STATISTICAL PROBLEMS IN FERROMAGNETISM, ANTIFERROMAGNETISM AND ADSORPTION

P. W. KASTELEIJN

STATISTICAL PROBLEMS IN FERROMAGNETISM, ANTIFERROMAGNETISM AND ADSORPTION

STATISTICAL PROBLEMS IN FERROMAGNETISM, ANTIFERROMAGNETISM AND ADSORPTION

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN
DOCTOR IN DE WIS- EN NATUURKUNDE
AAN DE RIJKSUNIVERSITEIT TE LEIDEN
OP GEZAG VAN DE RECTOR MAGNIFICUS
DR. A. E. VAN ARKEL, HOOGLERAAR IN
DE FACULTEIT DER WIS- EN NATUURKUNDE,
TEGEN DE BEDENKINGEN VAN
DE FACULTEIT DER WIS- EN NATUURKUNDE
TE VERDEDIGEN OP WOENSDAG
13 JUNI 1956 TE 16 UUR

DOOR

PIETER WILLEM KASTELEIJN

GEBOREN TE LEIDEN IN 1924

Promotor: PROF. DR. S. R. DE GROOT

AAN MIJN OUDERS
AAN MIJN VROUW

Promotor: PROF. DR. S. R. DE GROOT

CONTENTS

INTRODUCTION AND SUMMARY

The main purpose of this thesis is to investigate the statistical properties of several co-operative systems. The first three chapters are devoted to a discussion of ferro- and antiferromagnetic spin systems. Although the properties of these systems are qualitatively well described by the molecular field theory, there exist a number of important discrepancies between theory and experiment. Since the main assumption underlying this theory is that the spins of different atoms are statistically independent of each other, it is evident that for the construction of a better approximation one has to take into account correlations between different spins. In chapter I the statistical theory of a ferromagnetic spin system with isotropic coupling between nearest neighbour spins is formulated in such a way that correlations between neighbouring spins can be introduced in a simple way. To this end the partition function of the spin system is expressed in terms of a pair density matrix of a nearest neighbour pair of spins. With the help of this pair density matrix an effective Hamiltonian H_e for a spin pair is introduced, and it is shown that H_e contains only two coupling terms and a term representing an effective field acting on the two spins of the pair. The molecular field approximation is obtained by assuming that the coupling terms are zero. A next approximation is obtained by assuming that the effective coupling is equal to the actual isotropic coupling between the spins, which is shown to correspond to the limiting behaviour of H_e for high temperatures. This "constant coupling" approximation constitutes a straightforward generalization of the well-known quasi-chemical approximation for an Ising spin system to the case of Heisenberg coupling. The thermodynamic properties of the spin system, and in particular the critical data, are calculated on the basis of this constant coupling approximation and numerical results are given for various lattices. It is shown that in this theory the difficulty of an anti-Curie point does not occur.

In chapter II the theory developed in chapter I is extended to an antiferromagnetic spin system with a two-sublattice structure and isotropic nearest neighbour coupling. Both the case of a parallel and of a perpendicular external field are discussed. In contradistinction to the ferromagnetic case the contant coupling approximation for an antiferromagnetic spin system leads to the appearance of an anti-Curie point.

The application of the theory to ferro- and antiferromagnetic spin systems with Ising coupling is given in chapter III. In this case, the effective Hamiltonian contains only one coupling term. The constant coupling approximation which is obtained by assuming that this effective Ising coupling is constant is shown to be equivalent to the quasi-chemical approximation. For an antiferromagnetic spin system with an external field of arbitrary magnitude in the preferred direction this approximation is worked out in detail. The critical data are evaluated, and an explicit formula is derived for the critical curve in the B *vs* T plane.

In chapter IV the properties of a monolayer of particles, adsorbed on a surface with a two-sublattice structure, with a positive interaction energy (repulsive force) between neighbouring particles, are investigated by making use of the formal relation between the partition function of such a monolayer and that of an antiferromagnetic Ising spin system. On the basis of the constant coupling approximation it is shown that for temperatures lower than a certain critical temperature the adsorption isotherms show two discontinuities corresponding to transitions between the disordered state and a state with long-range order. The heat of adsorption is calculated as a function of the relative covering of the surface.

In chapter V the ground state of a one-dimensional crystal with antiferromagnetic coupling is investigated. Generalizing a procedure introduced by Slater and Hulthén, the energy of this state is approximated by means of a variational method. On the basis of this approximation it is found that, in contradiction to the predictions of the molecular field theory, the ground state of a chain with isotropic coupling shows no antiferromagnetic long-range order. If the coupling has uniaxial anisotropy, however, long-range order occurs if the anisotropy is large enough. The critical value of the anisotropy is calculated and the anisotropy dependence of various quantities is investigated.

Chapter I

CONSTANT COUPLING APPROXIMATION FOR HEISENBERG FERROMAGNETISM

§ 1. *Introduction.* Although the molecular field theory gives a good qualitative description of the statistical properties of ferro- and antiferromagnetic spin systems, there exist a number of important discrepancies between theory and experiment. As appears from the usual derivation of the molecular field theory from the Heisenberg model [1]), the main assumption underlying this theory is that the spins of different atoms are statistically independent of each other. The result is that in this approximation the short-range order and the specific heat vanish above the Curie temperature, contrary to the experimental findings. It is the purpose of this chapter to formulate the statistical theory of a ferromagnetic spin system with Heisenberg coupling between nearest neighbouring spins in such a way that the correlation between neighbouring spins can be taken into account in a simple way. To this end the partition function of the spin system is expressed in terms of the pair density matrix of a pair of nearest neighbouring spins (§ 2), in analogy to the expression of the partition function of a gas in terms of the pair distribution function. This pair density matrix is then used to define an effective Hamiltonian for a pair of nearest neighbouring spins, and this effective Hamiltonian is shown to be of a particularly simple form, which makes it a convenient quantity to work with. With the help of this formalism, the molecular field approximation can be derived in a natural way (§ 3), while the "constant coupling" approximation (§ 4) yields a straightforward generalization to the present case of Heisenberg interaction of Guggenheim's [2]) well-known quasi-chemical approximation for an Ising spin system. The details of this approximation are worked out in §§ 5, 6 and 7, and the extension of the theory to greater spin values is discussed in § 8. In § 9 the present theory is compared to other theories, in particular to the closely related cluster method of P. R. Weiss [3]). Finally, in § 10, some remarks are made about possible improvements of the theory.

§ 2. *General theory.* The Hamiltonian of a lattice of N spins $\frac{1}{2}$ with a ferromagnetic Heisenberg coupling between nearest neighbouring spins is given by

$$H = -2J\,\Sigma_{<i,j>}\,\mathbf{S}_i \cdot \mathbf{S}_j - 2\mu B\,\Sigma_i\,S_{iz}, \tag{1}$$

where $\Sigma_{<i,j>}$ is a sum over all pairs of nearest neighbours; J is the positive coupling constant; $\mathbf{S}_i$ is the vector spin operator, in units $\hbar$, of the i^{th} spin; μ is the magnetic moment per spin; B is the external magnetic field in the z-direction. The density operator of the canonical ensemble describing the equilibrium properties of the spin system (1) is given by

$$\varrho = Z^{-1} \exp(-\beta H), \tag{2}$$

where $\beta \equiv 1/kT$ and Z is the partition function

$$Z = \mathrm{Tr}\,[\exp(-\beta H)] = \Sigma_l \exp(-\beta E_l), \tag{3}$$

in which Tr indicates the trace, the sum runs over all the eigenstates l of the Hamiltonian (1), and E_l is the energy of the state l. The states l can be divided into $N+1$ groups corresponding to the $N+1$ possible values of the z-component of the total spin, $\Sigma_i S_{iz}$, which is a constant of the motion with the eigenvalues $m = -\frac{1}{2}N, -\frac{1}{2}N+1, \ldots, +\frac{1}{2}N$. Instead of m we shall use the long-range order parameter S defined by

$$S = 2m/N, \tag{4}$$

which in the limit $N \to \infty$ can assume all values between -1 and $+1$. Because S is an exact quantum number, we can write the partition function (3) in the form

$$Z = \Sigma_S Z_S, \tag{5}$$

where the sum runs over all values of S, and where Z_S is given by

$$Z_S = \Sigma'_l \exp(-\beta E_l), \tag{6}$$

in which the sum runs over all the eigenstates l of (1) belonging to the given value S of the long-range order parameter (4). Let $S^*(\beta, B)$ be that value of S which, for given values of β and B, makes (6) a maximum. In the limit $N \to \infty$, we may then put in the usual way:

$$\ln Z = \ln Z_{S^*}, \tag{7}$$

where care should be taken that in the limit of vanishing external field one first performs the limit $N \to \infty$ and then makes B tend to zero, ensuring in this way that only the states for which $S \geqslant 0$ are taken into account. From the definition (6) of Z_S, we see that

$$\partial(\ln Z_S)/\partial\beta = -E(\beta, S), \tag{8}$$

where E is the average value, at the temperature $T = 1/k\beta$, of the energy (1) over all the states with a given value of S:

$$E(\beta, S) = Z_S^{-1} \Sigma'_l E_l \exp(-\beta E_l). \tag{9}$$

Following Guggenheim[2]) we integrate eq. (8) at constant S (and B) over β, giving

$$\ln Z_S(\beta) = -\int_0^\beta E(\beta', S)\,\mathrm{d}\beta' + \ln g(S), \tag{10}$$

where use has been made of the fact that at infinitely high temperature ($\beta = 0$) the partition function (6) is equal to the number $g(S)$ of states belonging to the value S of the parameter (4):

$$Z_S(0) = g(S) = \frac{N!}{[\frac{1}{2}N(1+S)]!\,[\frac{1}{2}N(1-S)]!}. \tag{11}$$

From eqs. (7) and (10) we see that the partition function Z of the spin system can be calculated if the function $E(\beta, S)$ is known.

We shall now show that the average energy $E(\beta, S)$ can be expressed in terms of three quantities relating to a pair of nearest neighbouring spins, which have a simple physical significance. To this end we consider the density operator (2) in a representation in which all the S_{iz} are diagonal, *i.e.* we use as basic functions the spin product functions $|s_1 \ldots . s_N\rangle$, where each s_i can assume the two values $\pm\frac{1}{2}$. Consider an operator Ω of the form

$$\Omega = \Sigma_{<i,j>} \Omega_{ij}^{(2)}, \tag{12}$$

where $\Omega_{ij}^{(2)}$ acts only on the nearest neighbouring spins i and j. The average value $\bar{\Omega}(S)$ of Ω, in the ensemble (2), over the states belonging to a given value of S can be written in the form:

$$\bar{\Omega}(S) = \mathrm{Tr}'(\varrho\Omega)/\mathrm{Tr}'\,\varrho = \tfrac{1}{2}Nz\,\mathrm{Tr}\,(\varrho^{(2)}\,\Omega^{(2)}), \tag{13}$$

where $\varrho^{(2)}$ is given by

$$\langle s_1 s_2 \,|\, \varrho^{(2)} |\, s'_1 s'_2\rangle = (Z/Z_S)\,\Sigma^{(R)}_{s_3 \ldots s_N} \langle s_1 s_2 s_3 \ldots s_N\, |\varrho|\, s'_1 s'_2 s_3 \ldots s_N\rangle; \tag{14}$$

Tr′ means the trace over all the states corresponding to the given value of S, z is the co-ordination number, the superscript (R) indicates the restrictive condition

$$\Sigma_{i=3}^N s_i = \tfrac{1}{2}NS - (s_1 + s_2), \tag{15}$$

and the subscripts 1 and 2 refer to a definite pair of nearest neighbouring spins. In virtue of the symmetry properties of the Hamiltonian (1), the matrix (14) is the same for all pairs of nearest neighbouring spins, as can easily be proved.

The 4×4 matrix $\varrho^{(2)}$ defined by (14) can be interpreted as the density matrix of an ensemble of pairs of spins. We can then introduce an *effective energy* of the pair of spins described by $\varrho^{(2)}$ by putting

$$\varrho^{(2)} = \exp(-\beta H_e)/\mathrm{Tr}\,[\exp(-\beta H_e)]. \tag{16}$$

The effective Hamiltonian H_e defined by (16) depends in general parametrically on β and S, but not on B since in the right-hand member of eq. (14) the factors depending on B drop out. The most general form of H_e can be found by the following argument. Since $\Sigma_i S_{iz}$ commutes with H, all the matrix elements $\langle s_1 \ldots s_N|\varrho|\, s_1' \ldots s_N'\rangle$ of ϱ for which $\Sigma_i s_i \neq \Sigma_i s_i'$ vanish so that, according to the definition (14) of $\varrho^{(2)}$, we have

$$\langle s_1 s_2\, |\varrho^{(2)}|\, s_1' s_2'\rangle = 0, \text{ unless } s_1 + s_2 = s_1' + s_2'. \tag{17}$$

In virtue of the equivalence of all lattice sites, and hence of the two lattice sites involved in (14), we further have the symmetry relation

$$\langle s_1 s_2\, |\varrho^{(2)}|\, s_1' s_2'\rangle = \langle s_2 s_1\, |\varrho^{(2)}|\, s_2' s_1'\rangle. \tag{18}$$

From (17) and (18) it follows that, apart from an uninteresting constant term, the effective Hamiltonian must be of the following simple form:

$$H_e = -2A_1 \mathbf{S}_1 \cdot \mathbf{S}_2 - 2A_2 S_{1z} S_{2z} - 2\mu A_3 (S_{1z} + S_{2z}), \tag{19}$$

in which the quantities A_i are functions of β and S, but not of B, which, at least in principle, can be determined from the Hamiltonian (1) by means of the relations (2), (14) and (16). The quantities A_1 and A_2 are effective coupling constants, and the quantity A_3 is an effective field of the nature of the Weiss molecular field.

With the help of the expression (13) for the average value $\overline{\Omega}$ of an operator of the form (12), we can now express the average energy E defined by (9) in terms of the quantities A_i occurring in (19). The Hamiltonian (1) is of the form (12) with $H^{(2)}$ given by

$$H^{(2)} = -2J\, \mathbf{S}_1 \cdot \mathbf{S}_2 - (2\mu B/z)(S_{1z} + S_{2z}), \tag{20}$$

and according to (13) we then have

$$E = \overline{H} = \tfrac{1}{2} Nz\, \mathrm{Tr}\,(\varrho^{(2)} H^{(2)}). \tag{21}$$

Because $H^{(2)}$ and H_e commute, the operators $\varrho^{(2)}$ and $H^{(2)}$ also commute, and we can therefore introduce a representation in which $\varrho^{(2)}$ and $H^{(2)}$ are simultaneously diagonal. Writing out the trace in (21) in this representation the basic functions of which are the singlet and triplet functions of the spin pair, to be denoted by $|0\rangle$ and $|1\rangle$, $|2\rangle$, $|3\rangle$ respectively, we get:

$$E = \tfrac{1}{2} Nz\, \Sigma_{\nu=0}^{3} f_\nu E_\nu, \tag{22}$$

where the eigenvalues E_ν of $H^{(2)}$ are given by

$$\begin{aligned} E_0 &= \tfrac{3}{2}J; & E_1 &= -\tfrac{1}{2}J - 2\mu B/z; \\ E_2 &= -\tfrac{1}{2}J; & E_3 &= -\tfrac{1}{2}J + 2\mu B/z, \end{aligned} \tag{23}$$

and where, according to (16), the eigenvalues f_ν of $\varrho^{(2)}$ are equal to

$$f_\nu = \exp(-\beta\varepsilon_\nu)/\Sigma_\nu \exp(-\beta\varepsilon_\nu), \tag{24}$$

the ε_ν being the eigenvalues of the effective Hamiltonian (19), which are given by

$$\begin{aligned} \varepsilon_0 &= \tfrac{3}{2}A_1 + \tfrac{1}{2}A_2; \quad \varepsilon_1 = -\tfrac{1}{2}A_1 - \tfrac{1}{2}A_2 - 2\mu A_3; \\ \varepsilon_2 &= -\tfrac{1}{2}A_1 + \tfrac{1}{2}A_2; \quad \varepsilon_3 = -\tfrac{1}{2}A_1 - \tfrac{1}{2}A_2 + 2\mu A_3. \end{aligned} \tag{25}$$

We have thus succeeded in expressing in eq. (22) the average energy E defined by (9), and hence, according to (7) and (10), the partition function (3) of the spin system, in terms of the coefficients A_i occurring in the effective Hamiltonian (19). Since we have two relations between the f_ν, namely the normalizing relation

$$\Sigma_\nu f_\nu = 1, \tag{26}$$

which is a consequence of the definition (14) of $\varrho^{(2)}$, and the relation

$$\overline{S_{1z} + S_{2z}} = f_1 - f_3 = S, \tag{27}$$

which can be obtained by calculating by means of (13) the average value, for the given value of S, of $\Sigma_i S_{iz}$, we need to give only two additional relations between the f_ν to determine these quantities, and hence the partition function (3), completely. For instance, we can give $A_1(\beta, S)$ and $A_2(\beta, S)$; the value of $A_3(\beta, S)$ is then determined by eqs. (24)–(27). Examples will be given in the subsequent sections.

Finally we remark that the expression (22) for E may be written in a form that is completely analogous to the equation

$$E_{\text{Is}} = -\tfrac{1}{4}NzJ\sigma - N\mu BS,$$

expressing the average energy E_{Is} of an Ising spin system in terms of the long-range order parameter S and the short-range order parameter σ (*cf.* eq. (III. 2)). To this end we introduce the quantity

$$\tau \equiv 4\overline{\mathbf{S}_1 \cdot \mathbf{S}_2} = f_1 + f_2 + f_3 - 3f_0. \tag{28}$$

Making use of (27) and (28) the expression (22) can be written in the form

$$E = -\tfrac{1}{4}NzJ\tau - N\mu BS, \tag{29}$$

which shows that the quantity τ is the analogue in the Heisenberg case of the short-range order parameter σ used in the Ising case (*cf.* § 7f).

§ 3. *The molecular field approximation.* The molecular field approximation can be obtained in the present formalism by assuming that the effective coupling constants A_1 and A_2 appearing in (19) are zero:

$$A_1 = 0; \quad A_2 = 0. \tag{30}$$

The two relations (30) together with the relations (26) and (27) determine the f_ν completely, and it is not necessary to make any assumptions about the coefficient A_3 in (19). Eq. (30) is the precise formulation of the usual statement that in the molecular field approximation the spins are assumed to be statistically independent, or that the fluctuations in the local field are neglected. From (30) we find:

$$f_0 = f_2 = \tfrac{1}{4}(1 - S^2);\quad f_1 = \tfrac{1}{4}(1 + S)^2;\quad f_3 = \tfrac{1}{4}(1 - S)^2. \tag{31}$$

The quantity (28) is then equal to

$$\tau = S^2, \tag{32}$$

from which we see again that τ is the analogue in the Heisenberg case of the short-range order parameter σ in the Ising case for which eq. (32) is the Bragg-Williams equation.

Substituting (32) in the expression (29) for E, and evaluating the expression (10) for the partition function Z_S, we get:

$$\ln Z_S = \beta(\tfrac{1}{4}Nz\,JS^2 + N\mu BS) + \ln g(S). \tag{33}$$

That value of S which makes (33) a maximum, and which is therefore equal to the average value of S in the ensemble (2), is the solution of the equation

$$S = \tanh \beta(\mu B + \tfrac{1}{2}zJS), \tag{34}$$

which is the basic equation of the molecular field approximation [1]). For this value of S, the effective field A_3 occurring in (19) is given by

$$A_3 = B + \tfrac{1}{2}z\,JS\,\mu^{-1}, \tag{35}$$

which is identical to the usual expression for the Weiss effective field. From the fact that we have obtained (35) by maximizing the partition function (33), *i.e.* by minimizing the corresponding free energy, it is clear that the expression (35) for the effective field leads to the best approximation that can be constructed on the basis of the assumption (30) that the spins are statistically independent.

§ 4. *The constant coupling approximation.* In this section we shall discuss a generalization of the familiar quasi-chemical approximation for an Ising spin system [2]) to the present case of a spin system with Heisenberg coupling. This extension is made possible in virtue of our formulation of the problem in terms of the effective Hamiltonian (19), as we shall now show.

The present formalism can obviously be set up also for an Ising spin system, the only difference being, as can easily be seen, that the coupling constant A_1 appearing in (19) is then rigorously equal to zero, so that the effective Hamiltonian contains only an effective field and an effective Ising coupling, but no effective Heisenberg coupling. The molecular field approximation can then be obtained by putting the remaining coupling constant A_2 equal to zero. Further, as we shall see in chapter III, one can obtain the quasi-chemical approximation for the Ising spin system by assuming that A_2 is independent of β and S, as in the molecular field approximation, but not equal to zero but equal to the actual coupling constant J which can be shown to be the limiting value of A_2 for high temperatures and $S = 0$.

In virtue of this reformulation of the quasi-chemical approximation for an Ising spin system, we can now generalize this method to the case of Heisenberg coupling by assuming that the coupling constants A_1 and A_2 appearing in the effective Hamiltonian (19) for a Heisenberg spin system are both independent of β and S, but not equal to zero as in the molecular field approximation (30), but equal to the correct limiting values for high temperatures and $S = 0$. By means of an extension of the high-temperature expansion method of Opechowski[4]) one can show that for $S = 0$ the high-temperature expansions of A_1 and A_2 are of the following form:

$$A_1 = J + \Sigma_{k=1}^{\infty} a_{1k}(\beta J)^k; \quad A_2 = \Sigma_{k=1}^{\infty} a_{2k}\,(\beta J)^k, \tag{36}$$

where a_{1k} and a_{2k} are constants depending only on the lattice structure; for lattices which do not contain closed triangles of nearest neighbouring sites, a_{11} and a_{21} are equal to zero. We now introduce the assumption that for all values of β and S the quantities A_1 and A_2 are equal to the limiting values of (36) for $\beta \to 0$, *i.e.* we put:

$$A_1 = J; \qquad A_2 = 0. \tag{37}$$

In this approximation the effective coupling (19) is thus assumed to be a pure Heisenberg coupling with a coupling constant equal to the actual coupling constant J. We shall call this the "constant coupling" approximation, which is thus the analogue of the quasi-chemical approximation for Ising systems.

Closely related to the present approximation is the approximation method of P. R. Weiss[3]), which is the analogue for the case of a Heisenberg spin system of the Bethe-Peierls[5])[6]) cluster method developed for Ising spin systems and equivalent co-operative systems. In this theory it is assumed that a cluster of spins consisting of an arbitrary spin of the lattice and its z nearest neighbours behaves statistically as an isolated cluster which is placed in the given external field acting on all its spins, and in an "internal" field acting only on the outer spins. The magnitude of the internal field is fixed by the requirement that the average magnetic moments of the central spin and the outer spins are equal. In the case of an Ising spin system it can be proved[2]) that for lattices which do not contain triangles of nearest neighbouring sites this cluster method is completely equivalent to the quasi-chemical method in virtue of the fact that in those systems the energy of a group of spins is equal to the sum of the energies of the constitutive pairs of neighbouring spins. For Heisenberg systems, however, this is not the case because the pair energy operators $H_{ij}^{(2)}$ do not commute with each other; consequently the constant coupling approximation is not equivalent to the P. R. Weiss approximation. A discussion of the main differences between the two theories will be given in § 9.

§ 5. *Derivation of the basic equation of the constant coupling approximation.* From the basic assumption (37) of the constant coupling approximation we find with the aid of the equations (24) and (25) the following relations for the quantities f_ν:

$$f_1 f_3 / f_2^2 = 1, \tag{38}$$

$$f_2 / f_0 = x^2, \tag{39}$$

where $x \equiv \exp(\beta J)$. These equations, together with (26) and (27), determine the f_ν completely as functions of β and S. We can calculate τ from (28) and substitute the result into (29), after which the integration in (10) can be performed. In order to avoid unnecessary explicit calculations we observe that, according to (19), (27), (28) and (37), we have the relations

$$\begin{aligned} \mathrm{Tr}\,[\varrho^{(2)}\, H_e(\beta, S)] &= -\tfrac{1}{2} J \tau(\beta, S) - 2\mu\, A_3(\beta, S)\, S, \\ \mathrm{Tr}\,[\varrho^{(2)} (\partial H_e / \partial \beta)] &= -2\mu (\partial A_3 / \partial \beta)\, S. \end{aligned} \tag{40}$$

If we introduce the quantity

$$Z_e = \mathrm{Tr}\,[\exp(-\beta H_e)], \tag{41}$$

we find that

$$\begin{aligned} \partial(\ln Z_e)/\partial\beta = -\mathrm{Tr}\,[\varrho^{(2)} H_e] - \beta\, \mathrm{Tr}\,[\varrho^{(2)} (\partial H_e/\partial\beta)] = \\ = \tfrac{1}{2} J\tau + 2\mu S\, [\partial(\beta A_3)/\partial\beta]. \end{aligned} \tag{42}$$

Since S is kept constant, we may write

$$\tfrac{1}{2} J\tau = \partial(\ln Z_e - 2\beta\mu A_3 S)/\partial\beta = \partial(\tfrac{1}{2} J\tau\beta - \Sigma_\nu f_\nu \ln f_\nu)/\partial\beta, \tag{43}$$

where the last expression has been obtained by making use of the relation

$$\Sigma_\nu f_\nu \ln f_\nu = \overline{\ln \varrho^{(2)}} = \tfrac{1}{2}\beta J\tau + 2\beta\mu A_3 S - \ln Z_e, \tag{44}$$

which follows from (16), (40) and (41). We can now integrate eq. (43). At the lower boundary of the integral we have $\beta = 0$, *i.e.* $x = 1$; in this case the f_ν assume the same values (31) as in the molecular field approximation. We get the following expression for the partition function Z_S:

$$\ln Z_S = -\beta E(\beta, S) - \tfrac{1}{2} Nz\, \Sigma_\nu f_\nu \ln f_\nu - (z-1) \ln g(S), \tag{45}$$

where E is the average energy for constant S given by (29). The equilibrium value of S is found by maximizing the right-hand member of eq. (45) with respect to S at constant β. This is most readily done by making use of eq.

(44) again, and of the relation

$$\partial (\ln Z_e)/\partial S = -\beta \operatorname{Tr} [\varrho^{(2)} (\partial H_e/\partial S)] = 2\beta\mu S(\partial A_3/\partial S), \tag{46}$$

which is easily derived from (41) and (19). We find

$$0 = \beta\mu [B - zA_3(\beta, S)] + \tfrac{1}{2}(z-1) \ln [(1+S)/(1-S)], \tag{47}$$

where, according to (24), (25), (27) and (37), $A_3(\beta, S)$ is given by

$$\exp (2\beta\mu A_3) = (f_1/f_3)^{\frac{1}{2}} = [(S^2 + X^2 - X^2S^2)^{\frac{1}{2}} + S]/X(1-S), \tag{48}$$

where X stands for

$$X \equiv 2x^2/(1+x^2) = 2\,[1 + \exp(-2\beta J)]^{-1}. \tag{49}$$

In order to facilitate a comparison with the corresponding equation (34) from the molecular field theory we write (47) in the following form:

$$S = \tanh \beta\, [\mu B + \tfrac{1}{2}zJ\varphi(\beta, S)], \tag{50}$$

where $\varphi(\beta, S)$ is given by

$$\begin{aligned}\varphi(\beta, S) &= (\beta J)^{-1} [\ln \{(1+S)/(1-S)\} - 2\beta\mu A_3] = \\ &= (\beta J)^{-1} \ln \{[(S^2 + X^2 - X^2S^2)^{\frac{1}{2}} - S]/X(1-S)\}.\end{aligned} \tag{51}$$

According to eq. (50) the expression $B + \frac{1}{2}zJ\varphi\mu^{-1}$ can be interpreted as the effective field *for an individual spin,* which should not be confused with the quantity A_3 which is the effective field *for a pair of neighbouring spins* and in this approximation proves to be equal to $B + \frac{1}{2}(z-1)\, J\varphi\mu^{-1}$. In the molecular field approximation the two effective fields are both equal to $B + \frac{1}{2}zJS\mu^{-1}$, as appears from eqs. (34) and (35). Fig. 1 shows φ as a function of S for various values of the temperature; in the limit $T \to \infty$ we get $\varphi = S$.

With the aid of the expression (45) for the partition function Z_S we can now express the thermodynamic quantities of the spin system in terms of β, B and S; for given values of β and B we can then obtain the equilibrium value of S from (50). It is more convenient, however, to use the variables X, y and u, given by eq. (49) and by

$$\begin{aligned} y &\equiv \exp (2\beta\mu B), \\ u &\equiv \exp (2\beta\mu A_3), \end{aligned} \tag{52}$$

in the place of β, B and S. The equilibrium value, for given values of X and y, *i.e.* of T and B, of the quantity u is then given by the equation

$$y = u\left(\frac{u+X}{1+uX}\right)^{z-1}, \tag{53}$$

which is easily obtained from (47) or (50). It is remarkable that eq. (53) which we shall refer to as the basic equation of the constant coupling approximation is formally identical with the well-known Bethe-Peierls equation (III. 6) for an Ising system so that we can use many results derived from the latter equation, provided we take care of the difference between x and $X = 2x^2/(1 + x^2)$. In the next section we shall discuss the solutions of the basic equation (53).

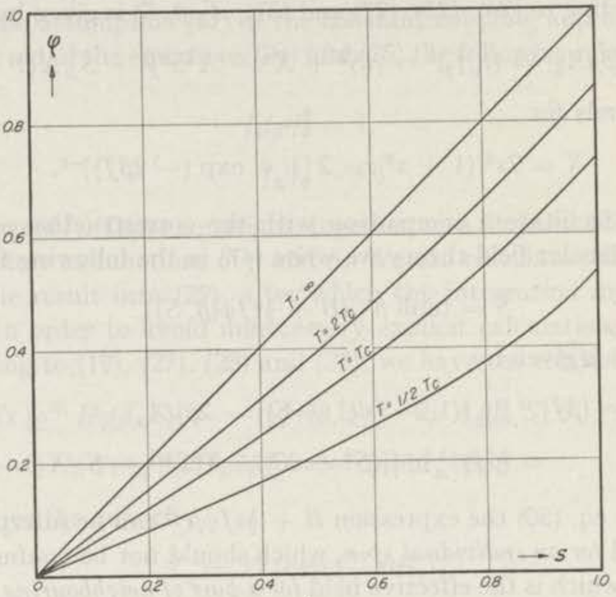

Fig. 1. φ as a function of the relative magnetization $S = M/N\mu$ for $z = 6$ and various values of the temperature.

§ 6. *The Curie temperature.* In the case of a vanishing external field, *i.e.* $y = 1$, eq. (53) has, as is well known, two solutions:

1) A solution $u = 1$ which exists for all temperatures. It can be shown that this solution is stable only for $X \leqslant X_c = z/(z - 2)$. According to (57) and (60) there is no spontaneous magnetization in this case; it is the paramagnetic solution.

2) A solution with $u \neq 1$, which can be written in the implicit form

$$X = \frac{u^{z/(z-1)} - 1}{u - u^{1/(z-1)}}. \tag{54}$$

This solution which corresponds to $S^* > 0$ exists only for $X \geqslant X_c$ and is stable in this temperature region. Hence there is a Curie temperature T_c, given by $X_c = z/(z - 2)$, *i.e.* by

$$x_c = \left(\frac{z}{z-4}\right)^{\frac{1}{2}}, \text{ or } kT_c = 2J\left(\ln \frac{z}{z-4}\right)^{-1}. \tag{55}$$

In table I the values of $T_c^* = 2kT_c/zJ$ are given in the molecular field approximation, the P. R. Weiss[3]) approximation and the constant coupling approximation.

TABLE I

Critical temperature $T_c^* = 2kT_c/zJ$ for various lattices

Lattice	z	molecular field approximation	P. R. Weiss approximation	constant coupling approximation
Linear chain	2	1	none	none
Honey-comb net	3	1	none	none
Quadratic layer	4	1	none	none
Hexagonal layer	6	1	none	0.607
Simple cubic lattice	6	1	0.617	0.607
Body-centered cubic l.	8	1	0.725	0.721
Face-centered cubic l.	12	1	?	0.822

The results listed in the fourth and fifth column are seen to differ only slightly except in the case of the hexagonal plane lattice for which the P. R. Weiss theory predicts the absence of a transition point, in agreement with the results of the spin wave theory, whereas the constant coupling approximation yields a transition point at $T^* = 0.607$. The probable reason for this discrepancy will be discussed in § 9. We observe that in lattices with $z \leqslant 4$ no transition occurs, while the corresponding restriction for Ising lattices is $z \leqslant 2$.

For $B \neq 0$ there occurs no transition; u is a continuous function of x with continuous derivatives.

It is a well-known fact that if we keep zJ constant and let z go to infinity the formulae of the quasi-chemical approximation for an Ising spin system approach to the corresponding formulae of the molecular field theory. We can readily verify that this is also true for lattices with Heisenberg interaction; eq. (55) for the Curie temperature, for instance, goes over into the equation $kT_c = \frac{1}{2}zJ$ which is the equation determining T_c in the molecular field theory. In the treatment given above this property goes back to the fact that in the given limit J tends to zero so that the assumption (37) about A_1 and A_2 becomes identical with the assumption (30) of the molecular field theory.

§ 7. *Calculation of the properties of the spin system.* For the calculation of the magnetic and caloric quantities of the spin system we first express the long-range order parameter S and the short-range order parameter τ given by (27) and (28) in terms of u and x; the equilibrium value of u is then given implicity by eq. (53). With the aid of (38), (39) and (52) we find

$$f_0 : f_1 : f_2 : f_3 = 1 : x^2u : x^2 : x^2/u. \tag{56}$$

Consequently S and τ are equal to

$$S = \frac{x^2(u^2-1)}{u + x^2(u^2+u+1)} = \frac{X(u^2-1)}{Xu^2+2u+X}, \tag{57}$$

$$\tau = \frac{-3u + x^2(u^2+u+1)}{u + x^2(u^2+u+1)}. \tag{58}$$

For $B = 0$, $T \geqslant T_c$ we have $u = 1$ and hence

$$S = 0, \qquad \tau = 3(x^2-1)/(3x^2+1); \tag{59}$$

at the Curie point τ is equal to $3(z-1)^{-1}$.

a. The magnetization is given by

$$M = kT[\partial\,(\ln Z)/\partial B] = N\mu S. \tag{60}$$

In virtue of the formal equivalence of the relations (53) and (57) between u, S, X and y, to the corresponding equations for an Ising spin system (*cf.* ch. III), the M *vs* T curve at $B = 0$ for a Heisenberg spin system (fig. 2) can be

Fig. 2. The long-range order parameter (relative magnetization) $S = M/N\mu$ for $z = 6$ as a function of the temperature in the molecular field approximation (m.f.) and the constant coupling approximation (c.c.).

obtained from the corresponding curve for an Ising system by means of a change in the temperature scale, which is determined by the replacement of

x by $X = 2x^2/(1 + x^2)$. This change of scale, however, has an important consequence: at the absolute zero we have $x = \infty$ and thus $X = 2$, which implies that the spontaneous magnetization has not its maximum value, as it should have, but is equal to the spontaneous magnetization of an Ising spin system at the finite temperature given by $x = 2$:

$$M_0 = N\mu S_0; \qquad S_0 = (u_0^2 - 1)/(u_0^2 + u_0 + 1), \tag{61}$$

where u_0 is that solution of

$$1 = u_0 \left(\frac{u_0 + 2}{2u_0 + 1} \right)^{z-1}, \tag{62}$$

which is > 1. For $z = 6$, 8 and 12 we find $S_0 = 0.955$, 0.991 and 0.9995 respectively. We see that the deviation from the exact value 1 decreases rapidly with increasing z; for $z \to \infty$, S_0 approaches to 1 (*cf.* eq. (64)). The foregoing can also be formulated in this way that for $x \to \infty$ the effective field A_3, and hence, according to eq. (25) the distance between the triplet levels ε_1, ε_2, ε_3, tends sufficiently rapidly to zero to prevent the pairs of the ensemble to fall into the lowest state, state $|1\rangle$; consequently the system as a whole is not in its lowest state even if $T = 0$. The origin of this difficulty must lie in the fact that for very low temperatures the constant coupling approximation is no longer a good approximation; this will be discussed in § 9.

For a non-vanishing external field, however, no difficulty arises. To understand this one should bear in mind that y depends not only on B but also on T so that by (53) u depends on T both through x and through y; therefore the M *vs* T curve is no longer obtained from the corresponding Ising curve by such a simple change of the temperature scale as in the case $B = 0$. If we write

$$y = x^{zB^*}, \qquad B^* \equiv 2\mu B/zJ, \tag{63}$$

we see that for $x \to \infty$ $(T \to 0)$, y tends to ∞ as x^{zB^*} so that the magnetization at the absolute zero is equal to the magnetization of an Ising spin system at $x = 2$ and $y = \infty$; in this case we have $u_0 = \infty$ and hence $S_0 = 1$; the case $B^* = 0$ appears to be a singular case.

For finite temperatures we can expand u, and hence S, in powers of $\gamma \equiv (yX^z)^{-1} = x^{-zB^*} (\frac{1}{2} + \frac{1}{2}x^{-2})^z$:

$$u = (X\gamma)^{-1} [1 - (z-1)(X^2-1)\gamma - \tfrac{1}{2}(z-1)(X^2-1)((z-2)X^2 - z)\gamma^2 \ldots],$$

$$S = 1 - 2\gamma - 2(zX^2 - z - 1)\gamma^2 \ldots. \tag{64}$$

Although this is an expansion for low temperatures only if $B^* \neq 0$, it yields also a useful approximation for $B^* = 0$; in the latter case we can interpret it as an expansion for large z, or, more precisely, as an expansion in powers of X^{-z}. For instance, for $T = 0$, *i.e.* $X = 2$, we have

$$S_0 = 1 - 2^{1-z} - (3z - 1)\, 2^{1-2z} \ldots . \tag{65}$$

Although from a theoretical point of view the singular behaviour of the spin system for $B=0$, $T=0$ ($\lim_{B\to 0} \lim_{T\to 0} S=1 \neq \lim_{T\to 0} \lim_{B\to 0} S=S_0$) is a shortcoming of the present method, the deviations from a regular behaviour are relatively small, the most unfavourable case being that of a lattice with $z = 6$ (lattices with $z < 6$ cannot be ferromagnetic), for which the discrepancy is less than 5 per cent.

b. The susceptibility per spin for $B = 0$ and $T > T_c$ is given by

$$\chi = 4\beta\mu^2x^2/[z - (z - 4)\, x^2]. \tag{66}$$

The $1/\chi$ *vs* T curve is not a straight line as it is in the molecular field approximation but shows a curvature which is most pronounced in the neighbourhood of the Curie point. For the paramagnetic Curie temperature which is the temperature for which the asymptote of the curve (66) intersects the T-axis, we find $kT = \frac{1}{2}zJ$ so that the ratio between the paramagnetic and the ferromagnetic Curie temperature appreciably exceeds the value 1.

c. The energy is calculated from (29), (57) and (58):

$$E = \frac{-\frac{1}{4}NzJ\,[-3u + x^2(u^2 + u + 1)] - N\mu Bx^2(u^2 - 1)}{u + x^2(u^2 + u + 1)}. \tag{67}$$

For $B = 0$, E is a continuous function of T with a discontinuity in the derivative at $T = T_c$ so that we have a transition of the second order; above the Curie temperature we have

$$E = -\tfrac{3}{4}NzJ\,(x^2 - 1)/(3x^2 + 1). \tag{68}$$

d. The jump in the specific heat can be calculated with the aid of the expansion of X in powers of $u - 1$ in the neighbourhood of $X = X_c$

$$X = \frac{z}{z-2}\left(1 + \frac{(u-1)^2}{6(z-1)} + \ldots\right);$$

we obtain

$$\Delta c = \frac{3}{32}\,\frac{z^2(z-4)^2}{(z-1)\,(z-2)}\left(\ln\frac{z}{z-4}\right)^2 Nk. \tag{69}$$

For $z = 6$, 8 and 12 this is equal to $0.81\,Nk$, $1.10\,Nk$ and $1.29\,Nk$ respectively. The only value calculated by P. R. Weiss, namely that for the body-centered cubic lattice ($z = 8$) is much larger ($\Delta c = 2.05\,Nk$), but this discrepancy can be traced back to the use by Weiss of a wrong expression for the mean energy.

e. The entropy S is given by $-kT \ln Z = E - T\mathsf{S}$ so that we get

$$\mathsf{S} = -\tfrac{1}{2}Nzk\,\Sigma_\nu\, f_\nu \ln f_\nu - (z-1)\,k \ln g(S). \tag{70}$$

For $B \neq 0$ we find that $\mathsf{S} \to 0$ for $T \to 0$, but for $B = 0$, S has a finite limit; this fact has the same origin as the fact that in this case M does not reach its saturation value for $T \to 0$.

At the Curie point S is equal to

$$\mathsf{S}_c = Nk\left[\ln 2 - \frac{z}{8(z-1)}\{3z \ln z + (z-4)\ln(z-4) - 4(z-1)\ln(z-1)\}\right], \tag{71}$$

so that, in contrast to the prediction of the molecular field theory the system does not reach its maximum entropy $Nk \ln 2$ at the Curie point; between $T = T_c$ and $T = \infty$ it acquires a relatively important contribution to its entropy, which for $z = 6$, 8 and 12 amounts to 0.314 $Nk \ln 2$, 0.199 $Nk \ln 2$ and 0.115 $Nk \ln 2$ respectively.

f. The short-range order. There are at least two possibilities of defining a short-range order parameter for a Heisenberg spin system. First one can define a parameter σ which is equal to the difference of the probabilities to find equal or opposite values for the z-components of spin of a pair of nearest neighbouring spins. This definition is identical with the definition for the short-range order parameter of an Ising spin system. In our case we get

$$\sigma \equiv 4\overline{S_{1z}S_{2z}} = -f_0 + f_1 - f_2 + f_3 = \frac{x^2(u^2+1) - u(x^2+1)}{x^2(u^2+1) + u(x^2+1)}. \tag{72}$$

The second definition is that of the quantity which measures the difference in probability for a pair of spins of being parallel (in a triplet state) or antiparallel (in a singlet state). This quantity is the quantity τ introduced in § 2, which in the present formalism plays the same role as the quantity σ in the theory of Ising systems. According to (28) and (58) we have

$$\tau \equiv 4\overline{\mathbf{S}_1 \cdot \mathbf{S}_2} = -3f_0 + f_1 + f_2 + f_3 = \frac{x^2(u^2+u+1) - 3u}{x^2(u^2+u+1) + u}.$$

For $B = 0$, $T \geqslant T_c$ we get

$$\sigma = (x^2-1)/(3x^2+1) = (X-1)/(X+1); \tag{73}$$

this is equal to $\frac{1}{3}\tau$, of course, since in the disordered state there is no preferred direction in space so that $S_{1x}S_{2x} = S_{1y}S_{2y} = S_{1z}S_{2z}$. At the Curie point we have $\sigma = (z-1)^{-1}$, just as in the case of an Ising system. For $B = 0$, $T \to 0$ we get: $\sigma \to \sigma_0$, $\tau \to \tau_0$, where

$$\begin{aligned}\sigma_0 &= (u_0-1)/(u_0+1),\\ \tau_0 &= 1;\end{aligned} \tag{74}$$

for $z = 6$, 8 and 12, σ_0 is equal to 0.919, 0.983 and 0.999 respectively. Fig. 3 shows σ and τ as functions of T for $z = 6$ and $B = 0$. We see that the two curves have a large "tail" above the Curie point, the critical values being $\frac{1}{5}$ and $\frac{3}{5}$ respectively.

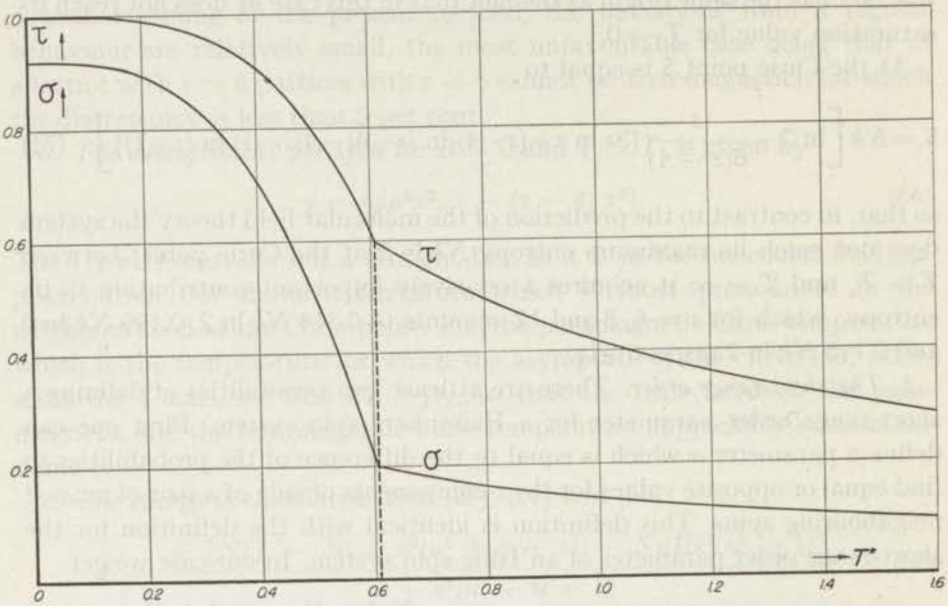

Fig. 3. The short-range order parameters σ and τ for $z = 6$ as functions of the reduced temperature $T^* = 2kT/zJ$ in the constant coupling approximation.

§ 8. *Extension of the theory to higher spin values.* The general theory developed in § 2, and in particular the constant coupling approximation, can be extended to spin values greater than $\frac{1}{2}$. This extension is straightforward and does not present any essentially new aspects. We mention only one result: in the constant coupling approximation the Curie temperature T_c of a system of spins s is determined by the equation

$$\Sigma_{n=0}^{2s} (2n+1) \left[(z-1)\, n(n+1) - 2zs(s+1)\right] \exp\left[n(n+1)\, J/kT_c\right] = 0, \quad (75)$$

which can easily be solved by numerical methods. The results may be compared with the values recently calculated by Brown and Luttinger [10]) using the P. R. Weiss [3]) cluster method. Table II shows the values of $T_c^* = [3kT_c/2zJs(s+1)]$ for the simple cubic (s.c.) and the body-centered cubic (b.c.c.) lattice, and for spin values up to 3, in the molecular field approximation [1]), the P. R. Weiss approximation and the constant coupling approximation.

In the limit $s \to \infty$, eq. (75) goes over into the following equation for the reduced Curie temperature T_c^*:

$$z(z-1)\, T_c^* \left[\exp(6/zT_c^*) - 1\right] - 3\left[(z-2) \exp(6/zT_c^*) + z\right] = 0. \quad (76)$$

The values of T_c^*, calculated from this equation, are also included in table II.

TABLE II

Critical temperatures $T_c^* = [3kT_c/2z\,Js(s+1)]$				
spin	lattice	molecular field approximation	P. R. Weiss approximation	constant coupling approximation
$^1/_2$	s.c.	1	0.617	0.607
	b.c.c.	1	0.727	0.721
1	s.c.	1	0.739	0.737
	b.c.c.	1	0.817	0.811
$^3/_2$	s.c.	1	0.773	0.773
	b.c.c.	1	0.836	0.836
2	s.c.	1	0.788	0.788
	b.c.c.	1	0.847	0.847
$^5/_2$	s.c.	1	—	0.796
	b.c.c.	1	—	0.853
3	s.c.	1	—	0.801
	b.c.c.	1	—	0.855
∞	s.c.	1	—	0.813
	b.c.c.	1	—	0.864

We see that in those cases where the calculations with the aid of the Weiss approximation have been carried out, the results of this approximation and the constant coupling approximation are nearly identical.

§ 9. *Discussion of results and comparison with other theories.* In the preceding sections we have seen that for vanishing external field the constant coupling approximation, in contradistinction to the molecular field theory, predicts the existence of an appreciable amount of short-range order above the Curie point, to which corresponds a "tail" in the energy, specific heat and entropy *vs* T curves, in agreement with experiment. Moreover the curve for the inverse susceptibility as a function of T shows a curvature, giving different values for the paramagnetic and the ferromagnetic Curie temperature, the latter of which was found to lie considerably lower than according to the molecular field theory.

Although much the same results were obtained by the P. R. Weiss[3]) cluster theory, there are several important differences. First the constant coupling approximation can be applied to the paramagnetic state as well as to the ferromagnetic state, while the cluster method is not suited to a treatment of the ferromagnetic state, the formulae being so complicated that even for the calculation of the critical data and of the thermodynamic quantities in the paramagnetic phase series expansions must be used. In our theory simple closed expressions are found both for $T \leqslant T_c$ and for $T > T_c$. The simplicity of the formulae is a consequence of the fact that in the present method a pair of spins is chosen as basic system whereas Weiss considers clusters of $z + 1$ spins. From a theoretical standpoint the present method has the advantage of giving a statistical foundation to the various approximations. In the Weiss method the internal field is introduced ad hoc so that one gets no insight into the nature and the accuracy of the approximation; in the general theory developed in § 2, however, the effective field arises in a natural way and can, at least in principle, be calculated rigorously.

Another advantage of the constant coupling approximation is the absence of an "anti-Curie point". As was pointed out by Anderson[7]), the Weiss method predicts the existence, for the case of vanishing external field, of an anti-Curie point below which no ferromagnetic ordering can exist, which indicates the failure of the theory at low temperatures. In the present approximation, however, no anti-Curie point occurs; the spontaneous magnetization increases continuously if the temperature is lowered from $T = T_c$ to $T = 0$. The only unsatisfactory point is that for $B = 0$ and $T = 0$ the magnetization does not reach its saturation value $N\mu$ but a value that is somewhat lower (at most a few per cent). This shows that for very low temperatures the constant coupling approximation is not a good approximation, which is not surprising because the basic assumption (37) makes it essentially a high-temperature approximation. The inadequacy of the assumption (37) in the lowest temperature region can be verified by calculating by means of the spin wave theory the low temperature behaviour of the parameters A_i; it is found *e.g.* that in the limit $T \to 0$, A_1 tends to zero, in contradiction to the assumption (37) that $A_1 = J$.

Finally the constant coupling approximation is a suitable starting point for the development of further, more accurate approximations (*cf.* § 10), whereas the next step in the P. R. Weiss approximation would be the analogue of the "second approximation" of Bethe[5]), in which the basic cluster is extended to include the next nearest neighbours of the central spin, which method is certainly too complicated to be of any practical use. The Weiss approximation itself, on the other hand, is almost certainly a better approximation than the present one in the case of lattices containing triangles of nearest neighbouring spins, such as the hexagonal plane lattice and the face-centered lattice, where there are pairs of nearest neighbouring spins among the nearest neighbours of a given spin; this is because the interactions between these pairs can be accounted for explicitly in a Bethe-Weiss cluster but not in a pair of spins. Even in this case, however, a higher approximation of the "constant coupling" type will yield more accurate results in a much simpler way.

For an Ising system there is, in addition to the quasi-chemical method and the Bethe-Peierls cluster method, a third equivalent approximation method, *viz.* that of Kikuchi[8]). In an unpublished report Kikuchi[9]) applies this method to a Heisenberg spin system. Postulating a certain expression for the entropy of the system he obtains a first approximation that can be proved to be equivalent to ours. He calculates the Curie temperature for various lattices in first and second approximation and observes that for $T = 0$ the spontaneous magnetization does not reach its saturation value. Further details of his work, however, are not given.

For high temperatures we can compare the results of the various approximation methods to the exact series expansions calculated by Opechowski and Zehler[4]). The only case where a sufficient number of terms has been given in the literature to make discrimination between the various methods possible is that of the susceptibility of a simple cubic lattice. If we expand the expression (66) for the susceptibility in the constant coupling approximation in powers of βJ, we get

$$\chi_{\text{c.c.}} = \beta\mu^2 [1 + \tfrac{1}{2}z(\beta J) + \tfrac{1}{4}z(z-2)(\beta J)^2 + \\ + \tfrac{1}{8}z(z^2 - 4z + \tfrac{8}{3})(\beta J)^3 + \tfrac{1}{16}z(z^3 - 6z^2 + \tfrac{28}{3}z - \tfrac{8}{3})(\beta J)^4 + \ldots]. \tag{77}$$

For the simple cubic lattice we have the following expansions:

$$\begin{aligned} \chi_{\text{m.f.}} &= \beta\mu^2 [1 + 3(\beta J) + 9(\beta J)^2 + 27(\beta J)^3 + 81(\beta J)^4 + \ldots], \\ \chi_{\text{P.R.W.}} &= \beta\mu^2 [1 + 3(\beta J) + 6(\beta J)^2 + \tfrac{2579}{224}(\beta J)^3 + \tfrac{10725}{448}(\beta J)^4 + \ldots], \\ \chi_{\text{c.c.}} &= \beta\mu^2 [1 + 3(\beta J) + 6(\beta J)^2 + 11(\beta J)^3 + 20(\beta J)^4 + \ldots], \\ \chi_{\text{ex.}} &= \beta\mu^2 [1 + 3(\beta J) + 6(\beta J)^2 + 11(\beta J)^3 + \tfrac{165}{8}(\beta J)^4 + \ldots], \end{aligned} \tag{78}$$

where the subscripts m.f., P.R.W. and ex. refer to the molecular field, P. R. Weiss and Opechowski method respectively. Whereas the Weiss method which is most related to our treatment results in an expansion which agrees with the exact one up to the coefficient of $(\beta J)^2$, the present method gives agreement up to the next coefficient and a much more satisfactory approximation to the coefficient of $(\beta J)^4$: for the latter we find the value 20 instead of the value $10725/448 = 23.94$ given by Weiss, while the exact value is $165/8 = 20.63$. This makes us belief that the value of the Curie temperature — which is the temperature for which $1/\chi$ becomes zero — calculated in § 6 on the basis of the constant coupling approximation, is a rather good approximation to the exact value.

§ 10. *Concluding remarks.* The foregoing analysis shows that it is possible to express the partition function of a spin system with Heisenberg interaction between nearest neighbours in terms of the density matrix of a pair of nearest neighbouring spins, and that this pair density matrix can be described in terms of an effective Hamiltonian of a pair of neighbouring spins, which is of a very simple form. This formalism proved to be particularly suited to the introduction of approximation methods, two of which were discussed: the molecular field approximation and the constant coupling approximation.

In principle there are two ways of improving the constant coupling approximation by means of this or a similar method. The first way consists in the replacement of the pair of spins as the basic system by a larger group of spins, *e.g.* a triangle or a square, followed by an assumption about the effective Hamiltonian of such a system, which is the analogue of the assumption (37) for a pair of spins (*cf.* the method developed by Kikuchi[8]) in the theory of order-disorder in Ising systems), the second one in the replacement of the assumption (37) by a more appropriate assumption. The former procedure is advantageous if we want to discriminate between lattices which have the same co-ordination number, but a different topology, such as the simple cubic lattice ($z = 6$) which according to the spin wave theory has a finite Curie temperature, and the hexagonal plane lattice which has the same co-ordination number but is not ferromagnetic. The second procedure is preferable from a general point of view, since the possibility of expressing the partition function (3) in terms of the pair density matrix is a natural consequence of the fact that the Hamiltonian (1) contains only pair interactions. Moreover, it is conjectured that the quantities A_1 and A_2 are continuous functions of T and S, without a jump in any of the derivatives at the Curie point. The reason for this conjecture is that even if one makes the drastic assumption that A_1 and A_2 are independent of T and S, as in the assumptions (30) and (37), one obtains a consistent theory with the appearance of a Curie point for $B = 0$. The singularity occurring in the temperature dependence of the thermodynamic quantities can be traced back to a singularity in the dependence of S^* on T, and is not due to any singularity in the functions $A_1(\beta, S)$ and $A_2(\beta, S)$, although the functions $A_1(\beta, S^*)$ and $A_2(\beta, S^*)$ of the temperature alone show, of course, a singularity at the Curie point. We think therefore that the quantities A_1 and A_2 are particularly well suited to the use of interpolation methods between low- and high-temperature expansions.

REFERENCES

1) Van Vleck, J. H., Rev. mod. Phys. **17** (1945) 27.
2) Fowler, R. H. and Guggenheim, E. A., Statistical thermodynamics, Camb. University Press (1939), Chapters VI, VIII, XIII.
3) Weiss, P. R., Phys. Rev. **74** (1948) 1493.
4) Opechowski, W., Physica **4** (1937) 181.
Zehler, V., Z. Naturf. **5a** (1950) 344.
5) Bethe, H. A., Proc. roy. Soc. A **150** (1935) 552.
6) Peierls, R., Proc. Camb. phil. Soc. **32** (1936) 477.
7) Anderson, P. W., Phys. Rev. **80** (1950) 922.
8) Kikuchi, R., Phys. Rev. **81** (1951) 988.
Kurata, M., Kikuchi, R. and Watari, T., J. chem. Phys. **21** (1953) 434.
9) Kikuchi, R., Quart. Progr. Rept., Solid-State and Molecular Theory Group, M.I.T., Oct. 1951 (unpublished).
10) Brown, H. A. and Luttinger, J. M., Phys. Rev. **100** (1955) 685.

CHAPTER II

CONSTANT COUPLING APPROXIMATION FOR ANTIFERROMAGNETISM

§ 1. *Introduction.* The purpose of this chapter is to extend the theory developed in the preceding chapter, hereafter referred to as I, to antiferromagnetic spin systems with an isotropic coupling between nearest neighbouring spins, and in particular to develop the constant coupling approximation for this case. We shall restrict ourselves to systems with a two-sublattice structure, but we shall treat both the case of a parallel and of a perpendicular external field. The general formalism for these two cases will be developed in § 2 and § 9 respectively, the molecular field approximation will be discussed in § 3 and § 10, while the constant coupling approximation will be considered in §§ 5–8 and § 11. A discussion of the results of the theory will be given in § 12.

A. PARALLEL EXTERNAL FIELD

§ 2. *General theory.* The Hamiltonian of a lattice of N spins $\frac{1}{2}$ with an isotropic, antiferromagnetic coupling between nearest neighbouring spins is given by

$$H = 2J\,\Sigma_{<i,j>}\,\mathbf{S}_i \cdot \mathbf{S}_j - 2\mu\,B\,\Sigma_i\,S_{iz}. \tag{1}$$

We restrict ourselves to lattice structures that can be divided into two sublattices A and B such that all the nearest neighbours of a site of the A-lattice are on the B-lattice, and vice versa.

As is well known, the case of an external field parallel to the sublattice magnetizations can be realized only if there is a preferred direction for the sublattice magnetizations. At a certain critical value of the external field, the threshold field, a transition occurs from the parallel to the perpendicular case. We shall not discuss these threshold phenomena, and we shall restrict ourselves throughout to vanishingly small external fields. We shall neglect the influence of the anisotropy on the properties of the spin system, although a very small anisotropy must, of course, always be present to stabilize the parallel arrangement in the presence of an external field, but this anisotropy is left understood in eq. (1) and the following. It is easy, however, to extend the theory given below to the case in which a staggered anisotropy field is present.

Using the spin product functions $|s_i\rangle \equiv |s_1 \ldots s_N\rangle$, we can write the partition function of the spin system in the form

$$Z = \Sigma_{s_1 \ldots s_N} \langle s_i \,|\, \exp\left(-\beta H\right) |\, s_i\rangle. \tag{2}$$

The sum runs over the 2^N sets of values $s_1, \ldots, s_N$, where each s_i can assume the values $\pm \frac{1}{2}$. If we classify the spin product functions according to the possible values of the order parameters S and s defined as

$$S = (2/N)\, \Sigma_i\, s_i, \tag{3}$$

$$s = (2/N)\, \Sigma_i\, \delta_i\, s_i, \tag{4}$$

where $\delta_i = +1$ for the spins on the A-lattice, and $\delta_i = -1$ for those on the B-lattice, we get

$$Z = \Sigma_{Ss}\, Z(S, s), \tag{5}$$

where the sum runs over all values of S and s, and where

$$Z(S, s) = \Sigma''_{s_i} \langle s_i \,|\, \exp\left(-\beta H\right) |\, s_i\rangle\,, \tag{6}$$

Σ'' being a sum over all the sets $s_1, \ldots, s_N$ satisfying the relations (3) and (4) for the given values of S and s. The quantity $Z(S, s)$ can be written in the invariant form proposed by Kubo 1) 2),

$$Z(S, s) = (2\pi i)^{-2} \iint \Lambda' \operatorname{Tr}\left[\Lambda \exp\left(-\beta H\right)\right] \mathrm{d}\lambda_1\, \mathrm{d}\lambda_2, \tag{7}$$

where

$$\Lambda \equiv \exp\left[\lambda_1 \left(\Sigma_i\, S_{iz}\right) + \lambda_2 \left(\Sigma_i\, \delta_i\, S_{iz}\right)\right], \tag{8}$$

$$\Lambda' \equiv \exp - \left[\lambda_1 \left(\tfrac{1}{2}NS\right) + \lambda_2 \left(\tfrac{1}{2}Ns\right)\right], \tag{9}$$

and in which the boundaries of the two integrations are $-i\infty$ and $+i\infty$. The ratio $Z(S, s)/Z$ is the probability distribution function for the quantities $\frac{1}{2}NS$ and $\frac{1}{2}Ns$ in the given canonical ensemble. Following Kubo 2), we shall assume that in the limit $N \to \infty$ we may put

$$\ln Z = \ln Z(S^*, s^*), \tag{10}$$

where S^* is the value of S which maximizes the quantity

$$Z(S) = \Sigma_s\, Z(S, s), \tag{11}$$

while s^* is the value of s which maximizes $Z(S^*, s)$. Because of the fact that s is not a constant of the motion, it is not possible to give a rigorous justification of the second step, since it is not obvious that the central limit theorem will apply to s. However, Anderson 3) and Kubo 1) 2) have shown on the basis of the spin wave approximation that the fluctuations in s reduce to the normal value if the anisotropy is taken into account, and it is therefore very likely that the assumption (10) does not involve any approximation. In addition to the precaution mentioned in I, § 2, we must take care that only the states with $s \geqslant 0$ are taken into account, since otherwise we would always find an average sublattice magnetization equal to zero. This can be effected by introducing a staggered anisotropy field which for the spins on the A-lattice points in the $+z$-direction and for the spins on the B-lattice in the $-z$-direction. This field can then be made to vanish after the performance of the limit $N \to \infty$.

By the same argument as we used in I, § 2, we now get:

$$\ln Z(S, s) = -\int_0^\beta E(S, s; \beta')\,\mathrm{d}\beta' + \ln g(S, s), \tag{12}$$

where $g(S, s)$ is the number of spin product functions for which the parameters (3) and (4) have the values S and s:

$$g(S, s) = \frac{(\frac{1}{2}N)!}{[\frac{1}{4}N(1 + S + s)]!\,[\frac{1}{4}N(1 - S - s)]!} \cdot \frac{(\frac{1}{2}N)!}{[\frac{1}{4}N(1 + S - s)]!\,[\frac{1}{4}N(1 - S + s)]!}, \tag{13}$$

and where E is the following average value of the energy:

$$E(S, s; \beta) = Z(S, s)^{-1}\, \Sigma''_{s_i} \langle s_i \mid H \exp(-\beta H) \mid s_i \rangle. \tag{14}$$

For an arbitrary operator Ω we define the following "average value for given S and s":

$$\overline{\Omega}(S, s) = [Z/Z(S, s)]\, \Sigma''_{s_i} \langle s_i \mid \tfrac{1}{2}(\varrho\Omega + \Omega\varrho) \mid s_i \rangle, \tag{15}$$

where ϱ is the density operator

$$\varrho = Z^{-1} \exp(-\beta H). \tag{16}$$

In contradistinction to the ferromagnetic case, *cf.* (I. 13), it is necessary to introduce here the symmetrized product $\frac{1}{2}(\varrho\Omega + \Omega\varrho)$ in the definition (15) of $\overline{\Omega}$. Because of the fact that s is not a constant of the motion, $\Sigma'' \langle s_i \mid \varrho\Omega \mid s_i \rangle$ is not necessarily equal to $\Sigma'' \langle s_i | \Omega\varrho | s_i \rangle$, and without the symmetrization, $\overline{\Omega}$ would not be real. For an operator of the form

$$\Omega = \Sigma_{<i,j>}\, \Omega^{(2)}_{ij}, \tag{17}$$

where the sum runs over all pairs of nearest neighbours, and $\Omega^{(2)}_{ij}$ acts only on the spins i and j, we can write (15) in the form

$$\overline{\Omega} = \tfrac{1}{2} Nz \operatorname{Tr}(\varrho^{(2)}\, \Omega^{(2)}), \tag{18}$$

where Tr indicates the trace, and where $\varrho^{(2)}$ is defined by its matrix in the $|s_i\rangle$ representation:

$$\langle s_1 s_2 |\varrho^{(2)}| s_1' s_2'\rangle = \tfrac{1}{2}[Z/Z(S,s)]\,[\Sigma^{(R)}_{s_3 \ldots s_N} \langle s_1 s_2 s_3 \ldots s_N |\varrho| s_1' s_2' s_3 \ldots s_N\rangle + \\ + \Sigma^{(R')}_{s_3 \ldots s_N} \langle s_1 s_2 s_3 \ldots s_N |\varrho| s_1' s_2' s_3 \ldots s_N\rangle]. \quad (19)$$

The superscripts (R) and (R') refer respectively to the restrictive conditions

$$\Sigma_{i=3}^{N} s_i = \tfrac{1}{2} NS - (s_1 + s_2);\ \Sigma_{i=3}^{N} \delta_i s_i = \tfrac{1}{2} Ns - (s_1 - s_2) \quad (20)$$

and

$$\Sigma_{i=3}^{N} s_i = \tfrac{1}{2} NS - (s_1' + s_2');\ \Sigma_{i=3}^{N} \delta_i s_i = \tfrac{1}{2} Ns - (s_1' - s_2'). \quad (21)$$

The density operator $\varrho^{(2)}$ is Hermitian in virtue of the fact that we used the symmetrized product in the definition (15), and the matrix (19) of $\varrho^{(2)}$ is the same for all pairs of nearest neighbouring spins, if 1 and 2 always refer to the A- and the B-lattice respectively. This restriction has the important consequence that the symmetry relation (I.18), *viz.*

$$\langle s_1 s_2 |\varrho^{(2)}| s_1' s_2'\rangle = \langle s_2 s_1 |\varrho^{(2)}| s_2' s_1'\rangle,$$

valid in the ferromagnetic case, does not hold now, and this circumstance constitutes an essential difference between the ferro- and the antiferromagnetic case. The relation (I.17), *viz.*

$$\langle s_1 s_2 |\varrho^{(2)}| s_1' s_2'\rangle = 0 \text{ unless } s_1 + s_2 = s_1' + s_2', \quad (22)$$

on the other hand, holds also in the present case, because S is still a constant of the motion.

As in the ferromagnetic case, we introduce the effective Hamiltonian H_e for a pair of nearest neighbouring spins in the ensemble of pairs described by $\varrho^{(2)}$ by putting

$$\varrho^{(2)} = \exp(-\beta H_e)/\mathrm{Tr}\,[\exp(-\beta H_e)]. \quad (23)$$

However, the most general expression for H_e now contains four rather than three terms:

$$H_e = -2A_1 \mathbf{S}_1 \cdot \mathbf{S}_2 - 2A_2 S_{1z} S_{2z} - 2\mu A_3 (S_{1z} + S_{2z}) - 2\mu A_4 (S_{1z} - S_{2z}). \quad (24)$$

In addition to the quantities A_1, A_2 and A_3 occurring also in the ferromagnetic case (*cf.* (I.19)), the effective Hamiltonian (24) contains a staggered effective field A_4 corresponding to the fact that the two sublattices are no longer equivalent. The quantities A_i are functions of S, s and β, but not of B. Since the Hamiltonian (1) is an operator of the type (17) with

$$H^{(2)} = 2J\mathbf{S}_1 \cdot \mathbf{S}_2 - (2\mu B/z)(S_{1z} + S_{2z}), \quad (25)$$

we can express the average energy E defined by (14) in terms of the quantities A_i. Writing out the trace in (18) in the representation in which $H^{(2)}$ is diagonal, the basic functions of which are the singlet and triplet spin functions to be denoted by $|0\rangle$ and $|1\rangle$, $|2\rangle$, $|3\rangle$ respectively, we get:

$$E = \bar{H} = \tfrac{1}{2} Nz \, \Sigma_{\nu=0}^{3} \, \varrho_{\nu\nu}^{(2)} \, E_\nu, \tag{26}$$

where the eigenvalues E_ν of $H^{(2)}$ are given by

$$\begin{aligned} E_0 &= -\tfrac{3}{2} J; \quad E_1 = \tfrac{1}{2} J - 2\mu B/z; \\ E_2 &= \tfrac{1}{2} J; \quad E_3 = \tfrac{1}{2} J + 2\mu B/z. \end{aligned} \tag{27}$$

For $A_4 \neq 0$, H_e does not commute with $H^{(2)}$, and in contradistinction to the ferromagnetic case, the matrix $\varrho^{(2)}$ is not diagonal in the $H^{(2)}$ representation, the matrix elements $\varrho_{02}^{(2)}$ and $\varrho_{20}^{(2)}$ in general being different from zero. In order to express the matrix elements of $\varrho^{(2)}$ in terms of the parameters A_i we make use of the fact that, according to (23), the eigenvalues of $\varrho^{(2)}$ are given by

$$f_\nu = \exp(-\beta\varepsilon_\nu)/\Sigma_\nu \exp(-\beta\varepsilon_\nu), \tag{28}$$

where the ε_ν are the eigenvalues of the effective Hamiltonian (24):

$$\begin{aligned} \varepsilon_0 &= \tfrac{1}{2} A_1 + \tfrac{1}{2} A_2 - [A_1^2 + 4\mu^2 A_4^2]^{\frac{1}{2}}; \; \varepsilon_1 = -\tfrac{1}{2} A_1 - \tfrac{1}{2} A_2 - 2\mu A_3; \\ \varepsilon_2 &= \tfrac{1}{2} A_1 + \tfrac{1}{2} A_2 + [A_1^2 + 4\mu^2 A_4^2]^{\frac{1}{2}}; \; \varepsilon_3 = -\tfrac{1}{2} A_1 - \tfrac{1}{2} A_2 + 2\mu A_3. \end{aligned} \tag{29}$$

The non-vanishing elements of $\varrho^{(2)}$ are then given by

$$\begin{aligned} \varrho_{00}^{(2)} &= f_0 \cos^2 \tfrac{1}{2}\omega + f_2 \sin^2 \tfrac{1}{2}\omega; \quad \varrho_{11}^{(2)} = f_1; \\ \varrho_{22}^{(2)} &= f_0 \sin^2 \tfrac{1}{2}\omega + f_2 \cos^2 \tfrac{1}{2}\omega; \quad \varrho_{33}^{(2)} = f_3; \\ \varrho_{02}^{(2)} &= \varrho_{20}^{(2)} = \tfrac{1}{2}(f_0 - f_2) \sin\omega, \end{aligned} \tag{30}$$

where ω is defined by

$$\sin\omega = 2\mu A_4 \, [A_1^2 + 4\mu^2 A_4^2]^{-\frac{1}{2}}; \; -\tfrac{1}{2}\pi \leqslant \omega \leqslant \tfrac{1}{2}\pi. \tag{31}$$

The quantity ω determines the eigenfunctions of H_e and is a measure for the difference in behaviour of the spins on the A-lattice from those on the B-lattice. For $\omega = 0$ the eigenfunctions of H_e are equal to those of $H^{(2)}$ and are given by the singlet and triplet functions $2^{-\frac{1}{2}}(\alpha\beta - \beta\alpha)$; $\alpha\alpha$; $2^{-\frac{1}{2}}(\alpha\beta + \beta\alpha)$; $\beta\beta$. For $\omega = \frac{1}{2}\pi$, on the other hand, the eigenfunctions of H_e are $\alpha\beta$; $\alpha\alpha$; $\beta\alpha$; $\beta\beta$. As remarked already in connection with the expression (24) for H_e, the occurrence of ω is characteristic for antiferromagnetism. In the ferromagnetic case H_e and $H^{(2)}$ commute, and we have $\omega \equiv 0$.

There are three relations between the five quantities f_ν and ω, one being the normalizing condition

$$\mathrm{Tr}\, \varrho^{(2)} = \Sigma_\nu f_\nu = 1, \tag{32}$$

which follows from the definition (19) of $\varrho^{(2)}$, while the other two can be obtained by calculating by means of (18) the average values of $\Sigma_i S_{iz}$ and

$\Sigma_i \delta_i S_{iz}$, which are equal to $\frac{1}{2}NS$ and $\frac{1}{2}Ns$ respectively:

$$\overline{S_{1z} + S_{2z}} = f_1 - f_3 = S, \tag{33}$$

$$\overline{S_{1z} - S_{2z}} = (f_0 - f_2) \sin \omega = s. \tag{34}$$

In virtue of the relations (32), (33) and (34), the f_ν and ω are completely determined if we give two additional relations between these quantities; for instance, it is sufficient to give $A_1(S, s, \beta)$ and $A_2(S, s, \beta)$. In the subsequent sections we shall discuss two examples.

Finally, introducing the quantity

$$\begin{aligned}\tau \equiv -4\,\overline{\mathbf{S}_1 \cdot \mathbf{S}_2} &= 3\varrho^{(2)}_{00} - \varrho^{(2)}_{11} - \varrho^{(2)}_{22} - \varrho^{(2)}_{33} = \\ &= f_0\,(1 + 2\cos\omega) - f_1 + f_2\,(1 - 2\cos\omega) - f_3,\end{aligned} \tag{35}$$

we can write the expression (26) in the form

$$E = -\tfrac{1}{4}\,NzJ\tau - N\mu BS, \tag{36}$$

which is formally identical with the corresponding expression (I.29) in the ferromagnetic case.

§ 3. *The molecular field approximation.* The molecular field approximation can be obtained in the present formalism by assuming that the spins are statistically independent, *i.e.* by putting

$$A_1 = 0; \quad A_2 = 0. \tag{37}$$

Making use of the relations (28), (29) and (31)–(34),we deduce from (37) that

$$\begin{aligned}&f_0 = \tfrac{1}{4}\,(1 + S + s)\,(1 - S + s);\ f_1 = \tfrac{1}{4}\,(1 + S + s)\,(1 + S - s);\\ &f_2 = \tfrac{1}{4}\,(1 + S - s)\,(1 - S - s);\ f_3 = \tfrac{1}{4}\,(1 - S + s)\,(1 - S - s);\end{aligned} \tag{38}$$

$$\sin\omega = 1.$$

The quantity (35) is equal to

$$\tau = s^2 - S^2, \tag{39}$$

and is independent of β. Substituting (39) into eq. (36), and evaluating the expression (12) for the partition function $Z(S, s)$, we get

$$\ln Z\,(S, s) = \beta\,[\tfrac{1}{4}\,NzJ\,(s^2 - S^2) + N\mu BS] + \ln g(S, s). \tag{40}$$

The values of S and s which make (40) a maximum, and which are therefore equal to the equilibrium values of these parameters, are the solutions of the equations

$$\begin{cases} S + s = \tanh \beta[\mu B - \tfrac{1}{2}zJ(S - s)], & (41)\\ S - s = \tanh \beta[\mu B - \tfrac{1}{2}zJ(S + s)]. & (42)\end{cases}$$

Since $S + s = 2\,\overline{S_{1z}}$, $S - s = 2\,\overline{S_{2z}}$, eqs. (41) and (42) are identical with the basic equations of the molecular field theory of antiferromagnetism 4). For these values of S and s, the effective fields A_3 and A_4 are given by

$$A_3 = B - \tfrac{1}{2} zJS\mu^{-1}, \tag{43}$$

$$A_4 = \phantom{B - {}} \tfrac{1}{2} zJs\mu^{-1}. \tag{44}$$

For $A_3 + A_4$ and $A_3 - A_4$ we find the usual expressions for the "effective fields" acting on the spins of the A- and B-lattice respectively.

§ 4. *The constant coupling approximation; derivation of the basic equations.* In analogy to the ferromagnetic case (I, § 4) we assume in this approximation that the effective coupling parameters A_1 and A_2 are independent of S, s and β, but not equal to zero as in the molecular field approximation, but equal to the correct limiting values for vanishing S, s and β:

$$A_1 = -J; \quad A_2 = 0. \tag{45}$$

With the help of (28) and (29), the relations (45) can be transformed into two relations between the quantities f_ν and ω, which, in conjunction with the relations (32), (33) and (34) determine the f_ν and ω completely as functions of S, s and β. By a reasoning similar to that given in I, § 4, we then find from (12), (35) and (36) the following expression for the partition function $Z(S, s)$:

$$\ln Z(S, s) = -\beta E(S, s) - \tfrac{1}{2} Nz \, \Sigma_\nu f_\nu \ln f_\nu - (z - 1) \ln g(S, s), \tag{46}$$

where E is given by (36).

The equations determining the equilibrium values of S and s, *i.e.* the values of S and s which maximize (46), can be written in the form

$$\begin{cases} S + s = \tanh \beta \, [\mu B - \tfrac{1}{2} zJ \{\Phi(S, s, \beta) - \varphi(S, s, \beta)\}], & (47) \\ S - s = \tanh \beta \, [\mu B - \tfrac{1}{2} zJ \{\Phi(S, s, \beta) + \varphi(S, s, \beta)\}], & (48) \end{cases}$$

where Φ and φ are defined as

$$\Phi = (\beta J)^{-1} [\tfrac{1}{2} \ln \{(1 - S - s)(1 - S + s)/(1 + S + s)(1 + S - s)\} + \\ + 2\beta\mu A_3], \tag{49}$$

$$\varphi = (\beta J)^{-1} [\tfrac{1}{2} \ln \{(1 + S + s)(1 - S + s)/(1 - S - s)(1 + S - s)\} - \\ - 2\beta\mu A_4]. \tag{50}$$

In (49) and (50) A_3 and A_4 are functions of S, s and β, which in principle can be obtained from the relations (28), (29) and (31)–(34). The quantities Φ and φ are therefore functions of S, s and β, but not of B. We have written the equilibrium conditions for S and s in a form that is closely analogous to that of the corresponding equations (41) and (42) of the molecular field approximation in order to facilitate a comparison between the two approximations.

The limiting values of the quantities Φ and φ for $\beta \rightarrow 0$ are S and s respectively. Fig. 1 shows φ as a function of s for $S = 0$ and various values of β. The function Φ is identically equal to zero for $S = 0$.

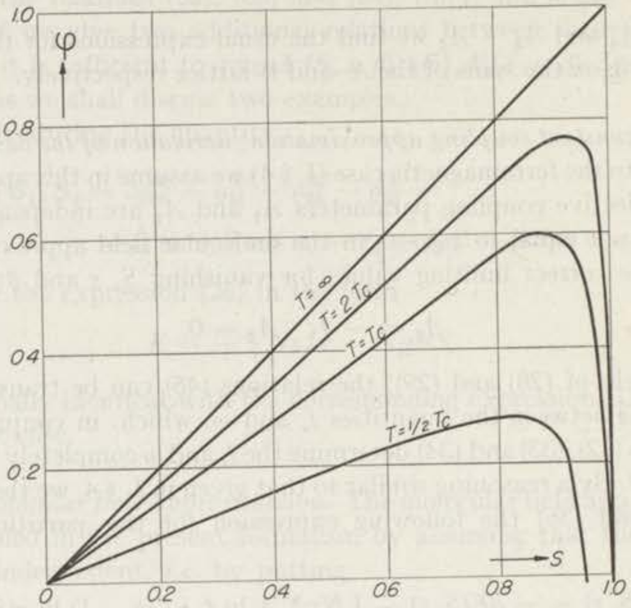

Fig. 1. φ as a function of s for $z = 6$, $S = 0$ and various values of the temperature.

For the calculation of the thermodynamic properties of the spin system it is convenient to introduce in the place of S and s the quantity

$$u \equiv \exp(2\beta\mu A_3), \tag{51}$$

which was also used in the ferromagnetic case, and the quantity ω, defined by (31), which in the constant coupling approximation is given by

$$\sin\omega = 2\mu A_4 \,[J^2 + 4\mu^2 A_4^2]^{-\frac{1}{2}}; \ -\tfrac{1}{2}\pi \leqslant \omega \leqslant \tfrac{1}{2}\pi. \tag{52}$$

In the place of β and B we shall use the quantities x and y defined by

$$x \equiv \exp(\beta J), \tag{53}$$

$$y \equiv \exp(2\beta\mu B). \tag{54}$$

From (28), (29), (45) and (51) we derive

$$f_0 : f_1 : f_2 : f_3 = xv : u : x/v : 1/u, \tag{55}$$

where v is given by

$$v \equiv x^{1/\cos\omega}. \tag{56}$$

The equations determining the equilibrium values of u and ω for given values of x and y, are obtained from (47) and (48) by making use of eqs. (32)–(34), (52) and (55). We get

$$y = u\,[W(1 + uX_-)/(u + X_+)]^{z-1}, \tag{57}$$

$$y = u\,[(1 + uX_+)/W(u + X_-)]^{z-1}, \tag{58}$$

where

$$W \equiv x^{[z/(z-1)]\tan\omega}, \tag{59}$$

$$X_\pm \equiv \tfrac{1}{2}\,[(1 \pm \sin\omega)\,xv + (1 \mp \sin\omega)\,x/v]. \tag{60}$$

The equations (57) and (58) will be referred to as the basic equations of the constant coupling approximation for an antiferromagnetic spin system. The solutions of these equations will be discussed in the subsequent sections.

§ 5. *The paramagnetic and the antiferromagnetic solution.* The basic equations (57), (58) have two solutions, for one of which the sublattices A and B have equal magnetizations, while for the other the sublattice magnetizations are different. This follows from an examination of the relation between u, ω and x obtained from (57) and (58) by elimination of y:

$$(u^2 + 1)\,[W^2X_- - X_+] + u\,[W^2(X_-^2 + 1) - (X_+^2 + 1)] = 0. \tag{61}$$

1) From (59) and (60) we see that for $\omega = 0$ we have $W = 1$, $X_+ = X_- = X$, where X is defined as

$$X \equiv \tfrac{1}{2}(1 + x^2). \tag{62}$$

In this case eq. (61) is satisfied for all values of u. The basic equations reduce to

$$y = u\,[(1 + uX)/(u + X)]^{z-1}, \tag{63}$$

which is to be compared with the basic equation (I.53) for the ferromagnetic case. For all values of y, u is now a continuous function of x without a singularity in the derivatives; for $B = 0$ we have $u \equiv 1$. According to (34), $\omega = 0$ implies that the two sublattices have equal magnetization; this is the paramagnetic solution.

2) If we take $\omega \neq 1$, we obtain a second solution by solving the quadratic equation (61) for u:

$$u = (Q_1Q_3 \pm Q_2Q_4)/(Q_1Q_3 \mp Q_2Q_4), \tag{64}$$

where the following abbreviations have been used:

$$\begin{aligned}
Q_1 &\equiv [(X_+ - 1) - W(X_- - 1)]^{\frac{1}{2}},\\
Q_2 &\equiv [(X_+ + 1) - W(X_- + 1)]^{\frac{1}{2}},\\
Q_3 &\equiv [(X_+ - 1) + W(X_- - 1)]^{\frac{1}{2}},\\
Q_4 &\equiv [(X_+ + 1) + W(X_- + 1)]^{\frac{1}{2}}.
\end{aligned} \tag{65}$$

Inserting (65) in (57) or (58) we get

$$y = \left(\frac{Q_1Q_3 \pm Q_2Q_4}{Q_1Q_3 \mp Q_2Q_4}\right)\left(\frac{Q_1Q_4 \pm Q_2Q_3}{Q_1Q_4 \mp Q_2Q_3}\right)^{z-1}. \tag{66}$$

The equation (64) yields a relation between u and ω depending parametrically on x, while eq. (66) determines ω implicitly as a function of x and y. Since in this case $s \neq 0$, we call this the antiferromagnetic solution. In (64) and (66) the upper sign corresponds to $y > 1$, *i.e.* to $B > 0$, the lower sign to $B < 0$. The case $B = 0$ will be treated in the next section.

§ 6. *Antiferromagnetic solution and critical temperature for vanishing external field.* It follows from eq. (66) that for $B = 0$ we have $Q_2Q_4 = Q_2Q_3 = 0$. Since $Q_4^2 > 0$, this implies $Q_2 = 0$, or

$$W = (X_+ + 1)/(X_- + 1). \tag{67}$$

Since Q_2^2 vanishes for $\omega = 0$ and tends to $-\infty$ for $\omega \to \frac{1}{2}\pi$, the equation $Q_2 = 0$, or (67), has a solution $\omega \neq 0$ if $\partial(Q_2^2)/\partial\omega$ is positive for $\omega = 0$. Now, we have

$$[\partial(Q_2^2)/\partial\omega]_{\omega=0} = 2(x^2 - 1) - \frac{z}{z-1}(x^2 + 3)\ln x, \tag{68}$$

so that there is an antiferromagnetic solution if $2(z - 1)(x^2 - 1) > z(x^2 + 3)\ln x$. For the temperature corresponding to $x = x_c$, where x_c satisfies

$$2(z - 1)(x_c^2 - 1) - z(x_c^2 + 3)\ln x_c = 0, \tag{69}$$

ω is a double solution of (61). The antiferromagnetic solution then coincides with the paramagnetic solution. Equation (69) has two roots x_c, whereas the equation (I.55) for the Curie point of a system with ferromagnetic Heisenberg interaction, and the corresponding equations for systems with ferromagnetic and antiferromagnetic Ising interaction have only one root. This means that, in addition to the upper limit T_c for the temperature region in which an antiferromagnetic phase can exist, there is also a lower limit $T_{c'}$. Calling the former temperature the Curie, or Néel temperature, as usual, the latter one can be called the anti-Curie temperature. In table I the values of $T_c^* = 2kT_c/zJ$ and $T_{c'}^* = 2kT_{c'}/zJ$ are given in the molecular field approximation, the cluster approximation of P. R. Weiss and Li [5]) and the constant coupling approximation.

TABLE I

Critical temperatures T_c^* and $T_{c'}^*$ for different lattices

Lattice	z	molecular field approximation		Weiss-Li approximation		constant coupling approximation	
		T_c^*	$T_{c'}^*$	T_c^*	$T_{c'}^*$	T_c^*	$T_{c'}^*$
simple cubic	6	1	0	0.668	0.327	0.677	0.287
body-centered cubic . . .	8	1	0	0.795	0.178	0.791	0.183

We see that the present approximation yields values of the Curie temperature which differ only slightly from those obtained by Li, whereas the differences in the anti-Curie temperatures are relatively larger. A closer examination shows that eq. (69) has no roots $x_c > 1$ for $z < 6$. The hexagonal plane lattice ($z = 6$) and the face-centered lattice have not been included in the table because the two-sublattice picture does not apply to these lattices. Finally, we see from (69) that, if we keep zJ constant, we have in the limit $z \to \infty : T_c^* = 1$, $T_{c'}^* = 0$, in accordance with the well-known results of the molecular field theory.

The equation for the critical curve in the *B vs T* plane separating the antiferromagnetic and paramagnetic phase, *i.e.* the curve along which $\omega = 0$ is a double solution of the basic equations, can be obtained by performing in eq. (66) the limit $\omega \to 0$. We get

$$y = \left(\frac{q_-q'_- \pm q_+q'_+}{q_-q'_- \mp q_+q'_+}\right)\left(\frac{q_+q'_- \pm q_-q'_+}{q_+q'_- \mp q_-q'_+}\right)^{z-1}, \tag{70}$$

where

$$q_\pm = [X \pm 1]^{\frac{1}{2}};\ q'_\pm = [2(X-1) - \{z/(z-1)\}\,(X \pm 1)\ln x]^{\frac{1}{2}} \tag{71}$$

For a given value of B, eq. (70) determines the transition temperature, while for a given temperature it yields the value of the "transition field", *i.e.* that value of the external field for which the antiferromagnetic long-range order vanishes. One can easily verify that for $y = 1$ eq. (70) reduces to eq. (69). For values of B and T corresponding to points "outside" the critical curve the paramagnetic solution is stable, for points inside the curve the antiferromagnetic solution has a lower free energy.

§ 7. *Calculation of the properties of the spin system.* In order to calculate the magnetic and caloric quantities of the spin system, we express these quantities in terms of u, ω and x with the aid of eq. (55). For the paramagnetic (p.m.) phase we have $\omega = 0$ so that all quantities are expressed in terms of u and x alone, where u is determined by (63). For the antiferromagnetic (a.f.m.) phase we can express u, with the aid of (64), in terms of ω and x, so that the quantities can be expressed in terms of ω and x, ω being determined by (66).

a) The total magnetization M of the lattice, and the difference m of the magnetizations of the A-lattice and the B-lattice are given by

$$M = N\mu S;\ m = N\mu s. \tag{72}$$

According to (33) and (34), the long-range order parameters S and s are equal to

$$\begin{aligned} S &= (u - 1/u)/(xv + u + x/v + 1/u), \\ s &= x(v - 1/v)\sin\omega/(xv + u + x/v + 1/u), \end{aligned} \tag{73}$$

where v is given by (56). We get

$$\text{p.m.: } S = \frac{u^2 - 1}{u^2 + 1 + u(x^2 + 1)}\ ;\ s = 0; \tag{74}$$

$$\text{a.f.m.: } S = \frac{Q_1 Q_2 Q_3 Q_4}{(X_+ X_- - 1)(W^2 - 1)}\ ;\ s = \frac{(X_+ - X_-)(W^2 X_- - X_+)}{(X_+ X_- - 1)(W^2 - 1)}. \tag{75}$$

For $B = 0$, eqs. (75) reduce to

$$S = 0;\ s = (W - 1)/(W + 1). \tag{76}$$

b) For *the susceptibility per spin* for $B = 0$ and $T < T_c$ (parallel susceptibility $\chi_{||}$) we find, after some lengthy but straightforward calculations

$$\chi_{||} = 8\beta\mu^2 W/(W + 1)\,[z(X_+ + X_- W) - (z - 2)(W + 1)]. \tag{77}$$

For the susceptibility for $T \geqslant T_c$ (paramagnetic susceptibility χ) we get

$$\chi = 2\beta\mu^2/[zX - (z - 2)] = 4\beta\mu^2/[zx^2 - (z - 4)]. \tag{78}$$

For $T = T_c$ ($\beta = \beta_c$) the two susceptibilities χ and $\chi_{||}$ are both equal to

$$\chi_c = 4\beta_c\mu^2/[zx_c^2 - (z - 4)], \tag{79}$$

so that the susceptibility is a continuous function of T with a discontinuity in the derivative at the Curie point. For $z = 6$ and $z = 8$ we have $\chi_c = 0.84\,\chi_0$ and $\chi_c = 0.92\,\chi_0$ respectively, where $\chi_0 = \mu^2/zJ$. According to the molecular field and the spin wave theory [2]) χ_0 is the value of the perpendicular susceptibility $\chi_\perp$ at the absolute zero. Whereas in the molecular field $\chi_\perp$ is independent of T, the spin wave theory predicts a slow decrease of $\chi_\perp$ with increasing temperature. The present results are in qualitative agreement with the latter prediction.

c) We shall consider the following *short-range order parameters:*

$$\begin{aligned}
\sigma_x &\equiv -\,4\overline{S_{1x}S_{2x}} = (f_0 - f_2)\cos\omega,\\
\sigma_y &\equiv -\,4\overline{S_{1y}S_{2y}} = (f_0 - f_2)\cos\omega,\\
\sigma_z &\equiv -\,4\overline{S_{1z}S_{2z}} = f_0 - f_1 + f_2 - f_3,\\
\tau^* &\equiv \tfrac{1}{3}(\sigma_x + \sigma_y + \sigma_z) = \tfrac{1}{3}\tau.
\end{aligned} \tag{80}$$

The quantities σ_x, σ_y and σ_z are equal to the difference of the probabilities of finding opposite or equal values for the x-, y- and z-components respectively of a pair of nearest neighbouring spins. The quantity $\tau^* = \frac{1}{3}\tau$ has been introduced as an order parameter rather than τ itself, because, in contrast to the corresponding quantity for a ferromagnetic spin system, the quantity τ can assume values larger than 1, the largest eigenvalue of $-4\mathbf{S}_1 \cdot \mathbf{S}_2$ being 3, and not 1. For $B = 0$ we find

$$\sigma_x = \sigma_y = (\cos\omega)\,(xv - x/v)/(xv + x/v + 2) =$$
$$= (\cot\omega)\,(W - 1)/(W + 1) = s\cot\omega, \tag{81}$$
$$\sigma_z = (xv + x/v - 2)/(xv + x/v + 2).$$

For temperatures above the Curie point we have

$$\sigma_x = \sigma_y = \sigma_z = \tau^* = (x^2 - 1)/(x^2 + 3) = (X - 1)/(X + 1). \tag{82}$$

As in the ferromagnetic case, there is a large amount of short-range order above the Curie point. For $z = 6$ and $z = 8$, the value of τ^* at the Curie point is equal to $\tau_c^* = 0.295$ and $\tau_c^* = 0.181$ respectively.

d) The energy E can be calculated from τ and S with the aid of eq. (36).

e) The jump in the specific heat for $B = 0$ can be calculated with the help of the expansion of x in powers of ω in the neighbourhood of $x = x_c$:

$$x = x_c\left[1 + \frac{z-1}{3z}\,\frac{x_c^2 - 1}{x_c^2 + 3}\,\frac{z(x_c^2 - 1)\,(x_c^2 - 13) + (36x_c^2 + 12)}{z(x_c^2 - 1)\,(x_c^2 - 9) + 16x_c^2}\,\omega^2 + \ldots\right].$$

We obtain

$$\Delta c = 3\,\frac{z-1}{z}\,\frac{(x_c^2 - 1)^2}{(x_c^2 + 3)^4}\,\frac{[z(x_c^2 - 1)\,(x_c^2 - 9) + 16x_c^2]^2}{z(x_c^2 - 1)\,(x_c^2 - 13) + 36x_c^2 + 12}\,Nk. \tag{83}$$

For $z = 6$ and $z = 8$ this is equal to 0.66 Nk and 1.16 Nk respectively. For a simple cubic lattice Li[5]) found a higher value, but his calculations were based on the expression for the mean energy E introduced by P. R. Weiss[5]), which can be shown to be incorrect for $T < T_c$.

f) The entropy S is given by $-kT \ln Z = E - T\mathsf{S}$. From (46) we find

$$\mathsf{S} = -\tfrac{1}{2}Nzk\,\Sigma_\nu\, f_\nu \ln f_\nu - (z - 1)\,k \ln g(S, s). \tag{84}$$

At the Curie point the entropy is equal to

$$\mathsf{S}_c = Nk\,[\ln 2 - z\,\{x_c^2\,(\ln x_c)/(x_c^2 + 3) - \tfrac{1}{2}\ln\tfrac{1}{4}\,(x_c^2 + 3)\}], \tag{85}$$

so that between $T = T_c$ and $T = \infty$ the entropy increases with an amount which for $z = 6$ and $z = 8$ is equal to 0.496 $Nk \ln 2$ and 0.257 $Nk \ln 2$ respectively. This is in qualitative agreement with the experimental results of Friedberg[6]), who found that in $CuCl_2.2H_2O$ nearly one third of the entropy of the spin system at $T = \infty$ is gained above the transition temperature.

§ 8. *Extension of the theory to higher spin values.* The theory developed in the previous sections can be extended to spin values greater than $\frac{1}{2}$. One finds that in the constant coupling approximation the Curie temperature T_c of a system of spins s is determined by the equation

$$[2zJ/kT_c - 4(z-1)]\, s(s+1) + [2zs(s+1)\, J/kT_c + \\ + (z-1)]\, \Sigma_{n=1}^{2s} (2n+1) \exp[-n(n+1)\, J/kT_c] = 0. \quad (86)$$

Table II shows the values of $T_c^* = [3kT_c/2zJs(s+1)]$, calculated by means of this equation, for the simple cubic (s.c.) and the body-centered cubic (b.c.c.) lattice and for values of s up to 3. For comparison, the values of T_c^* in the molecular field approximation [4]) and the Weiss-Li cluster approximation [7]) are also given.

In the limit $s \to \infty$ the values of $T_c{}^*$ are given by the same equation (eq. (I. 76)) as in the ferromagnetic case. These values have also been included in table II.

TABLE II

Critical temperatures $T_c{}^* = [3kT_c/2z\,Js(s+1)]$

spin	lattice	molecular field approximation	Weiss-Li approximation	constant coupling approximation
$^1/_2$	s.c.	1	0.670	0.677
	b.c.c.	1	0.790	0.791
1	s.c.	1	0.774	0.776
	b.c.c.	1	0.847	0.840
$^3/_2$	s.c.	1	0.794	0.794
	b.c.c.	1	0.852	0.852
2	s.c.	1	0.801	0.801
	b.c.c.	1	0.857	0.857
$^5/_2$	s.c.	1	—	0.806
	b.c.c.	1	—	0.860
3	s.c.	1	—	0.809
	b.c.c.	1	—	0.863
∞	s.c.	1	—	0.813
	b.c.c.	1	—	0.864

B. PERPENDICULAR EXTERNAL FIELD

§ 9. *General theory.* We shall now consider the case of a vanishingly small external field in the x-direction, *i.e.* perpendicular to the direction of easiest magnetization. The Hamiltonian of the spin system is given by

$$H = 2J\,\Sigma_{<i,j>}\, \mathbf{S}_i \cdot \mathbf{S}_j - 2\mu B\, \Sigma_i\, S_{ix}, \quad (87)$$

where the arbitrarily small anisotropy along the z-direction has again been left understood. Instead of the partition function $Z(S, s)$ for fixed values of S and s, given by (6) or (7), we now have to consider the following quantity

$$Z(S_x, s_z) = (2\pi i)^{-2} \iint \Lambda' \,\mathrm{Tr}\, [\Lambda \exp(-\beta H)]\, \mathrm{d}\lambda_1\, \mathrm{d}\lambda_2, \quad (88)$$

where

$$\Lambda \equiv \exp[\lambda_1 (\Sigma_i S_{ix}) + \lambda_2 (\Sigma_i \delta_i S_{iz})], \quad (89)$$

$$\Lambda' \equiv \exp - [\lambda_1(\tfrac{1}{2}NS_x) + \lambda_2(\tfrac{1}{2}Ns_z)]. \quad (90)$$

According to Kubo[1][2], (88) is the probability distribution for the quantities $\frac{1}{2}NS_x$ and $\frac{1}{2}Ns_z$, corresponding to the operators $\Sigma_i\, S_{ix}$ and $\Sigma_i\, \delta_i S_{iz}$ respectively. Since these operators do not commute we cannot find a representation which brings (88) into such a simple form as (6); it is necessary to use the invariant expression (88). The total partition function Z is given by

$$Z = \tfrac{1}{4}N^2 \iint Z(S_x, s_z)\, \mathrm{d}S_x\, \mathrm{d}s_z. \tag{91}$$

Following Kubo we assume that in the limit $N \to \infty$ we may put

$$\ln Z = \ln Z(S_x^*, s_z^*). \tag{92}$$

The average value of an operator Ω for given values of S_x and s_z is now defined as

$$\overline{\Omega}(S_x, s_z) = [Z/Z(S_x, s_z)]\, (2\pi i)^{-2} \iint \Lambda' \operatorname{Tr} [\tfrac{1}{2}\Lambda(\varrho\Omega + \Omega\varrho)]\, \mathrm{d}\lambda_1\, \mathrm{d}\lambda_2, \tag{93}$$

where ϱ is the density matrix (16) for the Hamiltonian (87). Using the same argument as in § 2 and in I, § 2, we get

$$\ln Z(S_x, s_z; \beta) = -\int_0^\beta E(S_x, s_z; \beta')\, \mathrm{d}\beta' + \ln Z(S_x, s_z; 0), \tag{94}$$

where $E = \overline{H}$ is the average, in the sense of (93), of the energy (87). Using the relation

$$\operatorname{Tr} \Lambda = \{\exp [\tfrac{1}{2}(\lambda_1^2 + \lambda_2^2)^{\frac{1}{2}}] + \exp [-\tfrac{1}{2}(\lambda_1^2 + \lambda_2^2)^{\frac{1}{2}}]\}^N, \tag{95}$$

we can calculate $\ln Z(S_x, s_z; 0)$ by means of the method of steepest descents. The result is

$$\ln Z(S_x, s_z; 0) = \ln g(S) =$$
$$= -N\, [\tfrac{1}{2}(1 + S) \ln \tfrac{1}{2}(1 + S) + \tfrac{1}{2}(1 - S) \ln \tfrac{1}{2}(1 - S)], \tag{96}$$

where $S = (S_x^2 + s_z^2)^{\frac{1}{2}}$.

For an operator of the form (17) we can write (93) in the form (18), where $\varrho^{(2)}$ is now defined as

$$\langle s_1 s_2 |\varrho^{(2)}| s_1' s_2' \rangle = [Z/Z(S_x, s_z)]\, (2\pi i)^{-2}\, .$$
$$. \iint \Lambda'\, \Sigma_{s_3 \ldots s_N} \langle s_1 s_2 s_3 \ldots s_N |\tfrac{1}{2}(\varrho\Lambda + \Lambda\varrho)| s_1' s_2' s_3 \ldots s_N \rangle\, \mathrm{d}\lambda_1\, \mathrm{d}\lambda_2. \tag{97}$$

The matrix elements (97) are real, and satisfy the symmetry relations

$$\langle s_1 s_2 |\varrho^{(2)}| s_1' s_2' \rangle = \langle s_1' s_2' |\varrho^{(2)}| s_1 s_2 \rangle \tag{98}$$

and

$$\langle s_1 s_2 |\varrho^{(2)}| s_1' s_2' \rangle = \langle -s_2, -s_1 |\varrho^{(2)}| -s_2', -s_1' \rangle, \tag{99}$$

which follow from the symmetry properties of the Hamiltonian (87) and the operators $\varrho\Lambda$ and $\Lambda\varrho$. The effective Hamiltonian defined by (23) now has the general form

$$H_e = -2A_1 \mathbf{S}_1 \cdot \mathbf{S}_2 - 2A_2 S_{1z} S_{2z} - 2A_2' S_{1x} S_{2x} - 2A_2''(S_{1x} S_{2z} - S_{1z} S_{2x}) -$$
$$- 2\mu A_3\, (S_{1x} + S_{2x}) - 2\mu A_4 (S_{1z} - S_{2z}), \tag{100}$$

where the quantities A_i are functions of S_x, s_z and β. The average E of the energy (87) can be written in the form

$$E = -\tfrac{1}{4}NzJ\tau - N\mu BS_x, \tag{101}$$

where τ is defined as

$$\tau \equiv -4\overline{\mathbf{S}_1\cdot\mathbf{S}_2} = -4\,\mathrm{Tr}\,[\varrho^{(2)}\,\mathbf{S}_1\cdot\mathbf{S}_2]. \tag{102}$$

With the aid of (94), (101) and (102) the partition function $Z(S_x, s_z)$ can then be expressed in terms of the quantities A_i. Two relations between the six A_i can be obtained by calculating by means of (18) the average values of $\Sigma_i\, S_{ix}$ and $\Sigma_i\, \delta_i\, S_{iz}$, which are known to be equal to $\tfrac{1}{2}NS_x$ and $\tfrac{1}{2}Ns_z$ respectively, so that we have

$$\mathrm{Tr}\,[\varrho^{(2)}(S_{1x} + S_{2x})] = S_x, \tag{103}$$

$$\mathrm{Tr}\,[\varrho^{(2)}(S_{1z} - S_{2z})] = s_z. \tag{104}$$

To determine the A_i completely, we must therefore give four additional relations between the A_i, two examples of which will be discussed in the subsequent sections.

§ 10. *The molecular field approximation.* In order to obtain the molecular field approximation for the case of a perpendicular external field, we assume that the spins are statistically independent, *i.e.* we put

$$A_1 = A_2 = A_2' = A_2'' = 0. \tag{105}$$

From (105) we can derive that

$$\tau = s_z^2 - S_x^2, \tag{106}$$

and, consequently, that

$$\ln Z(S_x, s_z) = \beta\,[\tfrac{1}{4}\,NzJ(s_z^2 - S_x^2) + N\mu BS_x] + \ln g(S). \tag{107}$$

The equilibrium values of S_x and s_z are given by

$$(S_x/S)\ln[(1+S)/(1-S)] = 2\beta\mu B - \beta zJS_x, \tag{108}$$

$$(s_z/S)\ln[(1+S)/(1-S)] = \beta zJs_z. \tag{109}$$

For the perpendicular susceptibility per spin we find

$$\chi_\perp = \mu^2/zJ, \tag{110}$$

in agreement with the well-known result of the molecular field theory [4]).

§ 11. *The constant coupling approximation.* The constant coupling approximation for the case of a perpendicular external field is obtained by assuming that the effective coupling in (100) is a purely isotropic, antiferromagnetic coupling with a coupling constant equal to J, *i.e.* that

$$A_1 = J;\ A_2 = A_2' = A_2'' = 0. \tag{111}$$

By a reasoning similar to that given in § 4 and in I, § 4, we find that

$$\ln Z(S_x, s_z) = -\beta E(S_x, s_z) - \tfrac{1}{2}Nz\,\Sigma_\nu f_\nu \ln f_\nu - (z-1)\ln g(S), \tag{112}$$

where the quantities f_ν are the eigenvalues of the pair density matrix (97). The equilibrium values of S_x and s_z are then determined by

$$(z-1)\,(S_x/S)\ln[(1+S)/(1-S)] + 2\beta\mu(B - zA_3) = 0, \tag{113}$$

$$(z-1)\,(s_z/S)\ \ln[(1+S)/(1-S)] - 2\beta\mu z A_4 \qquad = 0. \tag{114}$$

In these expressions A_3 and A_4 are functions of S_x, s_z and β, which in principle can be calculated from the equations (103) and (104). In the limit of small external fields, *i.e.* for $\mu B \ll J$, the quantity S is approximately equal to s_z, while in first order the eigenvalues ε_ν of H_e are given by

$$\begin{aligned}
\varepsilon_0 &= [-\tfrac{1}{2} - 1/\cos\omega - \tfrac{1}{2}(A_3/A_4)^2\,(1-\cos\omega)^2/\cos\omega]\,J,\\
\varepsilon_1 &= \tfrac{1}{2}J,\\
\varepsilon_2 &= [-\tfrac{1}{2} + 1/\cos\omega + \tfrac{1}{2}(A_3/A_4)^2\,(1+\cos\omega)^2/\cos\omega]\,J,\\
\varepsilon_3 &= [\tfrac{1}{2} - 2(A_3/A_4)^2]\,J,
\end{aligned} \tag{115}$$

where, in accordance with eq. (52), $\cos\omega$ is used as an abbreviation of $J[J^2 + 4\mu^2 A_4^2]^{-\frac{1}{2}}$. From (103) and (104) it follows that in this approximation

$$\begin{aligned}
S_x &= (2\beta\mu)^{-1}\,[\partial\,(\ln \Sigma_\nu \exp(-\beta\varepsilon_\nu))/\partial A_3] =\\
&= (A_3/A_4)\left[\frac{1-\cos\omega}{1+\cos\omega}f_0 - \frac{1+\cos\omega}{1-\cos\omega}f_2 + \frac{4\cos\omega}{\sin^2\omega}f_1\right]\sin\omega,
\end{aligned} \tag{116}$$

$$s_z = (f_0 - f_2)\sin\omega, \tag{117}$$

where the f_ν are the eigenvalues of $\varrho^{(2)}$ for $B = 0$. Since in the limit of vanishing external field there is no difference between the "perpendicular" and the "parallel" case, the f_ν are equal to the quantities given by eq. (55), with $u = 1$.

Eliminating A_3, A_4 and s_z from (113), (114), (116) and (117) we get the following expression for the perpendicular susceptibility per spin:

$$\begin{aligned}
\chi_\perp &= \frac{\mu S_x}{B} =\\
&= \frac{\mu^2}{zJ}\,\frac{(X_+ - X_-)[(1+\cos^2\omega)(X_+ - X_-) - 2\sin\omega\cos\omega(X_+ + X_- - 2)]}{(X_+ + X_- + 2)[\sin^2\omega(X_+ + X_- - 2) - \sin\omega\cos\omega(X_+ - X_-)]}.
\end{aligned} \tag{118}$$

For $T = T_c$, $\chi_\perp$ is equal to the critical value (79) of the paramagnetic and the parallel susceptibility, so that the susceptibility is continuous at the Curie point.

§ 12. *Concluding remarks.* A comparison of the molecular field and the constant coupling approximation shows that essentially the same differences exist between the two approximations in the antiferromagnetic case as in the ferromagnetic case discussed in I. According to the constant coupling approximation there is an appreciable amount of short-range order above the Curie point, giving rise to a "tail" in the energy-, specific heat-, and entropy *vs* T curves, and to a curvature in the $1/\chi$ *vs* T curve leading to a Curie point lying considerably lower than the molecular field value. Finally, the parallel, the perpendicular and the paramagnetic susceptibilities are equal to each other at the Curie point, the common value being smaller than the value of the perpendicular susceptibility at the absolute zero in the molecular field and the spin wave theory. On comparing these results with those of the cluster theory of Li[5]), we see that most of what has been remarked in I about the relation between the constant coupling approximation and the cluster theory of Weiss for the ferromagnetic case applies also to the present case of antiferromagnetism. The main advantages of the present method are the fact that the general formalism is based on an exact expression for the partition function of the complete spin system, and the fact that in the constant coupling approximation the relevant formulae are relatively simple, thus making the analysis tractable even below the Curie temperature. However, in contradistinction to the ferromagnetic case, the constant coupling approximation now leads to an anti-Curie temperature of the same order of magnitude as in the theories of Weiss and Li. Consequently, it is doubtful whether the constant coupling approximation gives a useful approximation below the Curie temperature. Thus, the high-temperature character of the constant coupling approximation is more pronounced in the antiferromagnetic case than in the ferromagnetic case, and to obtain more reliable results below the Curie temperature it will be necessary to replace the assumptions (45) and (111) of the constant coupling by more appropriate ones. Above the Curie temperature we can compare the results of the present and other approximations to the exact series expansion in powers of $1/T$. Apart from the replacement of J by $-J$ there is no difference with the ferromagnetic case discussed in I.

REFERENCES

1) Kubo, R., Phys. Rev. **87** (1952) 568.
2) Nagamiya, T., Yosida, K. and Kubo, R., Adv. in Phys. **4** (1955) 1.
3) Anderson, P. W., Phys. Rev. **86** (1952) 694.
4) Van Vleck, J. H., J. chem. Phys. **9** (1941) 85.
5) Weiss, P. R., Phys. Rev. **74** (1948) 1493.
Li, Y. Y., Phys. Rev. **80** (1950) 457.
6) Friedberg, S. A., Physica **18** (1952) 714.
7) Brown, H. A. and Luttinger, J. M., Phys. Rev. **100** (1955) 685.

CHAPTER III

CONSTANT COUPLING APPROXIMATION FOR ISING SPIN SYSTEMS

§ 1. *Introduction.* In this chapter the theory developed in the preceding chapters, hereafter referred to as I and II, is applied to spin systems with Ising interaction between nearest neighbours. As we shall show, the constant coupling approximation is in this case equivalent to the familiar quasi-chemical approximation [1]) [2]) [3]), and hence, for lattices which do not contain triangles of nearest neighbours, also to the cluster approximation of Bethe [4]). The latter approximation, originally introduced in the theory of superstructure in binary alloys, was applied to ferromagnetism by Peierls [5]); the theory was elaborated by Firgau [6]). Extending this treatment to antiferromagnetic systems with a magnetic field of arbitrary magnitude in the preferred direction, Firgau (*l.c.*) and Ziman [7]) derived the "basic equations" determining the temperature- and field dependence of the thermodynamic quantities, but they did not give an explicit solution of these equations for the antiferromagnetic phase. In this chapter we shall describe the antiferromagnetic equilibrium state in an analogous way as the paramagnetic state and the equilibrium state of a ferromagnetic system, namely by expressing the thermodynamic quantities in terms of one auxiliary quantity which has a simple physical significance, and by deriving a relation between the equilibrium value of this quantity, the temperature and the external field.

In § 2 we discuss the application of the theory developed in I, and in particular of the constant coupling approximation, to a ferromagnetic Ising spin system. In § 3 we apply this approximation to antiferromagnetic systems in an external field in the preferred direction, *i. e.* parallel to the spins. The Ising model is not suited to a treatment of the more general case where the external field has an arbitrary direction relative to the preferred direction. In §§ 3–5 this application is worked out in detail.

§ 2. *The constant coupling approximation for a ferromagnetic Ising spin system.* The Hamiltonian of a lattice of N spins $\frac{1}{2}$ with ferromagnetic Ising interaction between nearest neighbours is given by

$$H = -2J\,\Sigma_{<i,j>} S_{iz}S_{jz} - 2\mu B\,\Sigma_i S_{iz}. \tag{1}$$

The problem of determining the partition function Z of the system can be formulated in the same way as for a system with Heisenberg interaction (*cf.* I, § 2). In the representation of the spin product functions $|s_i\rangle$ the Hamiltonian (1) is diagonal; consequently the density matrix ϱ of the spin system, the pair density matrix $\varrho^{(2)}$, and the effective pair Hamiltonian H_e, connected to H by the relations (I.2), (I.14) and (I.16), are also diagonal in this representation. It follows that the isotropic-coupling parameter A_1 occurring in (I.19) is rigorously equal to zero so that the effective Hamiltonian contains only an effective Ising coupling and an effectieve field but no effective Heisenberg coupling. If the anisotropic-coupling parameter A_2 is given as a function of the long-range order parameter S and of $\beta \equiv 1/kT$, the partition function of the spin system can be calculated in the way indicated in I. For $\varrho^{(2)}$ and H_e we can use the representation of the spin product functions $|s_1s_2\rangle$ in the place of the singlet and triplet functions used in I; we let the indices $\nu = 0, 1,$ 2 and 3 refer to the functions $|+\frac{1}{2}, -\frac{1}{2}\rangle$, $|+\frac{1}{2}, +\frac{1}{2}\rangle$, $|-\frac{1}{2}, +\frac{1}{2}\rangle$ and $|-\frac{1}{2}, -\frac{1}{2}\rangle$ respectively. For the average energy E at a given value of S we find

$$E = -\tfrac{1}{2}\,NzJ\sigma - N\mu BS, \tag{2}$$

where, in accordance with the usual definition, the short-range order parameter σ is given by

$$\sigma \equiv 4\,\overline{S_{1z}S_{2z}} = -f_0 + f_1 - f_2 + f_3, \tag{3}$$

the f_ν being the eigenvalues of the pair density matrix $\varrho^{(2)}$.

It is readily verified that the assumption that there is no effective coupling between two neighbouring spins, *i.e.* that $A_2 = 0$, leads, also in this case, to the molecular field approximation [8]). In order to go beyond this approximation we assume that A_2 has a constant value which is not equal to zero but equal to the correct limiting value for $T \to \infty$ and $S = 0$. By expanding A_2, by means of a method analogous to the high-temperature expansion methods of Opechowski [9]) and Kirkwood [10]), into a power series in βJ we find that this value is equal to the actual coupling constant J. We thus put

$$A_2 = J. \tag{4}$$

The constant coupling approximation for an Ising spin system obtained in this way is equivalent to the well-known quasi-chemical approximation [1]) [2]) [3]), first introduced in the theory of liquid and solid mixtures. This can be seen from the relations

$$f_1 f_3 / f_0 f_2 = x^2; \quad f_0 = f_2, \tag{5}$$

which are easily derived from the assumption (4). Applying the procedure developed in I, §§ 2, 4 — which is now completely equivalent to Guggenheim's method [3]) of deriving the equation describing the equilibrium properties of an "Ising system" — we find the following relation between the quantities y, x and u, introduced in I:

$$y = u\,[(u + x)/(1 + ux)]^{z-1}. \tag{6}$$

Eq. (6) is identical to the equation obtained by Bethe [4]) and Peierls [5]) in a completely different way. It should be noted, however, that, whereas the quasi-chemical and the constant coupling approximation are equivalent for arbitrary lattices, the method of Bethe and Peierls is equivalent to the quasi-chemical, and hence to the present method only for lattices which do not contain triangles of nearest neighbours. For "triangle lattices" the Bethe-Peierls equation is different from (6).

For comparison with the basic equation (I.34) of the molecular field theory we write eq. (6) in the form

$$S = \tanh \beta\,[\mu B + \tfrac{1}{2} z J \varphi(S, \beta)], \tag{7}$$

where $\varphi(S, \beta)$ is given by

$$\varphi(S, \beta) = (\beta J)^{-1} \ln \{[(S^2 + x^2 - x^2 S^2)^{\frac{1}{2}} - S]/x(1 - S)\}. \tag{8}$$

Fig. 1 shows φ as a function of S for various values of the temperature. It follows from eq. (6) that there is a Curie temperature T_c given by

$$x_c = z/(z - 2), \text{ or } kT_c = J/\ln\,[z/(z - 2)]. \tag{9}$$

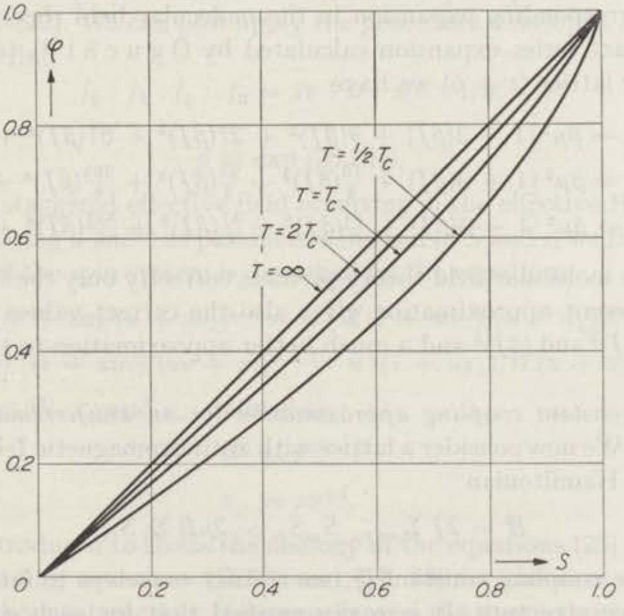

Fig. 1. φ as a function of the relative magnetization $S = M/N\mu$ for $z = 6$ and various values of the temperature.

For $B = 0$ and $T < T_c$, the "basic equation" (6) can be written in the form

$$x = (u^{z/(z-1)} - 1)/(u - u^{1/(z-1)}). \tag{10}$$

For the calculation of the magnetic and caloric quantities of the spin system we express these quantities in terms of u and x; the equilibrium value of the quantity u for given values of x and y (*i.e.* of T and B) is then determined by eq. (6). For the *magnetization* and the *energy* we find the expressions

$$M = N\mu x(u^2 - 1)/(xu^2 + 2u + x), \tag{11}$$

$$E = -\tfrac{1}{4} NzJ(xu^2 - 2u + x)/(xu^2 + 2u + x) - MB. \tag{12}$$

In addition to the results given in the literature [6]) we can also calculate the *entropy at the Curie point*. Making use of the relation $-kT \ln Z = E - TS$ we find

$$S_c = Nk\,[\ln 2 - \tfrac{1}{4} \frac{z}{z-1} (z \ln z + (z-2) \ln (z-2) - 2(z-1) \ln (z-1))]. \tag{13}$$

This differs from the maximum entropy $Nk \ln 2$ of the spin system by an amount which for $z = 6$, 8 and 12 is equal to 0.087 $Nk \ln 2$, 0.059 $Nk \ln 2$ and 0.036 $Nk \ln 2$ respectively.

Further we want to compare the high-temperature expansion of the expression for the *susceptibility* for $B = 0$ and $T > T_c$ in the constant coupling approximation (c.c.),

$$\chi = 2\beta\mu^2 x/[z - (z-2)\,x], \tag{14}$$

with the corresponding expansion in the molecular field theory (m.f.) and with the exact series expansion calculated by Oguchi [11]) (ex.). For the simple cubic lattice ($z = 6$) we have

$$\begin{aligned} &\text{m.f.: } \chi = \beta\mu^2\,[1 + 3(\beta J) + 9(\beta J)^2 + 27(\beta J)^3 + 81(\beta J)^4 + \ldots], \\ &\text{c.c.: } \chi = \beta\mu^2\,[1 + 3(\beta J) + \tfrac{15}{2}(\beta J)^2 + \tfrac{37}{2}(\beta J)^3 + \tfrac{365}{8}(\beta J)^4 + \ldots], \\ &\text{ex.: } \chi = \beta\mu^2\,[1 + 3(\beta J) + \tfrac{15}{2}(\beta J)^2 + \tfrac{37}{2}(\beta J)^3 + \tfrac{353}{8}(\beta J)^4 + \ldots]. \end{aligned} \tag{15}$$

Whereas the molecular field theory predicts correctly only the coefficient of (βJ), the present approximation gives also the correct values of the coefficients of $(\beta J)^2$ and $(\beta J)^3$ and a much better approximation to that of $(\beta J)^4$.

§ 3. *The constant coupling approximation for an antiferromagnetic Ising spin system.* We now consider a lattice with antiferromagnetic Ising coupling, *i.e.* with the Hamiltonian

$$H = 2J\, \Sigma_{<i,j>}\, S_{iz} S_{jz} - 2\mu B\, \Sigma_i\, S_{iz}, \tag{16}$$

with positive coupling constant J; we restrict ourselves to lattices with a two-sublattice structure. It is easily verified that for such a system the effective Hamiltonian H_e for a pair of neighbouring spins contains, in addition to the two terms occurring also in the ferromagnetic case, a term representing the interaction with a field pointing in opposite directions for the two spins. If, for this case, we define the short-range order parameter σ as

$$\sigma \equiv -4\,\overline{S_{1z}S_{2z}} = f_0 - f_1 + f_2 - f_3, \tag{17}$$

we get the same expression (2) for the energy as in the case of a ferromagnetic spin system. The partition function of the spin system can then be calculated in the usual way if the coupling parameter A_2 is given as a function of the long-range order parameters S and s, and of β. The quasi-chemical approximation is obtained by assuming that for all values of S, s and β, A_2 is equal to its limiting value for vanishing S, s and β. Evidently this value is equal to $-J$ for an antiferromagnetic system. We thus put

$$A_2 = -J. \tag{18}$$

From this assumption we easily find that the f_ν satisfy the relation

$$f_0 f_2 / f_1 f_3 = x^2. \tag{19}$$

The other relations between the f_ν are

$$\Sigma_\nu f_\nu = 1, \tag{20}$$

$$f_1 - f_3 = S, \tag{21}$$

$$f_0 - f_2 = s. \tag{22}$$

Eq. (22) comes in the place of the corresponding equation (II.34) for a spin system with isotropic interaction. The f_ν are completely determined by the relations (19)–(22). We can now apply the procedure developed in I, §§ 2, 4. For the f_ν we find

$$f_0 : f_1 : f_2 : f_3 = xv : u : x/v : 1/u, \tag{23}$$

where

$$v \equiv \exp\,(2\beta\mu A_4), \tag{24}$$

A_4 being the staggered effective field occurring in the effective Hamiltonian (*cf.* (II.24)). Using u and v as parameters instead of S and s, we find that the equilibrium of the spin system is described by the equations

$$y = uv\,[(v + xu)/(u + xv)]^{z-1} = u\,[W(1 + ux_-)/(u + x_+)]^{z-1}, \tag{25}$$

$$y = (u/v)\,[(1 + xuv)/(uv + x)]^{z-1} = u\,[(1 + ux_+)/W(u + x_-)]^{z-1}. \tag{26}$$

The quantities W, x_+ and x_-, defined as

$$W \equiv v^{z/(z-1)}, \tag{27}$$

$$x_\pm \equiv xv^{\pm 1}, \tag{28}$$

have been introduced to stress the analogy of the equations (25) and (26) to the corresponding equations (II.57) and (II. 58) for the case of isotropic nearest neighbour interaction. If we now express the thermodynamic quantities of the system in terms of u, v and x, we can, at least in principle, determine the equilibrium values of the quantities u and v for given values of T and B from (25) and (26). In the present form, however, these "basic equations" (which were earlier obtained, in a more complicated form, by Firgau[6]) and Ziman[7])) are not suited to actual calculations. Therefore we shall, in the next section, bring them into a more tractable form.

§ 4. *The paramagnetic and the antiferromagnetic solution; the Curie temperature.* The basic equations (25) and (26) have two solutions, for one of which the two sublattices have equal magnetizations, while for the other the sublattice magnetizations are different.

1) For $v = 1$ we have $W = 1$, $x_{\pm} = x$, and eqs. (25) and (26) reduce to

$$y = u\,[(1 + xu)/(u + x)]^{z-1}, \tag{29}$$

which is to be compared with the basic equation (6) for a ferromagnetic Ising spin system. For all values of y, u is now a continuous function of x without any singularity in the derivatives; for $y = 1$ (vanishing external field) we have $u \equiv 1$. From the definition of v we see that in this case A_4 vanishes so that the two sublattices are equivalent; this is the paramagnetic solution.

2) If we take $v \neq 1$, a second solution is obtained by eliminating y from (25) and (26) and solving the resulting quadratic equation for u:

$$u = (Q_1Q_3 \pm Q_2Q_4)/(Q_1Q_3 \mp Q_2Q_4), \tag{30}$$

where the following abbreviations have been used:

$$\begin{aligned} Q_1 &\equiv [(xv - 1) - v^{1/(z-1)}\,(x - v)]^{\frac{1}{2}}, \\ Q_2 &\equiv [(xv + 1) - v^{1/(z-1)}\,(x + v)]^{\frac{1}{2}}, \\ Q_3 &\equiv [(xv - 1) + v^{1/(z-1)}\,(x - v)]^{\frac{1}{2}}, \\ Q_4 &\equiv [(xv + 1) + v^{1/(z-1)}\,(x + v)]^{\frac{1}{2}}. \end{aligned} \tag{31}$$

In this case A_4 does not vanish; this is the antiferromagnetic solution. Inserting (30) into (25) or (26) we get

$$y = \left(\frac{Q_1Q_3 \pm Q_2Q_4}{Q_1Q_3 \mp Q_2Q_4}\right)\left(\frac{Q_1Q_4 \pm Q_2Q_3}{Q_1Q_4 \mp Q_2Q_3}\right)^{z-1}, \tag{32}$$

which is formally identical with the corresponding equation (II.66) for a spin system with isotropic interaction. The upper sign in (30) and (32) corresponds to $B > 0$ ($y > 1$), the lower sign to $B < 0$. For $B = 0$ we have $Q_2Q_4 = Q_2Q_3 = 0$; since Q_4^2 is positive, this implies $Q_2 = 0$, or

$$x = (v^{z/(z-1)} - 1)/(v - v^{1/(z-1)}), \tag{33}$$

which equation is the counterpart of eq. (10) valid for ferromagnetic systems, v coming in the place of u. We can conclude that for $B = 0$ the antiferromagnetic solution is possible only for temperatures lower than a certain temperature T_c, the Curie, or Néel temperature, which is given by (9).

The equation for the transition curve in the B *vs* T plane separating the antiferromagnetic and paramagnetic phase, *i.e.* the curve along which $v = 1$ is a double solution of the basic equations, can be obtained by performing in eq. (32) the limit $v \to 1$. We get

$$y = \left(\frac{(z-2)\,x^2 - z \pm q}{2x}\right)\left(\frac{(z-2)\,x^2 + z \pm q}{2(z-1)\,x}\right)^{z-1}, \tag{34}$$

$$q \equiv [(x^2-1)\,((z-2)^2\,x^2 - z^2)]^{\frac{1}{2}}. \tag{35}$$

For a given value of B, eq. (34) determines the corresponding transition temperature, while for a given temperature it yields the value of the "transition field", *i.e.* that value of the external field for which the antiferromagnetic ordering vanishes. One can easily verify that for $B = 0$, *i.e.* $y = 1$, eq. (34) reduces to eq. (9). For $T = 0$, *i.e.* $x = 1$, on the other hand, we find

$$2\mu\,|B| = zJ, \tag{36}$$

indicating that the maximum value of $|B|$, for which the state with complete *antiferromagnetic* ordering ($S = 0$, $s = 1$) is the ground state, is equal to $zJ/2\mu$. For $|B| > zJ/2\mu$, the ordering effect of the external field predominates over that of the antiferromagnetic coupling so that the state with complete *ferromagnetic* ordering ($|S| = 1$, $s = 0$) is the ground state. The transition curve for the simple cubic lattice ($z = 6$) is shown in fig. 2. For values of B and T corresponding to points "outside" the curve, the system is in the paramagnetic phase, for points inside the curve both the antiferromagnetic and the paramagnetic phase can exist, the former one having the lower free energy, as be can shown easily.

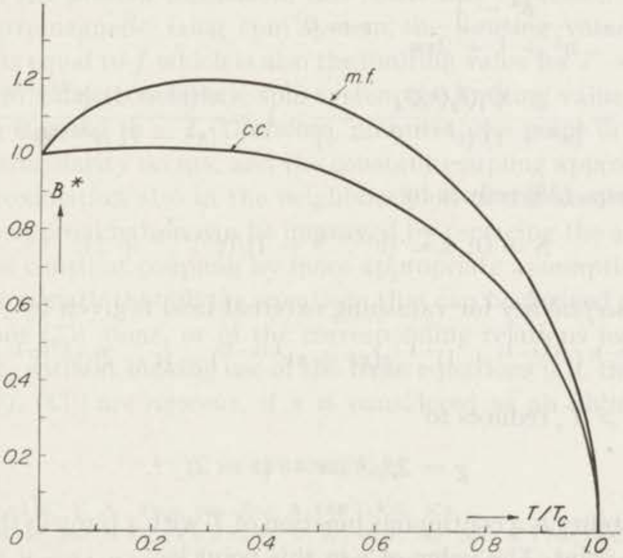

Fig. 2. The transition curve between the antiferromagnetic and the paramagnetic phase for the simple cubic lattice; $B^* = 2\mu\,|B|/zJ$.

§ 5. *Calculation of the properties of the spin system.* From the analysis in the preceding section it follows that both for the paramagnetic and for the antiferromagnetic phase we can express the thermodynamic quantities of the spin system in terms of x and one auxiliary quantity, and that the equilibrium is then described by one equation determining the temperature- and field dependence of this auxiliary quantity, just as in the ferromagnetic case. We first express the various quantities in terms of u, v and x with the aid of eq. (23). For the paramagnetic phase we have $v = 1$ so that all quantities are expressed in terms of u and x alone, where u is determined by (29). For the antiferromagnetic phase we can express u, with the aid of (30), in terms of v and x, so that the quantities can be expressed in terms of v and x, v being determined by (32).

a) *The total magnetization* of the lattice and the difference of the magnetizations of the A-lattice and the B-lattice are equal to $N\mu S$ and $N\mu s$ respectively. For the paramagnetic (p.m.) and the antiferromagnetic (a.f.m.) phase we find, respectively:

$$\text{p.m.: } S = \frac{u^2 - 1}{u^2 + 1 + 2xu}\ ;\ s = 0, \tag{37}$$

$$\text{a.f.m.: } S = \frac{Q_1 Q_2 Q_3 Q_4}{(x^2 - 1)\,(v^{2z/(z-1)} - 1)}\ ;\ \ s = \frac{x^2(v^2 - 1)\,(v^{2/(z-1)} - 1)}{(x^2 - 1)\,(v^{2z/(z-1)} - 1)}\,. \tag{38}$$

For $B = 0$, eqs. (38) reduce to

$$S = 0;\ s = (v^{z/(z-1)} - 1)/(v^{z/(z-1)} + 1). \tag{39}$$

b) *The susceptibility* for vanishing external field is given by

$$\chi = 8\beta\mu^2 v^{z/(z-1)}\,(v^{z/(z-1)} + 1)^{-1}\,[z(xv + xv^{1/(z-1)}) - (z-2)(v^{z/(z-1)} + 1)]^{-1}, \tag{40}$$

which for $T \geqslant T_c$ reduces to

$$\chi = 2\beta\mu^2\,[zx - (z - 2)]^{-1}. \tag{41}$$

The susceptibility is a continuous function of T with a jump in the derivative in the Curie point. The value of χ in this point is

$$\chi_c = \{[z(z - 2)/2(z - 1)]\ \ln\,[z/(z - 2)]\}\,\chi_0, \tag{42}$$

where $\chi_0 = \mu^2/zJ$ is the critical value of the susceptibility in the molecular field approximation [12]. For $z = 6$ and $z = 8$ we have $\chi_c = 0.973\ \chi_0$ and $\chi_c = 0.986\ \chi_0$ respectively.

c) For *the short-range order parameter* σ, given by (17), we find

$$\text{p.m.: } \sigma = (2xu - u^2 - 1)/(2xu + u^2 + 1), \tag{43}$$

$$\text{a.f.m.: } \sigma = [(x^2+1)(v^{2z/(z-1)} - 1) - 2x^2(v^2 - v^{2/(z-1)})]/[(x^2 - 1)\,(v^{2z/(z-1)} - 1)]. \tag{44}$$

For $B = 0$, these equations reduce to

$$\text{p.m.: } \sigma = (x - 1)/(x + 1), \tag{45}$$

$$\text{a.f.m.: } \sigma = [x(v^2 + 1) - 2v]/[x(v^2 + 1) + 2v]. \tag{46}$$

d) *The energy* E can be expressed in terms of σ and S with the aid of eq. (2). E is a continuous function of T, but the specific heat shows a jump in the transition point, which for $B = 0$ is equal to the jump in the specific heat for a ferromagnetic Ising spin system.

e) *The entropy at the Curie point* has, for $B = 0$, also the same value as in the ferromagnetic case, as can readily be verified.

§ 6. *Concluding remarks.* By its very nature the constant coupling approximation is a good approximation in the region of high temperatures and small external fields. For Ising spin systems, however, as contrasted with the systems with isotropic coupling discussed in I and II, it is also a good low-temperature approximation in virtue of the well-known relation between the high- and low-temperature behaviour of the partition function for an Ising system [13]). In the present formalism, this relation is expressed by the fact that for a ferromagnetic Ising spin system the limiting value of A_2 for $T \to 0$, $S = 1$ is equal to J which is also the limiting value for $T \to \infty$, $S = 0$, and that for an antiferromagnetic spin system the limiting value for $T \to 0$, $s = 1$, $S = 0$ is equal to $-J$. Therefore, no anti-Curie point or other low-temperature singularity occurs, and the constant coupling approximation is a useful approximation also in the neighbourhood of the absolute zero. In principle, the approximation can be improved by replacing the assumptions (4) and (18) of constant coupling by more appropriate assumptions. In this connection we remark that all the equations that can be derived on the basis of the relations (23) alone, or of the corresponding relations for the ferromagnetic case, without making use of the basic equations (*e.g.* the equations (11), (12), (37), (43)) are *rigorous*, if x is considered as an abbreviation of exp $(\beta|A_2|)$.

REFERENCES

1) Guggenheim, E. A., Proc. roy. Soc. **A 148** (1935) 304.

2) Fowler, R. H. and Guggenheim, E. A., Proc. roy. Soc. **A 174** (1940) 189.

3) Fowler, R. H. and Guggenheim, E. A., *Statistical Thermodynamics*, Cambridge University Press (1939), Chapters VI, VIII, XIII.

4) Bethe, H. A., Proc. roy. Soc. **A 150** (1935) 552.

5) Peierls, R., Proc. Camb. phil. Soc. **32** (1936) 471, 477.

6) Firgau, U., Ann. Phys. Lpz. **40** (1941) 295.

7) Ziman, J. M., Proc. phys. Soc. **64 A** (1951) 1108.

8) Van Vleck, J. H., Rev. mod. Phys. **17** (1945) 27.

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

10) Kirkwood, J. G., J. chem. Phys. **6** (1938) 70.

11) Oguchi, T., J. phys. Soc. Japan **6** (1951) 31.

12) Van Vleck, J. H., J. chem. Phys. **9** (1941) 85.

13) Domb, C., Proc. roy. Soc. **A 199** (1949) 199.
Wakefield, A. J., Proc. Camb. phil. Soc. **47** (1951) 419.

CHAPTER IV

THE STATISTICS OF AN ADSORBED MONOLAYER WITH REPULSIVE INTERACTION BETWEEN THE ADSORBED PARTICLES

§ 1. *Introduction.* The adsorption isotherms of a localized monolayer with negative interaction energy (attractive force) between neighbouring adsorbed particles were first investigated by Fowler[1]) and Peierls[2]), using, respectively, the approximation methods of Bragg and Williams[3]) and of Bethe[4]), which were originally introduced in the theory of superlattices in binary alloys. In this chapter we shall investigate the properties of an adsorbed layer with positive interaction energy by means of the method developed in the preceding chapters, hereafter to be referred to as I, II and III. In particular we shall develop the constant coupling approximation for this case, which, for a lattice with a two-sublattice structure, is completely equivalent to the Bethe-Peierls approximation. Earlier, the latter approximation was applied to the same case by Wang[5]). However, Wang discussed only the disordered phase, predicting that in monolayers of this kind no critical phenomena would occur. Although in later work[6]) [7]) he mentioned the possibility of a phase with sublattice ordering, he only discussed some symmetry properties of this phase without giving a detailed analysis. Such an analysis is given in this chapter. In § 2 the general method for the calculation of the grand partition function of the monolayer is discussed, and the constant coupling approximation is derived in § 3. With the aid of this approximation the adsorption isotherms are studied (§ 4), and the heat of adsorption is calculated (§ 5).

§ 2. *The grand partition function of an adsorbed monolayer with repulsion between neighbouring particles.* We consider an ideal gas of pressure p and temperature T in contact with an adsorbing surface. The following assumptions are made:

1) The surface has N sites on which a particle can be adsorbed. These sites

form a two-dimensional lattice which can be divided into two sublattices, A and B, in such a way that the z nearest neighbours of a site on the A-lattice are on the B-lattice, and vice versa.

2) The particles of the gas are adsorbed without being dissociated or associated; each adsorbed particle occupies one site of the lattice, and no site is occupied by more than one particle.

3) The adsorbed particles form a single layer ("monolayer").

4) There is a positive interaction energy (repulsive force) between particles adsorbed on neighbouring lattice sites, while particles that are not nearest neighbours do not interact.

The energy of such a monolayer is given by

$$V = V_r \Sigma_{<i,j>} \theta_i \theta_j - V_a \Sigma_i \theta_i, \tag{1}$$

where V_r is the interaction energy for a pair of neighbouring particles, and where V_a is the energy of adsorption per particle. The number θ_i ($i = 1, \ldots, N$) is equal to 1 or 0 accordingly as the i^{th} lattice site is occupied or empty, and the first sum runs over all pairs of nearest neighbours. The expressions $\Sigma\, \theta_i$ and $\Sigma\, \theta_i \theta_j$ are therefore equal to the number of adsorbed particles and the number of nearest neighbour pairs of adsorbed particles respectively. If the adsorbed phase is in thermodynamic equilibrium with the gas, the grand partition function of the monolayer is given by

$$Z_{gr} = \Sigma_{\theta_1 \ldots \theta_N} \exp\left[-\beta V(\theta_1, \ldots, \theta_N)\right] \exp\left[\beta \mu^{(g)} \Sigma_i \theta_i\right], \tag{2}$$

where the first sum runs over the 2^N sets of values $\theta_1, \ldots, \theta_N$; $\beta \equiv 1/kT$, and $\mu^{(g)}$ is the chemical potential of the gas:

$$\mu^{(g)} = kT \ln\left[(p/kT)\,(h^2/2\pi mkT)^{3/2}\right]. \tag{3}$$

The partition functions for the internal degrees of freedom in the gas phase and for the vibrations in the adsorbed state have been suppressed, assuming that the latter is independent of V_r and of the arrangement of the particles on the surface. In order to stress the analogy of the expression (2) to the expression for the partition function of an antiferromagnetic Ising spin system, we replace the variables θ_i by $s_i + \frac{1}{2}$, where s_i can assume the values $\pm\frac{1}{2}$. We then introduce two long-range order parameters, S and s, defined as

$$S = (2/N)\, \Sigma_i\, s_i; \quad s = (2/N)\, \Sigma_i\, \delta_i s_i, \tag{4}$$

where $\delta_i = +1$ for the sites of the A-lattice, and $\delta_i = -1$ for the sites of the B-lattice. These parameters are closely related to the fraction Θ of lattice sites occupied by adsorbed particles (relative covering) and the quantity θ defined as the difference of the fractions of occupied sites on the sublattices A and B. We have the following relations:

$$\Theta = (1/N)\, \Sigma_i\, \theta_i = \tfrac{1}{2}(1 + S); \quad \theta = (2/N)\, \Sigma_i\, \delta_i\, \theta_i = s. \tag{5}$$

The grand partition function (2) can now be written in the form

$$Z_{gr} = K\,\Sigma_{Ss}\,Z(S,s)\exp\left(\tfrac{1}{2}\beta N S \mu^{(g)}\right), \tag{6}$$

with

$$Z(S,s) = \Sigma''_{s_1 \ldots s_N} \exp\left[-\beta\left(V_r\,\Sigma_{<i,j>}\, s_i s_j - \tfrac{1}{2} N S\left(V_a - \tfrac{1}{2} z V_r\right)\right)\right], \tag{7}$$

where K is a constant factor, and where Σ'' is a sum over all the sets of values $s_1, \ldots, s_N$ satisfying the relations (4) for given values of S and s. The expression (7) is formally identical with the expression for the partition function at constant S and s of an antiferromagnetic Ising spin system with a coupling constant equal to $J = \tfrac{1}{2} V_r$ in an external field of magnitude $(V_a - \tfrac{1}{2} z V_r)/2\mu$. The statistical theory of an Ising spin system given in III can therefore be easily translated into a theory of the statistical properties of an adsorbed monolayer, and in particular we can apply the results of the constant coupling, or quasi-chemical approximation for the spin system to the adsorption case. This will be carried out in the next section.

§ 3. *The constant coupling approximation.* In order to introduce the constant coupling approximation for an adsorbed monolayer we express the partition function $Z(S,s)$ in terms of the effective energy V_e of a pair of neighbouring lattice sites. Applying the general theory developed in I we find

$$\ln Z(S,s) = -\textstyle\int_0^\beta E(S, s, \beta')\, \mathrm{d}\beta' + \ln g(S, s), \tag{8}$$

where $g(S, s)$ stands for the expression (II.13) and where the quantity $E(S, s, \beta)$, which is the average energy at given values of S and s, is given by

$$E = \tfrac{1}{2} N z V_r \frac{\Sigma_{\theta_1\theta_2}\,\theta_1\theta_2 \exp\left[-\beta V_e(\theta_1, \theta_2)\right]}{\Sigma_{\theta_1\theta_2} \exp\left[-\beta V_e(\theta_1, \theta_2)\right]} - \tfrac{1}{2} N V_a S, \tag{9}$$

in which expression the constant term has been omitted. The effective energy $V_e(\theta_1, \theta_2)$ is of the form

$$V_e(\theta_1, \theta_2) = C_1\theta_1\theta_2 - C_2'\,\theta_1 - C_2''\,\theta_2, \tag{10}$$

where the coefficients C_i are functions of S, s and β. C_1 is an effective repulsion energy, and C_2' and C_2'' are effective energies of adsorption for a site on the A- and the B-lattice respectively. Just as in the case of an antiferromagnetic spin system, one can easily find two relations between the coefficients in the effective pair energy and S, s and β so that the partition function can be calculated if one additional relation is given, for instance, if the effective repulsion energy C_1 is given as a function of S, s and β. The constant coupling approximation which, according to III, §§ 2, 3, is equivalent to the quasi-chemical, or Bethe-Peierls approximation is obtained by assuming that C_1 is equal to the actual repulsion energy V_r between two neighbouring particles in the monolayer,

$$C_1 = V_r. \tag{11}$$

Following the argument that was used in III we can then derive from eqs. (6) and (8)–(11) the equilibrium values of the order parameters S and s and of the thermodynamic quantities for given values of the pressure p and the temperature T. To this end we express these quantities in terms of two auxiliary quantities,

$$u \equiv \exp \tfrac{1}{2}\beta(C_2' + C_2''); \quad v \equiv \exp \tfrac{1}{2}\beta(C_2' - C_2''). \tag{12}$$

The equilibrium values of u and v are then determined by the equations (III.25) and (III. 26):

$$y = uv[(v + xu)/(u + xv)]^{z-1}, \tag{13}$$

$$y = (u/v)\,[(1 + xuv)/(uv + x)]^{z-1}, \tag{14}$$

where the quantities x and y which for spin systems are equal to exp (βJ) and exp $(2\beta\mu B)$ respectively, are now defined as

$$x \equiv \exp (\tfrac{1}{2}\beta V_r), \tag{15}$$

$$y \equiv \exp \beta(V_a - \tfrac{1}{2}zV_r + \mu^{(g)}) = p(h^2/2\pi m)^{3/2}\,(kT)^{-5/2} \exp \beta(V_a - \tfrac{1}{2}zV_r). \tag{16}$$

The appearance of the temperature dependent term $\mu^{(g)}$ in (16), which is not suggested by the correspondence of $2\mu B$ and $V_a - \frac{1}{2}zV_r$, finds its origin in the occurrence of the factor exp $(\frac{1}{2}\beta NS\mu^{(g)})$ in (6), *i.e.* in the fact that the expression (2), or (6), represents a *grand* partition function. It does not influence the analysis in an essential way, but it is important for the interpretation of the results , because it is through this term that the pressure p enters into the equations. The equations (13) and (14) have two solutions. For the first one, u and v are given by

$$y = u[(1 + xu)/(u + x)]^{z-1}, \tag{17}$$

$$v = 1. \tag{18}$$

This is the solution discussed by Wang [5]). It corresponds to the disordered phase of the monolayer (see below). For the other solution, u can be expressed in terms of v and x,

$$u = (Q_1Q_3 \pm Q_2Q_4)/(Q_1Q_3 \mp Q_2Q_4), \tag{19}$$

where the following abbreviations have been used:

$$\left.\begin{aligned} Q_1 &\equiv [(xv - 1) - v^{1/(z-1)}\,(x - v)]^{\frac{1}{2}},\\ Q_2 &\equiv [(xv + 1) - v^{1/(z-1)}\,(x + v)]^{\frac{1}{2}},\\ Q_3 &\equiv [(xv - 1) + v^{1/(z-1)}\,(x - v)]^{\frac{1}{2}},\\ Q_4 &\equiv [(xv + 1) + v^{1/(z-1)}\,(x + v)]^{\frac{1}{2}}. \end{aligned}\right\} \tag{20}$$

The value of v is then determined by

$$y = \left(\frac{Q_1Q_3 \pm Q_2Q_4}{Q_1Q_3 \mp Q_2Q_4}\right)\left(\frac{Q_1Q_4 \pm Q_2Q_3}{Q_1Q_4 \mp Q_2Q_3}\right)^{z-1}. \tag{21}$$

This solution corresponds to a phase with sublattice ordering.

Whereas the first solution exists for all values of x and y, *i.e.* of T and p, the second solution is possible only if the temperature T is lower than a certain critical temperature T_c, given by

$$kT_c = J/\ln[z/(z-2)], \tag{22}$$

and if the pressure p lies between two limits, p_- and p_+,

$$p_- < p < p_+, \tag{23}$$

which depend on T; the first solution is unstable in that case. For $p = p_\pm$ the two solutions become identical, $v = 1$ being a double solution of (13) and (14); hence we can obtain the relation between $p_\pm$, or the corresponding reduced pressures $p^*_\pm = p_\pm (h^2/2\pi m)^{3/2} V_r^{-5/2}$, and T by performing in eq. (21) the limit $v \to 1$. We get

$$p^*_\pm = \left(\frac{kT}{V_r}\right)^{5/2} \exp[(\tfrac{1}{2}zV_r - V_a)/kT] \left(\frac{(z-2)x^2 - z \pm q}{2x}\right)\left(\frac{(z-2)x^2 + z \pm q}{2(z-1)\,x}\right)^{z-1}, \tag{24}$$

where

$$q \equiv [(x^2-1)\,((z-2)^2\,x^2 - z^2)]^{\frac{1}{2}}. \tag{25}$$

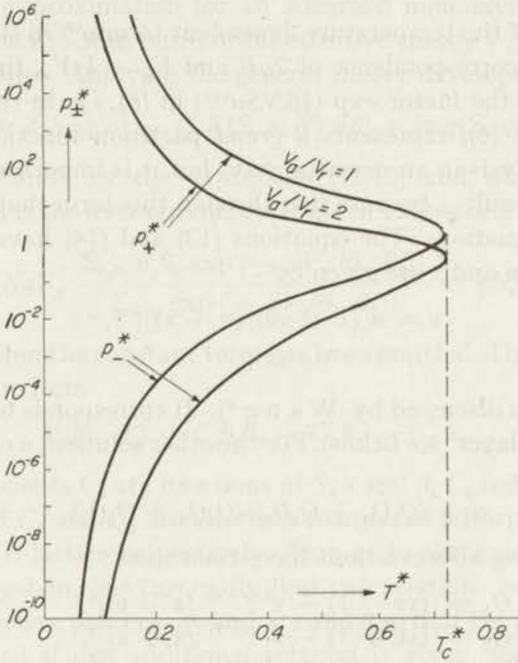

Fig. 1. The reduced transition pressures p^*_- and p^*_+ as functions of the reduced temperature $T^* = 4kT/zV_r$ for the quadratic lattice ($z = 4$) and for $V_a/V_r = 1$ and $V_a/V_r = 2$.

The "transition pressures" p_- and p_+ satisfy the relation

$$p_- p_+ = p_0^2(T), \tag{26}$$

where $p_0(T)$ is that value of the pressure for which y is equal to 1,

$$p_0(T) = (2\pi m/h^2)^{3/2}\,(kT)^{5/2} \exp[(\tfrac{1}{2}zV_r - V_a)/kT], \tag{27}$$

and which therefore corresponds to the value 0 of the external field in the case of a spin system. The curve in the p *vs* T plane represented by (24), for a quadratic lattice ($z = 4$) and for two values of V_a/V_r, is shown in fig. 1.

§ 4. *The adsorption isotherm.* The adsorption isotherm of the monolayer can be calculated with the help of eqs. (13) and (14). Making use of the relations (5) and (III.21)–(III.23), we first express Θ and θ in terms of u and v:

$$\Theta = \tfrac{1}{2}(xv + x/v + 2u)/(xv + x/v + u + 1/u), \tag{28}$$

$$\theta = (xv - x/v)/(xv + x/v + u + 1/u). \tag{29}$$

Solving u and v from these equations and inserting the resulting expressions into (13) and (14) we obtain the following equations

$$p = p_0(T)\left(\frac{\Theta + \frac{1}{2}\theta}{1 - \Theta - \frac{1}{2}\theta}\right)\left(\frac{2x^2\Theta + \theta - x^2 + [x^4 - 4x^2(x^2-1)\Theta(1-\Theta) - (x^2-1)\theta^2]^{\frac{1}{2}}}{2x(\Theta + \frac{1}{2}\theta)}\right)^z, \tag{30}$$

$$p = p_0(T)\left(\frac{\Theta - \frac{1}{2}\theta}{1 - \Theta + \frac{1}{2}\theta}\right)\left(\frac{2x^2\Theta - \theta - x^2 + [x^4 - 4x^2(x^2-1)\Theta(1-\Theta) - (x^2-1)\theta^2]^{\frac{1}{2}}}{2x(\Theta - \frac{1}{2}\theta)}\right)^z, \tag{31}$$

from which the relation between Θ and p can be derived by the elimination of θ. (It should be remarked that it is necessary to take the positive root in (30) and (31), because otherwise p would be negative.)

For temperatures above the critical point and, in the case $T < T_c$, for values of the pressure outside the range (p_-, p_+) we have, according to (18) and (29), $\theta = 0$ so that the equations (30) and (31) reduce to

$$p = p_0(T)\left(\frac{\Theta}{1 - \Theta}\right)\left(\frac{x(2\Theta - 1) + [x^2 - 4(x^2 - 1)\Theta\,(1 - \Theta)]^{\frac{1}{2}}}{2\Theta}\right)^z. \tag{32}$$

For $T < T_c$ and $p_- < p < p_+$, on the other hand, the adsorption isotherm is represented by eq. (30) or (31), where θ is a function of Θ, which in principle can be determined from (30) and (31). Making use of the equations (III.38) we can find a parametric representation of the relation between θ and Θ:

$$\theta = x^2(v^2 - 1)\,(v^{2/(z-1)} - 1)/(x^2 - 1)\,(v^{2z/(z-1)} - 1), \tag{33}$$

$$\Theta = \tfrac{1}{2}[1 \pm Q_1Q_2Q_3Q_4/(x^2 - 1)\,(v^{2z/(z-1)} - 1)]. \tag{34}$$

Fig. 2 shows θ as a function of Θ for $z = 4$ and for various values of the temperature. Since the value of θ is not changed by the replacement of Θ by $1 - \Theta$, the curves are symmetric with respect to the line $\Theta = \frac{1}{2}$. For $p \leqslant p_-$, $p \geqslant p_+$ we have $v = 1$, $\theta = 0$ so that, according to (30)–(32), Θ is continuous in the transition points, its value being equal to

$$\Theta = \Theta_\pm = \tfrac{1}{2}[1 \pm \{(z - 2)^2\,x^2 - z^2\}^{\frac{1}{2}}/\{z(x^2 - 1)^{\frac{1}{2}}\}], \tag{35}$$

as can be seen by performing in (34) the limit $v \to 1$. In the limit $T \to 0$ we get $\Theta_- = 1/z$, $\Theta_+ = (z - 1)/z$.

Concluding we may say that, in contradistinction to the adsorption isotherms for temperatures above the critical point, the isotherms for $T < T_c$ consist of three parts which cannot be described by one formula. In the limit of vanishing pressure, we have $\Theta = \theta = 0$, *i.e.* the surface is empty. If we then increase the pressure, keeping the temperature constant ($T < T_c$), the fraction Θ of the surface covered with adsorbed particles increases continuously according to (32), no preference being given to sites of the A- or the B-lattice ($\theta = 0$), until the pressure $p = p_-(T)$ is reached. By then the number of adsorbed particles has been so far increased that, under the influence of the mutual repulsion between the particles, the system goes over into a different phase in which more particles are adsorbed on one sublattice than on the other. A vanishingly small difference between the adsorption energies for the A- and the B-lattice is sufficient to determine which one of the two lattices has the greater occupation, *i.e.* to determine the sign of θ. We suppose θ to be positive. The change of the relative covering Θ of the surface and of the sublattice ordering θ with increasing p is now given by (30), (33) and (34). It can be shown that both Θ and θ increase with p until the pressure $p = p_0(T)$ is reached, *i.e.* until the surface is halfway filled up. As can easily be seen from fig. 2, θ increases so rapidly with Θ that the fraction of occupied sites in the B-lattice, which is equal to $\Theta - \frac{1}{2}\theta$, decreases abruptly with increasing Θ, and hence with increasing p, goes through a minimum, and is still very small for $p = p_0$ (unless $T \approx T_c$), practically all the adsorbed

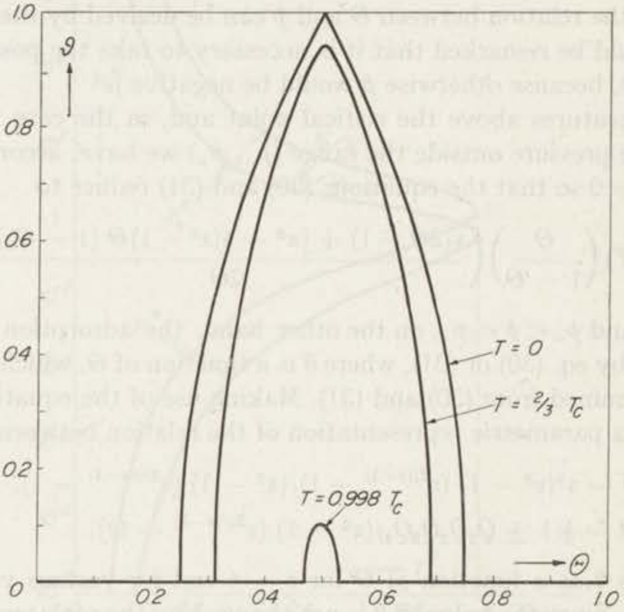

Fig. 2. The sublattice ordering θ as a function of the relative covering Θ for $z = 4$ and for various values of the temperature.

particles being located in the other sublattice. For $p > p_0$ the behaviour of the system is, in a sense, the symmetric image of the behaviour for $p < p_0$, as can be seen from (30) and (31). At a pressure $p = p'$ the numbers of empty sites in the A- and B-lattice are equal to the numbers of occupied sites in the B- and A-lattice respectively at the pressure $p = p_0^2/p'$. As a consequence, Θ is a monotonically increasing function of p, whereas θ decreases for $p > p_0$.

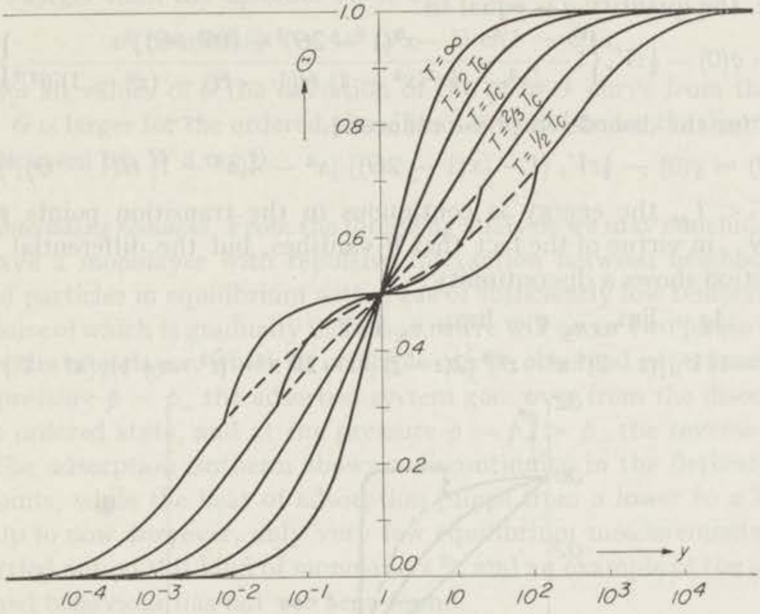

Fig. 3. The relative covering Θ as a function of $y = p/p_0(T)$ for the quadratic lattice ($z = 4$) and various values of the temperature.

At the pressure $p = p_+(T)$, θ has become equal to zero, and the system goes over into the initial phase again, in which there is no preferred sublattice. Thus it appears that, for temperatures below the critical temperature, the adsorption isotherms show *two* transition points. In these points, Θ and θ are continuous functions of p with a jump in the derivatives. For $T \geqslant T_c$, however, no transition occurs. Fig. 3 shows a number of isotherms for the quadratic lattice ($z = 4$); Θ has been plotted as a function of $y = p/p_0(T)$ for various values of the temperature.

§ 5. *The heat of adsorption.* In this section we shall calculate the differential heat of adsorption defined as the decrease of the total energy of the gas and the monolayer when one particle goes over from the gas phase into the adsorbed phase:

$$q = e^{(g)} - \partial E/\partial(N\Theta). \tag{36}$$

In this equation $e^{(g)}$ is the energy of a single gas particle,

$$e^{(g)} = \tfrac{3}{2}kT,$$

while E is the average energy of the monolayer given by (9). One can show, making use of eqs. (III.2), (III.17) and (III.19)–(III. 22), that, apart from a constant term, E is equal to

$$E = \tfrac{1}{2}NzV_r\{\Theta - (x^2 - [x^4 - 4x^2(x^2-1)\,\Theta(1-\Theta) - (x^2-1)\,\theta^2]^{\frac{1}{2}})/2(x^2-1)\} - NV_a\Theta. \quad (37)$$

Hence, the quantity q is equal to

$$q(\Theta) = q(0) - \tfrac{1}{2}zV_r\left\{1 - \frac{x^2(1-2\Theta) + (\partial\theta^2/\partial\Theta)_x}{[x^4 - 4x^2(x^2-1)\,\Theta(1-\Theta) - (x^2-1)\,\theta^2]^{\frac{1}{2}}}\right\}, \quad (38)$$

which for the disordered phase reduces to

$$q(\Theta) = q(0) - \tfrac{1}{2}zV_r\{1 - [x(1-2\Theta)]/[x^2 - 4(x^2-1)\,\Theta(1-\Theta)]^{\frac{1}{2}}\}. \quad (39)$$

For $T < T_c$, the energy is continuous in the transition points $p = p_{\pm}$, $\Theta = \Theta_{\pm}$, in virtue of the fact that θ vanishes, but the differential heat of adsorption shows a discontinuity:

$$\Delta q = \lim_{\Theta\downarrow\Theta_{\pm}} q - \lim_{\Theta\uparrow\Theta_{\pm}} q =$$

$$= 3z(z-1)\,V_r\,[(z-2)^2\,x^2 - z^2]^{\frac{1}{2}}\,\{2(z-2)\,[(z-2)^2\,x^2 - (z^2-z+1)]\,[x^2-1]\}^{-1}. \quad (40)$$

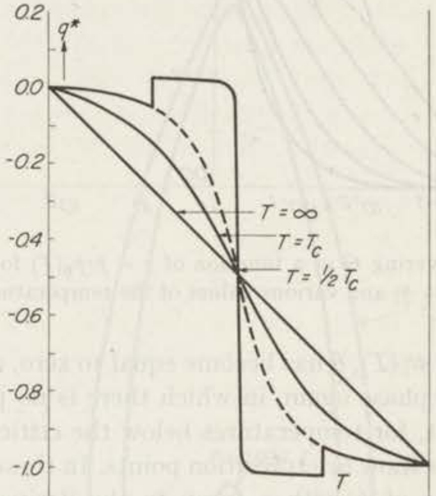

Fig. 4. The reduced differential heat of adsorption q^* as a function of the relative covering Θ for $z = 4$ and various values of the temperature.

For temperatures not to near the critical point, Δq is so large that, for values of Θ immediately above the value Θ_-, the heat of adsorption is even larger than for an empty surface. This is because a small increase of Θ causes a large increase of the sublattice ordering θ, and thus considerably lowers the energy. Fig. 4 shows $q^*(\Theta) = [q(\Theta) - q(0)]/zV_r$ as a function of Θ for

$z = 4$ and various values of the temperature; for $T = \infty$ we get $q^* = -\Theta$. Since θ is an even function of $\Theta - \frac{1}{2}$, $\partial\theta^2/\partial\Theta$ is an odd function of $\Theta - \frac{1}{2}$ and therefore the q^* *vs* Θ curves are symmetric with respect to the point $q^* = -\frac{1}{2}$, $\Theta = \frac{1}{2}$. Since $\partial\theta^2/\partial\Theta$ has the same sign as $1 - 2\Theta$, as can easily be proved, the absolute value of

$$[x^2(1 - 2\Theta) + \partial\theta^2/\partial\Theta]/[x^4 - 4x^2\,(x^2 - 1)\,\Theta(1 - \Theta) - (x^2 - 1)\,\theta^2]^{\frac{1}{2}}$$

is always larger than the absolute value of

$$x^2(1 - 2\Theta)/[x^4 - 4x^2\,(x^2 - 1)\,\Theta(1 - \Theta)],$$

so that for all values of Θ the deviation of the q^* *vs* Θ curve from the line $q^* = -\Theta$ is larger for the ordered phase than it would be for the disordered phase discussed by Wang[5]).

§ 6. *Concluding remarks.* From the foregoing analysis we may conclude that if we have a monolayer with repulsive interaction between neighbouring adsorbed particles in equilibrium with a gas of sufficiently low temperature, the pressure of which is gradually increased, there will occur two phase transitions in the monolayer, which in principle can be observed experimentally. At the pressure $p = p_-$ the adsorbed system goes over from the disordered into the ordered state, and at the pressure $p = p_+ > p_-$ the reverse takes place. The adsorption isotherm shows a discontinuity in the derivative at these points, while the heat of adsorption jumps from a lower to a higher value. Up to now, however, only very few equilibrium measurements have been carried out on this kind of monolayers[8]), and an example of the above-mentioned behaviour has not yet been found.

Finally, we remark that, if in the equations of the preceding sections we perform the limit $z \to \infty$, keeping zV_r constant, we obtain the corresponding equations of the Bragg-Williams-Fowler approximation for the repulsion case. This approximation can also be obtained directly by replacing the assumption (11) of constant coupling by the assumption that there is no effective interaction, *i.e.* that C_1 is equal to zero.

REFERENCES

1) Fowler, R. H., Proc. Camb. phil. Soc. **32** (1936) 144.
2) Peierls, R., Proc. Camb. phil. Soc. **32** (1936) 471.
3) Bragg, W. L. and Williams, E. J., Proc. roy. Soc. **A 145** (1934) 699; **A 151** (1935) 540.
4) Bethe, H. A., Proc. roy. Soc. **A 150** (1935) 552.
5) Wang, J. S., Proc. roy. Soc. **A 161** (1937) 127.
6) Wang, J. S., Proc. roy. Soc. **A 168** (1938) 56.
7) Wang, J. S., Phys. Rev. **67** (1945) 98.
8) Schuit, G. C. A. and De Boer, N. H., Rec. Trav. chim. **72** (1953) 909.

CHAPTER V

THE GROUND STATE OF AN ANTIFERROMAGNETIC LINEAR CHAIN

§ 1. *Introduction.* The ground state of a linear chain of spins $\frac{1}{2}$ with an isotropic, antiferromagnetic nearest neighbour coupling has been the subject of many discussions. In the molecular field theory [1]) the state of lowest energy shows a complete antiferromagnetic order, the lattice being divided into two sublattices with opposite magnetizations. However, Hulthén [2]) has shown in an approximate treatment that for the ground state the probability of finding a pair of neighbouring spins of opposite sign is much smaller than the value 1 predicted by the molecular field theory. The rigorous calculation of the lowest energy by Bethe [3]) and Hulthén [2]) also shows that the exact ground state differs considerably from the completely ordered state. Finally, it follows from the spin wave theory [4]) that at the absolute zero the antiferromagnetic *long-range order* vanishes, the two sublattices both having zero magnetization.

On the other hand, it is well-known that for a lattice with antiferromagnetic Ising interaction the completely ordered state is the ground state. However, the Ising interaction is anisotropic, and it is therefore interesting to examine how large the anisotropy must be in order that the spin system will have a ground state with non-vanishing antiferromagnetic long-range order. This investigation will be carried out in the following sections, by means of a variational method which is a generalization of a procedure introduced by Slater [5]) and refined by Hulthén [2]).

§ 2. *The variational problem.* Consider a linear lattice of N spins $\frac{1}{2}$ with the following Hamiltonian:

$$H = \Sigma_{<i,j>} H_{ij}^{(2)}, \qquad (1)$$

$$H_{ij}^{(2)} = 2J\,[(1-\alpha)\,(S_{ix}S_{jx} + S_{iy}S_{jy}) + S_{iz}S_{jz}], \qquad (2)$$

where J is positive, and where the anisotropy constant α lies between 0 and 1.

For $\alpha = 0$ the interaction is isotropic, for $\alpha \neq 0$ the coupling has uniaxial anisotropy, the most extreme case of which is that of Ising interaction ($\alpha = 1$).

The eigenfunctions of the operator (1) can be expanded in terms of the spin product functions $|s_i\rangle \equiv |s_1 \ldots s_N\rangle$. Since the z-component of the total spin, $\Sigma_i S_{iz}$, commutes with H, we can restrict our considerations to linear combinations of spin product functions with the same value of $\Sigma_i s_i$. The problem of determining the coefficients in this expansion is extremely difficult, but we can construct an approximate solution for the eigenfunction of the lowest energy state with a given value of $\Sigma_i s_i$ by means of a variational method. To this end we assume the lattice to be divided into two sublattices, and we classify the spin product functions by three parameters

$$K = 2\,\Sigma_i\, s_i, \tag{3}$$

$$k = 2\,\Sigma_i\, \delta_i\, s_i, \tag{4}$$

$$\varkappa = -4\,\Sigma_{<i,j>}\, s_i\, s_j, \tag{5}$$

which are proportional to the long-range order parameters S and s, and the short-range order parameter σ, introduced in the preceding chapters:

$$K = NS;\; k = Ns;\; \varkappa = N\sigma. \tag{6}$$

We then construct a trial function by taking a linear combination of spin product functions, in which functions with the same values of k and $\varkappa$ have the *same* coefficient, and we determine the best values of these coefficients by means of the variational principle.

Let the number of spin product functions $|s_i\rangle$ with given K, k and $\varkappa$ be $g(K, k, \varkappa)$. Introducing the normalized functions

$$|K\, k\, \varkappa\rangle = g^{-\frac{1}{2}}(K, k, \varkappa)\, \Sigma_{s_i}^{(K,k,\varkappa)}\, |s_i\rangle, \tag{7}$$

where the superscript $(K, k, \varkappa)$ indicates the restrictive conditions (3) — (5), we can write the trial function in the form

$$|K\rangle = \Sigma_{k,\varkappa}\, a(K, k, \varkappa)\, |K\, k\, \varkappa\rangle. \tag{8}$$

We have to minimize $\langle K|H|K\rangle$ with respect to variations of the $a(K, k, \varkappa)$ under the condition $\langle K|K\rangle = 1$. To this end we first examine how the Hamiltonian (1) operates on a spin product function $|s_i\rangle$ with given values of K, k and $\varkappa$. In the representation of the spin product functions the non-vanishing matrix elements of $H^{(2)}$ are given by

$$\begin{aligned} &\langle + + |H^{(2)}| + + \rangle = \langle - - |H^{(2)}| - - \rangle = \\ &\qquad = -\langle + - |H^{(2)}| + - \rangle = -\langle - + |H^{(2)}| - + \rangle = \tfrac{1}{2}J, \\ &\qquad \langle + - |H^{(2)}| - + \rangle = \langle - + |H^{(2)}| + - \rangle = (1 - \alpha)\, J, \end{aligned} \tag{9}$$

where $|++\rangle$ *etc.* are abbreviations for $|+\frac{1}{2}, +\frac{1}{2}\rangle$ *etc.* In the state described by the spin function in question there are $\frac{1}{2}(N - \varkappa)$ pairs of equal nearest neighbour spins ($++$ and $--$) and $\frac{1}{2}(N+\varkappa)$ pairs of opposite spins ($+-$ and $-+$). We conclude that $H|s_i\rangle$ is a linear combination of spin product functions, which, in addition to the original function $|s_i\rangle$ appearing with a coefficient $\frac{1}{2}(N-\varkappa)\,(\frac{1}{2}J)+\frac{1}{2}(N+\varkappa)\,(-\frac{1}{2}J)=-\frac{1}{2}\varkappa J$, contains $\frac{1}{2}(N+\varkappa)$ spin product functions differing from $|s_i\rangle$ by the interchange of two nearest neighbour spins with opposite signs (and hence, in general, by the values of k and $\varkappa$), and each one appearing with a coefficient $(1-\alpha)\,J$. The effect of such an interchange on k and $\varkappa$ can be investigated by considering the pair in question together with its two nearest neighbours, *i.e.* a configuration of four consecutive spins. Such a configuration is characterized by the values of the spin variables s_i of the four spins and by their position with respect to the two sublattices, which can be specified by giving the value of one of the δ_i; the order of the spins, however, is irrelevant. We can distinguish eight different configurations with opposite inner spins, which we shall denote by the signs of the spin variables, omitting the $\frac{1}{2}$'s and using the convention that the most left-hand symbol represents a spin on the A-lattice:

$$\begin{array}{l|l}
++--\quad k,\varkappa\rightarrow k+4,\ \varkappa+4 & --++\quad k,\varkappa\rightarrow k-4,\ \varkappa+4 \\
++-+\quad k,\varkappa\rightarrow k+4,\ \varkappa & --+-\quad k,\varkappa\rightarrow k-4,\ \varkappa \\
-+--\quad k,\varkappa\rightarrow k+4,\ \varkappa & +-++\quad k,\varkappa\rightarrow k-4,\ \varkappa \\
-+-+\quad k,\varkappa\rightarrow k+4,\ \varkappa-4 & +-+-\quad k,\varkappa\rightarrow k-4,\ \varkappa-4
\end{array}$$

It is evident that the interchange of the inner pair of these configurations changes k and $\varkappa$ in the way stated above, whereas K is left unaltered. Let $f_{\delta\Delta}(s_i)$ be the fraction of configurations responsible for a transition $k,\varkappa\rightarrow k+\delta,\varkappa+\Delta$, in the state $|s_i\rangle$ ($\delta=\pm 4$; $\Delta=0,\pm 4$). Then

$$f_{\delta\Delta}\,(K,k,\varkappa)=g^{-1}\,(K,k,\varkappa)\,\Sigma_{s_i}^{(K,k,\varkappa)}\,f_{\delta\Delta}\,(s_i) \tag{10}$$

is the fraction in which these configurations are present in the whole set of states with given values of K, k and $\varkappa$; their number, and hence the number of possible transitions $k,\varkappa\rightarrow k+\delta,\varkappa+\Delta$, is $Ng(K,k,\varkappa)\,f_{\delta\Delta}\,(K,k,\varkappa)$. After having performed a certain transition by exchanging the inner spins of one of these configurations, we can realize the reverse transition with $k+\delta$, $\varkappa+\Delta\rightarrow k,\varkappa$ by repeating the interchange just once; since *each* transition can be made in both directions in this way we must have

$$Ng(K,k,\varkappa)\,f_{\delta\Delta}\,(K,k,\varkappa)=Ng(K,k+\delta,\varkappa+\Delta)\,f_{\bar{\delta}\bar{\Delta}}\,(K,k+\delta,\varkappa+\Delta), \tag{11}$$

where $\bar{\delta}$ and $\bar{\Delta}$ in the index stand for $-\delta$ and $-\Delta$.

If we let H operate on $|K\,k\,\varkappa\rangle$ we get in the first place a term $-\frac{1}{2}\varkappa J\cdot\cdot|K\,k\,\varkappa\rangle$ having its origin in the contributions $-\frac{1}{2}\varkappa J|s_i\rangle$ of the various functions $H|s_i\rangle$. The coefficients of the other, modified functions can be found by counting the marks left behind by the operations which formed these functions from the original functions $|s_i\rangle$. A function $|s_i'\rangle$ with parameter values K, $k+4$, $\varkappa+4$, for instance, can be obtained from a function $|s_i\rangle$ with values K, k, $\varkappa$ only by the interchange of the inner pair of a configuration $++--$, giving $+-+-$. Since any configuration $+-+-$ in the new state can be considered as the mark of such an interchange and since the number of these configurations is equal to $Nf_{\bar{4}\bar{4}}(s_i')$, the function $|s_i'\rangle$ can be formed in $Nf_{\bar{4}\bar{4}}(s_i')$ ways. Its coefficient in the expression for $H|K\,k\,\varkappa\rangle$ is equal to the product of this number, the normalizing factor $g^{-\frac{1}{2}}$ and the constant $(1-\alpha)\,J$. In this way we find

$$H|K\,k\,\varkappa\rangle =$$
$$= -\tfrac{1}{2}\varkappa J|K\,k\,\varkappa\rangle + (1-\alpha)\,NJg^{-\frac{1}{2}}(K,k,\varkappa)\,\Sigma_{\delta,\Delta}\,\Sigma_{s_i}^{(K,k-\delta,\varkappa-\Delta)}\,f_{\delta\Delta}(s_i)|s_i\rangle. \quad (12)$$

Making use of the fact that the spin product functions, and hence also the functions $|K\,k\,\varkappa\rangle$, form an orthonormal set, and of the equations (10) and (11), we finally get for the expectation value of the energy in the state $|K\rangle$:

$$\langle K|H|K\rangle = -\tfrac{1}{2}J\,\Sigma_{k,\varkappa}\,\varkappa|a(K,k,\varkappa)|^2 +$$
$$+ (1-\alpha)\,NJ\,\Sigma_{k,\varkappa}\,a^*(K,k,\varkappa)\,\Sigma_{\delta,\Delta}\,F_{\delta\Delta}(K,k,\varkappa)\,a(K,k+\delta,\varkappa+\Delta), \quad (13)$$

where the "transition coefficients" $F_{\delta\Delta}$ are given by

$$F_{\delta\Delta}(K,k,\varkappa) = [f_{\delta\Delta}(K,k,\varkappa)\,f_{\bar{\delta}\bar{\Delta}}(K,k+\delta,\varkappa+\Delta)]^{\frac{1}{2}} =$$
$$= [g(K,k,\varkappa)/g(K,k+\delta,\varkappa+\Delta)]^{\frac{1}{2}}\,f_{\delta\Delta}(K,k,\varkappa). \quad (14)$$

By variation of (13) with respect to the $a(K,k,\varkappa)$ under the condition $\langle K|K\rangle = 1$ we get

$$-\tfrac{1}{2}J\varkappa a(K,k,\varkappa) + (1-\alpha)\,NJ\,\Sigma_{\delta,\Delta}\,F_{\delta\Delta}(K,k,\varkappa)\,a(K,k+\delta,\varkappa+\Delta) =$$
$$= Ea(K,k,\varkappa). \quad (15)$$

The lowest eigenvalue $E = E_0^{(K)}$ of this set of equations is an approximation to the lowest eigenvalue of H for the given value of K. Apart from an undetermined common factor the corresponding set of coefficients $a(K,k,\varkappa)$ is of the form

$$a(K,k,\varkappa) = (-1)^{\frac{1}{2}k}\,b(K,k,\varkappa), \quad (16)$$

where the $b(K,k,\varkappa)$ are real and positive, because otherwise we could obtain a solution with a lower eigenvalue than $E_0^{(K)}$ by replacing the $b(K,k,\varkappa)$ by $|b(K,k,\varkappa)|$ 6). This can be seen by substituting (16) into (15), dividing the left- and right-hand member by $b(K,k,\varkappa)$ and observing that the second term in the left-hand member gets a minus sign, in virtue of the fact that the allowed values of δ are ± 4, and that therefore $(-1)^{\frac{1}{2}(k+\delta)}/(-1)^{\frac{1}{2}k}$ is equal to -1.

In the limit $N\to\infty$ the variational problem for $E_0^{(K)}$ becomes of a particularly simple form. To show this we put

$$b(K,k,\varkappa) = \exp N\omega(S,s,\sigma,N), \quad (17)$$

where S, s and σ are given by (6); ω is real, because the $b(K, k, \varkappa)$ are positive. Substituting (16) and (17) into (15) and expanding into powers of $1/N$, we get, neglecting terms of order $1/N$:

$$E_0^{(S)}/N = -\tfrac{1}{2}J\sigma - (1-\alpha)\, J\, \Sigma_{\delta,\Delta}\, F_{\delta\Delta}(S, s, \sigma) \exp\left[2\delta(\partial\omega/\partial s) + 2\Delta\,(\partial\omega/\partial\sigma)\right], \quad (18)$$

in which the variables K, k and $\varkappa$ have been replaced by S, s and σ. Since N has disappeared from the equation we may conclude that in the limit $N \to \infty$, ω is independent of N:

$$b(K, k, \varkappa) = \exp N\omega(S, s, \sigma). \quad (19)$$

Now one can easily derive from (15) that $b(K, k, \varkappa)$ (the square of which is the probability, in the state $|K\rangle$, of finding the system with the parameter values k and $\varkappa$) must have at least one maximum somewhere inside the $(k, \varkappa)$-domain [6]). Therefore, ω has at least one maximum, say for $s = s^*$, $\sigma = \sigma^*$; in this maximum we have $\partial\omega/\partial s = 0$ and $\partial\omega/\partial\sigma = 0$, and hence

$$E_0^{(S)}/N = -\tfrac{1}{2}J\sigma^* - (1-\alpha)\, J\, \Sigma_{\delta,\Delta}\, F_{\delta\Delta}\,(S, s^*, \sigma^*). \quad (20)$$

If we minimize $E_0^{(S)}$ with respect to variations in the values of s^* and σ^* we find an approximation for the lowest energy level corresponding to the given value of the long-range order parameter S.

§ 3. *The calculation of the transition coefficients.* For the calculation of the transition coefficients $F_{\delta\Delta}$ we need the probabilities $f_{\delta\Delta}(S, s, \sigma)$ of finding certain configurations of four consecutive spins in an arbitrary state $|s_i\rangle$ with given values of S, s and σ (*i.e.* of K, k and $\varkappa$). These probabilities can be obtained in the following way: the chance of having a configuration $s_1s_2s_3s_4$ is the product of the chance of having an inner pair s_3s_2 (if for a pair we use the same convention as for a set of four spins, *viz.* that the left-hand symbol represents always a spin on the A-lattice), the chance of having a left-hand pair s_1s_2 when it is given that the second spin is s_2, and the chance of having a right-hand pair s_3s_4 when the third spin is s_3. The probability f_{44} of finding a configuration $+\,+\,-\,-$, for instance, is given by

$$f_{44} = f_{-+}\,[f_{++}/(f_{++} + f_{-+})]\,[f_{--}/(f_{-+} + f_{--})]. \quad (21)$$

The probabilities f_{++}, f_{+-}, f_{-+} and f_{--} of finding a $++$, $+-$, $-+$ or $--$ pair, which are identical to the quantities f_1, f_0, f_2 and f_3 respectively, introduced in chapter III, can be expressed easily in terms of the order parameters S, s and σ. We have

$$\begin{aligned} f_{++} &= \tfrac{1}{4}(1 + 2S - \sigma); \quad & f_{+-} &= \tfrac{1}{4}(1 + 2s + \sigma); \\ f_{--} &= \tfrac{1}{4}(1 - 2S - \sigma); \quad & f_{-+} &= \tfrac{1}{4}(1 - 2s + \sigma). \end{aligned} \quad (22)$$

In this way we find (observing that f_{40} and $f_{\bar{4}0}$ are the sum of two configuration probabilities):

$$
\begin{aligned}
f_{44} &= (1-2s+\sigma)\,(1+2S-\sigma)\,(1-2S-\sigma)/16(1+S-s)\,(1-S-s),\\
f_{\bar{4}\bar{4}} &= (1+2s+\sigma)\,(1+2S-\sigma)\,(1-2S-\sigma)/16(1+S+s)\,(1-S+s),\\
f_{40} &= (1-2s+\sigma)^2\,(1-\sigma)/8(1+S-s)\,(1-S-s),\\
f_{\bar{4}0} &= (1+2s+\sigma)^2\,(1-\sigma)/8(1+S+s)\,(1-S+s), \qquad (23)\\
f_{4\bar{4}} &= (1-2s+\sigma)^3/16(1+S-s)\,(1-S-s),\\
f_{\bar{4}4} &= (1+2s+\sigma)^3/16(1+S+s)\,(1-S+s),
\end{aligned}
$$

and hence, according to (14),

$$
\begin{aligned}
&F_{44} = F_{\bar{4}\bar{4}} =\\
&\quad = (1+2s+\sigma)\,[\{(1+\sigma)^2-4s^2\}\{(1-\sigma)^2-4S^2\}]^{\frac{1}{2}}/16[(1+S^2-s^2)^2-4S^2]^{\frac{1}{2}},\\
&F_{4\bar{4}} = F_{\bar{4}4} = \qquad (24)\\
&\quad = (1-2s+\sigma)\,[\{(1+\sigma)^2-4s^2\}\{(1-\sigma)^2-4S^2\}]^{\frac{1}{2}}/16[(1+S^2-s^2)^2-4S^2]^{\frac{1}{2}},\\
&F_{40} = F_{\bar{4}0} = (1-\sigma)\,\{(1+\sigma)^2-4s^2\}/8[(1+S^2-s^2)^2-4S^2]^{\frac{1}{2}}.
\end{aligned}
$$

§ 4. *The lowest-energy state.* We can now calculate the values of s^* and σ^* for which the expression (20) for $E_0^{(S)}$ is a minimum, at a given value of S, and the corresponding value of $E_0^{(S)}$. It is easily seen [6]) that the lowest possible value of $E_0^{(S)}$ is reached for $S = 0$. This energy, E_0, which is an approximation to the ground state of the linear chain, can be obtained by inserting (24) into (20), taking $S = 0$, and minimizing the expression in the right-hand member. We get, omitting asterisks,

$$E_0 = -\,NJ \max \eta(s, \sigma), \qquad (25)$$

where

$$\eta(s,\sigma) = \tfrac{1}{2}\sigma + (1-\alpha)\,[(1-\sigma^2)\,\{(1+\sigma)^2-4s^2\}^{\frac{1}{2}} + {} + (1-\sigma)\,\{(1+\sigma)^2-4s^2\}]/4(1-s^2). \qquad (26)$$

To find the maximum of $\eta(s, \sigma)$ we put

$$\partial\eta/\partial s = 0, \qquad (27)$$

$$\partial\eta/\partial\sigma = 0. \qquad (28)$$

The equation (27) has two solutions:

1) The solution $s = 0$ for which the two sublattices have equal magnetization. According to (28), the corresponding value of σ is

$$\sigma = \tfrac{1}{3}[7 + 3\alpha/(1-\alpha)]^{\frac{1}{2}} - \tfrac{1}{3}. \qquad (29)$$

For this value of s and σ, the extremum of $\eta(s, \sigma)$ is a maximum only if both

$$(\partial^2\eta/\partial s^2)_{s=0} = (1-\alpha)\,(1-\sigma)\,(\sigma^2 + 2\sigma - 2)$$

and

$$(\partial^2\eta/\partial\sigma^2)_{s=0} = -\,(1-\alpha)\,(4 + 12\sigma)$$

are negative, $(\partial^2\eta/\partial s\partial\sigma)_{s=0}$ being equal to zero. Since we consider only values of α in the region $0 \leqslant \alpha \leqslant 1$, σ is positive, in virtue of (29), and hence $\partial^2\eta/\partial\sigma^2$ is negative. We see, however, that $\partial^2\eta/\partial s^2 < 0$ holds only if $(\sigma + 1)^2 < 3$, *i.e.* for $\alpha < (24 - 4\sqrt{3})/33 = 0.517$.

2) For $\alpha \geqslant 0.517$, on the other hand, the position of the maximum of η is determined by the second solution of (27), which can be written in the form

$$\sigma = \tfrac{1}{2}(1+s)^{3/2}\,(3-s)^{1/2} + \tfrac{1}{2}(1-s)^{3/2}\,(3+s)^{1/2} - 1. \tag{30}$$

By substitution of (30) into (28) we obtain the relation between s and σ:

$$\alpha = 8\,\{8 + [(1-s^2)^{1/2}\,(9-s^2)^{1/2} - (1-s^2)]\,\cdot$$
$$\cdot\,[\tfrac{1}{2}(1+s)^{3/2}\,(3-s)^{1/2} + \tfrac{1}{2}(1-s)^{3/2}\,(3+s)^{1/2} + 2]\}^{-1}, \tag{31}$$

which is represented in fig. 1. For a given value of the anisotropy constant α,

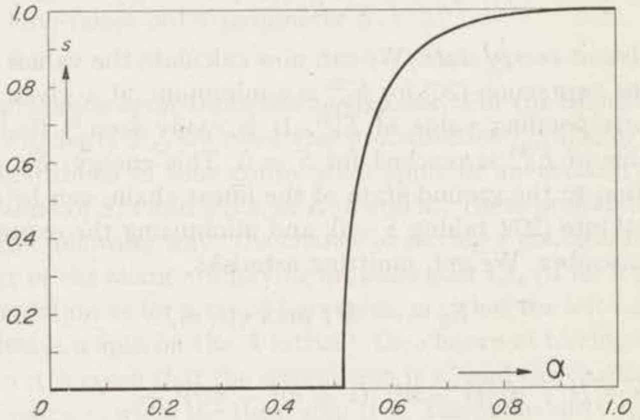

Fig. 1. The long-range order s in the ground state as a function of the anisotropy constant α.

the values of s and σ are determined by (30) and (31), while the corresponding value of E_0 is given by (25) and (26). The dependence of σ and E_0 on α, both for $\alpha < 0.517$ and $\alpha \geqslant 0.517$ is shown in fig. 2.

§ 5. *Concluding remarks.* From the foregoing analysis we may conclude that in the ground state of a linear lattice of spins ½ with an isotropic, antiferromagnetic coupling between nearest neighbours ($\alpha = 0$) no antiferromagnetic long-range order is present, the magnetization of each of the two sublattices being zero. This holds also for a chain with an anisotropic coupling of the form (1), (2) if the anisotropy constant α is not too large. For values of α larger than a certain critical value for which our approximative method gives $\alpha = 0.517$, long-range order occurs. The sublattice magnetizations which have opposite directions increase very rapidly with increasing anisotropy until finally, when α has reached the value 1 (Ising interaction), saturation is reached: $s = 1$. Although the minimum anisotropy required for long-range order is rather large in this case, one may expect that in two- and three-dimensional systems this anisotropy is much smaller, if not zero.

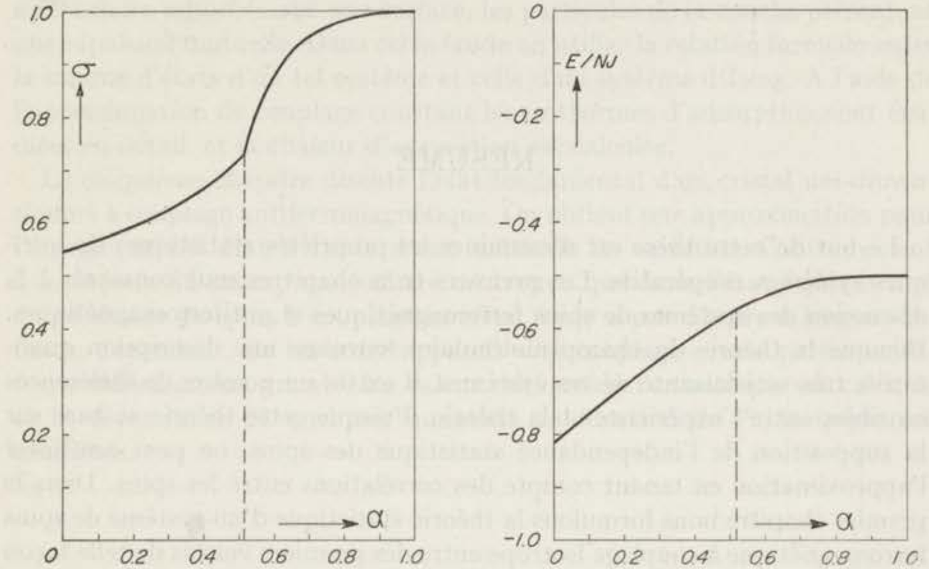

Fig. 2. The antiferromagnetic short-range order parameter σ in the ground state and the energy E_0 as functions of the anisotropy constant α.

REFERENCES

1) Van Vleck, J. H., J. chem. Phys. **9** (1941) 85.
2) Hulthén, L., Arkiv Mat. Astr. Fys. **26 A** (1938) No. 11, pp 36, 56.
3) Bethe, H. A., Z. Phys. **71** (1931) 205; *Handbuch der Physik* XXIV/2, J. Springer, Berlin (1933) 604, 618.
4) Anderson, P. W., Phys. Rev. **86** (1952) 694.
5) Slater, J. C., Phys. Rev. **35** (1930) 509.
6) Marshall, W., Proc. roy. Soc. **232 A** (1955) 48.

RÉSUMÉ

Le but de cette thèse est d'examiner les propriétés statistiques de quelques systèmes coopératifs. Les premiers trois chapitres sont consacrés à la discussion des systèmes de spins ferromagnétiques et antiferromagnétiques. Bienque la théorie du champ moléculaire fournisse une description qualitative très satisfaisante de ces systèmes, il existe un nombre de différences sensibles entre l'expérience et la théorie. Puisque cette théorie se base sur la supposition de l'indépendance statistique des spins, on peut améliorer l'approximation en tenant compte des corrélations entre les spins. Dans le premier chapitre nous formulons la théorie statistique d'un système de spins ferromagnétique à couplage isotrope entre des premiers voisins de telle façon que l'introduction d'une corrélation entre des spins voisins est très simple. La somme d'états du système est exprimée en termes d'une matrice densité d'une paire des premiers voisins. A l'aide de cette matrice densité nous introduisons l'énergie effective H_e d'une paire de spins, et nous montrons que H_e contient deux termes de couplage et un terme représentant un champ effectif agissant sur les deux spins de la paire. La supposition que les termes de couplage soient zéro conduit à l'approximation du champ moléculaire. On obtient une autre approximation en supposant que le couplage effectif est égal au couplage isotrope réel entre les spins; on peut démontrer que dans la limite $T \to \infty$ cette supposition est justifiée rigoureusement. Cette approximation „de couplage constant" est une généralisation directe de la méthode quasi-chimique pour un système de spins d'Ising. Les propriétés thermodynamiques du système sont calculées à l'aide de cette approximation, et des résultats numériques sont donnés pour plusieurs réseaux. Nous montrons que la difficulté d'un „anti-point de Curie" qui se présente dans la théorie de P. R. Weiss, est absente dans la susdite théorie.

Dans le deuxième chapitre nous étendons la théorie développée dans le premier chapitre aux systèmes de spins antiferromagnétiques à couplage isotrope entre des premiers voisins. Nous discutons le cas d'un champ extérieur parallèle et celui d'un champ perpendiculaire.

L'application de la théorie aux systèmes de spins ferro- et antiferromagné-

tiques à couplage d'Ising est donnée dans le troisième chapitre. Dans ce cas l'approximation de couplage constant apparaît être équivalente à l'approximation quasi-chimique. Nous développons cette approximation en détail pour un système antiferromagnétique dans un champ extérieur parallèle.

Le quatrième chapitre décrit l'étude des propriétés d'une couche monomoléculaire adsorbée sur une surface, les particules de la couche présentant une répulsion mutuelle. Dans cette étude on utilise la relation formelle entre la somme d'états d'un tel système et celle d'un système d'Ising. A l'aide de l'approximation de couplage constant les isothermes d'adsorption sont étudiées en détail, et la chaleur d'adsorption est calculée.

Le cinquième chapitre discute l'état fondamental d'un cristal uni-dimensionnel à couplage antiferromagnétique. On obtient une approximation pour l'énergie de cet état à l'aide d'une méthode de variation. Contrairement à la théorie du champ moléculaire on trouve que dans l'état fondamental d'une chaîne de spins à couplage isotrope il n'y a pas d'ordre à longue distance. Cependant, un tel ordre apparaît si le couplage a une partie anisotrope suffisamment grande. La valeur critique de l'anisotropie est calculée et la dépendance de quelques quantités en fonction de l'anisotropie est examinée.

SAMENVATTING

In dit proefschrift worden de statistische eigenschappen van enige coöperatieve systemen onderzocht met behulp van bepaalde benaderingsmethoden. De eerste drie hoofdstukken zijn gewijd aan een beschouwing van ferromagnetische en antiferromagnetische spinsystemen. Hoewel de eigenschappen van deze systemen kwalitatief goed worden beschreven door de theorie van het moleculaire veld, is er op een aantal belangrijke punten geen overeenstemming tussen experiment en theorie. Daar deze theorie berust op de veronderstelling, dat de spins statistisch onafhankelijk van elkaar zijn, kan men een betere benadering ontwikkelen door rekening te houden met het bestaan van correlaties tussen de verschillende spins. In hoofdstuk I wordt de statistische theorie van een ferromagnetisch spinsysteem met isotrope koppeling tussen naaste buren zodanig geformuleerd, dat men op eenvoudige wijze correlaties tussen naburige spins in rekening kan brengen. Hiertoe wordt de toestandssom van het spinsysteem uitgedrukt in de dichtheidsmatrix van twee naburige spins. Met behulp van deze paar-dichtheidsmatrix wordt de effectieve energie H_e van een spinpaar ingevoerd; H_e blijkt twee koppelingstermen te bevatten en een term die een effectief veld, werkend op de twee spins, voorstelt. De veronderstelling, dat de koppelingstermen nul zijn, blijkt te leiden tot de benadering van het moleculaire veld. Een volgende benadering wordt verkregen door te veronderstellen, dat de effectieve koppeling gelijk is aan de werkelijke, isotrope koppeling tussen de spins; bewezen wordt, dat in de limiet $T \rightarrow \infty$ aan deze veronderstelling streng voldaan is. Deze „benadering van constante koppeling" vormt een rechtstreekse veralgemening van de bekende quasi-chemische benadering voor een Ising-spinsysteem. Uitgaande van deze benadering van constante koppeling worden de thermodynamische eigenschappen van het spinsysteem berekend, terwijl voor een aantal roosters numerieke resultaten gegeven worden. Het blijkt, dat deze theorie, in tegenstelling tot die van P. R. Weiss, niet het bezwaar heeft, het bestaan van een anti-Curiepunt te voorspellen.

In hoofdstuk II wordt de theorie, ontwikkeld in hoofdstuk I, uitgebreid tot antiferromagnetische spinsystemen met isotrope koppeling tussen naaste buren. Zowel het geval van een evenwijdig als dat van een loodrecht uitwendig veld wordt behandeld. Hier blijkt de moeilijkheid van een anti-Curiepunt wel op te treden.

De toepassing van de theorie op ferro- en antiferromagnetische spinsystemen met Ising-wisselwerking wordt in hoofdstuk III gegeven. In dit geval is de benadering van constante koppeling equivalent met de quasi-chemische benadering. De benadering wordt in detail uitgewerkt voor een antiferromagnetisch spinsysteem in een evenwijdig uitwendig veld.

In hoofdstuk IV worden de eigenschappen van een monomoleculaire laag van geadsorbeerde deeltjes met een afstotende potentiaal tussen naburige deeltjes onderzocht door gebruik te maken van de formele relatie tussen de toestandssom van een dergelijk systeem en die van een Ising-systeem. Met behulp van de benadering van constante koppeling worden de adsorptieisothermen in detail bestudeerd en de adsorptiewarmte berekend.

Hoofdstuk V behandelt de grondtoestand van een eendimensionaal spinsysteem met antiferromagnetische koppeling tussen naaste buren. Door middel van een variatiemethode wordt een benadering voor de energie van deze grondtoestand verkregen. Anders dan in de theorie van het moleculaire veld vindt men, dat de grondtoestand van een keten van spins met isotrope koppeling geen antiferromagnetische lange-afstandsorde vertoont. Is de koppeling echter voldoende sterk anisotroop, dan treedt een dergelijke lange-afstandsorde wel op. De kritieke waarde van de anisotropie wordt berekend en de wijze waarop verschillende grootheden van de anisotropie afhangen wordt onderzocht.

Op verzoek van de Faculteit der Wis- en Natuurkunde volgt hier een kort overzicht van mijn academische studie.

In 1942 liet ik mij, na het afleggen van het eindexamen β aan het Stedelijk Gymnasium te Leiden, inschrijven aan de Universiteit van Amsterdam voor de studie in de chemie. Drie maanden later echter moest ik deze door de bekende oorlogsomstandigheden onderbreken. In 1945 hervatte ik mijn studie, nu aan de Rijksuniversiteit te Leiden. Na twee jaar scheikunde gestudeerd te hebben besloot ik mij te gaan voorbereiden op het candidaatsexamen letter A. Dit examen legde ik in juli 1948 af. Tijdens het practische gedeelte van mijn doctorale studie, van september 1948 tot maart 1949, was ik werkzaam bij het onderzoek van Dr. J. H. Mellink over de transportverschijnselen in vloeibaar helium II. In april 1951 legde ik het doctoraal examen met hoofdvak theoretische natuurkunde en bijvakken wiskunde en mechanica af.

Onder leiding van Prof. Dr. H. A. Kramers hield ik mij daarna bezig met onderzoekingen op het gebied van de theorie van ferromagnetisme en antiferromagnetisme, o.a. over de grondtoestand van een antiferromagnetische lineaire keten. Hierin kwam een onverwachte keer door het overlijden van Prof. Kramers in april 1952. Het wegvallen van zijn inspirerende leiding betekende ook voor mij een groot verlies. Hoeveel ik, ook naast mijn directe wetenschappelijke vorming, aan hem te danken heb, kan ieder begrijpen, die het voorrecht heeft gehad bij hem te werken, of op andere wijze de invloed van zijn persoonlijkheid heeft ondergaan.

Het onderzoek werd hierna voortgezet, eerst in nauw overleg met Prof. Dr. L. J. Oosterhoff, en na de benoeming van Prof. Dr. S. R. de Groot als opvolger van Prof. Kramers onder diens leiding en in samenwerking met Dr. J. van Kranendonk. De resultaten van dit werk zijn neergelegd in de eerste vier hoofdstukken van dit proefschrift.

Inmiddels was ik in october 1951 in dienst getreden bij het Koninklijke/Shell-Laboratorium te Amsterdam (N.V. De Bataafsche Petroleum Maatschappij). Ik bleef echter in de gelegenheid, mij geheel aan het onderzoek in Leiden te wijden.

PUBLICATIES

"Discussion of the properties of helium II on the basis of the two fluid model" (met C. J. Gorter en J. H. Mellink), Physica **16** (1950) 113.

"The lowest energy state of a linear antiferromagnetic chain", Physica **18** (1952) 104.

"Constant coupling approximation for Heisenberg ferromagnetism" (met J. van Kranendonk), Physica **22** (1956) 317.

"Constant coupling approximation for antiferromagnetism" (met J. van Kranendonk), Physica **22** (1956) 367.

"Constant coupling approximation for Ising spin systems", Physica **22** (1956) 387.

"The statistics of an adsorbed monolayer with repulsive interaction between the adsorbed particles", Physica **22** (1956) 397.

Instituut-Lorentz
Langebrug 111
Leiden

Op verzoek van de Faculteit der Wis- en Natuurkunde, volgt hier een kort overzicht van mijn academische studie.

In 1942 heb ik mij, na het afleggen van het eindexamen β aan het Stedelijk Gymnasium te Leiden, laten inschrijven aan de Universiteit van Amsterdam voor de studie in de chemie. Drie maanden later echter moest ik deze door de bekende oorlogsomstandigheden onderbreken. In 1945 hervatte ik mijn studie, nu aan de Rijksuniversiteit te Leiden. Na twee jaar scheikunde gestudeerd te hebben besloot ik mij te gaan voorbereiden op het candidaatsexamen letter A. Dit examen legde ik in juli 1948 af. Tijdens het praktische gedeelte van mijn doctorale studie, van september 1948 tot januari 1951, was ik werkzaam bij het onderzoek van Dr. J. H. Mellink over de transportverschijnselen in vloeibaar helium II. In april 1951 legde ik het doctoraal examen met hoofdvak theoretische natuurkunde en bijvakken wiskunde en mechanica af.

Onder leiding van Prof. Dr. H. A. Kramers hield ik mij daarna bezig met onderzoekingen op het gebied van de theorie van ferromagnetisme en antiferromagnetisme, o.a. over de grondtoestand van een antiferromagnetische lineaire keten. Hierin kwam een onverwachte keer door het overlijden van Prof. Kramers in april 1952. Het wegvallen van zijn inspirerende leiding betekende ook voor mij een groot verlies. Hoeveel ik, ook naast mijn directe wetenschappelijke vorming, aan hem te danken heb, kan ieder begrijpen, die het voorrecht heeft gehad bij hem te werken, of op andere wijze de invloed van zijn persoonlijkheid heeft ondergaan.

Het onderzoek werd hierna voortgezet, eerst in nauw overleg met Prof. Dr. L. J. Oosterhoff en na de benoeming van Prof. Dr. S. R. de Groot als opvolger van Prof. Kramers onder diens leiding en in samenwerking met Dr. J. van Kranendonk. De resultaten van dit werk zijn neergelegd in de eerste vier hoofdstukken van dit proefschrift.

Inmiddels was ik in october 1951 in dienst getreden bij het Koninklijke/Shell Laboratorium te Amsterdam (N.V. De Bataafsche Petroleum Maatschappij). Ik bleef echter in de gelegenheid mij geheel aan het onderzoek te Leiden te wijden.

PUBLICATIES

"Comparison of the properties of helium II as the basis of the two fluid model" (met C. J. Gorter en J. H. Mellink), Physica 16 (1950) 113.

STELLINGEN

I

Van een spinsysteem met antiferromagnetische Ising-wisselwerking kan men de susceptibiliteit bij het absolute nulpunt in een richting loodrecht op de voorkeursrichting streng berekenen. De aldus gevonden waarde is gelijk aan die van de loodrechte susceptibiliteit van een spinsysteem met isotrope wisselwerking volgens de spingolftheorie.

II

De ordeparameter, die Li gebruikt om de korte-afstandsorde in een antiferromagnetisch kristal te beschrijven, is niet de fysische grootheid, die bij de diffractie van neutronen aan een dergelijk kristal een rol speelt.

Y. Y. Li, Phys. Rev. **80** (1950) 457.

III

De verklaring, die Anderson geeft voor het falen van de door P. R. Weiss ontwikkelde theorie van het ferromagnetisme, voor temperaturen ver beneden het Curiepunt, is niet afdoende. Zij toont slechts aan, dat de theorie in dit temperatuurgebied geen goede benadering is, niet, dat de gemaakte verwaarlozingen het optreden van een anti-Curiepunt impliceren.

P. W. Anderson, Phys. Rev. **80** (1950) 922.

IV

Bij zijn berekening van de soortelijke warmte van een ferromagnetisch kristal leidt P. R. Weiss een uitdrukking af voor de gemiddelde energie van een „spin cluster". Men kan aantonen, dat deze uitdrukking niet de juiste is.

P. R. Weiss, Phys. Rev. **74** (1948) 1493.

Stellingen behorende bij proefschrift P. W. Kasteleijn

V

Men kan de statistische eigenschappen van een spinsysteem met isotrope wisselwerking tussen naaste buren beschrijven in termen van de effectieve energie van een systeem bestaande uit slechts enkele spins. Hierbij is een systeem bestaande uit twee naburige spins te verkiezen boven een systeem bestaande uit één spin tezamen met al zijn naaste buren.

P. R. Weiss, Phys. Rev. **74** (1948) 1493.
Hoofdstuk I van dit proefschrift.

VI

De isothermen van een monomoleculaire geadsorbeerde laag met een afstotende potentiaal tussen naburige deeltjes zullen beneden een zekere kritieke temperatuur twee knikken vertonen. Deze knikken corresponderen met overgangen tussen een toestand met en een toestand zonder onderrooster-ordening.

Hoofdstuk IV van dit proefschrift.

VII

De opvatting van Dingle, als zou de z.g. relativistische verjongingskuur geen reëel verschijnsel zijn, is onjuist.

H. Dingle, Nature **177** (1956) 782.

VIII

Bij pogingen om de triplettoestand van de door Evans door middel van susceptibiliteitsmetingen onderzochte optisch geëxciteerde moleculen met behulp van paramagnetische resonantie te onderzoeken zal men glasvormende koolwaterstoffen als oplosmiddel dienen te nemen in plaats van het door Evans gebruikte boorzuurglas.

D. F. Evans, Nature **176** (1955) 777.

IX

Bij de röntgenografische bepaling van de structuur van fluoreen is het principieel onjuist, de waterstofatomen mede in rekening te brengen door gebruik te maken van het verstrooiende vermogen, zoals dat is bepaald door Robertson.

D. M. Burns en J. Iball, Proc. roy. Soc. **A227** (1955) 200.
J. M. Robertson, Proc. roy. Soc. **A150** (1935) 106.

X

Het is onjuist, bij de beschrijving van de „paradox van Gibbs" te spreken van een „onaanvaardbare discontinuïteit van oorzaak en gevolg".

A. Landé, *Foundations of Quantum Mechanics* (Yale University Press, New Haven, 1955) p. 10–14.

XI

Uit de proeven van Schneider c.s. volgt overtuigend, dat de opvatting, dat de coëxistentiekromme van een gas-vloeistofsysteem een vlakke top heeft, op een verkeerde interpretatie der experimenten berust.

W. G. Schneider en H. W. Habgood, J. chem. Phys. **21** (1953) 2080.
O. K. Rice, J. chem. Phys. **23** (1955) 164.

XII

De resultaten van de metingen van Beun, Steenland, de Klerk en Gorter over de susceptibiliteit van chroommethylaminealuin in een transversaal veld stemmen niet overeen met die van de metingen van Hudson en McLane. Deze discrepantie kan worden verklaard door aan te nemen, dat tijdens de metingen van laatstgenoemden de in Leiden ontdekte anomalie in de richting van het uitwendig magneetveld was gelegen.

J. A. Beun, M. J. Steenland, D. de Klerk en C. J. Gorter, Physica **21** (1955) 767.
R. P. Hudson en C. K. McLane, Phys. Rev. **95** (1954) 932.

XIII

Het zou aanbeveling verdienen, metingen over de absorptie en dispersie van gewoon geluid in helium II uit te voeren bij temperaturen lager dan 1°K en frequenties lager dan 1 MHz.

XIV

De coëfficiënten in een lineaire regressievergelijking worden bepaald door het voorschrift, dat de residuvector volgens een bepaalde metriek zo klein mogelijk is. Bekend is, dat dan voldaan is aan de voorwaarde, dat de geschatte grootheden zo klein mogelijke spreiding hebben. Het is mogelijk, omgekeerd uit deze voorwaarde bovengenoemd voorschrift af te leiden, met inbegrip van de gebruikte metriek.

J. van IJzeren, Statistica **8** (1954).

XV

Het bewijs van de bewering van van Melsen, dat aan de relativiteitstheorie de vooronderstelling van het bestaan van een absolute gelijktijdigheid ten grondslag ligt, bevat een principiële onjuistheid.

A. G. M. van Melsen, *Natuurphilosophie* (N.V. Standaard-Boekhandel, Antwerpen–Amsterdam, 1955) pp. 234–236.

XVI

De karakterisering die van Melsen geeft van de opvatting van de natuurwetenschap der 20e eeuw inzake het probleem der elementaire kwaliteiten is onjuist. Zijn beschrijving van het standpunt der 19e eeuwse natuurwetenschap geeft echter een goede typering van dat der 20e eeuwse wetenschap.

A. G. M. van Melsen, id., pp., 261–263.