundary conditions,

$$\mathcal{H} = -\mathfrak{J} \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_j \sigma_j.$$

Recall that the notation $\langle i,j \rangle$ denotes the nearest-neighbor pair i and j . For this entire assignment, take $h = 0$.

Use or modify the program you wrote for an earlier problem to do the following simulations.

1. Simulate a 16×16 lattice with $\mathfrak{J} = 1.0$. Begin with giving each spin a random value of $+1$ or -1 . Scan the temperature from 5.0 to 0.5 at intervals of $\Delta T = -0.5$. Use about 100 or 1000 MC sweeps for each temperature, depending on the speed of your computer.

Make a table of the results and plots of m vs. T , E vs. T , c vs. T , and χ vs. T .

Does it look as if the critical temperature is at the mean-field value of $k_B T_c = 4\mathfrak{J}$?

2. Repeat the sweeps, but this time choose the temperature range to concentrate on the peaks in the specific heat and the magnetic susceptibility. Estimate the location and height of the peak in specific heat and magnetic susceptibility. These will be estimates of the critical temperature.
3. And now for something completely different.

Do a simulation of a 32×32 lattice with *negative* temperature, and print out the spins at the end of the run. Begin with the spins in a random state ($T = \infty$), and simulate for 100 sweeps at $T = -1.0$.

What do you see?

PROBLEM 31.6

A one-dimensional, spin-one model of magnetism

Consider a one-dimensional macroscopic model with spin-one quantum mechanical magnetic moments located on each of N sites. The Hamiltonian is

$$H = -\mathcal{J} \sum_{j=1}^{N-1} \sigma_j \sigma_{j+1} - h \sum_{j=1}^N \sigma_j - D \sum_{j=1}^N \sigma_j^2$$

where each σ_j takes on the values $-1, 0$, or $+1$, \mathcal{J} is an exchange interaction parameter, h is a magnetic field, and D is a parameter representing a ‘crystal field’. The entire system is in contact with a thermal reservoir at temperature T .

1. Write down a formal expression for the partition function of this system. You do not have to evaluate any sums or integrals that appear in your expression for this part of the problem.
2. For general values of the parameters, derive a formal expression for the quantity

$$Q = \left\langle \sum_{n=1}^N \sigma_n^2 \right\rangle$$

in terms of a derivative of the partition function with respect to an appropriate parameter.

3. For $\mathcal{J} = 0$, but arbitrary values of h and D , calculate the partition function for this system.
4. For $\mathcal{J} = 0$, but arbitrary values of h and D , calculate the free energy of this system.
5. For $\mathcal{J} = 0$ and $h = 0$, calculate the entropy of this system.

.....

