Statistics of the Two-Dimensional Ferromagnet. Part I

H. A. KRAMERS, *University of Leiden, Leiden, Holland*

AND

G. H. WANNIER, *University of Texas, Austin, Texas*[1]

(Received April 7, 1941)

In an effort to make statistical methods available for the treatment of cooperational phenomena, the Ising model of ferromagnetism is treated by rigorous Boltzmann statistics. A method is developed which yields the partition function as the largest eigenvalue of some finite matrix, as long as the manifold is only one dimensionally infinite. The method is carried out fully for the linear chain of spins which has no ferromagnetic properties. Then a sequence of finite matrices is found whose largest eigenvalue approaches the partition function of the two-dimensional square net as the matrix order gets large. It is shown that these matrices possess a symmetry property which permits location of the Curie temperature if it exists and is unique. It lies at

$$J/kT_c=0.8814$$

if we denote by J the coupling energy between neighboring spins. The symmetry relation also excludes certain forms of singularities at T_c, as, e.g., a jump in the specific heat. However, the information thus gathered by rigorous analytic methods remains incomplete.

TRANSITION temperatures of various types are a well-known phenomenon in the study of matter, and the statistical distribution laws form a generally accepted piece of theory. It is also generally believed that the former are a consequence of the latter. This is, however, by no means immediately obvious and an examination of the literature shows that there is not more than one case in which a proof of this fact has been attempted. The case which has received successful treatment is the condensation of vapors.[2] This paper wishes to carry out a similar treatment for the Curie transition of ferromagnets.

The problem has a mechanical and a statistical aspect. On the mechanical side we wish to improve our understanding of the responsible coupling forces. On the statistical side we wish to derive with certainty the thermal properties from a reasonably accurate mechanical model. Both aspects have received extensive attention. Quantum theory has explained satisfactorily the origin and nature of the coupling forces. There are also several theories available which explain in terms of them the thermal behavior of ferromagnets. Not one, however, applies just straight statistics to the mechanical data.[3] Generally some simplifying assumption is introduced to facilitate the evaluation of the partition function. It follows that the results obtained are not necessarily a consequence of the mechanical model, but may well be due to the statistical approximation.

The present paper is an attempt to gain sound statistical information about some model of a ferromagnet. The Ising model has been chosen because its extreme simplicity makes it particularly suitable for such a purpose.

In Part I, we shall show that the task of finding the state sum can be reduced to finding the largest eigenvalue of some matrix. The matrix will be very simple in the case of the linear chain and we shall re-derive by this method the results of Ising. No such simple solution will be possible for the two-dimensional square net where the matrix is infinite. However, some precise information can still be gained which will be collected in the latter sections of this paper.

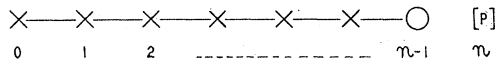

FIG. 1. Building up the linear chain. The elementary step consists in placing a spin in position $[P]$, whose sign depends only on the spin ○.

[1] Owing to communication difficulties, one of the authors (G. H. W.) is entirely responsible for the printed text.

[2] J. E. Mayer, J. Chem. Phys. **5**, 67 (1937); J. E. Mayer and Ph. G. Ackermann, J. Chem. Phys. **5**, 74 (1937); M. Born, Physica **4**, 1034 (1937); B. Kahn, Dissertation Utrecht, 1938. The last paper has the most rigorous treatment of the matter.

[3] For a detailed discussion of various approximations see Part II.

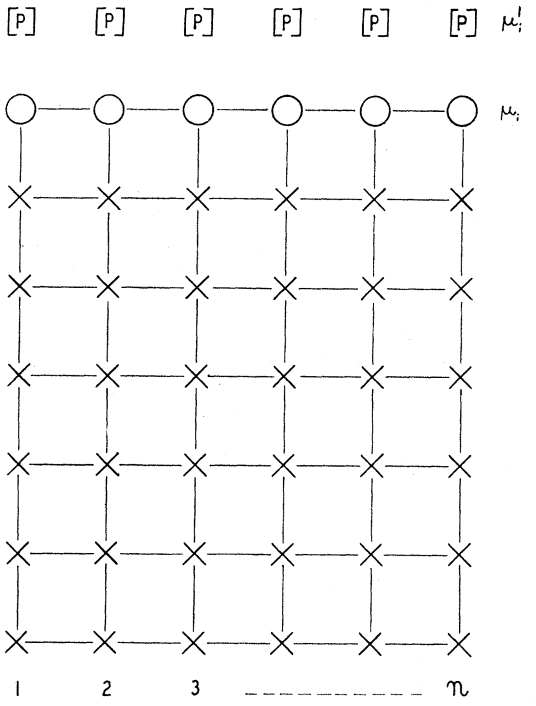

FIG. 2. Building up the infinite strip. Elementary step consists in filling all positions $[P]$ whose situation depends on each other and the spins ○.

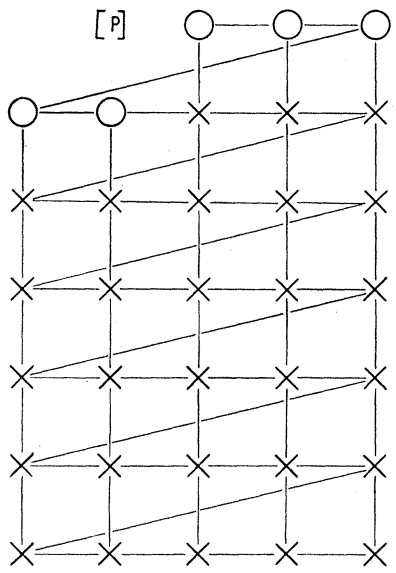

FIG. 3. Building up the infinite screw. Elementary step $[P]$ demands knowledge of spins ○, partly for the step itself, but mostly for later similar steps to follow.

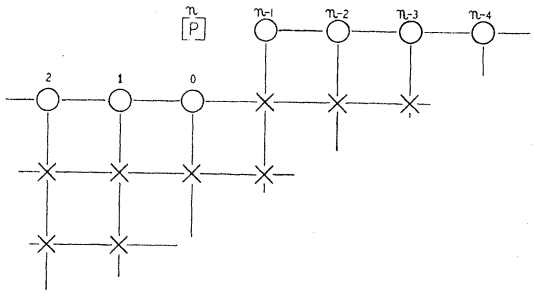

FIG. 4. Numbering of spins in the infinite screw problem.

In Part II, we shall complement this knowledge by approximate treatments. Some of them are already well known, as for instance the power series approximations, the Heisenberg method, the order-disorder method of Bethe. In addition, we wish to add two treatments of our own. Both are based on the matrix method. One will be a semi-numerical treatment to answer a specific question left open in Part I, the other a new approximation method giving results in closed form. It will be shown that this latter is very much superior to the older procedures.

1. THE MECHANICAL MODEL

The Ising model can be explained as follows. Assume a set of spins $\mu_1, \mu_2, \mu_3 \cdots \mu_N$ arranged in some regular order. Let each of the spins be capable of two orientations which we characterize by $\mu_i = +1$ and $\mu_i = -1$. Then the Ising model assumes that the forces on each spin depend only on the orientation of its immediate neighbors in addition to an eventually applied magnetic field. In particular, if all direct neighbors of a given spin are equivalent the model contains only two parameters, namely the magnetic moment m of each spin and a quantity J which is the energy gained if two neighbors change from an antiparallel to a parallel position. With these two definitions the total energy E takes the form

$$E = -(\tfrac{1}{2}J) \sum_{\langle i,k \rangle} \mu_i \mu_k - mH \sum_i \mu_i, \qquad (1)$$

where here as in the future $\sum_{\langle i,k \rangle}$ shall mean that the sum is carried out over all pairs (i, k) which are direct neighbors.

Most statistical questions concerning (1) can be considered solved if we can evaluate the

so-called partition function f

$$f=\sum_{\mu_i=\pm 1}\exp[K\sum_{\langle i,k\rangle}\mu_i\mu_k+C\sum_i\mu_i] \tag{2}$$

with

$$K=\tfrac{1}{2}J/kT \tag{3}$$

and

$$C=mH/kT. \tag{4}$$

The bold face summation sign $\sum_{\mu_i=\pm 1}$ is to be understood to extend over all possible states of the system, i.e., it would have to be written explicitly as

$$\sum_{\mu_i=\pm 1}\sim\sum_{\mu_1=\pm 1}\sum_{\mu_2=\pm 1}\sum_{\mu_3=\pm 1}\cdots\sum_{\mu_N=\pm 1}.$$

Once f is obtained most important physical consequences can be derived from it. We obtain for instance the total magnetization M and the total energy E:

$$M=m\,\partial\log f/\partial C \tag{5}$$

and

$$E=-MH-\tfrac{1}{2}J\,\partial\log f/\partial K. \tag{6}$$

2. The Linear Chain

Ising himself carried out the calculation (2) for the linear chain.[4] However, we shall reconsider this problem since it forms an easy introduction for the eigenvalue method of evaluating f.

Let us suppose first the chain to be finite with n members $\mu_0, \mu_1, \mu_2, \cdots\mu_{n-1}$, as indicated in Fig. 1. Then, by Boltzmann's theorem, the probability for a particular arrangement of spins $\mu_0=+1$, $\mu_1=+1$, $\mu_2=-1$, $\mu_3=+1$, $\cdots\mu_{n-1}=-1$ is proportional to $\exp(-E/kT)$, because every arrangement has weight 1. Since the energy E is given by (1) this probability is proportional to

$$\exp[K(\mu_0\mu_1+\mu_1\mu_2+\mu_2\mu_3+\cdots+\mu_{n-2}\mu_{n-1})+C(\mu_0+\mu_1+\mu_2+\cdots+\mu_{n-2}+\mu_{n-1})],$$

where K and C are given by (3) and (4). Exactly the same consideration is possible if we add one extra spin in the position $[P]$ of Fig. 1. The resulting expression is the same as above except that both sums extend to μ_n. It follows that the two probabilities differ from each other only through the factor

$$\exp[K\mu_{n-1}\mu_n+C\mu_n].$$

From these all-over probabilities others answering more simple questions can be obtained by summation. Let us determine for example from our first expression the probability $P(\mu_{n-1})$ that μ_{n-1} has either value regardless of the values of $\mu_0, \mu_1, \mu_2, \cdots\mu_{n-2}$. This is easily found to be

$$P(\mu_{n-1})\sim\sum_{\mu_0,\mu_1,\cdots\mu_{n-2}=\pm 1}\exp[K(\mu_0\mu_1+\mu_1\mu_2+\cdots\mu_{n-2}\mu_{n-1})+C(\mu_0+\mu_1+\cdots\mu_{n-1})].$$

By summing the second probability containing μ_n over the same μ's a probability $P(\mu_{n-1}, \mu_n)$ can be obtained giving the chance for any one of the four combinations $++$, $+-$, $-+$, $--$. The quantities $P(\mu_{n-1})$ and $P(\mu_{n-1}, \mu_n)$ still differ by the same factor, *viz.*,

$$\lambda P(\mu_{n-1}, \mu_n)=P(\mu_{n-1})\exp[K\mu_{n-1}\mu_n+C\mu_n],$$

the unknown factor λ entering because the Boltzmann exponentials are only proportional to probabilities.

If we sum both sides in the above expression over $\mu_{n-1}=\pm 1$ we get the probability $P(\mu_n)$ for μ_n having either value in terms of the same probabilities for μ_{n-1} before μ_n was added. However, *if the chain is very long*, the physical situation described by the two P's is identical. Hence $P(\mu_n)$ and $P(\mu_{n-1})$ must be the same mathematical functions of their argument

$$\lambda P(\mu_n)=\sum_{\mu_{n-1}=\pm 1}P(\mu_{n-1})\exp[K\mu_{n-1}\mu_n+C\mu_n].$$

These two linear equations have the form of a matrix eigenvalue problem. If we symmetrize the

[4] E. Ising, Zeits. f. Physik **31**, 253 (1925).

matrix by the substitution

$$a(\mu)=P(\mu)\exp\tfrac{1}{2}C\mu$$

the problem takes the form

$$\sum_{\mu'=\pm1}\mathfrak{K}(\mu,\mu')\,a(\mu')=\lambda\,a(\mu) \tag{7a}$$

with

$$\mathfrak{K}(\mu,\mu')=\exp[K\mu\mu'+\tfrac{1}{2}C\mu+\tfrac{1}{2}C\mu']. \tag{7b}$$

We shall consider the solutions of (7) later on.

Before doing so let us clear up the significance of the two eigenvalues λ_1 and λ_2.[5] This can be done most easily by using the fundamental theorem which develops any matrix in terms of its eigenvectors:

$$\mathfrak{K}(\mu_1,\mu_2)=\lambda_1 a_1(\mu_1)a_1(\mu_2)+\lambda_2 a_2(\mu_1)a_2(\mu_2),$$

where $a_1(\mu)$ and $a_2(\mu)$ denote the eigenvectors belonging to λ_1 and λ_2, respectively. They are orthogonal and may be assumed normalized

$$\sum_{\mu=\pm1}a_i(\mu)a_k(\mu)=\delta_{ik}.$$

This permits us to unite two $\mathfrak{K}$'s as follows

$$\sum_{\mu_2}\mathfrak{K}(\mu_1,\mu_2)\mathfrak{K}(\mu_2,\mu_3)=\lambda_1^2 a_1(\mu_1)a_1(\mu_3)+\lambda_2^2 a_2(\mu_1)a_2(\mu_3)$$

and next

$$\sum_{\mu_2}\sum_{\mu_3}\mathfrak{K}(\mu_1,\mu_2)\mathfrak{K}(\mu_2,\mu_3)\mathfrak{K}(\mu_3,\mu_4)=\lambda_1^3 a_1(\mu_1)a_1(\mu_4)+\lambda_2^3 a_2(\mu_1)a_2(\mu_4)$$

and so on until finally

$$\sum_{\mu_2,\,\mu_3\cdots\mu_N}\mathfrak{K}(\mu_1,\mu_2)\mathfrak{K}(\mu_2,\mu_3)\mathfrak{K}(\mu_3,\mu_4)\cdots\mathfrak{K}(\mu_N,\mu_{N+1})=\lambda_1^N a_1(\mu_1)a_1(\mu_{N+1})+\lambda_2^N a_2(\mu_1)a_2(\mu_{N+1}).$$

If we close the ring of spins by the assumption $\mu_{N+1}=\mu_1$ and sum over this last spin we get

$$\sum_{\mu_i}\mathfrak{K}(\mu_1,\mu_2)\mathfrak{K}(\mu_2,\mu_3)\cdots\mathfrak{K}(\mu_N,\mu_1)=\lambda_1^N+\lambda_2^N.$$

Substituting for the matrices $\mathfrak{K}$ their values as given by (7) we verify that the expression on the left-hand side is exactly the partition function f, as defined by (2), for a closed ring of N spins

$$f_N=\lambda_1^N+\lambda_2^N. \tag{8}$$

Finally, if the length N of the chain tends to infinity the smaller root λ_2 may be neglected.

What is the value of this larger root? The eigenvalue problem (7) reads

$$\begin{pmatrix}e^{K+C} & e^{-K}\\ e^{-K} & e^{K-C}\end{pmatrix}\begin{pmatrix}a(+)\\ a(-)\end{pmatrix}=\lambda\begin{pmatrix}a(+)\\ a(-)\end{pmatrix} \tag{7c}$$

and hence

$$\lambda=e^K\cosh C+(e^{2K}\sinh^2 C+e^{-2K})^{\frac{1}{2}}. \tag{9}$$

It has already been pointed out by Ising himself[4] that a linear chain of spins is not ferromagnetic. This can easily be verified by calculating the total magnetization with the help of (5) and (8):

$$M=mN\sinh C/(\sinh^2 C+e^{-4K})^{\frac{1}{2}}, \tag{10a}$$

an expression which, because of (4), vanishes with H. The initial molecular paramagnetic susceptibility comes out to be

$$\chi=(m^2/kT)\exp(J/kT). \tag{10b}$$

[5] The elegant form of procedure used here is due to Mr. E. Montroll who applied it first to the theory of molecular chains.

At low temperatures this value is very much larger than without the cooperative coupling J, but it is still finite.

In the absence of a magnetic field we have

$$\lambda = 2\cosh K,$$

which gives for the energy as a function of temperature

$$E = -\tfrac{1}{2}NJ\tanh K = -\tfrac{1}{2}NJ\tanh(J/2kT). \tag{11}$$

This is a smooth increase from $-\frac{1}{2}NJ$ to 0 as the temperature rises.

3. Matrix Form of the Square Net Problem

The successful calculation of Section 2 has unfortunately no bearing upon ferromagnetism since the model proves to be paramagnetic only. The situation is entirely different if we consider the still simplified case of a square net of spins having two infinite dimensions. Peierls[6] has proved that this model is ferromagnetic in the sense that it has a non-zero magnetization at absolute zero. We can conclude from that result that the spontaneous magnetization cannot possibly be represented by a single analytic function of temperature since it vanishes identically in the high temperature range. This in turn would make us suspect the existence of a singular point at which the magnetization ceases to vanish identically. It follows that the two-dimensional Ising model is a fair test case for the general statistical theory of ferromagnets.

The reduction of the linear chain problem can be described in a qualitative way as follows. It is possible to build up a chain by repeating constantly one and the same operation, namely adding another spin beyond the one just placed previously. In fact, if the chain is really very long no physical change takes place through the addition of one more spin. The successful mathematical treatment is based on this identity on the one hand and on the other on the fact that the *state of the last spin μ_n is only dependent upon the state of its predecessor μ_{n-1}*. It follows that the function $P(\mu_n)$ depends operationally on $P(\mu_{n-1})$, yet is the same mathematical function of its argument. Exactly this is expressed in Eq. (7).

A strip of spins having infinite length but finite width can obviously be built up in an analogous manner. Let us recall briefly the steps with the help of Fig. 2. First we write down two probability expressions, one referring to all spins marked by crosses and rings (we assume the positions $[P]$ to be vacant), and the other to all crosses, rings and positions $[P]$. Both expressions follow directly from Boltzmann's theorem, that is they are proportional to $\exp(-E/kT)$, E being given by (1). In interpreting (1) we must remember that the neighborhood in $\sum_{\langle i,k\rangle}$ is now both horizontal and vertical, as expressed by the connecting rods in Fig. 2. As a consequence, the factor by which the probabilities differ is equal to

$$\exp[K(\mu_1'\mu_2'+\mu_2'\mu_3'+\cdots+\mu_{n-1}'\mu_n')+K(\mu_1\mu_1'+\mu_2\mu_2'+\cdots\mu_n\mu_n')+C(\mu_1'+\mu_2'+\cdots+\mu_n')],$$

where the indices of the μ's refer to the column number. Only the μ's of the two top lines are involved in the expression, the very top one being indicated by μ_i', the next one by μ_i (see Fig. 2). Probabilities referring only to the two lines just mentioned can be gained by summing over all crossed spins. We are left with a function $P(\mu_i)$ of n variables and $P(\mu_i, \mu_i')$ of $2n$ variables. The two probabilities still differ from each other by the same factor

$$\rho\, P(\mu_i, \mu_i') = P(\mu_i)\exp[K\sum_{i=1}^{n-1}\mu_i'\mu_{i+1}'+K\sum_{i=1}^{n}\mu_i\mu_i'+C\sum_{i=1}^{n}\mu_i'],$$

[6] R. Peierls, Proc. Camb. Phil. Soc. **32**, 477 (1936). Ising's erroneous conclusions for the two-dimensional case are still quoted in some papers, e.g., Lamek Hulthen, Arkiv för Matematik Astronomi och Fysik **26A**, No. 11 (1940).

the constant ρ appearing for the same reason as above (§2). Finally, if the strip is very long $P(\mu_i')=\sum_{\mu_i} P(\mu_i, \mu_i')$ must be the same function of its variables as $P(\mu_i)$. The substitution

$$a(\mu_i)=P(\mu_i)\exp[\tfrac{1}{2}K\sum_{i=1}^{n-1}\mu_i\mu_{i+1}+\tfrac{1}{2}C\sum_{i=1}^{n}\mu_i]$$

will make the resulting eigenvalue problem symmetrical. It reads then

$$\sum_{\mu_i'}\mathfrak{K}(\mu_i, \mu_i')\ a(\mu_i')=\rho\ a(\mu_i) \tag{12a}$$

with

$$\mathfrak{K}(\mu_i, \mu_i')=\exp\ [K\sum_{i=1}^{n}\mu_i\mu_i'+\tfrac{1}{2}K\sum_{i=1}^{n-1}\mu_i\mu_{i+1}+\tfrac{1}{2}K\sum_{i=1}^{n-1}\mu_i'\mu_{i+1}'+\tfrac{1}{2}C\sum_{i=1}^{n}\mu_i+\tfrac{1}{2}C\sum_{i=1}^{n}\mu_i']. \tag{12b}$$

To find the significance of ρ we turn again to the matrix development theorem. We have this time 2^n roots ρ_i since the matrix is of that order. As a starting point, we use the two basic statements

$$\sum_{\mu i} a_p(\mu_i)a_q(\mu_i)=\delta_{pq},$$

$$\mathfrak{K}(\mu_i, \mu_i')=\sum_{p=1}^{2^n}\rho_p\ a_p(\mu_i)\ a_p(\mu_i').$$

If we consider now a finite ring which is n spins wide and m spins in circumference we find that its partition function can be written in terms of the "nuclei" $\mathfrak{K}$ thus

$$f=\sum_{\mu_i^k}\mathfrak{K}(\mu_i^1, \mu_i^2)\mathfrak{K}(\mu_i^2, \mu_i^3)\mathfrak{K}(\mu_i^3, \mu_i^4)\cdots\mathfrak{K}(\mu_i^{m-1}, \mu_i^m)\mathfrak{K}(\mu_i^m, \mu_i^1).$$

Putting in the development for each matrix $\mathfrak{K}$, interchanging summations and remembering the orthogonality conditions we find

$$f=\sum_{p=1}^{2^n}\rho_p{}^m. \tag{13}$$

If we make the strip very long while keeping the width the same, m will become very large and all but the largest root ρ can be neglected. It is well known from the theory of matrices that the largest root has an eigenvector $a(\mu)$ with only positive components. This must be so, of course, because of their probability significance.

The doubly infinite square net results from (12) and (13) only in the limit when the width n of the strip tends to infinity as well as m. It has therefore been our endeavor to find a way of connecting n and $n+1$ just as our Eq. (12) carries out the reduction from $m+1$ to m. It has been impossible so far to do this in a rigorous manner. However, a very powerful approximation method can be based upon Eq. (12) which will be discussed in Part II, Section 7.

In spite of this failure some exact results can be obtained for the two-dimensional case; they are based on certain matrix identities. In order to obtain them we have to take up once more our basic steps which led to the eigenvalue problem (12). For it is possible to use instead of the matrix $\mathfrak{K}$ another equivalent one which does not have some of its inconveniences. The matrix $\mathfrak{K}$ is of order 2^n; it is filled solidly with elements which take up various values in a more or less haphazard way. In addition, its largest eigenvalue changes its meaning as n increases because it does not refer to one spin but to a whole line of them, i.e., we do not have $\rho^{nm}=f$, but only $\rho^m=f$. The reason for these features is that we built up the strip in large steps, filling a large number of vacant positions $[P]$ at one time (Fig. 2).

This situation can be amended partly by arranging the spins along the thread of a screw instead of a simple strip. For this purpose we dispose of the free right- and left-hand edges in Fig. 2 by bringing together each left end spin with the right end spin of the next line; the resulting order is

visualized in Fig. 3. Its advantage is that the whole net can now be built up through the simple operation of placing a new spin immediately beside its predecessor as we move along the thread of the screw. Figure 3 indicates this operation by a $[P]$ as in the two previous figures.

The mathematical build-up of the eigenvalue problem follows the two previous procedures. Boltzmann's theorem may be used to give the probabilities for all μ's with or without a spin being placed at $[P]$. As a second step we may eliminate, just as before, all explicit reference to the crossed spins by summation. If we denote by n the number of spins making up one pitch of the screw we can re-label the ring spins as indicated on Fig. 4. Let us call $A(\mu_{n-1}, \mu_{n-2}\cdots\mu_1, \mu_0)$ the probability referring to them alone and $P(\mu_n, \mu_{n-1}\cdots\mu_0)$ the one including μ_n in position $[P]$ as well. The factor by which the two quantities differ is now much simpler than it was in our previous treatment in §3. It contains only the two couplings which link up μ_n with μ_{n-1} and μ_0 in addition to the action of the field H on μ_n

$$\lambda\, P(\mu_n, \cdots\mu_0) = A(\mu_{n-1}, \cdots\mu_0)\, \exp[K\mu_n(\mu_{n-1}+\mu_0)+C\mu_n].$$

The eigenvalue problem follows from this equation if we sum $P(\mu_n, \cdots\mu_1, \mu_0)$ over μ_0 and notice that the resulting situation is identical with the one described by $A(\mu_{n-1}, \cdots\mu_0)$, provided the screw is very long. The only difference is that μ_1 now occupies the place of μ_0, μ_2 of μ_1, and so forth. It follows that we get the equation

$$\sum_{\mu_0} \exp[K\mu_n(\mu_{n-1}+\mu_0)+C\mu_n]\, A(\mu_{n-1}, \cdots\mu_0) = \lambda\, A(\mu_n, \cdots\mu_1). \tag{14}$$

If we compare (14) with (12) we notice that the matrix has become essentially asymmetric and that we have not achieved any reduction of its order. But we have realized two improvements: Each line contains now only two non-zero elements and the eigenvalue λ is now the partition function per individual spin regardless of the order of the matrix.

The proof of this latter proposition repeats the pattern outlined after Eq. (12), except that we have to take into account the asymmetry of the matrix. Let λ_i be its 2^n eigenvalues and $A_i(\mu_n\cdots\mu_1)$ the corresponding eigenvectors to the right. Then we have to consider in addition the eigenvectors to the left which we call $B_i(\mu_n, \cdots\mu_1)$ and which satisfy the equations

$$\sum_{\mu_n} \exp[K\mu_n(\mu_{n-1}+\mu_0)+C\mu_n]\, B(\mu_n, \cdots\mu_1) = \lambda\, B(\mu_{n-1}, \cdots\mu_0). \tag{15}$$

Before applying the basic equations it is advantageous to eliminate the appearance of identical μ's on either side of our Eqs. (14) and (15). This can be done by repeating n times the matrix operation indicated. If we discard the terms in C, i.e., let the magnetic field be zero for simplicity, we get

$$\sum_{\mu_1, \mu_2\cdots\mu_n} \exp[K\sum_{i=n}^{2n-1}\mu_i\mu_{i+1}+K\sum_{i=1}^{n}\mu_i\mu_{i+n}]A(\mu_n\cdots\mu_1) = \lambda^n A(\mu_{2n}\cdots\mu_{n+1})$$

and

$$\sum_{\mu_{n+1}, \mu_{n+2}, \cdots\mu_{2n}} \exp[K\sum_{i=n}^{2n-1}\mu_i\mu_{i+1}+K\sum_{i=1}^{n}\mu_i\mu_{i+n}]B(\mu_{2n}\cdots\mu_{n+1}) = \lambda^n B(\mu_n\cdots\mu_1).$$

It follows that with proper normalization

$$\sum_{\mu_i} B_p(\mu_n\cdots\mu_1)A_q(\mu_n\cdots\mu_1) = \delta_{pq}$$

and

$$\exp[K\sum_{i=n}^{2n-1}\mu_i\mu_{i+1}+K\sum_{i=1}^{n}\mu_i\mu_{i+n}] = \sum_{p=1}^{2^n}\lambda_p{}^n A_p(\mu_{2n}\cdots\mu_{n+1})B_p(\mu_n\cdots\mu_1).$$

We can write down this same formula for the next pitch of the screw

$$\exp[K\sum_{i=2n}^{3n-1}\mu_i\mu_{i+1}+K\sum_{i=n+1}^{2n}\mu_i\mu_{i+n}] = \sum_{p=1}^{2^n}\lambda_p{}^n A_p(\mu_{3n}\cdots\mu_{2n+1})B_p(\mu_{2n}\cdots\mu_{n+1}).$$

Now we multiply these two formulas with each other and sum over the spins of the middle row $\mu_{n+1}, \mu_{n+2}\cdots\mu_{2n}$. We get

$$\sum_{\mu_{n+1},\mu_{n+2}\cdots\mu_{2n}} \exp[K\sum_{i=n}^{3n-1}\mu_i\mu_{i+1}+K\sum_{i=1}^{2n}\mu_i\mu_{i+n}]=\sum_{p=1}^{2^n}\lambda_p{}^{2n}A_p(\mu_{3n}\cdots\mu_{2n+1})B_p(\mu_n\cdots\mu_1).$$

We can continue building up that way on the left the complete partition function of the problem. If we dispose of the ends of our screw by making it endless on a torus-shaped body, i.e., if we put

$$\mu_{mn+1}=\mu_1,$$

then we get finally

$$f=\sum_{p=1}^{2^n}\lambda_p{}^{mn}=\sum_{p=1}^{2^n}\lambda_p{}^{N}, \tag{16}$$

where $mn=N$ is the total number of spins present. If m is made sufficiently large while keeping n fixed, all but the largest eigenvalue λ may be neglected.

It has proved necessary in the study of the matrix (14) to avoid all unnecessary complications of the problem. One such complication is the dependence of λ on two parameters K and C, corresponding physically to the temperature and the magnetic field (Eqs. (3), (4)). It is an obvious simplification to study the square net at zero field only. Unfortunately the disadvantage of such a step is greater than is immediately obvious, for Eq. (5) shows us that it will be impossible under those circumstances to study even zero-field magnetization as a function of temperature. This reduces us henceforth entirely to the study of the energy and the specific heat. It is fortunate that these quantities show singularities concurrently with the magnetization at the Curie point, and their behavior as functions of T should be sufficient to discern the complete behavior of our model. From (6) and (16) we can thus deduce the following simplified expressions for the energy E and the molar specific heat C

$$E=-\tfrac{1}{2}NJ\,d\log\lambda/dK \tag{17}$$

and

$$C/R=K^2\,d^2\log\lambda/dK^2, \tag{18}$$

where R is the gas constant $R=Nk$.

If we make this assumption of zero magnetic field the parameter C in (14) is zero and the matrix operator is invariant with respect to inversion of all spins. It follows that the eigenvectors are either symmetric or antisymmetric with respect to that operation. Since the largest eigenvalue has only positive components it must belong to the symmetric class

$$A(\mu_n,\mu_{n-1}\cdots\mu_1)=A(-\mu_n,-\mu_{n-1},\cdots-\mu_1). \tag{19}$$

This follows also from the significance of the A's as probabilities. Equation (19) reduces the number of components of **A** from 2^n to 2^{n-1}.

It has proved to be convenient to write the matrix in (14) as a thing with lines and columns. This implies some ordering of the components of **A**. The following scheme has been adopted.

Take a given arrangement of μ's, for instance

$$+\ +\ -\ -\ +\ +\ -\ +,$$

replace all the plus signs by zeros and all the minus signs by ones, thus:

$$0\ 0\ 1\ 1\ 0\ 0\ 1\ 0;$$

read the number thus obtained as if 2 were our decimal unit and let it be the order number of the arrangement. The above arrangement for instance has the order number 50. As an example let us write down the components for the case $n=5$:

+	+	+	+	+	0		+	−	+	+	+	8
+	+	+	+	−	1		+	−	+	+	−	9
+	+	+	−	+	2		+	−	+	−	+	10
+	+	+	−	−	3		+	−	+	−	−	11
+	+	−	+	+	4		+	−	−	+	+	12
+	+	−	+	−	5		+	−	−	+	−	13
+	+	−	−	+	6		+	−	−	−	+	14
+	+	−	−	−	7		+	−	−	−	−	15

Because of the symmetry condition (19) the components 16 to 31 are identical with 15 to 0. We find generally

$$\text{component } (i)=\text{component } (2^n-i-1).$$

From here on it is a straightforward matter to bring the matrix (14) into its customary square. form. In the case $n=5$ we find, for example,

$$\mathfrak{M}(K)=\begin{vmatrix}
\alpha & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \alpha & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & \beta & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \beta & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \beta & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
\beta & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{vmatrix} \tag{20}$$

where $\alpha=e^{2K}$ and $\beta=e^{-2K}$. For other values of n the order of the matrix is a different power of two. Apart from that the matrix stays the same. The non-zero elements occupy a characteristic, turned-over V-shape. On the top arm e^{2K} and 1 are alternating, on the bottom e^{-2K} and 1. Altogether each line and each column has two non-zero elements.

Our statistical problem is now reduced to finding the largest eigenvalue $\lambda(K)$ of the matrix $\mathfrak{M}$:

$$\mathfrak{M}(K)\ \mathbf{A}(K)=\lambda(K)\ \mathbf{A}(K) \tag{21}$$[7]

and to studying the behavior of $\lambda(K)$ as the order of $\mathfrak{M}(K)$ increases indefinitely. For only in that limit does the solution correspond to the doubly infinite square net. The parameter K represents the temperature variable as defined in Eq. (3).

4. Properties of the Square Net Solution

It has been mentioned already that we have failed to find an exact solution in closed form for our problem. This does not mean, however, that no information concerning λ can be obtained. For some of the properties of the matrix $\mathfrak{M}$ have a bearing upon the properties of λ.

It is possible, for instance, to replace K by $-K$ in (20) through a simple re-ordering of terms. Using the case $n=5$ again as a test case this re-ordering matrix $\mathfrak{R}$ takes the following form

$$\mathfrak{R}=\begin{vmatrix}
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0
\end{vmatrix} \tag{22}$$

and its effect on $\mathfrak{M}(K)$ is given by

$$\mathfrak{R}\ \mathfrak{M}(K)\ \mathfrak{R}=\mathfrak{M}(-K). \tag{23}$$

It is clear that the transformation does not alter the eigenvalue spectrum and hence

$$\lambda(K)=\lambda(-K). \tag{24}$$

The substitution $K\rightarrow -K$ means replacing the ferromagnetic coupling J by an antiferromagnetic coupling of equal strength. It follows that such a coupling would also produce the thermal effects of ferromagnetism. This holds in particular for an eventual Curie point. Many approximate treatments destroy this basic symmetry.

It is easy to extend the definition (22) of $\mathfrak{R}$ to other orders. For if we compare that definition with the list of vector components on p. 259, we see that it simply changes the sign of every other spin. It should be mentioned in this connection that this invariance does not exist for $n=2,\ 4,\ 6\cdots$ because in those cases a completely antiferromagnetic pattern cannot be fitted into our screw arrangement of Fig. 3. For $n=1,\ 3,\ 5\cdots$ the difficulty does not arise and this partial sequence is sufficient to establish the invariance of λ for $n\rightarrow\infty$.

A more interesting result is obtained by operating on $\mathfrak{M}(K)$ with the unitary symmetric matrix $\mathfrak{T}$ which for $n=5$ has the following form:

$$\mathfrak{T}=\tfrac{1}{4}\begin{vmatrix}
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&-1&-1&-1&-1&-1&-1&-1&-1\\
1&1&1&1&-1&-1&-1&-1&-1&-1&-1&-1&1&1&1&1\\
1&1&1&1&-1&-1&-1&-1&1&1&1&1&-1&-1&-1&-1\\
1&1&-1&-1&-1&-1&1&1&1&1&-1&-1&-1&-1&1&1\\
1&1&-1&-1&-1&-1&1&1&-1&-1&1&1&1&1&-1&-1\\
1&1&-1&-1&1&1&-1&-1&-1&-1&1&1&-1&-1&1&1\\
1&1&-1&-1&1&1&-1&-1&1&1&-1&-1&1&1&-1&-1\\
1&-1&-1&1&1&-1&-1&1&1&-1&-1&1&1&-1&-1&1\\
1&-1&-1&1&1&-1&-1&1&-1&1&1&-1&-1&1&1&-1\\
1&-1&-1&1&-1&1&1&-1&1&-1&-1&1&1&-1&-1&1\\
1&-1&-1&1&-1&1&1&-1&-1&1&1&-1&-1&1&1&-1\\
1&-1&1&-1&1&-1&1&-1&1&-1&-1&1&-1&1&-1&1\\
1&-1&1&-1&-1&1&-1&1&-1&1&1&-1&1&-1&1&-1\\
1&-1&1&-1&-1&1&-1&1&1&-1&1&-1&-1&1&-1&1
\end{vmatrix} \tag{25}$$

[7] When using matrix notation in this and the following paper we shall use uniformly the following convention: matrices will be printed in German type, vectors in bold face Latin type.

For arbitrary values of n the matrix can be formed with the help of the following prescription: Suppose you have the matrix in the case $n-1$. Then you get the matrix for the case n by writing down twice each line of the previous matrix and then continuing on the right-hand side alternately in a symmetric and antisymmetric fashion. In addition, divide by $\sqrt{2}$ to preserve unitary character. The matrix of order 1 you start out with is the number 1.

If we transform $\mathfrak{M}(K)$ with this matrix $\mathfrak{T}$ we find that it goes over from its form (20) into a matrix having upright V-shape. Upon closer inspection we can write it as follows

$$\mathfrak{T}\ \mathfrak{M}(K)\ \mathfrak{T}=\sinh 2K\ \mathfrak{M}^{+}(K^*),$$

where the cross on $\mathfrak{M}$ stands for transposition, i.e., exchange of lines and columns and K^* is an auxiliary quantity defined through

$$e^{2K^*}=\coth K$$

or more symmetrically

$$\sinh 2K \sinh 2K^*=1 \tag{26a}$$

or also

$$\frac{\sinh 2K}{\cosh^2 2K}=\frac{\sinh 2K^*}{\cosh^2 2K^*}. \tag{26b}$$

The matrix relation can also be brought in a more symmetric form through the substitution

$$\mathfrak{V}(K)=\frac{1}{\cosh 2K}\mathfrak{M}(K), \tag{27}$$

which yields

$$\mathfrak{T}\ \mathfrak{V}(K)\ \mathfrak{T}=\mathfrak{V}^{+}(K^*) \tag{28}$$

a relation which, because of

$$\mathfrak{T}=\mathfrak{T}^{+},\quad \text{and}\quad \mathfrak{T}^2=1$$

is obviously reversible.

Equation (28) has an important bearing upon λ. Neither transformation with $\mathfrak{T}$ nor transposition of $\mathfrak{M}$ does change its eigenvalues and hence we find for λ

$$\lambda(K)/\cosh 2K=\lambda(K^*)/\cosh 2K^*.$$

We shall have occasion later to study this invariant quotient which we define as $x(K)$:

$$x(K)=\lambda(K)/\cosh 2K \tag{29}$$

and for which

$$x(K)=x(K^*). \tag{30}$$

If we interpret (26) with the help of (3) we see that it associates two temperatures with each other. As one of them rises from 0 to ∞ the other one drops from ∞ to 0. The significance of Eq. (30) is then that singular temperatures can only arise in pairs, since every singularity at K will be matched by one at K^*. The only exception to this rule is the temperature for which

$$\sinh 2K_c=1, \tag{31a}$$

since K_c is its own mate. The numerical value of K_c is found to be

$$K_c=0.44069. \tag{31b}$$

We conclude therefrom that if λ possesses one singularity only it must occur at the temperature given by (31). It is therefore the only possible location of the Curie point.

It is, of course, impossible to determine by such a symmetry argument the nature of the singularity, since we are not even certain of its existence. For the symmetry property in question is common to the whole sequence of matrices of order 1, 2, 4, 8, 16$\cdots$. All these finite matrices have solutions λ which are continuous throughout. But we can use our information in a negative way to exclude with certainty certain types of possible singularities. Using the definition (29) we find from (17) and (18)

$$E=-NJ\{\tanh 2K-\tfrac{1}{2}\,d\log x/dK\} \tag{32}$$

and

$$C/R=K^2\{4/\cosh^2 2K+d^2\log x/dK^2\}. \tag{33}$$

Now from (26) we can derive the following relations at the Curie point

$$(dK^*/dK)_{K=K_c}=-1,\quad (d^2K^*/dK^2)_{K=K_c}=2\sqrt{2}$$

and therefore from (30)

$$(dx/dK)_{K_c+0}+(dx/dK)_{K_c-0}=0 \tag{34}$$

and

$$\left(\frac{d^2x}{dK^2}\right)_{K_c+0}-\left(\frac{d^2x}{dK^2}\right)_{K_c-0}=-\sqrt{2}\left\{\left(\frac{dx}{dK}\right)_{K_c+0}-\left(\frac{dx}{dK}\right)_{K_c-0}\right\}. \tag{35}$$

Because of (32) Eq. (34) tells us that if the energy is continuous at the Curie point it must have the value

$$E(K_c) = -\tfrac{1}{2}\sqrt{2}NJ \tag{36}$$

which is a rather slow growth from the low temperature minimum $E=-NJ$ to the high temperature value $E=0$. In case of a phase transition (36) would at least represent the arithmetic mean of the values it has in the two phases.

Equation (35) tells us that *if the energy at K_c is continuous then the specific heat is also continuous unless it is infinite.* This infinity, if existing, would have to be of a rather symmetric nature since $C(K)-C(K^*)$ must tend to zero as we approach the Curie point. This result will be of great importance in Part II since it is in flat contradiction to most approximate solutions, which show a jump of the specific heat at the Curie point.

Both these results can be united into a single statement. If we study the sequence of solutions x_n for $n=1, 2, 3\cdots$ the specific heat as given by (33) must either tend to infinity at $K=K_c$ or else both energy and specific heat are continuous. The question thus formulated is specific enough to permit numerical treatment. It will occupy Section 6 in Part II and will give strong evidence that the specific heat is actually infinite at the Curie temperature.

We hope that the matrix method of solving statistical problems will be of use to other workers in this field. It may be mentioned in this connection that the treatment of the three-dimensional Ising model can also be reduced to the solution of a sequence of V-shaped matrices. It seems altogether as if a better understanding of such V-matrices might be helpful for statistics. One would expect this not to be too difficult in view of the small number of non-zero elements and their periodic structure. Their main drawback as compared to other simple matrices seems to be that the linear system to which they belong does not have the structure of a system of recursion relations.

In conclusion we want to express our thanks to Mr. E. Montroll for a helpful discussion.