

# An Introduction to the Ising Model

BARRY A. CIPRA, *St. Olaf College, Northfield, Minnesota 55057*

*Barry Cipra* received a Ph.D. in mathematics from the University of Maryland in 1980. He has been an instructor at M.I.T. and at the Ohio State University, and is currently an assistant professor of mathematics at St. Olaf College in Northfield, Minnesota. He is the author of *Mistakes and How to Find Them Before the Teacher Does*, published by Birkhauser Boston.



## Introduction

This article is an invitation, or advertisement, for readers to work on a problem which is apparently very difficult, yet certainly extremely important. The problem is known generically as the *Ising model*, named after Ernst Ising, who did the first work on it in the early 1920s. Although unpromising in its initial results, the Ising model has turned out to be an exceptionally rich idea. The number of papers written on the subject is staggering; the number which remain to be written is conceivably even more staggering.

The Ising model is concerned with the physics of phase transitions, which occur when a small change in a parameter such as temperature or pressure causes a large-scale, *qualitative* change in the state of a system. Phase transitions are common in physics and familiar in everyday life: we see one, for instance, whenever the temperature drops below 32° F, and another whenever we put a kettle of water on the stove. Other examples include the formation of binary alloys and the phenomenon of ferromagnetism. The latter is also of interest historically: an understanding of ferromagnetism—and especially “spontaneous magnetization”—was the original purpose of the Ising model and the subject of Ising’s doctoral dissertation. Partly for this historical significance, we shall use ferromagnetism as a reference point later on for interpreting various features of the model.

In spite of their familiarity, phase transitions are not well understood. One purpose of the Ising model is to explain how short-range interactions between, say, molecules in a crystal give rise to long-range, correlative behavior, and to predict in some sense the potential for a phase transition. The Ising model has also been applied to problems in chemistry, molecular biology, and other areas where “cooperative” behavior of large systems is studied. These applications are possible because the Ising model can be formulated as a *mathematical* problem. Although we shall refer frequently to the physics of ferromagnetism and use language from statistical mechanics, it is the mathematical aspects of the model which will concern us in this article. In particular we shall see that the Ising model has a combinatorial interpretation which is powerful enough in itself to establish some of the basic results concerning phase transitions. There are many other approaches and aspects to the

Ising model, but the combinatorial one makes an especially suitable introduction to the subject.

**1. Lattices and the Partition Function**

Our starting point for the Ising model is a *lattice*, which for us will be a finite set of regularly spaced points in a space of dimension  $d = 1, 2,$  or  $3$ . In dimension 1 we simply have a string of points on a line, which we can enumerate from 1 to  $N$  (" $N$ " will always denote the number of lattice sites, regardless of dimension):

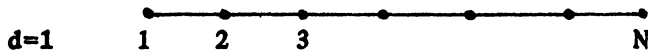


FIG. 1.

In dimension 2 we shall consider the lattice of squares as below:

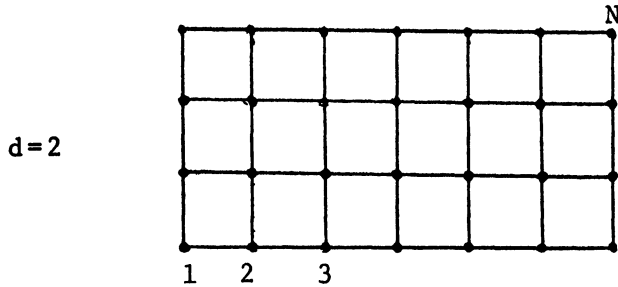


FIG. 2.

In dimension 3 we shall consider the lattice whose repeating units are cubes:

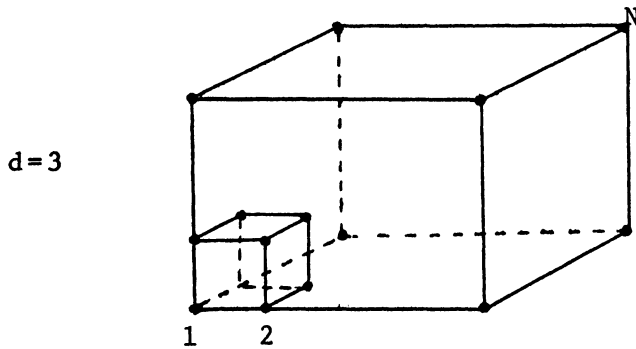


FIG. 3.

In our pictures, each line segment between lattice sites is called a *bond*, and lattice sites are called *nearest neighbors* if there is a bond connecting them. In general, except for lattice sites on the “boundary” of the lattice, each lattice site in a  $d$ -dimensional lattice has  $2d$  nearest neighbors:

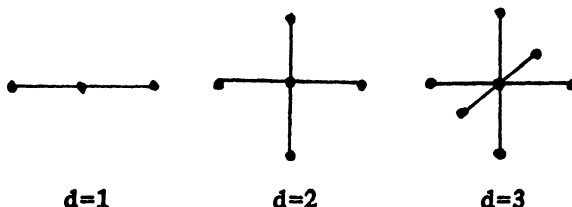


FIG. 4.

The difference between lattice sites on the “boundary” and those in the “interior” of the lattice is mildly annoying. One way to deal with this annoyance is to get rid of the boundary by adopting what might be called a “wrap-around” model: we simply introduce extra bonds connecting lattice sites on opposite sides of the boundary. This amounts to wrapping the one-dimensional lattice into a necklace, the two-dimensional lattice into a doughnut, and the three-dimensional lattice into who knows what.

Although this introduces physically unrealistic long-range interactions (or else requires us to bend a three-dimensional crystal in an impossible manner), physicists will be the first to go along with the idea: intuitively, an extra condition imposed on a “negligibly small” percentage of lattice sites should not affect the overall behavior of the system. (There is an important alternative which we shall mention again later: one can impose “boundary conditions” which *do* influence the behavior of the system, by establishing a preferred direction for the spontaneous magnetization.) Eliminating the boundary also introduces an appealing symmetry into the problem: in the wrap-around Ising model, there are  $dN$  bonds connecting the  $N$  lattice sites, and every lattice site “looks like” every other lattice site. Therefore we shall henceforth assume that the lattice has been wrapped around.

Our first step is to assign an independent variable  $\sigma_i$  to each lattice site  $i = 1, \dots, N$ . The variables  $\sigma_i$  take on only two values,  $\sigma_i = \pm 1$ , which we shall call the two possible *states* of the lattice site. This reflects the physical assumption that only two possibilities exist at each lattice site, such as up/down or occupied/vacant, as we shall explain below. An assignment of values  $(\sigma_1, \sigma_2, \dots, \sigma_N)$  to each lattice site is called a *configuration* of the system. An essential ingredient in the Ising model will be a *sum* over all possible configurations. Since there are  $2^N$  configurations, this sum clearly has an enormous number of terms if  $N$  is at all large. For a macroscopic crystal, with  $N \sim 10^{23}$ , one should not even contemplate carrying out such a calculation numerically!

In the model of ferromagnetism—Ising’s original study—one thinks of the lattice sites as being occupied by atoms of a magnetic material. Each atom has a magnetic moment which is allowed to point either “up” or “down.” In a model for binary alloys, the lattice sites are again occupied by atoms, which may be one or the other of the two constituents of the alloy. A third interpretation has the paradoxical name “lattice gas”: the lattice sites are points in space which are either occupied or vacant. (The sought-for phase transition here is between a “solid,” which has segregated regions of occupied and vacant space, and a “gas” for which the lattice is a homogeneous mixture of the two.) In all cases, the variable  $\sigma_i$  is used to designate which state the  $i$ th lattice site is in. Of course one of the many generalizations of the model is to increase the number of states, say to 1, 0, and  $-1$ , or to a continuum of states.

We next form what is called the *Hamiltonian* of the system. In mathematical physics, the Hamiltonian is the total energy of a system, and it governs the dynamics. For the Ising model, the Hamiltonian is defined after an ideal and apparently very severe assumption is made: we assume that only short-range, “nearest-neighbor” interactions and interactions of the lattice sites with an “external field” contribute to the energy level of the system. For each configuration  $\sigma = (\sigma_1, \dots, \sigma_N)$  we have

$$H = H(\sigma) = - \sum_{\langle i, j \rangle} E \sigma_i \sigma_j - \sum_i J \sigma_i, \quad (1.1)$$

where  $E$  and  $J$  are parameters, the second sum is over all lattice sites, and the first sum is over all pairs of nearest neighbors in the lattice. The parameters  $E$  and  $J$  correspond to the “energies” associated with nearest-neighbor interactions and interactions with the external field, respectively. For a ferromagnet,  $E$  is positive, so that a “magnetized” configuration (with most nearest-neighbor pairs having parallel moments,  $\sigma_i = \sigma_j$ ) has a lower energy level than a non-magnetized configuration. The parameter  $J$  corresponds to the presence of an “external magnetic field”, which will tend to line up the magnetic moments in the direction of the field, again “favoring” configurations with lower energy levels. Fighting against this, as we shall see below, is thermal agitation. At sufficiently low temperatures, there is not much random motion, and configurations lined up with an external field are highly favored, while at sufficiently high temperatures, the random thermal motion destroys much of the effect of the field.

Partly for its historic interest, let us now explain the nature of the ferromagnetic phase transition which Ising originally sought in his dissertation. The phase transition occurs with the appearance of what is called *spontaneous magnetization*.

Suppose a lattice of magnetic material is placed in a magnetic field and held at a constant temperature. The field will induce a certain amount of magnetization into the lattice—i.e., it will create a tendency for the magnetic moments to point in, say, the “up” direction. The amount of magnetization depends on the strength of the external field and on the (constant) temperature.

Now suppose the external field is slowly turned off. What happens to the lattice? Not surprisingly, for high temperatures, the lattice returns to an unmagnetized condition. But for low temperatures, the lattice retains a degree of magnetism; there is a non-negligible residual tendency for the moments to stay in the “up” position. This is called *spontaneous magnetization*. (Note: it seems that “residual magnetization” would have been a better term, but so be it.) There is a *critical temperature* at which spontaneous magnetization begins to appear, and this is where the phase transition occurs. The figure below shows an (idealized) graph of induced magnetization versus external field strength for three temperatures, including the critical temperature. The curve for the critical temperature is characterized by its having a vertical tangent line at the origin.

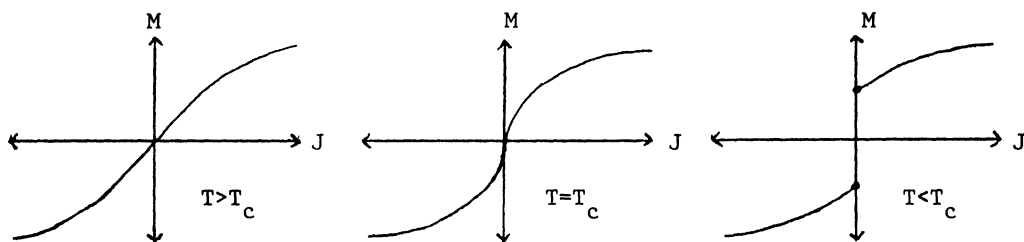


FIG. 5.

As we shall show later, the one-dimensional Ising model does not exhibit a phase transition at any temperature. This negative result, plus some arguments that the same thing would happen in three dimensions, discouraged Ising from pursuing the subject. The Ising model lay dormant for about a decade, until Rudolf Peierls [37] in 1936 showed by a very simple argument that, in two dimensions, a phase transition was *guaranteed for some temperature*. In 1941, Hendrick Kramers and Gregory Wannier [26] located the phase transition precisely for the two-dimensional model, under the assumption that there is a unique such value. In 1944, Lars Onsager [36] gave a complete solution to the two-dimensional Ising model in the “zero-field” ( $J = 0$ ) case. To date, no one has solved any three-dimensional model.

Returning to the Ising model, our third step brings us to the central object in statistical mechanics: the partition function. This is formed by exponentiating the Hamiltonian and then summing over all configurations, which here involves  $2^N$  possible assignments of  $\pm 1$  to the  $N$  variables  $\sigma_1, \dots, \sigma_N$ :

$$Z = Z(\beta, E, J, N) = \sum_{\pm 1} e^{-\beta H(\sigma)}. \quad (1.2)$$

The parameter  $\beta$  cancels whatever dimensions the Hamiltonian may have. In statistical mechanics, we typically have  $\beta = 1/kT$ , where  $k$  is Boltzmann’s constant and  $T$  is temperature (in absolute degrees).

A simple example may clarify some of the notation. Let’s take a very small one-dimensional lattice, consisting of  $N = 3$  lattice sites with no wrap-around:

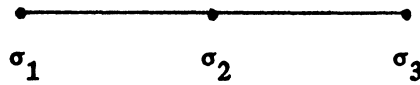


FIG. 6.

The Hamiltonian is

$$H = -E(\sigma_1\sigma_2 + \sigma_2\sigma_3) - J(\sigma_1 + \sigma_2 + \sigma_3).$$

To simplify matters further, we shall set  $J = 0$  (the “zero-field” case). The partition function is now

$$\begin{aligned} Z &= e^{-\beta H(1,1,1)} + e^{-\beta H(1,1,-1)} + e^{-\beta H(1,-1,1)} + e^{-\beta H(1,-1,-1)} \\ &\quad + e^{-\beta H(-1,1,1)} + e^{-\beta H(-1,1,-1)} + e^{-\beta H(-1,-1,1)} + e^{-\beta H(-1,-1,-1)} \\ &= e^{\beta E(1+1)} + e^{\beta E(1-1)} + e^{\beta E(-1-1)} + e^{\beta E(-1+1)} \\ &\quad + e^{\beta E(-1+1)} + e^{\beta E(-1-1)} + e^{\beta E(1-1)} + e^{\beta E(1+1)} \\ &= 2e^{2\beta E} + 4 + 2e^{-2\beta E} \\ &= 2^3 \cosh^2 \beta E. \end{aligned}$$

(The final formula in the example is suggestive of what is to come. The reader may want to pause at this point and work out the partition function for the one-dimensional, zero-field model with  $N$  lattice sites.)

The partition function plays a fundamental role in statistical mechanics. Essentially, it is the “denominator” in the calculation of probabilities. More precisely, the probability of being in a particular configuration  $\sigma = (\sigma_1, \dots, \sigma_N)$  is given by the formula

$$\text{Prob}(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z}. \tag{1.3}$$

The negative sign confers a higher probability on states with lower energy. A small value of  $\beta$  (corresponding to a high temperature, since  $\beta = 1/kT$ ) tends to “flatten out” the distribution, making all configurations more or less equally likely, while a large value of  $\beta$  (corresponding to a low temperature) tends to accentuate the probabilities of the lowest energy states.

From the partition function, one may in principle derive all of the important thermodynamical features of the physical system being modeled: internal energy, specific heat, magnetization and magnetic susceptibility, and so forth. For example, the internal energy is defined as

$$U = \frac{1}{Z} \sum_{\pm 1} H(\sigma) e^{-\beta H(\sigma)}$$

and we easily see that this can be re-expressed as

$$U = -\frac{\partial}{\partial \beta} \log Z.$$

(For more on applications of the partition function, see [42] or [44].)

Many of the quantities one computes from the partition function turn out to depend on the logarithm of  $Z$ . This is natural, since  $Z$ , being a sum over  $2^N$  configurations, tends to grow exponentially with the size of the lattice. This brings us to our last step in setting up the Ising model; we define the “free energy per lattice site” to be

$$F = F(\beta, E, J) = \lim_{N \rightarrow \infty} \frac{1}{N} \log Z(\beta, E, J, N). \quad (1.4)$$

The limit as  $N \rightarrow \infty$  is called the “thermodynamic limit.” The main problem of the Ising model is this: *Find a closed-form, analytic expression for the function  $F$ .* The idea is that phase transitions will show up as discontinuities in  $F$  or in one of its derivatives: a phase transition occurs when some aspect of the system changes radically at certain values of the parameters.

There is, of course, no *a priori* guarantee that the thermodynamic limit  $F$  exists. There is also some question as to how the limit is meant to be taken in two or three dimensions, since the lattice can grow at different rates in different directions. We shall not consider these questions any further, but merely assume that the appropriate limits do exist.

For the rest of this article, we shall introduce some elementary steps for analyzing the Ising model and describe what is known about the exact solutions. We shall also present the arguments due to Peierls and to Kramers and Wannier for the existence of phase transitions in two dimensions. The results of these arguments were superseded by Onsager’s complete solution (which we do not exposit here), but the techniques and ideas continue to be important. Peierls’ argument, in particular, generalizes fairly easily to higher dimensions, where very little else is rigorously known.

## 2. Elementary Analysis—Some Combinatorics

We shall begin by converting the partition function from transcendental exponentials into a polynomial in two variables with integer coefficients. This is based on the simple observation

$$e^{\pm x} = \cosh x \pm \sinh x = \cosh x(1 \pm \tanh x). \quad (2.1)$$

Since the variables  $\sigma_i$  take on the values  $\pm 1$ , we have

$$\begin{aligned} Z &= \sum_{\pm 1} e^{\sum_{\langle i, j \rangle} \beta E \sigma_i \sigma_j + \sum_i \beta J \sigma_i} = \sum_{\pm 1} \left( \prod_{\langle i, j \rangle} e^{\beta E \sigma_i \sigma_j} \right) \left( \prod_i e^{\beta J \sigma_i} \right) \\ &= \sum_{\pm 1} \left( \prod_{\langle i, j \rangle} \cosh(\beta E)(1 + \sigma_i \sigma_j T) \right) \left( \prod_i \cosh(\beta J)(1 + \sigma_i U) \right) \\ &= (\cosh(\beta E))^B (\cosh(\beta J))^N \sum_{\pm 1} \left( \prod_{\langle i, j \rangle} (1 + \sigma_i \sigma_j T) \right) \left( \prod_i (1 + \sigma_i U) \right), \end{aligned} \quad (2.2)$$

where  $B$  is the number of bonds,  $T = \tanh(\beta E)$  and  $U = \tanh(\beta J)$ . If we use the “wrap-around” lattice, then  $B = dN$  where  $d = 1, 2$ , or  $3$  is the dimension of the

model. It is also convenient to make the sum into an *average* over all configurations by introducing a factor  $2^N$ :

$$Z = (2 \cosh^d(\beta E) \cosh(\beta J))^N \frac{1}{2^N} \sum_{\pm 1} \left( \prod_{\langle i, j \rangle} (1 + \sigma_i \sigma_j T) \right) \left( \prod_i (1 + \sigma_i U) \right). \quad (2.3)$$

The thermodynamic limit is now viewed as consisting of two pieces:

$$F = \lim_{N \rightarrow \infty} \frac{1}{N} \log Z = \log(2 \cosh^d(\beta E) \cosh(\beta J)) + \lim_{N \rightarrow \infty} \frac{1}{N} \log Z', \quad (2.4)$$

where

$$Z' = \frac{1}{2^N} \sum_{\pm 1} \left( \prod_{\langle i, j \rangle} (1 + \sigma_i \sigma_j T) \right) \left( \prod_i (1 + \sigma_i U) \right). \quad (2.5)$$

The first piece,  $\log[2 \cosh^d(\beta E) \cosh(\beta J)]$ , is always analytic for real (i.e., physical) values of  $\beta$ ,  $E$ , and  $J$ ; hence it is a “trivial” contribution exhibiting no discontinuities. We are left with the task of analyzing the “modified” partition function  $Z'$ .

Because  $\sigma_i^2 = 1$  for all  $i$ , we may write

$$\begin{aligned} & \prod_{\langle i, j \rangle} (1 + \sigma_i \sigma_j T) \prod_i (1 + \sigma_i U) \\ &= P(T, U) + \sigma_1 P_1(T, U, \sigma_2, \dots, \sigma_N) \\ & \quad + \sigma_2 P_2(T, U, \sigma_3, \dots, \sigma_N) + \dots + \sigma_N P_N(T, U) \end{aligned} \quad (2.6)$$

for polynomials  $P$  and  $P_1, \dots, P_N$ . Note that  $P$  is of degree  $dN$  in  $T$  and  $N$  in  $U$ , assuming again the wrap-around lattice. When we sum over all configurations, however, each  $\sigma_k P_k$  term vanishes by trivial cancellation:

$$\begin{aligned} \sum_{\pm 1} \sigma_k P_k(T, U, \sigma_{k+1}, \dots, \sigma_N) &= \left( \sum_{\sigma_k = \pm 1} \sigma_k \right) \left( \sum_{\pm 1} P_k(T, U, \sigma_{k+1}, \dots, \sigma_N) \right) \\ &= (0) (\text{whatever}) = 0. \end{aligned}$$

This leaves the modified partition function

$$Z' = \frac{1}{2^N} \sum_{\pm 1} P(T, U) = P(T, U), \quad (2.7)$$

which is a polynomial in two variables with integer coefficients.

At this point, we shall simplify our discussion by setting  $U = 0$ . This is called the “zero magnetic field case.” In this case the coefficients of the polynomial

$$P(T, 0) = 1 + c(1)T + c(2)T^2 + \dots + c(dN)T^{dN}$$

have a simple combinatorial interpretation. If the lattice is thought of as a graph with lattice sites as the vertices and bonds between nearest neighbors as the edges, then  $c(n)$  counts the number of “even” subgraphs with  $n$  edges, where “even” means that each vertex has positive, *even* degree. This can be seen by letting the



presence or absence of the bond  $\langle i, j \rangle$  in a subgraph correspond to the choice of  $\sigma_i \sigma_j T$  or 1 in the expansion of  $\prod_{\langle i, j \rangle} (1 + \sigma_i \sigma_j T)$ . Each subgraph corresponds to a term in the expansion:  $(\prod \sigma_i^{\delta_i}) T^n$ , where  $\delta_i =$  degree of vertex  $i$  and  $n = \frac{1}{2} \sum \delta_i =$  number of edges. Only even subgraphs raise each  $\sigma_i$  to an even power, hence only even subgraphs contribute to the modified partition function  $Z' = P(T)$  in the zero magnetic field case.

Each connected component of an even subgraph is a closed path in the original lattice. This enables us to solve completely the one-dimensional, zero-field Ising model: in the wrap-around model, there is only one closed path, namely, the complete circuit of length  $N$ . Thus  $Z' = 1 + T^N$ , so that

$$\begin{aligned} F &= \log(2 \cosh^d(\beta E)) + \lim_{N \rightarrow \infty} \frac{1}{N} \log(1 + T^N) \\ &= \log(2 \cosh^d(\beta E)), \end{aligned}$$

since  $|T| = |\tanh \beta E| < 1$  implies  $\lim_{N \rightarrow \infty} (1/N) \log(1 + T^N) = 0$ . (Note that in the "non-wrap-around" case, there are *no* closed paths, so that  $\log(Z') = 0$  directly.)

In dimensions 2 and 3, closed paths obviously do exist, but they must be of even length, unless they are long enough to make use of wrap-around. Also, the shortest paths are of length 4, so we have

$$Z' = 1 + c(4)T^4 + c(6)T^6 + c(8)T^8 + \dots$$

if the lattice is sufficiently large. For any given  $n$ , we can also work out explicitly the coefficient  $c(n)$ ; however this is practical only for small values of  $n$ . We shall show this computation (really a counting and bookkeeping argument) for  $n = 4, 6$ , and 8, and leave  $n = 10, 12$ , and any higher degrees for the interested (and industrious) reader.

To distinguish between dimensions, let us write

$$Z'_d = 1 + c_d(4)T^4 + c_d(6)T^6 + c_d(8)T^8 + \dots \quad (2.8)$$

for the modified partition function for the  $d$ -dimensional Ising model ( $d = 1, 2, 3$ ). As we pointed out before,  $c_1(n) = 0$  for all  $n \ll N$ . We shall henceforth consider only dimensions  $d = 2$  and 3.

An even subgraph with  $n = 4$  edges is simply a square. For  $d = 2$ , the square may be located with a specified (say, lower-left-hand) corner at any lattice site (using again the wrap-around model), so that  $c_2(4) = N$ . For  $d = 3$ , we have in addition a choice of orientation, so that  $c_3(4) = 3N$ .

For  $d = 2$ , an even subgraph with  $n = 6$  edges is a  $2 \times 1$  rectangle, which can be located at any of the  $N$  lattice sites and oriented in two possible ways. Hence  $c_2(6) = 2N$ . For  $d = 3$ , in addition to  $6N$  "flat" rectangles, there are another  $12N$  "bent" rectangles and  $4N$  more "twisted" rectangles, for a total  $c_3(6) = 22N$ .

For  $n = 8$ , the situation becomes more complicated. For one thing, the subgraphs no longer need to be connected. A disconnected subgraph, however, must consist of two disjoint squares. For  $d = 2$ , the "first" square may be placed with its lower-

left-hand corner at any of the  $N$  lattice sites, while the same corner of the “second” square need only avoid nine lattice sites (see Figure 7).

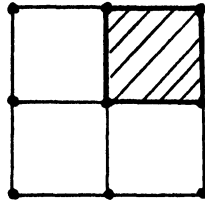


FIG. 7.

Thus for  $d = 2$ , there are  $N(N - 9)/2$  disconnected even subgraphs with 8 edges. (We divide by 2 to eliminate the distinction between the “first” and “second” square.) For  $d = 3$ , a similar argument shows that there are  $3N(3N - 33)/2$  disconnected subgraphs with 8 edges.

The connected paths of length 8 in dimension 2 are easy to count. There are four different types, with a total of 9 orientations, giving, in all,  $9N$  connected subgraphs with 8 edges. Thus,

$$c_2(8) = N(N - 9)/2 + 9N = N(N + 9)/2.$$

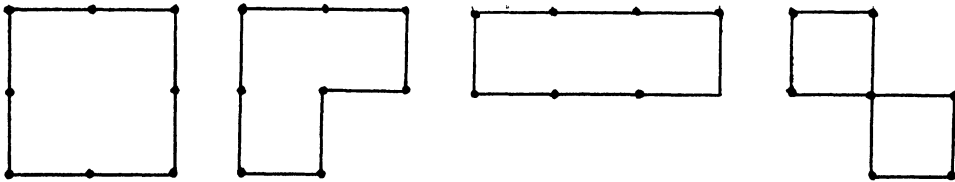


FIG. 8.

The real complication appears for  $d = 3$ : there are suddenly a lot of different paths of length 8. Classifying them in a manner analogous to the paths of length 6, we have  $27N$  ( $= 3 \times 9N$ ) “flat” graphs,  $108N$  graphs with one “bend”,  $48N$  with two bends, and  $48N$  “twisted” graphs, for a total of  $231N$  possibilities. Adding in the disconnected subgraphs, we find

$$c_3(8) = 3N(3N - 33)/2 + 231N = (9N^2 + 363N)/2.$$

The reader is invited to look for simpler means of computing these coefficients. This is as far as we shall pursue the matter.

Knowing the first few terms of the partition function allows us to compute corresponding terms in the power series for the thermodynamic limit. We proceed as follows. Since

$$\log(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots \quad (|x| < 1),$$

we have

$$\begin{aligned} \log(Z') &= \log[1 + c(4)T^4 + c(6)T^6 + c(8)T^8 + c(10)T^{10} + \dots] \\ &= [c(4)T^4 + c(6)T^6 + c(8)T^8 + c(10)T^{10}] \\ &\quad - \frac{1}{2}[c(4)T^4 + c(6)T^6]^2 + O(T^{12}) \\ &= c(4)T^4 + c(6)T^6 + \left[ c(8) - \frac{1}{2}c(4)^2 \right] T^8 \\ &\quad + [c(10) - c(4)c(6)]T^{10} + O(T^{12}). \end{aligned}$$

From our computations above, we find

$$\frac{1}{N} \log(Z'_2) = T^4 + 2T^6 + \frac{9}{2}T^8 + \dots \quad (2.9)$$

and

$$\frac{1}{N} \log(Z'_3) = 3T^4 + 22T^6 + \frac{363}{2}T^8 + \dots \quad (2.10)$$

Note how the lattice size  $N$  has vanished on the right-hand side (at least for the terms we have shown—we expect it to happen for all terms). Taking the limit as  $N \rightarrow \infty$  gives us a power series expansion for the (modified) free energy  $F'$ .

The reader who has had an introductory course in hard analysis should be appalled at what we've just done. In particular, we have not proved the validity of truncating the power series expansion for  $\log(1+x)$  and then letting  $N$  tend to infinity. We've also not proved that the  $N$  has vanished from all terms on the right-hand side. These objections can be dealt with, however, by taking a formal power series point of view.

That leaves the question of the radius of convergence of the power series as an analytic function around  $T=0$ . This is an important question, because it is non-analytic behavior that we look for as the defining characteristic of a phase transition. What we hope will happen is that there will be a phase transition corresponding to some “physical” value of  $T$  in the interval  $(0, 1)$ , and that this will be the closest singularity to the origin. Of course we have no right to think this is what's going to happen. But for  $d=2$  it does.

### 3. Exact Solutions

To review, we have set  $T = \tanh(\beta E)$  and  $U = \tanh(\beta J)$ , and defined

$$Z'_d = Z'_d(T, U, N) = \frac{1}{2^N} \sum_{\pm 1} \left[ \prod_{\langle i, j \rangle} (1 + \sigma_i \sigma_j T) \prod_i (1 + \sigma_i U) \right]$$

for the modified partition function of the  $d$ -dimensional Ising model. Let us also

introduce the “modified free energy” function

$$F'_d(T, U) = \lim_{N \rightarrow \infty} \frac{1}{N} \log Z'_d(T, U, N)$$

(Recall that the original free energy is

$$F_d = \log(2 \cosh^d(\beta E) \cosh(\beta J)) + F'_d(T, U).$$

We have seen that  $Z'_d$  is a polynomial in  $T$  and  $U$  with integer coefficients (of degree  $dN$  in  $T$  and  $N$  in  $U$ , for the wrap-around model). Assuming that the limit exists,  $F'_d$  is, therefore, a power series (at least formally) in  $T$  and  $U$ , with rational coefficients. If we fix  $E$  and  $J$ , we may consider  $T$  and  $U$  as functions of the parameter  $\beta$ . Our objective here is to realize  $F'_d$  as an analytic function of  $\beta$  for small  $\beta$ . (Note that  $T$  and  $U$  are small if  $\beta$  is small.) Remembering that  $\beta$  is inversely proportional to temperature in the physical model, we call the power series in  $T$  and  $U$  a “high-temperature expansion” for  $F'_d$ . Phase transitions occur at the positive real values of  $\beta$  at which  $F'_d$  is nonanalytic.

This objective has been met only “halfway.” The results are given below, organized according to the dimension of the model and the absence or presence of an “external magnetic field”  $U$ .

$$F'_1(T, 0) = 0.$$

$$F'_1(T, U) = \log \left[ \frac{1 + T + [(1 + T)^2 - 4T(1 - U^2)]^{1/2}}{2} \right]. \quad (\text{Ising, 1925})$$

$$F'_2(T, 0) = \frac{1}{2} \int_0^1 \int_0^1 \log[(T^2 + 1)^2 - 2T(1 - T^2)[\cos(2\pi x) + \cos(2\pi y)]] dx dy.$$

(Onsager, 1944)

$$F'_2(T, U) = ?$$

$$F'_3(T, 0) = ??$$

$$F'_3(T, U) = ???$$

In the next section we shall derive Ising’s result for  $F'_1(T, U)$ . A derivation of Onsager’s famous result for  $F'_2(T, 0)$  is beyond the scope of this article. It has been written up in many forms, and we refer the reader to any or all of [12], [33], and [42]. We shall, however, present the beautiful arguments of Peierls [37] and Kramers and Wannier [26], which establish the existence of spontaneous magnetization in two dimensions and the precise location of the phase transition for  $F'_2(T, 0)$  under a mild (and physically reasonable) assumption.

#### 4. Ising’s Result—The Transfer Matrix Method

In this section we shall obtain the complete solution to the one-dimensional Ising model. We begin by looking at the “linear” rather than the “wrap-around” model. (As remarked earlier, it should make no difference in the thermodynamic limit

anyway.) Then

$$Z'(N) = \frac{1}{2^N} \sum_{\pm 1} (1 + \sigma_N U) \prod_{i=1}^{N-1} [(1 + \sigma_i \sigma_{i+1} T)(1 + \sigma_i U)]. \quad (4.1)$$

Suppose we write  $Z'_+(N)$  for that portion of the partition function summation for which  $\sigma_N = +1$ , and  $Z'_-$  for that portion for which  $\sigma_N = -1$ . Clearly  $Z'(N) = Z'_+(N) + Z'_-(N)$ . But also,

$$\begin{aligned} Z'_+(N) &= \frac{1}{2^N} \sum_{\pm 1} (1 + U)(1 + \sigma_{N-1} T)(1 + \sigma_{N-1} U) \\ &\quad \times \prod_{i=1}^{N-2} [(1 + \sigma_i \sigma_{i+1} T)(1 + \sigma_i U)] \\ &= \frac{1}{2} [(1 + U)(1 + T)Z'_+(N-1) + (1 + U)(1 - T)Z'_-(N-1)], \end{aligned} \quad (4.2)$$

and, likewise,

$$Z'_-(N) = \frac{1}{2} [(1 - U)(1 - T)Z'_+(N-1) + (1 - U)(1 + T)Z'_-(N-1)]. \quad (4.3)$$

We can put this in matrix form:

$$\begin{bmatrix} Z'_+(N) \\ Z'_-(N) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} (1 + U)(1 + T) & (1 + U)(1 - T) \\ (1 - U)(1 - T) & (1 - U)(1 + T) \end{bmatrix} \begin{bmatrix} Z'_+(N-1) \\ Z'_-(N-1) \end{bmatrix}. \quad (4.4)$$

Iterating this, and paying careful attention to the initial case  $N = 2$ , we obtain the formula

$$\begin{bmatrix} Z'_+(N) \\ Z'_-(N) \end{bmatrix} = \frac{1}{2^{N-1}} \begin{bmatrix} (1 + U)(1 + T) & (1 + U)(1 - T) \\ (1 - U)(1 - T) & (1 - U)(1 + T) \end{bmatrix}^{N-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (4.5)$$

The matrix in these expressions is called the *transfer matrix*. If we denote it by  $M$ ,

$$M = M(U, T) = \frac{1}{2} \begin{bmatrix} (1 + U)(1 + T) & (1 + U)(1 - T) \\ (1 - U)(1 - T) & (1 - U)(1 + T) \end{bmatrix}, \quad (4.6)$$

then we have

$$Z'(N) = [1 \quad 1] M^{N-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (4.7)$$

We have derived this expression for  $Z'(N)$  using the “linear” model because the derivation is especially simple to explain. However, we now prefer to replace it with the corresponding formula for the “wrap-around” model, but shall leave the derivation of that formula as an exercise for the reader. To distinguish the two

models, we shall write  $Z''(N)$  for the wrap-around model in this section:

$$Z''(N) = \frac{1}{2^N} \sum_{\pm 1} \prod_{i=1}^N [(1 + \sigma_i \sigma_{i+1} T)(1 + \sigma_i U)], \tag{4.8}$$

where it is understood that  $\sigma_{N+1} = \sigma_1$ . The analogue to equation (4.7) is much nicer:

$$Z''(N) = \text{Tr}(M^N), \tag{4.9}$$

where “Tr” denotes the trace and  $M$  is still the transfer matrix.

It is now clear from elementary linear algebra what to do: we *diagonalize*  $M$ , by finding its eigenvalues,  $\lambda_1$  and  $\lambda_2$ , and conclude that

$$Z''(N) = \lambda_1^N + \lambda_2^N. \tag{4.10}$$

Furthermore, assuming that the eigenvalues are positive real numbers with  $\lambda_1 > \lambda_2$ , then

$$\begin{aligned} F_1' &= \lim_{N \rightarrow \infty} \frac{1}{N} \log Z''(N) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log(\lambda_1^N (1 + (\lambda_2/\lambda_1)^N)) \\ &= \log(\lambda_1). \end{aligned} \tag{4.11}$$

(Note: the main idea here is to diagonalize  $M$ ; this leads to the result  $F_1' = \log(\lambda_1)$  even if one sticks with the “linear” model. Our reason for preferring the wrap-around model is purely aesthetic.)

The rest of the solution is routine: One easily sees that  $\text{Tr}(M) = 1 + T$  and  $\det(M) = T(1 - U^2)$ , and, therefore,  $M$  has the characteristic equation

$$\lambda^2 - (1 + T)\lambda + T(1 - U^2) = 0,$$

so that the eigenvalues are

$$\lambda = \frac{1 + T \pm [(1 + T)^2 - 4T(1 - U^2)]^{1/2}}{2}.$$

Note that  $|U| = |\tanh(\beta J)| < 1$  for real values of the parameters, and therefore,

$$(1 - T)^2 < [(1 + T)^2 - 4T(1 - U^2)] < (1 + T)^2.$$

In any case, the eigenvalues are positive real numbers when  $0 < T < 1$ .

The partition function is of less interest at this point than the free energy. We find

$$F_1' = \log \left[ \frac{1 + T + [(1 + T)^2 - 4T(1 - U^2)]^{1/2}}{2} \right]. \tag{4.12}$$

As a function of  $\beta$ ,  $F_1'$  is analytic on the positive real axis. We interpret this as

meaning that the one-dimensional Ising model does not exhibit a phase transition at any temperature: a string of iron atoms will not spontaneously magnetize (according to this model, anyway).

That was the discouraging result of Ising's doctoral dissertation. The lack of a phase transition can be understood by thinking of spontaneous magnetization as a *cooperative* phenomenon of the lattice, which requires "communication" between lattice sites. But in the one-dimensional lattice, a single defect destroys the only line of communication. For example, a configuration  $\cdots + + + - - - \cdots$  is only negligibly more energetic (i.e., less "favorable") than  $\cdots + + + + + + \cdots$ : only one term in the Hamiltonian changes.

According to Brush [4], Ising "gave some approximate calculations purporting to show that his model could not exhibit a phase transition in three dimensions either." However, the higher-dimensional models do have phase transitions. The "single-defect" argument does not apply: there are many "lines of communication" connecting each pair of lattice sites.

### 5. Spontaneous Magnetization in Two Dimensions

In this section we shall present a proof originally due to Peierls [37], which shows that the two-dimensional Ising model does have a phase transition—i.e., it exhibits spontaneous magnetization at sufficiently low temperatures. For this purpose we shall forsake some of our previous notation and also return to the "flat" model which has a boundary. We shall exploit the boundary to create a preference for the magnetic moments throughout the lattice.

Recall that spontaneous magnetization is the tendency for the magnetic moments to remain in, say, the "up" position after an external magnetic field has been turned off. One way to imagine turning off the field is to "impose" a magnet on the boundary of the lattice by *setting* all  $\sigma_i = +1$  on the boundary—then letting the boundary "move off to infinity," which is what happens anyway when  $N \rightarrow \infty$ . We can then ask the following question: For a lattice site "O" "deep in the interior", what is the probability that  $\sigma_0 = -1$ ?

If there were no magnetic field, this probability would simply be  $1/2$ . But the fixed "+" signs on the boundary tend to make the lattice sites near them be positive also, and this creates a "ripple effect" that goes some distance into the lattice. When the temperature is high, this effect is quickly dissipated, but for low temperatures it is possible that the "ripple" will travel a considerable distance inward. What we shall show is that the effect can in fact travel all the way through the lattice; i.e., for sufficiently low temperatures, the probability that  $\sigma_0 = -1$  is less than  $1/2$  by an amount which is independent of the lattice size. The proof is quite beautiful in its elegant use of crude estimates to bound the probability.

Recall that the probability for a given configuration  $\sigma = (\sigma_1, \dots, \sigma_N)$  is given by the formula

$$\text{Prob}(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z}, \quad (5.1)$$

where  $H$  is the Hamiltonian and

$$Z = \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}, \tag{5.2}$$

where  $\Omega$  is the set of all configurations which are positive on the boundary. (The “ $\pm 1$ ” notation is not sufficient here; the heart of Peierls’ proof is to consider the sum over various subsets of configurations.) Suppose we label the lattice so that  $\sigma_0$  corresponds to a lattice site somewhere in the “middle” of the lattice. Then

$$\text{Prob}(\sigma_0 = -1) = \frac{1}{Z} \sum_{\sigma \in \Omega_0} e^{-\beta H(\sigma)}, \tag{5.3}$$

where  $\Omega_0 \subseteq \Omega$  is the set of configurations  $\sigma$  for which  $\sigma_0 = -1$ .

Consider a typical configuration in the set  $\Omega_0$ , such as the one shown in Figure 9:

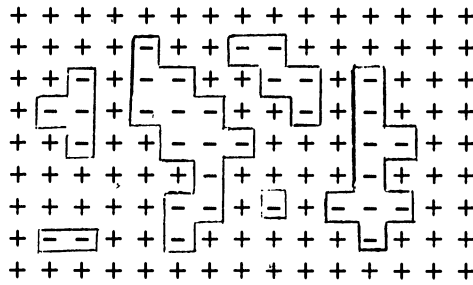


FIG. 9.

Because of the boundary condition, we can think of any configuration as consisting of “islands” of negative signs in a positive “ocean.” Some of the islands may have interior “lakes”, but they all have “shoreslines.” Finally, one of the islands contains the site “0.”

Now a “shoreline” is a closed path consisting of line segments connecting the midpoints of adjacent squares in the lattice. The main characteristic is that each segment of shoreline separates a positive sign (ocean) from a negative sign (land). Thus a given shoreline corresponds to a set of bonds  $\langle i, j \rangle$  for which  $\sigma_i \sigma_j = -1$ .

Suppose now we draw a shoreline  $S$ , creating an island around “0,” and say that its length is  $n(S)$ . Let’s consider the set  $\Omega_S$  of configurations in  $\Omega_0$  having  $S$  as a shoreline. Then

$$\begin{aligned} \text{Prob}(\Omega_S) &= \frac{1}{Z} \sum_{\sigma \in \Omega_S} e^{-\beta H(\sigma)} \\ &= \frac{1}{Z} \sum_{\sigma \in \Omega_S} e^{-\beta E n(S)} e^{\beta E \sum_{\langle i, j \rangle \notin S} \sigma_i \sigma_j} \\ &= e^{-\beta E n(S)} \frac{1}{Z} \sum_{\sigma \in \Omega_S} e^{\beta E \sum_{\langle i, j \rangle \notin S} \sigma_i \sigma_j}. \end{aligned} \tag{5.4}$$



Given  $\sigma \in \Omega_S$ , we can form another configuration,  $\sigma'$ , by changing all the signs inside the shoreline  $S$ . For a fixed shoreline  $S$ , the map  $\sigma \rightarrow \sigma'$  is one-to-one. We shall let  $\Omega'_S$  denote the image of  $\Omega_S$  under this mapping. One easily sees that

$$\sum_{\langle i, j \rangle \notin S} \sigma_i \sigma_j = \sum_{\langle i, j \rangle} \sigma'_i \sigma'_j - n(S), \tag{5.5}$$

and, therefore,

$$\sum_{\langle i, j \rangle \notin S} \sigma_i \sigma_j < \sum_{\langle i, j \rangle} \sigma'_i \sigma'_j. \tag{5.6}$$

Plugging this inequality into the previous computation (noting that  $\beta E > 0$ ), we have

$$\text{Prob}(\Omega_S) < e^{-\beta E n(S)} \frac{1}{Z} \sum_{\sigma' \in \Omega'_S} e^{\beta E \sum_{\langle i, j \rangle} \sigma'_i \sigma'_j} = e^{-\beta E n(S)} \frac{1}{Z} \sum_{\sigma' \in \Omega'_S} e^{-\beta H(\sigma')}. \tag{5.7}$$

(We have also used the fact that  $\sigma \rightarrow \sigma'$  is one-to-one, so that the sum over  $\Omega_S$  can be replaced by the sum over  $\Omega'_S$ .) But now, since  $e^{-\beta H(\sigma)} > 0$  for *all* configurations  $\sigma$ , we can replace the sum over  $\Omega'_S$  by a sum over *all* configurations! Thus,

$$\text{Prob}(\Omega_S) < e^{-\beta E n(S)} \frac{1}{Z} \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)} = e^{-\beta E n(S)}. \tag{5.8}$$

since  $Z = \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}$ !

Now consider the set  $\mathcal{S}$  of all shorelines which surround the lattice site “0.” Then

$$\begin{aligned} \text{Prob}(\sigma_0 = -1) &= \sum_{S \in \mathcal{S}} \text{Prob}(\Omega_S) \\ &< \sum_{S \in \mathcal{S}} e^{-\beta E n(S)} \\ &= \sum_{n=4}^{\infty} s(n) e^{-\beta E n}, \end{aligned} \tag{5.9}$$

where  $s(n)$  denotes the number of shorelines of length  $n$  which surround the lattice site “0.” Thus we have one last chore before the denouement: we have to bound  $s(n)$ . We shall do this in a wonderfully crude manner.

A shoreline, remember, is simply a path in the lattice connecting the midpoints of adjacent squares. Since we required our shorelines to surround the lattice site “0,” the path cannot wander too far away from “0”: if the shoreline is of length  $n$ , it must be contained in a square with sides of length  $n/\sqrt{2}$ . (See Figure 10.)

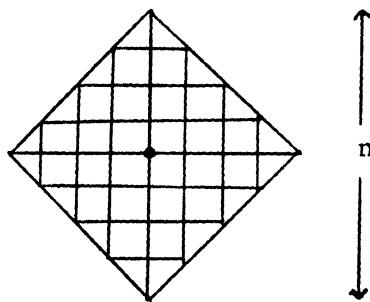


FIG. 10.

Now let  $r(n)$  denote the number of “random walks” of length  $n$  which originate inside this square. (Minor remark: the random walk, like the shoreline, goes from midpoint to midpoint, not lattice site to lattice site.) It is easy to see that

$$s(n) < \frac{1}{n}r(n). \tag{5.10}$$

(The factor  $1/n$  comes from the fact that each shoreline gets counted  $n$  times as a random walk, since any point along it can be considered as the origin.) But the random walk has  $(n/\sqrt{2})^2 = n^2/2$  possible starting points, and then  $4^n$  possible paths. (This can be reduced to  $4 \cdot 3^{n-1}$  if you disallow “backtracking,” but there’s no real gain in doing so.) Thus

$$s(n) < \frac{1}{2}n4^n, \tag{5.11}$$

and, therefore,

$$\begin{aligned} \text{Prob}(\sigma_0 = -1) &< \sum_{n=4}^{\infty} \frac{1}{2}n4^n e^{-\beta E n} \\ &= \frac{1}{2} \sum_{n=4}^{\infty} n(4e^{-\beta E})^n \\ &< \frac{1}{2} \sum_{n=1}^{\infty} n(4e^{-\beta E})^n. \end{aligned} \tag{5.12}$$

The denouement is at hand: recall that

$$\frac{1}{1-x} = 1 + x + x^2 + \dots = \sum_{n=0}^{\infty} x^n \text{ if } |x| < 1.$$

Hence, by differentiating and multiplying by  $x$ ,

$$\frac{x}{(1-x)^2} = x(1 + 2x + 3x^2 + \dots) = \sum_{n=1}^{\infty} nx^n.$$

Thus, letting  $x = 4e^{-\beta E}$ , we have

$$\text{Prob}(\sigma_0 = -1) < \frac{1}{2} \left[ \frac{4e^{-\beta E}}{(1 - 4e^{-\beta E})^2} \right]. \tag{5.13}$$

The conclusion is clear: by taking  $\beta$  sufficiently large (which corresponds to low temperature), the right-hand side of the inequality can be made arbitrarily small, in a way which is independent of the size of the lattice. Thus spontaneous magnetization is guaranteed at some temperature.

### 6. The Critical Point in Two Dimensions

Peierls' proof that a phase transition exists for the two-dimensional model can put a lower bound on the critical temperature, but cannot locate it exactly. In this section we shall present a lovely combinatorial argument due to Kramers and Wannier [26] which proves the following result for the two-dimensional Ising model in the zero-field case: if there is a unique phase transition for  $F_2'(T, 0)$  on the interval  $(0, 1)$ , then it occurs precisely at  $T_c = \sqrt{2} - 1$ . (The historical progression of results is thus Peierls' 1936 proof that a phase transition exists; Kramers and Wanniers' 1941 proof that it occurs at  $T_c = \sqrt{2} - 1$ ; and Onsager's complete solution in 1944.)

The starting point is the combinatorial interpretation of  $c_2(n)$  as the number of "even subgraphs with  $n$  edges," whose connected components are closed paths on the lattice. In this section it will be convenient to refer to such subgraphs as "closed paths of length  $n$ ," even when the "path" has several components.

In general, a closed path of finite length in the plane may be associated with the bounded region which it encloses or, alternatively, with the unbounded region outside of it. If we consider the set of bounded regions and their complements, there is a two-to-one correspondence between such "shaded" regions and closed paths in the plane. (See Figure 11.)

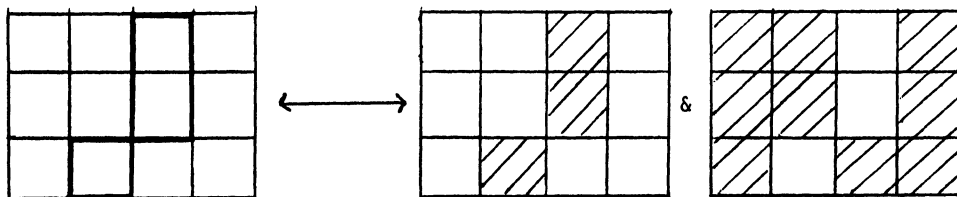


FIG. 11.

On the lattice, a "shaded" region can be designated by enumerating all the *squares* of the lattice,  $i = 1, 2, 3, \dots$ , and assigning an independent variable, say,  $\tau_i$ , to each square:  $\tau_i = 1$  if the square  $i$  is shaded, and  $\tau_i = -1$  if square  $i$  is unshaded. Unfortunately, when we restrict regions to a finite lattice, the correspondence is no longer precisely two-to-one: the "zero" path and the path around the boundary both correspond to both the "empty" region and the full lattice region. This can be

fixed by eliminating the boundary with the wrap-around lattice, but then even worse things happen: a simple loop around the lattice does not correspond to *any* region (much less to two of them). One should note, however, that such paths are of very long length— $\sqrt{N}$ , if one takes the overall lattice to be a square. The contribution of these paths to the partition function is thus far out in the power series, and hence we expect it to vanish in the thermodynamic limit. In keeping with our disregard for mathematical rigor (but only when its safe to do so!) we shall take it for granted that this is what happens.

In spite of these drawbacks, we shall use the wrap-around model since it simplifies some of the notation. In particular, there are as many squares as there are lattice sites, and we can enumerate them according to, say, the lattice site  $i$  in the lower left-hand corner. Each square has four “nearest neighbors” with which it shares an edge: squares  $i$  and  $j$  are nearest neighbors if and only if lattice sites  $i$  and  $j$  are nearest neighbors.

Suppose we are given a configuration for the *squares*,  $\tau = (\tau_1, \dots, \tau_N)$ . How long is the closed-path boundary of the corresponding shaded region on the lattice? We answer this by first noting that the edge joining squares  $i$  and  $j$  is part of the boundary if and only if  $\tau_i\tau_j = -1$ —i.e., if and only if one square is shaded and the other is not. Thus the length,  $n$ , of the closed path is given by

$$n(\tau) = \sum_{\langle i, j \rangle} \delta(i, j) \quad \text{where} \quad \delta(i, j) = \begin{cases} 1 & \text{if } \tau_i\tau_j = -1 \\ 0 & \text{if } \tau_i\tau_j = 1. \end{cases}$$

Our interest is actually in  $T^n$ . We write  $T^{n(\tau)} = \prod_{\langle i, j \rangle} T^{\delta(i, j)}$  to begin with, but then notice that we can rewrite

$$\begin{aligned} T^{\delta(i, j)} &= \frac{1}{4} [(\tau_i + \tau_j)^2 + (\tau_i - \tau_j)^2 T] \\ &= \frac{1}{2} [(1 + T) + \tau_i\tau_j(1 - T)] \\ &= \frac{1 + T}{2} \left[ 1 + \tau_i\tau_j \left( \frac{1 - T}{1 + T} \right) \right]. \end{aligned}$$

Therefore,

$$T^{n(\tau)} = \left( \frac{1 + T}{2} \right)^{2N} \prod_{\langle i, j \rangle} \left[ 1 + \tau_i\tau_j \left( \frac{1 - T}{1 + T} \right) \right]. \tag{6.1}$$

Now remembering the combinatorial interpretation of  $c_2(n)$ , remembering that there is a two-to-one correspondence between closed paths and shaded regions, and forgetting that this correspondence breaks down at some point, we have

$$Z'(T) = \sum_{n=0}^{\infty} c_2(n) T^n \cong \frac{1}{2} \sum_{\pm 1} T^{n(\tau)}, \tag{6.2}$$

where the sum is now over all configurations  $\tau = (\tau_1, \dots, \tau_N)$ . (The approximate

equality ( $\cong$ ) reflects the breakdown of the correspondence between closed paths and shaded regions; more precisely, it means that the formal power series are identical out to a power determined by the size of the lattice.) Plugging in (6.1), we have

$$\begin{aligned} Z'(T) &\cong \frac{1}{2} \left( \frac{1+T}{2} \right)^{2N} \sum_{\pm 1} \prod_{\langle i,j \rangle} \left[ 1 + \tau_i \tau_j \left( \frac{1-T}{1+T} \right) \right] \\ &\cong \frac{1}{2} \left( \frac{(1+T)^2}{2} \right)^N Z' \left( \frac{1-T}{1+T} \right). \end{aligned} \quad (6.3)$$

The partition function reappears on the right-hand side! When we now take the limit  $N \rightarrow \infty$ , the approximate equality becomes exact, and we obtain the result

$$F'(T) = \log \left( \frac{(1+T)^2}{2} \right) + F' \left( \frac{1-T}{1+T} \right). \quad (6.4)$$

Recall that values of  $T$  near 0 correspond to high temperatures, while values near 1 correspond to low temperatures. Observe now that when  $T$  is near 0,  $(1-T)/(1+T)$  is near 1 and vice versa:  $T \rightarrow (1-T)/(1+T)$  maps the interval onto itself. Thus (6.4) is a formula—or functional equation, if you will—relating high and low temperatures. Viewing  $F'$  as an analytic function, (6.4) provides an analytic continuation of  $F'$ . In particular, if  $F'$  is nonanalytic at  $T$ , then it is also nonanalytic at  $(1-T)/(1+T)$ ; i.e., phase transitions will occur in *pairs*. Thus if we assume that there is a unique (physical) phase transition in the interval  $(0, 1)$ , then it can only occur at the solution of the equation

$$T = \frac{1-T}{1+T},$$

which is obviously at  $T = T_c = \sqrt{2} - 1$ .

We conclude by observing that this result is in agreement with Onsager's solution

$$F'(T, 0) = \frac{1}{2} \int_0^1 \int_0^1 \log \left[ (T^2 + 1)^2 - 2T(1 - T^2) [\cos(2\pi x) + \cos(2\pi y)] \right] dx dy.$$

For  $0 < T < 1$ , we have

$$\begin{aligned} (T^2 + 1)^2 - 2T(1 - T^2)(\cos(2\pi x) + \cos(2\pi y)) &\geq (T^2 + 1)^2 - 4T(1 - T^2) \\ &= T^4 + 4T^3 + 2T^2 - 4T + 1 \\ &= (T^2 + 2T - 1)^2 \end{aligned}$$

with equality only when  $\cos(2\pi x) = \cos(2\pi y) = 1$ . Thus the integrand in Onsager's solution has a singularity if and only if  $T^2 + 2T - 1 = 0$ —i.e.,  $T = \sqrt{2} - 1$ .

## 7. Concluding Remarks

The Ising model has become a vast subject. This article has touched only on portions of it, and the simplest ones at that. We have not spoken, for instance, of critical exponents, correlation functions, or renormalization. We have adhered rather strictly to a combinatorial approach, ignoring important algebraic and representation-theoretic techniques. Our purpose here has been to introduce the Ising model to a wider audience, not to expound on what the experts already know; the combinatorial interpretation seems to be the most accessible avenue, and has indeed led to several of the advances in the subject. The author hopes that this article may encourage some of its readers to dig more deeply into the Ising model. There is a lot of gold left in the mine.

The author would like to thank Lynn Steen at St. Olaf and the referee for their helpful comments.

## REFERENCES

1. R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, 1982.
2. N. L. Biggs, *Interaction Models*, Cambridge University Press, 1977.
3. A. B. Bortz, Proofs of Spontaneous Magnetization Using the Peierls Method, *Am. J. Phys.*, 40 (1972) 1524–1528.
4. S. G. Brush, History of the Lenz-Ising Model, *Rev. Mod. Phys.*, 39 (1967) 883–893.
5. \_\_\_\_\_, *Statistical Physics and the Atomic Theory of Matter*, from Boyle and Newton to Landau and Onsager, Princeton University Press, 1983.
6. J. C. Dash, Two-Dimensional Matter, *Scientific American*, 228 (1973) 30–40.
7. R. L. Dobrushin, The existence of a phase transition in the two- and three-dimensional Ising models, *Theory Prob. Appl.* 10 (1965) 193–213.
8. C. Domb, Series Expansions for Ferromagnetic Models, *Adv. Phys.*, 19 (1970) 339–370.
9. \_\_\_\_\_ and M. S. Green, eds., *Phase Transitions and Critical Phenomena*, Academic Press, vols. 1–6, 1972–1976.
10. F. J. Dyson, Existence and Nature of Phase Transitions in One-Dimensional Ising Ferromagnets, in *Mathematical Aspects of Statistical Mechanics* (SIAM-AMS Proc. vol. 5, J. C. T. Poole, ed.), AMS, 1972.
11. R. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, 1985.
12. R. P. Feynman, *Statistical Mechanics*, Benjamin, 1972.
13. M. E. Fisher, Simple Ising Models Still Thrive, *Physica* 106A (1981), 28–47.
14. J. Frohlich, On the Mathematics of Phase Transitions and Critical Phenomena, in *Proc. of the International Congress of Mathematicians, Helsinki (1978)*, 896–904.
15. M. L. Glasser, Exact Partition Function for the Two-Dimensional Ising Model, *Am. J. Phys.*, 38 (1970) 1033–1036.
16. R. B. Griffiths, Peierls' proof of spontaneous magnetization in a two-dimensional Ising ferromagnet, *Phys. Rev.* 136A (1964), 437–439.
17. F. Harary, ed., *Graph Theory and Theoretical Physics*, Academic Press, 1967.
18. E. Ising, Beitrag zur Theorie des Ferromagnetismus, *Z. Physik*, 31 (1925) 253–258.
19. M. Kac, *Enigmas of Chance*, Harper-Row, 1985.
20. \_\_\_\_\_, Phase Transitions, in *The Physicist's Conception of Nature* (J. Mehra, ed.), Reidel (1973) 814–826.
21. \_\_\_\_\_, Probability in Classical Physics; in *Proc. of Symposia in Applied Mathematics*, vol. 7, AMS (1957) 73–86.

22. \_\_\_\_\_ and J. C. Ward, A Combinatorial Solution of the Two Dimensional Ising Model, *Phys. Rev.* 88 (1952) 1332–1337.
23. B. Kaufman and L. Onsager, Crystal Statistics II. Partition Function Evaluated by Spinor Analysis, *Phys. Rev.*, 76 (1949) 1232–1243.
24. \_\_\_\_\_, Crystal Statistics III. Short-Range Order in a Binary Ising Model, *Phys. Rev.*, 76 (1949) 1244–1252.
25. J. B. Kogut, An Introduction to Lattice Gauge Theory and Spin Systems, *Rev. Mod. Phys.*, 51 (1979) 659–713.
26. H. A. Kramers and G. H. Wannier, Statistics of the two-dimensional ferromagnet, I and II, *Phys. Rev.*, 60 (1941) 252–262, 263–276.
27. L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pergamon Press, 1980.
28. T. D. Lee and C. N. Yang, Statistical Theory of Equations of State and Phase Transitions I. Theory of Condensation, *Phys. Rev.*, 87 (1952) 404–409.
29. \_\_\_\_\_, Statistical Theory of Equations of State and Phase Transitions II. Lattice Gas and the Ising Model, *Phys. Rev.*, 87 (1952) 410–419.
30. W. Lenz, Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern, *Z. Physik* 21 (1920), 613–615.
31. T. M. Liggett, *Interacting Particle Systems*, Springer-Verlag, 1985.
32. S.-K. Ma, *Statistical Mechanics*, World Scientific, 1985.
33. B. M. McCoy and T. T. Wu, *The Two-Dimensional Ising Model*, Harvard University Press, 1973.
34. O. G. Mouritsen, *Computer Studies of Phase Transitions and Critical Phenomena*, Springer-Verlag, 1984
35. G. F. Newell and E. W. Montroll, On the theory of the Ising Model of Ferromagnetism, *Rev. Mod. Phys.*, 25 (1953) 353–389.
36. L. Onsager, Crystal statistics I. A Two-Dimensional Model With an Order-Disorder Transition, *Phys. Rev.*, 65 (1944) 117–149.
37. R. Peierls, Ising's Model of Ferromagnetism, *Proc. Camb. Phil. Soc.*, 32 (1936), 477–481.
38. D. Ruelle, *Statistical Mechanics: Rigorous Results*, Benjamin, 1969.
39. Y. G. Sinai, *Theory of Phase Transitions: Rigorous Results*, Pergamon Press, 1982.
40. H. E. Stanley, ed., *Cooperative Phenomena near Phase Transitions (A bibliography with selected readings)*, M.I.T. Press, 1973.
41. \_\_\_\_\_, *Introduction to Phase Transitions and Critical Phenomena*, Oxford University Press, 1985.
42. C. Thompson, *Mathematical Statistical Mechanics*, Princeton University Press, 1972.
43. G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics*, AMS, 1963.
44. G. H. Wannier, *Statistical Physics*, Wiley, 1966.
45. K. G. Wilson, Problems in Physics with Many Scales of Length, *Scientific American*, 241 (1979) 158–179.
46. C. N. Yang, The Spontaneous Magnetization of a Two-Dimensional Ising Model, *Phys. Rev.*, 85 (1952) 808–816.
47. J. M. Ziman, *Models of Disorder*, Cambridge University Press, 1979.