Statistical Mechanics of Nearest Neighbor Systems II. General Theory and Application to TwoDimensional Ferromagnets

Elliott W. Montroll

Additional information on J. Chem. Phys.

by the addition of a few cc of alcohol which has no particular affinity for either component?

When the heat of mixing of polymer and solvent is positive, i.e., in a poor solvent medium, the polymer molecules should tend to prefer more compact configurations where intramolecular contacts are more numerous. As is well known, globular or spherical polymer molecules enhance the viscosity of their solutions relatively little in contrast to the marked effects of uncoiled long polymer chains which assume more or less random configurations. Therefore, the low relative viscosities in poor solvents very probably reflect changes in average configuration of the polymer molecules, dependent upon the solvent medium. Intrinsic viscosities at incipient precipitation are of the order of one-half or less of the values for the same polymers in a good solvent medium.[22] This would lead to the conclusion that the polymer molecule at the critical point pervades about the same volume as another molecule of half its chain length in a good solvent.

CONCLUSIONS

Quantitative statistical treatment of Meyer's model accounts for the following hitherto unexplained peculiar properties of high polymer solutions: (a) the shape of the partial molal free energy-composition diagram throughout the composition range (i.e., the large deviation from "ideal" entropies), (b) the virtual independence of partial pressure on molecular weight of the polymer except in dilute solutions, and (c) the extreme dissymmetry of the mutual solubility-temperature diagram and the limited high swelling capacity of linear polymers in certain media. These achievements confirm the correctness of the general point of view regarding these systems. On the other hand, quantitative agreement between calculations and observations is not particularly good, presumably due to imperfections of the model itself and to limitations of the present treatment imposed by assumption 4.

Statistical Mechanics of Nearest Neighbor Systems

II. General Theory and Application to Two-Dimensional Ferromagnets

Elliott W. Montroll*

Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut

(Received October 2, 1941)

This is a continuation of an effort to reduce to the solution of a characteristic value problem the rigorous calculation of thermodynamic properties of systems in which the intermolecular forces are sufficiently short ranged so that practically the entire potential energy of the system results from interactions between nearest neighbors. The partition function of such a system can be expressed in terms of the largest characteristic value of a linear operator equation and finally as the ratio of partition functions of systems with relatively few particles whose potential energy functions differ somewhat from those in the original system. A method of evaluating grand partition functions and one of introducing interactions between more distant neighbors are discussed. The general theory is applied to the calculation of magnetization, internal energy, and specific heat of two-dimensional ferromagnets on the basis of the Ising model. There seems to exist a λ-point phase transition in the change from ferromagnetic to nonferromagnetic states.

I. INTRODUCTION

THE intermolecular forces in many solids are sufficiently short ranged to permit the

* Sterling Research Fellow in Chemistry, Yale University. Present address, Cornell University, Ithaca, New York.

total potential energy of the solid to be represented by the sum of the interactions of each molecule with its nearest neighbors. With this approximation the evaluation of the potential energy contribution to the partition function and the thermodynamic properties of a solid can

be reduced to the solution of a characteristic value problem.[1,2] Inasmuch as such problems have been much studied while the summation of highly multiple sums has been a rather neglected subject, the reduction is an advantageous one.

Suppose a system can be divided into $L \times M \times N$ cells so that there will be L layers, each composed of $N \times M$ cells and each layer will have the same average behavior as any other layer. The molecules in each cell may have several or even a continuum of possible configurations. Since the same is true of each layer, let us define a set $\{\alpha_i\}$ of all those configurations available to a physical coordinate α of a layer. For example, in a binary substitutional alloy α_i might represent an ith arrangement of the two kinds of atoms in a layer.

If $v(\alpha_m)$ is the total potential energy with respect to a coordinate α between all neighboring cells of a layer in the mth configuration, and if

$v(\alpha_m, \alpha_n)$ is the total potential of interaction between cells of one layer in the mth configuration with their nearest neighbors in the nth configuration, the potential energy factor of the partition function with respect to the coordinate α is[3]

$$Z = \sum_{\{\alpha_1\}} \cdots \sum_{\{\alpha_L\}} \exp -\theta[v(\alpha_1) + \cdots + v(\alpha_L)$$
$$+ v(\alpha_1, \alpha_2) + \cdots + v(\alpha_{L-1}, \alpha_L)]. \quad (1)$$

Making the substitutions

$$V(\alpha_i, \alpha_j) = v(\alpha_i)/2 + v(\alpha_i, \alpha_j) + v(\alpha_j)/2;$$
$$\theta = 1/kT \quad (2)$$

(1) becomes

$$Z = \sum_{\{\alpha_1\}} \cdots \sum_{\{\alpha_L\}} \prod_{i=1}^{L-1} \exp -\theta V(\alpha_i, \alpha_{i+1}). \quad (3)$$

If $\{\psi_n\}$ is the set of orthonormal characteristic vectors and $\{\lambda_n\}$ the corresponding set of characteristic values of the matrix equation

$$\lambda \psi = \exp [-\theta V] \cdot \psi \quad (4)$$

$$\lambda \begin{bmatrix} \psi(\alpha_1) \\ \psi(\alpha_2) \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} \exp -\theta V(\alpha_1, \alpha_1) & \exp -\theta V(\alpha_1, \alpha_2) & \cdots \\ \exp -\theta V(\alpha_2, \alpha_1) & \exp -\theta V(\alpha_2, \alpha_2) & \cdots \\ \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot & \cdots \end{bmatrix} \begin{bmatrix} \psi(\alpha_1) \\ \psi(\alpha_2) \\ \cdot \\ \cdot \end{bmatrix}$$

(when the $\exp -\theta V$ matrix is symmetrical), by expanding $\exp -\theta V(\alpha_i, \alpha_j)$ as a bilinear combination of $\{\psi_n\}$ it can be shown that[1]

$$Z = \lambda_{max}^{L-1} [\sum_{\{\alpha_1\}} \psi_{max}(\alpha_1)]^2 \quad (5)$$

where λ_{max} is the largest characteristic value of (4).

II. PARTITION FUNCTION BY METHOD OF SUCCESSIVE APPROXIMATIONS

The secular equation of the matrix in (4) corresponding to a real crystal is of too great an order to be solved directly, so it seems necessary to seek a somewhat indirect method of solution that will lead to correct results. In studying long, narrow (three and four atoms wide), two-dimensional ferromagnetic strips[1] it was found that the method of successive approximations leads to asymptotic expressions for λ_{max} which show excellent agreement with the exact values

as calculated numerically. This suggests that the same approach might give good results when applied to real crystals.

In the method of successive approximations one first guesses a reasonable vector ψ_0 and calculates

$$\psi_1{}^u = (\exp -V\theta) \cdot \psi_0,$$
i.e.,
$$\psi_1{}^u(\alpha_1) = \sum_{\alpha_2} \psi_0(\alpha_2) \exp -\theta V(\alpha_1, \alpha_2).$$

Then some value of α_1, say $\alpha_1 = 0$, is chosen and one sets[4]

$$\psi_1(0) = \psi_0(0) = 1,$$
$$\psi_1(\alpha_1) = \psi_1{}^u(\alpha_1)/\psi_1{}^u(0)$$
$$= \sum_{\alpha_2} \psi_0(\alpha_2) \exp -\theta V(\alpha_1, \alpha_2)/$$
$$\sum_{\alpha_2} \psi_0(\alpha_2) \exp -\theta V(0, \alpha_2).$$

[1] E. Montroll, J. Chem. Phys. 9, 706 (1941). Henceforth this paper will be referred to as SM-I.

[2] H. Kramers and G. Wannier, Phys. Rev. 60, 252 and 263 (1941). These papers will be designated by KW-I and KW-II.

[3] Note the implicit assumption of independence of the various kinds of coordinates. For example, in a ferromagnetic system this would be equivalent to the assumption that spin orientations are independent of lattice vibrations.

[4] The superscript u is used to imply that $\psi_1{}^u(0)$ has not yet been normalized to unity. For a detailed mathematical discussion of the method of successive approximations see Frazier, Duncan, and Collar, *Elementary Matrices* (Cambridge University Press, 1938), p. 140.

$\psi_1(\alpha_1)$ is the second approximation of $\psi(\alpha_1)$, and to find the $(i+1)$st approximation one proceeds according to the general scheme

$$\psi_i{}^u(\alpha_1) = \sum_{\alpha_2} \psi_{i-1}(\alpha_2) \exp -\theta V(\alpha_1, \alpha_2),$$
$$\psi_i(\alpha_1) = \psi_i{}^u(\alpha_1)/\psi_i{}^u(0). \tag{6}$$

The corresponding characteristic value is

$$\lambda_i = \sum_{\alpha_1, \alpha_2} \psi_i(\alpha_1)\psi_i(\alpha_2) \exp -\theta V(\alpha_1, \alpha_2) / \sum_{\alpha_1, \alpha_2} \psi_i(\alpha_1)\psi_i(\alpha_2). \tag{7}$$

Using this procedure we shall now develop a power series in θ for $\lambda_{\max}$ and $\psi_{\max}(\alpha)$ of a real crystal. Since this development is the appropriate one for high temperatures, the logical first approximation ψ_0 of ψ should be that which corresponds to a system in a state of complete disorder; that is, all configurations of a given layer should be chosen as being equally probable. Such a ψ (unnormalized) is

$$\psi_0 = (1, 1, 1, \cdots, 1),$$
i.e., $\quad \psi_0(\alpha) = 1$ for all α.

Then if there are ν possible configurations per lattice point and ν^{MN} per layer

$$\psi_1{}^u(\alpha_1) = \sum_{\{\alpha_2\}} 1 \cdot [1 - \theta V(\alpha_1, \alpha_2) + \theta^2 V^2(\alpha_1, \alpha_2)/2 - \cdots]$$
$$= \nu^{MN} - \theta \sum_{\{\alpha_2\}} V(\alpha_1, \alpha_2) + \theta^2 \sum_{\{\alpha_2\}} V^2(\alpha_1, \alpha_2)/2 - \cdots$$

and

$$\psi_1(\alpha_1) = \psi_1{}^u(\alpha_1)/\psi_1{}^u(0)$$
$$= 1 - \theta\nu^{-MN} \sum_{\{\alpha_2\}} [V(\alpha_1, \alpha_2) - V(0, \alpha_2)] + \theta^2\nu^{-MN} \sum_{\{\alpha_2\}} \{[V^2(\alpha_1, \alpha_2) - V^2(0, \alpha_2)]/2 - V(0, \alpha_3)[V(\alpha_1, \alpha_2) - V(0, \alpha_2)]\} + \cdots.$$

Note that
$$\psi_1(\alpha_1) - \psi_0(\alpha_1) = 0(\theta).$$

Now
$$\psi_2(\alpha_1) = \psi_2{}^u(\alpha_1)/\psi_2{}^u(0) = \frac{\sum \psi_1(\alpha_2)\{1 - \theta V(\alpha_1, \alpha_2) + \theta^2 V^2(\alpha_1, \alpha_2)/2 - \cdots\}}{\sum \psi_1(\alpha_2)\{1 - \theta V(0, \alpha_2) + \theta^2 V^2(0, \alpha_2)/2 - \cdots\}}.$$

So
$$\psi(\alpha_1) \sim \psi_2(\alpha_1) = 1 - \theta\nu^{-MN} \sum_{\{\alpha_2\}} [V(\alpha_1, \alpha_2) - V(0, \alpha_2)] + \theta^2\nu^{-MN} \{ \sum_{\{\alpha_2\}} [V^2(\alpha_1, \alpha_2) - V^2(0, \alpha_2)]/2$$
$$+ \nu^{-2MN} \sum_{\{\alpha_2, \alpha_3, \alpha_4\}} [V(\alpha_1, \alpha_2) - V(0, \alpha_2)][V(\alpha_2, \alpha_3) - V(0, \alpha_3) - V(\alpha_3, \alpha_4)]\} - \cdots. \tag{8}$$

Here we see that
$$\psi_2(\alpha_1) - \psi_1(\alpha_1) = 0(\theta^2).$$

The coefficients of higher powers of θ are rather long expressions so we will not record them here. However, by continuing the process started above one can show, using the same sort of reasoning as that used in the appendix, that

$$\psi_l(\alpha) - \psi_{l-1}(\alpha) = 0(\theta^l). \tag{9}$$

Since by definition
$$\psi_l{}^u(\alpha_1) = \sum_{\alpha_2} \psi_{l-1}(\alpha_2) \exp -\theta V(\alpha_1, \alpha_2)$$
$$= \psi_l{}^u(0)\psi_l{}^u(\alpha_1)/\psi_l{}^u(0) = \psi_l(\alpha_1)\psi_l{}^u(0)$$

and $\psi_l(\alpha)$ is equal to $\psi_{l-1}(\alpha)$ to within terms of order θ^l, $\psi_l{}^u(0)$ is equal to λ_{l-1} (the lth approximation to $\lambda_{\max}$) to within terms of order θ^l. Thus

$$\lambda_0 = \nu^{MN},$$
$$\lambda_1 = \nu^{MN} - \nu^{-MN}\theta \sum_{\{\alpha_1, \alpha_2\}} V(\alpha_1, \alpha_2),$$

$$\lambda_{max} \sim \lambda_2 = \nu^{MN} - \nu^{-MN}\theta \sum_{\{\alpha_1, \alpha_2\}} V(\alpha_1, \alpha_2) + \theta^2 \nu^{-MN} \{ \sum_{\{\alpha_1, \alpha_2\}} V^2(\alpha_1, \alpha_2)/2$$

$$+ \nu^{-MN} \sum \cdots \sum_{\{\alpha_1, \alpha_2, \alpha_3\}} V(\alpha_1, \alpha_2) V(\alpha_2, \alpha_3) - \nu^{-2MN} [\sum_{\{\alpha_1, \alpha_2\}} V(\alpha_1, \alpha_2)]^2 \}. \quad (10)$$

At this point let us search for a physical interpretation of the iteration process described above. If the successive approximations are made without bothering to expand in a power series in θ,

$$\psi_0(\alpha) = 1 \text{ for all } \alpha.$$

Then

$$\psi_1(\alpha_1) = \sum_{\alpha_2} \exp - \theta V(\alpha_1, \alpha_2) / \sum_{\alpha_2} \exp - \theta V(0, \alpha_2),$$

$$\psi_2(\alpha_1) = \sum_{\alpha_2} \psi_1(\alpha_2) \exp - \theta V(\alpha_1, \alpha_2) / \sum_{\alpha_2} \psi_1(\alpha_2) \exp - \theta V(0, \alpha_2)$$

$$= \sum_{\alpha_2, \alpha_3} \exp - \theta [V(\alpha_1, \alpha_2) + V(\alpha_2, \alpha_3)] / \sum_{\alpha_2, \alpha_3} \exp - \theta [V(0, \alpha_2) + V(\alpha_2, \alpha_3)].$$

In general

$$\psi_l(\alpha_1) = \frac{\sum \cdots \sum \exp - \theta [V(\alpha_1, \alpha_2) + V(\alpha_2, \alpha_3) + \cdots + V(\alpha_l, \alpha_{l+1})]}{\sum \cdots \sum \exp - \theta [V(\alpha_1, \alpha_2) + V(\alpha_2, \alpha_3) + \cdots + V(\alpha_{l-1}, \alpha_l)]}, \quad (11)$$

$$\psi_{max}(\alpha) - \psi_l(\alpha) = 0(\theta^{l+1}). \quad (12)$$

The characteristic value corresponding to $\psi_l(\alpha)$ is

$$\lambda_l = \psi_l{}^*(\exp \ -\theta V)\psi_l / \psi_l{}^* \cdot \psi_l$$

$$= \frac{\sum \cdots \sum \exp - \theta [V(\alpha_1, \alpha_2) + \cdots + V(\alpha_{2l}, \alpha_{2l+1})]}{\sum \cdots \sum \exp - \theta [V(\alpha_1, \alpha_2) + \cdots + V(\alpha_{2l-1}, \alpha_{2l})]} = \frac{S_{2l+1}}{S_{2l}},$$

but in the appendix it is shown that

$$S_{2l+1}/S_{2l} = S_{l+1}/S_l + 0(\theta^{l+1}).$$

Therefore, remembering the remarks made following Eq. (9),

$$\lambda_{max} = S_{l+1}/S_l + 0(\theta^{l+1}). \quad (13)$$

Now

$$S_l = \sum_{\alpha_1} \cdots \sum_{\alpha_l} \exp - \theta [V(\alpha_1, \alpha_2) + \cdots + V(\alpha_{l-1}, \alpha_l)] \quad (14)$$

is the partition function of a system l layers thick in which the interaction between lattice points in the first (and in the lth) layer is one-half that between those of other layers of the system. Thus as $L \to \infty$, the $(L-1)$st root of Z (see Eq. (5)), the partition function of a system L layers thick, is equal, to within terms of $0[(1/kT)]^{l+1}$, to the partition function of a system of $(l+1)$ layers divided by that of a system of l layers:

$$Z^{1/(L-1)} = S_{l+1}/S_l + 0(\theta^{l+1}). \quad (15)$$

Since S_l is the partition function of a system $l \times M \times N$,

$$S_l = [\mu(l)_{max}]^{M+1} [\sum_\beta \varphi(l, \beta)_{max}]^2, \quad (16)$$

where $\mu(l)_{max}$ is the largest characteristic value of

$$\mu(l)\varphi(l, \beta_1) = \sum_{\beta_2} \varphi(l, \beta_2) \exp - W_l(\beta_1, \beta_2)\theta \quad (17a)$$

with

$$W_l(\beta_1, \beta_2) = w_l(\beta_1)/2 + w_l(\beta_1, \beta_2) + w_l(\beta_2)/2, \quad (17b)$$

$w_l(\beta_i) =$ potential energy of interaction between neighboring elements in the ith $l \times N$ layer in configuration β_i,

and

$w_l(\beta_i, \beta_j) =$ potential energy of interaction between elements in the ith $l \times N$ layer in configuration β_i, and their nearest neighbors in the jth layer in configuration β_j.

In the energy term $W_l(\beta_i, \beta_j)$ the interaction between the N elements in the first row of the ith layer and their nearest neighbors in the first row of the jth layer (likewise for interactions between lth rows of the two layers) is reduced by half.

If we let

$$T_{m,l} = \sum_{\beta_1} \cdots \sum_{\beta_m} \exp -\theta[W_l(\beta_1, \beta_2) + \cdots + W_l(\beta_{m-1,m})] = T_{l,m}, \tag{18a}$$

by the same arguments as those used in the derivation of (13):

$$\mu(l)_{\max} = T_{m+1,l}/T_{m,l} + 0(\theta^{m+1}) \tag{18b}$$

and

$$Z^{1/(L-1)(M-1)} = \frac{T_{l+1,l+1}/T_{l,l+1} + 0(\theta^{l+1})}{T_{l+1,l}/T_{l,l} + 0(\theta^{l+1})}$$

$$= T_{l+1,l+1}/T_{l,l} + 0(\theta^{l+1}). \tag{19}$$

Now $T_{l,l}$ is exactly the partition function of a system of $l \times l \times N$ lattice points in which the potential energy of interaction between all pairs of lattice points is the same as that between pairs in the original system, except: (1) for pairs both of whose components lie on one of the four faces of width l and length N—the interaction between such pairs in one-half that between normal pairs; (2) for pairs both of whose components lie on one of the four edges (formed by the intersection of two of the $l \times N$ planes) of length N—the interaction between such pairs is one-fourth that between normal pairs. This means that the $l \times l \times N$ system can be "compressed" in the N direction by the characteristic value method, and if $\omega(l,l)_{\max}$ is the largest characteristic value of the operator equation

$$\omega(l,l)\chi(l,l;\gamma_1) = \sum_{\gamma_2} \chi(l,l;\gamma_2) \exp -\theta X_{l,l}(\gamma_1, \gamma_2), \tag{20}$$

$$T_{l,l} = [\omega(l,l)_{\max}]^{N-1}[\sum_{\gamma} \chi(l,l;\gamma)_{\max}]^2. \tag{21}$$

Neglecting 1 as compared with L, M, N

$$Z^{1/LMN} = \omega(l+1,l+1)_{\max}/\omega(l,l)_{\max} + 0(\theta^{l+1}). \tag{22}$$

In Eq. (20)

$$X_{l,l}(\gamma_1, \gamma_2) = x_{l,l}(\gamma_1)/2 + x_{l,l}(\gamma_1, \gamma_2) + x_{l,l}(\gamma_2)/2.$$

$x_{l,l}(\gamma_i) =$ total potential energy of interaction between elements in the first layer while it is in configuration γ_i. The interaction between pairs of elements on the border of the $l \times l$ square is only one-half that between other pairs of elements.

$x_{l,l}(\gamma_1, \gamma_2) =$ total potential energy of interaction between elements in the first layer in configuration γ_1 and their nearest neighbors in the second in configuration γ_2. The interaction between the four elements on the corners of the $l \times l$ square in layer 1 and their nearest neighbors on the corners of layer 2 in one-fourth that between pairs of internal neighbors. The interaction between elements on the $l \times 2$ faces is one-half that between pairs of internal elements.

Thus we have shown that in order to calculate the partition function correct to terms of order $(1/kT)^l$ of a system of $L \times M \times N$ cells (as $L, M, N \rightarrow \infty$) it is sufficient to find the largest characteristic value of an operator equation whose kernel involves the potential energy of a system of $l \times l \times 2$ cells.

The low temperature calculation of λ_{max} can also be effected by the use of the method of successive approximations. Suppose there are two possible ordered configurations for a given layer[5] and that these are represented by 0_1 and 0_2. As a first approximation to ψ we choose a vector which defines a completely ordered state:

$$\psi_0(0_2) = \psi_0(0_1) = 1; \quad \psi_0(\alpha) = 0 \text{ if } 0_1 \neq \alpha \neq 0_2; \quad \psi_1{}^u(\alpha_1) = \exp\left[-\theta V(0_1, \alpha_1)\right] + \exp\left[-\theta V(0_2, \alpha_1)\right].$$

Here we shall keep

$$\psi_0(0_1) = \psi_1(0_1) = \cdots = 1.$$

Then

$$\psi_1(\alpha_1) = \frac{\exp\left[-\theta V(0_1, \alpha_1)\right] + \exp\left[-\theta V(0_2, \alpha_1)\right]}{\exp\left[-\theta V(0_1, 0_1)\right] + \exp\left[-\theta V(0_1, 0_2)\right]},$$

$$\psi_2(\alpha_1) = \frac{\sum_{\alpha_2}\left[\exp -\theta V(\alpha_2, 0_1) + \exp -\theta V(\alpha_2, 0_2)\right]\left[\exp -\theta V(\alpha_1, \alpha_2)\right]}{\sum_{\alpha_2}\left[\exp -\theta V(\alpha_2, 0_1) + \exp -\theta V(\alpha_2, 0_2)\right]\left[\exp -\theta V(0_1, \alpha_2)\right]}.$$

By continuing the approximation process we finally obtain

$$\psi_n(\alpha_1) = \frac{\sum \cdots \sum \{\exp -\theta V(\alpha_n, 0_1) + \exp -\theta V(\alpha_n, 0_2)\} \{\exp -\theta[V(\alpha_1, \alpha_2) + \cdots + V(\alpha_{n-1}, \alpha_n)]\}}{\sum \cdots \sum \{\exp -\theta V(\alpha_n, 0_1) + \exp -\theta V(\alpha_n, 0_2)\} \{\exp -\theta[V(0_1, \alpha_2) + \cdots + V(\alpha_{n-1}, \alpha_n)]\}} \quad (23a)$$

and

$$\lambda_n = \sum_{\alpha_1, \alpha_2} \psi_n(\alpha_1)\psi_n(\alpha_2) \exp -\theta V(\alpha_1, \alpha_2) / \sum_{\alpha_1, \alpha_2} \psi_n(\alpha_1)\psi_n(\alpha_2). \quad (23b)$$

It is apparent from Eqs. (23a) and (23b) that the low temperature expression for λ_{max} is the quotient of the partition functions of two smaller systems. Proceeding in a manner analogous to that used for the high temperature expressions, one can finally evaluate Z as the ratio of characteristic values of operator equations whose kernels are functions of the potential energy between two small layers of lattice points. The details of this calculation will not be carried out here, but they will be apparent from the discussion of the special case of two-dimensional ferromagnetism at the end of this paper.

III. GRAND PARTITION FUNCTION AND THE CHARACTERISTIC VALUE METHOD

In systems of more than one component it is usually necessary or at least it is usually more convenient to use the grand partition function for the calculation of thermodynamic quantities. So, we shall now investigate the applicability of the characteristic value method to the evaluation of grand partition functions.

The grand partition function[6] of a k component system is defined as the sum

$$P = \sum_{n_1, n_2, \cdots, n_k, m} \exp (\mu_1 n_1 + \mu_2 n_2 + \cdots \mu_k n_k - E_m)\theta \quad (24)$$

over-all compositions n_1, n_2, $\cdots$, n_k and over-all energy states E_m. The terms μ_i are the chemical potentials of the various components and are so chosen that

$$N_i = \frac{1}{P} \sum_{i=1, 2, \cdots, k} \cdots \sum n_i \exp \theta(\mu_1 n_1 + \cdots + \mu_k n_k - E_m), \quad (25)$$

equals the number of particles of the ith component of the system.

---

[5] This would be the case if each lattice point could achieve either of two configurations. The method can without difficulty be extended to the case of more than two ordered configurations per layer.

[6] For discussions of grand partition functions and their uses cf. Tolman, *Principles of Statistical Mechanics* (Oxford University Press, 1938), p. 619; Gibbs, *Collected Works* (Longmans, Green and Company, 1931), p. 187.

In a nearest neighbor system of the type discussed in Part I, retaining the notation introduced there,

$$P = \sum_{\alpha_1 \cdots \alpha_L} \cdots \sum \exp \theta [\mu_1 n_1(\alpha_1) + \cdots + \mu_k n_k(\alpha_1) - v(\alpha_1) + \mu_1 n_1(\alpha_2) + \cdots + \mu_k n_k(\alpha_2) - v(\alpha_2)$$

$$\cdots \cdots \cdots \cdots \cdots$$

$$+ \mu_1 n_1(\alpha_L) + \cdots + \mu_k n_k(\alpha_L) - v(\alpha_L) - v(\alpha_1, \alpha_2) - v(\alpha_2, \alpha_3) - \cdots - v(\alpha_{L-1}, \alpha_L)]$$

with $n_i(\alpha_j)$ being defined as the number of particles of kind i in the jth layer when it is in configuration α_j.

If we make the substitution

$$V(\alpha_i, \alpha_j) = -[\mu_1 n_1(\alpha_i) + \cdots + \mu_k n_k(\alpha_i) - v(\alpha_i)]/2 + v(\alpha_i, \alpha_j) - [\mu_1 n_1(\alpha_j) + \cdots + \mu_k n_k(\alpha_j) - v(\alpha_j)]/2, \qquad (26)$$

$$P = \sum_{\alpha_1} \cdots \sum_{\alpha_L} \prod_{i=1}^{L-1} \exp - \theta v(\alpha_i \alpha_{i+1})$$

and if the **exp** $- \mathbf{\theta V}$ matrix is symmetrical the problem is formally the same as that of the evaluation of Z

$$P = \lambda_{\max}^{L-1} [\sum_\alpha \psi_{\max}(\alpha)]^2. \qquad (27)$$

The calculating devices introduced in studying Z are valid for studying P.

As an example let us consider the grand partition function of a one-dimensional binary substitutional alloy. Let there be N_0 particles of kind "0" and N_1 particles of kind "1," and further let E_{ij} be the energy of interaction of a particle of kind "i" with a nearest neighbor of kind "j." Since any lattice point can be occupied by either kind of particle,

$$P = \lambda_{\max}^{N_0 + N_1 - 1} [\sum_\alpha \psi_{\max}(\alpha)]^2$$

and the **exp** $- \mathbf{\theta V}$ matrix is

$$\mathbf{exp} - \mathbf{\theta V} = \begin{pmatrix} \exp - \theta V(0, 0) & \exp - \theta V(0, 1) \\ \exp - \theta V(1, 0) & \exp - \theta V(1, 1) \end{pmatrix}$$

$$= \begin{pmatrix} \exp \theta(\mu_0 - E_{00}) & \exp \theta[(\mu_0 + \mu_1)/2 - E_{10}] \\ \exp \theta[(\mu_0 + \mu_1)/2 - E_{10}] & \exp \theta(\mu_1 - E_{11}) \end{pmatrix}.$$

By simple algebra the roots of the secular equation of **exp** $- \mathbf{\theta V}$ are

$$\lambda_1 = \tfrac{1}{2} \{ \exp \theta(\mu_0 - E_{00}) + \exp \theta(\mu_1 - E_{11}) + ([\exp \theta(\mu_0 - E_{00}) - \exp \theta(\mu_1 - E_{11})]^2 + 4 \exp \theta(\mu_1 + \mu_0 - 2E_{12}))^{\frac{1}{2}} \},$$

$$\lambda_2 = \tfrac{1}{2} \{ \exp \theta(\mu_0 - E_{00}) + \exp \theta(\mu_1 - E_{11}) - ([\exp \theta(\mu_0 - E_{00}) - \exp \theta(\mu_1 - E_{11})]^2 + 4 \exp \theta(\mu_1 + \mu_0 - 2E_{12}))^{\frac{1}{2}} \}.$$

Thus

$$P^{1/(N_0 + N_1 - 1)} = \lambda_1$$

and the chemical potentials μ_0, μ_1 must be chosen to make simultaneously

$$N_0 = \partial \log \lambda_1 / \partial (\mu_0 \theta),$$

$$N_1 = \partial \log \lambda_1 / \partial (\mu_1 \theta).$$

Inasmuch as this problem was introduced merely as an example, we shall carry it no further. It might be mentioned however that there is no phase transition in this linear system.

IV. INTRODUCTION OF INTERACTIONS WITH MORE DISTANT NEIGHBORS AS APPLIED TO LINEAR FERROMAGNETS

The characteristic value theory of the partition function developed in SM-I and in further studies of the theory as reported so far in this paper are adapted to "systems in which the molecular forces are sufficiently short ranged to permit the total energy of the system to be represented by the sum of the interactions of each molecule with its nearest neighbors." One might now ask if it is possible to include direct interactions between more distant neighbors (as well as the correlation effects which are automatically taken into account in the present method) and still retain some remnants of the characteristic value theory. Formally it is made possible by placing more than one molecule in each of the cells mentioned in the introduction. The potential energy terms in (1) would then represent interactions between multiply populated cells. Of course the $\exp -\theta V$ matrices would be of a higher order than those of the corresponding nearest neighbor systems. A further difficulty arises in that the $\exp -\theta V$ matrix of the extended problem may not be symmetrical even though that of the original problem is, so we cannot blindly use (5).

In order to understand an approach that is satisfactory for evaluating some partition functions in which interactions between more distant neighbors have been introduced, and in order to see what effect the introduction has on thermodynamic properties let us consider in detail the statistics of a linear ferromagnetic chain.

The seat of the magnetic properties of a ferromagnetic element is in the unpaired electrons of the d shell of the atoms at each lattice point. We shall let $-\epsilon$ be the energy of interaction of two neighboring atoms whose unpaired spins are parallel and $+\epsilon$ that when they are antiparallel. Furthermore, we shall let $-f$ be the energy of

interaction of two "next nearest neighbors" with parallel excess spins, and f that with antiparallel spins. Also, we write $J=\epsilon/kT$ and $K=f/kT$. A linear chain of N lattice points can be divided into $N/2$ cells containing two lattice points per cell

$$
\underset{\text{I}}{\overset{1}{|\times}} \underset{\text{II}}{\overset{2}{\times|}} \underset{\text{III}}{\overset{3}{\times|}} \overset{4}{\times|} \overset{5}{\times|} \overset{6}{\times|} \cdots | \underset{N/2}{\overset{N-1}{\times}} \overset{N}{\times|}.
$$

By representing a lattice point with excess $+$spin by a "0" and one with excess $-$spin by a "1," there exists a unique two digit number, α, in the binary system which represents a given cell configuration. If we define $v(\sigma_i, \sigma_j)$ as the spin interaction between the ith lattice point in the configuration σ_i and the jth lattice point in configuration σ_j, the partition function of our linear system is

$$
Z=\sum_{\sigma_1}\cdots\sum_{\sigma_N} \exp -\theta[v(\sigma_1, \sigma_2)+v(\sigma_1, \sigma_3) \\ +v(\sigma_2, \sigma_3)+v(\sigma_2, \sigma_4)+\cdots+v(\sigma_N, \sigma_{N-1})].
$$

Then substituting

$$
V(\alpha_i, \alpha_{i+1})=v(\sigma_{2i-3}, \sigma_{2i-2})/2+v(\sigma_{2i-3}, \sigma_{2i-1}) \\ +v(\sigma_{2i-2}, \sigma_{2i-1})+v(\sigma_{2i-1}, \sigma_{2i}) \\ +v(\sigma_{2i-1}, \sigma_{2i})/2,
$$

$$
Z=\sum_{\alpha_1}\cdots\sum_{\alpha_{N/2}} \prod_{i=2}^{N/2-1} \exp -\theta V(\alpha_i, \alpha_{i+1}).
$$

Now let us examine the unsymmetrical matrix

$$
\exp -\theta V =
\begin{array}{c}
\alpha_2\backslash\alpha_1 \\
00 \\
01 \\
11 \\
10
\end{array}
\begin{array}{cccc}
00 & 01 & 11 & 10 \\
\end{array}
\left[
\begin{array}{cccc}
e^{2J+2K} & e^{J} & e^{-2K} & e^{-J} \\
e^{-J} & e^{-2J+2K} & e^{J} & e^{-2K} \\
e^{-2K} & e^{-J} & e^{2J+2K} & e^{J} \\
e^{J} & e^{-2K} & e^{-J} & e^{-2J+2K}
\end{array}
\right].
$$

To demonstrate the calculation of the matrix elements let us consider the element in the 01th row and the 11th column. This is the term $\exp -\theta V(01, 11)$ and corresponds to the configuration

$$
\frac{0 \quad 1 \ | \ 1 \quad 1}{1 \quad 2 \quad 3 \quad 4}.
$$

Since

$$
v(\sigma_1, \sigma_2)=\epsilon, \quad v(\sigma_1, \sigma_3)=f, \quad v(\sigma_2, \sigma_3)=-\epsilon, \quad v(\sigma_2, \sigma_4)=-f, \quad v(\sigma_3, \sigma_4)=-\epsilon,
$$

$$
V(01, 11)=\epsilon/2+(f-\epsilon-f)-\epsilon/2=-\epsilon=-JkT
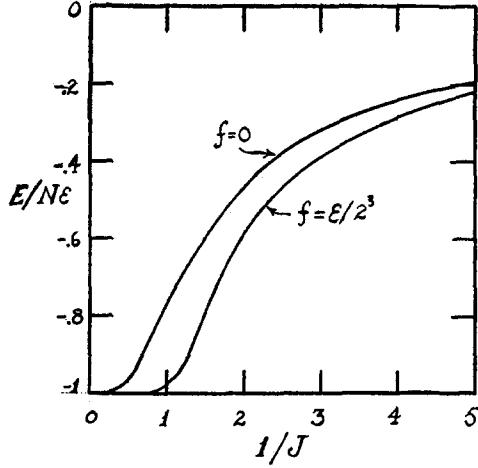$$

and

$$
\exp -\theta V(01, 11)=e^{J}.
$$

FIG. 1. Average internal energy of spin per lattice point (in units of ϵ) for a nearest and next nearest neighbor system.

FIG. 2

The matrix

$$\mathbf{R} = 2^{-\frac{1}{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} = \mathbf{R}^{-1}$$

transforms $\exp -\theta V$ into

$$\mathbf{R} \exp(-\theta V)\mathbf{R}^{-1} = \begin{bmatrix} e^{2J+2K}+e^{-2K} & e^{J}+e^{-J} & 0 & 0 \\ e^{J}+e^{-J} & e^{-2K}+e^{2K-2J} & 0 & 0 \\ 0 & 0 & e^{2J+2K}-e^{-2K} & e^{J}-e^{-J} \\ 0 & 0 & e^{-J}-e^{J} & e^{-2J+2K}-e^{-2K} \end{bmatrix}$$

and the roots of the symmetrical upper left factor of the secular equation of $\mathbf{R} \exp(-\theta V)\mathbf{R}^{-1}$ are

$$\lambda_1 = e^{-2K}+e^{2K}\cosh 2J + 2\cosh J(e^{4K}\sinh^2 J+1)^{\frac{1}{2}},$$

$$\lambda_2 = e^{-2K}+e^{2K}\cosh 2J - 2\cosh J(e^{4K}\sinh^2 J+1)^{\frac{1}{2}}.$$

The vectors

$$\phi_1^* = (A_1, B_1, 0, 0), \quad \phi_2^* = (A_2, B_2, 0, 0),$$

with

$$A_s = \frac{\lambda_s - e^{-2J+2K} - e^{-2K}}{[\lambda_s^2 - 2\lambda_s(e^{-2J+2K}+e^{-2K}) + (e^{J}+e^{-J})^2 + (e^{-2J+2K}+e^{-2K})^2]^{\frac{1}{2}}},$$

$$B_s = (1 - A_s^2)^{\frac{1}{2}},$$

are the pair of orthonormal characteristic vectors of $\mathbf{R} \exp(-\theta V)\mathbf{R}^{-1}$, whose characteristic values are λ_1 and λ_2, respectively. The corresponding characteristic vectors of $\exp -\theta V$ are

$$\psi_1^* = \phi_1^* R = (A_1, B_1, A_1, B_1)/\sqrt{2}, \quad \psi_2^* = \phi_2^* R = (A_2, B_2, A_2, B_2)/\sqrt{2},$$

also with characteristic values λ_1 and λ_2.

At finite temperatures >0, the roots λ_3 and λ_4 of the secular equation of the lower right sub-matrix of $\mathbf{R} \exp(-\theta V)\mathbf{R}^{-1}$ have absolute values less than λ_1. Since their parent matrix is not symmetrical,

$$N \begin{cases} \begin{bmatrix} 0 & \times & \times & \cdots & \times & 0 & \beta'_1 \\ 0 & \times & \times & \cdots & \times & 0 & \beta'_2 \\ 0 & \times & \times & \cdots & \times & 0 & \beta'_3 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \times & \times & \cdots & \times & 0 & \beta'_{N-1} \\ 0 & \times & \times & \cdots & \times & 0 & \beta'_N \end{bmatrix} \end{cases} \qquad N \begin{cases} \begin{bmatrix} 0 & \times & \times & \cdots & \times & 1 & \gamma'_1 \\ 0 & \times & \times & \cdots & \times & 1 & \gamma'_2 \\ 0 & \times & \times & \cdots & \times & 1 & \gamma'_3 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \times & \times & \cdots & \times & 1 & \gamma'_{N-1} \\ 0 & \times & \times & \cdots & \times & 1 & \gamma'_N \end{bmatrix} \end{cases}$$

with column headings $\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n$ over both matrices.

<div align="center">FIG. 3</div>

the characteristic vectors differ fundamentally from ϕ_1 and ϕ_2. If $\phi_3{}^R = (0, 0, C_3, D_3)$ is a right characteristic vector with the property

$$[R \exp -\theta V R^{-1}]\phi_3{}^R = \lambda_3 \phi_3{}^R$$

it does not follow that $\phi_3{}^R[\ ] = \lambda_3 \phi_3{}^R$; however, in our special case $\phi_3{}^L = (0, 0, C_3, -D_3)$ satisfies the equation $\phi_3{}^L[\ ] = \lambda_3 \phi_3{}^L$. These vectors are biorthogonal to the other characteristic vectors (including $\phi_4{}^L$ and $\phi_4{}^R$ which correspond to λ_4) in the manner

$$\phi_j{}^L \cdot \phi_i = 0; \quad \phi_i{}^* \cdot \phi_j{}^R = 0,$$
$$\phi_4{}^L \cdot \phi_3{}^R = 0; \quad \phi_3{}^L \cdot \phi_4{}^R = 0, \qquad (i = 1, 2; j = 3, 4)$$

and can be normalized so that

$$\phi_j{}^L \cdot \phi_j{}^R = 1.$$

The corresponding characteristic vectors of $\exp -\theta V$ are

$$\psi_3{}^L = \phi_3{}^L R = (C_3, D_3, -C_3, -D_3)/\sqrt{2}; \quad \psi_3{}^R = (C_3, -D_3, -C_3, D_3)/\sqrt{2}$$

and

$$\psi_4{}^L = \phi_4{}^L R = (C_4, D_4, -C_4, -D_4)/\sqrt{2}; \quad \psi_3{}^R = (C_4, -D_4, -C_4, D_4)/\sqrt{2}.$$

By expanding

$$\exp -\theta V(\alpha_1, \alpha_2) = \sum_i \lambda_i \psi_i{}^L(\alpha_1)\psi_i{}^R(\alpha_2)$$

(with $\psi_1{}^L(\alpha) = \psi_1{}^R(\alpha) = \psi_1(\alpha)$ and $\psi_2{}^L(\alpha) = \psi_2{}^R(\alpha) = \psi_2(\alpha)$)

and applying the orthonormality of the characteristic vectors,

$$Z = \lambda_1{}^{N/2-1}[\sum_\alpha \psi_1(\alpha)]^2 + \lambda_2{}^{N/2-1}[\sum_\alpha \psi_2(\alpha)]^2 + \lambda_3{}^{N/2-1}[\sum \psi_3(\alpha)]^2 + \lambda_4{}^{N/2-1}[\sum \psi_4(\alpha)]^2.$$

Since

$$\lambda_1 > \lambda_2, \lambda_3, \lambda_4; \quad \lambda_1{}^{N/2-1} \gg \lambda_2{}^{N/2-1}, \lambda_3{}^{N/2-1}, \lambda_4{}^{N/2-1}$$

$$Z = \lambda_1{}^{N/2-1}[\sum \psi_1(\alpha)]^2 \quad \text{and} \quad Z^{1/N} = \lambda_1{}^{\frac{1}{2}}$$

is the partition function per lattice point. Neglecting interactions between "next nearest" neighbors, $f = 0$ and

$$\lambda_1 = \lambda_{\max} = (e^J + e^{-J})^2,$$

yielding

$$Z^{1/N} = 2 \cosh J,$$

which agrees with the known partition function of nearest neighbor linear chains.[1,2]

Since the average internal energy of spin interaction is

$$E = \partial \log Z / \partial(-\theta),$$

it can be found from our expression for Z. $E/N\epsilon$ is plotted in Fig. 1 for $f = 0$ and $f = \epsilon/2$.[3]

V. APPLICATION TO TWO-DIMENSIONAL FERROMAGNETS

As an application of Parts I and II, we shall calculate the high and low temperature partition function, internal energy, specific heat, and magnetization of two-dimensional $N \times M$ (as N, $M \to \infty$) ferromagnets. The definitions of interaction energies introduced in Part IV will be retained here. If our ferromagnet is inserted in a magnetic field H, the total energy of spin-field interaction is

$$-\mu(n_+ - n_-)H = -\mu H \sum_{i=1}^{N} [n_+(\alpha_i) - n_-(\alpha_i)], \tag{28}$$

where

μ = magnetic moment of each spin,

$n_+(\alpha_i)$ = number of "+" spins in the ith row when it is in configuration α_i,

$n_-(\alpha_i)$ = number of "−" spins in the ith row when it is in configuration α_i,

$n_+ = \sum n_+(\alpha_i)$ and $n_- = \sum n_-(\alpha_i)$.

We shall henceforth abbreviate $C = \mu H / kT$.

1. High Temperature Development

To derive a power series in $\theta = 1/kT$ for Z that is valid at high temperatures, the two-dimensional analogue of (13) will be used. This necessitates the calculation of

$$S_n = \sum \cdots \sum \exp -[V(\alpha_1, \alpha_2) + V(\alpha_2, \alpha_3) + \cdots + V(\alpha_{n-1}, \alpha_n)]/kT \tag{29a}$$

and S_{n+1} with the value of n determined by the number of terms desired in the resulting power series.

These S_n's correspond to partition functions of narrow strips of the sort exhibited in Fig. 2 in which the interactions between neighboring elements in chain α_1 (and likewise those between neighboring elements of chain α_n) are only one-half those between other adjacent elements in the lattice. Instead of expressing S_n as a sum over configurations of the vertical chains (i.e., α's) in Fig. 3, one can equally well express it as a sum over configurations of horizontal chains (i.e., β's)

$$S_n = \sum_{\beta_1} \cdots \sum_{\beta_N} \exp -\theta[W(\beta_1, \beta_2) + \cdots + W(\beta_{N-1}, \beta_N)], \tag{29b}$$

$$W(\beta_i, \beta_j) = \tfrac{1}{2} \sum_{k=1}^{n-1} v(\sigma_{i,k}, \sigma_{i,k+1}) + \tfrac{1}{2}[v(\sigma_{i,1}, \sigma_{j,1}) + v(\sigma_{i,n}, \sigma_{j,n})] + \sum_{k=1}^{n} v(\sigma_{i,k}, \sigma_{j,k})$$

$$+ \tfrac{1}{2} \sum_{k=1}^{n-1} v(\sigma_{j,k}, \sigma_{j,k+1}) - \frac{C}{2} \sum_{k=2}^{n-1} [n_+(\sigma_{i,k}) + n_+(\sigma_{j,k}) - n_-(\sigma_{i,k}) - n_-(\sigma_{j,k})]$$

$$- \frac{C}{4}[n_+(\sigma_{i,1}) + n_+(\sigma_{j,1}) + n_+(\sigma_{i,n}) + n_+(\sigma_{j,n}) - n_-(\sigma_{i,1}) - n(\sigma_{j,1}) - n_-(\sigma_{i,n}) - n_-(\sigma_{j,n})]$$

and σ_{il} defines to the spin configuration of the lth element in the ith row.

$$n_+(\sigma) = 1 \quad \text{if} \quad \sigma = 0 \quad \text{(i.e., + spin),}$$
$$n_+(\sigma) = 0 \quad \text{if} \quad \sigma = 1 \quad \text{(i.e., − spin),}$$
$$n_-(\sigma) = 0 \quad \text{if} \quad \sigma = 0,$$
$$n_-(\sigma) = 1 \quad \text{if} \quad \sigma = 1.$$

By expanding $\exp -W_n(\beta_i \beta_j)$ as a bilinear combination of the characteristic vectors of the equation[7]

$$\mu(n)\varphi(n, \beta_i) = \sum_{\beta_j} \exp -\theta W_n(\beta_i, \beta_j)\varphi(n, \beta_j), \tag{29c}$$

$$S_n = \mu(n)_{\max}^{N-1}[\sum_{\beta} \varphi(n, \beta)]^2.$$

[7] $\mu(n)$ is a characteristic value corresponding to strips of width n.

If $n = 2$, the exp $-\theta W$ matrix will be constructed so that the element in the ith row and jth column is the term

$$\exp -W(i-1, j-1)\theta.$$

The interaction energy in this term is that between a row in a configuration which corresponds to the two digit number $(i-1)$ in the binary system and a neighboring row in a configuration corresponding to the two digit number $(j-1)$. For example,

$$W(0, 1) = [v(0, 0) + v(0, 0) + v(0, 1) + v(0, 1)]/2 - C[3n_+(0) - n_-(1)]/4$$

$$\begin{matrix} 0{-}0 & \beta_1 = 0 \\ | \quad | \\ 0{-}1 & \beta_2 = 1 \end{matrix}$$

but

$$v(0, 1) = \epsilon, \quad v(0, 0) = -\epsilon, \quad \epsilon = JkT.$$

Therefore $W(0, 1) = -C/2$. Similar calculations of the other matrix elements yield

$$\exp -W\theta = \begin{pmatrix} e^{2J+C} & e^{C/2} & c^{C/2} & 1 \\ e^{C/2} & 1 & e^{-2J} & 1 \\ e^{C/2} & e^{-2J} & 1 & 1 \\ 1 & 1 & 1 & e^{2J-C} \end{pmatrix}.$$

Expanding each matrix element as a power series in θ (i.e., in powers of J and C), choosing as a first approximation to the characteristic vector $\exp -W\theta$ the unnormalized vector of complete disorder $(1, 1, 1, 1)$ and finally proceeding by the method of successive approximations discussed in S.M.–I, Part–2, we obtain for the unnormalized characteristic vector $\phi(2)_{max}$ of $\exp -W\theta$ and for its characteristic value $\mu(2)_{max}$

$$\phi(2)_{max} = \begin{pmatrix} 1 \\ 1 - J - C/2 + C^2/8 + J^2/2 \\ 1 - J - C/2 + C^2/8 + J^2/2 \\ 1 - C + C^2/2 - CJ \end{pmatrix},$$

$$\mu(2)_{xam} = 4 + 3J^2 + C^2 + 0(\theta^3).$$

When $n = 3$ the $\exp -W\theta$ matrix is

$$\exp -W\theta = \left(\begin{array}{cccc|cccc} e^{2C+4J} & e^{2J+3C/2} & e^{C} & e^{C/2} & e^{2J+3C/2} & e^{C} & e^{C/2} & 1 \\ e^{2J+3C/2} & e^{C+2J} & e^{-2J+C/2} & 1 & e^{C} & e^{C/2} & e^{-2J} & e^{-C/2} \\ e^{C} & e^{-2J+C/2} & 1 & e^{-C/2} & e^{-2J+C/2} & e^{-4J} & e^{-C/2} & e^{-C} \\ e^{C/2} & 1 & e^{-C/2} & e^{2J-C} & e^{-2J} & e^{-2J-C/2} & e^{-C} & e^{2J-3C/2} \\ \hline e^{2J+3C/2} & e^{C} & e^{-2J+C/2} & e^{-2J} & e^{2J+C} & e^{C/2} & 1 & e^{-C/2} \\ e^{C} & e^{C/2} & e^{-4J} & e^{-2J-C/2} & e^{C/2} & 1 & e^{-2J-C/2} & e^{-C} \\ e^{C/2} & e^{-2J} & e^{-C/2} & e^{-C} & 1 & e^{-2J-C/2} & e^{2J-C} & e^{2J-3C/2} \\ 1 & e^{-C/2} & e^{-C} & e^{2J-3C/2} & e^{-C/2} & e^{-C} & e^{2J-3C/2} & e^{4J-2C} \end{array} \right).$$

The characteristic vector $\phi(3)_{max}$ which corresponds to the largest characteristic value $\mu(3)_{max}$

(unnormalized)

$$\phi(3)_{max} = \begin{vmatrix} 1 \\ 1-C/2-J+C^2/8+J^2/2 \\ 1-2J-C+C^2/2+2J^2 \\ 1-3C/2-J+9C^2/8-CJ+J^2/2 \\ \overline{1-C/2-J+C^2/8+J^2/2} \\ 1-2J-C+C^2/2+JC+2J^2 \\ 1-3J/2-J+9C^2/8-CJ+J^2/2 \\ 1-2C+2C^2-3CJ \end{vmatrix}$$

and

$$\mu(3)_{max} = 8+14J^2+6C^2+\cdots.$$

Then from Eqs. (13) and (29c) the first approximation to the partition function is (as M, $N \to \infty$ so that $M-1 \to M$ and $N-1 \to N$)

$$Z^{1/(N-1)} \sim S_3/S_2 + 0(\theta^3)$$

and

$$Z^{1/MN} \sim \mu(3)_{max}/\mu(2)_{max} + 0(\theta^3)$$

$$\sim 2+2J^2+C^2+0(\theta^3). \tag{30}$$

By the well-known relation between magnetization and the partition function

$$\mathfrak{M}/MN = \mu \partial \log Z/\partial C = \mu C + \cdots.$$

As the external field $C \to 0$, the magnetization per lattice point becomes zero and therefore in the temperature range where (30) is valid, a two-dimensional "ferromagnetic" system is not ferromagnetic.

In the absence of a magnetic field we have calculated by the above method

$$Z_0^{1/MN} \sim \mu(5)_{max}/\mu(4)_{max} + 0(\theta^5) = 2(1+J^2+4J^4/3+\cdots),$$

which agrees with the first three terms of the result of Kramers and Wannier[2,8] obtained by another method

$$Z_0^{1/MN} = 2(1+J^2+4J^4/3+77J^6/45+\cdots). \tag{31}$$

Since

$$\bar{E} = \partial \log Z/\partial(-\theta) \quad \text{and} \quad J = \epsilon/kT = \epsilon\theta,$$

and

$$\bar{E}/MN \sim -\epsilon(2J+10J^3/3+64J^5/15+\cdots) \tag{32a}$$

$$C_v/kMN \sim J^2(2+10J^2+64J^4/3+\cdots). \tag{32b}$$

2. Low Temperature Development

To derive the low temperature exponential development for Z we shall use Eq. (23) and therefore need

and

$$S_{0,n} = \sum_{\alpha_1} \cdots \sum_{\alpha_n} \exp -[V(0, \alpha_1)+V(\alpha_1, \alpha_2)+\cdots+V(\alpha_n, 0)]/kT$$

$$S_{1,n} = \sum_{\alpha_1} \cdots \sum_{\alpha_n} \exp -[V(0, \alpha_1)+V(\alpha_1, \alpha_2)+\cdots+V(\alpha_n, 1)]/kT.$$

These S's correspond to partition functions of narrow strips as shown in Fig. 3, in which the interactions between neighboring elements in the chains of 0's and in the chain of 1's is only one-half that between other adjacent elements in the lattice. Instead of expressing the S's as sums over configurations of vertical chains (i.e., in terms of α's) we shall express them as sums over horizontal (i.e., β' and γ') configurations. If we proceed in a manner analogous to that used in the low temper-

---

[8] See also W. Opechowski, Physica **4**, 181 (1937).

ature development

$$S_{0,n} + S_{1,n} = {}_{n,0}\mu_{\max}[\sum \varphi_{\max}(\beta')]^2 + {}_{n,1}\mu_{\max}[\sum \varphi_{\max}(\gamma')]^2,$$

where ${}_{n,0}\mu_{\max}$ and ${}_{n,1}\mu_{\max}$ are, respectively, the largest characteristic values of the matrices of inter-actions of two adjacent β' rows and between two adjacent γ' rows. Now

$$_{n,0}\mu_{\max} = \exp\left[(2n+2)J + (n+1)C\right] + O(\exp\left[(2n-6)J + nC\right]),$$

$$_{n,1}\mu_{\max} = \exp\left(2nJ + nC\right) + O(\exp\left[(2n-8)J + (n-1)C\right]).$$

So,

$$({}_{n,1}\mu_{\max}/{}_{n,0}\mu_{\max})^N = \exp\,-(2NJ + NC) + \cdots \rightarrow 0$$

as $N \rightarrow \infty$, and we can neglect $S_{1,n}$ as compared with $S_{0,n}$:

$$S_{0,n} + S_{1,n} = S_{0,n}[1 + O(e^{-2NJ - NC})].$$

To find ${}_{1,0}\mu_{\max}$ we need the $\exp\,-\theta W(\beta'_1, \beta'_2)$ matrix of the configuration

	1	2	3	
0	x_1	0	β'_1
0	x_2	0	β'_2

with

$$W(\beta_1, \beta_2) = v(0, x_1)/2 + v(x_1, 0)/2 + v(0, x_2)/2 + v(x_2, 0)/2 + v(0, 0)/2 + v(0, 0)/2 + v(x_1, x_2) - C$$

$$- [n_+(x_1) - n_-(x_1) + n_+(x_2) - n_-(x_2)]/2.$$

This matrix is

$$\exp\,-\mathbf{W}\theta = \begin{matrix} & \begin{matrix} x_2\backslash x_1 & \quad 0 & \quad 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} e^{4J+2C} & e^C \\ e^C & 1 \end{pmatrix} \end{matrix}.$$

By the method of successive approximations (choosing the "ordered" vector $(1, 0)$ as the first approximation)

$$\begin{pmatrix} e^{4J+2C} & e^C \\ e^C & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = e^{4J+2C}\begin{pmatrix} 1 \\ e^{-4J-C} \end{pmatrix},$$

$$\begin{pmatrix} e^{4J+2C} & e^C \\ e^C & 1 \end{pmatrix}\begin{pmatrix} 1 \\ e^{-4J-C} \end{pmatrix} = (e^{4J+2C} + e^{-4J})\begin{pmatrix} 1 \\ e^{-4J-C} + e^{-8J-3C} \end{pmatrix},$$

until

$$_{1,0}\mu_{\max} = e^{4J+2C} + e^{-4J} + e^{-8J-2C} + \cdots.$$

To continue further we must know ${}_{2,0}\mu_{\max}$ and therefore develop the $\exp\,-\theta W(\beta'_1, \beta'_2)$ matrix of the configuration

$$0 \quad x_{11} \quad x_{12} \quad 0,$$
$$0 \quad x_{21} \quad x_{22} \quad 0.$$

Again all interactions except those between x's of adjacent rows are one-half their normal value. The matrix is

$x_{21}x_{22}\backslash x_{11}x_{12}$	00	01	10	11
00	e^{6J+3C}	e^{2J+2C}	e^{2J+2C}	e^C
01	e^{2J+2C}	e^{2J+C}	e^{C-2J}	1
10	e^{2J+2C}	e^{-2J+C}	e^{2J+C}	1
11	e^C	1	1	e^{2J-C}

$$= \exp\,-\mathbf{V}\theta.$$

Starting with the "ordered" vector $(1, 0, 0, 0)$ as a first approximation the method of successive approximations yields the unnormalized characteristic vector

$$\begin{bmatrix} 1 \\ e^{-4J-C} + e^{-8J-3C} + 2e^{-12J-5C} - e^{-12J-3C} + \cdots \\ e^{-4J-C} + e^{-8J-3C} + 2e^{-12J-5C} - e^{-12J-3C} + \cdots \\ e^{-6J-2C} + 2e^{-10J-4C} + e^{-10J-6C} + \cdots \end{bmatrix}$$

and the corresponding characteristic value $_{2,0}\mu_{\max}$ is

$$_{2,0}\mu_{\max} = e^{6J+3C} + 2e^{-2J+C} + 3e^{-6J-C} - 2e^{-10J-C} + 6e^{-10J-3C} + e^{-10J-5C} + \cdots.$$

From Eq. (23) it is apparent that

$$Z^{1/(M-1)} = S_{0, n+1}/S_{0, n}$$

and

$$Z^{1/(M-1)(N-1)} = {}_{n+1, 0}\mu_{\max}/{}_{n, 0}\mu_{\max}$$

as $n \to \infty$.

Therefore as a first approximation

$$Z^{1/(M-1)(N-1)} \sim {}_{2,0}\mu_{\max}/{}_{1,0}\mu_{\max} = e^{2J+C}(1 + e^{-8J-2C} + 2e^{-12J-4C} + \cdots).$$

To find the next approximation we must calculate $_{3,0}\mu_{\max}$ from the matrix of the system

$$0 \quad x_{11} \quad x_{12} \quad x_{13} \quad 0,$$

$$0 \quad x_{21} \quad x_{22} \quad x_{23} \quad 0,$$

with all interactions except those between x's of adjacent rows having one-half their normal values. It turns out by method of successive approximations that

$$_{3,0}\mu_{\max} = e^{8J+4C} + 3e^{2C} + 5e^{-4C} + e^{-8J}(-3 + 12e^{-2C} + 3e^{-4C}) + \cdots.$$

Thus as a second approximation (neglecting the "1" in the terms $N-1$ and $M-1$ since N, $M \to \infty$)

$$Z^{1/NM} \sim {}_{3,0}\mu_{\max}/{}_{2,0}\mu_{\max} = e^{2J+C}[1 + e^{-8J-2C} + 2e^{-12J-4C} + e^{-16J-4C}(-3 + 6e^{-2C} + 2e^{-4C}) + \cdots]. \quad (33a)$$

In the absence of an external magnetic field, $C = 0$ and

$$Z_0^{1/MN} = e^{2J}(1 + e^{-8J} + 2e^{-12J} + 5e^{-16J} + \cdots), \quad (33b)$$

which agrees with the results of Wannier and Kramers, KW-II.

For higher approximations it is necessary to find $_{4,0}\mu_{\max}$, $_{5,0}\mu_{\max}$, etc. by the same methods used for smaller values of n. In field free space $_{4,0}\mu_{\max}$ was of a value to make the next term in the $Z^{1/MN}$ series $14e^{-18J}$. The low temperature magnetization per lattice point is

$$\mathfrak{M}/MN = \mu \partial \log Z/\partial C$$

$$= \mu[1 - 2e^{-8J-2C} - 8e^{-12J-4C}$$

$$+ 2e^{-16J-4C}(7 - 18e^{-2C} - 8e^{-4C}) + \cdots], \quad (34)$$

which shows that a two-dimensional "ferromagnet" is actually ferromagnetic in the temperature range for which (34) is valid, in the sense that its magnetization is greater than zero in the absence of an external magnetic field.

In zero external magnetic field and at low temperatures the average internal energy per lattice point is

$$\bar{E}/MN = -\epsilon(2 - 8e^{-8J}$$

$$-24e^{-12J} - 72e^{-16J} - \cdots) \quad (35a)$$

and the specific heat per lattice point is

$$C/kMN = 64J^2e^{-8J}(1 + 9e^{-4J}/2$$

$$+ 18e^{-8J} + \cdots). \quad (35b)$$
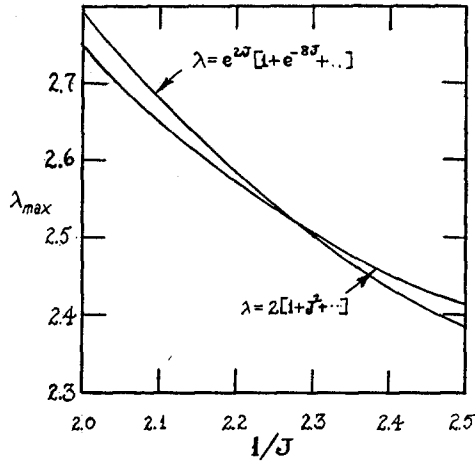
FIG. 4. λ_{max} as a function of $1/J$ in the neighborhood of J_c.

3. Discussion

Equations (32) and (33) provide us with expressions for the average energy and specific heat due to spin interactions at high and low temperatures. A reasonable problem now is to find the temperature ranges in which each expression is valid and to locate the temperature of a phase transition if one exists. In SM–I, p. 710, it is pointed out that a necessary condition for the existence of a phase transition at a temperature T_c is that the largest characteristic value of (4) be doubly degenerate. This condition of necessity results from supposing that the characteristic vector of (4) describing the most probable state of a layer of the system when $T > T_c$ is ψ_2 and when $T < T_c$ that it is ψ_1. Then if λ_1 is the characteristic value corresponding to ψ_1 and λ_2 is the characteristic value corresponding to ψ_2
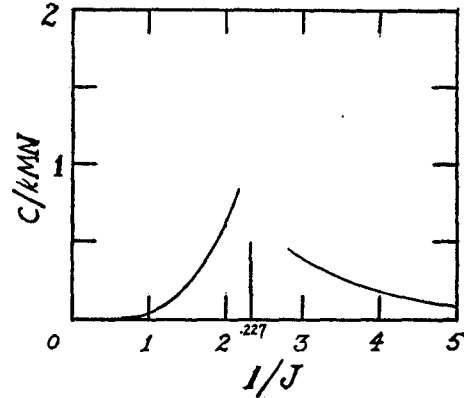
$$\lambda_1 = \lambda_{max} > \lambda_2 \text{ if } T < T_c,$$

$$\lambda_2 = \lambda_{max} > \lambda_1 \text{ if } T > T_c.$$

Since at a transition point both phases (if there are two) are equally probable

$$\lambda_1 = \lambda_2 = \lambda_{max} \text{ if } T = T_c.$$

Applying this criterion to the two-dimensional ferromagnet problem, if there is a phase transition it must occur at the temperature at which the $Z_0^{1/MN}$ of Eq. (33b) is equal to the $Z_0^{1/MN}$ of Eq. (31). That is, at the temperature which

FIG. 5. Specific heat per lattice point in units of k.

makes

$$2(1 + J^2 + 4J^4/3 + 77J^6/45 + \cdots)$$
$$= e^{2J}(1 + e^{-8J} + 2e^{-12J} + 5e^{-16J} + 24e^{-20J} + \cdots).$$

From Fig. 4 where the functions on both sides of this equation are plotted separately, the point of intersection of the two curves, $J_c = 0.44 = \epsilon/kT_c$, gives the temperature of a phase transition if one exists. At this point $Z_0^{1/MN} = 2.52$. Using a necessary condition that is a special property of a two-dimensional ferromagnet for the existence of a sharp phase transition, Kramers and Wannier KW–I, II have calculated $J_c = 0.4407$ and $Z_0^{1/MN} = 2.532$ which agrees closely with our results.

Unfortunately it is impossible from our equations to establish definitely the existence of a sharp phase transition at $J_c = 0.44$. Since $Z_0^{1/MN}$ as defined by (31) is equal to λ_{max} when $J > J_c$, (32) gives the correct asymptotic expression for the average internal energy and specific heat in this temperature range. When $J < J_c$ (35) is a valid asymptotic expression. If the specific heat is plotted as a function of J (Fig. 5) the curves seem to be approaching a peak of the λ-point variety at $J = J_c$. Inasmuch as the series in (32) and (35) are probably asymptotic divergent series in the neighborhood of J_c whose smallest terms (and therefore whose possible error[9]) are quite large, the calculated specific heat would probably be inaccurate around J_c. As has been mentioned before, the magnetization at zero external field is zero in the region of

[9] Knopp, *Theory and Applications of Infinite Series* (Blackie and Sons, 1928), p. 54.

validity of (30) while that in the region of validity of (34) is positive. Also, as $J \rightarrow J_c$, $\mathfrak{M}$, as defined by (34) decreases; but (34) does not have enough terms to decide whether or not it decreases continuously to zero at J_c. Approximate calculations of higher order terms indicate that the $\mathfrak{M}$ series is probably an asymptotic divergent one around J_c, so as in the case of the specific heat, this series gives there only a qualitative idea of the behavior of $\mathfrak{M}$. In spite of the inaccuracies near J_c, our results as do those of Kramers and Wannier seem to suggest the existence of a λ-point phase transition at J_c that accompanies a change from a ferromagnetic state to an unferromagnetic state.

In conclusion, it seems apparent that the characteristic value method of treating nearest neighbor problems is adequate to give asymptotic formulae for physical properties at high and low temperatures and to locate the temperature of a phase transition if one exists, but at its present state of development the method cannot provide accurate values of physical quantities which are derivatives of the partition function, in the immediate neighborhood of a transition point.

The author wishes to thank Professor Lars Onsager for his enlightening discussions concerning this problem.

APPENDIX[10]

Proof that $S_{2l+1}/S_{2l} = S_{l+1}/S_l + O(\theta^{l+1})$

By definition

$$S_l = \sum_{\alpha_1} \cdots \sum_{\alpha_l} \prod_{i=1}^{l-1} \exp - \theta V(\alpha_i, \alpha_{i+1}).$$

Therefore, if $\{\lambda_i\}$ is the set of characteristic values corre-

---

[10] The notation $g(\theta) = O[f(\theta)]$ as $\theta \rightarrow 0$ is used here in the usual sense that there exists some positive finite constant A independent of θ such that as $\theta \rightarrow 0$, $|g(\theta)/f(\theta)|$ finally becomes and remains less than A.

sponding to the set of characteristic vectors $\{\psi_j\}$ of the set of linear equations

$$\lambda \psi(\alpha_1) = \sum_{\alpha_2=1}^{\nu^{MN}} \psi(\alpha_2) \exp - \theta V(\alpha_1, \alpha_2)$$

and if there are ν configurations available to each lattice point

$$S_l = \sum_{j=1}^{\nu^{MN}} \lambda_j{}^{l-1} [\Sigma_\alpha \psi_j(\alpha)]^2,$$

$$S_{2l+1} S_l = \Sigma_i \lambda_i{}^{3l} [\Sigma_\alpha \psi_1(\alpha)]^2 + \Sigma'_{i,j} \lambda_i{}^{2l} \lambda_j{}^{l-1} [\Sigma_\alpha \psi_i(\alpha)]^2 [\Sigma_\alpha \psi_j(\alpha)]^2$$

(the prime on the Σ is introduced to denote that the terms with $i=j$ are omitted). Using a similar expression for $S_{2l} S_{l+1}$

$$S_{2l+1} S_l - S_{2l} S_{l+1}$$
$$= \sum_{i,l=1}^{\nu^{MN}} [\lambda_i{}^{2l} \lambda_j{}^{l-1} - \lambda_i{}^{2l-1} \lambda_j{}^l][\Sigma_\alpha \psi_i(\alpha)]^2 [\Sigma_\alpha \psi_j(\alpha)]^2. \quad (36)$$

Suppose $\lambda_i = \Sigma a_{ij} \theta^j$; $\lambda_1 = \lambda_{max}$. Then as $\theta \rightarrow 0$ in the matrix $\exp - \theta V$, $\lambda_i \rightarrow a_{i0}$. Since as $\theta \rightarrow 0$ all the elements of $\exp - \theta V$ approach unity it is easily verified that $\lambda_1 = \lambda_{max} \rightarrow \nu^{MN}$ and $\lambda_i \rightarrow 0$ if $i>1$, or $\lambda_i = 0(\theta)$. Also, when $i>1$, either $\psi_i(\alpha) = 0(\theta)$ or for every α which makes $\psi_i(\alpha) = c$ there exists a unique β which makes $\psi_i(\beta) = -c$; for if a single α yielded $\psi_i(\alpha) = c$ and there existed no corresponding β to make $\psi_i(\beta) = -c$,

$$\lambda_i = \sum_{\alpha_1 \alpha_2} \psi_i(\alpha_1) \psi_i(\alpha_2) \exp - \theta V(\alpha_1 \alpha_2)$$

$$= c^2 + 0(\theta)$$

contrary to our above conclusions. This implies

$$\Sigma_\alpha \psi_i(\alpha) = \begin{cases} 0(1) & \text{if } i=1 \\ 0(\theta) & \text{if } i>1. \end{cases}$$

Since all $\lambda_i = 0(\theta)$ except λ_1 which is $0(1)$, the dominant term in (36) as $\theta \rightarrow 0$ is

$$[\lambda_1{}^{2l} \lambda_j{}^{l-1} - \lambda_1{}^{2l} \lambda_j{}^l][\Sigma_\alpha \psi_1(\alpha)]^2 [\Sigma_\alpha \psi_j(\alpha)]^2 \quad (j \neq 1)$$
$$= [0(1)0(\theta^{l-1}) - 0(1)0(\theta^l)][0(1)]^2 [0(\theta)]^2$$
$$= 0(\theta^{l+1}) - 0(\theta^{l+2}) \rightarrow 0(\theta^{l+1}).$$

So we finally have

$$S_{2l+1} S_l - S_{2l} S_{l+1} = 0(\theta^{l+1})$$

or

$$S_{2l+1}/S_{2l} - S_{l+1}/S_l = 0(\theta^{l+1})/S_l S_{2l} = 0(\theta^{l+1}),$$

since $S_l S_{2l} \rightarrow 0(1)$ as $\theta \rightarrow 0$, and our theorem is proven.