PHYSICAL REVIEW VOLUME 64, NUMBERS 5 AND 6 SEPTEMBER 1 AND 15, 1943

The Propagation of Order in Crystal Lattices*

J. Ashkin and W. E. Lamb, Jr.
Pupin Physics Laboratories, Columbia University, New York, New York

(Received June 10, 1943)

Zernike's problem of the propagation of order in a binary crystal alloy is discussed by means of the matrix formalism recently developed for treating cooperative phenomena. It is proved generally that the existence of long range order over any temperature range implies a degeneracy of the maximum characteristic value of the fundamental matrix over the same temperature range. For a two-dimensional crystal, the special form of the matrix theory developed by Kramers and Wannier has been used to obtain explicitly the probabilities for finding an A or B atom at any lattice site if it is known that there is an A (or B) atom at a certain site (correlation probabilities, or intermediate range order). The results are obtained in the form of power series valid at low temperatures by a perturbation treatment of the highest characteristic values and characteristic vectors. It is seen that the maximum characteristic value is doubly degenerate from $T=0$ to some finite temperature, consistent with the existence of long range order at low temperature. The absence of long range order at high temperatures is proved by showing that the maximum characteristic value is non-degenerate at sufficiently high temperatures. By comparing the solutions of Zernike's approximate equations for the correlation probabilities in two dimensions with our exact solution, and his expressions for the long range order and energy (short range order) in three dimensions with those given by van der Waerden, we find that Zernike's approximation is better for a two- than a three-dimensional crystal. The problem of propagation of order is generalized by an investigation of the correlation probabilities for more complicated configurations of fixed atoms. For example, the ordering influence of an adjacent pair of disordered atoms ("dipole") is found as a function of position in the lattice. It is found here, as might be expected, that the ordering influence falls off rapidly with the distance of the site from the dipole. Other cases are treated as well.

In order to interpret the diffuse scattering of x-rays by a partially disordered crystal alloy it is necessary to know how the presence of one atom at a certain lattice point affects the probability of finding the various kinds of atoms at the other lattice points. The older theories of the order-disorder problem, principally those of Bragg and Williams,[1] and Bethe,[2] gave only the long and short range order and did not consider the more general problem of finding the correlation proabilities for *all* distances, intermediate as well as long and short. Experimentally, several of these correlation probabilities were measured by Wilchinsky[3] for the alloy Cu_3Au, while the problem of calculating them theoretically was first attacked by Zernike[4] in a paper entitled "The Propagation of Order in Cooperative Phenomena."

In Zernike's work, the propagation of order is described by a partial finite difference equation relating the probability of finding a certain kind of atom at a given lattice point with the probabilities for various atoms at the nearest neighboring lattice points. However, his equation is necessarily approximate because in order to obtain a recurrence relation, he has to treat the probabilities for the nearest neighboring atoms as at least formally independent of one another. Nevertheless, this method is so far the only really practicable one for the three-dimensional propagation problem and gives reasonable results for the thermal properties and the correlation probabilities at low and high temperatures. One of the results of the present investigation will be to obtain a comparison in two dimensions between Zernike's solution and the exact solution valid at low temperatures. In three dimensions, exact series expansions for the long range order and the energy at low temperatures have been given by van der Waerden[5] and have been compared with the series of Zernike. The comparison indicates that Zernike's approximation is better in two dimensions than it is in three. In one dimension Zernike's solution can be shown to be exact. It therefore appears that the approximation

---

* Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University, New York, New York. Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] W. L. Bragg and E. J. Williams Proc. Roy Soc. London **145A**, 699 (1934).

[2] H. A. Bethe, Proc. Roy Soc. **150A**, 552 (1935).

[3] Z. W. Wilchinsky, Phys. Rev. **63**, 223 (1943).

[4] F. Zernike, Physica **7**, 565 (1940).

[5] v. d. Waerden, Zeits. f. Physik **118**, 473 (1941).

becomes successively poorer as the dimension is increased.

Although there was no further work directly following up that of Zernike, a new technique has been introduced into the study of the order-disorder phenomenon which shows some promise of providing a solution to the propagation problem. Recent work by Montroll,[6] Kramers and Wannier,[7] Lassettre and Howe,[8] and Onsager,[9] has shown how the statistical treatment of binary alloys can in principle be reduced to the solution of the characteristic value problem of a certain matrix, the characteristic values being related to the partition function for the crystal and the characteristic vectors to the probability for finding various configurations of the atoms in the lattice. It is the object of this paper to show how these matrix methods can be used to give at least a partial solution to the propagation problem.

The application of the matrix theory to real three-dimensional crystals has been held up unfortunately by considerable mathematical difficulties. The matrix whose characteristic value problem is to be solved is very difficult to handle even by perturbation methods, the essential complication arising from the fact that all its elements are different from zero and do not appear to vary in any regular manner from one position in the matrix to the next. These difficulties have so far compelled us to restrict our investigations to two-dimensional crystals, where the general matrix theory has been developed by Kramers and Wannier in a more tractable form. Using a perturbation method, valid at low temperatures, we have obtained power series for the maximum characteristic value and for the corresponding characteristic vectors. By inserting these series into the appropriate formalism of the matrix theory we then discuss the propagation problem by giving the correlation probabilities as functions of temperature and position in the lattice, obtaining series for the long range order, the short range order, and the order of intermediate range. The treatment is restricted to low temperatures, however, the perturbation theory at high temperatures being very complicated and difficult to carry through.

The paper is divided into two parts. Part I gives a brief outline of the matrix theory in three dimensions and contains a simple demonstration of the fact that states of long range order are not possible over a range of temperatures unless the maximum characteristic value of the associated matrix is degenerate over this same range. Part II is devoted to the discussion of the propagation problem in two-dimensional crystals.

I. LONG RANGE ORDER AND THE MATRIX METHOD

1. Results of the General Matrix Theory in Three Dimensions

In this section we shall collect the results of the general matrix theory which are applicable to the problem under consideration. The notation is identical with that used by Montroll.[6]

Consider a binary substitutional alloy AB whose crystal form is simple cubic. Let the edges of the crystal be L, M, N in units of the lattice distance. The crystal can then be regarded as made up of L layers of $M\times N$ atoms each, with each layer capable of 2^{MN} distinct configurations, corresponding to the fact that each of the $M\times N$ sites in the layer can be occupied by an A or a B atom. With the symbol α_j used for a typical one of the configurations available to the atoms in the jth layer, $v(\alpha_j)$ is to represent the total potential energy of interaction between all the nearest neighboring atoms of the jth layer in configuration α_j, and $v(\alpha_j, \alpha_{j+1})$ is taken as the representative of the total energy of interaction between atoms in the jth layer in configuration α_j with their nearest neighboring atoms in the $(j+1)$th layer in configuration α_{j+1}. With this notation and with the periodic boundary conditions imposed by requiring layer 1 to interact with layer L,[10] the partition function has the form,

$$Z=\sum_{\{\alpha_1\}}\cdots\sum_{\{\alpha_L\}}\exp -\{v(\alpha_1)+\cdots+v(\alpha_L)+v(\alpha_1, \alpha_2)+\cdots+v(\alpha_{L-1}, \alpha_L)+v(\alpha_L, \alpha_1)\}/kT. \quad (1.1)$$

---

[6] E. W. Montroll, J. Chem. Phys. 9, 706 (1941).

[7] H. A. Kramers and G. H. Wannier, Phys. Rev. 60, 252 (1941).

[8] E. N. Lassettre and J. P. Howe, J. Chem. Phys. 9, 747 (1941); and 9, 801 (1941).

[9] Onsager's work is not yet published. He has, among other things, obtained a closed expression for the partition function of the two-dimensional lattice.

[10] This restriction is not necessary, serving merely to simplify the final formulae.

Making the substitution

$$V(\alpha_i,\alpha_j)=\tfrac{1}{2}v(\alpha_i)+v(\alpha_i,\alpha_j)+\tfrac{1}{2}v(\alpha_j),$$

we find that (1.1) becomes

$$Z=\sum_{\{\alpha_1\}}\cdots\sum_{\{\alpha_L\}}\left\{\prod_{i=1}^{L}\exp\left(-V(\alpha_i,\alpha_{i+1})/kT\right)\right\}, \quad (1.2)$$

with the understanding that

$$\alpha_{L+1}\equiv\alpha_1.$$

Writing

$$\mathfrak{M}(\alpha,\alpha')=\exp\left[-V(\alpha,\alpha')/kT\right] \quad (1.3)$$

we define a symmetrical matrix $\mathfrak{M}$ whose $2^{MN}\times 2^{MN}$ rows and columns are labeled by $\alpha,\alpha',\alpha'',\cdots$ in accordance with the 2^{MN} possible configurations α. The characteristic values λ_r and characteristic vectors ψ_r ($r=1,2,\cdots,2^{MN}$) of this matrix $\mathfrak{M}$ are fundamental for the theory. For, as we shall see, if we can solve the characteristic value problem

$$\mathfrak{M}\psi=\lambda\psi, \quad (1.4)$$

we shall have the answer to all questions of thermodynamic and statistical interest for the crystal, including the order-disorder phenomenon.

A characteristic vector ψ_r has 2^{MN} components, $\psi_r(\alpha)$, one for each configuration α. If we assume the λ_r and $\psi_r(\alpha)$ are known, the matrix element (3) can be written as

$$\exp\left[-V(\alpha,\alpha')/kT\right]=\sum_{r=1}^{2^{MN}}\lambda_r\psi_r(\alpha)\psi_r(\alpha'), \quad (1.5)$$

provided that the ψ_r are normalized to unity;

$$\textstyle\sum_\alpha\psi_r(\alpha)\psi_s(\alpha)=\delta_{rs}. \quad (1.6)$$

Substituting (1.5) into (1.2) and applying (1.6) one can easily prove that

$$Z=\sum_{r=1}^{2^{MN}}\lambda_r{}^L.$$

Thus the evaluation of Z, and consequently the investigation of thermodynamic quantities like entropy, energy, and specific heat, are reduced to the characteristic value problem (1.4). Since the number of layers L is very large for an actual crystal, the problem (1.4) can be restricted to that of obtaining the highest characteristic value $\lambda_{\max}$ and its corresponding characteristic vectors. If $\lambda_{\max}$ is d-fold degenerate[11] (1.7) becomes[12]

$$Z=d\cdot\lambda_{\max}^{L}. \quad (1.8)$$

In the work of Montroll, and Kramers and Wannier it was assumed that $\lambda_{\max}$ was non-degenerate for all temperatures with the exception of possible isolated points. It was first pointed out by Lassettre and Howe[8] in their papers on binary solid solutions, that the maximum characteristic value is actually doubly degenerate for a range of temperatures beginning at $T=0$, and they were able to demonstrate that within this range, a separation into two phases takes place. In the next section we shall show directly that this degeneracy of $\lambda_{\max}$ is a necessary condition for the existence of long range order.

2. Long Range Order and the Degeneracy of $\lambda_{\max}$

Suppose that a state of long range order exists in the AB alloy and that it is known that the atoms in layer 1 have a certain configuration α_1. Then the probability for finding the atoms of layer n, very far from layer 1, in a certain configuration α_n, should depend explicitly on the configuration α_1. We shall show that if $\lambda_{\max}$ is non-degenerate, the probability for the configuration α_n is entirely independent of what configuration exists in layer 1, so that the non-degeneracy of $\lambda_{\max}$ will imply the absence of long range order.

The probability $P(\alpha_1,\alpha_n)$ of finding simultaneously the configuration α_1 in layer 1 and α_n in layer n, is given by an expression similar to (1.2) where, however, the configurations α_1 and α_n of layers 1 and n are held fixed. Using (1.5),

---

[11] Although $\lambda_{\max}$ itself may not be degenerate, there may exist other characteristic values λ whose negatives are equal to $\lambda_{\max}$. As is proved for a special case in part II, L will be even under such circumstances. A formula identical with (1.8) will then apply with d having an obvious meaning. In the future a statement that $\lambda_{\max}$ is degenerate will be understood to include the possibility mentioned here.

[12] It must be assumed here that the sum of the Lth powers of the remaining characteristic values is negligible compared with $\lambda_{\max}^L$. Thus even if the ratios of the other characteristic values to $\lambda_{\max}$ are small, say less than $1-x$ with x close to unity, there are approximately 2^{MN} of them and there is no assurance that $2^{MN}(1-x)^L$ is small compared to unity.

(1.6), and (1.7) we find

$$P(\alpha_1, \alpha_n) = \frac{1}{\sum_r \lambda_r^L} \sum_{s=1}^{2^{MN}} \sum_{t=1}^{2^{MN}} \lambda_s^{n-1} \lambda_t^{L-n+1} \times \psi_s(\alpha_1)\psi_s(\alpha_n)\psi_t(\alpha_n)\psi_t(\alpha_1). \quad (2.1)$$

To discuss the long range order, n and $L-n$ must be taken as very large integers, the latter condition being necessary because of the boundary condition that layers 1 and L interact. This permits us to neglect all the characteristic values with the exception of $\lambda_{\max}$, and if the latter is assumed to be non-degenerate, (2.1) reduces to

$$P(\alpha_1, \alpha_n) = \psi_{\max}(\alpha_1)^2 \cdot \psi_{\max}(\alpha_n)^2. \quad (2.2)$$

If $P(\alpha_1, \alpha_n)$ is summed over all possible configurations α_n, we obtain the probability $P(\alpha_1)$ for finding layer 1 in configuration α_1. Since $\psi_{\max}$ is normalized to unity, this summation gives $P(\alpha_1) = \psi_{\max}(\alpha_1)^2$. Similarly, $P(\alpha_n) = \psi_{\max}(\alpha_n)^2$. Thus $P(\alpha_1, \alpha_n) = P(\alpha_1) \cdot P(\alpha_n)$. The probability for finding the configuration α_n when it is known that the atoms of layer 1 are in configuration α_1, is equal to $P(\alpha_1, \alpha_n)$ divided by $P(\alpha_1)$. This quotient is completely independent of α_1, being just equal to the probability for finding α_n when no restrictions at all are put on the other layers. There is therefore no long range order under these circumstances.

A further remark[13] concerning the interpretation of this result must be made, in the light of a theorem of Frobenius to the effect that if all the elements of a matrix are positive, the maximum characteristic value is likewise positive and in addition is non-degenerate. The elements of the matrix $\mathfrak{M}$, namely, exp $[-V(\alpha, \alpha')/kT]$, are all greater than zero and will evidently remain bounded away from zero so long as M and N remain finite. Consequently, for any crystal in which the dimension L is very large compared with the dimensions M and N, $\lambda_{\max}^L$ will exceed by far the Lth power of any other characteristic value and states of long range order will not exist. On the other hand, if at least one of M and N is very large, and comparable to L in magnitude, it may happen that some of the matrix elements of $\mathfrak{M}$ are close enough to zero to cause a practical, if not actual, degeneracy in $\lambda_{\max}$; that is, there may exist other characteristic values λ, whose Lth powers are not appreciably less than $\lambda_{\max}^L$. If $\lambda_{\max}$ is d-fold degenerate, (2.1) becomes

$$P(\alpha_1, \alpha_n) = \frac{1}{d} \sum_{s=1}^{d} \sum_{t=1}^{d} \psi_{\max, s}(\alpha_1)\psi_{\max, s}(\alpha_n) \times \psi_{\max, t}(\alpha_n)\psi_{\max, t}(\alpha_1), \quad (2.3)$$

where the $\psi_{\max, s}$ are the characteristic vectors corresponding to $\lambda_{\max}$. Long range order will then in general be possible, since $P(\alpha_1, \alpha_n)$ no longer factors.

II. THE PROPAGATION OF ORDER IN TWO-DIMENSIONAL LATTICES

As was already pointed out in the introduction, the characteristic value problem (1.4) for three dimensions, is very difficult to handle. Therefore, in the hope of obtaining useful qualitative information which might possibly be carried over into three dimensions, we turn to the two-dimensional case where the corresponding problem is solvable. The matrix theory in two dimensions was developed mainly by Kramers and Wannier[7] in a paper on the statistics of two-dimensional ferromagnets.

In Section 3 we write down the matrix of Kramers and Wannier and review the properties of it which we shall find useful. Section 4 is devoted to a perturbation theory of the characteristic value problem, the main results of which are to show that the maximum characteristic value is doubly degenerate from $T=0$ to some finite temperature, and to obtain power series for the maximum characteristic value and the corresponding characteristic vectors valid at low temperatures. In the remaining sections, the solution is applied to various problems of physical interest. The long range order as a function of the temperature is obtained in Section 5, and in Section 6 a proof is given that an order-disorder transition actually occurs in the two-dimensional crystal. In Section 7 we discuss the general problem of the propagation of order through the lattice, showing in detail the gradual transition from the short range order to the influence at very great distances. In Section 8 Zernike's approximation is compared with the

---

[13] E. W. Montroll, J. Chem. Phys. 9, 711 (1941).

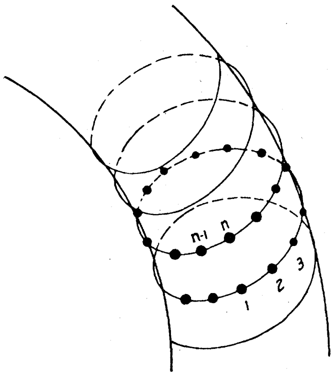

FIG. 1. A two-dimensional crystal on the surface of a torus.

exact solution. In Section 9, we treat the ordering influence of more complicated configurations of fixed atoms.

3. The Matrix of Kramers and Wannier

Kramers and Wannier have introduced a simplification into the theory of two-dimensional square lattices by making use of a special connectivity scheme for their crystal. This scheme can best be described as being equivalent to having the lattice sites regularly distributed along a continuous line twisting its way in screw-like fashion over the surface of a torus as in Fig. 1.

m pitches of n atoms each in this case take the place of L layers of $M\times N$ atoms each in the three-dimensional case. For convenience the atoms are distinguished by $+$ and $-$ instead of A and B, coordinates $\mu_1, \mu_2, \cdots, \mu_{mn}$ capable of assuming the values $+1$ and -1 being assigned to the respective lattice sites. The configuration of the atoms on a pitch will thus be given by a set of n values $(\mu_n\cdots\mu_2\mu_1)$. which we denote by α for brevity. In the magnetic case μ represents a spin capable of two orientations; for the case of alloys, the μ_j are still convenient for describing the various configurations. The energies of interaction between neighboring pairs are written $V(++)=V(--)=-\frac{1}{2}J$ and $V(+-)=\frac{1}{2}J$, so that the total energy is expressible as

$$E=-\tfrac{1}{2}J\sum \mu_i\mu_k, \tag{3.1}$$

where the sum is over all pairs of nearest neighbors. If J is positive, like atoms will attract one another and the ordered arrangement at low temperatures will be one in which the $+$ and $-$ atoms are separated into two phases. We shall then say that we are in the "ferromagnetic case." The "antiferromagnetic case," where J is negative, will lead to ordered states at low temperatures in which $+$ and $-$ atoms alternate on the lattice sites.

As in the three-dimensional case, a matrix $\mathfrak{M}$ is found whose characteristic value problem is fundamental for the statistical theory. $\mathfrak{M}$ has $2^n\times 2^n$ rows and columns labeled by the 2^n different possible configurations of the atoms on a pitch, $\alpha, \alpha', \alpha'', \cdots$. However, the matrix is not symmetrical, and it thus becomes necessary to introduce right- and left-handed characteristic vectors, say $\mathbf{A}_q$ and $\tilde{\mathbf{B}}_p$, where

$$\mathfrak{M}\mathbf{A}_q=\lambda_q\mathbf{A}_q \quad\text{and}\quad \tilde{\mathbf{B}}_p\mathfrak{M}=\lambda_p\tilde{\mathbf{B}}_p. \tag{3.2}$$

With proper normalization the $\mathbf{A}_q$ and $\tilde{\mathbf{B}}_p$ can be made to satisfy orthogonality conditions.

$$\tilde{\mathbf{B}}_p\mathbf{A}_q=\sum\nolimits_\alpha \mathbf{B}_p(\alpha)\mathbf{A}_q(\alpha)=\delta_{pq}. \tag{3.3}$$

It actually turns out that the analogue of the matrix (1.3) in the three-dimensional case, is not $\mathfrak{M}$, but the nth power of $\mathfrak{M}$. Thus one finds, upon expressing the matrix element $(\alpha_1|\mathfrak{M}^n|\alpha_2)$ as a bilinear combination of the components $\mathbf{A}_q(\alpha_2)$ and $\mathbf{B}_q(\alpha_1)$.

$$\sum_{q=1}^{2^n} \lambda_q{}^n\mathbf{A}_q(\alpha_2)\mathbf{B}_q(\alpha_1)=\exp\left[-\frac{v(\alpha_2)+v(\alpha_1,\alpha_2)}{kT}\right], \tag{3.4}$$

where $v(\alpha_2)$ is the total energy of interaction between the nearest neighboring atoms of pitch 2 and $v(\alpha_1,\alpha_2)$ is the energy of interaction between the pitches 1 and 2. This relation, just as (1.5), plays a fundamental role.

By repeated use of (3.4) and (3.3) one finds for the partition function, the expression

$$Z=\sum_{p=1}^{2^n}\lambda_p{}^{mn} \tag{3.5}$$

and it appears, by comparison with (1.7), that λ^n is the analogue of the λ of the three-dimensional case. The advantage of this result of Kramers and Wannier is that now λ has a meaning which is independent of the size of the crystal, the logarithm of $\lambda_{\max}$ being essentially the free energy per particle. The variation of the properties of the crystal as it becomes infinite in two directions can therefore be studied conveniently by looking at $\lambda_{\max}$.

To write the matrix $\mathfrak{M}$ as a square array, it is necessary to arrange the 2^n configurations $\alpha=(\mu_n\cdots\mu_2\mu_1)$, in some definite order. The method chosen by Kramers and Wannier is the following: for any configuration, e.g., $++--++-+$, one replaces every $+$ by a 0 and every $-$ by a 1 (giving in our example 00110010), reads the resulting number in the dual number system and takes that as the order number of the configuration. For the configuration above the order number is 50. It proves convenient to separate the configurations into two classes in accordance with the sign of μ_n. The configurations of the class $\mu_n=1$ consist of those with order numbers 0, 1, $\cdots$, $2^{n-1}-1$ and are arranged in that order. The remaining half of the configurations belong to the class $\mu_n=-1$ and are arranged in the order 2^n-1, 2^n-2, $\cdots$, 2^{n-1}. The reason for the arrangement is the following. Configurations α and $\bar{\alpha}$ in corresponding places in the two classes have order numbers which add up to 2^n-1 and are conjugate to one another in the sense that α can be obtained from $\bar{\alpha}$ by changing every $+$ to a $-$ and vice versa.

Arranging the rows and columns to correspond to the arrangement of the configurations in the two classes, and introducing the new variable K and the functions α and β by

$$K=\frac{1}{2}\frac{J}{kT}, \quad \alpha=e^{2K}, \quad \beta=e^{-2K}$$

the matrix $\mathfrak{M}$ has the form

$$\mathfrak{M}(K)=\left\|\begin{array}{cccc|cccc}
\alpha 1 & & & & & & & \\
 & \alpha 1 & & & & & & \\
 & & \ddots & & & & & \\
 & & & \alpha 1 & & & & 00 \\
 & & & 00 & & & & \beta 1 \\
 & & & & & & \iddots & \\
 & & & & & \beta 1 & & \\
 & & & & \beta 1 & & & \\
\hline
 & & & & \alpha 1 & & & \\
 & & & & & \alpha 1 & & \\
 & & & & & & \ddots & \\
 & & & 00 & & & & \alpha 1 \\
 & & & \beta 1 & & & & 00 \\
 & & \iddots & & & & & \\
 & \beta 1 & & & & & & \\
\beta 1 & & & & & & &
\end{array}\right\|, \tag{3.6}$$

where blank spaces indicate zeros. $\mathfrak{M}(K)$ is easily seen to be reducible. Thus, with a matrix $\mathbf{H}$ defined by

$$\mathbf{H}=\frac{1}{\sqrt{2}}\left\|\begin{array}{c|c} \mathbf{I} & \mathbf{I} \\ \hline \mathbf{I} & -\mathbf{I} \end{array}\right\|. \tag{3.7}$$

$\mathbf{I}$ being the identity matrix of order 2^{n-1}, we have

$$\mathbf{H}\mathfrak{M}(K)\mathbf{H}^{-1}=\left\|\begin{array}{c|c} \mathbf{V}_+(K) & 0 \\ \hline 0 & \mathbf{V}_-(K) \end{array}\right\|, \tag{3.8}$$

where

$$\mathbf{V}_+(K)=\left\|\begin{array}{ccccc}
\alpha 1 & & & & \\
 & \alpha 1 & & & \\
 & & \alpha 1 & & \\
 & & & \ddots & \\
 & & & & \alpha 1 \\
 & & & & \beta 1 \\
 & & & \iddots & \\
 & & \beta 1 & & \\
 & \beta 1 & & & \\
\beta 1 & & & &
\end{array}\right\|,$$

$$\mathbf{V}_-(K)=\left\|\begin{array}{ccccc}
\alpha\;\;1 & & & & \\
 & \alpha\;\;1 & & & \\
 & & \alpha\;\;1 & & \\
 & & & \ddots & \\
 & & & & \alpha\;\;1 \\
 & & & & -\beta-1 \\
 & & & \iddots & \\
 & & -\beta-1 & & \\
 & -\beta-1 & & & \\
-\beta-1 & & & &
\end{array}\right\|. \tag{3.9}$$

We shall call $\mathbf{V}_+(K)$ the "plus $\mathbf{V}$ matrix" and $\mathbf{V}_-(K)$ the "minus V matrix." The characteristic values and characteristic vectors of $\mathfrak{M}(K)$ fall into two classes. If $\psi_+(K)$ and $\psi_-(K)$ are right-handed characteristic vectors of $\mathbf{V}_+(K)$ and $\mathbf{V}_-(K)$, respectively, corresponding to the characteristic values $\lambda_+(K)$ and $\lambda_-(K)$, the associated characteristic vectors of $\mathfrak{M}(K)$ are

$$\mathbf{A}_+(K)=\mathbf{H}^{-1}\left\|\begin{array}{c} \psi_+(K) \\ 0 \end{array}\right\|=\left\|\begin{array}{c} \psi_+(K) \\ \psi_+(K) \end{array}\right\|,$$

$$\mathbf{A}_-(K)=\mathbf{H}^{-1}\left\|\begin{array}{c} 0 \\ \psi_-(K) \end{array}\right\|=\left\|\begin{array}{c} \psi_-(K) \\ -\psi_-(K) \end{array}\right\|. \tag{3.10}$$

The components are arranged in the order of the two groups $\mu_n = 1$ and $\mu_n = -1$. Similarly, if $\tilde{\phi}_+(K)$ and $\tilde{\phi}_-(K)$ are the left-handed characteristic vectors of $\mathbf{V}_+(K)$ and $\mathbf{V}_-(K)$, the corresponding vectors for $\mathfrak{M}(K)$ are

$$\begin{aligned} \tilde{\mathbf{B}}_+(K) &= \|\tilde{\phi}_+(K),\ \tilde{\phi}_+(K)\|, \\ \tilde{\mathbf{B}}_-(K) &= \|\tilde{\phi}_-(K),\ -\tilde{\phi}_-(K)\|. \end{aligned} \tag{3.11}$$

As a final property of $\mathfrak{M}(K)$ which should be mentioned, there is a simple device by which one can treat the antiferromagnetic case once the ferromagnetic case has been discussed. In the former case, the interaction energy J is negative and the parameter K is therefore negative. In the ferromagnetic case, $K > 0$. If $\mathbf{R}$ is a permutation matrix which changes every alternate atom on a pitch from + to − and vice versa, it is readily verified that

$$\mathbf{R}\mathbf{H}\mathfrak{M}(-K)\mathbf{H}^{-1}\mathbf{R}^{-1} = \left\| \begin{array}{c|c} \mathbf{V}_+(K) & 0 \\ \hline 0 & -\mathbf{V}_-(K) \end{array} \right\|, \tag{3.12}$$

so that the characteristic values of $\mathfrak{M}(-K)$ are $\lambda_+(K)$ and $-\lambda_-(K)$. The characteristic vectors $\mathbf{A}_+(-K)$, etc., are obtained from $\mathbf{A}_+(K)$, $\tilde{\mathbf{B}}_+(K)$, $\cdots$ by permuting the components according to $\mathbf{R}$.

The problem with which we are now faced is that of finding the characteristic values and characteristic vectors of the plus and minus $\mathbf{V}$ matrices as functions of the temperature. Kramers and Wannier have neglected the minus $\mathbf{V}$ matrix in their treatment, but this is unimportant for the discussion of the thermal properties. We are, however, interested in the degeneracy of $\lambda_{\max}$ and shall find in the next section that the maximum characteristic value of $\mathbf{V}_-$ is equal to the maximum characteristic value of $\mathbf{V}_+$ in the limiting case of an infinitely large crystal.

4. Solution of the Matrix Problem for Low Temperatures

For the low temperature ferromagnetic case, K is very large and positive, so that the parameter $\beta = e^{-2K}$ is very small, approaching 0 as the temperature tends to zero. It then becomes possible to carry out an expansion in powers of β. Thus, by defining a new matrix, $\mathbf{U}_+(\beta)$, the plus $\mathbf{V}$ matrix can be written

$$\mathbf{V}_+(K) = \alpha \begin{Vmatrix} 1\ \beta & & & & \\ & 1\ \beta & & & \\ & & \ddots & & \\ & & & & 1\ \beta \\ & & & & \beta^2\beta \\ & & & \cdot^{\cdot^{\cdot}} & \\ & & \beta^2\beta & & \\ \beta^2\beta & & & & \end{Vmatrix} \equiv \frac{1}{\beta}\mathbf{U}_+(\beta). \tag{4.1}$$

$\mathbf{U}_+(\beta)$ can be expanded in the form

$$\mathbf{U}_+(\beta) = \begin{Vmatrix} 10 & & & & \\ & 10 & & & \\ & & \ddots & & \\ & & & 10 & \\ & & & & 10 \\ & & & & \\ & & & & \\ & & & & \end{Vmatrix} + \beta \begin{Vmatrix} 01 & & & & \\ & 01 & & & \\ & & \ddots & & \\ & & & 01 & \\ & & & & 01 \\ & & & & 01 \\ & & & 01 & \\ & & \cdot^{\cdot^{\cdot}} & & \\ & 01 & & & \\ 01 & & & & \end{Vmatrix} + \beta^2 \begin{Vmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & 10 \\ & & & 10 & \\ & & \cdot^{\cdot^{\cdot}} & & \\ & 10 & & & \\ 10 & & & & \end{Vmatrix}. \tag{4.2}$$

Defining the constant matrices $\mathbf{U}_0$, $\mathbf{U}_1$, and $\mathbf{U}_2$ by (4.2) we may write for brevity

$$\mathbf{U}_+(\beta) = \mathbf{U}_0 + \beta\mathbf{U}_1 + \beta^2\mathbf{U}_2. \tag{4.3}$$

If there are n atoms per pitch, these matrices will have 2^{n-1} rows and columns corresponding to the configurations 0, 1, 2, $\cdots$, $2^{n-1}-1$ which all have $\mu_n=1$.

Our procedure will consist in obtaining power series in the parameter β for the maximum characteristic value $\lambda_{\max}$ and corresponding characteristic vector $\psi_{\max}$ of the matrix $\mathbf{U}_+(\beta)$. It is evident that $\mathbf{U}_0$ has a single non-degenerate characteristic value of unity and a $2^{n-1}-1$-fold degenerate value of zero. Therefore, $\lambda_{\max}$ will be well separated from the other characteristic values at low temperatures and will be an analytic function of β in the neighborhood of $\beta=0$. One will thus be justified in trying series expansions for $\lambda_{\max}$ and $\psi_{\max}$ of the form

$$\begin{aligned}\lambda&=\lambda_{\max}=\lambda_0+\beta\lambda_1+\beta^2\lambda_2+\cdots+\beta^p\lambda_p+\cdots,\\ \psi&=\psi_{\max}=\psi_0+\beta\psi_1+\beta^2\psi_2+\cdots+\beta^p\psi_p+\cdots.\end{aligned}\tag{4.4}$$

Inserting (4.4) into the equation $\mathbf{U}_+(\beta)\psi=\lambda\psi$, and equating coefficients of like powers of β, we find a set of equations for the λ_p and ψ_p:

$$\begin{aligned}&0)\quad U_0\psi_0=\lambda_0\psi_0,\\ &1)\quad U_0\psi_1+U_1\psi_0=\lambda_0\psi_1+\lambda_1\psi_0,\\ &2)\quad U_0\psi_2+U_1\psi_1+U_2\psi_0=\lambda_0\psi_2+\lambda_1\psi_1+\lambda_2\psi_0,\\ &p)\quad U_0\psi_p+U_1\psi_{p-1}+U_2\psi_{p-2}=\sum_{q=0}^{p}\lambda_q\psi_{p-q},\quad p\geqq 3.\end{aligned}\tag{4.5}$$

The characteristic vector ψ_0 corresponding to $\lambda_0=1$ evidently has only one non-vanishing component, that corresponding to configuration number zero. We therefore take $\psi_0(0)=1$ and $\psi_0(\alpha)=0$ for $\alpha=1, 2, \cdots, 2^{n-1}-1$. Thus for ψ_0 and its adjoint vector, $\tilde{\psi}_0$, we have

$$\psi_0=\left\|\begin{matrix}1\\0\\ \cdot\\ \cdot\\ \cdot\\0\end{matrix}\right\|,\quad \tilde{\psi}_0=\|\,1\ \ 0\ \cdots\ 0\,\|.\tag{4.6}$$

Since $\mathbf{U}_1\psi_0\equiv 0$, the equation for ψ_1 and λ_1 is

$$\mathbf{U}_0\psi_1=\psi_1+\lambda_1\psi_0.$$

Taking the inner product on the left with the vector $\tilde{\psi}_0$ and noticing that $\tilde{\psi}_0\mathbf{U}_0=\psi_0$, we find that $\lambda_1=0$. The vector ψ_1 then satisfies the equation $\mathbf{U}_0\psi_1=\psi_1$. It is clear, however, that we are at liberty to choose ψ_1 and all the remaining ψ_p orthogonal to ψ_0, saving the final normalization of ψ until the end. With this restriction, ψ_1 must vanish identically.

Multiplying the equation for ψ_2 on the left by $\tilde{\psi}_0$, we find in this case also that $\lambda_2=0$. To find ψ_2 we must therefore solve the equation

$$(\mathbf{I}-\mathbf{U}_0)\psi_2=\mathbf{U}_2\psi_0,$$

where $\mathbf{I}$ is the identity matrix.

Denoting the components of ψ_2 by $\psi(\alpha)$ ($\alpha=0, 1, 2, \cdots, 2^{n-1}-1$), we find for $(\mathbf{I}-\mathbf{U}_0)\psi_2$ and $\mathbf{U}_2\psi_0$ the vectors

$$(\mathbf{I}-\mathbf{U}_0)\psi=\left\|\begin{matrix}0\\ \psi(1)-\psi(2)\\ \psi(2)-\psi(4)\\ \vdots\\ \psi(j)-\psi(2j)\\ \vdots\\ \psi(2^{n-2}-1)-\psi(2^{n-1}-2)\\ \psi(2^{n-2})\\ \vdots\\ \psi(2^{n-2}+p)\\ \vdots\\ \psi(2^{n-1}-1)\end{matrix}\right\|,\tag{4.7}$$

$$\mathbf{U}_2\psi_0=\left\|\begin{matrix}0\\0\\ \vdots\\0\\0\\ \vdots\\0\\1\end{matrix}\right\|\tag{4.8}$$

Equation (4.7) is of the type $(\mathbf{I}-\mathbf{U}_0)\psi=\xi$, where ξ is a known vector and ψ is to be found. The entire perturbation procedure consists in repeatedly solving equations of this kind. The method used is the following. Since the components 2^{n-2}, $2^{n-2}+1$, $\cdots$, $2^{n-1}-1$ of $(\mathbf{I}-\mathbf{U}_0)\psi$ are identical with those of ψ, they are obtained immediately from the corresponding components of ξ. The remaining half of the components of ψ are obtained with the help of those already known. For example,

$$\begin{aligned}\psi(2^{n-2}-1)&=\psi(2^{n-1}-2)+\xi(2^{n-2}-1),\\ \psi(2^{n-2}-2)&=\psi(2^{n-1}-4)+\xi(2^{n-2}-2),\end{aligned}$$

etc.

A component of even index j is coupled directly to the component of index $\frac{1}{2}j$ and indirectly to those with indices $\frac{1}{4}j$, $\frac{1}{8}j$ . . . if j is also divisible by 4, 8. . . . One continues to follow the even components $j/2^k$ until further division by 2 is no longer possible.

Applied to $\boldsymbol{\psi}_2$, the procedure gives $\boldsymbol{\psi}_2(2^{n-1}-1)=1$ and $\boldsymbol{\psi}_2(\alpha)=0$ for $\alpha=0, 1, 2, \cdots, 2^{n-1}-2$. Multiplying the equation for $\boldsymbol{\psi}_3$ on the left by $\tilde{\boldsymbol{\psi}}_0$, we find $\lambda_3=\tilde{\boldsymbol{\psi}}_0\mathbf{U}_1\boldsymbol{\psi}_2$. In general, from the fact that $\boldsymbol{\psi}_0$ is orthogonal to the other $\boldsymbol{\psi}_j$ and that $\tilde{\boldsymbol{\psi}}_0\mathbf{U}_2=0$, we find $\lambda_r=\tilde{\boldsymbol{\psi}}_0\mathbf{U}_1\boldsymbol{\psi}_{r-1}$. However, since

$$\tilde{\boldsymbol{\psi}}_0\mathbf{U}_1=\|0\ 1\ 0\ \cdots\ 0\|,$$

we have

$$\lambda_r=\boldsymbol{\psi}_{r-1}(1). \tag{4.9}$$

Since the component number 1 of $\boldsymbol{\psi}_2$ vanishes we see that $\lambda_3=0$. $\boldsymbol{\psi}_3$ is then to be found from the equation $(\mathbf{I}-\mathbf{U}_0)\boldsymbol{\psi}_3=U_1\boldsymbol{\psi}_2$. The non-zero components of $\mathbf{U}_1\boldsymbol{\psi}_2$ are

$$(\mathbf{U}_1\boldsymbol{\psi}_2)(2^{n-2}-1)=1,\quad (\mathbf{U}_1\boldsymbol{\psi}_2)(2^{n-2})=1.$$

The non-zero components of $\boldsymbol{\psi}_3$ are therefore

$$\boldsymbol{\psi}_3(2^{n-2})=1,\quad \boldsymbol{\psi}_3(2^{n-2}-1)=1,$$
$$\boldsymbol{\psi}_3(2^k)=1\quad (k=1, 2, \cdots, n-3)$$

arising from the connection with the component 2^{n-2}, and $\boldsymbol{\psi}_3(1)=1$ arising from the connection with $\boldsymbol{\psi}_3(2)$.

From (4.9) we then find $\lambda_4=\boldsymbol{\psi}_3(1)=1$. The vector $\boldsymbol{\psi}_4$ therefore satisfies the equation

$$(\mathbf{I}-\mathbf{U}_0)\boldsymbol{\psi}_4=\mathbf{U}_1\boldsymbol{\psi}_3+\mathbf{U}_2\boldsymbol{\psi}_2-\boldsymbol{\psi}_0.$$

However, $\mathbf{U}_2\boldsymbol{\psi}_0=0$. Since

$$(\mathbf{U}_1\boldsymbol{\psi})(j)=(\mathbf{U}_1\boldsymbol{\psi})(2^{n-1}-1-j)=\boldsymbol{\psi}(2j+1),$$

we find for the non-vanishing components of $\mathbf{U}_1\boldsymbol{\psi}_3$ the following:

$$(\mathbf{U}_1\boldsymbol{\psi}_3)(0)=(\mathbf{U}_1\boldsymbol{\psi}_3)(2^{n-1}-1)=1$$

and

$$(\mathbf{U}_1\boldsymbol{\psi}_3)(2^{n-3}-1)=(\mathbf{U}_1\boldsymbol{\psi}_3)(3.2^{n-3})=1.$$

Therefore,

$$\boldsymbol{\psi}_4(2^{n-1}-1)=1,\quad \boldsymbol{\psi}_4(3.2^{n-3})=1$$

and from this

$$\boldsymbol{\psi}_4(3.2^k)=1\quad (k=1\cdots n-4),\quad \boldsymbol{\psi}_4(3)=1;$$

finally $\boldsymbol{\psi}_4(2^{n-3}-1)=1$.

This perturbation procedure was carried as far as the 12th power of β and gives for the maximum characteristic value of the plus **V** matrix[14]

$$\lambda_+=\frac{1}{\beta}(1+\beta^4+2\beta^6+5\beta^8+14\beta^{10}+44\beta^{12}+\cdots). \tag{4.10}$$

The convergence of this series will be inferred later on. More complete series for $\boldsymbol{\psi}_+$ will also be given later together with the characteristic vectors of the minus **V** matrix. However, up to terms of the order β^5, the non-zero components of the approximating vectors are given in Table I.

The expression (4.10) for λ_+, if carried out completely, would essentially give the free energy per particle for an infinite crystal (i.e., one for which $n=\infty$). For any finite crystal the corresponding series will agree with the limiting one, (4.10), up to a certain power of β, after which they will differ. It is not difficult to see from the perturbation procedure that the agreement will persist to higher and higher powers of β as n increases. The only reason for a difference between the limiting series and the one for finite n, is that at some stage in the calculation two or more components of a vector which are labeled differently actually become identical. For example, in the case $n=5$ the component number 3 of the fourth approximating vector $\boldsymbol{\psi}_4$ does not have the value 1 as one would find by the above methods, but actually has the value 2. This comes about because the component labeled $2^{n-3}-1$ is also the component number 3. Because of these coincidences the components of the approximating vector $\boldsymbol{\psi}_r$ and hence also the coefficients λ_r for a given n, will eventually differ, as r increases, from the $\boldsymbol{\psi}_r$ and λ_r for larger values of n. However, these coincidences can be postponed by taking larger and larger values for n, so that at any fixed stage r in the approximation a given component of $\boldsymbol{\psi}_r$ or the value of λ_r will be the same for all sufficiently large finite problems.

The same type of perturbation procedure is applicable to the problem of the minus **V** matrix. Carried as far as terms of the order β^{12} we find that its maximum characteristic value λ_- is

---

[14] The terms up to β^{10} of this series were previously obtained by Bloch, using a different method. F. Bloch, Zeits. f. Physik **61**, 206 (1930).

TABLE I. The approximating vectors ψ_0 to ψ_5.

ψ_0		ψ_2		ψ_3		ψ_4		ψ_5	
Component	Value	Component	Value	Component	Value	Component	Value	Component	Value
0	1	$2^{n-1}-1$	1	1	1	3	1	1	2
				$2^k \quad k=1\cdots n-3$	1	$3.2^k \quad k=1\cdots n-4$	1	7	1
				$2^{n-2}-1$	1	$2^{n-3}-1$	1	$2^k \quad k=1\cdots n-2$	1
				2^{n-2}	1	3.2^{n-3}	1	$7.2^k \quad k=1\cdots n-4$	1
						$2^{n-1}-1$	1	$2^{n-4}-1$	1
								$2^{n-2}-1$	3
								$2^{n-1}-2^k-1 \quad k=1\cdots n-3$	1
								$2^{n-1}-2$	2

identical with λ_+ for sufficiently large n. Moreover we find to this approximation that the components of the vector ψ_- are identical with those of ψ_+ except for characteristic differences in sign. For any finite value of n, the characteristic values λ_+ and λ_- will differ. However, experience with the perturbation method leads us to believe that as n becomes increasingly large, the difference between λ_+ and λ_- is postponed to higher and higher powers of β, so that in the limiting case of an infinite crystal the degeneracy between λ_+ and λ_- will be expected to hold as far as the series converge. Really to give a proof of the degeneracy, one would have to devise some inductive argument of the type given by Lassettre and Howe[15] to show the degeneracy of the maximum characteristic value of the matrix for the three-dimensional problem.

To find the left-handed characteristic vectors $\tilde{\mathbf{B}}_+$ and $\tilde{\mathbf{B}}_-$ the problems of obtaining the left-handed characteristic vectors of the plus and minus $\mathbf{V}$ matrices are treated by the same perturbation methods. The same series for λ_+ and λ_- are found as previously and the equality, except for characteristic sign differences, of the components of $\tilde{\boldsymbol{\phi}}_+$ and $\tilde{\boldsymbol{\phi}}_-$ is also observed.

For a proof of the convergence of the series for λ_+ and λ_- we refer to the work of van der Waerden.[5] By a consideration of the probabilities for finding polygons of various lengths enclosing atoms which are not ordered with respect to one another, he calculates the long range order and the energy of the crystal at low temperatures, obtaining series identical with ours. Although his method would be very cumbersome for treating the general propagation problem, it is more powerful than the matrix method in the respect that it provides simple convergence proofs. From the convergence of the series for the energy we can conclude that the series for λ_+ and λ_- also converge. The energy is obtained from the partition function, or from the partition function per particle, by

$$\begin{aligned} E &= -\frac{\partial \log Z}{\partial(1/kT)} = NJ\beta\frac{d\log\lambda_+}{d\beta} \\ &= -NJ+NJ\beta\frac{d\log(\beta\lambda_+)}{d\beta} \qquad (4.11) \\ &= -NJ+NJ\{4\beta^4+12\beta^6+36\beta^8 \\ &\qquad +120\beta^{10}+448\beta^{12}+\cdots\}. \end{aligned}$$

This is the same as the series for the energy which is given by van der Waerden and from his proof of its convergence we see that the logarithmic derivative of $\beta\lambda_+$ is analytic in the neighborhood of $\beta=0$. The logarithm of $\beta\lambda_+$ is therefore analytic at $\beta=0$ and hence the same is true of $\beta\lambda_+$ itself.

5. The Long Range Order

The probability $P(\alpha_1, \alpha_r)$ of finding simultaneously pitch 1 in configuration α_1, and pitch r in configuration α_r is given by an expression analagous to (2.1) of Section 2. Using (3.3), (3.4), and (3.5) one finds

$$P(\alpha_1, \alpha_r) = \frac{1}{\sum_p \lambda_p{}^{mn}} \sum_p \sum_q \lambda_p{}^{(r-1)n} \lambda_q{}^{(m-r+1)n} \times \mathbf{B}_p(\alpha_1)\mathbf{A}_q(\alpha_1)\mathbf{B}_q(\alpha_r)\mathbf{A}_p(\alpha_r). \quad (5.1)$$

In the ferromagnetic case, for sufficiently low temperatures, we have found the maximum

---

[15] E. N. Lassettre and J. P. Howe, J. Chem. Phys. 9, 801 (1941).

characteristic value to be two-fold degenerate. Therefore, for two pitches far enough apart so that $r-1$ and $m-r+1$ are large integers, we have

$$P(\alpha_1, \alpha_r)=\tfrac{1}{2}\{\mathbf{B}_+(\alpha_1)\mathbf{A}_+(\alpha_r)+\mathbf{B}_-(\alpha_1)\mathbf{A}_-(\alpha_r)\} \times\{\mathbf{B}_+(\alpha_r)\mathbf{A}_+(\alpha_1)+\mathbf{B}_-(\alpha_r)\mathbf{A}_-(\alpha_1)\}. \quad (5.2)$$

Let us now ask for the probability of finding two + atoms simultaneously, one in a given position in pitch 1 and the other at a given place in pitch r. In particular we might ask that the two end atoms of the separate pitches be both +. Then the probability P in question is obtained by summing $P(\alpha_1, \alpha_r)$ over those configurations consistent with having + atoms at the end of the pitch. These comprise the first half of all the configurations, namely, the ones numbered from 0 to $2^{n-1}-1$.

From (3.10) and (3.11) the normalized forms of $\mathbf{A}_+$ and $\mathbf{B}_+$ are

$$\begin{aligned}\mathbf{A}_+&=(2\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+)^{-\frac{1}{2}}\left\|\begin{matrix}\boldsymbol{\psi}_+\\ \boldsymbol{\psi}_+\end{matrix}\right\|,\\ \mathbf{B}_+&=(2\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+)^{-\frac{1}{2}}\|\tilde{\boldsymbol{\phi}}_+, \tilde{\boldsymbol{\phi}}_+\|.\end{aligned} \quad (5.3)$$

In the same way, the normalizing factor for the minus vectors is $(2\tilde{\boldsymbol{\phi}}_-\boldsymbol{\psi}_-)^{-\frac{1}{2}}$. However, the series expansions show that $\tilde{\boldsymbol{\phi}}_-\boldsymbol{\psi}_-=\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+$. It also turns out that $\mathbf{B}_+(\alpha)\mathbf{A}_+(\alpha)=\mathbf{B}_-(\alpha)\mathbf{A}_-(\alpha)$ and that $\mathbf{B}_+(\alpha)\mathbf{A}_-(\alpha)=\mathbf{B}_-(\alpha)\mathbf{A}_+(\alpha)$. Using these relations, multiplying out (5.2), and summing over the first half of the configurations, we obtain for P the expression

$$P=\tfrac{1}{4}\{1+(\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_-/\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+)^2\}. \quad (5.4)$$

It can be verified without very much difficulty that this same expression remains valid if the two + atoms under consideration do not necessarily occupy corresponding positions in the two distant pitches. P is therefore the chance of finding simultaneously two + atoms very far apart in a ferromagnetic crystal. In the absence of long range order, $P=\frac{1}{4}$.

If the crystal is antiferromagnetic, this probability will of course depend on the odd or even character of the number of steps between the two atoms. Instead of a degeneracy in $\lambda_{\max}$, we have, from (3.12), $\lambda_+(-K)=-\lambda_-(-K)$. The total number of atoms in the crystal, mn is even, however, since there are as many + atoms as − atoms. Thus, $\lambda_+(-K)^{mn}=\lambda_-(-K)^{mn}$, and the effect is the same as though there were a real degeneracy. Also, from (3.12), we see that the characteristic vectors $\tilde{\boldsymbol{\phi}}_\pm(-K)$, $\boldsymbol{\psi}_\pm(-K)$ are obtained from $\tilde{\boldsymbol{\phi}}_\pm(K)$, $\boldsymbol{\psi}_\pm(K)$ by permuting the components in accordance with a certain permutation matrix $\mathbf{R}$. The components of $\tilde{\boldsymbol{\phi}}_\pm(K)$ and $\boldsymbol{\psi}_\pm(K)$ are permuted in the same way, however. Consequently the inner products of the vectors are invariant; that is, for example,

$$\tilde{\boldsymbol{\phi}}_+(K)\cdot\boldsymbol{\psi}_-(K)=\tilde{\boldsymbol{\phi}}_+(-K)\cdot\boldsymbol{\psi}_-(-K)$$

and

$$\tilde{\boldsymbol{\phi}}_+(K)\cdot\boldsymbol{\psi}_+(K)=\tilde{\boldsymbol{\phi}}_+(-K)\cdot\boldsymbol{\psi}_+(-K).$$

Carrying out the computation for P, we find an expression

$$P=\tfrac{1}{4}\{1+(-1)^k(\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_-/\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+)^2\}, \quad (5.5)$$

where k is, say, the smallest number of steps between the two + atoms entering into the definition of P. If k is even, the two + atoms form an ordered pair, and P exceeds $\frac{1}{4}$. If k is odd, the pair of + atoms is not ordered, and P is less than $\frac{1}{4}$.

To make connection with the usual long range order S we shall obtain an expression for P in terms of S and compare with (5.4) and (5.5). For an antiferromagnetic alloy, one divides the sites of the crystal into two groups, α and β sites, the former being occupied exclusively by + atoms at the zero of temperature and the latter by − atoms. At a finite temperature one denotes by r the probability that an atom be right, or on a site which is appropriate to it, and by w the probability that a + or a − atom be wrong, that is, be on a site not appropriate to it. The long range order S is then defined by writing w and r in the form

$$w=\tfrac{1}{2}(1-S), \quad r=\tfrac{1}{2}(1+S). \quad (5.6)$$

In terms of S, the probability, P, for finding two atoms, of which one is +, and which are correctly ordered with respect to one another is given by

$$P=\frac{1}{2}\left(\frac{1+S}{2}\right)^2+\frac{1}{2}\left(\frac{1-S}{2}\right)^2=\frac{1}{4}(1+S^2). \quad (5.7)$$

This can be seen as follows. If one takes the factors in order, the first $\frac{1}{2}$ is the chance that an α site is selected for the first atom of the pair, which we may take as $+$. The chance of getting a $+$ atom on this site and also of getting the right atom at the distant site, is then $\frac{1}{4}(1+S)^2$. If the site chosen originally is a β site, the two atoms of the pair would both have to be wrong from the point of view of the $\alpha-\beta$ sites in order to be correct with respect to one another. This would give the factor $\frac{1}{4}(1-S)^2$. In the same way the probability for finding two distant atoms which are not correctly ordered with respect to one another, is $\frac{1}{4}(1-S^2)$.

From (5.4) and (5.5) it therefore follows that the long range order S is given in terms of the characteristic vectors by

$$S=\tilde{\phi}_+\psi_-/\tilde{\phi}_+\psi_+. \tag{5.8}$$

The computation of S up to terms of order β^8 is given in Table II. All configurations α which give the same series expansion for $\phi_+(\alpha)\psi_-(\alpha)$, to terms of order β^8, are grouped together and listed in the first column. The number of configurations in each group is given in column two. The third column gives the coefficients of β^0, β^4, β^6, and β^8 appearing in the series for $\phi_+(\alpha)\psi_-(\alpha)$. To obtain $\tilde{\phi}_+\psi_-$, we must sum $\phi_+(\alpha)\psi_-(\alpha)$ over all α from 0 to $2^{n-1}-1$. This is conveniently done by multiplying the coefficients in the third column by the number of configurations in the second column and then adding. The results are indicated in the last column. We have

$$\tilde{\phi}_+\psi_-=1+(n-2)\beta^4+(3n-7)\beta^6 +\left(\frac{1}{2}n^2+\frac{17}{2}n-30\right)\beta^8+\cdots. \tag{5.9a}$$

The product $\tilde{\phi}_+\psi_+$ is gotten in the same way. Its table differs from that of $\tilde{\phi}_+\psi_-$ only in that all the signs of the $\phi_+(\alpha)\psi_+(\alpha)$ are positive. Otherwise the numerical values are the same as for the $\phi_+(\alpha)\psi_-(\alpha)$. Making the appropriate changes we find

$$\tilde{\phi}_+\psi_+=1+n\beta^4+(3n+1)\beta^6 +\left(\frac{1}{2}n^2+\frac{21}{2}n+4\right)\beta^8+\cdots. \tag{5.9b}$$

Both $\tilde{\phi}_+\psi_-$ and $\tilde{\phi}_+\psi_+$ depend on the size of the crystal, n, but their quotient, which is the long range order, must be independent of n for very large n. Performing the division, we obtain to terms of order β^{12}

$$S=\tilde{\phi}_+\psi_-/\tilde{\phi}_+\psi_+=1-2(\beta^4+4\beta^6+17\beta^8 +76\beta^{10}+357\beta^{12}+\cdots) \tag{5.10}$$

and

$$P=\tfrac{1}{4}(1+S^2)=\tfrac{1}{2}-\beta^4-4\beta^6-16\beta^8 -68\beta^{10}-307\beta^{12}-\cdots. \tag{5.11}$$

This same series for S was found by van der Waerden and a proof was given that it is convergent for $\beta<\frac{1}{3}$. The question of the actual radius of convergence of the series is still open, however. A highly probable conjecture can be made in the light of the investigations of Kramers and Wannier. On the basis of an important symmetry property of the plus $\mathbf{V}$ matrix, they are able to show that if it is assumed that the crystal can have only one point of singular behavior, that point must be given by

$$\beta_c=\sqrt{2}-1=0.414. \tag{5.12}$$

As we shall prove in the next section, there is no long range order at high temperatures. It therefore seems very probable that the point $\beta_c=0.414$ represents the Curie point for the transition from order to disorder and that the series for S has radius of convergence 0.414. A real proof of this fact and an investigation of the limiting value of $S(\beta)$ as β approaches β_c from below, would be very much desired, however.

6. Proof That an Order-Disorder Transition Actually Takes Place

Although the location of the transition point between states with finite long range order and those with no long range order has not been rigorously obtained by the matrix method, it is easy to give a proof that such a transition really occurs. Such proofs have been given by Peierls[16] and van der Waerden[5] using the properties of polygons enclosing regions of disordered atoms.

---

[16] R. Peierls, Proc. Camb. Phil. Soc. **32**, 477–481 (1936).

TABLE II. The calculation of $\tilde{\phi}_+\psi_-$.

Configuration, α		Number in group	$\phi_+(\alpha)\psi_-(\alpha)$				$\tilde{\phi}_+\psi_-$			
			β^0	β^4	β^6	β^8	β^0	β^4	β^6	β^8
0		1	1				1			
2^k	$k=1\cdots n-2$	$n-2$		1	2	5		$n-2$	$2n-4$	$5n-10$
3.2^k	$k=1\cdots n-3$	$n-3$			1	6			$n-3$	$6n-18$
2^j+2^k	$j=3\cdots n-2,\ k=1\cdots j-2$	$\frac{1}{2}(n-4)(n-3)$				1				$\frac{1}{2}n^2-\frac{7}{2}n+6$
7.2^k	$k=1\cdots n-4$	$n-4$				1				$n-4$
$2^{n-1}-2$		1				-4				-4
1		1		1	3	8		1	3	8
3		1			1	7			1	7
7		1				1				1
2^k+1	$k=2\cdots n-2$	$n-3$				1				$n-3$
$2^{n-3}-1$		1				-1				-1
$2^{n-2}-1$		1			-1	-7			-1	-7
$2^{n-1}-2^k-1$	$k=1\cdots n-3$	$n-3$				-1				$-n+3$
$2^{n-1}-1$		1		-1	-3	-8		-1	-3	-8
							1	$n-2$	$3n-7$	$\frac{1}{2}n^2+\frac{17}{2}n-30$

From the point of view of the matrix theory all that is necessary is to show that for sufficiently high temperatures the maximum characteristic value of $\mathfrak{M}(K)$ is no longer degenerate.

For high temperatures $K(=\frac{1}{2}J/kT)$ tends to zero and both α and β tend to unity. Thus for very high temperatures the matrices $\mathbf{V}_+(K)$ and $\mathbf{V}_-(K)$ are practically

$$\mathbf{V}_+(0)=\left\|\begin{array}{cccc} 11 & & & \\ & 11 & & \\ & & \ddots & \\ & & & 11 \\ & & & 11 \\ & & \cdot^{\cdot^{\cdot}} & \\ & 11 & & \\ 11 & & & \end{array}\right\|,$$

$$\mathbf{V}_-(0)=\left\|\begin{array}{cccc} 1\ 1 & & & \\ & 1\ 1 & & \\ & & \ddots & \\ & & & 1\ 1 \\ & & & -1-1 \\ & & \cdot^{\cdot^{\cdot}} & \\ & -1-1 & & \\ -1-1 & & & \end{array}\right\|. \tag{6.1}$$

Taking a finite problem for which the matrices are of the order 2^r, one finds upon successive multiplication that the rth powers of $\mathbf{V}_+(0)$ and $\mathbf{V}_-(0)$ have the forms

$$\mathbf{V}_+(0)^r=\left\|\begin{array}{ccc} 111 & \cdots & 111 \\ 111 & \cdots & 111 \\ 111 & \cdots & 111 \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ 111 & \cdots & 111 \\ 111 & \cdots & 111 \end{array}\right\|,$$

$$\mathbf{V}_-(0)^r=\left\|\begin{array}{ccccccc} 1 & 1 & 1 & \cdots & 1 & 1 & 1 \\ -1 & -1 & -1 & \cdots & -1 & -1 & -1 \\ 1 & 1 & 1 & \cdots & 1 & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & -1 & -1 & \cdots & -1 & -1 & -1 \\ 1 & 1 & 1 & \cdots & 1 & 1 & 1 \end{array}\right\|. \tag{6.2}$$

$\mathbf{V}_+(0)^r$ consists entirely of ones while $\mathbf{V}_-(0)^r$ has alternate rows of ones and minus ones. Since the matrices are of rank one, their characteristic equations reduce to the form

$$\lambda^{2^r}-(\text{trace }\mathbf{V}_\pm(0)^r)\cdot\lambda^{2^r-1}=0. \tag{6.3}$$

The trace of $\mathbf{V}_+(0)^r$ is 2^r. Therefore the characteristic values of $\mathbf{V}_+(0)^r$ are 2^r, 0, 0, $\cdots$, 0 and hence the maximum characteristic value of $\mathbf{V}_+(0)$ is 2 and all the rest are 0. The trace of $\mathbf{V}_-(0)^r$ vanishes, however. As a result all of its characteristic values and hence all the characteristic values of $\mathbf{V}_-(0)$ vanish.

This proves that for sufficiently high temperatures the maximum characteristic value of $\mathfrak{M}$ is non-degenerate and that there is therefore no long range order.

7. The Propagation of Order

The principal advantage that can be claimed for the matrix method is that it provides a solution not only to the problem of the long range order, but also to the more general problem of the propagation of order throughout the lattice. For example, suppose that a certain site in the crystal is known to be occupied by a + atom. We can then investigate how this knowledge affects the probability for + (or −) atoms at *all* other sites and see the gradual transition from the strong short range influence of the fixed + atom to the weaker long range influence which is described by the long range order. This is the problem originally proposed by Zernike. Moreover, using the matrix method we can investigate the propagation of order for more complicated initial configurations consisting of several atoms. Thus we shall find the ordering effect of a pair of fixed atoms, of combinations of three, four atoms, and so on, in each case observing the change from the short to the long range effect.

Zernike's Problem. The Transition from Short to Long Range Order

If the expression (5.1) for $P(\alpha_1, \alpha_r)$ is summed over all configurations α_r, one finds for the probability of a configuration α_1, in pitch number 1

$$P(\alpha_1)=\frac{1}{\sum_p \lambda_p{}^{mn}}\sum_p \lambda_p{}^{mn}\mathbf{B}_p(\alpha_1)\mathbf{A}_p(\alpha_1).$$

At low temperatures and for ferromagnetic crystals, the twofold degeneracy of $\lambda_{\max}$ leads to the expression (dropping the subscript 1)

$$\begin{aligned}P(\alpha)&=\tfrac{1}{2}(\mathbf{B}_+(\alpha)\mathbf{A}_+(\alpha)+\mathbf{B}_-(\alpha)\mathbf{A}_-(\alpha))\\&=\mathbf{B}_+(\alpha)\mathbf{A}_+(\alpha).\qquad(7.1)\end{aligned}$$

This is in marked analogy with the quantum mechanics, where the probability density is proportional to the wave function multiplied by its complex conjugate. From (7.1) we can also see that the theory is entirely symmetrical in the two kinds of atoms, + and −. Thus if we make a transformation which takes a configuration α into one $\bar{\alpha}$ obtained from it by changing all + atoms to − atoms, and conversely, we find from (3.10) and (3.11) that $\mathbf{B}_+(\alpha)=\mathbf{B}_+(\bar{\alpha})$ and $\mathbf{A}_+(\alpha)=\mathbf{A}_+(\bar{\alpha})$, so that $P(\alpha)=P(\bar{\alpha})$. This symmetry is evidently necessary, since + and − atoms have been treated equally from the start, the energy of interaction between a pair of + atoms being equal to the energy of interaction of a pair of − atoms.

Any question concerning the probability of a certain disposition of atoms on lattice sites all in one pitch is answered by summing $P(\alpha)$ over all configuration α consistent with the restrictions proposed. We shall, therefore, always have to deal with the expressions of the type

$$\sum\nolimits'_\alpha \mathbf{B}_+(\alpha)\mathbf{A}_+(\alpha)=\frac{\sum'_\alpha \boldsymbol{\phi}_+(\alpha)\boldsymbol{\psi}_+(\alpha)}{2\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+}=\frac{\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+-\mathbf{V}}{2\tilde{\boldsymbol{\phi}}_+\boldsymbol{\psi}_+},\qquad(7.2)$$

where the prime indicates that the summation is to be taken over the configurations allowed by the restrictions, and V is the sum of $\boldsymbol{\phi}_+(\alpha)\boldsymbol{\psi}_+(\alpha)$ over the suppressed configurations. The sum V is generally easier to carry out than the direct sum over the allowed configurations.

Let us now apply (7.2) to the investigation of the ordering influence of one atom on its neighbors in the same line of the crystal. Let $P(+_1r_2)$, $P(+_1r_3)$, $\cdots$, $P(+_1r_p)$,[17] be the probability of finding a + atom at a site conveniently called 1 and then an atom rightly ordered with respect to it at a distance of one site away, two sites away, . . . , $(p-1)$ sites away, respectively. We can write $P(+_1r_p)$ as equal to the probability for finding a + atom at the site 1 in the lattice multiplied by the probability that the right atom is found $p-1$ sites away when it is already known that a + atom exists at site 1. The latter correlation probability we denote by $C(+_1|r_p)$ and take it as a measure of the ordering influence of a single + atom. Of course

$$P(+_1r_p)=\tfrac{1}{2}C(+_1|r_p).$$

---

[17] This r notation is used so as not to distinguish between the ferromagnetic and antiferromagnetic cases. The calculations will be made for the ferromagnetic case, but it will be shown that the results are valid in the other as well. For the ferromagnetic case, r_2, for example, would be $+_2$, while for the antiferromagnetic case it would be $-_2$.

These correlation probabilities are easily calculated from a table of values of $\phi_+(\alpha)\psi_+(\alpha)$ similar to that given in Table II. For example, if, in (7.2), we make the restriction that the last two atoms, numbers $n-1$ and $n-2$ are both $+$ we have the conditions necessary for the calculation of $C(+_1|r_2)$. Similarly, for $C(+_1|r_3)$ we would hold the atoms $n-1$ and $n-3$ fixed as $+$. We find

$$\begin{aligned}
C(+_1|r_2)&=1-2\beta^4-6\beta^6-18\beta^8\\
&\qquad -60\beta^{10}-224\beta^{12}-\cdots,\\
C(+_1|r_3)&=1-2\beta^4-8\beta^6-30\beta^8\\
&\qquad -112\beta^{10}-430\beta^{12}-\cdots,\\
C(+_1|r_4)&=1-2\beta^4-8\beta^6-32\beta^8\\
&\qquad -134\beta^{10}-576\beta^{12}-\cdots,\\
C(+_1|r_5)&=1-2\beta^4-8\beta^6-32\beta^8\\
&\qquad -136\beta^{10}-612\beta^{12}-\cdots,\\
C(+_1|r_6)&=1-2\beta^4-8\beta^6-32\beta^8\\
&\qquad -136\beta^{10}-614\beta^{12}-\cdots,\\
C(+_1|r_7)&=1-2\beta^4-8\beta^6-32\beta^8\\
&\qquad -136\beta^{10}-614\beta^{12}-\cdots,
\end{aligned}\tag{7.4}$$

and $C(+_1|r_p)=C(+_1|r_6)$ to terms of order β^{12} for $p\geqq 7$.

The correlation probability, $C(+_1|r_p)$, has its maximum value of unity at the absolute zero of temperature ($\beta=0$) and for finite temperatures ($\beta>0$) falls off monotonically as one goes further from the central site at which the $+$ atom is situated. Examining the sequence (7.4) we find that the coefficients of the various powers of β increase monotonically in absolute value until they reach a certain limit, the approach to the limit being rapid at first and then slower at the end. For the lower powers of β this approach to the limiting value occurs earlier in the sequence than it does for the higher powers. Thus, from (7.4),

$$\begin{aligned}
&C(+_1|r_2)>C(+_1|r_3)>C(+_1|r_4)\\
&\qquad >C(+_1|r_5)>C(+_1|r_6)>\cdots.
\end{aligned}\tag{7.5}$$

For $p\geqq 6$ we see that to terms of the order β^{12}, $C(+_1|r_p)$ becomes identical with the long range correlation (5.11). Denoting the long range correlation by $C(+_1|r_\infty)$ we have

$$C(+_1|r_\infty)=\tfrac{1}{2}(1+S^2).\tag{7.6}$$

In this approximation the long range effect is already attained five sites away from the central atom. This holds of course only for those temperatures for which terms of order higher than β^{12} can be neglected. As the temperature increases the transition to the long range limit takes place only after one goes considerably further from the central site.

The correlation probability $C(+_1|r_2)$ is closely related to the average energy of the crystal. Thus if Q_r stands for the probable number of right pairs of atoms each having an energy of interaction $-\frac{1}{2}J$ and Q_w for the number of wrong pairs of atoms each having an energy of $\frac{1}{2}J$, we have for the average energy

$$E=-\tfrac{1}{2}JQ_r+\tfrac{1}{2}JQ_w.$$

Remembering that for a crystal containing N atoms there are $2N$ pairs of nearest neighbors, we find

$$Q_r=2NC(+_1|r_2),\quad Q_w=2N\{1-C(+_1|r_2)\}$$

and

$$\begin{aligned}
E&=NJ\{1-2C(+_1|r_2)\}\\
&=-NJ+NJ(4\beta^4+12\beta^6+36\beta^8\\
&\qquad +120\beta^{10}+448\beta^{12}+\cdots).
\end{aligned}\tag{7.7}$$

This series for E has been given before in (4.11) and was obtained there from the expansion of the maximum characteristic value λ_+. The present method is based on the characteristic vectors and the agreement between the two provides a check on the internal consistency of the results.

It is not necessary to confine oneself entirely to atoms on the same line with the fixed $+$ atom. For in virtue of the screwlike construction of the lattice, the atom at the end of the pitch, $n-1$, is situated near the atoms 0, 1, 2, $\cdots$ at the beginning of the pitch, as shown in Fig. 2. We can find the influence of a fixed $+$ atom at sites on the next line by applying (7.2) with the restrictions that atoms $n-1$ and 0, $n-1$ and 1, $n-1$ and 2, etc. are to be successively held fixed as $+$ atoms. The notation $C({}^{+1}|r_p)$ suggests

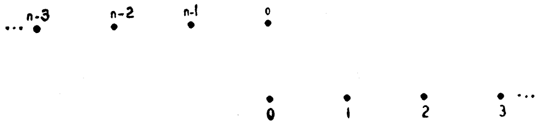

FIG. 2. The n sites of a single pitch.

itself for the correlation probabilities. The results are

$$\begin{aligned} C({}^{+1}|r_2) &= 1-2\beta^4-8\beta^6-26\beta^8 \\ &\qquad -88\beta^{10}-326\beta^{12}-\cdots, \\ C({}^{+1}|r_3) &= 1-2\beta^4-8\beta^6-32\beta^8 \\ &\qquad -124\beta^{10}-484\beta^{12}-\cdots, \\ C({}^{+1}|r_4) &= 1-2\beta^4-8\beta^6-32\beta^8 \\ &\qquad -136\beta^{10}-594\beta^{12}-\cdots, \\ C({}^{+1}|r_5) &= 1-2\beta^4-8\beta^6-32\beta^8 \\ &\qquad -136\beta^{10}-614\beta^{12}-\cdots, \\ C({}^{+1}|r_6) &= 1-2\beta^4-8\beta^6-32\beta^8 \\ &\qquad -136\beta^{10}-614\beta^{12}-\cdots. \end{aligned} \tag{7.8}$$

This sequence has the same properties as the sequence (7.4). A comparison of the two shows that

$$\begin{aligned} C(+_1|r_2) &> C({}^{+1}|r_2) > C(+_1|r_3) \\ &> C({}^{+1}|r_3) > C(+_1|r_4) > C({}^{+1}|r_4) > \cdots \end{aligned} \tag{7.9}$$

indicating again a monotonic decrease of the correlation probability as the distance from the fixed atom increases.

The correlation probabilities (7.4) and (7.8) do not constitute a complete solution of Zernike's problem since there are many distances between pairs of atoms which are not included. However, this is as far as one can go without a knowledge of the entire spectrum of characteristic values of the $\mathbf{V}_+$ matrix. This is because the expression for $P(\alpha_1, \alpha_r)$, where r is a small integer, cannot be simplified to the point where only the maximum characteristic value of $\mathbf{V}_+$ enters. Nevertheless the range of distances covered by (7.4) and (7.8) is fairly complete and for those that are left out, one can get good estimates of the coefficients of the smaller powers of β by applying the condition for monotonic decrease with distance.

It should be pointed out that the expressions found for the correlation probabilities are valid for antiferromagnetic crystals as well. As we see from (3.12), the vectors $\tilde{\mathbf{B}}_\pm(-K)$, $\mathbf{A}_\pm(-K)$ applicable in this case have components which are identical except for the order with those of $\tilde{\mathbf{B}}_\pm(K)$ and $\mathbf{A}_\pm(K)$ appropriate to the ferromagnetic case. The permutation matrix $\mathbf{R}$, carrying the one set of vectors into the other, carries each configuration into the one obtained from it by changing the sign of every alternate atom in it. Thus for $C(+_1|+_2)$, calculated in the ferromagnetic case, we summed over all the configurations for which atoms $n-1$ and $n-2$ were both $+$. The permutation matrix $\mathbf{R}$ carries these configurations into all those for which the atoms $n-1$ and $n-2$ are, respectively, $+$ and $-$. Therefore $C(+_1|+_2)$, as calculated in the ferromagnetic case, is equal to $C(+_1|-_2)$ for the antiferromagnetic case.

8. Comparison with the Results of Zernike

Carrying out Zernike's method for treating the propagation problem, we find in the case of two dimensions for the long range order and the correlation probabilities

$$\begin{aligned} S &= 1-2\beta^4-8\beta^6-20\beta^8 \\ &\qquad -48\beta^{10}-122\beta^{12}-\cdots, \\ C(+_1|r_2) &= 1-2\beta^4-6\beta^6-11\beta^8-\cdots, \\ C({}^{+1}|r_2) &= 1-2\beta^4-8\beta^6-14\beta^8-\cdots, \\ C(+_1|r_3) &= 1-2\beta^4-8\beta^6-16\beta^8-\cdots, \\ C(+_1|r_4) &= 1-2\beta^4-8\beta^6-18\beta^8-\cdots, \\ C(+_1|r_p) &= C(+_1|r_4) \quad p\geqq 5 \quad \text{(to terms} \\ &\qquad \text{of order } \beta^8). \end{aligned} \tag{8.1}$$

Comparison with (5.10), (7.4), and (7.8) shows that Zernike's approximation at low temperatures is valid up to terms of the order β^6.

In three dimensions we make use of the series for S and $C(+_1|r_2)$ found by van der Waerden.[5] The comparison with Zernike is given in (8.2).

van der Waerden:

$$S=1-2\beta^6-12\beta^{10}+14\beta^{12}-90\beta^{14}+192\beta^{16}-792\beta^{18}+\cdots,$$

Zernike:

$$S=1-2\beta^6+6\beta^8+\cdots, \tag{8.2}$$

van der Waerden:

$$C(+_1|r_2)=1-2\beta^6-10\beta^{10}+14\beta^{12}-70\beta^{14}+176\beta^{16}-626\beta^{18}+\cdots,$$

Zernike:

$$C(+_1|r_2)=1-2\beta^6+6\beta^8+\cdots.$$

Zernike's method therefore seems to be better in two dimensions than it is in three. This can perhaps be understood from the following consideration. The essential approximation in Zernike's difference equation comes from the assumption that the probabilities for the atoms forming the nearest neighbors of a given atom, are independent of one another. Although no two of the atoms at these surrounding sites are nearest neighbors, the probabilities for them are not independent because of the possibility of interaction through intermediate atoms. The magnitude of the correlation will in some way be connected with the number of short paths of influence between the two atoms. In three dimensions, the correlation will presumably be greater because the number of available paths is certainly much greater than in two dimensions. We might therefore expect the approximation to get poorer as the dimension is increased. In one dimension, where the assumption of Zernike is least serious, it can be shown[18] that his solution is actually exact, although the method of deriving the equation involves an approximation.

9. The Ordering Influence of More Complicated Configurations

Next to a single + atom, the simplest initial arrangement to consider is an ordered pair of adjacent atoms, $+_1r_2$. Denoting by $P(+_1r_2r_j)$ the probability for finding a plus atom at a given site and two atoms rightly ordered with respect to it at sites adjacent and $(j-1)$ steps away, we have for the correlation probability

$$C(+_1r_2|r_j)=P(+_1r_2r_j)/P(+_1r_2).$$

Application of the methods of Section 7 gives

$$\begin{aligned}
C(+_1r_2|r_3)&=1-\beta^4-4\beta^6-17\beta^8-70\beta^{10}-291\beta^{12}-\cdots,\\
C(+_1r_2|r_4)&=1-\beta^4-5\beta^6-24\beta^8-109\beta^{10}-487\beta^{12}-\cdots,\\
C(+_1r_2|r_5)&=1-\beta^4-5\beta^6-25\beta^8-121\beta^{10}-580\beta^{12}-\cdots,\\
C(+_1r_2|r_6)&=1-\beta^4-5\beta^6-25\beta^8-122\beta^{10}-599\beta^{12}-\cdots,\\
C(+_1r_2|r_7)&=1-\beta^4-5\beta^6-25\beta^8-122\beta^{10}-600\beta^{12}-\cdots
\end{aligned} \tag{9.1}$$

and to terms of order β^{12}

$$C(+_1r_2|r_p)=C(+_1r_2|r_7),\quad p\geqq 8.$$

Comparison of (7.10) and (7.4) shows that in general

$$C(+_1r_2|r_j)>C(+_1|r_{j-1}).$$

Thus, as expected, an ordered pair of atoms is a more effective ordering influence than a single atom. The sequence (9.1) shows the same kind of approach to the long distance limit as do the sequences (7.4) and (7.8).

It will be interesting to compare the correlation probabilities for the ordered pair $+_1r_2$ with those for the ordered pair ${}^{+1}r_2$ where the atoms are separated by a distance $\sqrt{2}$ times the lattice distance. (See Fig. 3.) By imposing, in (7.2), the condition that the atoms $n-1$, 0, and $j-2$, of a single pitch be held fixed as + we find the probability $P({}^{+1}r_2r_j)$. Dividing this by $P({}^{+1}r_2)$ gives

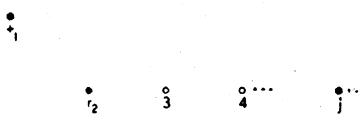

FIG. 3. An ordered pair of atoms separated by a distance $\sqrt{2}$ (in units of the lattice distance).

---

[18] A. Nordsieck (unpublished).

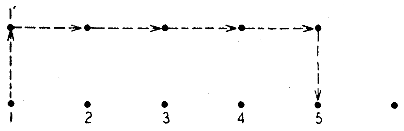

FIG. 4. The shortest path of influence from site 1 to site 5 when site 2 is occupied by a known atom.

the correlation probability $C({}^{+1}r_2|r_j)$. The result is

$$\begin{aligned}
C({}^{+1}r_2|r_3) &= 1-\beta^4-3\beta^6-14\beta^8 \\
&\quad -62\beta^{10}-269\beta^{12}-\cdots, \\
C({}^{+1}r_2|r_4) &= 1-\beta^4-4\beta^6-20\beta^8 \\
&\quad -96\beta^{10}-447\beta^{12}-\cdots, \\
C({}^{+1}r_2|r_5) &= 1-\beta^4-4\beta^6-21\beta^8 \\
&\quad -107\beta^{10}-532\beta^{12}-\cdots, \\
C({}^{+1}r_2|r_6) &= 1-\beta^4-4\beta^6-21\beta^8 \\
&\quad -108\beta^{10}-550\beta^{12}-\cdots, \\
C({}^{+1}r_2|r_7) &= 1-\beta^4-4\beta^6-21\beta^8 \\
&\quad -108\beta^{10}-551\beta^{12}-\cdots
\end{aligned} \tag{9.2}$$

and to terms of order β^{12}

$$C({}^{+1}r_2|r_p)=C({}^{+1}r_2|r_7), \quad p\geqq 8.$$

The comparison of (9.1) with (9.2) shows that the pair ${}^{+1}r_2$ is a more effective ordering influence than the pair $+_1r_2$; that is

$$C({}^{+1}r_2|r_j)>C(+_1r_2|r_j).$$

The reason for this can be easily understood. In the diagram of Fig. 4 let us consider the correlation probability at site 5, perhaps, for an ordered pair of atoms at sites 1 and 2 and then for an ordered pair at 1′ and 2. With the atom at 2 held fixed, the shortest path of influence from 1 to 5 consists of six steps and is shown by the dotted line. However, for the atom at site 1′, the shortest path of influence consists of only five steps. Thus, if site 2 is occupied by a known atom, placing a properly ordered atom at site 1′ will be more effective at 5 than placing an ordered atom at site 1. We should, therefore, expect $C({}^{+1}r_2|r_5)$ to be larger than $C(+_1r_2|r_5)$.

If the correlation probabilities arising from an ordered pair of atoms are greater than the correlation probabilities arising from a single atom, we should find that the correlations for a disordered pair will be smaller than those for a single atom. We shall denote such a disordered pair, or "dipole," by $+_1w_2$. The correlation probabilities $C(+_1w_2|r_j)$ can be obtained from results already given. Thus, expressing the identity

$$P(+_1w_2r_j)+P(+_1r_2r_j)\equiv P(+_1r_j)$$

in terms of correlation probabilities, we have

$$C(+_1w_2|r_j)=\frac{C(+_1|r_j)-C(+_1|r_2)C(+_1r_2|r_j)}{1-C(+_1|r_2)}. \tag{9.3}$$

Here r_j means right with respect to the + atom at site 1.

From (7.4), (9.1), and (9.3) we have then

$$\begin{aligned}
C(+_1w_2|r_3) &= \tfrac{1}{2}(1-\beta^2-3\beta^4-8\beta^6-22\beta^8-\cdots), \\
C(+_1w_2|r_4) &= \tfrac{1}{2}(1\qquad -\beta^4-8\beta^6-40\beta^8-\cdots), \\
C(+_1w_2|r_5) &= \tfrac{1}{2}(1\qquad\qquad -\beta^6-15\beta^8-\cdots), \\
C(+_1w_2|r_6) &= \tfrac{1}{2}(1\qquad\qquad\qquad -\beta^8-\cdots), \\
C(+_1w_2|r_7) &= \tfrac{1}{2}(1\qquad\qquad\qquad\qquad -\cdots)
\end{aligned} \tag{9.4}$$

and to terms of order β^8

$$C(+_1w_2|r_p)=\tfrac{1}{2} \quad p\geqq 7.$$

These series show that at large distances the correlation probability arising from a dipole, $+_1w_2$, approaches the value $\frac{1}{2}$. The dipole therefore has no ordering influence at all at large distances, since the probability of finding a + or a − atom at a given lattice point when the dipole is not present is also equal to $\frac{1}{2}$. The atoms of the disordered pair thus have the tendency to annul each other's ordering effect, the cancellation being complete at infinite distance.

Many further examples of sequences of correlation probabilities could be given. Instead, however, we shall give in Table III a list of correlation probabilities for a line of the lattice from which all other correlations can be obtained by suitable combination. The table gives the correlation probabilities as a function of position in a line of the lattice for a sequence of initial configurations beginning first with one fixed +

atom, then an ordered pair of atoms, then 3, 4, 5, 6, etc. . . . consecutive ordered atoms.

For each fixed initial configuration, the correlation probability decreases monotonically with distance and tends to a limiting function at very large distance. This is the behavior exhibited by the sequences (7.4), (7.8), (9.1), (9.2), and (9.4). However, there is also an approach to a limit for sequences like

$$C(+_1|r_2),\quad C(+_1r_2|r_3),\quad C(+_1r_2r_3|r_4),\quad \cdots,$$
$$C(+_1\cdots r_j|r_{j+1}),\quad \cdots.$$

TABLE III. Correlation probabilities for the atoms in one line of the lattice.

	β^0	β^4	β^6	β^8	β^{10}	β^{12}
$C(+_1\|r_2)$	1	−2	−6	−18	−60	−224
$C(+_1\|r_3)$	1	−2	−8	−30	−112	−430
$C(+_1\|r_4)$	1	−2	−8	−32	−134	−576
$C(+_1\|r_5)$	1	−2	−8	−32	−136	−612
$C(+_1\|r_6)$	1	−2	−8	−32	−136	−614
$C(+_1\|r_7)$	1	−2	−8	−32	−136	−614
$C(+_1r_2\|r_3)$	1	−1	−4	−17	−70	−291
$C(+_1r_2\|r_4)$	1	−1	−5	−24	−109	−487
$C(+_1r_2\|r_5)$	1	−1	−5	−25	−121	−580
$C(+_1r_2\|r_6)$	1	−1	−5	−25	−122	−599
$C(+_1r_2\|r_7)$	1	−1	−5	−25	−122	−600
$C(+_1r_2\|r_8)$	1	−1	−5	−25	−122	−600
$C(+_1r_2r_3\|r_4)$	1	−1	−3	−11	−50	−241
$C(+_1r_2r_3\|r_5)$	1	−1	−4	−17	−78	−376
$C(+_1r_2r_3\|r_6)$	1	−1	−4	−18	−89	−451
$C(+_1r_2r_3\|r_7)$	1	−1	−4	−18	−90	−469
$C(+_1r_2r_3\|r_8)$	1	−1	−4	−18	−90	−470
$C(+_1r_2r_3\|r_9)$	1	−1	−4	−18	−90	−470
$C(+_1\cdots r_4\|r_5)$	1	−1	−3	−10	−38	−170
$C(+_1\cdots r_4\|r_6)$	1	−1	−4	−16	−65	−286
$C(+_1\cdots r_4\|r_7)$	1	−1	−4	−17	−76	−360
$C(+_1\cdots r_4\|r_8)$	1	−1	−4	−17	−77	−378
$C(+_1\cdots r_4\|r_9)$	1	−1	−4	−17	−77	−379
$C(+_1\cdots r_4\|r_{10})$	1	−1	−4	−17	−77	−379
$C(+_1\cdots r_5\|r_6)$	1	−1	−3	−10	−37	−150
$C(+_1\cdots r_5\|r_7)$	1	−1	−4	−16	−64	−265
$C(+_1\cdots r_5\|r_8)$	1	−1	−4	−17	−75	−339
$C(+_1\cdots r_5\|r_9)$	1	−1	−4	−17	−76	−357
$C(+_1\cdots r_5\|r_{10})$	1	−1	−4	−17	−76	−358
$C(+_1\cdots r_5\|r_{11})$	1	−1	−4	−17	−76	−358
$C(+_1\cdots r_6\|r_7)$	1	−1	−3	−10	−37	−149
$C(+_1\cdots r_6\|r_8)$	1	−1	−4	−16	−64	−264
$C(+_1\cdots r_6\|r_9)$	1	−1	−4	−17	−75	−338
$C(+_1\cdots r_6\|r_{10})$	1	−1	−4	−17	−76	−356
$C(+_1\cdots r_6\|r_{11})$	1	−1	−4	−17	−76	−357
$C(+_1\cdots r_6\|r_{12})$	1	−1	−4	−17	−76	−357
$C(+_1\cdots r_7\|r_8)$	1	−1	−3	−10	−37	−149
$C(+_1\cdots r_7\|r_9)$	1	−1	−4	−16	−64	−264
$C(+_1\cdots r_7\|r_{10})$	1	−1	−4	−17	−75	−338
$C(+_1\cdots r_7\|r_{11})$	1	−1	−4	−17	−76	−356
$C(+_1\cdots r_7\|r_{12})$	1	−1	−4	−17	−76	−357
$C(+_1\cdots r_7\|r_{13})$	1	−1	−4	−17	−76	−357

These probabilities show the ordering influence at a site immediately adjacent to stretches of ordered atoms of different lengths. The correlation appears to increase as the length of the ordered stretch increases. For stretches of six or more atoms, however, the correlation probabilities are identical, to terms of the order β^{12}. We can conclude, then, that to this order of accuracy the correlation probability at a point immediately adjacent to a completely ordered half-line, $r_{-\infty}\cdots r_{-2}r_{-1}+_0$ is given by

$$C(r_{-\infty}\cdots r_{-2}r_{-1}+_0|r_1)=1-\beta^4-3\beta^6-10\beta^8-37\beta^{10}-149\beta^{12}-\cdots. \tag{9.5}$$

In the same way, the correlation probabilities for sites 2, 3, 4, 5, $\cdots$ steps away from the ordered half-line are given by

$$\begin{aligned}C(r_{-\infty}\cdots r_{-1}+_0|r_2)&=1-\beta^4-4\beta^6-16\beta^8-64\beta^{10}-264\beta^{12}-\cdots,\\C(r_{-\infty}\cdots r_{-1}+_0|r_3)&=1-\beta^4-4\beta^6-17\beta^8-75\beta^{10}-338\beta^{12}-\cdots,\\C(r_{-\infty}\cdots r_{-1}+_0|r_4)&=1-\beta^4-4\beta^6-17\beta^8-76\beta^{10}-356\beta^{12}-\cdots,\\C(r_{-\infty}\cdots r_{-1}+_0|r_5)&=1-\beta^4-4\beta^6-17\beta^8-76\beta^{10}-357\beta^{12}-\cdots.\end{aligned}\tag{9.6}$$

To terms of order β^{12}

$$C(r_{-\infty}\cdots r_{-1}+_0|r_p)=C(r_{-\infty}\cdots r_{-1}+_0|r_5)\quad p\geqq 6.$$

For the long range effect of the ordered half-line we therefore have the expression

$$C(r_{-\infty}\cdots r_{-1}+_0|r_\infty)=1-\beta^4-4\beta^6-17\beta^8-76\beta^{10}-357\beta^{12}-\cdots. \tag{9.7}$$

Comparing (9.7) and (5.10) we obtain the interesting result that

$$C(r_{-\infty}\cdots r_{-1}+_0|r_\infty)=\tfrac{1}{2}(1+S). \qquad (9.8)$$

The explanation of this identity would seem to be that the constraint of keeping an entire half-line ordered provides sufficient knowledge to decide which sites in the lattice are α sites and which are β sites, in the sense of Section 5. The probability for finding a right atom would then be given in terms of the long range order by (9.8) or (5.6).

PHYSICAL REVIEW VOLUME 64, NUMBERS 5 AND 6 SEPTEMBER 1 AND 15, 1943

Statistics of Two-Dimensional Lattices with Four Components†

J. Ashkin and E. Teller
Pupin Physics Laboratories, Columbia University, New York, New York

(Received June 10, 1943)

We have considered a two-dimensional square net consisting of four kinds of atoms supposing that only nearest neighbors interact and that there are only two distinct potential energies of interaction, one between like and one between unlike atoms. In extension of a method due to Onsager it is found that for the case where like atoms attract one another a simple "reciprocity" relation exists between the partition functions at pairs of temperatures "reciprocally" related to one another. As one temperature T tends to zero, the other T^* tends to infinity. If one further assumes that only one "Curie" transition point exists, the relation between T and T^* enables one to locate the Curie temperature. Predictions can be made concerning the nature of the transition point with results similar to those of Kramers and Wannier. The reciprocity relation for the case of attraction between like atoms is found to be not valid for the case where unlike atoms attract one another.

INTRODUCTION

In a recent paper[1] on the statistics of two-dimensional ferromagnets, Kramers and Wannier have given a treatment of a model for cooperative phenomena. They discovered an important property of their system which expressed itself in the form of a simple symmetrical relation between the partition functions Z and Z^* at pairs of temperatures T and T^*, related in a certain symmetrical way. Any knowledge of the partition function Z at temperature T implies an equivalent amount of knowledge about Z^* at the related temperature T^*. The relation between T and T^* is a one to one relation; T decreases monotonically if T^* increases so that $T\rightarrow 0$ as $T^*\rightarrow\infty$. There is only one temperature at which $T=T^*$. Assuming long range order at $T=0$ and absence of long range order at $T=\infty$, one will expect a Curie point. If the Curie point is the only temperature at which Z becomes singular, this temperature must be the one at which T and T^* become equal. Furthermore, one finds that the types of singularities which might occur at this transition are restricted by the relation between Z and Z^*. For example, if one assumes that there is no latent heat at the Curie point, the specific heat must be either continuous through the transition or infinite on both sides of the transition temperature.

Onsager[2] has been able to obtain the reciprocity relation discussed above from a more direct consideration of the form of the partition function. Moreover, his ingenious methods appear to lend themselves readily to generalization. Onsager proceeds by giving two formally different expressions for the partition function at a tempera-

---

† Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University, New York, New York. Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252 (1941).

[2] This method has not yet appeared in print. We are greatly indebted to Dr. Onsager for his permission to make use of it.