

Statistics of the Two-Dimensional Ferromagnet. Part II

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

AND

G. H. WANNIER, *University of Texas, Austin, Texas*

(Received June 12, 1941)

The study of the two-dimensional Ising model is continued. Its specific heat at the Curie point is investigated. The quantity in question is computed for six successive finite matrix problems and the conclusion is drawn that the specific heat is infinite at the Curie point. A new closed form approximation of the partition function λ is then developed by using the matrix method in its variational form. The two power series for λ at extreme temperatures are used as a test for this and various other approximations, and it is found that the new result is a considerable improvement over the existing solutions. Finally it is pointed out that these closed form solutions support our conclusion as to the place and nature of the Curie point transition.

AFTER having collected in Part I⁸ some exact information concerning the partition function f of the two-dimensional Ising model we wish to present in this paper some approximate methods of our own and compare their results with the exact information available and other well-known approximate schemes.

Before proceeding let us recall the notation. We denote by λ the partition function per spin, i.e., for N spins we have

$$f = \lambda^N(K). \quad (37)$$

The parameter K is the only variable on which λ depends. It combines the coupling energy J and the temperature T in the form

$$K = J/2kT. \quad (38)$$

The knowledge of $\lambda(K)$ is not sufficient for the computation of the magnetic properties of the model, but it permits calculation of the thermal quantities, particularly the total energy E and the molar specific heat C [Eqs. (17) and (18)].

5. POWER SERIES DEVELOPMENTS OF λ

The energy of our system can be obtained by elementary reasoning in both the very high and very low temperature region.

At high temperatures (i.e., $K=0$) the spins orient themselves at random regardless of coupling forces. We conclude, therefore, by

direct inspection of Eq. (1) that

$$E(0) = 0.$$

By a similar reasoning we find for large K

$$E(\infty) = -NJ.$$

Equations (2) and (37) then permit the computation of λ in these two extreme cases. We find

$$\lambda(0) = 2, \quad (39)$$

$$\lambda(\infty) \approx e^{2K}. \quad (40)$$

Either one of these two limiting formulas can be continued by a power series. The continuation of (39) is the well-known development of λ in powers⁹ of $1/T$ which in our case means powers of K . If we carry out this development in Eq. (2) we get

$$\begin{aligned} f &= \sum_{\mu_i = \pm 1} [1 + K \sum_{\langle i,k \rangle} \mu_i \mu_k + \frac{1}{2} K^2 \langle \sum_{\langle i,k \rangle} \mu_i \mu_k \rangle^2 + \dots] \\ &= 2^N [1 + K \langle \sum_{\langle i,k \rangle} \mu_i \mu_k \rangle_{Av} + \frac{1}{2} K^2 \langle (\sum_{\langle i,k \rangle} \mu_i \mu_k)^2 \rangle_{Av} + \dots]. \end{aligned}$$

The averages are quite elementary to evaluate because they are to be taken at infinite temperature, that is, regardless of coupling. They are expressions containing various powers of N . But when we raise f to the power $1/N$ in accordance with (37) these powers disappear. Thus we find for λ

$$\lambda = 2 \left(1 + K^2 + \frac{4}{3} K^4 + \frac{77}{45} K^6 + \frac{1009}{315} K^8 + \dots \right). \quad (41)$$

⁸ H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252 (1941), this issue.

⁹ W. Opechowski, Physica **4**, 181 (1937).

TABLE I. Values of eigenvectors.

VECTOR A_2	VECTOR A_5	VECTOR A_6	
1.000000	1.000000	1.000000	.094527
.249038	.145548	.137739	.017936
	.100846	.098916	.009995
VECTOR A_3	.129103	.112698	.017578
	.098804	.097652	.010148
1.000000	.019608	.016676	.002709
.186931	.075422	.071758	.008510
.105677	.148388	.114363	.022371
.231108	.096230	.096454	.065628
	.020594	.016343	.017673
VECTOR A_4	.011008	.010661	.008320
	.023617	.016817	.022060
1.000000	.070174	.069010	.056461
.159434	.023648	.016516	.022385
.103346	.066754	.062185	.060217
.160827	.218677	.141720	.216286
.099340			
.026148			
.079123			
.222893			

It is also possible to find a power series valid for low temperatures.¹⁰ In the very lowest state all spins are oriented parallel to each other.⁶ The first excited state is given by the reversal of a single spin and requires an excitation energy $4J$. The next higher state corresponds to the reversal of two neighboring spins (excitation energy $6J$), the following to the reversal of two independent spins, and so forth. The development proceeds in inverse powers of

$$k = e^K = e^{J/2kT} \quad (42)$$

and yields, after extraction of the N th root,
 $\lambda = k^2(1 + k^{-8} + 2k^{-12} + 5k^{-16} + 14k^{-20} + \dots)$. (43)

Equations (41) and (43) must be different aspects of one and the same power series, in view of the symmetry property expressed through (26), (29) and (30). This is, in fact, true and can be made explicit through the introduction of the following parameter

$$\kappa = \frac{\sinh 2K}{2 \cosh^2 2K}. \quad (44a)$$

Its invariance property

$$\kappa(K) = \kappa(K^*) \quad (44b)$$

has already been pointed out in (26b). It vanishes at either very high or very low temperatures, and behaves as K in the former and as k^{-2} in the latter region. Its maximum value at $K = K_c$ is

¹⁰ F. Bloch, *Zeits. f. Physik* **61**, 206 (1930).

equal to $\frac{1}{4}$. With the help of this parameter we may unite the two series into

$$x(K) = \lambda(K) / \cosh 2K = 2(1 - \kappa^2 - 4\kappa^4 - 29\kappa^6 - 265\kappa^8 - 2745\kappa^{10} - \dots). \quad (45)$$

For our problem these series solutions are of no direct interest, for they seem to diverge in the critical region. They are, however, a very agreeable criterion to test the accuracy of approximate solutions by comparing their series development with (41) and (43).

6. THE SPECIFIC HEAT AT THE CURIE POINT

If our model leads to one singular temperature only (which is the standard idea associated with the Curie temperature) then the Eqs. (26), (30), and (31) permit us to limit our attention to the point $K = 0.44069$. In addition, the possibilities to be expected at that point are essentially twofold only: Either the specific heat tends to infinity as we approach the infinite problem over a sequence of finite matrix problems, or else it stays finite, in which case both energy and specific heat must be continuous [Eqs. (34) and (35)].

We have applied this test to the sequence of V matrices (20) of order 2, 4, 8, 16, 32. This corresponds physically to the arrangement of spins along the thread of a screw with 2, 3, 4, 5, 6 spins per pitch. In addition, we added the solution λ_1 having one spin per pitch as the first member. This case is equivalent to a linear chain with doubly strong coupling.

There is no difficulty of principle in computing these λ 's as functions of K , and in particular their second derivative at $K = K_c$. In practice, however, it is found that the secular equation derived from (20) becomes rapidly unmanageable as the order of the matrix increases. Another, more elegant, method has therefore been followed, which uses a good number of special properties of such V matrices.

In the first place, it is possible to obtain the largest eigenvalue λ for a given value of K by operating repeatedly with the matrix (20) on an arbitrary vector (iteration method).¹¹ After a

¹¹ Frazer, Duncan and Collar, *Elementary Matrices and some Applications to Dynamics and Differential Equations* (Cambridge University Press and the Macmillan Company, 1938), p. 138.

It is immediately generalized to other orders with the help of the list of components in I, p. 259. The matrix inverts the sequence of signs, that is $+++-+$ goes over into $+-+++$. In addition, it multiplies with a factor $\exp(-K\sum_i\mu_i\mu_{i+1})$. \mathfrak{B} , \mathfrak{I} and \mathfrak{S} are related by a fundamental identity. Using (50) and (58) we get

$$\mathfrak{I}\mathfrak{S}(K^*)\mathfrak{I}\mathfrak{S}(K)\mathfrak{B}(K) = \mathfrak{B}(K)\mathfrak{I}\mathfrak{S}(K^*)\mathfrak{I}\mathfrak{S}(K),$$

i.e., the expression

$$\mathfrak{I}\mathfrak{S}(K^*)\mathfrak{I}\mathfrak{S}(K)$$

commutes with \mathfrak{B} and hence, assuming no multiple roots, is expressible in terms of it. We find, by direct inspection,

$$\mathfrak{I}\mathfrak{S}(K^*)\mathfrak{I}\mathfrak{S}(K) = \{2[\kappa(K)]^{\frac{1}{2}}\mathfrak{B}^{-1}(K)\}^{n-1}, \quad (61)$$

where n is, as usual, the number of spins per pitch, that is, 2^{n-1} is the matrix order. Because of (59), already $\mathfrak{I}\mathfrak{S}(K)$ commutes with \mathfrak{B} at $K=K_c$ and must equal some square root of (61). The sign varies for different eigenvectors and is found to be positive for our vector $\mathbf{A}(K_c)$. Remembering that $\kappa(K_c) = \frac{1}{4}$ and (46) we get

$$\mathfrak{I}\mathfrak{S}(K_c)\mathbf{A}(K_c) = (1/x(K_c))^{(n-1)/2}\mathbf{A}(K_c). \quad (62)$$

With the help of $\mathfrak{I}\mathfrak{S}(K_c)$ we define a new matrix \mathfrak{B} which shall be a function of \mathfrak{B} not depending directly on K

$$\mathfrak{B} = \alpha + \beta\mathfrak{B} + \gamma\mathfrak{B}^2 + \delta\mathfrak{B}^3 + \dots \quad (63)$$

and for which

$$\mathfrak{B}(K_c) = \mathfrak{I}\mathfrak{S}(K_c). \quad (64)$$

Because of (50) and (54), the derivative of \mathfrak{B} obeys the equation

$$\mathfrak{I}\mathfrak{B}'(K_c) + \mathfrak{B}'(K_c)\mathfrak{I} = 0,$$

which, considering (51), means that

$$\mathfrak{A} = \mathfrak{I}\mathfrak{B}'(K_c) \quad (65)$$

$$\mathfrak{A} = \begin{pmatrix} 0 & -2^{\frac{1}{2}}k^{-2} & -2k^{-2} & -2^{\frac{1}{2}}k^{-2} & -2k^{-2} & 0 & -2k^{-2} & -2^{\frac{1}{2}}k^{-2} & -2k^{-2} & 0 & 0 & 0 & -2k^{-2} & 0 & -2k^{-2} & -2^{\frac{1}{2}}k^{-2} \\ 2^{\frac{1}{2}}k^{-2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2^{\frac{1}{2}}k^{-2} & -2 & 0 & -2 & -2^{\frac{1}{2}}k^{-2} & -2 & -2^{\frac{1}{2}}k^{-2} & 0 \\ 2k^{-2} & 0 & 0 & 0 & 0 & 0 & 0 & 2^{\frac{1}{2}}k^{-2} & 0 & -2^{\frac{1}{2}}k^2 & -2k^2 & -2^{\frac{1}{2}}k^2 & 0 & 0 & 0 & -2^{\frac{1}{2}}k^{-2} \\ 2^{\frac{1}{2}}k^{-2} & 0 & 0 & 0 & -2^{\frac{1}{2}}k^{-2} & -2 & -2^{\frac{1}{2}}k^{-2} & 0 & 2^{\frac{1}{2}}k^{-2} & 2 & 0 & 2 & 0 & 0 & 0 & 0 \\ 2k^{-2} & 0 & 0 & 2^{\frac{1}{2}}k^{-2} & 0 & -2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^{-2} & 0 & 0 & -2k^2 & -2^{\frac{1}{2}}k^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2^{\frac{1}{2}}k^2 & 0 & 0 & -2 & 2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^2 & 0 & 0 & 2 \\ 2k^{-2} & 0 & 0 & 2^{\frac{1}{2}}k^{-2} & 0 & 0 & 0 & 0 & 0 & 0 & 2k^2 & 2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^{-2} \\ 2^{\frac{1}{2}}k^{-2} & 0 & -2^{\frac{1}{2}}k^{-2} & 0 & 2^{\frac{1}{2}}k^{-2} & 2 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 2^{\frac{1}{2}}k^{-2} & 2 & 0 & 0 \\ 2k^{-2} & 2^{\frac{1}{2}}k^{-2} & 0 & -2^{\frac{1}{2}}k^{-2} & 0 & -2^{\frac{1}{2}}k^2 & 0 & 0 & 0 & 0 & -2k^2 & 0 & 0 & -2^{\frac{1}{2}}k^2 & 0 & 0 & 0 \\ 0 & 2 & 2^{\frac{1}{2}}k^2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & -2^{\frac{1}{2}}k^2 & 0 & 2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^2 & 2 & 0 \\ 0 & 0 & 2k^2 & 0 & 2k^2 & 2^{\frac{1}{2}}k^2 & -2k^2 & 0 & 2k^2 & 2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^2 & -2k^2 & -2^{\frac{1}{2}}k^2 & 2k^2 & 2 & 0 \\ 0 & 2 & 2^{\frac{1}{2}}k^2 & -2 & 2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^2 & 2 & 0 & 0 & 2^{\frac{1}{2}}k^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2k^{-2} & 2^{\frac{1}{2}}k^{-2} & 0 & 0 & 0 & 2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^{-2} & 0 & -2^{\frac{1}{2}}k^2 & 2k^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 2^{\frac{1}{2}}k^2 & -2 & 2^{\frac{1}{2}}k^2 & 0 & 2^{\frac{1}{2}}k^2 & 0 & 0 & 0 & -2^{\frac{1}{2}}k^2 & 2 & 0 \\ 2k^{-2} & 2^{\frac{1}{2}}k^{-2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2^{\frac{1}{2}}k^2 & -2k^2 & 0 & 0 & 2^{\frac{1}{2}}k^2 & 0 & -2^{\frac{1}{2}}k^{-2} & 0 \\ 2^{\frac{1}{2}}k^{-2} & 0 & 2^{\frac{1}{2}}k^{-2} & 0 & 0 & -2 & 2^{\frac{1}{2}}k^{-2} & 0 & 0 & -2 & 0 & 0 & 0 & -2 & 2^{\frac{1}{2}}k^{-2} & 0 & 0 \end{pmatrix} \quad (69)$$

is an antisymmetric matrix. Because of (63), \mathfrak{B} admits \mathbf{A} as an eigenvector; if we differentiate this relation as well as (59) and remember (64) we arrive at

$$\begin{aligned} \mathfrak{B}'(K_c)\mathbf{A}(K_c) + \mathfrak{I}\mathfrak{S}(K_c)\mathbf{A}'(K_c) &= (1/x(K_c))^{(n-1)/2}\mathbf{A}'(K_c), \\ \mathfrak{I}\mathfrak{S}'(K_c)\mathbf{A}(K_c) + \mathfrak{I}\mathfrak{S}(K_c)\mathbf{A}'(K_c) &= -(1/x(K_c))^{(n-1)/2}\mathbf{A}'(K_c). \end{aligned}$$

Subtracting and simplifying with (51) and (65) we get

$$\mathbf{A}'(K_c) = \frac{1}{2}(x(K_c))^{(n-1)/2}\mathfrak{I}\{\mathfrak{A} - \mathfrak{S}'(K_c)\}\mathbf{A}(K_c). \quad (66)$$

Uniting finally (57) and (66), simplifying with (52) and (54) and leaving off the now superfluous argument K_c we get the final formula

$$x'' = \frac{1}{\mathbf{B}+\mathbf{A}}[\mathbf{B}+\mathfrak{B}''\mathbf{A} + x^{(n-1)/2}\mathbf{A}+\mathfrak{B}'+\{\mathfrak{S}'-\mathfrak{A}\}\mathbf{A}]. \quad (67)$$

Expression (67) for x'' has the desired form. The eigenvalue x and the eigenvectors \mathbf{A} and \mathbf{B} are numerically given on page 264 and through (52); the matrices involved are obtainable by differentiation from (20), (27), and (60). The matrix \mathfrak{A} is the only exception, being given through implicit equations. It is easy, however, to obtain for it a number of equations far in excess of the number of its components, particularly since \mathfrak{A} is restricted to be antisymmetric. For we have, for any power p of \mathfrak{B} , because of (63)

$$\mathfrak{B}\mathfrak{B}^p - \mathfrak{B}^p\mathfrak{B} = 0,$$

which, because of (50), (54), (64), (65), gives

$$\mathfrak{B}^{p+1}\mathfrak{A} - \mathfrak{A}\mathfrak{B}^p = \mathfrak{S}(\mathfrak{B}^p)' + (\mathfrak{B}^p)'+\mathfrak{S}. \quad (68)$$

In the case $n=5$, the matrix obtained for \mathfrak{A} looks as follows

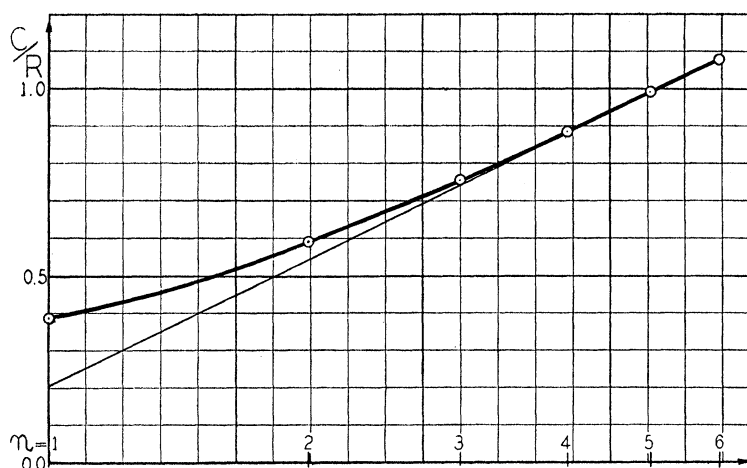


FIG. 5. Behavior of specific heat at T_c in successive "screw" approximations. The logarithm of the number of spins making up one "pitch" of the screw is plotted on abscissa.

It has, unfortunately, not been possible to find a general rule of formation for the matrix for all values of n . But the number of regularities is sufficiently large to permit guessing of the elements with the help of some of the equations (68). Thus our purpose of computing $x''(K_c)$ without the knowledge of the secular equation is achieved. The latter method was used as a cross check for the smaller matrices.

We find, in this fashion, the values for $x(K_c)$ and $x''(K_c)$ as functions of n given in Table II. Because of (56), these numbers are the only variables to enter into the specific heat as given by (33):

$$S = \frac{C}{R} = K_c^2 \left(2 + \frac{x''}{x} \right). \quad (70)$$

The fundamental question of this section is: does the quantity S tend to infinity or not with increasing n ? Since this involves extrapolating to the limit of an infinite sequence from its first six members the answer is somewhat a matter of judgment. This is particularly true in this case because the growth of S is decelerating with n . A careful analysis of the data, however, still indicates strongly that C tends to infinity. The conclusion is reached by comparing S with $\log n$.

TABLE II. Values of x and x'' .

n	x	x''	n	x	x''
1	2.00000	0.00000	4	1.81984	4.66665
2	1.88320	1.97715	5	1.81002	5.62246
3	1.83929	3.48493	6	1.80450	6.41935

The resulting curve is plotted on Fig. 5. It indicates that the mutual relation approaches linearity. This is borne out even more exactly by the analysis shown in Table III. The first two columns contain the argument n and the function S , then follows the difference quotient dividing the increment of S by the increment of $\log n$. That this quotient tends to a constant is shown in the next two columns. The first shows that the differences between successive quotients decrease rapidly, and the last, by forming the ratios between those differences, shows that the decrease is exponential and therefore sufficient to guarantee convergence of the difference quotients to a constant value.

We conclude therefrom that the specific heat of the two-dimensional square net of spins is infinite at $K_c = 0.4407$, insofar as numerical extrapolation methods are able to decide such a question. Whether this means a latent heat or not is beyond the scope of the method.

TABLE III.

n	$S = C/R$	$\frac{\Delta S}{\Delta \log n}$	$\Delta \frac{\Delta S}{\Delta \log n}$	RATIOS OF Δ 'S
1	0.38841	0.2942		
2	0.59230	0.4046	0.1105	0.43
3	0.75637	0.4520	0.0474	0.42
4	0.88641	0.4717	0.0197	0.45
5	0.99167	0.4805	0.0088	
6	1.07928			

7. A NEW APPROXIMATION OF λ FOR ALL TEMPERATURES

The calculations of the preceding section indicate that we are not able, at present, to dispense with approximate evaluations of λ . But even in this field better results can be obtained with the eigenvalue method. We depart in this case from the eigenvalue problem (12) referring to an infinite strip (Fig. 2). For simplicity we may assume that the free ends of each strip are linked up with each other. This will make it an infinite cylinder in which each level circle consists of n equivalent spins. Our unknown λ is then the n th root of the eigenvalue ρ .

The largest eigenvalue of a matrix can always be expressed in a variational form. In our case we find from (12)

$$\lambda = \text{Max}_{\text{all } a\text{'s}} \frac{\sum_{\mu_i \mu_{i'}} \mathfrak{H}(\mu_i, \mu_{i'}) a(\mu_i) a(\mu_{i'})}{\sum_{\mu_i} [a(\mu_i)]^2} \quad (71)$$

Expression (71) obviously suggests a variation method in which a restricted set of competing functions are used for the a 's. In the first place, the correct a 's themselves possess the symmetries of the nucleus \mathfrak{H} as may be verified from Section 3. In the second place, we may restrict ourselves to combinations of the μ 's which are linear in each of them, since $\mu_i^2 = 1$. Let us note down a few

of them

$$\begin{aligned} &\mu_1 + \mu_2 + \mu_3 + \mu_4 + \dots \\ &\mu_1 \mu_2 + \mu_2 \mu_3 + \mu_3 \mu_4 + \dots \\ &\mu_1 \mu_3 + \mu_2 \mu_4 + \mu_3 \mu_5 + \dots \\ &\mu_1 \mu_2 \mu_3 + \mu_2 \mu_3 \mu_4 + \dots \\ &\mu_1 \mu_4 + \dots \end{aligned}$$

Among them, the first two are outstanding because \mathfrak{H} contains them already explicitly. Let us, therefore, put

$$\sum_{i=1}^n \mu_i = nm, \quad (72)$$

$$\sum_{i=1}^n \mu_i \mu_{i+1} = nq \quad (73)$$

and evaluate the maximum of (71) under the restriction that the a 's be of the form

$$a(\mu_i) = a(q, m). \quad (74)$$

We shall prove now that there is a form of a equivalent to (74), which has, however, great practical advantages. This form is

$$a(\mu_i) = \exp n \{ H(K, C)q + A(K, C)m \}, \quad (75)$$

where H and A are constants as far as q and m are concerned, but depend upon temperature and field, i.e., K and C as parameters. This dependence has to be determined by the maximum condition (71).

The first step of the proof is to introduce explicitly the assumption (74) into (71). For this purpose we denote by

$$\exp ng(q, m) dq dm, \quad (76)$$

the number of arrangements of the μ 's lying in a rectangle of sides dq and dm around q and m . Let us also introduce the notation

$$\sum_{\substack{\mu_i, \mu_{i'} \\ q, m, q', m'}} \exp [nK \sum_i \mu_i \mu_{i'}] = \exp [nf(q, n, q', m')] dq dm dq' dm', \quad (77)$$

where the summation on the left is extended only over those combinations that lie within a volume $dq dm dq' dm'$ around the point q, m, q', m' . Then (71) takes the form

$$\lambda^n = \text{Max} \frac{\iiint \exp n \{ f(q, m, q', m') + \frac{1}{2}K(q+q') + \frac{1}{2}C(m+m') \} a(q, m) a(q', m') dq dm dq' dm'}{\iint \exp ng(q, m) a^2(q, m) dq dm} \quad (78)$$

Now transforming

$$a(q, m) = b(q, m) \exp [-\frac{1}{2}ng(q, m)],$$

we write

$$\lambda^n = \text{Max} \frac{\iiint \exp n \{ f(q, m, q', m') - \frac{1}{2}g(q, m) - \frac{1}{2}g(q', m') + \frac{1}{2}K(q+q') + \frac{1}{2}C(m+m') \} b(q, m) b(q', m') dq dm dq' dm'}{\iint b^2(q, m) dq dm}$$

We can study this maximum problem in two parts; first we may assume the denominator to be normalized to unity and then we may look for the maximum of the numerator. If we do that we see that we shall get λ^n as large as possible if we let b be large wherever the exponent reaches its maximum. If we evaluate then the integral with the saddle point method, $b(q, m)$ will make no contribution to it because of its normalization and we get

$$\log \lambda = \text{Max}_{q, m, q', m'} [f(q, m, q', m') - \frac{1}{2}g(q, m) - \frac{1}{2}g(q', m') + \frac{1}{2}K(q+q') + \frac{1}{2}C(m+m')]. \tag{79}$$

We can eliminate the unknown functions f and g with the help of simpler ones. We introduce the solution χ^n of a two-strip problem

$$\chi^n(I, B, I', B') = \sum_{\mu_i, \mu'_i} \exp[K \sum_{i=1}^n \mu_i \mu'_i + I \sum_{i=1}^n \mu_i \mu_{i+1} + I' \sum_{i=1}^n \mu'_i \mu'_{i+1} + B \sum_{i=1}^n \mu_i + B' \sum_{i=1}^n \mu'_i] \tag{80}$$

or because of (72), (73), and (77)

$$\chi^n = \int \int \int \int \exp n \{f(q, m, q', m') + Iq + I'q' + Bm + B'm'\} dq dm dq' dm'.$$

For large n we can evaluate χ by the saddle point method

$$\log \chi(I, B, I', B') = \text{Max}_{q, m, q', m'} [f(q, m, q', m') + Iq + I'q' + Bm + B'm']. \tag{81}$$

Similarly we introduce two one-strip solutions

$$\psi^n(H, A) = \sum_{\mu_i} \exp[2H \sum_{i=1}^n \mu_i \mu_{i+1} + 2A \sum_{i=1}^n \mu_i], \tag{82}$$

which can be written as $\log \psi(H, A) = \text{Max}_{q, m} [g(q, m) + 2Hq + 2Am]$ (83a)

and similarly $\log \psi(H', A') = \text{Max}_{q', m'} [g(q', m') + 2H'q' + 2A'm']$. (83b)

It looks incorrect at first sight to substitute (81) and (83) into (79) because the former are already maximized with respect to q, m, q', m' . However, the parameters $I, B, I', B', H, A, H', A'$ are arbitrary, and, as they vary, the quantities q, m, q', m' cover their full range of variability. We may therefore carry out the substitution and write

$$\log \lambda = \text{Max}_{q, m, q', m'} [\log \chi(I, B, I', B') - \frac{1}{2} \log \psi(H, A) - \frac{1}{2} \log \psi(H', A') + (\frac{1}{2}K - I + H)q + (\frac{1}{2}K - I' + H')q' + (\frac{1}{2}C - B + A)m + (\frac{1}{2}C - B' + A')m']. \tag{84}$$

In studying (84) we have to distinguish three types of relations between the quantities involved:

(a) The extremal conditions on χ and ψ indicated in (81) and (83). We shall use these conditions only in an indirect form. They are sufficient in number to make the expressions $I, B \dots A'$ functions of q, m, q', m' , or vice versa. Another consequence is that first derivatives of χ and ψ may be formed as if q, m, q', m' , appearing explicitly in (81) and (83), were constants.

(b) The relations expressing the identity of q, m, q', m' in the various parts. They read

$$\frac{\partial \log \chi}{\partial I} = \frac{1}{2} \frac{\partial \log \psi}{\partial H} = q, \quad \frac{\partial \log \chi}{\partial I'} = \frac{1}{2} \frac{\partial \log \psi'}{\partial H'} = q', \quad \frac{\partial \log \chi}{\partial B} = \frac{1}{2} \frac{\partial \log \psi}{\partial A} = m, \quad \frac{\partial \log \chi}{\partial B'} = \frac{1}{2} \frac{\partial \log \psi'}{\partial A'} = m'.$$

(c) Finally, the extremal conditions on λ

$$\frac{\partial \log \lambda}{\partial q} = \frac{\partial \log \chi}{\partial I} \frac{\partial I}{\partial q} + \frac{\partial \log \chi}{\partial B} \frac{\partial B}{\partial q} + \frac{\partial \log \chi}{\partial I'} \frac{\partial I'}{\partial q} + \frac{\partial \log \chi}{\partial B'} \frac{\partial B'}{\partial q} - \frac{1}{2} \frac{\partial \log \psi}{\partial H} \frac{\partial H}{\partial q} - \frac{1}{2} \frac{\partial \log \psi}{\partial A} \frac{\partial A}{\partial q} - q \left(\frac{\partial I}{\partial q} - \frac{\partial H}{\partial q} \right) - q' \frac{\partial I'}{\partial q} - m \left(\frac{\partial B}{\partial q} - \frac{\partial A}{\partial q} \right) - m' \frac{\partial B'}{\partial q} + \frac{1}{2}K - I + H = 0$$

and three similar equations. The relations under (b) simplify them to

$$\begin{aligned} I &= \frac{1}{2}K + H, & I' &= \frac{1}{2}K + H', \\ B &= \frac{1}{2}C + A, & B' &= \frac{1}{2}K + A'. \end{aligned}$$

If we eliminate I, I', B, B' with the help of these four equations, (84) takes the form

$$\lambda = \frac{\chi(\frac{1}{2}K + H, \frac{1}{2}C + A, \frac{1}{2}K + H', \frac{1}{2}C + A')}{[\psi(H, A)\psi(H', A')]^{\frac{1}{2}}}$$

and the equations under (b) read

$$\frac{\partial \log \lambda}{\partial H} = \frac{\partial \log \lambda}{\partial H'} = \frac{\partial \log \lambda}{\partial A} = \frac{\partial \log \lambda}{\partial A'} = 0,$$

i.e., the four parameters H, H', A, A' must make λ a stationary value. Should there be several sets of solutions leading to several possible sets of q, m, q', m' , then the original maximum condition would demand the largest one. This gives, then,

$$\lambda = \text{Max}_{H, A, H', A'} \frac{\chi(\frac{1}{2}K + H, \frac{1}{2}C + A, \frac{1}{2}K + H', \frac{1}{2}C + A')}{[\psi(H, A)\psi(H', A')]^{\frac{1}{2}}}. \quad (85)$$

Our last step is to prove that

$$H = H', \quad A = A'.$$

This is done by substituting (80) and (82) into (85). Remembering the definition (12) of \mathcal{H} we get it in the form

$$\lambda^n = \text{Max}_{c, d} \frac{\sum_{\mu_i \mu_i'} \mathcal{H}(\mu_i, \mu_i') c(\mu_i) d(\mu_i')}{\{[\sum_{\mu_i} c^2(\mu_i)][\sum_{\mu_i'} d^2(\mu_i')]\}^{\frac{1}{2}}}, \quad (86)$$

where the competing c 's and d 's are a restricted class of functions obeying (75). This maximum problem does not seem to agree in form with our basic equation (71). The greater generality is only apparent, however, for we can prove the following theorem:

If we have a set of c 's and d 's defining λ as in (86) a larger λ can always be obtained either by replacing the c 's by the d 's or else the d 's by the c 's. We prove, first, the lemma that

$$\sum_{\mu_i, \mu_i'} \mathcal{H}(\mu_i, \mu_i') a(\mu_i) a(\mu_i')$$

is a positive definite form. We proceed by induction, starting from the fact that the statement is manifestly true if the a 's depend only on one spin μ . Now suppose it to be true for $n-1$ spins then we can prove it for n . The nucleus \mathcal{H} is of the form

$$\mathcal{H}(\mu_i, \mu_i') = \phi(\mu_i) \phi(\mu_i') \exp K \sum_{i=1}^n \mu_i \mu_i'.$$

Absorbing the ϕ 's into the a 's, which are arbitrary, we must only prove that

$$F = \sum_{\mu_i, \mu_i'} \exp(K \sum_{i=1}^n \mu_i \mu_i') a(\mu_i) a(\mu_i')$$

is a positive definite form. Now we single out μ_n and μ_n' and denote by $\sum'_{\mu_i, \mu_i'}$ summation over the remaining μ 's. Since μ_n is only capable of two values, the dependence of a on μ_n can be made explicit by writing

$$a(\mu_i) = \alpha(\mu_i) + \mu_n \beta(\mu_i),$$

where α and β are independent of μ_n . We can thus carry out the summation over μ_n and μ_n' and get F in the form

$$F = 4 \cosh 2K \sum_{\mu_i, \mu_i'} \exp(K \sum_{i=1}^{n-1} \mu_i \mu_i') \alpha(\mu_i) \alpha(\mu_i') + 4 \sinh 2K \sum_{\mu_i, \mu_i'} \exp(K \sum_{i=1}^{n-1} \mu_i \mu_i') \beta(\mu_i) \beta(\mu_i').$$

This proves our lemma, for the two right-hand sides are, by assumption, positive. To prove the theorem itself we apply the Schwartz inequality to (86); assuming the c 's and d 's normalized we get

$$2 \sum_{\mu_i, \mu_i'} \mathfrak{R}(\mu_i, \mu_i') c(\mu_i) d(\mu_i') \leq \sum_{\mu_i, \mu_i'} \mathfrak{R}(\mu_i, \mu_i') c(\mu_i) c(\mu_i') + \sum_{\mu_i, \mu_i'} \mathfrak{R}(\mu_i, \mu_i') d(\mu_i) d(\mu_i'),$$

which proves the theorem. It transforms (85) into the final form

$$\lambda = \text{Max}_{H, A} \frac{\chi(\frac{1}{2}K + H, \frac{1}{2}C + A)}{\psi(H, A)} \tag{87}$$

Because of (12), (80), and (82), it is that particular case of (71) in which the a 's have the form (75). Hence (75) is a consequence of (74). The approximation could be pushed further, in principle, by making a depend on more than just two parameters among the ones listed on p. 268. One verifies easily that such an expression could still be transformed to the form (75). Other parameters beside H and A would then have to be added to (87).

To evaluate (87) we observe that χ and ψ can be easily obtained by the methods of Sections 2 and 3. The explicit form of ψ is actually given in (9) and χ is obtained from a fourth-order matrix problem of the type (12). Using the definitions

$$e^K = k, \quad e^C = c, \quad e^{2H} = h, \quad e^{2A} = a, \tag{88}$$

we find for ψ

$$\psi = \frac{1}{2}h \left(a + \frac{1}{a} \right) + \frac{1}{2} \left[h^2 \left(a - \frac{1}{a} \right)^2 + \frac{4}{h^2} \right]^{\frac{1}{2}} \tag{89a}$$

and χ

$$\chi^3 - \chi^2 \left\{ k^2 h \left(ac + \frac{1}{ac} \right) + \frac{1}{k} \left(kh + \frac{1}{kh} \right) \right\} + \chi \left(kh - \frac{1}{kh} \right) \left\{ kh \left(ac + \frac{1}{ac} \right) + k^2 \left(kh + \frac{1}{kh} \right) \right\} - k \left(kh - \frac{1}{kh} \right)^3 = 0. \tag{89b}$$

The computation was only carried through for zero field, i.e., $c=0$. In this case $a=1$ is a possible solution since both $\partial\psi/\partial a$ and $\partial\chi/\partial a$ vanish at that point. The equation in χ then factorizes into

$$\left[\chi^2 - \chi \left(k + \frac{1}{k} \right) \left(kh + \frac{1}{hk} \right) + \left(kh - \frac{1}{kh} \right)^2 \right] \left[\chi - k \left(kh - \frac{1}{kh} \right) \right] = 0$$

and the last solution may be discarded. The equation may be considered an equation in λ . The substitution

$$\frac{kh - (1/kh)}{h + (1/h)} = s\lambda$$

brings it in the form

$$\lambda = \text{Max}_s \frac{2}{(s - \sinh K)^2 + 1 - \sinh^2 K},$$

which is immediately maximized to

$$\lambda = \frac{2}{1 - \sinh^2 K}. \tag{90}$$

It is obvious that (90) cannot be valid for high values of K . In that region another solution having $a \neq 1$ becomes the correct one (nonvanishing "inner field"). The substitutions

$$\left(kh - \frac{1}{kh}\right)_{\psi}^{-1} = s\lambda, \quad \left(h + \frac{1}{h}\right)_{\psi}^{-1} = y$$

transform (87) into

$$\lambda = \text{Max}_{s, y} \frac{k\left(k + \frac{1}{k}\right)(k-s) + \frac{2}{k}(1-k^3s)y - k\left(k - \frac{1}{k}\right)(k-s)y^2}{k + \frac{1}{k} - \frac{s}{k}\left(k - \frac{1}{k}\right) + s^2k^2\left(k - \frac{1}{k}\right) - s^3k\left(k + \frac{1}{k}\right) - 2ks(k-s)y}$$

Forming $\partial\lambda/\partial y = 0$ and eliminating y we find

$$(\lambda - k^2)^2 s^2 (k-s)^2 - (\lambda - k^2)(k-s) \left(\frac{1}{k^2} - s^2\right) \left[k\left(k^2 - \frac{1}{k^2}\right) - s\left(k^2 + \frac{1}{k^2}\right) \right] + \left(\frac{1}{k^2} - s^2\right)^2 = 0.$$

After restoring linearity by the substitution

$$(1/k^2) - s^2/s(k-s) = (\lambda - k^2)/z,$$

we may carry out the second differentiation $\partial\lambda/\partial z = 0$. We get

$$(1 + (2/k^2)z + z^2)^2 = [k^2 - (1/k^2)]^2 z \quad (91a)$$

and

$$\lambda = k^2 + \frac{(1 + k^2 z)(1 - z^2)}{k^4 [k^2 - (1/k^2)]}. \quad (91b)$$

The two pieces (90) and (91) make up the complete solution. They join with a common tangent at the point

$$K_c = \frac{1}{2} \log \frac{5 + 17^{1/2}}{4} = 0.4122, \quad (92)$$

which means a transition without latent heat, but with a discontinuity in the specific heat curve. Other features of the solution will be discussed in the next section.

8. COMPARISON WITH OTHER TREATMENTS

Several approximate methods have been developed in the past to arrive at the partition function of a set of regularly arranged spins. The first evaluation of λ is due to Heisenberg in his original paper on ferromagnetism.¹² His method has been systematized recently by Kirkwood¹³ who developed a scheme of successive

approximations of which Heisenberg's is the first step. It is based on the following consideration. Expression (2) for λ depends on two combinations involving the μ 's, namely

$$\sum_{\langle i, k \rangle} \mu_i \mu_k = 2Nq, \quad \sum_i \mu_i = Nm.$$

If we denote the density of arrangements in $q-m$ space by $D^N(q, m)$ we have

$$\lambda^N = \int \int dq dm \exp N \{ \log D(q, m) + 2Kq + Cm \}. \quad (93)$$

Now $D^N(q, m)$ is unknown, but the expression

¹²W. Heisenberg, *Zeits. f. Physik* **49**, 619 (1928). See also: R. H. Fowler, *Statistical Mechanics* (Cambridge University Press and Macmillan Company, 1936), second edition, p. 485.

¹³John G. Kirkwood, *J. Chem. Phys.* **6**, 70 (1938); see also F. C. Nix and W. Shockley, *Rev. Mod. Phys.* **10**, 27, 65 (1938).

that results from it by integration over q is known. It is given by

$$\log D(m) = \log 2 - \frac{1}{2}(1-m) \log(1-m) - \frac{1}{2}(1+m) \log(1+m).$$

Hence if the term Kq can be developed in terms of K and m only, the expression can be integrated. This is done by identifying the power series in powers of K resulting from (93) with the one resulting from a hypothetical expression of the form

$$\lambda^N = \int dm \exp N \left\{ \log D(m) + Cm + \tau_1 2K + \frac{1}{2!} \tau_2 (2K)^2 + \frac{1}{3!} \tau_3 (2K)^3 \dots \right\}. \quad (94a)$$

The q averages which result can be evaluated along the lines which led to formula (41). The first three τ 's have the following values

$$\begin{aligned} \tau_1 &= q_{Av} = m^2, \\ \tau_2 &= N(q_{Av}^2 - q_{Av}^2) = \frac{1}{2}(1-m^2)^2, \\ \tau_3 &= N^2(q_{Av}^3 - 3q_{Av}^2 q_{Av} + 2q_{Av}^3) = m^2(1-m^2)^2. \end{aligned} \quad (94b)$$

The evaluation of m as a function of K follows then from the usual extremum condition of the saddle point method.

The result of the Heisenberg approximation which proceeds only to τ_2 is somewhat surprising. It gives a phase transition at

$$K_c = 0.3912$$

with a jump of $-E/NJ$ from 0.6370 to 0.3912. The result should not be taken too seriously, however, for the next stage of the Kirkwood method, which is being evaluated numerically for the first time in this paper, brings about a regular Curie transition with specific heat jump, which takes place at

$$K_c = 0.3389.$$

A more simple and straightforward treatment is the one developed by Bethe.¹⁴ In its simplest form it treats only one spin with its direct neighbors by statistical methods, but the piece to be studied can be enlarged at will to approach the true solution. At low temperatures this

¹⁴ H. A. Bethe, Proc. Roy. Soc. **A150**, 552 (1935); see also Nix and Shockley, reference 13, pp. 17, 63.

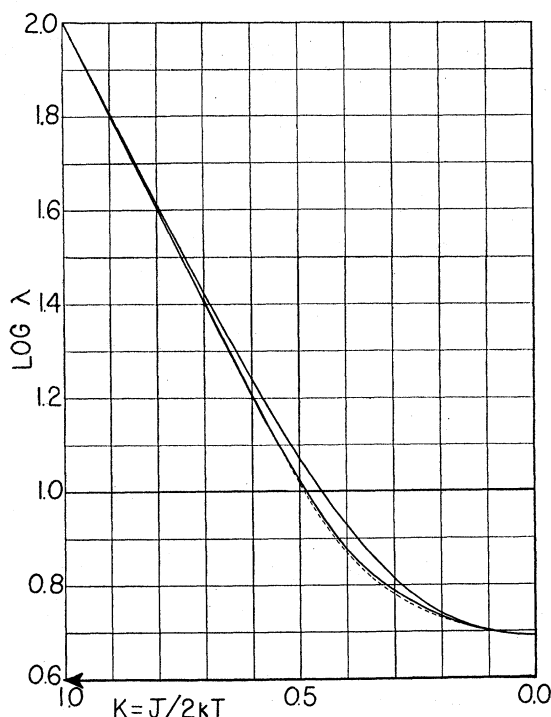


FIG. 6. Logarithm of partition function against reciprocal temperature. — 2-screw approximation (upper curve); variational solution (lower curve). - - - - Bethe solution. The slope of the curve is proportional to the energy at that temperature.

simple approximation is improved by assuming an ordering potential acting on the outer spins. It is determined by demanding symmetry in the result between them and the center spin. As a result of this calculation the following expressions are obtained for λ :

$$\begin{aligned} \lambda &= 2 \cosh^2 K && \text{below } K_c, \\ \lambda &= e^{2K} + \frac{e^{-2K}}{e^{4K} - 2} && \text{above } K_c. \end{aligned} \quad (95)$$

The two solutions join with a specific heat jump at

$$K_c = \frac{1}{2} \log 2 = 0.3466.$$

Results on the so-called second approximation of Bethe have been communicated to us by Mr. Groen. Its singular point is located at

$$K_c = 0.381.$$

Zernike¹⁵ has indicated a more dynamical approach to the problem by studying the propa-

¹⁵ F. Zernike, Physica **7**, 565 (1940).

gation of order in a lattice. By Boltzmann's theorem we can express the probability that a given spin has either value in terms of the probabilities for its neighbors. If the latter are treated as independent, a system of difference equations is the result. The unknowns are the probabilities $P_{hk}(\mu)$ for each lattice position h, k and the system is such that it permits the computation of any P from one given one. If we now take a certain spin as given, say $P_{00}(+) = 1$, we can ask for the probabilities for a spin very far away. At high temperatures, this limiting probability is $\frac{1}{2}$; Zernike is able to show, however, that below a certain temperature limiting probabilities different from $\frac{1}{2}$ will result. In our case this critical temperature is given by

$$\tanh 4K_c + 2 \tanh 2K_c - 2 = 0 \quad (96)$$

or

$$K_c = 0.3236.$$

The method is capable of giving all standard thermodynamic results because $P_{01}(\mu)$ is directly connected with the energy. It is found that the contact between the two branches is rather smooth, giving a discontinuity in the derivative of the specific heat only. It is interesting to notice that these calculations lead to a symmetry between very high and very low temperatures which is similar to (26). The propagation

at high temperature of short range order in a generally disordered lattice leads to a difference equation of the same form as the propagation at low temperature of local disorder in a generally ordered lattice [Zernike, Eqs. (5) and (44)].

Among the approximations, the λ 's of Section 6 (referred to below as "screw" solutions) should not be forgotten. They are, by their very nature, continuous functions of K , but compare not unfavorably with other approximate solutions.

In comparing various approximations it should be borne in mind that most of them are primarily approximating the partition function λ . It follows that an improved method of approach does not necessarily yield better energy and specific heat data, although it will be the case as a general rule. The differences in λ we have to account for are not large. In Fig. 6 curves having $\log \lambda$ plotted against K are given for the Bethe solution, the variation solution (Section 7) and the 2 screw. Most solutions discussed in this section are too small because they are essentially of the type (71); the screw solutions, on the other hand, are too large.

All λ 's agree for very large and very small K and their usefulness may be tested by comparing their series expansions with (41) and (43). Ordering the solutions according to magnitude, we find for small K

$$\begin{aligned} 3 \text{ screw: } \lambda &= 2 \left[1 + K^2 + \frac{7}{3} K^4 + \dots \right] \\ \text{Zernike: } \lambda &= 2 \left[1 + K^2 + \frac{4}{3} K^4 + \frac{227}{45} K^6 + \dots \right] \\ 5 \text{ screw: } \lambda &= 2 \left[1 + K^2 + \frac{4}{3} K^4 + \frac{122}{45} K^6 + \dots \right] \\ \text{correct: } \lambda &= 2 \left[1 + K^2 + \frac{4}{3} K^4 + \frac{77}{45} K^6 + \frac{1009}{315} K^8 + \dots \right] \\ \text{variation: } \lambda &= 2 \left[1 + K^2 + \frac{4}{3} K^4 + \frac{77}{45} K^6 + \frac{694}{315} K^8 + \dots \right] \\ \text{Heis-Kirk: } \lambda &= 2 \left[1 + K^2 + \frac{1}{2} K^4 + \dots \right] \\ \text{Bethe I: } \lambda &= 2 \left[1 + K^2 + \frac{1}{3} K^4 + \dots \right] \end{aligned} \quad (97)$$

For large K we find

$$\begin{aligned}
 \text{3 screw: } \lambda &= k^2[1 + 2k^{-8} + \dots \\
 \text{5 screw: } \lambda &= k^2[1 + k^{-8} + 3k^{-12} + \dots \\
 \text{correct: } \lambda &= k^2[1 + k^{-8} + 2k^{-12} + 5k^{-16} + 14k^{-20} + \dots \\
 \text{variation: } \lambda &= k^2[1 + k^{-8} + 2k^{-12} + 5k^{-16} + 14k^{-20} + \dots \\
 \text{Bethe I: } \lambda &= k^2[1 + k^{-8} + 2k^{-12} + 4k^{-16} + \dots \\
 \text{Kirkwood: } \lambda &= k^2[1 + k^{-8} + k^{-16} \left(\frac{64}{3}K^3 + 16K^2 + 8K \right) + \dots \\
 \text{Heisenberg: } \lambda &= k^2[1 + k^{-8} + k^{-16}(16K^2 + 8K) + \dots \\
 \text{Zernike: } \lambda &= k^2[1 + \frac{1}{2}k^{-8} + \dots
 \end{aligned} \tag{98}$$

It may be noticed that the Bethe and Kirkwood methods switch places in the two tabulations. The reason is that the Bethe method is more elaborate for low temperatures. The Kirkwood approximation, on the other hand, is poor for very low temperatures because (94) is indirectly based on a $1/T$ expansion.⁹ A similar situation exists for the Zernike solution because of his assumption of independent probabilities. Compared to other methods, our variational solution is very satisfactory. Being of the type (71), it lies below the true solution, but extremely close to it. A similar conclusion is reached if we extrapolate the sequence on p. 265 to find the value of λ at $K = 0.4407$. We find about $\lambda = 2.5335$, which is very close to the variational value $\lambda = 67 - 47\sqrt{2} = 2.5320$, but quite above the Bethe value $\lambda = 2.5224$.

This simple situation is complicated somewhat when we form the first and second deriva-

tives of λ which give the energy and the specific heat through (17) and (18). A solution whose λ is too large has an energy curve which is too smooth; and if λ is too small the energy passes too rapidly from one extreme to the other. This conclusion is verified by the energy *versus* temperature curves in Fig. 7. Beside various approximations, Fig. 7 has the exact location of K_c marked off together with its energy value (36). The correct curve passes through that point very probably in a vertical direction (Section 6).

If we disregard the Heisenberg solution which carries little weight for reasons mentioned above, we may divide the approximations in two classes. The ones proceeding from finite problems and the one given by Zernike have energy and specific heat continuous throughout. The others, using some kind of an inner field assumption, arrive at one singularity with a jump in the specific heat. As examples, specific heat curves

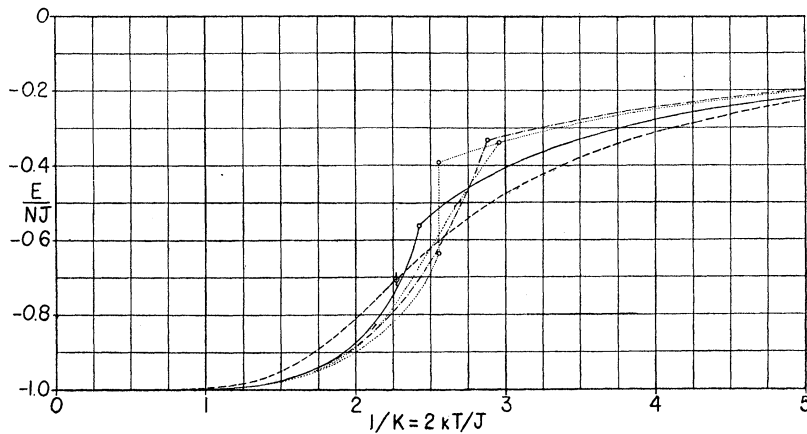
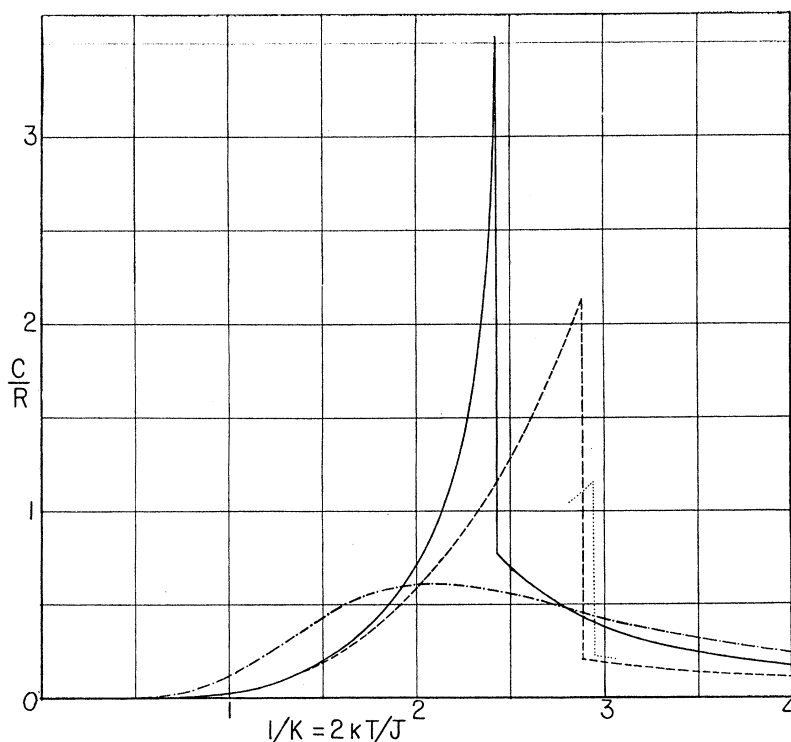


FIG. 7. Energy *versus* temperature in various approximations. - - - - 3-screw approximation. — Variation method. - · - · - Bethe method. · · · · · Heisenberg and Kirkwood. ○ Singular points. Point with vertical bar belongs to exact curve (with vertical slope).

FIG. 8. Specific heat *versus* temperature curves. — · · · · 2-screw approximation. — Variation method. - - - - Bethe method. · · · · Kirkwood method.



are plotted in Fig. 8 for the 2-screw and the variational Bethe and Kirkwood solutions. The large discontinuity in the specific heat is the dominant feature of the latter three. However, this result cannot possibly be correct, for it could be disproved by exact methods in Eq. (35). In view of the results of Section 6, it seems, therefore, that the main significance of these approximate results is not to be sought in the jump of C , but rather in its numerical magnitude at the Curie point. This magnitude increases as

the approximation improves. Simultaneously, the position of the singularity approaches the value (31), as can be verified from the data of this section and (92). Thus the approximate solutions of the last two sections are not in contradiction to the conclusions of Sections 4 and 6, that the Curie point lies at $K_c = 0.4407$ and that the specific heat at that point is infinite.

In conclusion, we want to express our thanks to Mr. P. Groen who has assisted us in some of the calculations.