THE STATISTICAL MECHANICS OF THREE DIMENSIONAL FERROMAGNETS

being a thesis presented by

Edward Dempsey, B.Sc.,

to the University of St. Andrews

in application for the degree of

Doctor of Philosophy.

DECLARATION

I hereby declare that this thesis is entirely based on the results of work carried out by myself in the Department of Natural Philosophy, St. Andrews University. The thesis is my own composition and it has not previously been presented for a higher degree.

CERTIFICATE

I hereby certify that Edward Dempsey has spent nine terms engaged in research work under my direction, that he has fulfilled the conditions of Ordinance 16 (St. Andrews) and that he is qualified to submit this thesis in application for the degree of Doctor of Philosophy.

Research Supervisor.

25-7-57

PERSONAL PREFACE

I matriculated in the United College of the University of St. Andrews in October 1948 and in July 1952 graduated with First Class Honours in Natural Philosophy. In October 1952 I commenced research in the Department of Natural Philosophy under the supervision of Dr. D. ter Haar, Lecturer in Theoretical Physics in the University of St. Andrews. It is the result of this research which is described in this thesis.

CONTENTS

I

INTRODUCTION

§ 1 The problem of providing a theoretical description of the experimentally observed behaviour of ferromagnets is a particular example of the broader problem of the theoretical treatment of co-operative phenomena in general. Fortunately in a first approach, a treatment may be given which applies generally; in fact we shall demonstrate in the next paragraph, that a discussion of ferromagnetism say, applies equally well under certain circumstances, to the problems of antiferromagnetism and of substitutional alloys. We make the transition to different problems simply by altering the significance of our unit of energy to suit a particular problem. Comparatively recently Domb and Potts [1]) have included in their general discussion of a two dimensional model, as well as the above problems, that of the adsorbed monolayer. This has also been discussed by Fowler and Guggenheim [2]).

Co-operative phenomena are characterised by the fact that within a system of atoms or molecules showing the phenomena, groups of atoms say, combine to form sub-systems which maintain their physical character despite the disrupting influence of thermal agitation. This implies the existence of a dynamical equilibrium within the sys-

tem; as atoms leave the sub-systems, due to thermal agitation, other atoms are brought into them in order to maintain the extent of the sub-systems peculiar to the temperature of the system. In general, the extent of the sub-systems, or the degree of co-operation, varies uniquely with temperature, increasing with decreasing temperature and vice versa. This is only true however, below a certain temperature, which is a characteristic of any particular substance and above which it may be said, in a first approximation, no degree of co-operation exists.

The variation of the degree of co-operation with temperature is not linear. At low temperatures it is difficult to decrease the degree of co-operation. As the temperature is increased however, it becomes relatively easier until near the characteristic temperature, an avalanche effect sets in and the co-operation apparently disappears abruptly. On closer investigation it is observed that the co-operation has not completely gone; a small residual effect remains, which is due to a slight degree of local co-operation. We shall disregard this for the time being.

In the following discussion, we shall normally identify the degree of co-operation in a system with the degree of order in the system, while the characteristic temperature will normally be referred to, on analogy with the theory of ferromagnetism, as the Curie or transition temperature.

The transition is then understood to be that from the state of finite order to that of zero order (strictly speaking, that of zero long range order).

In any theory of co-operative phenomena, we must ultimately derive results which we can compare with results obtained by experiment. We must now, therefore consider the physical manifestations of the state of order of any system we choose to investigate.

We should expect on the third law of thermodynamics, that as we approach the absolute zero of temperature, the degree of order in a system will increase, having a maximum value at the absolute zero. This is the state of perfect order and minimum energy. If we now supply heat to the system, its temperature will be raised; but since at any finite temperature below the Curie temperature the system must be disordered to a certain extent, some of the energy supplied as heat must be taken up in producing the equilibrium value of the degree of order appropriate to the final temperature of the system. We see then that, if the only anomalous effect we have to consider in our system is that of the order-disorder transition, any energy supplied to the system will be distributed in the way we have suggested. As we increase the temperature still further, the amount of energy necessary to produce the corresponding degree of disorder will increase until at the Curie

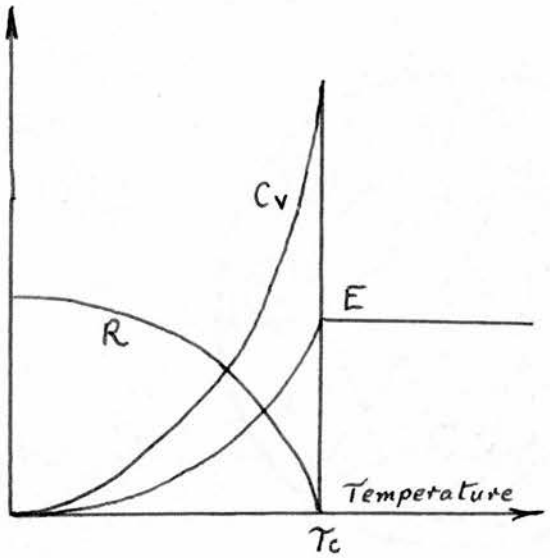


Figure 1.

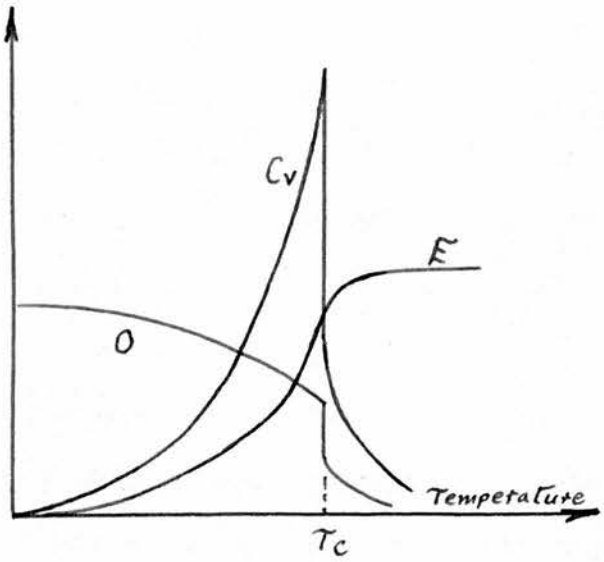


Figure 2.

temperature Tc, where the degree of disorder is a maximum the energy producing the disordering will remain constant. Any further energy supplied will merely heat up still further the disordered system. The energy which it is necessary to put into the system to produce the equilibrium degree of disorder for a particular temperature, we shall call the configurational energy. Its variation is shown qualitatively in figure 1. Corresponding to this anomalous disordering or configurational energy in the system, we shall have an anomalous specific heat contribution - this we shall call the configurational specific heat. Its course is also shown in figure 1. For completeness we have included in the figure the curve showing the variation of order (curve R) with temperature.

If we take account of the small degree of residual order which remains above the Curie temperature, our various curves are modified in accordance with figure 2. In this sketch we have designated order by the symbol O, since we wish to reserve R for the long range order - which does disappear completely at Tc. We see from figure 2, that the specific heat curve now shows a true lambda point.

From this discussion then, it seems reasonable to expect that any theory of co-operative phenomena may be judged as to the validity of the approximations inherent in its formulation on the accuracy with which it predicts first of

all the Curie temperature and secondly the course of the variation of configurational specific heat with temperature. In order to be able to study the variation of specific heat, or in fact of any of the usual thermodynamic functions, with temperature, we shall use the methods of statistical mechanics. In fact, from the nature of the problem, a statistical approach is the only possible one.

It is well known that in any statistical problem once we have succeeded in evaluating the partition function, Z for a system, it is relatively simple to obtain most of the thermodynamic functions. We may mention in particular, E, the internal energy of the system, and C_v, the specific heat at constant volume - since these are the two functions with which we shall be concerned. The partition function is defined as

$$Z = \sum_{\substack{\text{all states} \\ j}} e^{-E_j \beta} = e^{-F\beta} \qquad (1,1)$$

where $\beta = \frac{1}{kT}$, in which k is Boltzmann's constant and T is the absolute temperature. E_j is the energy of the system in state j and F is the free energy.

For the total energy E and the specific heat C_v of the system, we have

$$E = -\frac{\partial \ln Z}{\partial (\frac{1}{kT})} = -\frac{\partial \ln Z}{\partial \beta}$$

$$C_v = \frac{1}{kT^2} \frac{\partial^2 \ln Z}{\partial (1/kT)^2} = k\beta^2 \frac{\partial^2 \ln Z}{\partial \beta^2}$$

These are equations (1,2)

Actually, in this work, we shall not be concerned so much with Z, which we might call the macroscopic partition function, as with λ, the partition function per particle. We define

$$\ln \lambda = \frac{\ln Z}{N}$$

since ln Z is an extensive quantity. N is the total number of particles in the system. Our configurational energy and specific heat will then refer to the average values of these quantities per particle and will be calculated with respect to ln λ.

As we have mentioned, the main physical problems to which our theory of co-operative phenomena will apply are those of ferromagnetism, antiferromagnetism and substitutional solid solutions or alloys. Before going further, we should consider the properties of each of these types of material from the point of view of the order-disorder transition in each.

For ferromagnetism and antiferromagnetism we shall use a simple model due to Ising [3]. This assumes that in a crystal, the lattice points are occupied by spins which may spin in only two possible directions - either parallel

or antiparallel to a possible field direction H (spin quantum number s = 1/2). The degree of magnetisation shown by the crystal is then proportional to the degree of ordering of the spins. From our point of view, the only difference between a ferromagnetic and an antiferromagnetic substance lies in the state of perfect order. For the former, at absolute zero, we assume that the lattice spins are all lined up parallel to each other, while for the latter, the spins are lined up alternately parallel and antiparallel. For a ferromagnetic substance, for instance, as the temperature is raised, more and more of the spins turn over and the degree of ferromagnetism decreases, until at the Curie temperature it disappears completely. Of course in this discussion, which we have said is to be generally applicable to co-operative phenomena, we shall discuss magnetic phenomena only in the absence of an external magnetic field.

For substitutional solid solutions, we shall consider our lattice sites occupied by different types of atoms. We have just mentioned the no-field restriction which must be imposed upon the theory as it will apply to magnetic phenomena if this is to be equivalent to a general theory. Similarly we must impose corresponding limitations on our model of a substitutional alloy. It is obvious that in the state of complete disorder in either of the magnetic models

there will be an equal number of spins parallel and anti-parallel. These will be arranged at random throughout the lattice. To correspond to this we must, in our substitutional alloy, consider only two types of atom (corresponding to the fact that the magnetic spins can have only two orientations). These we designate A and B atoms. Also we must have equal numbers of A and B atoms which, in the state of complete disorder, will be distributed at random through the lattice.

A somewhat more general treatment than ours has been discussed by Rushbrooke [4]) in which he does not restrict his binary solid solution to equal numbers of A and B atoms. He reduces his problem to the evaluation of a function $\Lambda(\lambda, \zeta)$ which for the case of the ferromagnet, represents the partition function of the assembly, while for the case of the binary solid it represents the grand partition function. In Rushbrooke's paper $\lambda = e^{2mH\beta}$ and $\zeta = e^{-\omega\beta}$ for the case of the ferromagnet, while $\lambda = e^{\mu\beta}$ and $\zeta = e^{-\omega'\beta}$ for the case of the binary solid. Here m is the magnetic moment per spin, H is an external magnetic field, ω is the difference in energy of two neighbouring dipoles when they change from the antiparallel to the parallel state and $\omega' = \omega_{AB} - \frac{1}{2}(\omega_{AA} + \omega_{BB})$. ω_{AB}, ω_{AA} and ω_{BB} are the energies associated with AB, AA, and BB pairs of atoms in the binary solid solution. μ is the

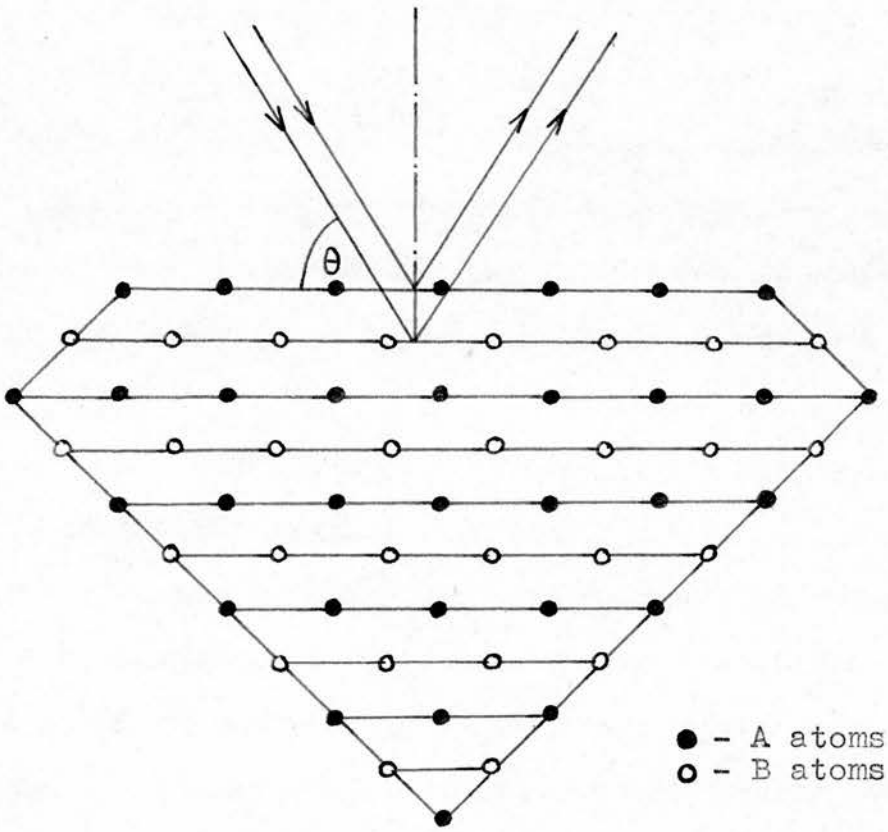


Figure 3.

chemical potential. Some of these quantities we shall have to refer to once again in the next section.

It is perhaps appropriate at this point, to mention two other manifestations of the state of order in a substitutional alloy. Tammann [5], investigating copper-gold alloys of varying composition, found that if the alloy contained a maximum of 50 % of copper atoms, nitric acid did not affect it, whereas copper was dissolved if the atomic constitution contained more than 50 % of copper. He interpreted this to mean that equal numbers of gold and copper atoms could form an ordered structure, while any excess copper atoms (and presumably also gold atoms) could not be fitted into the structure.

X-ray and neutron diffraction experiments have also given indication of the existence of ordered structures within crystals containing different types of atoms [6]. The sketch - figure 3 - shows an ordered arrangement of a crystal plane containing equal numbers of A and B atoms. We assume, of course, preference for unlike nearest neighbours i.e. the state of lowest energy is that where every A atom is surrounded by B atoms and vice versa.

If the elements A and B are not too near to each other in the periodic table, their reflection properties for X-rays will be different and so for various suitable angles θ , where the path length difference is a multiple

of $\lambda/2$, we may get a line in the X-ray spectrum. If the reflecting planes have the same reflection properties, as for instance if all the atoms are the same, no line will appear. If on the other hand, we have A and B atoms which, however, are arranged at random, we shall on the average have equal numbers of A and B atoms on any reflecting plane and the reflection properties of any two lattice planes will thus be the same. We see then, that in a crystal of an alloy, we shall get additional X-ray spectrum lines (superstructure lines) when a state of order exists, which will be absent above the Curie temperature for the crystal. Neutron diffraction may give results similar to the above in the case where the A and B atoms have X-ray properties which are too similar to each other.

§ 2 Before giving a short historical account of the development of the theory of co-operative phenomena, we must first show that results for any one of the problems we have mentioned are immediately applicable to the remaining problems. This is important not merely from the point of view of unifying the theory; but also since, of the different methods which have been developed for handling the problems, some have been applied to the theory of substitutional solid solutions and some have been applied to the theory of ferromagnetism.

Consider a lattice of equal numbers of A and B atoms and let there be a total of N atoms in the lattice. Let us assume a co-ordination number z - that is, if we neglect surface effects, the number of nearest neighbours of any atom in the lattice is z. We assume also that we have only nearest neighbour interactions between the atoms on the lattice sites. Now if we take for the energies associated with AA, BB, and AB pairs of atoms at nearest neighbour positions, the values v_{AA}, v_{BB}, and v_{AB} and if, of the total of Q nearest neighbour pairs in the lattice, Q_{AA}, Q_{BB}, and Q_{AB} are the numbers of AA, BB, and AB pairs respectively, then for the configurational energy of the lattice we may write

$$E' = Q_{AA} v_{AA} + Q_{BB} v_{BB} + Q_{AB} v_{AB} \qquad (1,3)$$

but, neglecting surface effects, $Q_{AA} = Q_{BB}$ and so

$$E' = Q_{AA} (v_{AA} + v_{BB}) + Q_{AB} v_{AB} \qquad (1,4)$$

Since $Q = Q_{AA} + Q_{BB} + Q_{AB} = 2 Q_{AA} + Q_{AB}$

the last equation may be written

$$E' = \tfrac{1}{2} (Q - Q_{AB})(v_{AA} + v_{BB}) + Q_{AB} v_{AB} \qquad (1,5)$$

or

$$E' = \frac{Q}{2} (v_{AA} + v_{BB}) - Q_{AB} \left[\tfrac{1}{2} (v_{AA} + v_{BB}) - v_{AB}\right]$$

$$= E_o - Q_{AB} v \qquad (1,6)$$

where E_0 is a constant, and

$$v = \tfrac{1}{2}(v_{AA} + v_{BB}) - v_{AB} = \text{constant} \qquad (1,7)$$

We shall rewrite equation (1,6) as

$$E'' = E' - E_0 = -Q_{AB}\, v \qquad (1,8)$$

Now from our discussion of antiferromagnetism, we know that the energy of an antiferromagnetic lattice is a minimum when the spins are lined up alternately parallel and antiparallel to each other. We can assign to each spin a parameter μ_i' where i refers to the i^{th} spin, such that $\mu_i' = +1$ if the spin is in the parallel direction (to an external field H say, where H may subsequently be diminished to zero) and $\mu_i' = -1$ if the spin is in the antiparallel direction. We also define the energy J to be the energy gained when two isolated nearest neighbour spins change from the parallel-antiparallel state to the both-parallel state. If this is so, then we may write the configurational energy of an antiferromagnetic lattice as

$$E = \tfrac{1}{2} J \sum \mu_i' \mu_j' \qquad (1,9)$$

where the summation is over all nearest neighbours. Here we see that -E is a maximum $= \frac{1}{4} J N z$ in the state of perfect order, since each product $\mu_i \mu_j$ gives -1, so that E is a minimum as we should expect. The factor 1/2

comes in since the summation alone makes each energy change twice as great as it should be. We can see this if we consider the state of perfect order, with minimum energy, and then turn one spin over. If there are z nearest neighbours, it is obvious that the energy increase should be $z\mathcal{J}$. However doing this via the summation, we first have to subtract the pairs affected by the flipping over of the spin, i.e. z pairs, and then add them on again with the sign changed. This altogether is equivalent to subtracting 2z pairs and gives an energy increase of $2z\mathcal{J}$ units.

We can now, in our AB lattice, define a parameter μ_i'' referring to the atom on the i^{th} site, such that μ_i'' is +1 if the atom is an A atom, -1 if it is a B atom. Thus we may write

$$\sum \mu_i'' \mu_j'' = 2Q_{AA} - Q_{AB} \qquad (1,10)$$

But as $Q = 2Q_{AA} + Q_{AB}$, we have

$$\sum \mu_i'' \mu_j'' = Q - 2Q_{AB} \qquad (1,11)$$

or $$Q_{AB} = \tfrac{1}{2}(Q - \sum \mu_i'' \mu_j'') = \tfrac{1}{4} zN - \tfrac{1}{2} \sum \mu_i'' \mu_j'' \qquad (1,12)$$

where we have used the fact that $Q = \tfrac{1}{2} zN$ - each of our N lattice points has z nearest neighbours; but since this counts each pair twice, we introduce the factor 1/2.

Substituting into (1,8), equation (1,12), we have

$$E'' = \tfrac{1}{2} v \sum \mu_i'' \mu_j'' - \tfrac{1}{4} z N v$$

or shifting our zero of energy

$$E = E'' + \tfrac{1}{4} v z N = \tfrac{1}{2} v \sum \mu_i'' \mu_j'' \qquad (1,13)$$

which has the same form as has (1,9).

We see then that we can reduce the problem of the substitutional solid solution to that of the antiferromagnet quite simply.

Now consider the problem of the ferromagnet. At absolute zero, the spins are all lined up parallel to each other. If we want the energy at the absolute zero to be a minimum, we must write

$$E = -\tfrac{1}{2} J \sum \mu_i \mu_j \qquad (1,14)$$

where the μ_i can be taken to be the same as the μ_i' except that they refer to the ferromagnet. The minus sign is necessary since in the state of perfect order, each product $\mu_i \mu_j$ will give +1. J is the energy gained when two isolated nearest neighbours in a ferromagnet change from the low to the high energy state. The summation, of course, is again over nearest neighbour pairs.

In the AB lattice, we now label our lattice sites calling them alternately α and β sites, so that each α site

is surrounded by z β sites and vice versa. Then in the state of perfect order, we shall have all the A atoms on α sites and all the B atoms on β sites. This introduces the notion of sites being correctly and incorrectly occupied. As we heat up the lattice, some of the atoms will go 'wrong' and we shall find A atoms on β sites and B atoms on α sites. It must be borne in mind, of course, that this labelling of sites is merely a convenient fiction adopted for the purpose of analysing the problem - in particular, when we reach the state of complete disorder, the labelling becomes completely meaningless.

We introduce once again, a parameter, say μ_i''' such that μ_i''' is +1 if the i^{th} site is an α site occupied by an A atom or a β site occupied by a B atom - a site correctly occupied in other words. $\mu_i''' = -1$ if the i^{th} site is an α site occupied by a B atom or a β site occupied by an A atom. Thus we see that an AB pair will give us $\mu_i'''\mu_j''' = +1$ and an AA or a BB pair will give $\mu_i'''\mu_j''' = -1$. Using this we find

$$\sum \mu_i'''\mu_j''' = Q_{AB} - 2Q_{AA} = 2Q_{AB} - Q \qquad (1,15)$$

and we see that the introduction of μ_i''' has had the effect of changing the right hand side of equation (1,11) making it negative.

We can rewrite (1,15) as $Q_{AB} = \frac{1}{2}\left[\sum \mu_i'''\mu_j''' + \frac{1}{2} \times N\right]$

and substituting into (1,8), we get

$$E'' = -\tfrac{1}{2} v \sum \mu_i''' \mu_j''' - \tfrac{1}{4} v z N$$

or $$E = E'' + \tfrac{1}{4} v z N = -\tfrac{1}{2} v \sum \mu_i''' \mu_j''' \qquad (1,16)$$

(compare (1,13))

This time we have reduced the problem of the substitutional solid solution to that of the ferromagnet, since (1,16) has the same form as has (1,14).

Altogether then, we can see that within the limitations we have set to our problems, we can discuss any one of them and immediately apply the results obtained to the others. A point brought out in this discussion, incidentally, is the fact that in the substitutional alloy problem, the important energy unit is not one of the basic units, v_{AA}, v_{BB}, or v_{AB}, but is the combination of them denoted by v, and given by equation (1,7).

§ 3. Finally, in order to show how the method we propose to use fits into the historical development of the theory of co-operative phenomena, we shall give a brief outline of the development and indicate the contributions made by the different workers in the field.

The first attempt to develop a theory of the order-disorder transition, was made by Bragg and Williams [7], who treated the question of the substitutional lattice containing equal numbers of A and B atoms. They labelled the

lattice sites α and β as we have done earlier and introduced the quantities r_α, r_β, w_α, and w_β such that:

r_α = fraction of α sites correctly occupied = $N_{A\alpha}/(\frac{N}{2})$

r_β = " " β " " " = $N_{B\beta}/(\frac{N}{2})$

w_α = fraction of α sites incorrectly occupied = $N_{B\alpha}/(\frac{N}{2})$

w_β = " " β " " " = $N_{A\beta}/(\frac{N}{2})$

Here $N_{A\alpha}$, $N_{B\beta}$, etc. denote the numbers of A atoms on α sites, the numbers of B atoms on β sites, etc. Bragg and Williams then define an order parameter which we shall term the long range order, and denote by R. In the original paper R is defined by an equation equivalent to

$$R = 2r_\alpha - 1 \qquad (1,17)$$

By long range order we mean the extent to which the lattice sites are correctly occupied. From the definition, we see that, for the state of perfect order, $r_\alpha = 1$ and $R = 1$, while for the state of complete dis-order, $r_\alpha = 1/2$ and $R = 0$. (This is the more obvious when we recollect that r_α may also be looked upon as being the probability that any α site is correctly occupied. In the state of randomness of distribution of atoms it is reasonable to suppose $r_\alpha = 1/2$).

The basic assumption of the Bragg-Williams theory may be summarised in the equation

$$J = R J_0 \qquad (1,18)$$

where J is the average energy required to exchange an A atom on an α site with a B atom on a β site at any particular temperature. J_0 is the value of J at the absolute zero of temperature. Obviously $J_0 = 2z\mathcal{J}$. It is very simple now, to derive an expression for the energy of the system, and we find that

$$E = E_0 (1 - R^2) + E_1 \qquad (1,19)$$

E_1, E_0 constants

Following out the usual statistical method, we write for the free energy of the system

$$F = E - TS = E - kT \ln W(R) \qquad (1,20)$$

where k and T have the usual meanings. S is the entropy and W (R) is the number of ways of arranging the system in accordance with a particular value of R. It may easily be seen that

$$W(R) = \binom{\frac{1}{2}N}{\frac{1}{2}N r_\alpha}\binom{\frac{1}{2}N}{\frac{1}{2}N w_\beta} \qquad (1,21)$$

where $\binom{a}{b} = {}_aC_b$

After some manipulation of (1,21) we can write the free energy in terms of R only, and then find the equilibrium value of R at a particular temperature in the usual manner from $\frac{\partial F}{\partial R} = 0$.

This leads us ultimately to the Bragg-Williams equations

$$R = \tanh X$$

$$X = \frac{2 E_0 R}{N k T} \qquad (1,22)$$

from which we can estimate T_c.

It is not really necessary to go into the discussion of this treatment of the problem much further, since it forms at best, a rather crude first approximation to the theory; we have outlined it in some detail for completeness and for its historical interest. Before going on, however, we must point out two shortcomings of the theory which are corrected in subsequent discussions.

First of all the results of the theory correspond to the curves of figure 1. - at T_c all order disappears. The reason for this of course, lies in our assumption of the existence of a long range order only. To bring the curves closer to physical reality, we must take account of the fact that above T_c there is still a tendency for A atoms to surround themselves with B atoms and vice versa. This involves the introduction of a short range order which is effectively included in all of the later theories.

The other point may be seen if we consider the summation

$$\sum_i \mu_i''' = N - 2\frac{N}{2}\omega_\beta - 2\frac{N}{2}\omega_\alpha$$

$$= N(2r_\alpha - 1) = NR \qquad (1,23)$$

where we have used the fact that $r_\alpha = r_\beta$.

In other words, $R = 1/N \sum_i \mu_i'''$.

But in the Bragg-Williams theory, we have an equation of the form

$$E = -\tfrac{1}{2} v \sum \mu_i''' \mu_j''' \quad \text{— from (1,16)}$$

$$\propto R^2 \propto \left(\sum_i \mu_i''' \right)^2$$

We see then that in effect we have used the approximation

$$\sum \mu_i''' \mu_j''' = \left(\sum_i \mu_i''' \right)^2$$

which implies that every atom in the lattice has all the other atoms as its nearest neighbours. It seems reasonable therefore, to assume that the Bragg-Williams approximation is a limiting case of other more accurate theories. This is in fact the case.

The work of Bragg and Williams naturally stimulated further work on the subject, since it was soon realised wherein the shortcomings of the theory lay.

Bethe [8]) was the first to introduce a short range parameter to measure how well on the average A atoms are surrounded by B atoms. He developed his method on a probability basis and in such a way that successive approximations could be taken in order to approach closer and closer to an exact solution of the problem. Unfortunately, going to only the second approximation increased the work

necessary for a solution by a considerable amount compared to that necessary for a solution on his first approximation. Furthermore, the improvement in the solution was not so very great. Bethe showed that the Bragg-Williams theory was a limiting case of his own theory.

Some time later, Fowler and Guggenheim [9]) published a general treatment which they termed the Quasi-Chemical method and which in some ways, may be looked upon as being a fairly straightforward extension of the Bragg-Williams method to include a short range order parameter. We shall not describe this method for reasons which we shall give shortly. Chang [10]) has shown that Bethe's method yields equations which are entirely equivalent to those of the Quasi-Chemical method and so we may look upon Bethe's method as being a particular case of the latter method.

Within the next few years, other methods were developed including one due to Kirkwood [11]), and one due to Zernike [12]) All of these were of course approximate methods, the approximation being capable, in principle, of being pushed as far as was desired.

In 1941, Kramers and Wannier [13]) introduced into the discussion of these problems, the so-called Variational Method. It is this method which will primarily concern us. They applied the method to a lattice consisting of a linear chain of atoms and to a square net lattice - a two

dimensional ferromagnet. Their calculations indicated that the method was considerably more powerful than any previous treatment; but they did not extend their work to include a third dimension.

It was not until 1952 that Martin and ter Haar [14]) applied the variational method to a three dimensional model. In their paper they treated the problem of the simple cubic lattice and found, in the region above the Curie temperature at least, that the variational method gave a much better result than did any of the other approximate methods [15]). They did not tackle the low temperature region.

In view of the success of the method, we decided to apply it to the problems of the face centred and body centred lattices and to attempt solutions for these models in the regions below their Curie temperatures as well as above them. Before going on to a description of the variational method, however, we must first point out that in order to assess the value of the method, we shall require results from the older methods which are comparable to those of the variational method. Unfortunately, most of the older methods were used to estimate the Curie temperatures of the various models discussed, whereas the variational method leads directly to series for the partition functions per spin in the regions above and below the Curie

temperature. Thus, in order to be able to compare our results from the variational treatment with those of the other treatments, we have had to calculate series for the partition functions per spin in the two temperature regions for the more powerful of the older methods.

It is for this reason, that we have merely mentioned in passing, the various methods following upon the Bragg-Williams treatment. We thought it more suitable to defer a description of these methods till a later chapter, when we can include at the same time, an outline of the way in which we have derived the partition functions per spin for the different methods. To the best of our knowledge, these series have not previously been calculated.

II.

THE VARIATIONAL METHOD

§ 1. We shall now consider the variational method as developed in 1941 by Kramers and Wannier for solving the problem of finding the partition function per spin of the two dimensional ferromagnet. Since the method, applied to three dimensional structures, represents the main part of this work, we wish to consider it in some detail.

Throughout the whole of this and the next chapter, we shall make use of the Ising model of a ferromagnet which we now want to define precisely.

We assume that our regular lattice structure has all of its sites occupied by spins, each of which we denote by a parameter μ_i as mentioned earlier. These spins may be found in only two possible orientations which we signify by allowing the spin parameter to take on the two possible values +1 and -1. The Ising model then states that interaction between spins is confined to that between nearest neighbours. In other words, the force on each spin, and consequently the energy associated with each spin, is due only to the orientation of its z nearest neighbours in the lattice - and possibly also to an external field which may be applied. We neglect surface effects of course. z we have defined earlier, to be the co-ordination number of the lattice. We should mention also that,

in the next chapter, we shall have to extend our definition of the model to include other than nearest neighbour interactions - which can be done without much difficulty.

If we assume that each spin is equivalent to all the others, then the total energy of the lattice will depend upon the interaction energy $\mathcal{J}$ between nearest neighbouring spins and upon the magnetic moment m per spin. Taking into account an external magnetic field H, we may write for the energy E of a ferromagnet (see equation (1,14))

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

The notation $\sum_{\langle i,j\rangle}$ we shall define to mean summation over all nearest neighbour pairs (i, j).

This is the energy in a particular state of the lattice and consequently the state sum, or as we have called it, the macroscopic partition function, is given by

$$Z = \sum_{\text{all states}} e^{-E\beta} = \sum_{\mu_i = \pm 1} e^{K\sum_{\langle i,j\rangle}\mu_i\mu_j + C\sum_i \mu_i} \qquad (2,2)$$

where $K = \tfrac{1}{2}\mathcal{J}\beta \quad ; \quad C = mH\beta$

As we have mentioned in chapter I, once we have found Z, the problem is virtually solved.

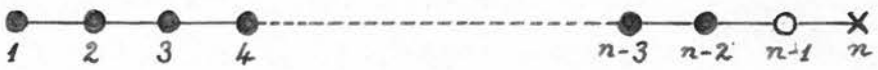


Figure 4.

§ 2 As a simple introduction to the method of obtaining Z, we shall consider the equivalent one dimensional problem of the linear chain of spins. We cannot refer to this as a one dimensional ferromagnet since, as Ising himself showed, it has no ferromagnetic properties.

Figure 4. shows the linear chain of spins - finite in length, with n components $\mu_1, \mu_2 \cdots \mu_n$. From Boltzmann's theorem, the probability of finding a particular arrangement of the spins $\mu_1, \mu_2 \cdots \mu_{n-1}$ is proportional to the Boltzmann factor of the energy associated with that arrangement, since all arrangements have the same weight. Using (2,1), we see that this probability is

$$P(\mu_1 \cdots \mu_{n-1}) = P_{n-1} e^{K(\mu_1\mu_2 + \cdots + \mu_{n-2}\mu_{n-1}) + c(\mu_1 + \cdots + \mu_{n-1})} \qquad (2,3)$$

p_{n-1} is a proportionality factor or normalising constant. Adding on the n^{th} spin, we can correspondingly write for the probability that the spins $\mu_1, \mu_2 \cdots \mu_n$ have a certain arrangement

$$P(\mu_1 \cdots \mu_{n-1}, \mu_n) = P_n e^{K(\mu_1\mu_2 + \cdots + \mu_{n-1}\mu_n) + c(\mu_1 + \cdots + \mu_n)} \qquad (2,4)$$

Now, for the probability that μ_{n-1} has a particular value irrespective of the values of the preceding n-2 spins, we sum (2,3) over all values of the first n-2 spins, i.e.

$$P(\mu_{n-1}) = \sum_{\mu_1=\pm1} \sum_{\mu_2=\pm1} \cdots \sum_{\mu_{n-2}=\pm1} p_{n-1}\, e^{K\sum_{\langle i,j\rangle}^{n-1} \mu_i\mu_j + C\sum_{i}^{n-1}\mu_i} \qquad (2,5)$$

Summing in the same way over (2,4), we find the probability that μ_{n-1} and μ_n have a particular pair of values irrespective of the values of the others, i.e.

$$P(\mu_{n-1},\mu_n) = \sum_{\mu_1=\pm1} \sum_{\mu_2=\pm1} \cdots \sum_{\mu_{n-2}=\pm1} p_n\, e^{K\sum_{\langle i,j\rangle}^{n} \mu_i\mu_j + C\sum_{i}^{n}\mu_i} \qquad (2,6)$$

Taking the ratio of (2,6) and (2,5), we get

$$P(\mu_{n-1},\mu_n) = \frac{p_n}{p_{n-1}} P(\mu_{n-1})\, e^{K\mu_{n-1}\mu_n + C\mu_n} \qquad (2,7)$$

Now if we put $p_{n-1}/p_n = \lambda$ and sum both sides of the resulting expression over the two values of μ_{n-1}, we get the probability that μ_n has a particular value irrespective of the preceding n-1 spins, i.e.

$$\lambda\, P(\mu_n) = \sum_{\mu_{n-1}=\pm1} P(\mu_{n-1})\, e^{K\mu_{n-1}\mu_n + C\mu_n} \qquad (2,8)$$

We assume at this point, that the chain is very long, in which case $P(\mu_n)$ and $P(\mu_{n-1})$ must be the same function of their arguments. In other words, we may write equation (2,8) in the form of a matrix eigenvalue problem. First

of all however, we symmetrise the matrix using the substitution:

$$a(\mu) = P(\mu)\, e^{-c/2\,\mu}$$

where for simplicity, we denote μ_n by μ and μ_{n-1} by μ'. We have then

$$\lambda\, a(\mu) = \sum_{\mu'=\pm 1} a(\mu')\, e^{K\mu'\mu + c/2\,(\mu+\mu')} \qquad (2,9)$$

or $\lambda\, a(\mu) = \mathcal{H}(\mu,\mu')\, a(\mu')$

where $\mathcal{H}(\mu,\mu')$ is the matrix whose elements are the exponentials $\exp[K\mu\mu' + c/2\,(\mu+\mu')]$

Before going further, we must investigate the significance of the latent roots or eigenvalues λ. Since $\mathcal{H}(\mu,\mu')$ is a second order matrix, we shall have two eigenvalues λ_1 and λ_2 and corresponding to these, two eigenvectors $a_1(\mu)$ and $a_2(\mu)$. If now we assume that our eigenvectors are orthonormal, then

$$\sum_{\mu=\pm 1} a_i(\mu)\, a_k(\mu) = \delta_{ik} \qquad (2,10)$$

where $\delta_{i,k}$ is Kronecker's delta. From this it is easily shown that for the elements of $\mathcal{H}(\mu,\mu')$, we may write

$$\mathcal{H}(\mu_1,\mu_2) = \lambda_1 a_1(\mu_1) a_1(\mu_2) + \lambda_2 a_2(\mu_1) a_2(\mu_2) \qquad (2,11)$$

Using this and the relationship (2,10), we can prove

$$\sum_{\mu_2=\pm 1} \mathcal{H}(\mu_1,\mu_2)\,\mathcal{H}(\mu_2,\mu_3) = \lambda_1^2\, a_1(\mu_1)\, a_1(\mu_3) + \lambda_2^2\, a_2(\mu_1)\, a_2(\mu_3)$$

and

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

etc. until we reach the end of our chain with spin n + 1

$$\sum_{\mu_2,\mu_3,\ldots,\mu_n} \mathcal{H}(\mu_1,\mu_2)\,\mathcal{H}(\mu_2,\mu_3)\cdots\mathcal{H}(\mu_n,\mu_{n+1}) = \lambda_1^n\, a_1(\mu_1)\, a_1(\mu_{n+1}) + \lambda_2^n\, a_2(\mu_1)\, a_2(\mu_{n+1})$$

We now close the ring of spins by assuming $\mu_{n+1} = \mu_1$, and sum over this spin. If we also use the relationship (2,10) we get

$$\sum_{\mu_2,\mu_3,\ldots,\mu_n} \mathcal{H}(\mu_1,\mu_2)\,\mathcal{H}(\mu_2,\mu_3)\cdots\mathcal{H}(\mu_n,\mu_1) = \lambda_1^n + \lambda_2^n \qquad (2,12)$$

From the definition of $\mathcal{H}(\mu,\mu')$, we see that the left hand side of this equation is the macroscopic partition function of our one dimensional lattice.

We have already assumed that our chain is very long, i.e. n is a large number. If this is so, and if λ_2 is the smaller of our two eigenvalues, we may write, from (2,12) and (2,2),

$$Z = \lambda_1^n \qquad (2,13)$$

This is the fundamental equation of the variational method - showing as it does, the relationship between the physically important quantity Z and the largest

eigenvalue of our matrix equation. We see that this largest eigenvalue is the partition function per spin of our linear lattice.

Returning to equation (2,9), we see that we have to solve

$$\begin{vmatrix} e^{K+C} - \lambda & e^{-K} \\ e^{-K} & e^{K-C} - \lambda \end{vmatrix} = 0$$

for our eigenvalues.

Doing this, we find for the largest eigenvalue

$$\lambda = e^{K} \cosh C + \left(e^{2K} \sinh^2 C + e^{-2K}\right)^{1/2} \qquad (2,14)$$

where we have dropped the subscript on λ.

For the magnetisation M of the lattice, we have

$$M = m \frac{\partial \ln Z}{\partial C} = mN \frac{\partial \ln \lambda}{\partial C} = mN \frac{\sinh C}{\left[\sinh^2 C + e^{-4K}\right]^{1/2}} \qquad (2,15)$$

so that when H = C = 0, we have M = 0. Since on removing the magnetic field, no magnetisation is left, we see that our linear chain is not ferromagnetic. This is not really surprising, since having one spin wrongly orientated in the chain will completely upset the state of order in the chain, a fact which is not true of two and three dimensional models.

If there is no magnetic field present, we get

$$\lambda = 2 \cosh K \qquad (2,16)$$

and for the configurational energy and specific heat we find

$$E = -\tfrac{1}{2}\mathcal{J}\frac{\partial \ln Z}{\partial K} = -\tfrac{1}{2}N\mathcal{J}\frac{\partial \ln \lambda}{\partial K} = -\tfrac{1}{2}N\mathcal{J}\tanh K \qquad (2,17)$$

$$C_v = \frac{\partial E}{\partial T} = NK^2 k \frac{\partial^2 \ln \lambda}{\partial K^2} = kNK^2 \operatorname{Sech}^2 K \qquad (2,18)$$

From which we again see that the linear chain does not have ferromagnetic properties - since both E and C_v are smooth functions of temperature showing no Curie point.

§ 3. The one dimensional case of our problem which we have just treated, is rather trivial; its value, so far as we are concerned, lies in that it gives a good example of the method of setting up the matrix equation for this type of problem. In the case of the linear chain, the question of finding the largest eigenvalue of our matrix was very easily resolved; however in this respect, the next case, that of the two dimensional ferromagnet, is fundamentally different.

Since it is only in the two dimensional case that the variational method is introduced, we shall consider this case in some detail, particularly as it contains a formal proof of the validity of what is probably the basic

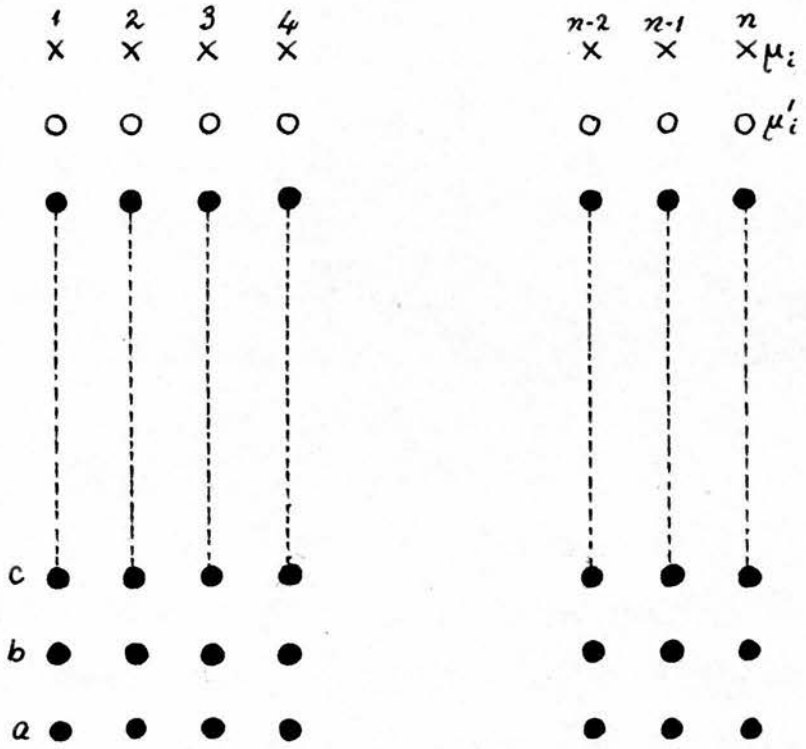


Figure 5.

assumption of the variational method as such. Henceforth, in proceeding in the next chapter to a consideration of the three dimensional analogue of the cases considered in the present chapter, we shall take this assumption for granted. The assumption lies in our choice of eigenvectors.

In the two dimensional case then, our first object is to set up the matrix problem or equation. From a consideration of the linear chain problem, we see that we managed to effect a reduction of the problem in effect, by assuming the chain to be built up by adding one spin at a time beyond the last one placed. We made the chain very long and in working out the probability for the state of the last spin added, succeeded in expressing this in terms of the probability for the state of the second last spin only. Since the chain was very long, we could assume that our probability functions were the same function with different arguments, so that consequently equation (2,8) took on the form of a matrix equation.

Applying this procedure to the square net of spins (see figure 5.) we see that we may build up the net by adding one row of spins at a time to the existing rows.

We consider first the probability that the spins in row μ_i' have a certain set of values irrespective of the preceding n-2 rows and with row μ_i absent. This is given

by an expression of the form of (2,5), where the summations on the right hand side of the equation refer to summations over all possible arrangements of the spins in a particular row - row a, row b, etc. The summation $\sum_{\langle i,j\rangle}$ in the exponential factor of (2,5) now includes contributions in the vertical as well as the horizontal directions. We can then add row μ_i and form an expression similar to (2,6) giving the probability of getting a particular arrangement of the spins in the rows μ'_i and μ_i irrespective of the others.

Taking the ratio of these two expressions, we get the two dimensional analogue of (2,7), where the exponential factor is now

$$\exp\left[K\sum_{\langle i,j\rangle}^{n}\mu_i\mu_j + K\sum_{\langle i,j\rangle}^{n}\mu_i\mu'_i + c\sum_{i}^{n}\mu_i\right] \qquad (2,19)$$

In which μ_i refers to the n spins in the top row only and μ'_i to those in the second top row only. We see that in this equation, the first and third summations are over those spins in the top row while the second is is over the nearest neighbours in both rows.

Eventually we shall get the equation

$$\sum_{\mu'_i} P(\mu_i) = \sum_{\mu'_i} P(\mu'_i)\, e^{K\sum_{\langle i,j\rangle}^{n}\mu_i\mu_j + K\sum_{\langle i,j\rangle}^{n}\mu_i\mu'_i + c\sum_{i}^{n}\mu_i} \qquad (2,20)$$

where ς comes in for the same reason that λ did previously. Once again we symmetrise the equation, using this time

$$a(\mu_i) = P(\mu_i)\, e^{-\left[\kappa/2 \sum_{\langle i,j \rangle}^{n} \mu_i\mu_j + c/2 \sum_{i}^{n} \mu_i\right]} \tag{2,21}$$

giving ultimately

$$\varsigma a(\mu_i) = \sum_{\mu_i'} \mathcal{H}(\mu_i,\mu_i')\, a(\mu_i') \tag{2,22a}$$

where

$$\mathcal{H}(\mu_i,\mu_i') = e^{\kappa \sum_{i=1}^{n} \mu_i\mu_i' + \frac{\kappa}{2} \sum_{i=1}^{n-1} (\mu_i\mu_{i+1} + \mu_i'\mu_{i+1}') + \frac{c}{2} \sum_{i=1}^{n} (\mu_i + \mu_i')} \tag{2,22b}$$

The subscript i in the summations is of course assumed to run along the rows.

Having found our matrix equation, we now want to discover the meaning of the latent roots ς - of which there will be 2^n.

We again assume our eigenvectors to be orthonormal

$$\sum_{\mu_i} a_p(\mu_i) a_q(\mu_i) = \delta_{pq} \tag{2,23}$$

and at the same time use the expansion

$$\mathcal{H}(\mu_i,\mu_i') = \sum_{p=1}^{2n} \varsigma_p\, a_p(\mu_i)\, a_p(\mu_i') \tag{2,24}$$

We can assume that our spin system forms a ring of spins

n spins wide and m in circumference and we find that its partition function may be written as

$$Z = \sum_{\mu_i^k} \mathcal{H}(\mu_i^1, \mu_i^2)\,\mathcal{H}(\mu_i^2, \mu_i^3) \cdots \mathcal{H}(\mu_i^{m-1}, \mu_i^m)\,\mathcal{H}(\mu_i^m, \mu_i^1) \qquad (2,25)$$

μ_i^k referring to the spins in the k^{th} row.

If we put the condition (2,24) into (2,25) and use (2,25) we find

$$Z = \sum_{p=1}^{2^n} \rho_p^m \doteq \rho^m \qquad (2,26)$$

if as before, we let m become very large while keeping n fixed. By ρ we mean the largest eigenvalue of our matrix.

Since in each row or ring of spins we have n components we may write

$$\rho = \lambda^n \qquad \text{or}$$

$$Z = \lambda^{mn} \qquad (2,27)$$

where once again λ is the partition function per spin of our model.

Our next object is to try to effect a solution of equation (2,22) - that is, find an expression for λ . Unfortunately, as we shall see, this problem cannot be solved without making some sort of approximation.

§ 4. In the second part of the original paper by Kramers and Wannier, the authors note that the largest eigenvalue of a matrix may always be expressed in a variational form. This is in fact an adaptation of the Ritz variational principle which occurs in quantum mechanics. From our matrix equation, we can write

$$\rho = \lambda^n = \max_{\text{all } a(\mu_i)} \frac{\sum_{\mu_i \mu_i'} \mathcal{H}(\mu_i, \mu_i') a(\mu_i) a(\mu_i')}{\sum_{\mu_i} [a(\mu_i)]^2} \qquad (2,28)$$

Our object being to find $a(\mu_i)$'s which maximise this equation.

Up to this point, our method is exact. However we now introduce restrictions upon the a's which allow us to obtain only an approximation to ρ. In $\mathcal{H}(\mu_i, \mu_i')$ the μ_i's occur in the combinations

$$\begin{aligned} \sum_{i=1}^{n} \mu_i &= nm \\ \sum_{i=1}^{n} \mu_i \mu_{i+1} &= nq \end{aligned} \qquad (2,29)$$

Our first restriction on the $a(\mu_i)$ therefore lies in assuming

$$a(\mu_i) = a(q, m) \qquad (2,30)$$

Kramers and Wannier then prove - and we shall consider this proof shortly - that (2,30) is equivalent to

$$a(\mu_i) = e^{n[H(K,C)q + A(K,C)m]} \qquad (2,31)$$

where H and A are as yet undefined except that they depend upon temperature and field.

Since we are going to assume this form for our eigenvectors in later work, we shall give the proof of the equivalence of equations (2,31) and (2,30).

The first step is to introduce into equation (2,28), assumption (2,30). Let

$$e^{ng(q,m)}\, dq \cdot dm \qquad (2,32)$$

represent the number of arrangements of the μ's contained in a rectangle dq.dm drawn about (q,m), and let

$$\sum_{\substack{\mu_i,\mu_i' \\ q,m,q',m'}} e^{nK\sum_i \mu_i\mu_i'} = e^{nf(q,m,q',m')}\, dq \cdot dm \cdot dq' \cdot dm' \qquad (2,33)$$

where the summation on the left hand side is extended only over the combinations lying within the volume dq.dm.dq'.dm' about the point q,m,q',m'.

Substituting into (2,28), we get

$$\lambda'' = \max \frac{\iiiint e^{n\left[f(q,m,q',m') + \frac{K}{2}(q+q') + \frac{C}{2}(m+m')\right]} a(q,m)\, a(q',m')\, dq.dm.dq'.dm'}{\iint e^{ng(q,m)}\, a^2(q,m)\, dq.dm.}$$

Now assume $a(q,m) = b(q,m)\, e^{-\frac{1}{2} n g(q,m)}$ giving

$$\lambda^n = \max \frac{\iiiint \exp n \left[f(q,m,q',m') - \frac{1}{2} g(m,q) - \frac{1}{2} g(m',q') + \frac{1}{2} K(q+q') + \frac{1}{2} c(m+m') \right] b(q,m)\, b(q',m')\, dq.dm.dq'.dm'}{\iint b^2(q,m)\, dq.dm.}$$

If we assume that the denominator of this expression is normalised to unity, we can look for the maximum of the numerator. We see that we get our largest ρ if we let b be large wherever the exponent reaches its maximum. Evaluating the integral by the saddle point method, b will make no contribution due to the normalisation we have assumed. We find

$$\ln \lambda = \max_{q,m,q',m'} \left[f(q,m,q',m') - \frac{1}{2} g(q',m') - \frac{1}{2} g(q,m) + \frac{K}{2}(q+q') + \frac{c}{2}(m+m') \right] \qquad (2,34)$$

In order to eliminate the unknown functions f and g, we introduce the solution of a two strip problem

$$\chi^n(I, B, I', B') = \sum_{\mu_i \mu_i'} \exp \left[K \sum_{i=1}^{n} \mu_i \mu_i' + I \sum_{i=1}^{n} \mu_i \mu_{i+1} + I' \sum_{i=1}^{n} \mu_i' \mu_{i+1}' + B \sum_{i=1}^{n} \mu_i + B' \sum_{i=1}^{n} \mu_i' \right] \qquad (2,35)$$

Writing this in terms of q and m, we find

$$\chi^n = \iiiint \exp n \left[f(q,m,q',m') + I q + I' q' + B m + B' m' \right] dq.dm.dq'.dm'$$

Again we use the saddle point method for evaluating this and find

$$\ln \chi(I,B,I',B') = \max_{q,m,q',m'} \left[F(q,m,q',m') + Iq + I'q' + Bm + B'm' \right] \qquad (2,36)$$

Similarly, we can introduce two one strip solutions

$$\psi^n(H,A) = \sum_{\mu_i} \exp \left[2H \sum_{i=1}^{n} \mu_i \mu_{i+1} + 2A \sum_{i=1}^{n} \mu_i \right. \qquad (2,37)$$

or

$$\ln \psi(H,A) = \max_{q,m} \left[g(q,m) + 2Hq + 2Am \right] \qquad (2,38)$$

and $\ln \psi(H',A') = \max_{q',m'} \left[g(q',m') + 2Hq' + 2Am' \right]$

Now substitute (2,36) and (2,38) into (2,34)

$$\ln \lambda = \max_{q,m,q',m'} \left[\ln \chi(I,B,I',B') - \tfrac{1}{2} \ln \psi(H,A) - \tfrac{1}{2} \ln \psi(H',A') \right.$$

$$\left. + (\tfrac{1}{2}K - I + H)q + (\tfrac{1}{2}K - I' + H')q' + (\tfrac{1}{2}C - B + A)m + (\tfrac{1}{2}C - B' + A')m' \right] \qquad (2,39)$$

Due to our maximisation conditions, we must take account of three distinct types of relationship among our various quantities. We have (a) maximisation conditions on χ and ψ as indicated in equations (2,36) and (2,38). These allow us to make the functions I, B, ... A', functions of q, m, q', m', or vice versa. They also allow us to take first derivatives of χ and ψ as if q, m, q', m', appearing in (2,36) and (2,38) were constants.

(b) equations of the form

$$\frac{\partial \ln \chi}{\partial I} = \frac{1}{2}\frac{\partial \ln \psi}{\partial H} = q \; ; \; \frac{\partial \ln \chi}{\partial B} = \frac{1}{2}\frac{\partial \ln \psi}{\partial A} = m$$

and two others for q' and m'.

(c) maximisation conditions on λ - again four in number one each for q, q', m, m', - of the form (see equation (2,39))

$$\frac{\partial \ln \lambda}{\partial \theta} = \frac{\partial \ln \chi}{\partial I}\frac{\partial I}{\partial \theta} + \frac{\partial \ln \chi}{\partial B}\frac{\partial B}{\partial \theta} + \frac{\partial \ln \chi}{\partial I'}\frac{\partial I'}{\partial \theta} + \frac{\partial \ln \chi}{\partial B'}\frac{\partial B'}{\partial \theta}$$

$$- \frac{1}{2}\frac{\partial \ln \psi}{\partial H}\frac{\partial H}{\partial \theta} - \frac{1}{2}\frac{\partial \ln \psi}{\partial A}\frac{\partial A}{\partial \theta} - q\left(\frac{\partial I}{\partial \theta} - \frac{\partial H}{\partial \theta}\right)$$

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

It is easily verified that these equations may be simplified if we use the conditions listed under (b). We then get for the conditions (c)

$$I = 1/2K + H \qquad B = 1/2C + A$$

$$I' = 1/2K + H' \qquad B' = 1/2C + A'$$

We can now use these relations to simplify (2,39). We see that

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

and the relations under (b) read

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

This indicates that H, H′, A, A′, must be chosen so as to give λ a stationary value. If there are several solutions leading to a number of sets of q, m, q′, m′, our original maximisation condition requires that the largest one be chosen, i.e.

$$\lambda = \max_{H,A,H';A'} \frac{\chi\left[\tfrac{1}{2}K+H,\ \tfrac{1}{2}C+A,\ \tfrac{1}{2}K+H',\ \tfrac{1}{2}C+A'\right]}{\left[\psi(H,A)\cdot\psi(H';A')\right]^{1/2}} \qquad (2,40)$$

Finally, we must show that H = H′ and A = A′. In order to do this, we introduce functions $c(\mu_i)$ and $d(\mu_i)$ which obey (2,31). We substitute into (2,40), the expressions given by (2,35) and (2,37) and find

$$\lambda^n = \max_{c,d} \frac{\sum_{\mu_i\mu_i'} \mathcal{H}(\mu_i,\mu_i')\, c(\mu_i)\, d(\mu_i')}{\left\{\left[\sum_{\mu_i} c^2(\mu_i)\right]\left[\sum_{\mu_i'} d^2(\mu_i')\right]\right\}^{1/2}} \qquad (2,41)$$

Here we have used the definition of $\mathcal{H}$ given by (2,22). We can now prove that if we have a set of c's and d's defining λ as in (2,41), we can find a larger λ by replacing the c's by d's or vice versa. ..... (S)

First of all, we prove that $\sum_{\mu_i\mu_i'} \mathcal{H}(\mu_i,\mu_i')\, a(\mu_i)\, a(\mu_i')$ is positive definite. This we do by induction, since it is obviously true if the a's depend on only one spin μ. If we suppose it to be true for n-1 spins, we can prove it to be so for n spins. $\mathcal{H}$ has the form

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

Since the a's are arbitrary, we can absorb the φ's into them and we see that we must prove only that

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

is positive definite.

To do this, we single out μ_n and μ_n' and denote by $\sum'_{\mu_i\mu_i'}$ summation over the remaining μ's . But μ_n can take on only two values so we can write

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

where α and β do not depend on μ_n . Now we can carry out the summation over μ_n and μ_n' and get

$$F = 4\cosh 2K \sum'_{\mu_i\mu_i'} \exp\left[K\sum_{i=1}^{n-1}\mu_i\mu_i'\right]\alpha(\mu_i)\,\alpha(\mu_i')$$

$$+ 4\sinh 2K \sum'_{\mu_i\mu_i'} \exp\left[K\sum_{i=1}^{n-1}\mu_i\mu_i'\right]\beta(\mu_i)\,\beta(\mu_i')$$

Since both terms on the right hand side are positive, this proves our result.

Finally we apply the Schwartz inequality to equation (2,41) assuming that the c's and d's are normalised. This gives

$$2\sum_{\mu_i\mu_i'} \mathcal{H}(\mu_i,\mu_i')\, c(\mu_i)\, \alpha(\mu_i') \leq \sum_{\mu_i\mu_i'} \mathcal{H}(\mu_i,\mu_i')\, c(\mu_i)\, c(\mu_i') + \sum_{\mu_i\mu_i'} \mathcal{H}(\mu_i,\mu_i')\, \alpha(\mu_i)\, \alpha(\mu_i')$$

Thus proving the statement denoted by (S).

Using this result, we can rewrite (2,40) as

$$\lambda = \max_{H,A} \frac{\chi\left[\tfrac{1}{2}K+H\,,\,\tfrac{1}{2}C+A\right]}{\psi(H,A)} \qquad (2,42)$$

As we have suggested earlier, the transition from (2,28) to (2,42) or from (2,30) to (2,31) is one of the basic assumptions of the method of Kramers and Wannier and we shall frequently use it in the next chapter.

In order to evaluate (2,42), we refer to results obtained earlier in this chapter. The form of ψ is given by equation (2,14), while χ can be obtained as the largest eigenvalue of a matrix problem of the form of (2,22). This matrix will be of the fourth order; but since similar matrices will arise in the next chapter, we shall not write the present one out in its array.

Kramers and Wannier, using the definitions

$$e^K = k : e^C = c : e^{2H} = h : e^{2A} = a \qquad (2,43)$$

find for ψ and χ ,

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

$$\chi^3 - \chi^2\left[h^2 h(ac+a^{-1}c^{-1}) + k^{-1}(kh + k^{-1}h^{-1})\right]$$

$$+ \chi(kh-k^{-1}h^{-1})\left[hk(ac+a^{-1}c^{-1}) + k^2(kh+k^{-1}h^{-1})\right] - k(kh-k^{-1}h^{-1})^3 = 0$$

These are equations (2,44a) and (2,44b).

As in all of this type of work, the calculations were carried out for zero external field, i.e. c = 1. If this is so, then a = 1 is a possible solution, since the first derivatives of χ and ψ, with respect to a, vanish at a = 1. Using this, we find for (2,44b)

$$\left[\chi^2 - \chi(k+k^{-1})(kh+k^{-1}h^{-1}) + (kh-k^{-1}h^{-1})^2\right]\left[\chi - k(kh-k^{-1}h^{-1})\right] = 0 \qquad (2,45)$$

the last solution of which may be discarded as being small.

If we now make the substitution $(kh-k^{-1}h^{-1})/(h+h^{-1}) = s\lambda$ we get from (2,45), after substituting $\lambda = \chi/\psi$ and $\psi = h + h^{-1}$,

$$\lambda = \max_s \frac{2}{(s-\sinh K)^2 + 1 - \sinh^2 K}$$

which maximises to

$$\lambda = \frac{2}{1-\sinh^2 K} \qquad (2,46)$$

This solution, however, is not valid for high values of K - it is only true forthe temperature region above the Curie point. In the region of high K, Kramers and Wannier

using various substitutions manage to find a solution with $a \neq 1$. Without going into details of their calculation we may give their result as

$$\lambda = k^2 + \frac{(1+k^2 z)(1-z^2)}{k^4 [k^2 - k^{-2}]} \qquad (2,47)$$

where $\frac{1}{k^2} - \frac{s^2}{s(k-s)} = \frac{\lambda - k^2}{z}$ defines z

and $(kh - h^{-1}h^{-1})\psi^{-1} = s\lambda$ defines s as before.
The two parts of λ given by (2,46) and (2,47) make up the complete solution which is of the form shown in figure 2. of the last chapter.

Again without going into details, we may mention that Kramers and Wannier, in part 8 of their paper, give series expansions for λ in powers of K for (2,46) and in powers of k^{-1} for (2,47). They then supply the corresponding series as calculated from various other approximate methods and demonstrate that in both the high and low temperature regions, the variational method gives a far better solution than is given by any of the other methods.

It should be noted, in connection with the two dimensional ferromagnet on this treatment, that Kramers and Wannier succeed in finding the solutions for λ in closed forms. For the three dimensional cases, we have not been able to do this in view of the extreme complexity of our equations.

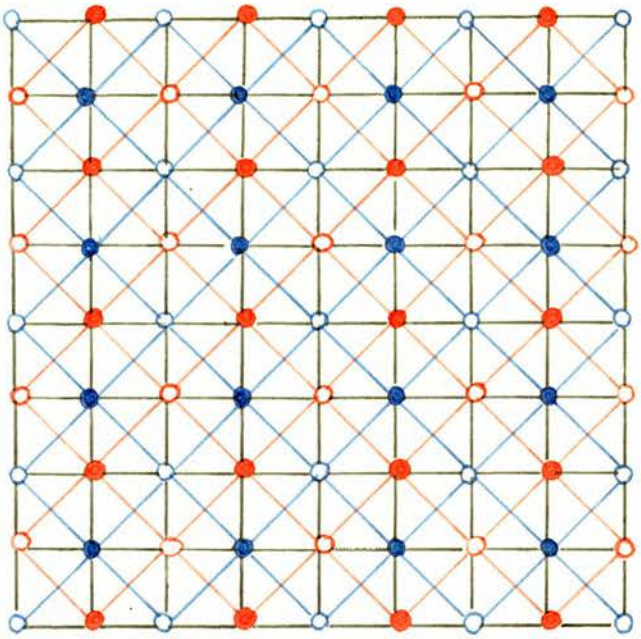
III

THE VARIATIONAL METHOD

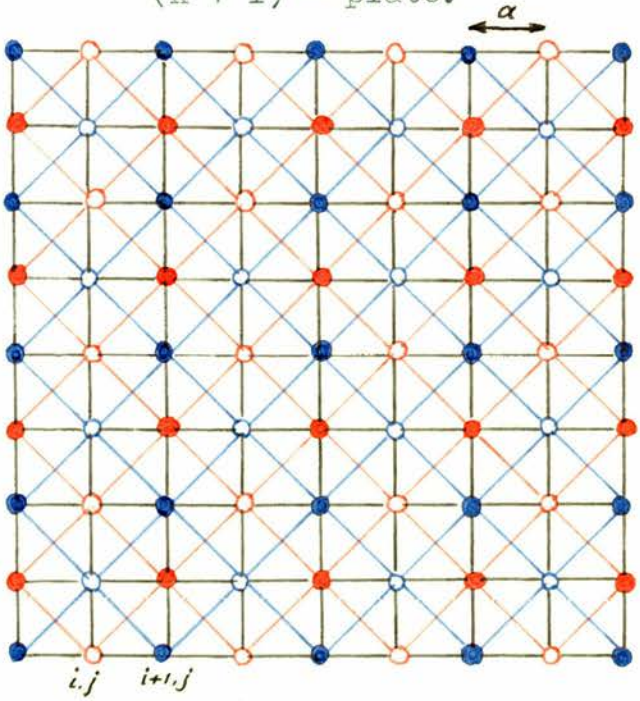
APPLIED TO THREE DIMENSIONAL STRUCTURES

§1 In this chapter, we shall consider the application of the variational method developed in the last chapter, to the problem of cubic lattice structures. Martin and ter Haar have already discussed the simple cubic case at high temperatures; we, however, shall develop a general approach which will be capable of modification to yield results for any one of the three cubic cases. Furthermore, we shall extend our calculations to include low temperature series for our significant quantities - in particular for the partition function per spin.

As a preliminary, we attempted to apply the variational method directly to the problem of finding the partition function per spin of the face centred cubic lattice. If it had been successful, this approach would have given the best approximation to the required partition function that the variational method could yield; we soon found, however, that the method led to a matrix which was far too large to handle. (We shall give a short discussion of this direct approach, in an appendix.) In consequence of this, we turned to an alternative approach which has the advantage, mentioned above, that in order to tackle any of the



$(k + 1)^{th}$ plate.



k^{th} plate.

Figure 6.

three three dimensional problems - simple cubic, face centred cubic, or body centred cubic - we merely had to modify our main equations, provided we had made the approach sufficiently general to include all three cases.

Before describing the method, we should first of all repeat that in this work we are concerned with finding the partition functions of the various ferromagnetic lattice structures where only nearest neighbour interactions between the atoms are taken into account. The extension to include second or third nearest neighbour interactions will not be considered for solution.

§ 2 In order to set up our general case, we consider figure 6. This depicts two adjacent plates of the three dimensional simple cubic lattice with first, second, and third nearest neighbour interactions. The separation of the plates is the lattice constant 'a'. The black lines indicate nearest neighbour interactions, the blue and the red lines indicate next nearest neighbour interactions, while the third nearest neighbour interactions are not shown since these do not lie in the plane of the paper.

From the sketch, we see that nearest neighbour interactions link all atoms irrespective of whether they are represented as dots or as circles or as being red or blue. Thus the set of all points of intersection of nearest

neighbour interactions represents one system - the simple cubic lattice structure. Also, on displacing one of the plates a distance 'a' in the i or j direction, we see that next nearest neighbour interactions link only all red atoms or all blue atoms, irrespective of whether they are represented as dots or as circles, so that, on making the displacement, the set of all red atoms - intersections of red next nearest neighbour interaction lines - represents one system, and the set of intersections of all blue next nearest neighbour interaction lines represents another (equivalent) system - the face centred cubic lattice structure. Finally, considering only red circles and the corresponding third order interaction lines, we see that the red circles represent one of four independent but equivalent systems - the body centred cubic lattice structure. The other systems are represented by red dots, blue circles and blue dots.

From these considerations, it is easy to see that this particular model may eventually be modified in various ways in order to discuss different problems. These are:

1) Simple cubic lattice with first, second, and third order interactions.

2) Simple cubic lattice with first and second order interactions - by putting the third order interaction energy equal to zero.

3) Simple cubic lattice with first order interaction only - by putting second and third order interaction energies equal to zero. This is the case considered by Martin and ter Haar at high temperatures.

4) Face centred cubic lattice with first order interactions only - by putting first and third order interaction energies equal to zero.

5) Body centred cubic lattice with first order interactions only - by putting first and second order interaction energies equal to zero.

Of these various problems, we shall discuss only the last two in detail. The third problem, we shall consider only to the extent of demonstrating that we can obtain the equations of Martin and ter Haar from our general equations. We have not attempted to solve the simple cubic lattice problem at low temperatures although this would have been of considerable interest, especially as the method we are going to use is, in the problems we shall treat, an approximation to the variational method. By this, we mean that that our application of the variational method does not allow the method its full scope so far as problems 4) and 5) are concerned. We might say that it does not do full justice to the power of the method. This is of course bound up with the fact that we found that we could not apply the variational method directly to these two problems.

(The difficulty discussed in the appendix for the face centred lattice exists also for the body centred lattice as is mentioned there.) In this respect, however, the simple cubic lattice problem is different from the other two since our method, suitably modified in the later stages, represents a direct approach to the simple cubic lattice problem. Obviously, then, it would be of great interest to effect a solution of this case at low temperatures.

Our reason for leaving the simple cubic case uncompleted is twofold. First of all, there was the practical difficulty of the magnitude of the task. As we shall see, although the simple cubic case is the simplest of the cubic structures, the basic variational method equations appropriate to it on our treatment are not only more numerous, but are much more complicated than are those for the other cases. This can be inferred from the results of the next chapter, where we shall see that the face centred case yields equations which, although they give a much better approximation to λ than do those for the body centred case, are very much longer (in the low temperature region especially) than are the corresponding body centred case equations. This is not to imply that the equations cannot be solved; but merely that once the equations have been written out, which in itself will be a tedious task requiring considerable checking, a great deal of work

will be necessary to effect a solution, especially since it will be necessary to work out a large number of terms for λ before deviation from the exact series is observed.

The other reason for neglecting this case lies in the fact that of the three cubic cases, the simple cubic structure is the least important from the point of view of occurrence in nature.

In developing the theory of our present method, we shall take it as far as possible while it is completely general i.e. we shall develop the theory for the simple cubic lattice with first, second, and third order interactions and only in turning to the special cases we require, shall we modify our equations by putting particular interactions equal to zero.

§ 3 As in previous work, we take for our model, that of a crystal structure with spins situated on the lattice sites according to figure 6. All spins are equivalent. We consider each spin to be capable of taking up one or other of two orientations indicated by giving one or other of the values ± 1 to a spin parameter denoted by $\mu_{i,j,k}$. Our lattice is built up from plates of the form of those in figure 6, the separation of the plates being the lattice parameter 'a'. We assume that we have m spins in the x direction, n in the y direction, and p in the z direction

the plates lying in ~~the~~ xy planes and the lattice being built up in the z direction. At absolute zero, all the spins will be ordered relative to each other, i.e. all nearest neighbours will be parallel to each other.

Our first aim is to find equations leading to the partition function per spin with all three orders of interaction, for the simple cubic structure. From Boltzmann's theorem, the probability of getting a certain state or configuration of the system is proportional to $e^{-E/\beta}$ where, as before, E is the energy of the state. The partition function Z is defined to be $\sum_{\text{states } j} e^{-E_j/\beta}$ as we have seen earlier. This summation in the present case contains 2^{mnp} terms.

For the energy of such a ferromagnetic crystal, in a magnetic field H, we have

$$E = -\tfrac{1}{2}\mathcal{J}\sum\mu\mu' - \tfrac{1}{2}\mathcal{J}'\sum\mu\mu'' - \tfrac{1}{2}\mathcal{J}''\sum\mu\mu''' - mH\sum\mu \qquad (3,1)$$

where $\mathcal{J}$ is the co-operational energy of nearest neighbouring spins in the lattice, i.e. it is the energy gained when two isolated nearest neighbours change from the parallel to the antiparallel alignment, and $\mathcal{J}'$ and $\mathcal{J}''$ are the co-operational energies of next and third nearest neighbouring spins respectively. m is the magnetic moment per spin. The first summation is over all nearest neighbours in the lattice, the second is over all next nearest

neighbours, the third is over all third nearest neighbours, and the fourth is over all spins.

From Boltzmann's theorem, considering the lattice to have p-1 plates, the probability of getting a certain arrangement of the spins in these plates (i.e. the probability of getting a certain state of the system) is given by

$$P(\mu_1,\mu_2 \cdots \mu_{p-1}) = \psi_{p-1} \exp\Big[K \sum_{}^{p-1} \mu_{rst}\mu_{r+1,s,t} + K' \sum_{}^{p-1} (\mu_{rst}\mu_{r+1,s+1,t} + \mu_{r+1,s,t}\mu_{r,s+1,t}) + K'' \sum_{}^{p-1} (\mu_{rst}\mu_{r+1,s+1,t+1} + \mu_{r+1,s+1,r}\mu_{r,s,t+1}) + c \sum_{}^{p-1} \mu_{rst} \Big] \quad (3,2)$$

In this equation, $\mu_1, \mu_2 \cdots \mu_{p-1}$ represent all spins in plates 1, 2, ... p-1 respectively and take account of all spins up to and including those in the $(p-1)^{th}$ plate. In expressions like $\mu_{r+1,s,t}$, the order of r, s, and t is of no significance, since the expression is meant to mean $\mu_{x+1,y,z} + \mu_{x,y+1,z} + \mu_{x,y,z+1}$. Of course, having decided upon a particular meaning for r, s, and t, for the first factor of a product of μ's, this meaning applies to the second factor also. The notation is merely meant to cover every possibility for a particular order of interaction. Thus, under the first summation, we have three general terms - as r + 1 runs through the values x + 1, y + 1, and z + 1. Under the second summation we have six general terms and under the third we have four general terms. $K = J/2kT = \frac{1}{2}\beta J : K' = \frac{1}{2}J'\beta : K'' = \frac{1}{2}J''\beta : c = m\beta H.$

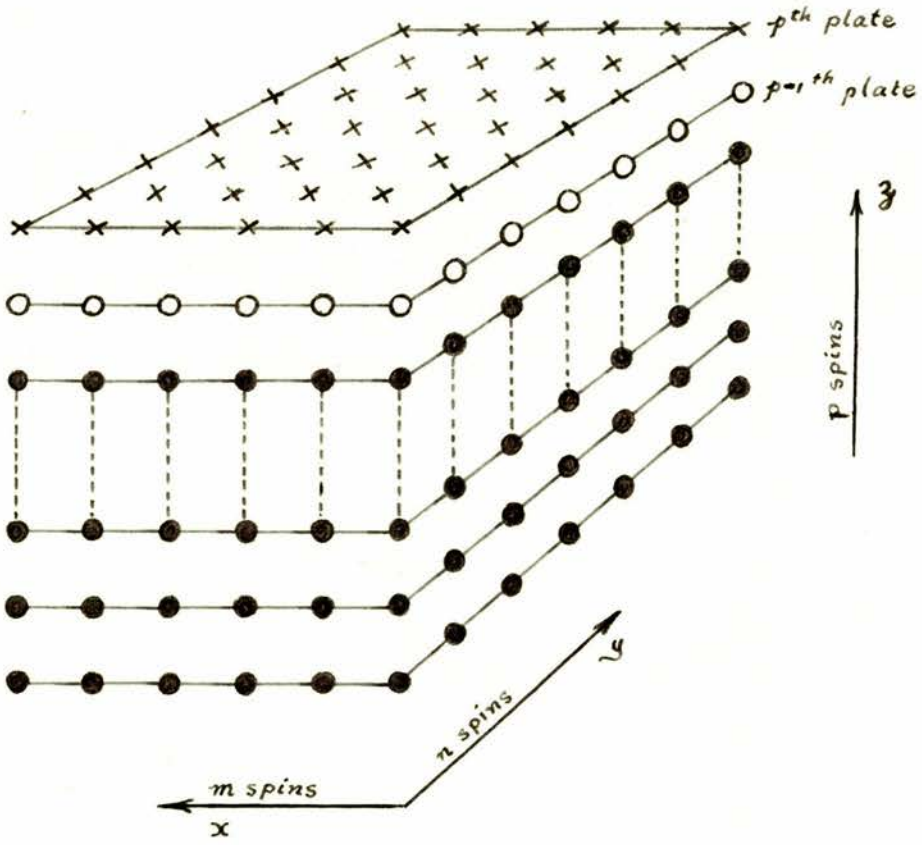


Figure 7.

p_{p-1} is a proportionality factor.

Following the usual procedure, we now add the p^{th} plate to our array of spins and enquire as to the probability of this new system having a particular arrangement of its spins. This process of adding successive plates is illustrated in figure 7. The new probability is

$$P(\mu_1 \cdots \mu_{p-1}, \mu_p) = p_p \exp\Big[K \sum_{r,s,t}^{p} \mu_{rst}\mu_{r+1,st} + K' \sum_{r,s,t}^{p} (\mu_{rst}\mu_{r+1,s+1,t} + \mu_{r+1,st}\mu_{r,s+1,t}) + K'' \sum_{r,s,t}^{p} (\mu_{rst}\mu_{r+1,s+1,t+1} + \mu_{r+1,s+1,t}\mu_{r,s,t+1}) + C \sum_{r,s,t}^{p} \mu_{rst}\Big] \qquad (3,3)$$

Now the probability of getting a certain arrangement of the spins in the $(p-1)^{th}$ plate irrespective of the arrangement of the spins in the previous p-2 plates is obtained by summing (3,2) over all possible arrangements of the spins in the previous plates i.e.

$$P(\mu_{p-1}) = \sum_{\Sigma\mu_1=\pm1} \sum_{\Sigma\mu_2=\pm1} \cdots \sum_{\Sigma\mu_{p-2}=\pm1} P(\mu_1, \mu_2 \cdots \mu_{p-1})$$

$$= \sum_{\Sigma\mu_1=\pm1} \sum_{\Sigma\mu_2=\pm1} \cdots \sum_{\Sigma\mu_{p-2}=\pm1} p_{p-1} \exp\Big[K \sum_{}^{p-1} \mathcal{A} + K' \sum_{}^{p-1} \mathcal{B} + K'' \sum_{}^{p-1} \mathcal{L} + C \sum_{}^{p-1} \mathcal{U}\Big] \qquad (3,4)$$

where $\mathcal{A}, \mathcal{B}, \mathcal{L}, \mathcal{U}$ represent the various sums of products of μ's occurring in the exponential factors of such equations as (3,3).

Similarly, the probability of a certain arrangement of spins in the p^{th} and $(p-1)^{th}$ plates taken together and irrespective of the spins in the previous plates is

obtained by summing (3,3) in the same way

$$P(\mu_{p-1},\mu_p) = \sum_{\Sigma\mu_1=\pm1}\sum_{\Sigma\mu_2=\pm1}\cdots\sum_{\Sigma\mu_{p-2}=\pm1} P(\mu,\mu_2\cdots\mu_{p-1},\mu_p)$$

$$= \sum_{\Sigma\mu_1=\pm1}\sum_{\Sigma\mu_2=\pm1}\cdots\sum_{\Sigma\mu_{p-2}=\pm1} p_p \exp\Big[K\sum_{1}^{p}\alpha + K'\sum_{1}^{p}\mathcal{G} + K''\sum_{1}^{p}\mathcal{L} + C\sum_{1}^{p}\mathcal{V}\Big] \qquad (3,5)$$

If we next divide (3,5) by (3,4) and put $\varsigma = p_{p-1}/p_p$, we get

$$\varsigma P(\mu_{p-1},\mu_p) = P(\mu_{p-1}) \exp\Big[K\sum_{p-1}^{p}\alpha + K'\sum_{p-1}^{p}\mathcal{G} + K''\sum_{p-1}^{p}\mathcal{L} + C\sum_{p-1}^{p}\mathcal{V}\Big] \qquad (3,6)$$

We now denote the p^{th} plate by μ' ($\mu' \equiv \mu_p$) and the $(p-1)^{th}$ plate by μ to give

$$\varsigma P(\mu,\mu') = P(\mu) \exp\Big[K\sum_{p-1}^{p}\alpha + K'\sum_{p-1}^{p}\mathcal{G} + K''\sum_{p-1}^{p}\mathcal{L} + C\sum_{p-1}^{p}\mathcal{V}\Big] \qquad (3,7)$$

In order to find the probability for a certain arrangement of the spins in the p^{th} plate only, we can sum (3,7) over all spin arrangements possible in the $(p-1)^{th}$ plate

$$\varsigma P(\mu') = \varsigma\sum_{\Sigma\mu_{p-1}=\pm1} P(\mu,\mu') = \sum_{\Sigma\mu=\pm1} P(\mu) \exp\Big[K\sum_{p-1}^{p}\alpha + K'\sum_{p-1}^{p}\mathcal{G} + K''\sum_{p-1}^{p}\mathcal{L} + C\sum_{p-1}^{p}\mathcal{V}\Big] \qquad (3,8)$$

or

$$\varsigma P(\mu') = \sum_{\Sigma\mu=\pm1} P(\mu)\,\mathcal{Y} \qquad (3,9)$$

When we let p - the number of spins in the z direction - go to infinity, $P(\mu')$ and $P(\mu)$ become the same function with different arguments (that is, the addition of the

p^{th} plate has not altered the physical properties of the crystal) and the last equation takes the form of a matrix equation. We see that the ρ will be eigenvalues of the matrix which we have denoted by $\mathcal{H}(\mu,\mu')$.

The next step consists in writing out an exact expression for the matrix $\mathcal{H}(\mu,\mu')$ in terms of the interactions in the top plate and between the two top plates. In order to see how this expression arises, we may refer to figure 6. Doing this, we find

$$\mathcal{H}(\mu,\mu') = \exp\Big[K\sum(\mu'_{ij}\mu'_{i+1,j} + \mu'_{ij}\mu'_{i,j+1} + \mu'_{ij}\mu_{ij})$$
$$+ K'\sum(\mu'_{ij}\mu'_{i+1,j+1} + \mu'_{i+1,j}\mu'_{i,j+1} + \mu'_{ij}\mu_{i,j+1} + \mu'_{ij}\mu_{i+1,j} + \mu_{ij}\mu'_{i+1,j} + \mu_{ij}\mu'_{i,j+1})$$
$$+ K''\sum(\mu'_{ij}\mu_{i+1,j+1} + \mu'_{i+1,j}\mu_{i,j+1} + \mu'_{i+1,j+1}\mu_{ij} + \mu'_{i,j+1}\mu_{i+1,j})$$
$$+ C\sum\mu'_{ij}\Big] \qquad (3,10)$$

In order to symmetrise $\mathcal{H}$ to give, say $\mathcal{H}(\mu,\mu')$, we use the substitution

$$P(\mu) = a(\mu)\exp\Big[\tfrac{K}{2}\sum\overbrace{(\mu_{ij}\mu_{i+1,j} + \mu_{ij}\mu_{i,j+1})}^{\alpha}$$
$$+ \tfrac{K'}{2}\sum\overbrace{(\mu_{ij}\mu_{i+1,j+1} + \mu_{i+1,j}\mu_{i,j+1})}^{\beta} + \tfrac{C}{2}\sum\overbrace{\mu_{ij}}^{\gamma}\Big] \qquad (3,11)$$

This, when substituted into (3,9), gives the matrix with elements of the form

$$\mathcal{H}(\mu,\mu') = \exp\Big[\tfrac{K}{2}\sum(\mu'_{ij}\mu'_{i+1,j} + \mu'_{ij}\mu'_{i,j+1} + \mu_{ij}\mu_{i+1,j} + \mu_{ij}\mu_{i,j+1})$$

$$+ K\sum\mu'_{ij}\mu_{ij} + \tfrac{K'}{2}\sum(\mu'_{ij}\mu'_{i+1,j+1} + \mu'_{i+1,j}\mu'_{i,j+1} + \mu_{ij}\mu_{i+1,j+1} + \mu_{i+1,j}\mu_{i,j+1})$$

$$+ K'\sum(\mu'_{ij}\mu_{i+1,j} + \mu'_{ij}\mu_{i,j+1} + \mu'_{i+1,j}\mu_{ij} + \mu'_{i,j+1}\mu_{ij})$$

$$+ K''\sum(\mu'_{ij}\mu_{i+1,j+1} + \mu'_{i+1,j}\mu_{i,j+1} + \mu'_{i+1,j+1}\mu_{ij} + \mu_{i+1,j}\mu'_{i,j+1})$$

$$+ c/2\sum(\mu'_{ij} + \mu_{ij})\Big] \qquad (3,12)$$

From an extension of the theory of Kramers and Wannier, we can show that the largest eigenvalue ρ of this matrix is equal to the mn^{th} power of the partition function per spin of the lattice. The largest eigenvalue outweighs all the others in importance if we let the height of the structure (p) be very great.

Our theory up to this point has effectively reduced our problem from that involving a crystal of three dimensions to one involving a crystal of essentially two dimensions since we are now considering only two plates, the p^{th} and $(p-1)^{th}$, of our original crystal. If we denote by λ the partition function per spin of the crystal, then we have $\rho_{max} = \lambda^{mn}$ where ρ_{max} is the largest eigenvalue of our matrix.

In order to find ρ_{max}, we invoke the Ritz variational principle

$$\rho_{max} = \lambda^{mn} = \underset{\text{orthog a's}}{\text{maximum}} \frac{\sum \mathcal{H}(\mu,\mu')\, a(\mu)\, a(\mu')}{\sum [a(\mu)]^2} \qquad (3,13)$$

where as indicated, we vary the eigenvectors in order to maximise the expression.

Adapting the results of Kramers and Wannier, we now assume that it is possible to prove that we may choose $a(\mu)$ to be of the form

$$a(\mu) = \exp\left[\frac{A}{2}\sum\alpha + \frac{A'}{2}\sum\beta + \frac{B}{2}\sum\gamma\right] \qquad (3,14)$$

where α, β, and γ are the expressions indicated in equation (3,11) and $A = A(K:K':c) : A' = A'(K:K':c) : B = B(K:K':c)$ This is the assumption the proof of which we have given at length for the two dimensional case in the last chapter.

If we substitute this form of the $a(\mu)$ into equation (3,13) and take the mn^{th} root, we find

$$\lambda = \max_{A,A',B} \frac{\chi\left[\frac{K+A}{2} : \frac{K'+A'}{2} : \frac{c+B}{2}\right]}{\psi\left[A : A' : B\right]} \qquad (3,15)$$

which is approximately equal to the partition function per spin of our three dimensional ferromagnetic lattice structure.

On writing out the expressions for χ^{mn} and ψ^{mn} in full we find that these functions have the form of partition

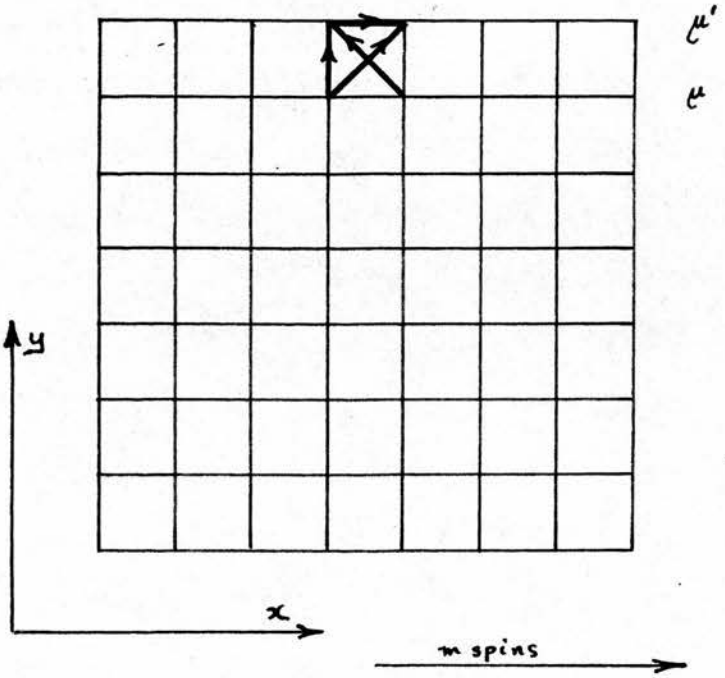


Figure 8.

functions of crystals containing respectively two plates and one plate of spins. In fact, for ψ^{mn} we have

$$\psi^{mn} = \sum \exp\left[A\sum(\mu_{ij}\mu_{i+1,j} + \mu_{ij}\mu_{i,j+1}) + A'\sum(\mu_{ij}\mu_{i+1,j+1} + \mu_{i+1,j}\mu_{i,j+1}) + B\sum\mu_{ij}\right] \quad (3,16)$$

which is a summation over all possible arrangements of the spins in one plate; the first summation inside the square bracket referring to all nearest neighbour interactions in the plate, the second to all next nearest neighbour interactions in the plate, and the third to all spins in the plate.

In the following part of the work, we adopt the notation of figure 8, the dashes here having no connection with those used previously.

Our procedure now is as follows. On consideration of equation (3,16), we see that it has the form of the partition function of a square net of spins - n spins by m spins - and with nearest and next nearest neighbours involved. Starting from the beginning once again, we would consider the probability of getting a certain arrangement of spins taking only the first n-1 strips of spins and then the probability of a certain arrangement taking in all the strips including the n^{th} one. We would then take in each of these probability equations, summations over the spins in the first n-2 strips in order to find the probabilities for particular arrangements of the spins in the top strip

only and of the spins in the two top strips only. Taking the ratio of these expressions and letting the number of strips go to infinity, we would eventually obtain a matrix equation of the form

$$\sigma P(\mu_i') = \sum_{\sum \mu_{m-1} = \pm 1} P(\mu_i)\, \mathcal{Y}(\mu_i, \mu_i') \qquad (3,17)$$

where $\mathcal{Y}$ has the form

$$\exp\left[A \sum (\mu_i \mu_{i+1}' + \mu_i' \mu_i) + A' \sum (\mu_i' \mu_{i+1} + \mu_{i+1}' \mu_i) + B \sum \mu_i'\right] \qquad (3,18)$$

As before, we symmetrise equation (3,18) to give $\mathcal{H}(\mu_i \mu_i')$ using $P(\mu_i) = a(\mu_i) \exp\left[A/2 \sum \mu_i \mu_{i+1} + B/2 \sum \mu_i \right]$ giving

$$\mathcal{H}(\mu_i, \mu_i') = \exp\Big[A \sum \mu_i' \mu_i + \frac{A}{2} \sum (\mu_i' \mu_{i+1}' + \mu_i \mu_{i+1})$$

$$+ A' \sum (\mu_i \mu_{i+1}' + \mu_{i+1} \mu_i') + B/2 \sum (\mu_i + \mu_i') \Big] \qquad (3,19)$$

The largest eigenvalue σ_{max} of this will be the partition function per spin to the power m of the square net of spins to which Ψ^{mn} corresponds.

Again, for the largest eigenvalue σ_{max} of $\mathcal{H}(\mu_i \mu_i')$ we have

$$\sigma_{max} = \underset{\text{all } a(\mu_i)\text{'s}}{\text{maximum}} \frac{\sum \mathcal{H}(\mu_i \mu_i') a(\mu_i) a(\mu_i')}{\sum [a(\mu_i)]^2} \qquad (3,20)$$

This time, for our eigenvectors, we choose the form

$$a(\mu_i) = \exp\left[P/2 \sum \mu_i \mu_{i+1} + Q/2 \sum \mu_i \right] \qquad (3,21)$$

where $P = P(K:c)$, $Q = Q(K:c)$

Putting this into the variational equation, and taking the m^{th} root, we find

$$\psi = \max_{P,Q} \frac{\mathcal{B}\left[\frac{A+P}{2} : \frac{B+Q}{2}\right]}{\mathcal{F}[P:Q]} \qquad (3,22)$$

for the partition function per spin of the subsidiary two dimensional lattice structure.

On writing out $\mathcal{B}^m$ and $\mathcal{F}^m$ in full, we find that these functions have the form of partition functions of crystals containing respectively two strips and one strip of spins. For $\mathcal{B}^m$, for instance, we have

$$\mathcal{B}^m = \sum \exp\Big[A\sum \mu_i\mu_i' + \frac{A+P}{2}\sum(\mu_i'\mu_{i+1}' + \mu_i\mu_{i+1}) + A'\sum(\mu_i\mu_{i+1}' + \mu_{i+1}\mu_i') + \frac{B+Q}{2}\sum(\mu_i + \mu_i')\Big] \qquad (3,23)$$

which is a summation over all possible arrangements or states of the spins in two strips of spins, the first summation within the square brackets referring to all nearest neighbour interactions between the strips, the second to all nearest neighbour interactions in each strip, the third to all next nearest neighbour interactions between the strips (it being impossible to have such interactions in the strips) and the final summation to all spins in the strips.

We now treat this problem on its own merits, as it were, considering first the probability of a certain arrangement of the spins up to but excluding the last pair μ_m and μ'_m i.e. up to and including μ_{m-1} and μ'_{m-1}, and then the probability, taking the last two spins into account. If we follow out the usual procedure, we shall ultimately arrive at the expression

$$\mathcal{H}(\mu_m, \mu'_m) = \exp\left[A(\mu'_m \mu_m) + \frac{A+P}{2}(\mu'_m \mu'_{m-1} + \mu_m \mu_{m-1}) + A'(\mu_m \mu'_{m-1} + \mu'_m \mu_{m-1}) + \frac{B+Q}{2}(\mu_m + \mu'_m) \right] \quad (3,24)$$

This matrix, when written out in full, will be square and of order 4×4 - obtained by varying μ_m, μ'_m, μ_{m-1} and μ'_{m-1} through their values ± 1. The largest eigenvalue of the matrix will be $\mathcal{V}$, the partition function per spin of our crystal of two linear chains. This matrix, in its array, is written out on page 75.

For ζ^m we have $\sum \exp\left[P \sum \mu_i \mu_{i+1} + Q \sum \mu_i \right]$

This has the form of the partition function of a single chain of m spins, the probability treatment for which has already been given by Kramers and Wannier and which has been discussed in chapter 2. We see easily, or from chapter 2 (see equation (2,9)) that we shall find a matrix

$$\mathcal{H}(\mu_m, \mu_{m-1}) = \exp\left[P \mu_m \mu_{m-1} + \frac{Q}{2}(\mu_m + \mu_{m-1}) \right] \quad (3,25)$$

The matrix is of order 2×2 with a largest eigenvalue ϕ which is the partition function per spin of the linear chain.

To summarise this piece of work, we have found that in order to evaluate ψ , the partition function of our derived plate of spins, we must evaluate $\mathcal{Z}$ and ϕ - the partition functions per spin of systems of two chains and of one chain respectively. We have found also that $\mathcal{Z}$ and ϕ can be derived directly from matrices of order 4×4 and 2×2.

$$\psi = \mathcal{Z}/\phi \qquad (3,26)$$

There remains now, the function χ to be dealt with. We first of all write out the expression for χ^{mn} obtainable from equation (3,15).

$$\chi^{mn} = \sum \exp\Big[\frac{K+A}{2} \sum (\mu'_{ij}\mu'_{i+1,j} + \mu'_{ij}\mu'_{i,j+1} + \mu_{ij}\mu_{i+1,j} + \mu_{ij}\mu_{i,j+1})$$

$$+ K \sum \mu'_{ij}\mu_{ij} + \frac{K'+A'}{2} \sum (\mu'_{ij}\mu'_{i+1,j+1} + \mu'_{i+1,j}\mu'_{i,j+1} + \mu_{ij}\mu_{i+1,j+1} + \mu_{i+1,j}\mu_{i,j+1})$$

$$+ K' \sum (\mu'_{ij}\mu_{i+1,j} + \mu'_{ij}\mu_{i,j+1} + \mu_{ij}\mu'_{i+1,j} + \mu_{ij}\mu'_{i,j+1})$$

$$+ K'' \sum (\mu'_{ij}\mu_{i+1,j+1} + \mu'_{i+1,j}\mu_{i,j+1} + \mu'_{i+1,j+1}\mu_{ij} + \mu_{i+1,j}\mu'_{i,j+1})$$

$$+ \frac{C+B}{2} \sum (\mu'_{ij} + \mu_{ij}) \Big] \qquad (3,27)$$

It is easy to see that this has the form of the partition function of a crystal consisting of two plates of m X n

spins each. χ is the partition function per spin of the crystal.

Following out the usual probability procedure for this problem from first principles once again, or (since we can now see how the reduction of the problem at various stages proceeds) from equation (3,27), we can reduce these two plates. We find a matrix of the form

$$\mathcal{H}(\mu,\mu';\mu^*,\mu'^*) = \exp\Big[\frac{K+A}{2}\sum(\mu_i'^*\mu_{i+1}'^* + \mu_i^*\mu_{i+1}^*)$$
$$+ \frac{K+A}{2}\sum(\mu_i'\mu_i'^* + \mu_i\mu_i^*) + K\sum\mu_i^*\mu_i'^* + K'\sum(\mu_i'\mu_i^* + \mu_i'^*\mu_i)$$
$$+ \frac{K'+A'}{2}\sum(\mu_i'\mu_{i+1}'^* + \mu_{i+1}'\mu_i'^* + \mu_i\mu_{i+1}^* + \mu_{i+1}\mu_i^*) + K''\sum(\mu_i'^*\mu_{i+1}^* + \mu_{i+1}'^*\mu_i^*)$$
$$+ K''\sum(\mu_i'\mu_{i+1}^* + \mu_{i+1}'\mu_i^* + \mu_{i+1}'^*\mu_i + \mu_i'^*\mu_{i+1}) + \frac{C+B}{2}\sum(\mu_i'^* + \mu_i^*)\Big]$$

where

$$\Pi P(\mu^*,\mu'^*) = \sum_{\mu_i'\mu_i} P(\mu,\mu')\,\mathcal{H}(\mu,\mu';\mu^*\mu'^*) \qquad (3,28)$$

To symmetrise this matrix, we use the substitution

$$Q(\mu_i,\mu_i') = P(\mu_i,\mu_i')\exp\Big[\frac{K+A}{4}\sum(\mu_i'\mu_{i+1}' + \mu_i\mu_{i+1}) + \frac{K}{2}\sum\mu_i\mu_i'$$
$$+ \frac{K'}{2}\sum(\mu_i'\mu_{i+1} + \mu_{i+1}'\mu_i) + \frac{C+B}{4}\sum(\mu_i' + \mu_i)\Big] \qquad (3,29)$$

and obtain in this way

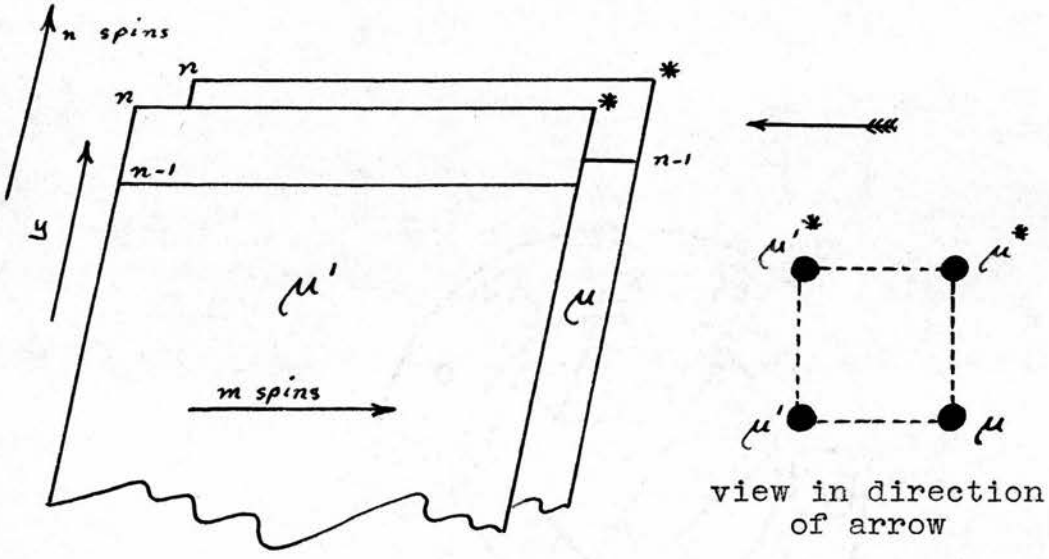


Figure 9.

$$\mathcal{N}(\mu,\mu';\mu^*,\mu^*) = \exp\Big[\frac{K+A}{4}\sum(\mu_i'\mu_{i+1}' + \mu_i\mu_{i+1} + \mu_i'^*\mu_{i+1}'^* + \mu_i^*\mu_{i+1}^*)$$

$$+ \frac{K+A}{2}\sum(\mu_i'\mu_i'^* + \mu_i\mu_i^*) + \frac{K}{2}\sum(\mu_i'\mu_i + \mu_i'^*\mu_i^*)$$

$$+ \frac{K'+A'}{2}\sum(\mu_i'\mu_{i+1}^* + \mu_{i+1}'\mu_i'^* + \mu_i\mu_{i+1}^* + \mu_{i+1}\mu_i^*) + K'\sum(\mu_i'\mu_i'^* + \mu_i'^*\mu_i)$$

$$+ \frac{K'}{2}\sum(\mu_i'\mu_{i+1} + \mu_{i+1}'\mu_i + \mu_{i+1}'^*\mu_i'^* + \mu_i'^*\mu_{i+1}'^*)$$

$$+ K''\sum(\mu_i\mu_{i+1}^* + \mu_{i+1}'\mu_i'^* + \mu_{i+1}'^*\mu_i + \mu_i'^*\mu_{i+1}) + \frac{C+B}{4}\sum(\mu_i + \mu_i' + \mu_i'^* + \mu_i'^*)\Big] \qquad (3,30)$$

Figure 9 shows how this equation is obtained on the probability approach; we write down the probability for an arrangement with the two end strips (labelled n) missing and then with them included and proceed as usual.

The largest eigenvalue of this last expression - equation (3,30) - will be the partition function per spin χ to the power m of our two plate crystal. In order to proceed further, we again use

$$\Pi_{max} = \chi^m \cdot \max_{\text{all } a\text{'s}} \frac{\sum \mathcal{N}(\mu,\mu';\mu^*,\mu'^*)\, a(\mu^*,\mu'^*)\, a(\mu,\mu')}{\sum[a(\mu,\mu')]^2} \qquad (3,31)$$

This time, we choose $a(\mu,\mu')$ to be of the form

$$a(\mu,\mu') = \exp\Big[\frac{E}{4}\sum(\mu_i'\mu_{i+1}' + \mu_i\mu_{i+1}) + F/2\sum\mu_i'\mu_i$$

$$+ \frac{E'}{2}\sum(\mu_i'\mu_{i+1} + \mu_{i+1}'\mu_i) + G/4\sum(\mu_i' + \mu_i)\Big] \qquad (3,32)$$

Substituting into (3,31), we find, after taking the m^{th} root

$$\chi = \max_{E,F,E',G} \frac{\varphi\left[\frac{K+A+E}{4} : \frac{K+F}{2} : \frac{K'+E'}{2} : \frac{C+B+G}{4}\right]}{\omega\left[\frac{E}{2} : F : E' : G/2\right]} \qquad (3,33)$$

From which we have

$$\omega^m = \sum \exp\Big[\frac{E}{2}\sum(\mu_i'\mu_{i+1}' + \mu_i\mu_{i+1}) + F\sum\mu_i'\mu_i + E'\sum(\mu_i'\mu_{i+1} + \mu_{i+1}'\mu_i) + \frac{G}{2}\sum(\mu_i' + \mu_i)\Big] \qquad (3,34)$$

which has the form of the partition function of two strips of m spins. In fact we can see that this equation has the same form as that from which (3,24) was derived, but with different interaction energies. It is obvious then, from (3,24) or otherwise, that on reducing this stage of the problem further, we shall reach the expression

$$\mathcal{N}(\mu_m, \mu_m') = \exp\Big[F\mu_m\mu_m' + \frac{E}{2}(\mu_m'\mu_{m-1}' + \mu_{m-1}\mu_m) + E'(\mu_m'\mu_{m-1} + \mu_{m-1}'\mu_m) + \frac{G}{2}(\mu_m' + \mu_m)\Big] \qquad (3,35)$$

which when written out in full as a matrix, will be of order 4×4 and of the same form as that corresponding to equation (3,24). Its largest eigenvalue ω will be the partition function per spin of the second subsidiary crystal of two strips of spins.

Finally we must consider our expression for φ . From equation (3,33), we see that φ^m has the form

$$\sum \exp\Big[\frac{K+A+E}{4} \sum (\mu_i'\mu_{i+1}' + \mu_i\mu_{i+1} + \mu_i'^{*}\mu_{i+1}'^{*} + \mu_i^{*}\mu_{i+1}^{*})$$

$$+ \frac{K+A}{2} \sum (\mu_i'\mu_i'^{*} + \mu_i\mu_i^{*}) + \frac{K+F}{2} \sum (\mu_i'\mu_i + \mu_i'^{*}\mu_i^{*})$$

$$+ \frac{K'+A'}{2} \sum (\mu_i'\mu_{i+1}'^{*} + \mu_{i+1}'\mu_i'^{*} + \mu_i\mu_{i+1}^{*} + \mu_{i+1}\mu_i^{*}) + K' \sum (\mu_i'\mu_i^{*} + \mu_i'^{*}\mu_i)$$

$$+ \frac{K'+E'}{2} \sum (\mu_i'\mu_{i+1} + \mu_{i+1}'\mu_i + \mu_i'^{*}\mu_{i+1}^{*} + \mu_{i+1}'^{*}\mu_i^{*})$$

$$+ K'' \sum (\mu_i'\mu_{i+1}^{*} + \mu_{i+1}'\mu_i^{*} + \mu_{i+1}'^{*}\mu_i + \mu_i'^{*}\mu_{i+1}) + \frac{C+B+G}{4} \sum (\mu_i' + \mu_i + \mu_i'^{*} + \mu_i^{*}) \Big] \quad (3,36)$$

It is easy to see that this expression has the form of the partition function for a crystal consisting of four linear chains of spins each containing m spins.

For the last time we use the probability treatment to reduce the problem, considering the chains to have first of all m-1 spins and then m spins. This final reduction leads us to the following expression for the elements of a matrix.

$$\mathcal{A}(\mu_m, \mu_m' : \mu_m^{*}, \mu_m'^{*}) = \exp\Big[\frac{K+A+E}{4} (\mu_m'\mu_{m-1}' + \mu_m\mu_{m-1} + \mu_m'^{*}\mu_{m-1}'^{*} + \mu_m^{*}\mu_{m-1}^{*})$$

$$+ \frac{K+A}{2} (\mu_m'\mu_m'^{*} + \mu_m\mu_m^{*}) + \frac{K+F}{2} (\mu_m'\mu_m + \mu_m'^{*}\mu_m^{*})$$

$$+ \frac{K'+A'}{2} (\mu_m'\mu_{m-1}'^{*} + \mu_m\mu_{m-1}^{*} + \mu_m'^{*}\mu_{m-1}' + \mu_m^{*}\mu_{m-1})$$

$$+ K' (\mu_m'\mu_m^{*} + \mu_m'^{*}\mu_m) + \frac{K'+E'}{2} (\mu_m'\mu_{m-1} + \mu_{m-1}'\mu_m + \mu_m'^{*}\mu_{m-1}^{*} + \mu_{m-1}'^{*}\mu_m^{*})$$

$$+ K'' (\mu_m'\mu_{m-1}^{*} + \mu_{m-1}'\mu_m^{*} + \mu_{m-1}'^{*}\mu_m + \mu_m'^{*}\mu_{m-1}) + \frac{C+B+G}{4} (\mu_m' + \mu_m + \mu_m'^{*} + \mu_m^{*}) \Big] \quad (3,37)$$

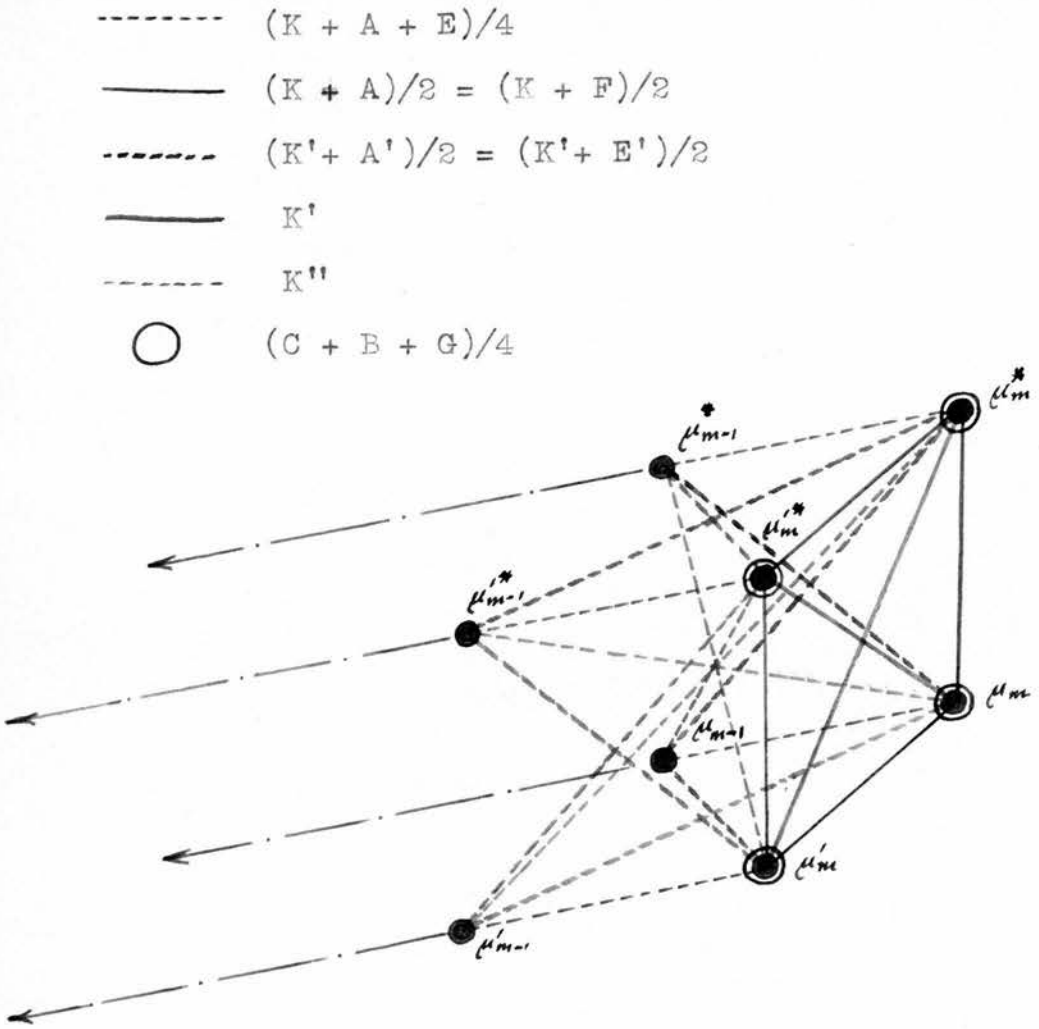


Figure 10.

The matrix represented by this expression is of order 16×16 and its largest eigenvalue $\mathcal{P}$ will be the partition function per spin of our 4 - chain system of spins. Figure 10 shows the reduction of the four linear chains of spins leading to equation (3,37). This sketch shows also the various interactions upon which equation (3,37) is based. This is of particular interest since, from the sketch, we see that due to the symmetry of the system, we can put the interaction corresponding to A equal to that corresponding to F and that corresponding to A' equal to that corresponding to E' - thus reducing the number of our unknown parameters.

It should be noted also that, in going from the simple cubic lattice and the corresponding matrices as set up in the foregoing theory, to say, the face centred lattice, although in putting nearest and third nearest neighbour interactions equal to zero, we get two superposed face centred lattices in place of the original simple cubic one, our equations need no further modification. We do not, for instance, have to take the square root of our various subsidiary partition functions, as might at first seem to be the case. This is of course due to the fact that the two face centred lattices are entirely independent of each other and to the fact that our expressions for the matrix elements lead to partition functions per spin for the

particular lattice types under consideration.

We see then, that the original partition function for which we set up our equation including the three orders of interaction, can be obtained approximately as λ , the partition function per spin in the form

$$\lambda = \frac{\chi}{\psi} = \frac{\varphi}{\omega} \Big/ \frac{\mathfrak{B}}{\zeta} = \frac{\varphi \zeta}{\omega \mathfrak{B}} \qquad (3,38)$$

where φ , $\mathfrak{B}$, ω , and ζ are partition functions per spin for various subsidiary lattices. It remains now to write out from the expressions which we have derived for their elements, the matrices corresponding to the significant quantities in equation (3,38).

§ 4. Below the Curie Temperature.

Our matrices have the following general expressions for their elements:

$$\exp\left[P \mu_m \mu_{m-1} + \tfrac{Q}{2} (\mu_m + \mu_{m-1}) \right] \qquad (3,25)$$

from which we want the largest eigenvalue ζ_{max}.

$$\exp\Big[A \mu'_m \mu_m + \tfrac{A+P}{2} (\mu'_m \mu'_{m-1} + \mu_m \mu_{m-1}) + A' (\mu_m \mu'_{m-1} + \mu'_m \mu_{m-1}) + \tfrac{B+Q}{2} (\mu_m + \mu'_m) \qquad (3,24)$$

from which we want the largest eigenvalue $\mathfrak{B}_{max}$.

$$\exp\Big[A\mu_m'\mu_m + \frac{E}{2}(\mu_m'\mu_{m-1}' + \mu_m\mu_{m-1}) + A'(\mu_m\mu_{m-1}' + \mu_m'\mu_{m-1}) + \frac{G}{2}(\mu_m + \mu_m')\Big] \qquad (3,35)$$

from which we want the largest eigenvalue ω_{max} .

$$\exp\Big[\frac{K+A+E}{4}(\mu_m'\mu_{m-1}' + \mu_m\mu_{m-1} + \mu_m'^{*}\mu_{m-1}'^{*} + \mu_m^{*}\mu_{m-1}^{*}) + \frac{K+A}{2}(\mu_m'\mu_m'^{*} + \mu_m\mu_m^{*} + \mu_m'\mu_m + \mu_m'^{*}\mu_m^{*}) + \frac{K'+A'}{2}(\mu_m'\mu_{m-1}'^{*} + \mu_m\mu_{m-1}^{*} + \mu_{m-1}'\mu_m'^{*} + \mu_{m-1}\mu_m^{*} + \mu_m'\mu_{m-1} + \mu_{m-1}'\mu_m + \mu_m'^{*}\mu_{m-1}^{*} + \mu_{m-1}^{*}\mu_m^{*}) + K''(\mu_m'\mu_{m-1}^{*} + \mu_{m-1}'\mu_m^{*} + \mu_{m-1}^{*}\mu_m + \mu_m'^{*}\mu_{m-1}) + K'(\mu_m'\mu_m^{*} + \mu_m'^{*}\mu_m) + \frac{C+B+G}{4}(\mu_m' + \mu_m + \mu_m'^{*} + \mu_m^{*})\Big] \qquad (3,37)$$

from which we want the largest eigenvalue φ_{max} .
These are respectively 2 × 2, 4 × 4, 4 × 4, and 16 × 16 matrices .

Let us first of all consider the largest matrix. Its array written as an eigenvalue equation (equation (3,39)) is shown on the following page. From inspection of the array, we see that it is indeed highly symmetrical. This we would expect of course, from our method of treatment of the problem. Incidentally, the symmetry of the array was of great assistance during the tedious process of writing out the 256 elements.

A more important feature of this symmetry lies in the

The starred row immediately below, gives the factors of the respective columns.

$K+A+E = 2\ln y$ *
$2(K'+A') = \ln\alpha$
$K+A = \ln x$
$C+B+G = 4\ln z$
$K' = \ln\beta$
$2K'' = \ln\mu$

	$x^2\beta^2z^4$	z^2	z^2	β^{-2}	z^2	β^{-2}	$x^{-2}\beta^2$	z^{-2}	z^2	$x^{-2}\beta^2$	β^{-2}	z^{-2}	β^{-2}	z^{-2}	z^{-2}	$x^2\beta^2z^{-4}$
μ_m	$+1$								-1							
$\mu_{m'}$	$+1$				-1				$+1$				-1			
μ^*_m	$+1$		-1		$+1$		-1		$+1$		-1		$+1$		-1	
$\mu^*_{m'}$	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1
$\mu_{m-1}=+1$, $\mu'_{m-1}=+1$, $\mu^*_{m-1}=+1$	$\alpha^2y^2\mu^{-2}\varphi$	$\alpha y\mu$	$\alpha y\mu$	1	$\alpha y\mu$	1	1	$\alpha^{-1}y^{-1}\mu^{-1}$	$\alpha y\mu$	1	1	$\alpha^{-1}y^{-1}\mu^{-1}$	1	$\alpha^{-1}y^{-1}\mu^{-1}$	$\alpha^{-1}y^{-1}\mu^{-1}$	$\alpha^{-2}y^{-2}\mu^{2}$
$\mu_{m-1}=+1$, $\mu'_{m-1}=+1$, $\mu^*_{m-1}=-1$	$\alpha y\mu$	$y^2-\varphi$	α	$y\mu^{-1}$	α	$y\mu^{-1}$	$\alpha y^{-1}\mu^{-1}$	μ^{-2}	μ^2	$\alpha^{-1}y\mu$	$y^{-1}\mu$	α^{-1}	$y^{-1}\mu$	y^{-2}	$\alpha^{-1}y^{-1}\mu^{-1}$	
$\mu_{m-1}=+1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=+1$	$\alpha y\mu$	α	$y^2-\varphi$	$y\mu^{-1}$	μ^2	$y^{-1}\mu$	$\alpha^{-1}y\mu$	α^{-1}	α	$\alpha y^{-1}\mu^{-1}$	$y\mu^{-1}$	μ^{-2}	$y^{-1}\mu$	y^{-2}	α^{-1}	$\alpha^{-1}y^{-1}\mu^{-1}$
$\mu_{m-1}=+1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=-1$	1	$y\mu^{-1}$	$y\mu^{-1}$	$y^2\mu^{-2}-\varphi$	$y^{-1}\mu$	1	1	$y\mu^{-1}$	$y^{-1}\mu$	1	1	$y\mu^{-1}$	$y^{-2}\mu^2$	$y^{-1}\mu$	$y^{-1}\mu$	1
$\mu_{m-1}=-1$, $\mu'_{m-1}=+1$, $\mu^*_{m-1}=+1$	$\alpha y\mu$	α	μ^2	$y^{-1}\mu$	$y^2-\varphi$	$y\mu^{-1}$	$\alpha^{-1}y\mu$	α^{-1}	α	$\alpha y^{-1}\mu^{-1}$	$y^{-1}\mu$	y^{-2}	$y\mu^{-1}$	μ^{-2}	α^{-1}	$\alpha^{-1}y^{-1}\mu^{-1}$
$\mu_{m-1}=-1$, $\mu'_{m-1}=+1$, $\mu^*_{m-1}=-1$	1	$y\mu^{-1}$	$y^{-1}\mu$	1	$y\mu^{-1}$	$y^2\mu^{-2}-\varphi$	1	$y\mu^{-1}$	$y^{-1}\mu$	1	$y^{-2}\mu^2$	$y^{-1}\mu$	1	$y\mu^{-1}$	$y^{-1}\mu$	1
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=+1$	1	$\alpha y^{-1}\mu^{-1}$	$\alpha^{-1}y\mu$	1	$\alpha^{-1}y\mu$	1	$\alpha^{-2}y^2\mu^2\varphi$	$\alpha^{-1}y\mu$	$\alpha y^{-1}\mu^{-1}$	$\alpha^2y^{-2}\mu^{-2}$	1	$\alpha y^{-1}\mu^{-1}$	1	$y^{-1}\alpha\mu^{-1}$	$y\alpha^{-1}\mu$	1
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=-1$	$\alpha^{-1}y^{-1}\mu^{-1}$	μ^{-2}	α^{-1}	$y\mu^{-1}$	α^{-1}	$y\mu^{-1}$	$\alpha^{-1}y\mu$	$y^2-\varphi$	y^{-2}	$\alpha y^{-1}\mu^{-1}$	$y^{-1}\mu$	α	$y^{-1}\mu$	α	μ^2	$\alpha y\mu$
$\mu_{m-1}=-1$, $\mu'_{m-1}=+1$, $\mu^*_{m-1}=+1$	$\alpha y\mu$	μ^2	α	$y^{-1}\mu$	α	$y^{-1}\mu$	$\alpha y^{-1}\mu^{-1}$	y^{-2}	$y^2-\varphi$	$\alpha^{-1}y\mu$	$y\mu^{-1}$	α^{-1}	$y\mu^{-1}$	α^{-1}	μ^{-2}	$\alpha^{-1}y^{-1}\mu^{-1}$
$\mu_{m-1}=-1$, $\mu'_{m-1}=+1$, $\mu^*_{m-1}=-1$	1	$\alpha^{-1}y\mu$	$\alpha y^{-1}\mu^{-1}$	1	$\alpha y^{-1}\mu^{-1}$	1	$\alpha^2y^{-2}\mu^{-2}$	$\alpha y^{-1}\mu^{-1}$	$\alpha^{-1}y\mu$	$\alpha^{-2}y^2\mu^{-2}-\varphi$	1	$\alpha^{-1}y\mu$	1	$\alpha y^{-1}\mu^{-1}$	1	1
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=+1$	1	$y^{-1}\mu$	$y\mu^{-1}$	1	$y^{-1}\mu$	$y^{-2}\mu^2$	1	$y^{-1}\mu$	$y\mu^{-1}$	1	$y^2\mu^{-2}-\varphi$	$y\mu^{-1}$	1	$y^{-1}\mu$	$y\mu^{-1}$	1
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=-1$	$\alpha^{-1}y^{-1}\mu^{-1}$	α^{-1}	μ^{-2}	$y\mu^{-1}$	y^{-2}	$y^{-1}\mu$	$\alpha y^{-1}\mu^{-1}$	α	α^{-1}	$\alpha^{-1}y\mu$	$y\mu^{-1}$	$y^2-\varphi$	$y^{-1}\mu$	μ^2	α	$\alpha y\mu$
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=+1$	1	$y^{-1}\mu$	$y^{-1}\mu$	$y^{-2}\mu^2$	$y\mu^{-1}$	1	1	$y^{-1}\mu$	$y\mu^{-1}$	1	1	$y^{-1}\mu$	$y^2\mu^{-2}-\varphi$	$y\mu^{-1}$	$y\mu^{-1}$	1
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=-1$	$\alpha^{-1}y^{-1}\mu^{-1}$	α^{-1}	y^{-2}	$y^{-1}\mu$	y^{-2}	$y^{-1}\mu$	$\alpha y^{-1}\mu^{-1}$	α	α^{-1}	$\alpha^{-1}y\mu$	μ^2	$y^2-\varphi$	$y^{-1}\mu$	μ^2	α	$\alpha y\mu$
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=+1$	1	$y^{-1}\mu$	$y^{-1}\mu$	$y^{-2}\mu^2$	$y\mu^{-1}$	1	1	$y^{-1}\mu$	$y\mu^{-1}$	1	1	$y^{-1}\mu$	$y^2\mu^{-2}-\varphi$	$y\mu^{-1}$	$y\mu^{-1}$	1
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=-1$	$\alpha^{-1}y^{-1}\mu^{-1}$	α^{-1}	y^{-2}	$y^{-1}\mu$	y^{-2}	$y^{-1}\mu$	$\alpha y^{-1}\mu^{-1}$	α	α^{-1}	$\alpha^{-1}y\mu$	μ^2	$y^2-\varphi$	$y^{-1}\mu$	μ^2	α	$\alpha y\mu$
$\mu_{m-1}=-1$, $\mu'_{m-1}=-1$, $\mu^*_{m-1}=+1$	$\alpha^{-2}y^{-2}\mu^{-2}$	$\alpha^{-1}y^{-1}\mu^{-1}$	$\alpha^{-1}y^{-1}\mu^{-1}$	1	$\alpha^{-1}y^{-1}\mu^{-1}$	1	1	$\alpha y\mu$	$\alpha^{-1}y^{-1}\mu^{-1}$	1	1	$\alpha y\mu$	1	$\alpha y\mu$	$\alpha y\mu$	$\alpha^2y^2\mu^{-2}-\varphi$

Equation (3,39)

possibility of being able to reduce the order of the determinant since the 16 X 16 determinant is far too large to handle. A point to be considered in this connection is the fact that if, after reducing the matrix, its order is still high - say order 4 or greater - then it will be of importance to try to preserve the shape of the array to the extent of keeping the φ's on the principal diagonal. We shall then be able to multiply out using the spur method i.e. directly in powers of φ . In other words, we want to make our reducing operations symmetrical.

In carrying out the reduction, we found that the following series of operations were successful.

Permutate first the rows and then the columns in the order:

1, 16, 7, 10, 2, 3, 5, 9, 8, 12, 14, 15, 4, 6, 11, 13.

Add columns as:

4 + 3, 6 + 5, 8 + 7, 10 + 9, 12 + 11, 16 + 13, 15 + 14.

Subtract rows as:

4 - 3, 6 - 5, 8 - 7, 10 - 9, 12 - 11, 16 - 13, 15 - 14.

Add columns as:

7 + 5, 8 + 6, 11 + 9, 12 + 10, 14 + 13, 15 + 16.

Subtract rows as:

7 - 5, 8 - 6, 11 - 9, 12 - 10, 14 - 13, 15 - 16.

Permutate first the columns and then the rows in the order:

1, 2, 3, 5, 9, 13, 4, 8, 12, 6, 10, 15, 7, 11, 16, 14.

in which subtractions are made from the underlined rows and additions are made to the underlined columns.

The result of these operations is shown on page 74.

Martin and ter Haar have shown that the largest eigenvalue f in a 4 × 4 matrix eigenvalue problem is contained in a 2 × 2 determinant of the form

$$\left| a_{ik} - b_i \delta_{ik} f \right| = 0 \qquad (3,41)$$

where δ_{ik} is the Kronecker delta, provided the original 4 × 4 matrix has certain transformation properties among its rows and columns. For a 16 × 16 matrix it is much more difficult to show that the determinant having the largest eigenvalue is of a form analogous to (3,41), but we may take it that the determinant in question in our case, is that one of order 6 × 6 contained in the top left hand corner of (3,40), i.e.

$\mu^2\alpha^2y^2 - \varphi/x^2\beta^2z^4$	$\alpha^{-2}y^{-2}\mu^{-2}$	2	$4\alpha y\mu$	$4\alpha^{-1}y^{-1}\mu^{-1}$	4	
$\alpha^{-2}y^{-2}\mu^{-2}$	$\mu^2y^2\alpha^2 - \varphi/x^2\beta^2z^4$	2	$4\alpha^{-1}y^{-1}\mu^{-1}$	$4\alpha y\mu$	4	
1	1	$\mu^2y^2\alpha^{-2} + \mu^{-2}y^{-2}\alpha^2 - \varphi/x^2\beta^2$	$2(\alpha y\mu^{-1} + \alpha^{-1}y\mu)$	$2(\alpha^{-1}y\mu + \alpha y^{-1}\mu^{-1})$	4	$= 0$
$\alpha y\mu$	$\alpha^{-1}y^{-1}\mu^{-1}$	$\mu^{-1}\alpha y^{-1} + \alpha^{-1}\mu y$	$y^2 + 2\alpha + \mu^2 - \varphi/z^2$	$y^{-2} + 2\alpha^{-1} + \mu^{-2}$	$2(\mu^{-1}y + y^{-1}\mu)$	
$\alpha^{-1}y^{-1}\mu^{-1}$	$\alpha y\mu$	$\mu^{-1}\alpha y^{-1} + \mu y\alpha^{-1}$	$y^{-2} + 2\alpha^{-1} + \mu^{-2}$	$y^2 + 2\alpha + \mu^2 - \varphi/z^{-2}$	$2(\mu^{-1}y + y^{-1}\mu)$	
1	1	2	$2(y\mu^{-1} + y^{-1}\mu)$	$2(\mu^{-1}y + y^{-1}\mu)$	$\mu^{-2}y^2 + \mu^2y^{-2} + 2 - \varphi/\beta^{-2}$	

This is equation (3,42).

Equation (3,40) = 0

$\mu^2\alpha^2\gamma^2 - \varphi/x^2\beta^2 z^4$	$\alpha^{-2}\gamma^{-2}\mu^{-2}$	2	$4\alpha\gamma\mu$	$4\alpha^{-1}\gamma^{-1}\mu^{-1}$	4	1	$\mu\gamma\alpha$	$\alpha^{-1}\gamma^{-1}\mu^{-1}$	$2\alpha\gamma\mu$	$2\alpha^{-1}\gamma^{-1}\mu^{-1}$	1	$2\alpha\gamma\mu$	$2\alpha^{-1}\gamma^{-1}\mu^{-1}$	2	2
$\alpha^{-2}\gamma^{-2}\mu^{-2}$	$\mu^2\alpha^2\gamma^2 - \varphi/x^2\beta^2 z^4$	2	$4\alpha^{-1}\gamma^{-1}\mu^{-1}$	$4\alpha\gamma\mu$	4	1	$\mu^{-1}\alpha^{-1}\gamma^{-1}$	$\alpha\gamma\mu$	$2\alpha^{-1}\gamma^{-1}\mu^{-1}$	$2\alpha\gamma\mu$	1	$2\alpha^{-1}\gamma^{-1}\mu^{-1}$	$2\alpha\gamma\mu$	2	2
1	1	$\mu^2\alpha^{-2}\gamma^2 + \alpha^2\gamma^{-2}\mu^{-2} - \varphi/x^{-2}\beta^2$	$2(\alpha\gamma^{-1}\mu^{-1} + \mu\alpha^{-1}\gamma)$	$2(\alpha^{-1}\gamma\mu + \mu^{-1}\alpha\gamma^{-1})$	4	$\alpha^2\gamma^{-2}\mu^{-2}$	$\mu^{-1}\alpha\gamma^{-1}$	$\alpha^{-1}\gamma\mu$	$\mu\alpha^{-1}\gamma + \alpha\gamma^{-1}\mu^{-1}$	$\mu^{-1}\gamma^{-1}\alpha + \alpha^{-1}\gamma\mu$	1	$\mu\alpha^{-1}\gamma + \alpha\gamma^{-1}\mu^{-1}$	$\mu^{-1}\alpha\gamma^{-1} + \alpha\gamma^{-1}\mu$	2	2
$\alpha\gamma\mu$	$\alpha^{-1}\gamma^{-1}\mu^{-1}$	$\mu^{-1}\alpha\gamma^{-1} + \alpha^{-1}\gamma\mu$	$\gamma^{-2} + 2\alpha^{-1} - \mu^{-2} - \varphi/z^2$	$\gamma^{-2} + 2\alpha^{-1} + \mu^{-2}$	$2(\mu^{-1}\gamma + \gamma^{-1}\mu)$	$\alpha^{-1}\gamma\mu$	μ^{-2}	γ^{-2}	$\mu^{-2} + \alpha$	$\gamma^{-2} + \alpha^{-1}$	$\gamma^{-1}\mu$	$\alpha + \mu^2$	$\alpha^{-1} + \gamma^{-2}$	$2\gamma^{-1}\mu$	$\mu^{-1}\gamma - \gamma^{-1}\mu$
$\alpha^{-1}\gamma^{-1}\mu^{-1}$	$\alpha\gamma\mu$	$\mu\alpha^{-1}\gamma + \alpha\gamma^{-1}\mu^{-1}$	$\gamma^{-2} + 2\alpha^{-1} + \mu^{-2}$	$\gamma^2 + 2\alpha + \mu^2 - \frac{\varphi}{z^{-2}}$	$2(\mu^{-1}\gamma + \gamma^{-1}\mu)$	$\alpha\gamma^{-1}\mu^{-1}$	γ^{-2}	μ^2	$\gamma^{-2} + \alpha^{-1}$	$\alpha + \mu^2$	$\gamma^{-1}\mu$	$\alpha^{-1} + \gamma^{-2}$	$\alpha + \mu^2$	$2\gamma^{-1}\mu$	$\mu^{-1}\gamma - \gamma^{-1}\mu$
1	1	2	$2(\mu^{-1}\gamma + \gamma^{-1}\mu)$	$2(\mu^{-1}\gamma + \gamma^{-1}\mu)$	$\mu^{-2}\gamma^2 + \gamma^{-2}\mu^2 - \varphi/\beta^{-2}z^{-2}$	1	$\mu\gamma^{-1}$	$\gamma^{-1}\mu$	$\mu^{-1}\gamma + \gamma^{-1}\mu$	$\mu^{-1}\gamma + \gamma^{-1}\mu$	1	$2\gamma^{-1}\mu$	$2\gamma^{-1}\mu$	$\mu^2\gamma^{-2} + 1$	2
0	0	0	0	0	0	$\mu^2\alpha^{-2}\gamma^2 - \alpha^2\mu^{-2}\gamma^{-2} - \varphi/x^{-2}\beta^2$	$\mu\alpha^{-1}\gamma - \alpha\gamma^{-1}\mu^{-1}$	$\mu^{-1}\alpha\gamma^{-1} - \alpha^{-1}\gamma\mu$	0	0	0	0	0	0	0
0	0	0	0	0	0	$2(\alpha^{-1}\gamma\mu - \mu^{-1}\gamma^{-1}\alpha^{-2})$	$\gamma^2 - 2\alpha + \mu^2 - \frac{\varphi}{z^2}$	$\mu^{-2} - 2\alpha^{-1} + \gamma^{-2}$	0	0	0	0	0	0	0
0	0	0	0	0	0	$2(\mu^{-1}\gamma^{-1}\alpha^{-1} - \alpha^{-1}\gamma\mu)$	$\mu^{-2} - 2\alpha^{-1} + \gamma^{-2}$	$\gamma^2 - 2\alpha + \mu^2 - \frac{\varphi}{z^{-2}}$	0	0	0	0	0	0	0
0	0	0	0	0	0	$\mu^{-2}\alpha\gamma^{-1} - \alpha^{-1}\mu\gamma$	$\alpha - \mu^2$	$\alpha^{-1} - \gamma^{-2}$	$\gamma^2 - \mu^2 - \varphi/z^2$	$\mu^{-2} - \gamma^{-2}$	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0	0	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0
0	0	0	0	0	0	$\mu\alpha^{-1}\gamma - \alpha\gamma^{-1}\mu^{-1}$	$\alpha^{-1} - \gamma^{-2}$	$\alpha - \mu^2$	$\mu^{-2} - \gamma^{-2}$	$\gamma^2 - \mu^2 - \varphi/z^{-2}$	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0	0	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0
0	0	0	0	0	0	0	0	0	$2(\mu^{-1}\gamma - \gamma^{-1}\mu)$	$2(\mu^{-1}\gamma - \gamma^{-1}\mu)$	$\mu^{-2}\gamma^2 - \gamma^{-2}\mu^2 - \varphi/\beta^{-2}$	$2(\gamma^{-1}\mu - \mu^{-1}\gamma)$	$2(\gamma^{-1}\mu - \mu^{-1}\gamma)$	0	0
0	0	0	0	0	0	$\mu^{-1}\alpha\gamma^{-1} - \alpha^{-1}\gamma\mu$	$\alpha - \mu^2$	$\alpha^{-1} - \gamma^{-2}$	0	0	0	$\gamma^2 - \mu^2 - \varphi/z^2$	$\mu^{-2} - \gamma^{-2}$	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0
0	0	0	0	0	0	$\mu\alpha^{-1}\gamma - \alpha\gamma^{-1}\mu^{-1}$	$\alpha^{-1} - \gamma^{-2}$	$\alpha - \mu^2$	0	0	0	$\mu^{-2} - \gamma^2$	$\gamma^2 - \mu^2 - \varphi/z^{-2}$	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0
0	0	0	0	0	0	0	$\mu^{-1}\gamma - \mu\gamma^{-1}$	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0	0	0	$2(\gamma\mu^{-1} - \mu\gamma^{-1})$	$2(\gamma\mu^{-1} - \gamma^{-1}\mu)$	$\mu^{-2}\gamma^2 - \gamma^{-2}\mu^2 - \varphi/\beta^{-2}$	0
0	0	0	0	0	0	0	0	0	$\mu\gamma^{-1} - \gamma\mu^{-1}$	$\mu\gamma^{-1} - \gamma\mu^{-1}$	$\mu^2\gamma^{-2} - 1$	$\mu^{-1}\gamma - \gamma^{-1}\mu$	$\mu^{-1}\gamma - \gamma^{-1}\mu$	0	$\mu^{-2}\gamma^2 - \gamma^{-2}\mu^2 - 2 - \varphi/\beta^{-2}$

For the determinant corresponding to equation (3,24), we find

		$\mu_m = 1$		$\mu_m = -1$		$= 0$
		$\mu_{m'} = 1$	$\mu_{m'} = -1$	$\mu_{m'} = 1$	$\mu_{m'} = -1$	
$\mu_{m-1} = 1$	$\mu'_{m-1} = 1$	$\alpha\rho\alpha'^2\sigma - \nu\beta$	α^{-1}	α^{-1}	$\alpha\rho^{-1}\alpha'^{-2}\sigma^{-1}$	
	$\mu'_{m-1} = -1$	$\alpha\sigma$	$\alpha^{-1}\rho\alpha'^{-2} - \nu\beta$	$\alpha^{-1}\rho^{-1}\alpha'^{2}$	$\alpha\sigma^{-1}$	
$\mu_{m-1} = -1$	$\mu'_{m-1} = 1$	$\alpha\sigma$	$\alpha^{-1}\rho^{-1}\alpha'^{2}$	$\alpha^{-1}\rho\alpha'^{-2} - \nu\beta$	$\alpha\sigma^{-1}$	
	$\mu'_{m-1} = -1$	$\alpha\rho^{-1}\alpha'^{-2}\sigma$	α^{-1}	α^{-1}	$\alpha\rho\alpha'^{2}\sigma^{-1} - \nu\beta$	

- equation (3,43).

And for that corresponding to (3,35)

		$\mu_m = 1$		$\mu_m = -1$		$= 0$
		$\mu_{m'} = 1$	$\mu_{m'} = -1$	$\mu_{m'} = 1$	$\mu_{m'} = -1$	
$\mu_{m-1} = 1$	$\mu'_{m-1} = 1$	$\alpha\varepsilon\alpha'^2\gamma - \omega$	α^{-1}	α^{-1}	$\alpha\varepsilon^{-1}\alpha'^{-2}\gamma^{-1}$	
	$\mu'_{m-1} = -1$	$\alpha\gamma$	$\alpha^{-1}\varepsilon\alpha'^{-2} - \omega$	$\alpha^{-1}\varepsilon^{-1}\alpha'^{2}$	$\alpha\gamma^{-1}$	
$\mu_{m-1} = -1$	$\mu'_{m-1} = 1$	$\alpha\gamma$	$\alpha^{-1}\varepsilon^{-1}\alpha'^{2}$	$\alpha^{-1}\varepsilon\alpha'^{-2} - \omega$	$\alpha\gamma^{-1}$	
	$\mu'_{m-1} = -1$	$\alpha\varepsilon^{-1}\alpha'^{-2}\gamma$	α^{-1}	α^{-1}	$\alpha\varepsilon\alpha'^{2}\gamma^{-1} - \omega$	

- equation (3,44).

In these equations, $e^A = a$: $e^{A+P} = p$: $e^E = e$: $e^{A'} = a'$: $e^{B+Q} = q$: $e^G = g$.

For the determinant corresponding to (3,25), we have

$rs - \beta$	r^{-1}
r^{-1}	$rs^{-1} - \beta$

(3,45)

$e^P = r$: $e^Q = s$.

With these last four equations to hand, we are now in a position to solve any of the problems listed at the beginning of this chapter; as they stand, they relate to the simple cubic lattice with three orders of interaction. In order to effect a solution of any particular problem, we shall need other equations of course, since in general, we shall have more unknowns than we have equations . These other equations can be obtained from our maximisation conditions - this point will be dealt with later when it arises.

Since we cannot hope to find closed solutions of our equations, we shall look for series expansions for our various partition functions. Obviously, if our models have any physical significance, they should show a transition temperature so we shall follow the usual procedure and attempt to find series expansions for our various physical quantities valid above and below the Curie temperature.

In the rest of the work, we shall consider the case of zero magnetic field - that is, H, and therefore also C, we put equal to zero. A consequence of this (see the discussion of the square net of spins) is that at high temp-

eratures, B, Q, and G will all be zero for λ to be a maximum. At low temperatures, these quantities will be zero when λ has a minimum value and this has no physical significance.

Thus we see that our equations as given by the determinants (3,42) ..... (3,45) apply as they stand, to the simple cubic ferromagnet in the region below the Curie temperature if C = 0. Putting C = 0 means simply that z is equal to $e^{(B+G)/4}$ instead of $e^{(B+G+C)/4}$.

Above the Curie point however, we have C = B = Q = G = 0, so that in our determinants, q = g = z = s = 1, and as a consequence, we can reduce the determinants still further.

§ 5 Above the Curie Temperature.

In the 6 X 6 determinant - equation (3,42) - we may put z = 1 and perform the following operations:

Add column 2 to column 1 and column 5 to column 4

Subtract row 1 from row 2 and row 4 from row 5.

Permutate rows and columns in the order 1, 3, 4, 6, 2, 5.

This effectively reduces our determinant to order 4 X 4 on making use of the discussion attendant on equation (3,41). We find the determinant shown at the top of the following page

$\mu^2\alpha^2y^2+\mu^{-2}\alpha^{-2}y^{-2}-\varphi/x^2\beta^2$	2	$4(\alpha y\mu+\alpha^{-1}y^{-1}\mu^{-1})$	4
2	$\mu^2\alpha^{-2}y^2+y^{-2}\alpha^2\mu^{-2}-\varphi/x^{-2}\beta^2$	$4(\alpha y^{-1}\mu^{-1}+\alpha^{-1}y\mu)$	4
$\alpha y\mu+\alpha^{-1}y^{-1}\mu^{-1}$	$\mu^{-1}\alpha y^{-1}+\alpha^{-1}y\mu$	$y^2+y^{-2}+2\alpha+2\alpha^{-1}+\mu^2+\mu^{-2}-\varphi$	$2(\mu^{-1}y+y^{-1}\mu)$
2	2	$4(y\mu^{-1}+y^{-1}\mu)$	$\mu^{-2}y^2+\mu^2y^{-2}+2-\varphi/\beta^{-2}$

$= 0$

- equation (3,46).

In the 4 X 4 determinants - equations (3,43) and (3,44), we can put q = g = 1 and perform the operations:

Subtract column 3 from column 2 and column 4 from column 1.

Add row 2 to row 3 and row 1 to row 4.

This reduces these determinants to

$a(\rho^{-1}b^{-1}+\rho b)-2\beta$	a^{-1}
$4a$	$a^{-1}(\rho b^{-1}+\rho^{-1}b)-2\beta$

$= 0$ (3,47)

and

$a(e^{-1}b^{-1}+eb)-\omega$	a^{-1}
$4a$	$a^{-1}(eb^{-1}+e^{-1}b)-\omega$

$= 0$ (3,48)

where we have put $a'^2 = b$.

Incidentally, the determinants (3,43) and (3,44) can be

reduced to order 3 × 3 as they stand.

In the 2 × 2 determinant, we put s = 1 and get

$$\begin{vmatrix} \gamma - 6 & \gamma^{-1} \\ \gamma^{-1} & \gamma - 6 \end{vmatrix} = 0 \qquad (3,49)$$

Thus, for our general simple cubic case, we have the four equations (3,42), (3,43), (3,44), and (3,45) referring to the low temperature region for our model and, for the high temperature region, we have the corresponding equations (3,46), (3,47), (3,48), and (3,49). We turn now to the question of the specific problems we started out to investigate.

IV

THE BODY CENTRED AND FACE CENTRED CUBIC FERROMAGNETS

§ 1 Body Centred Cubic Ferromagnet at High Temperatures.

In order to treat this problem, we put our first and second order interaction energies equal to zero; that is, we put

$$a = p = e = a' \text{ (or } b) = y = x = r = \alpha = \beta = 1. \qquad (4,1)$$

This transforms our simple cubic lattice into four superimposed body centred lattices which are quite independent of each other. The third order interactions of the simple cubic lattice become the first order interactions of the body centred cubic lattices. For the high temperature region, we apply equation (4,1) to the equations (3,46), (3,47), (3,48), and (3,49).

In (3,46), we find that the determinant can be further reduced by the operations:
Add column 1 to column 2 and subtract row 2 from row 1. This reduces our determinant to order 3 × 3. We then subtract row 1 from row 3 and add column 3 to column 1. We are left with the following determinant of order 2 × 2.

$$\begin{vmatrix} \mu^2+\mu^{-2}+6-\varphi & 4(\mu+\mu^{-1}) \\ 4(\mu+\mu^{-1}) & \mu^2+\mu^{-2}+6-\varphi \end{vmatrix} = 0$$

which has for its largest root or eigenvalue

$$\varphi_{max} = \mu^2 + \mu^{-2} + 4\mu + 4\mu^{-1} + 6$$
$$= \frac{(\mu+1)^4}{\mu^2} = (\mu^{1/2} + \mu^{-1/2})^4 = (e^{K} + e^{-K})^4 \qquad (4,2)$$

where since $K = K' = 0$, we have put $K'' = K$ for convenience. We see that we have no undetermined parameters left in this equation.

Now consider the equations (3,47) and (3,48). Using the same conditions as before, we find that these equations reduce straightaway to

$$\nu = \omega = 4 \qquad (4,3)$$

Similarly, equation (3,49) reduces to

$$\psi = 2 \qquad (4,4)$$

Thus, for the partition function per spin of the body centred cubic ferromagnet at high temperatures, we have on this treatment,

$$\lambda = \varphi\psi/\omega\nu = \frac{2}{16}(e^{K}+e^{-K})^4 = 2\cosh^4 K. \qquad (4,5)$$

We shall see in the next chapter, that this result is of considerable interest when we compare it to analogous results derived on other approximate methods. For the present, however, it is sufficient to point out that obviously the body centred case is a limiting case of our approach since all our parameters go out. The question of maximisation, which is fundamental to the method, does

not arise in this problem. We might expect then, that the result expressed by equation (4,5) will not be a particularly good one - or at least, that it will not be as good as other results derived using our method.

Two other points of interest here are the following. First of all, we see that we have found our result in a closed form; however, in view of the above considerations, this is not of much significance or importance. It might in fact be said to be due to the crudity of our approximations in this case.

The other point is seen if we let the temperature go to infinity - i.e. $K \longrightarrow 0$. Under these circumstances we find our limit for λ to be $\lambda = 2$. This has been pointed out by Kramers and Wannier and is general for these problems.

§ 2 Body Centred Cubic Ferromagnet at Low Temperatures.

For this part of the problem, we use the body centred cubic substitutions - (4,1) - in the determinantal equations (3,42), (3,43), (3,44), and (3,45). This gives

$$\begin{vmatrix} q-2 & 1 & 1 & q^{-1} \\ q & 1-2 & 1 & q^{-1} \\ q & 1 & 1-2 & q^{-1} \\ q & 1 & 1 & q^{-1}-2 \end{vmatrix} = 0 \qquad \begin{vmatrix} g-\omega & 1 & 1 & g^{-1} \\ g & 1-\omega & 1 & g^{-1} \\ g & 1 & 1-\omega & g^{-1} \\ g & 1 & 1 & g^{-1}-\omega \end{vmatrix} = 0$$

$\mu^2 - \varphi/z^4$	μ^{-2}	2	4μ	$4\mu^{-1}$	4
μ^{-2}	$\mu^2 - \varphi/z^4$	2	$4\mu^{-1}$	4μ	4
1	1	$\mu^2+\mu^{-2}-\varphi$	$2(\mu+\mu^{-1})$	$2(\mu+\mu^{-1})$	4
μ	μ^{-1}	$\mu+\mu^{-1}$	$3+\mu^2-\frac{\varphi}{z^2}$	$3+\mu^{-2}$	$2(\mu+\mu^{-1})$
μ^{-1}	μ	$\mu^{-1}+\mu$	$3+\mu^{-2}$	$3+\mu^2-\frac{\varphi}{z^2}$	$2(\mu^{-1}+\mu)$
1	1	2	$2(\mu+\mu^{-1})$	$2(\mu^{-1}+\mu)$	$\mu^{-2}+\mu^2+2-\varphi$

$= 0$

and

$s - \psi$	1
1	$s^{-1} - \psi$

$= 0$

In which all the symbols refer to 'magnetic' terms except μ which is $e^{2K''}$ - our original third order interaction term.

In the determinants in ψ and ω, we carry out the operations:

Add column 2 to column 1 and column 4 to column 3.

Subtract row 1 from row 2 and row 3 from row 4. This results in determinants of the form

$\frac{q}{g} + 1 - \frac{2\psi}{\omega}$	$1 + \frac{g^{-1}}{g^{-1}}$
$\frac{q}{g} + 1$	$1 + \frac{g^{-1}}{g^{-1}} - \frac{2\psi}{\omega}$

$= 0$

having maximum eigenvalues

$$\nu = 2 + q^{-1} + q \qquad (4,6)$$

$$\omega = 2 + \sigma^{-1} + \sigma \qquad (4,7)$$

We have also $$\psi = s + s^{-1} \qquad (4,8)$$

as our maximum eigenvalue of the ψ determinant.

For the determinant in φ, by subtracting row 3 from row 6 and adding column 6 to column 3, we can reduce to order 5 X 5:

$\mu^2 z^4 - \varphi$	$\mu^{-2} z^{-4}$	6	$4\mu z^2$	$4\mu^{-1} z^{-2}$	$= 0$
$\mu^{-2} z^4$	$\mu^2 z^{-4} - \varphi$	6	$4\mu^{-1} z^2$	$4\mu z^{-2}$	
z^4	z^{-4}	$\mu^2 + \mu^{-2} + 4 - \varphi$	$2z^2(\mu^{-1} + \mu)$	$2z^{-2}(\mu + \mu^{-1})$	
μz^4	$\mu^{-1} z^{-4}$	$3(\mu + \mu^{-1})$	$z^2(3 + \mu^2) - \varphi$	$z^{-2}(3 + \mu^{-2})$	
$\mu^{-1} z^4$	μz^{-4}	$3(\mu + \mu^{-1})$	$z^2(3 + \mu^{-2})$	$z^{-2}(3 + \mu^2) - \varphi$	

We can expand this by the spur method, which in this case involves the evaluation of 26 subsidiary determinants. On doing so, we find

$$\varphi^5 - \varphi^4\left[\mu^2(z^4+z^{-4}) + (3+\mu^2)(z^2+z^{-2}) + \mu^2+\mu^{-2}+4\right]$$

$$+\varphi^3\left[\mu^2-1\right]\left[\mu^2(z^6+z^{-6}) + (5+\mu^2)(z^4+z^{-4}) + \mu^{-2}(2\mu^4+6\mu^2+7)(z^2+z^{-2})\right.$$

$$\left. + \mu^{-4}(\mu^2+1)(2\mu^4+6\mu^2+2)\right]$$

$$-\varphi^2\left[\mu^2-1\right]^3\left[(z^6+z^{-6}) + \mu^{-2}(\mu^2+5)(z^4+z^{-4}) + \mu^{-4}(2\mu^4+6\mu^2+7)(z^2+z^{-2})\right.$$

$$\left. + \mu^{-6}(\mu^2+1)(2\mu^4+6\mu^2+2)\right]$$

$$+\varphi\left[\mu^2-1\right]^6\left[\mu^{-4}(z^4+z^{-4}) + \mu^{-6}(3+\mu^2)(z^2+z^{-2}) + \mu^{-8}(\mu^4+4\mu^2+1)\right] - \mu^{-10}(\mu^2-1)^{10} = 0 \quad (4,9)$$

It is easy to see that this equation has the form

$$\varphi^5 - A\varphi^4 + BC\varphi^3 - B^3C\varphi^2 + B^6A\varphi - B^{10} = 0 \quad (4,10)$$

where $B = \mu^{-1}(\mu^2-1)$

C = contents of the larger square bracket in the coefficient of φ^3 with a factor μ^{-1} taken out.

A = coefficient of φ^4.

(4,10) has a factor $\varphi - B^2$ as may easily be verified, and we find that (4,9) reduces for our purposes to

$$\varphi^4 - \varphi^3\left[\mu^2(z^4+z^{-4}) + (\mu^2+3)(z^2+z^{-2}) + 6\right]$$

$$+\varphi^2\left[(\mu^2-1)\mu^{-1}\right]\left[\mu^3(z^6+z^{-6}) + 6\mu(z^4+z^{-4}) + \mu^{-1}(\mu^4+4\mu^2+10)(z^2+z^{-2})\right.$$

$$\left. + \mu^{-3}(2\mu^6+2\mu^4+14\mu^2+2)\right]$$

$$-\varphi\left[\mu^{-1}(\mu^2-1)\right]^4\left[\mu^2(z^4+z^{-4}) + (\mu^2+3)(z^2+z^{-2}) + 6\right] + \mu^{-8}(\mu^2-1)^8 = 0 \quad (4,11)$$

Our object now is to attempt to solve equations (4,6), (4,7), (4,8), and (4,11) simultaneously in such a way as

to obtain ν, ω, ζ, and φ in terms of μ. Before attempting this, however, we first write our equations in terms of the actual variables we wish to work with. Let us take as our variables, $z^2 = e^{(B+G)/2}$, $s = e^{Q}$. Instead of $g = e^{G}$ take z^4b^{-1} and instead of $q = e^{B+Q}$ take sb. Doing this, equations (4,3) and (4,11) remain unaltered while, instead of (4,6) and (4,7), we have

$$\nu - sb - s^{-1}b^{-1} - 2 = 0 \qquad (4,12)$$

$$\omega - z^4b^{-1} - z^{-4}b - 2 = 0 \qquad (4,13)$$

§ 3. In order to solve our equations, we first of all used a rather simple method. This consisted in trying to pick out from the equations, the important terms in each, taking the correct ratios of the terms and then trying to maximise the result by inspection. To make this clearer, we have as our most important terms in each of our equations:

$$\varphi^4 - \varphi^3\mu^2z^4 = 0 \qquad \text{or} \qquad \varphi = \mu^2z^4$$

$$\nu - sb = 0 \qquad\qquad \nu = sb$$

$$\omega = z^4b^{-1}$$

$$\zeta = s$$

To find λ, we first of all take $\varphi/\omega = \mu^2b = \chi$
This should, according to equation (3,33), be maximised with respect to G or z^4b^{-1}.

We have also $\psi = \nu/\zeta = b$ which according to equation

(3,22) should be maximised with respect to Q or s.

Finally, we have $\lambda = \frac{\chi}{\psi}$ to be maximised with respect to B or b - equation (3,15). Taking this ratio, we get $\lambda = \mu^2$ for our first approximation to a solution for λ. We have not yet had to use the maximisation conditions to which we have referred.

For our next approximation, we substitute into equation (4,11), the expression

$$\varphi = \mu^2 z^4 (1 + \varphi_1)$$

and, after cancelling a number of terms, find, as the most important term for φ_1, the expression $4\mu^{-2}z^{-2}$. The corresponding second terms for ω, ϑ, and ψ are obvious and we have eventually

$$\varphi = \mu^2 z^4 (1 + 4\mu^{-2}z^{-2})$$

$$\omega = z^4 b^{-1} (1 + 2z^{-4}b)$$

$$\vartheta = sb (1 + 2s^{-1}b^{-1})$$

$$\psi = s(1 + s^{-2})$$

For χ, we have $\varphi/\omega = \mu^2 b(1 + 4\mu^{-2}z^{-2} - 2z^{-4}b)$ which must be maximised with respect to z^4b^{-1}. This equation can be written

$$\chi = \left[1 + 2\mu^{-4}b^{-1} - 2(\mu^{-2}b^{-1/2} - z^{-2}b^{1/2})^2 \right] \mu^2 b$$

which maximises immediately to $\chi = \mu^2 b[1 + 2\mu^{-4}b^{-1}]$

with the condition $\mu^2 = z^2 b^{-1}$ (4,14)

For ψ , we have χ/ψ to be maximised with respect to s. We find

$$\psi = b\left[1 + 2s^{-1}b^{-1} - s^{-2}\right] = b\left[1 + b^{-2} - (b^{-1} - s^{-1})^2\right]$$

maximising to $\psi = b\left[1 + b^{-2}\right]$

with the condition $b = s$ (4,15)

Finally $\lambda = \chi/\psi$ to be maximised with respect to b. The values above, give $\lambda = \mu^2\left[1 + 2\mu^{-4}b^{-1} - b^{-2}\right]$

$$= \mu^2\left[1 + \mu^{-8} - (\mu^{-4} - b^{-1})^2\right]$$

which maximises to $1 + \mu^{-8}$, with the condition

$$b = \mu^4 \qquad (4,16)$$

Using equations (4,14), (4,15), (4,16), we can find first approximations for our parameters b, z, s, in terms of μ . These, when substituted into our main equations, allow us to find λ as being

$$\lambda = \mu^2\left(1 + \mu^{-8} + 4\mu^{-14} - 4\mu^{-16} + 4\mu^{-20}\right) \qquad (4,17)$$

At this point, we must mention that as a check on our approximate series for λ calculated by any approximate method, we have exact series calculated by Trefftz [16]. We shall discuss Trefftz's series later; but, for the present, we shall merely mention that the series (4,17) does not agree with that given by Trefftz for the body centred lattice at low temperatures. Trefftz gives $28\mu^{-20}$ instead of $4\mu^{-20}$.

It would seem then, that either our method fails at the term in μ^{-20} or, as is more likely, our approximations for our maximisation variables are inadequate at this point. We should have to assume series for b, s, and z and attempt to find the unknown coefficients at each stage by maximisation in order to carry this method further. We decided instead, to use a slightly more sophisticated method.

§ 4 Our four main equations have the form

$$\begin{aligned} \mathcal{L}_1(\varphi, \mu, z^2) &= 0 \\ \mathcal{L}_2(\psi, s) &= 0 \\ \mathcal{L}_3(\vartheta, s, b) &= 0 \\ \mathcal{L}_4(\omega, z^2, b) &= 0 \end{aligned} \qquad (4,18)$$

These have seven unknowns φ, ψ, ϑ, ω, z^2, s, b, for which we should like to develop series in negative powers of μ. To effect a solution, we require other three equations. These we obtain from our maximisation conditions as follows.

We have $\chi = \varphi/\omega$ which has to be maximised by our choice of z. In other words, we must have

$$\omega \frac{\partial \varphi}{\partial z} - \varphi \frac{\partial \omega}{\partial z} = 0$$

Similarly, $\psi = \xi\eta/\zeta$ must be a maximum as a result of correct choice of s, which gives

$$\zeta \frac{\partial \xi\eta}{\partial s} - \xi\eta \frac{\partial \zeta}{\partial s} = 0$$

Finally, $\lambda = \chi/\psi = \varphi\zeta/\omega\xi\eta$ must be maximised with respect to b. This condition yields

$$\omega\xi\eta\varphi \frac{\partial \zeta}{\partial b} + \omega\xi\zeta\eta \frac{\partial \varphi}{\partial b} - \varphi\zeta\omega \frac{\partial \xi\eta}{\partial b} - \varphi\zeta\xi\eta \frac{\partial \omega}{\partial b} = 0$$

or

$$\omega \frac{\partial \xi\eta}{\partial b} + \xi\eta \frac{\partial \omega}{\partial b} = 0$$

since ζ and φ are not explicit functions of b.

It is more convenient to write these equations in terms of the $\mathcal{L}$'s , and we find

$$\begin{aligned}
&\omega \frac{\partial \mathcal{L}_1}{\partial z} \frac{\partial \mathcal{L}_4}{\partial \omega} - \varphi \frac{\partial \mathcal{L}_1}{\partial \varphi} \frac{\partial \mathcal{L}_4}{\partial z} = 0 \\
&\zeta \frac{\partial \mathcal{L}_3}{\partial s} \frac{\partial \mathcal{L}_2}{\partial \zeta} - \xi\eta \frac{\partial \mathcal{L}_3}{\partial \xi\eta} \frac{\partial \mathcal{L}_2}{\partial s} = 0 \qquad (4,19) \\
&\omega \frac{\partial \mathcal{L}_3}{\partial b} \frac{\partial \mathcal{L}_4}{\partial \omega} - \xi\eta \frac{\partial \mathcal{L}_3}{\partial \xi\eta} \frac{\partial \mathcal{L}_4}{\partial b} = 0
\end{aligned}$$

The two sets of equations, (4,18) and (4,19) should be sufficient to enable us to solve the problem. In order to do so, we assume the following series

$$\varphi = \mu^{14}\sum_{i=1}^{\infty}(1+\varphi_i\mu^{-2i})$$

$$b = \mu^{4}\sum_{i=1}^{\infty}(1+b_i\mu^{-2i}) \qquad \omega = \mu^{8}\sum_{i=1}^{\infty}(1+\omega_i\mu^{-2i})$$

$$s = \mu^{4}\sum_{i=1}^{\infty}(1+s_i\mu^{-2i}) \qquad \mathcal{S} = \mu^{8}\sum_{i=1}^{\infty}(1+\mathcal{S}_i\mu^{-2i}) \qquad (4,20)$$

$$z^2 = \mu^{6}\sum_{i=1}^{\infty}(1+z_i\mu^{-2i}) \qquad \zeta = \mu^{4}\sum_{i=1}^{\infty}(1+\zeta_i\mu^{-2i})$$

For convenience, we shall collect together our equations.

Eigenvalue Equations

$$\varphi^4 - \varphi^3\left[\mu^2 z_+^4 + (\mu^2+3)z_+^2 + 6\right] + \varphi^2\mu_-\left[\mu^3 z_+^6 + 6\mu z_+^4 + \mu^{-1}(\mu^4+4\mu^2+10)z_+^2 + \mu^{-3}(2\mu^6+2\mu^4+14\mu^2+2)\right] - \varphi(\mu_-)^4\left[\mu^2 z_+^4 + (\mu^2+3)z_+^2 + 6\right] + (\mu_-)^8 = 0 \qquad (4,11)$$

$$\zeta - s_+ = 0 \qquad (4,8)$$

$$\mathcal{S} - (sb)_+ - 2 = 0 \qquad (4,12)$$

$$\omega - (z^4 b^{-1})_+ - 2 = 0 \qquad (4,13)$$

Maximisation Equations (from equations (4,19))

$$\omega\left\{-\varphi^3\left[4\mu^2 z_-^4 + 2(\mu^2+3)z_-^2\right] + \varphi^2\mu_-\left[6\mu^3 z_-^6 + 24\mu z_-^4 + 2\mu^{-1}(\mu^4+4\mu^2+10)z_-^2\right] - \varphi(\mu_-)^4\left[4\mu^2 z_-^4 + 2(\mu^2+3)z_-^2\right]\right\} - 4\varphi\left\{4\varphi^3 - 3\varphi^2\left[\mu^2 z_+^4 + (\mu^2+3)z_+^2 + 6\right] + 2\varphi\mu_-\left[\mu^3 z_+^6 + 6\mu z_+^4 + \mu^{-1}(\mu^4+4\mu^2+10)z_+^2 + \mu^{-3}(2\mu^6+2\mu^4+14\mu^2+2)\right] - (\mu_-)^4\left[\mu^2 z_+^4 + (\mu^2+3)z_+^2 + 6\right]\right\}\left\{(z^{-4}b)_-\right\} = 0 \qquad (4,21)$$

$$\varphi [sb]_- - \vartheta [s]_- = 0 \qquad (4,22)$$

$$\omega [sb]_- + \vartheta [z^{-4}b]_- = 0 \qquad (4,23)$$

In which, to facilitate calculation, we have multiplied throughout each equation by the unknown variable with respect to which we have differentiated, and where we have used the substitution $p_{\pm} = p \pm p^{-1}$ (for any p).

On making use of the series assumed in (4,20), in these equations, we eventually find, after lengthy and tedious calculation, the following coefficients for the series

$$b = M^4 (1 - 4M^{-6} + 4M^{-8} - 10M^{-12})$$

$$s = M^4 (1 - 4M^{-6} + 4M^{-8} - 18M^{-12})$$

$$z^2 = M^6 (1 - 6M^{-6} + 6M^{-8} - 13M^{-12})$$

$$\varphi = M^{14}\left[1 + \varphi_3^{\times} M^{-6} + (\varphi_4^{\times} + 4) M^{-8} + \varphi_6^{\times} M^{-12} + (\varphi_7^{\times} + 4z_3 + 4) M^{-14} + (\varphi_8^{\times} + 4z_4 + 2) M^{-16} + \varphi_9^{\times} M^{-18} + (\varphi_{10}^{\times} + 4z_6 + 4) M^{-20} + (\varphi_{11}^{\times} + 4z_7) M^{-22} + (\varphi_{12}^{\times} + 4z_8) M^{-24}\right]$$

$$\omega = M^8\left[1 + \omega_3^{\times} M^{-6} + (\omega_4^{\times} + 2) M^{-8} + \omega_6^{\times} M^{-12} + \omega_7^{\times} M^{-14} + (\omega_8^{\times} + 1) M^{-16} + \omega_9^{\times} M^{-18} + \omega_{10}^{\times} M^{-20} + (\omega_{11}^{\times} + 8) M^{-22} + (\omega_{12}^{\times} - 8) M^{-24}\right]$$

$$\vartheta = M^8\left[1 + \vartheta_3^{\times} M^{-6} + (\vartheta_4^{\times} + 2) M^{-8} + \vartheta_6^{\times} M^{-12} + \vartheta_7^{\times} M^{-14} + (\vartheta_8^{\times} + 1) M^{-16} + \vartheta_9^{\times} M^{-18} + \vartheta_{10}^{\times} M^{-20} + (\vartheta_{11}^{\times} + 8) M^{-22} + (\vartheta_{12}^{\times} - 8) M^{-24}\right]$$

$$\zeta = \mu^2\left[1 + \zeta_3^* \mu^{-6} + (\zeta_4^*+1)\mu^{-8} + \zeta_6^* \mu^{-12} + (\zeta_7^* - s_3)\mu^{-14} + (\zeta_8^* - s_4)\mu^{-16}\right.$$
$$\left. + \zeta_9^* \mu^{-18} + (\zeta_{10}^* - s_6 + 16)\mu^{-20} + (\zeta_{11}^* - s_7 - 32)\mu^{-22} + (\zeta_{12}^* - s_8 + 16)\mu^{-24}\right]$$

We shall refer to these equations as equations (4,24).

In these series, we have only taken those for b, s, and z as far as was necessary to obtain the correct coefficients of μ^{-24} in our other series. The starred terms in the elements of the last four series refer to a first approximation solution. That is, if we had approximated our eigenvalue equations by taking, in each case, only the first two important terms as we did in § 3, then our solution would have contained only the starred terms. In calculating λ, the starred terms disappear since at each stage $\varphi^* + \zeta^* - 2\beta^* - \omega^* = 0$

Calculating λ from our series, we find

$$\lambda = \mu^2\left[1 + \mu^{-8} + 4\mu^{-14} - 4\mu^{-16} + 28\mu^{-20} - 60\mu^{-22} + 32\mu^{-24}\right] \qquad (4,25)$$

which deviates from the series given by Trefftz for the partition function per spin of the body centred cubic ferromagnet at low temperatures, in the last term given.

As we shall see, this is a much better result than the corresponding high temperature series which we derived earlier for the same model. The improvement is, however, not surprising since, in this part of the work, we have had three unknown parameters to work with, whereas before

we had none. It is worth mentioning also that, despite the existence of these three maximisation parameters, we must still look upon equation (4,25) as representing, along with equation (4,5), a limiting case of the variational method. These parameters are necessary to distinguish between the two temperature regions in our model, and it is the absence of the other type of unknown parameter, that associated with the lattice interaction energies, which indicates that the present case is a limiting one. We now turn to the other case we wish to consider, that of the face centred cubic lattice with first order interactions between the spins.

§ 5 Face Centred Cubic Ferromagnet at High Temperatures

For the equations appropriate to this problem, we put our first order and third order interactions equal to zero. That is, we put

$$a = e = p = y = x = r = \kappa = 1 \qquad (4,26)$$

in our main equations. This transforms our simple cubic lattice into two superimposed face centred lattices which are independent of each other. The second order interactions of the original simple cubic lattice have become the first order interactions of the face centred lattices. Applying this to equations (3,46), (3,47), (3,48), and (3,49), we get equations appropriate to the region above

the Curie temperature.

Our modification of equation (3,46) can now be reduced by the following operations on its rows and columns: Add column 1 to column 2 and subtract row 2 from row 1. This gives a 3 × 3 determinant which multiplies out to give

$$\varphi^3 - \varphi^2\left[\beta^2(\alpha+\alpha^{-1})^2 + 2(\alpha+\alpha^{-1}+2) + 4\beta^{-2}\right] + 2\varphi\left[\beta^2(\alpha+\alpha^{-1})^2 + 2(\alpha+\alpha^{-1}+2) + 4\beta^{-2}\right](\alpha+\alpha^{-1}-2) - 8(\alpha+\alpha^{-1}-2)^3 = 0 \qquad (4,27)$$

This has the form

$$\varphi^3 - A\varphi^2 + 2AB\varphi - (2B)^3 = 0 \qquad (4,28)$$

where the meanings of A and B are obvious.

A factor of (4,28) is $\varphi = 2B = 2(\alpha^{-1}+\alpha-2)$ which, when taken out of (4,27), leaves

$$\varphi^2 - \varphi\left[k(kb+k^{-1}b^{-1})^2 + 8 + 4k^{-1}\right] + 4\left[bk + b^{-1}k^{-1} - 2\right]^2 = 0 \qquad (4,29)$$

on substituting for α and β. We have put $e^{2K'} = \beta^2 = k$. We shall refer to this equation as $L_1(\varphi, k, b)$.

The equations in ω and $\nu\beta$, i.e. (3,47) and (3,48), after substitution of the condition (4,26), reduce to one equation of the form

$$\omega^2 - 2\omega(b + b^{-1}) + (b + b^{-1})^2 - 4 = 0$$

which has as its largest root

$$\omega = b^{-1} + b + 2 = 2\beta \qquad (4,30)$$

This we shall call $\mathcal{L}_2(\omega, b)$.

Finally, we find that the equation (3,49) reduces to

$$\beta = 2 \qquad (4,31)$$

For λ we now shall have

$$\lambda = \frac{2\varphi}{\omega^2} \qquad (4,32)$$

In equations (4,29), (4,30), and (4,32), we have three equations for the four unknown quantities φ, ω, b, and λ. A fourth equation we obtain from the maximisation condition (3,15) i.e. $\lambda = \chi / \psi$

to be maximised with respect to A' or b.

From (4,32), we see that this is equivalent to the condition

$$2\varphi \frac{\partial \omega}{\partial b} - \omega \frac{\partial \varphi}{\partial b} = 0 \qquad \text{or, in terms of the } \mathcal{L}\text{'s},$$

$$2\varphi \frac{\partial \mathcal{L}_2}{\partial b} \frac{\partial \mathcal{L}_1}{\partial \varphi} - \omega \frac{\partial \mathcal{L}_1}{\partial b} \frac{\partial \mathcal{L}_2}{\partial \omega} = 0 \qquad (4,33)$$

Carrying out the differentiations and substituting into this equation, we find

$$2(b^{-1}-b)\left[2\varphi^2 - \varphi(k^3b^2 + 2k + k^{-1}b^{-2} + 8 + 4k^{-1})\right]$$

$$- \omega\left[2\varphi(k^{-1}b^{-2} - k^3b^2) + (8b^2k^2 - 16bk + 16b^{-1}k^{-1} - 8b^{-2}k^{-2})\right] = 0 \qquad (4,34)$$

In which we have multiplied through by b.

We can now solve equations (4,29), (4,30), (4,32), and

(4,34) for our unknown quantities and develop power series for them in terms of K'. We assume the series

$$k = e^{2K'} = 1 + 2K + 2K^2 + 4/3K^3 + 2/3K^4 + \cdots$$

$$b = \mathcal{B}\left(1 + \sum_i b_i K^i\right)$$

$$\omega = \Omega\left(1 + \sum_i \omega_i K^i\right) \qquad (4,35)$$

$$\varphi = \Phi\left(1 + \sum_i \varphi_i K^i\right)$$

where since the first order interaction energies are zero, we have put K' = K.

Solving by successive approximation, we find for our series

$$b = 1 + 2K + 10K^2 + \frac{172}{3}K^3 + \frac{1090}{3}K^4 + \frac{37084}{15}K^5 + \frac{159124}{9}K^6 + b_7 K^7$$

$$\omega = 4\left[1 + K^2 + 8K^3 + \frac{169}{3}K^4 + 400K^5 + \frac{131642}{45}K^6 + \frac{329968}{15}K^7\right]$$

$$\varphi = 16\left[1 + 5K^2 + 24K^3 + \frac{461}{3}K^4 + 1040K^5 + \frac{66758}{9}K^6 + \frac{5764528}{105}K^7\right]$$

On substituting into (4,32), we find for the partition function per spin for the face centred cubic lattice at high temperatures, the series

$$\lambda = 2\left[1 + 3K^2 + 8K^3 + 34K^4 + 160K^5 + \frac{38391}{45}K^6 + \frac{523096}{105}K^7\right] \qquad (4,36)$$

In deriving this series, we have had one maximisation variable to vary, which represents an improvement upon

the conditions under which the high temperature partition function for the body centred cubic lattice was derived. We might expect then, that the approximation involved in equation (4,36) is better than that in equation (4,5). This is borne out, as we shall see, by a comparison of both equations with the corresponding series derived by Trefftz.

§ 6. Face Centred Cubic Ferromagnet at Low Temperatures.

On making the substitutions (4,26) into the general low temperature equations, we find the following modifications of the various determinants

$\alpha^2\beta^2 z^4 - \varphi$	$\alpha^{-2}\beta^2 z^{-4}$	$2\beta^2$	$4\alpha z^2$	$4\alpha^{-1} z^{-2}$	$4\beta^{-2}$
$\alpha^{-2}\beta^2 z^4$	$\alpha^2\beta^2 z^{-4} - \varphi$	$2\beta^2$	$4\alpha^{-1} z^2$	$4\alpha z^{-2}$	$4\beta^{-2}$
$\beta^2 z^4$	$\beta^2 z^{-4}$	$(\alpha^{-2}+\alpha^2)\beta^2 - \varphi$	$2(\alpha+\alpha^{-1}) z^2$	$2(\alpha+\alpha^{-1}) z^{-2}$	$4\beta^{-2}$
$\alpha\beta^2 z^4$	$\alpha^{-1}\beta^2 z^{-4}$	$(\alpha+\alpha^{-1})\beta^2$	$2z^2(\alpha+1) - \varphi$	$2z^{-2}(\alpha^{-1}+1)$	$4\beta^{-2}$
$\alpha^{-1}\beta^2 z^4$	$\alpha\beta^2 z^{-4}$	$(\alpha+\alpha^{-1})\beta^2$	$2z^2(\alpha^{-1}+1)$	$2z^{-2}(\alpha+1) - \varphi$	$4\beta^{-2}$
$\beta^2 z^4$	$\beta^2 z^{-4}$	$2\beta^2$	$4z^2$	$4z^{-2}$	$4\beta^{-2} - \varphi$

$= 0$

$$\begin{vmatrix} bg-\omega & 1 & 1 & b^{-1}g^{-1} \\ g & b^{-1}-\omega & b & g^{-1} \\ g & b & b^{-1}-\omega & g^{-1} \\ b^{-1}g & 1 & 1 & bg^{-1}-\omega \end{vmatrix} = 0 \qquad \begin{vmatrix} bq-\chi & 1 & 1 & b^{-1}q^{-1} \\ q & b^{-1}-\chi & b & q^{-1} \\ q & b & b^{-1}-\chi & q^{-1} \\ b^{-1}q & 1 & 1 & bq^{-1}-\chi \end{vmatrix} = 0$$

$$\begin{vmatrix} s-\zeta & 1 \\ 1 & s^{-1}-\zeta \end{vmatrix} = 0$$

In the determinants in ω and χ, by subtracting column 3 from column 2 and adding row 2 to row 3, we can reduce to order 3 X 3 and on multiplying out, we find

$$\chi^3 - \chi^2\left[b(q+q^{-1}) + b + b^{-1}\right] + \chi\left[(q+q^{-1})(b^2-1) + b^2 - b^{-2}\right]$$
$$-\left[b(b^2-3) - b^{-1}(b^{-2}-3)\right] = 0 \qquad (4,37)$$

$$\omega^3 - \omega^2\left[b(g+g^{-1}) + b + b^{-1}\right] + \omega\left[(g+g^{-1})(b^2-1) + b^2 - b^{-2}\right]$$
$$-\left[b(b^2-3) - b^{-1}(b^{-2}-3)\right] = 0 \qquad (4,38)$$

For the 6 X 6 determinant in φ, there is no obvious method of effecting a reduction of the order and so the spur method must be used to multiply it out as it stands. This is a rather tedious procedure involving the evaluation of 57 subsidiary determinants. Since the equation is

so lengthy, we shall defer writing it out until page 109 where we shall collect together all the equations necessary to solve this part of the problem. For the sake of continuity, however, we shall refer to this equation as (4,39) - it is a function of φ , k, b, and z^2.

Before trying to solve our equations, we must again consider our variables. In order to save work, we shall leave equation (4,39) as it stands, and instead of q and g in equations (4,37) and (4,38), take respectively cs and z^4c^{-1}, where $c = e^B$. Our equations then become

$$\varphi^3 - \varphi^2\left[b\,(cs+c^{-1}s^{-1})+b+b^{-1}\right] + \varphi\left[(cs+c^{-1}s^{-1})(b^2-1)+b^2-b^{-2}\right] - \left[b(b^2-3) - b^{-1}(b^{-2}-3)\right] = 0 \qquad (4,40)$$

$$\omega^3 - \omega^2\left[b(z^4c^{-1}+z^{-4}c)+b+b^{-1}\right] + \omega\left[(z^4c^{-1}+z^{-4}c)(b^2-1)+b^2-b^{-2}\right] - \left[b(b^2-3) - b^{-1}(b^{-2}-3)\right] = 0 \qquad (4,41)$$

and

$$\varphi - s - s^{-1} = 0 \qquad (4,42)$$

§ 7. As in the case of the body centred lattice, we first of all attempted to solve our four eigenvalue equations without deriving maximisation equations. This was in order to find the multiplying factors of our various series to be derived for our unknown quantities (9 in number, if we include λ) and in order to see how far a first approximation for our maximisation variables would carry us. The method used was that of paragraph 3 of this chapter

although, in the present case, it was more difficult to apply in view of (a) the complexity of the equation in φ and (b) the fact that the most important terms in the equations in ω and ν after the first, (bcs and $bc^{-1}z^4$ respectively) are not immediately obvious. (On substituting expressions of the form $\nu_\omega = b^{cs}_{c^{-1}z^4}(1 + \nu_{\omega_1})$ into (4,40) and (4,41) it is found that a number of terms cancel out.)

Instead of describing this procedure, which in any case is adequately described in § 3, it is of more interest to mention another method of finding the successive approximations for φ, ω, ν and ζ. This method is worth outlining, since it refers back to the more physical side of the problem.

Going back to the stage in the development of the problem i.e. in the development of our main equations, where we had reduced our original lattice to subsidiary ones involving strips of spins, we have (equations (3,23) and (3,34))

$$\nu^m = \sum \exp\left[A' \sum (\mu_i\mu_{i+1} + \mu_i'\mu_{i+1}) + \frac{\beta+\alpha}{2} \sum (\mu_i + \mu_i')\right]$$

$$\omega^m = \sum \exp\left[A' \sum (\mu_i\mu_{i+1} + \mu_i'\mu_{i+1}) + \frac{\alpha}{2} \sum (\mu_i + \mu_i')\right]$$

in which we have put first order interactions equal to zero and where we have, in (3,34), put E' = A'.(see page 68).

These are partition functions of two linear chains of

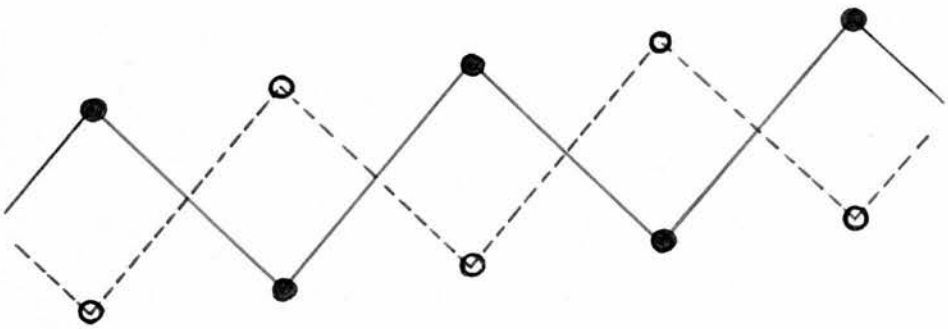


Figure 11.

Figure 12.

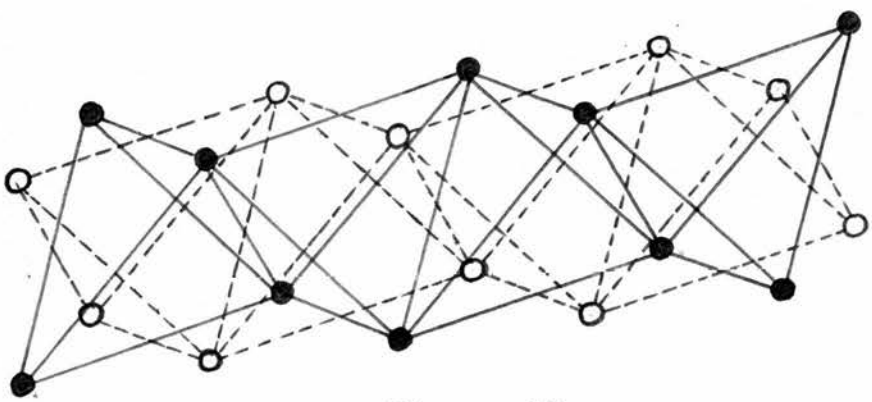


Figure 13.

spins each divided so as to form two independent interpenetrating lattices as shown in figure 11. (The two lattices are shown respectively as dots and circles.)

We have also, from the equation preceding (3,25)

$$\gamma^{M} = \sum \exp\left[Q \sum \mu_i \right]$$

which is the partition function of a chain of spins looked upon as being two superposed chains of the form of those in figure 12 - with nearest neighbour interactions again equal to zero.

Finally, from equation (3,36), we have

$$\varphi^{M} = \sum \exp\Big[\tfrac{K'+A'}{2} \sum (\mu_i \mu_{i+1} + \mu_i \mu_{i+1}^{*} + \mu_{i+1}' \mu_i^{*} + \mu_{i+1}^{*} \mu_i^{*} + \mu_i' \mu_{i+1}' + \mu_{i+1}' \mu_i + \mu_i' \mu_{i+1}^{*} + \mu_{i+1}^{*} \mu_i^{*}) + K' \sum (\mu_i' \mu_i^{*} + \mu_i^{*} \mu_i) + \tfrac{B+G}{4} \sum (\mu_i + \mu_i' + \mu_i^{*} + \mu_i^{*}) \Big]$$

- the partition function of four strips of spins, with first and third order interactions equal to zero - figure 13 shows this.

These four equations have the form of series; the summations over the whole of the right hand sides of the equations indicating that each term of each series arises by considering a unique number of spins in a particular arrangement, to be 'wrongly' orientated with respect to the orientation at absolute zero. In fact, by considering the successive stages in the heating up of the various one dimensional crystals of figures 11 to 13, from the point

of view of spin orientation, we can build up their various partition functions.

At the absolute zero of temperature, for instance, all our spins will be parallel to each other; this state of the crystals will give us our first terms in the series for the partition functions. These are, after taking the m^{th} root in each case,

$$\vartheta = \exp[2A' + (B+Q)] = bcs$$

$$\omega = \exp[2A' + G] = bz^4c^{-1}$$

$$\psi = \exp Q = s$$

$$\varphi = \exp[4(K'+A') + 2K' + (B+G)] = k^3b^2z^4$$

These are the same expressions that we find by taking first approximations from the eigenvalue equations for ϑ, ω, ψ and φ in § 6 as we did in § 3. Taking the appropriate ratios, we find, as our first approximation for λ the value

$$\lambda = k^3$$

For our second approximation we consider one spin in each crystal to be antiparallel to the general spin direction. However, since in each crystal we have two independent lattices, we must consider one spin from each lattice reversed. Doing this, we find

$$\vartheta^m = \exp\left[2mA' + (B+Q)m\right] + 2m\exp\left[(2m-4)A' + (2m-2)\frac{B+Q}{2}\right]$$

$$\omega^m = \exp\left[2mA' + Gm\right] + 2m\exp\left[(2m-4)A' + (2m-2)G/2\right]$$

$$\varphi^m = \exp mQ + m\exp Q(m-2)$$

$$\Phi^m = \exp\left[(K'+A')4m + 2mK' + (B+G)m\right]$$
$$+ 4m\exp\left[\frac{K'+A'}{2}(8m-8) + K'(2m-2) + \frac{B+G}{4}(4m-2)\right]$$

or, taking the m^{th} root

$$\vartheta = bcs\,(1 + 2b^{-2}c^{-1}s^{-1})$$

$$\omega = bz^4c^{-1}(1 + 2b^{-2}z^{-4}c)$$

$$\varphi = s(1 + s^{-2})$$

$$\Phi = k^3b^2z^4(1 + 4k^{-3}b^{-2}z^{-2})$$

Again, by using the method of § 3, we should find these equations as second approximations from our eigenvalue equations.

We can now take $\chi = \Phi/\omega$ and by completing the square, maximise the resulting expression. This gives

$$\chi = k^3c^2b^{-1}(b^2c^{-1} + 2k^{-6}c^{-2})$$

with the condition $z^2 = k^3c$ (compare (4,14)) (4,43)

Similarly, $\psi = \vartheta/\varphi$ yields

$$\psi = bc\,(1 + b^{-4}c^{-2})$$

with the condition $s = b^2c$ (compare (4,15)) (4,44)

Finally, $\lambda = \chi/\psi = k^3[1 + k^{-12}]$

and $c = k^6 b^{-2}$ (compare (4,16)) (4,45)

The conditions (4,43) - (4,45) unfortunately are not sufficient to enable us to find first approximations for b, c and z in terms of k and so we must carry our method a stage further.

Using either the method of this paragraph and considering two adjacent spins inverted in each of our subsidiary lattices, or the method of § 3, we find as our next approximations for our subsidiary partition functions:

$$\mathcal{B} = bcs(1 + 2b^{-2}c^{-1}s^{-1} + 2b^{-2}c^{-2}s^{-2})$$

$$\omega = bz^4c^{-1}(1 + 2b^{-2}z^{-4}c + 2b^{-2}z^{-8}c^2)$$

$$\mathcal{G} = s(1 + s^{-2})$$

$$\varphi = k^3b^2z^4(1 + 4k^{-3}b^{-2}z^{-2} + 2k^{-4}b^{-4}z^{-4} + 8k^{-5}b^{-3}z^{-4})$$

Proceeding as before, we find

$$\chi = k^3bc(1 + 4k^{-3}b^{-2}z^{-2} - 2b^{-2}cz^{-4} - 2b^{-2}c^2z^{-8} + 2k^{-4}b^{-4}z^{-4} + 8k^{-5}b^{-3}z^{-4})$$

$$\psi = bc(1 + 2b^{-2}c^{-1}s^{-1} - s^{-2} + 2b^{-2}c^{-2}s^{-2})$$

Taking the ratio of χ to ψ and using the conditions (4,43) - (4,45), we find

$$\lambda = k^3[1 + k^{-12} + 2k^{-24}(k^2 + 4kb - 2b^2)]$$

$$= k^3[1 + k^{-12} + 2k^{-24}(3k^2 - 2(k-b)^2)]$$

On maximising this equation with respect to b, we see that this approximation for λ gives

$$\lambda = k^3 \left[1 + k^{-12} + 6 k^{-22} \right]$$

with the additional condition $k = b$ (4,46)

The conditions (4,43) - (4,46) enable us to find first approximations for our maximisation variables in terms of k; these will be the multiplying factors of the series for these variables which we shall eventually develop.

Using these first approximations in our eigenvalue equations, we find for λ the series

$$\lambda = k^3 \left[1 + k^{-12} + 6 k^{-22} - 6 k^{-24} + 8 k^{-30} + 12 k^{-32} \right]$$

which deviates from the exact series in the last term. Trefftz gives 42 instead of 12 as the coefficient of k^{-32}.

In order to find out whether the failure of our series is due to our approximations for b, c_s and z, we shall use our maximisation conditions to obtain other four equations which, in conjunction with our eigenvalue equations, will enable us to solve our problem.

§8 We can write our eigenvalue equations as

$$\begin{aligned} L_1(\varphi, k, b, z) &= 0 \qquad L_3(\vartheta, b, c, s) = 0 \\ L_2(\omega, b, z, c) &= 0 \qquad L_4(\psi, s) = 0 \end{aligned} \tag{4,47}$$

Since $\chi = \varphi/\omega$ must be maximised with respect to z, we have the condition $\omega \frac{\partial \varphi}{\partial z} - \varphi \frac{\partial \omega}{\partial z} = 0$

Similarly, $\psi = \eta/\psi$ must be maximised with respect to s giving $\psi \frac{\partial \eta}{\partial s} - \eta \frac{\partial \psi}{\partial s} = 0$

Finally, the variables c and b must be chosen so as to make λ a maximum. Thus we have

$$\omega \frac{\partial \eta}{\partial c} + \eta \frac{\partial \omega}{\partial c} = 0$$

and

$$\omega \eta \frac{\partial \varphi}{\partial b} - \varphi \eta \frac{\partial \omega}{\partial b} - \varphi \omega \frac{\partial \eta}{\partial b} = 0$$

The other terms go out in these last two equations since φ is not an explicit function of c and ψ is not a function of b or c.

It is more convenient to write our maximisation equations in terms of the L's and we find, on making the appropriate substitutions

$$\omega \frac{\partial L_1}{\partial z} \frac{\partial L_2}{\partial \omega} - \varphi \frac{\partial L_2}{\partial z} \frac{\partial L_1}{\partial \varphi} = 0$$

$$\psi \frac{\partial L_3}{\partial s} \frac{\partial L_4}{\partial \psi} - \eta \frac{\partial L_4}{\partial s} \frac{\partial L_3}{\partial \eta} = 0$$

$$\eta \frac{\partial L_2}{\partial c} \frac{\partial L_3}{\partial \eta} + \omega \frac{\partial L_3}{\partial c} \frac{\partial L_2}{\partial \omega} = 0 \qquad (4,48)$$

$$\eta\omega \frac{\partial L_1}{\partial \varphi} \frac{\partial L_2}{\partial \omega} \frac{\partial L_3}{\partial \eta} - \varphi\eta \frac{\partial L_2}{\partial \varphi} \frac{\partial L_1}{\partial \varphi} \frac{\partial L_3}{\partial \eta} - \varphi\omega \frac{\partial L_3}{\partial \varphi} \frac{\partial L_1}{\partial \varphi} \frac{\partial L_3}{\partial \omega} = 0$$

The two sets of equations, (4,47) and (4,48), are sufficient in principle, to allow us to find series for our variables in terms of k^{-1}. We assume series of the form

$$b = k \sum_i (1 + b_i k^{-2i}) \qquad \varphi = k^{19} \sum_i (1 + \varphi_i k^{-2i})$$

$$c = k^4 \sum_i (1 + c_i k^{-2i}) \qquad \omega = k^{11} \sum_i (1 + \omega_i k^{-2i})$$

$$s = k^6 \sum_i (1 + s_i k^{-2i}) \qquad \varpi = k^{11} \sum_i (1 + \varpi_i k^{-2i}) \qquad (4,49)$$

$$z = k^7 \sum_i (1 + z_i k^{-2i}) \qquad \psi = k^6 \sum_i (1 + \psi_i k^{-2i})$$

By successive approximation into our two sets of equations, we can find the various coefficients of these series. We shall now write out the equations.

Eigenvalue Equations:

(In these equations and in the following ones, we again use the notation $p_\pm = p \pm p^{-1}$).

$$\psi - s_+ = 0 \qquad (4,42)$$

$$-\omega^3 + \omega^2[b(z^4c^{-1})_+ + b_+] - \omega[(b^2-1)(z^4c^{-1})_+ + b_-^2]$$
$$+[b_-^3 - 3b_-] = 0 \qquad (4,41)$$

$$-\varpi^3 + \varpi^2[b(cs)_+ + b_+] - \varpi[(b^2-1)(cs)_+ + b_-^2]$$
$$+[b_-^3 - 3b_-] = 0 \qquad (4,40)$$

$$\varphi^6 - \varphi^5\{k^3b^2z_+^4+2(kb+1)z_+^2+k(k^2b^2)_+ +4k^{-1}\}$$
$$+ \varphi^4[bk-1]\{2k^3b^2z_+^6+[4(kb+1)+k^2(kb+1)(k^2b^2+1)]z_+^4+[2k^{-1}b^{-2}(k^4b^4$$
$$+2k^3b^3+2k^2b^2+2kb+2)+2b^{-1}(k^3b^3+k^2b^2+kb-1)+8k^{-1}]z_+^2+8k^{-1}b^{-1}(kb+1)^2$$
$$+k^{-2}b^{-4}(kb+1)(k^2b^2+1)(k^4b^4+1)\}$$
$$- \varphi^3\{kb-1\}^3\{[2k^3b(kb+1)+8]z_+^6+4k^{-1}b^{-2}(kb+1)[2k^2b^2+4kb+3]z_+^4$$
$$+[2k^{-2}b^{-4}(kb+1)(2k^5b^5+3k^4b^4+6k^3b^3+6k^2b^2+7kb+3)+8k^{-2}b^{-2}(2k^2b^2$$
$$+3kb+4)]z_+^2+4k^{-3}b^{-4}(kb+1)[2k^4b^4+4k^3b^3+6k^2b^2+4kb+2]$$
$$+16k^{-3}b^{-2}(kb+1)+k^{-3}b^{-6}(kb+1)^3(k^2b^2+1)^3\}$$
$$+ \varphi^2\{(kb-1)^5(kb+1)k^{-1}b^{-1}\}\{8kz_+^6+[16k^{-1}b^{-1}+4k^{-1}b^{-2}(kb+1)(k^2b^2$$
$$+kb+2)]z_+^4+[8k^{-2}b^{-3}(2k^3b^3+3k^2b^2+5kb+1)+2k^{-1}b^{-4}(kb+1)^2(1$$
$$+k^2b^2)^2]z_+^2+4k^{-4}b^{-5}(kb+1)^4(k^2b^2+1)+16k^{-3}b^{-3}(k^2b^2+1)$$
$$+4k^{-3}b^{-5}(k^2b^2-1)(k^4b^4+2k^3b^3+4k^2b^2+2kb+1)\}$$
$$- \varphi\{4(kb-1)^8(kb+1)^2k^{-4}b^{-4}\}\{4kz^4+2b^{-2}(kb+1)(k^2b^2+kb+2)z_+^2$$
$$+k^{-1}b^{-4}(kb+1)^2(k^2b^2+1)^2+4k^{-1}b^{-2}(k^2b^2+1)\}$$
$$+16k^{-6}b^{-8}(kb-1)^{12}(kb+1)^4 = 0 \qquad (4,39)$$

<u>Maximisation Equations</u> (from equations (4,48))

(In these equations we shall multiply each one through by the maximisation variable with respect to which we differentiate - as was done in the corresponding equations for the body centred case)

$$\mathcal{S}\{\omega^2b(z^{-4}c)_- - \omega(b^2-1)(z^{-4}c)_-\}\{-3\mathcal{S}^2+2\mathcal{S}[b(cs)_+ +b_+]$$
$$-[(b^2-1)(cs)_+ +b_-^2]\} +$$
$$\omega\{\mathcal{S}^2b(cs)_- - \mathcal{S}(b^2-1)(cs)_-\}\{-3\omega^2+2\omega[b(z^4c^{-1})_+ +b_+]$$
$$-[(b^2-1)(z^4c^{-1})_+ +b_-^2]\} = 0 \qquad (4,50)$$

$$\omega\{-\varphi^5[k^3b^2z_-^4+(kb+1)z_-^2]+\varphi^4[3(k^4b^3-k^3b^2)z_-^6+(k^6b^4+4k^2b^2-k^2-4)z_-^4+(2k^4b^3+k^3b^2-2k^{-1}b^{-2}-2k-4k^{-1}+b^{-1}+4b)z_-^2]-\varphi^3[3(k^7b^5-2k^6b^4+2k^4b^2+4k^3b^3-k^3b-12k^2b^2+12kb-4)z_-^6+(8k^5b^4-20k^3b^2-8k^2b+24k+8b^{-1}-12k^{-1}b^{-2})z_-^4+(2k^7b^5-k^6b^4-2k^4b^2-k^3b-2k^2+8b^{-2}-k^{-1}b^{-3}-3k^{-2}b^{-4}+8k^3b^3-12k^2b^2+4kb-20+36k^{-1}b^{-1}-16k^{-2}b^{-2})z_-^2]+\varphi^2[3(4k^6b^5-16k^5b^4+20k^4b^3-20k^2b+16k-4b^{-1})z_-^6+(16k^4b^4-64k^3b^3+80k^2b^2-80+64k^{-1}b^{-1}-16k^2b^{-2}+4k^7b^6-8k^6b^5+8k^3b^2+16k^2b-32k+20k^{-1}b^{-2}-8k^{-2}b^{-3})z_-^4+(8k^6b^5-20k^5b^4+12k^4b^3-16k^3b^2+44k^2b-8k-60b^{-1}+48k^{-1}b^{-2}-4k^{-2}b^{-3}-4k^{-3}b^{-4}+k^{10}b^7-2k^9b^6+2k^7b^4-3k^6b^3+4k^5b^2-4k^3+3k^2b^{-1}-2kb^{-2}+2k^{-1}b^{-4}-k^{-2}b^{-5})z_-^2]-\varphi[4(kb-1)^8(kb+1)^2k^{-4}b^{-4}][4kz_-^4+(k^3b+2k^2+3kb^{-1}+2b^{-2})z_-^2]\}\{-3\omega^2+2\omega[bc^{-1}z^4+bz^{-4}c+b+b^{-1}]-[b^2c^{-1}z^4+b^2c\,z^{-4}-z^4c^{-1}-z^{-4}c+b^2-b^{-2}]\}-\varphi\{\omega^2[bz^4c^{-1}-bz^{-4}c]-\omega[b^2z^4c^{-1}-b^2z^{-4}c-z^4c^{-1}+z^{-4}c]\}\{6\varphi^5-5\varphi^4[k^3b^2z_+^4+2(kb+1)z_+^2+k^3b^2+k^{-1}b^{-2}+4k^{-1}]+4\varphi^3[(2k^4b^3-2k^3b^2)z_+^6+\cdots\cdots\cdots+(8k^2b^2+8kb-8-8k^{-1}b^{-1}+k^6b^4-k^{-2}b^{-4})]-3\varphi^2[(2k^7b^5-4k^6b^4+4k^4b^2+8k^3b^3-2k^3b-24k^2b^2+24kb-8)z_+^6\cdots\cdots\cdots+32k^{-2}b^{-1}-16k^{-3}b^{-2}+k^9b^6-3k^5b^2+3kb^{-2}-k^{-3}b^{-6}]+2\varphi[(8k^6b^5-32k^5b^4+40k^4b^3-40k^2b+32k-8b^{-1})z_+^6+\cdots\cdots+28b^{-2}-8k^{-1}b^{-3}-8k^{-3}b^{-5}+4k^{-4}b^{-6}]-[4(kb-1)^8(kb+1)^2k^{-4}b^{-4}][4kz_+^4+\cdots\cdots\cdots+k^{-1}b^{-4}+4k+4k^{-1}b^{-2}]\}=0 \qquad (4,51)$$

In this equation the coefficients of φ^3, φ^2, φ, and the term without φ, occurring in lines 14, 16, 17 and 19 have not been completed since they are the same coefficients that occur in equation (4,39) for φ^4, φ^3, etc.

$$\}\{\varkappa^2b(cs)_--\varkappa(b^2-1)(cs)_-\}+\varkappa s_-\{-3\varkappa^2+2\varkappa[b(cs)_++b_+]-[(b^2-1)(cs)_++b_-^2]\}=0 \qquad (4,52)$$

$$\omega\vartheta\{-\varphi^5[2k^3b^2z_+^4+2kbz_+^2+2k^3b^2-2k^{-1}b^{-2}]+\varphi^4[(6k^4b^3-4k^3b^2)z_+^6$$
$$+(4k^6b^4+8k^2b^2)z_+^4+(12k^4b^3+4k^3b^2+8k^{-1}b^{-2}-2b^{-1}+8b)z_+^2+(16k^2b^2$$
$$+8kb+8k^{-1}b^{-1}+4k^6b^4+4k^{-2}b^{-4})]-\varphi^3[(10k^7b^5-16k^6b^4+8k^4b^2+24k^3b^3$$
$$-2k^3b-48k^2b^2+24kb)z_+^6+(32k^5b^4-40k^3b^2-8k^2b-8b^{-1}+24k^{-1}b^{-2})z_+^4$$
$$+(20k^7b^5-8k^6b^4-8k^4b^2-2k^3b-32b^{-2}+6k^{-1}b^{-3}+24k^{-2}b^{-4}+48k^3b^3$$
$$-48k^2b^2+8kb-72k^{-1}b^{-1}+64k^{-2}b^{-2})z_+^2+32k^5b^4-16k^3b^2-16k^2b-16b^{-1}$$
$$-16k^{-1}b^{-2}+32k^{-3}b^{-4}+32kb^2-32b-32k^{-2}b^{-1}+32k^{-3}b^{-2}+6k^9b^6-6k^5b^2$$
$$-6kb^{-2}+6k^{-3}b^{-6}]+\varphi^2[(40k^6b^5-128k^5b^4+120k^4b^3-40k^2b+8b^{-1})z^6_+$$
$$+(64k^4b^4-192k^3b^3+160k^2b^2-64k^{-1}b^{-1}+32k^{-2}b^{-2}+24k^7b^6-40k^6b^5$$
$$+16k^3b^2+16k^2b-40k^{-1}b^{-2}+24k^{-2}b^{-3})z_+^4+(80k^6b^5-160k^5b^4+72k^4b^3$$
$$-64k^3b^2+88k^2b+120b^{-1}-192k^{-1}b^{-2}+24k^{-2}b^{-3}+32k^{-3}b^{-4}+14k^{10}b^7$$
$$-24k^9b^6+16k^7b^4-18k^6b^3+16k^5b^2-6k^2b^{-1}+8kb^{-2}-16k^{-1}b^{-4}$$
$$+10k^{-2}b^{-5})z_+^2+24k^7b^6-64k^5b^4+40k^3b^2+40k^{-1}b^{-2}-64k^{-3}b^{-4}+24k^{-5}b^{-6}$$
$$+64k^4b^4-192k^3b^3+192k^2b^2-64kb-64k^{-1}b^{-1}+192k^{-2}b^{-2}-192k^{-3}b^{-3}$$
$$+64k^{-4}b^{-4}+24k^8b^6-40k^7b^5-24k^5b^3+56k^4b^2+16k^3b-16kb^{-1}-56b^{-2}$$
$$+24k^{-1}b^{-3}+40k^{-3}b^{-5}-24k^{-4}b^{-6}]-\varphi[32(kb-1)^7(kb+1)^2k^{-3}b^{-3}$$
$$+8(kb-1)^8(kb+1)k^{-3}b^{-3}-16(kb-1)^8(kb+1)^2k^{-4}b^{-4}][4kz_+^4+(2k^3b$$
$$+4k^2+6kb^{-1}+4b^{-2})z_+^2+k^5b^2+2k^4b+3k^3+4k^2b^{-1}+3kb^{-2}+2b^{-3}+k^{-1}b^{-4}$$
$$+4k+4k^{-1}b^{-2}]-\varphi[4(kb-1)^8(kb+1)^2k^{-4}b^{-4}][(2k^3b-6kb^{-1}-8b^{-2})z_+^2$$
$$+2k^5b^2+2k^4b-4k^2b^{-1}-6kb^{-2}-6b^{-3}-4k^{-1}b^{-4}-8k^{-1}b^{-2}]$$
$$-128k^{-6}b^{-8}(kb-1)^{12}(kb+1)^4+192k^{-5}b^{-7}(kb-1)^{11}(kb+1)^4$$
$$+64k^{-5}b^{-7}(kb-1)^{12}(kb+1)^3\}\{-3\omega^2+2\omega[bz^4c^{-1}+bz^{-4}c+b+b^{-1}]$$
$$-[b^2z^4c^{-1}+b^2z^{-4}c-z^4c^{-1}-z^{-4}c+b^2-b^{-2}]\}\{-3\vartheta^2+2\vartheta[bcs+bc^{-1}s^{-1}$$
$$+b+b^{-1}]-[b^2cs+b^2c^{-1}s^{-1}-cs-c^{-1}s^{-1}+b^2-b^{-2}]\}-\varphi\vartheta\{\omega^2[z^4c^{-1}b$$
$$+bz^{-4}c+b-b^{-1}]-\omega[2b^2z^4c^{-1}+2b^2z^{-4}c+2b^2+2b^{-2}]+[3b^3-3b+3b^{-3}$$

$$-3b^{-1}]\}\{-3\vartheta^2+2\vartheta[bcs+bc^{-1}s^{-1}+b+b^{-1}]-[b^2cs+b^2c^{-1}s^{-1}-cs-c^{-1}s^{-1}$$
$$+b^2-b^{-2}]\}\{6\varphi^5-5\varphi^4[k^3b^2z_+^4+ \cdots\cdots +4k^{-1}]+4\varphi^3[(2k^4b^3$$
$$+2k^3b^2]z_+^6+ \cdots\cdots +(8k^2b^2+8kb-8-8k^{-1}b^{-1}+k^6b^4-k^{-2}b^{-4})]$$
$$-3\varphi^2[\cdots\cdots -3k^5b^2+3kb^{-2}-k^{-3}b^{-6}]+2\varphi[(\cdots\cdots -8k^{-3}b^{-5}$$
$$+4k^{-4}b^{-6}]-[4(kb-1)^8(kb+1)^2k^{-4}b^{-4}][4kz_+^4+ \cdots\cdots +4k^{-1}b^{-2}]\}$$
$$-\varphi\omega\{\vartheta^2[bcs+bc^{-1}s^{-1}+b-b^{-1}]-\vartheta[2b^2cs+2b^2c^{-1}s^{-1}+2b^2+2b^{-2}]$$
$$+[3b^3-3b+3b^{-3}-3b^{-1}]\}\{-3\omega^2+2\omega[bz^4c^{-1}+bz^{-4}c+b+b^{-1}]-[b^2z^4c^{-1}$$
$$+b^2z^{-4}c-z^4c^{-1}-z^{-4}c+b^2-b^{-2}]\}\{6\varphi^5-5\varphi^4[k^3b^2z_+^4+ \cdots\cdots$$
$$+4k^{-1}]+4\varphi^3[(2k^4b^3+2k^3b^2)z_+^6+ \cdots\cdots +(8k^2b^2+8kb-8-8k^{-1}b^{-1}$$
$$+k^6b^4-k^{-2}b^{-4})]-3\varphi^2[\cdots\cdots -3k^5b^2+3kb^{-2}-k^{-3}b^{-6}]+2\varphi[\cdots\cdots$$
$$-8k^{-3}b^{-5}+4k^{-4}b^{-6}]-[4(kb-1)^8(kb+1)^2k^{-4}b^{-4}][4kz_+^4+ \cdots\cdots,\cdots$$
$$+4k^{-1}b^{-2}]\} = 0 \qquad (4,53)$$

The remarks at the end of equation (4,51) apply also to this equation.

The solution of these equations proved to be a very lengthy business; however, we eventually arrived at the following series:

$$b = k(1-10k^{-10}+12k^{-12}+60k^{-18}+\beta k^{-20})$$

$$c = k^4(1+16k^{-10}-20k^{-12}-90k^{-18}+\gamma k^{-20})$$

$$s = k^6(1-6k^{-10}+6k^{-12}+30k^{-18}+\sigma k^{-20})$$

$$z^2 = k^7(1+13k^{-10}-17k^{-12}-75k^{-18}+\chi k^{-20})$$

$$\varphi = k^3 b^2 z^4 \left[1 + 4k^{-12} + 38 k^{-22} - 32 k^{-24} - 172 k^{-30} + (76 - 8\beta - 4\lambda) k^{-32} \right]$$

$$\omega = b z^4 c^{-1} \left[1 + 2k^{-12} + 22 k^{-22} - 21 k^{-24} - 120 k^{-30} + (2 + 2\gamma - 4\beta - 4\chi) k^{-32} \right]$$

$$\vartheta = bcs \left[1 + 2k^{-12} + 22 k^{-22} - 21 k^{-24} - 120 k^{-30} + (2 - 4\beta - 2\sigma - 2\gamma) k^{-32} \right]$$

$$\psi = s \left[1 + k^{-12} + 12 k^{-22} - 12 k^{-24} - 60 k^{-30} - 26 k^{-32} \right]$$

In which we have omitted those parts of each term - starred in equations (4,24) - which cancel out in the calculation of λ. These are equations (4,54).

Taking the appropriate ratios of these quantities, we find

$$\lambda = k^3 \left[1 + k^{-12} + 6 k^{-22} - 6 k^{-24} + 8 k^{-30} + 72 k^{-32} \right] \qquad (4,55)$$

where the undetermined quantities, β, γ, σ and χ, have cancelled out. This series deviates from that given by Trefftz, in the last term.

By comparing this series for λ with equation (4,25), we see that, as we might expect, the series for the face centred lattice is better than that for the body centred lattice. This is again due to the fact that we have an increased number of maximisation parameters - resulting from the improvement in our approximations in the face centred case.

§9. Simple Cubic Ferromagnet at High Temperatures

The purpose of this section is to show that from our general equations, we can derive the results obtained by Martin and ter Haar for the simple cubic case.

In our general high temperature equations - (3,46) to (3,49) - we put our second and third order interactions equal to zero i.e.

$$\alpha = \beta = a' = \mu = 0 \qquad (4,56)$$

This procedure leaves us with one simple cubic lattice involving only first order interactions between the spins. The resulting determinants are

$y^2+y^{-2}-\frac{\varphi}{x^2}$	2	$4(y+y^{-1})$	4	$=0$
2	$y^2+y^{-2}-\frac{\varphi}{x^{-2}}$	$4(y+y^{-1})$	4	
$y+y^{-1}$	$y^{-1}+y$	$y^2+y^{-2}+6-\varphi$	$2(y+y^{-1})$	
2	2	$4(y+y^{-1})$	$y^2+y^{-2}+2-\varphi$	

$a(\rho^{-1}+\rho)-2\beta$	a^{-1}	$=0$
$4a$	$a^{-1}(\rho^{-1}+\rho)-2\beta$	

$a(e^{-1}+e)-\omega$	a^{-1}	$=0$
$4a$	$a^{-1}(e^{-1}+e)-\omega$	

$\tau-\gamma$	τ^{-1}	$=0$
τ^{-1}	$\tau-\gamma$	

Since these determinants are the same as those given by Martin [17]) it follows that we shall find the result that he gives.

Before giving this result, however, we should first mention that, since we have three maximisation parameters in excess of the magnetisation maximisation parameters, we should expect an even better approximation to Trefftz's exact series than we found in either of the other two cases. This is, in fact, what we find. Martin and ter Haar give for the partition function per spin at high temperatures, the series

$$\lambda_{/2} = 1 + \tfrac{3}{2}K^2 + \tfrac{31}{8}K^4 + \tfrac{5461}{240}K^6 + \tfrac{2180851}{13440}K^8 \qquad (4,57)$$

which deviates from the exact series in the last term.

So far, the low temperature series for the simple cubic case has not been calculated, although it has been verified that the equations following from the present treatment check with the unpublished equations for this case derived by Martin and ter Haar.

V

DISCUSSION OF APPROXIMATE METHODS OTHER THAN THE VARIATIONAL METHOD

§ 1. In order to gain some idea of the relative power of the variational method as compared to other approximate methods, we shall calculate series for the partition function per spin of our two three dimensional models on the basis of these other methods. As has been mentioned in our introduction, these series do not appear previously to have been calculated; consequently, we shall give a somewhat detailed account of our derivations. This chapter will also serve to complete the historical background to our work, outlined in paragraph 3 of the introduction. Before proceeding to the discussion of these methods, however, we have certain reservations to make with regard to the applicability of our model.

These reservations apply to face centred lattices. We saw in chapter I that we can reduce the problems of substitutional solid solutions, ferromagnetism and antiferromagnetism to the same general discussion - say the discussion of substitutional solid solutions. However, this is so only for the case of an AB lattice in which the sites can be labelled alternately α and β - i.e. at the absolute zero of temperature, all A atoms are on α sites, all B atoms are on β sites

and each A atom is surrounded by B atoms and vice versa. Obviously then, a general discussion on the above lines will be applicable to the simple cubic lattice and to the body centred cubic lattice and to the analogous models of a ferromagnet and an antiferromagnet, but not to the face centred substitutional solid lattice. (Since, in this case, we have to abandon one of our assumptions - either we have an AB lattice in which, at the absolute zero, each A atom is not entirely surrounded by B nearest neighbours, or we have, at the absolute zero, each A atom surrounded by B nearest neighbours, but our lattice is of the type AB_3.) It is also evident that our general model is not applicable to the face centred cubic antiferromagnet, since, at the absolute zero, this is equivalent to an AB lattice. It remains to show some justification of the treatment of the face centred cubic ferromagnet on our general model of a substitutional solid solution; if this can be done, we can proceed on the understanding that our work is applicable to all cubic structures of our three examples of co-operative phenomena, with the exceptions of the face centred cubic substitutional solid solution and the face centred cubic antiferromagnet.

Ferromagnets differ from the other two cases in that they show preference for like neighbours in the low energy states - that is, at the absolute zero, we have all our spins

parallel to each other and our sites are all α sites, say. In other words, the trouble which arose with face centred cubic lattices in the substitutional solid case - the impossibility of surrounding each α site with β sites, while, at the same time, keeping the number of α sites equal to the number of β sites - does not arise with ferromagnets. It is reasonable then, to say that if a general theory is applicable to ferromagnets in general, then it is applicable to the face centred ferromagnet in particular. Our treatment is thus justified, since we know from chapter I that a general discussion applies to ferromagnets.

§ 2. The Quasi-Chemical Method. (Bethe's First Approximation)

To improve upon the Bragg-Williams treatment, we require to introduce, explicitly or implicitly, into our calculations a short range order parameter, to measure the average extent to which, in the case of an AB lattice for example, the A atoms are surrounded by B atoms and vice versa.

Let us consider the case of a ferromagnet with co-ordination number z; ter Haar [18]) gives the application to a substitutional solid.

Since all our sites are equivalent, all α sites, say, we have three possibilities for the occupation of any pair of neighbouring sites. As before, we associate $\mu = +1$ with a spin aligned with the 'positive' direction and $\mu = -1$

with a spin in the direction antiparallel to this reference direction. Our possibilities are

	μ_r	μ_{r+1}		
↑↑	+	+	say we have Q_{++} of these pairs	(i)
↓↓	−	−	say we have Q_{--} of these pairs	(ii)
↑↓	+	−	say we have Q'_{+-} of these pairs	(iii)

Now the basic assumption of this method is that the pairs of spins in the lattice are all independent of each other. If this is so, we can split up the last of our above possibilities into two and write

$$Q'_{+-} = Q_{+-} + Q_{-+}$$

$$\text{where} \qquad Q_{+-} = Q_{-+} \tag{5,1}$$

since, for independent pairs, it is equally probable that the spin in the positive direction will be on either site. For the ferromagnet, this distinction has no real significance, but it ensures that the present discussion will be applicable to co-operative phenomena in general.

We have, if Q is the total number of pairs,

$$Q = Q_{++} + Q_{+-} + Q_{-+} + Q_{--} \tag{5,2}$$

Our short range order, we define by

$$\sigma Q = Q_{++} + Q_{--} - Q_{-+} - Q_{+-} \tag{5,3}$$

This gives preference to equal neighbours, since, at the absolute zero, $\sigma = 1$ if all pairs are ++ (or −−). At infinite temperature, σ should be equal to zero, which is the case if all our Q's are equal. (We see that our device

of splitting Q'_{+-} into two equal parts has ensured that possibility (iii) has double weight compared to the other two possibilities, at infinite temperature.)

From (5,3), we see also that

$$\sigma = \frac{1}{Q}\sum \mu_i\mu_j = -2E/Q\mathcal{J} \qquad \text{(from (1,14))}$$

$$\text{or} \qquad E = -\tfrac{1}{2}Q\mathcal{J}\sigma \qquad (5,4)$$

Subtracting (5,3) from (5,2), we find

$$Q_{+-} = Q_{-+} = 1/4Q(1 - \sigma) \qquad (5,5)$$

Let r be the fraction of sites correctly occupied, i.e., the fraction of spins which are parallel to their direction at T = 0. We have then,

$$2Q_{++} + Q_{+-} + Q_{-+} = 2rQ \qquad (5,6)$$

As before, we may define our long range order parameter by the equation

$$R = 2r - 1 \qquad (5,7)$$

From these two equations together with (5,2) and (5,5), it is easy to show that

$$\begin{aligned} Q_{++} &= Q/4(1 + \sigma + 2R) \\ Q_{--} &= Q/4(1 + \sigma - 2R) \end{aligned} \qquad (5,8)$$

Now, for the free energy in our lattice, we have

$$F = E(\sigma) - kT\ln W(R,\sigma) \qquad (5,9)$$

where $W(R,\sigma)$ is the number of ways of realising a certain pair of values of R and σ, i.e. it is the number of ways of realising the state of the crystal at the temperature appropriate to the values R and σ of the two order

parameters. Since the pairs in the lattice are independent, we have

$$W(R,\sigma) = \mathcal{W}(R)\,\frac{Q!}{Q_{++}!\,Q_{+-}!\,Q_{-+}!\,Q_{--}!} \qquad (5,10)$$

where $\mathcal{W}(R)$ depends only on R.

Substituting this into (5,9) and using (5,4) and Stirling's formula $\ln p! = p\ln p - p$, we find

$$F = -kT\left[\ln \mathcal{W}(R) + Q\ln Q - Q_{++}\ln Q_{++} - Q_{+-}\ln Q_{+-} - Q_{-+}\ln Q_{-+} - Q_{--}\ln Q_{--}\right] - \tfrac{1}{2}\mathcal{J}Q\sigma \qquad (5,11)$$

To find the equilibrium value of σ we use $\partial F/\partial\sigma = 0$ which gives eventually

$$\ln\frac{Q_{++}\,Q_{--}}{Q_{+-}\,Q_{-+}} = \frac{2\mathcal{J}}{kT} = \ln x^{-2} \qquad (5,12)$$

if we define $e^{-\mathcal{J}/kT} = x$. Taking antilogarithms of (5,12) and using (5,5) and (5,8) for the Q's, we find for σ in terms of R

$$\sigma = \frac{1+x^2-2x\sqrt{1-R^2+R^2x^2}}{1-x^2} \qquad (5,13)$$

We now want to use this result to find the equilibrium value of R. To this end, we define a free energy E'(R) by the equation

$$F(R) = E'(R) - kT\ln W(R) \qquad (5,14)$$

This equation is comparable to (1,20), the W(R) appearing here being similar to that in (1,21).

Rearranging (5,14), we have for the macroscopic partition function Z

$$Z = W(R)\, e^{-\frac{E'(R)}{kT}} \qquad (5,15)$$

But we know that, in equilibrium, the energy of the system is given by

$$E = -\frac{\partial \ln Z}{\partial\left(\frac{1}{kT}\right)}$$

which, from (5,15), is equal to $\frac{\partial}{\partial\left(\frac{1}{T}\right)}\left[\frac{E'(R)}{T}\right]$ (5,16)

As $T \longrightarrow \infty$, $E'(R) \longrightarrow 0$, so we have eventually, from (5,14)

$$F(R) = -kT \ln W(R) + T\int_{\frac{1}{T}=0}^{\frac{1}{T}} E(\sigma)\, d\left(\frac{1}{T}\right)$$

$$= -kT \ln W(R) + \tfrac{1}{2} kQT \int_{1}^{x} \sigma \frac{dx}{x} \qquad (5,17)$$

on making use of (5,4) and the definition of x.

We can now use the expression (5,13) for σ in this last equation and take for W(R)

$$W(R) = \binom{N}{Nr} = \frac{N!}{Nr!\, Nw!} \quad ; \quad \{(1 - r) = w\} \qquad (5,18)$$

since we have N lattice sites of which Nr have to be correctly occupied. Using Stirling's formula and the definition (5,7) of R, it is easy to show that (5,18) (and also (1,21)) is equivalent to

$$W(R) = N\left[\ln 2 - \tfrac{1}{2}(R+1)\ln(R+1) - \tfrac{1}{2}(1-R)\ln(1-R)\right]$$

Substituting this and (5,13) into (5,17) and using the definition

$$\chi u = \sqrt{1 - R^2 + R^2 x^2} \qquad (5,19)$$

we find

$$F(R) = \tfrac{1}{2} NkT \Big[(1+R)\ln(1+R) + (1-R)\ln(1-R) - 2\ln 2$$

$$+ \tfrac{z}{2}\Big[(1+R)\ln\frac{u+R}{1+R} + (1-R)\ln\frac{u-R}{1-R} - 2\ln\frac{u+1}{2} - \tfrac{1}{2}\ln\frac{u^2-R^2}{1-R^2}\Big\}\Big] \qquad (5,20)$$

where we have put Q = 1/2zN (see foot of page 13.) Incidentally, this is the correct form of equation (12.428) in reference 18.

We can now find the equilibrium value of R from

$$\frac{\partial F}{\partial R} = \left(\frac{\partial F}{\partial u}\right)_R \frac{\partial u}{\partial R} + \left(\frac{\partial F}{\partial R}\right)_u \qquad (5,21)$$

taken, of course, for x = constant. Carrying out these differentiations, we find

$$\left(1 - \frac{z}{2}\right)\ln\frac{1+R}{1-R} + \frac{z}{2}\ln\frac{u+R}{u-R} = 0 \qquad (5,22)$$

In order to find series expansions for the partition functions per spin of our two models, we shall use equations (5,19), (5,20) and (5,22) - which are quite general (compare those in reference 18).

§ 3 High and Low Temperature Series for λ on the Quasi-Chemical Method

At high temperatures, R = 0 and our three equations reduce to $u = 1/x$

$$\text{and } \frac{F}{NkT} = \tfrac{1}{2}\left[-2\ln 2 + \frac{z}{2}\left\{2\ln u - 2\ln\frac{u+1}{2} - \tfrac{1}{2}\ln u^2\right\}\right] \tag{5,23}$$

On substituting for u, the second of these becomes

$$-\frac{F}{NkT} = \ln \frac{2}{\left[\frac{2x^{1/2}}{1+x}\right]^{z/2}}$$

or

$$\frac{\lambda}{2} = \tfrac{1}{2} e^{-F/NkT} = \left[\frac{1+x}{2x^{1/2}}\right]^{z/2} = \left[\cosh K\right]^{z/2} \tag{5,24}$$

if we use the definition of x and put $J/2kT = K$.

At low temperatures, we can rearrange (5,20) and, making use of (5,22), put it in the form

$$\lambda = e^{-F/NkT} = x^{-z/4}\left[\frac{u+1}{u-R}\right]^{z/2}\left[\frac{1-R}{2}\right]^{z/2-1} \tag{5,25}$$

For the body centred cubic lattice, we substitute z = 8 into equations (5,22) and (5,25) to give

$$\begin{aligned}(u-R)^4 \lambda x^2 &= 1/8\,(u+1)^4(1-R)^3 \\ (1+R)^3(u-R)^4 &= (u+R)^4(1-R)^3\end{aligned} \tag{5,26}$$

We have also $1 - R^2 = x^2(u^2 - R^2)$ (5,19)

If we assume the series

$$\lambda = \left(1 + \sum_i \lambda_i x^{2i}\right) x^{-3/4}$$
$$R = 1 + \sum_i R_i x^{2i} \qquad (5,27)$$
$$u = 1 + \sum_i u_i x^{2i}$$

and solve our equations by successive approximation, we find eventually

$$\lambda = x^{-2}\left[1 + x^8 + 4x^{14} - 4x^{16} + 28x^{20} - 60x^{22} + 32x^{24}\right]$$

$$R = 1 - 2x^8 - 16x^{14} + 18x^{16} - 168x^{20} + 384x^{22} - 218x^{24} - 2016x^{26} + 6960x^{28} - 7968x^{30}$$

$$u = 1 + 2x^6 - 2x^8 + 14x^{12} - 32x^{14} + 18x^{16} + 140x^{18} - 492x^{20} + 570x^{22} + 1420x^{24} - 7728x^{26} + 13560x^{28}$$

- equations (5,28).

For the face centred cubic lattice we put z = 12 giving, from our main equations

$$(u - R)^6 \lambda x^3 = \tfrac{1}{32}(u+1)^6 (1-R)^5$$
$$(1+R)^5 (u-R)^6 = (u+R)^6 (1-R)^5 \qquad (5,29)$$

$$1 - R^2 = x^2 (u^2 - R^2) \qquad (5,19)$$

Assuming the same form of series as those in (5,27) we find

$$\lambda = x^{-3}\left[1 + x^{12} + 6x^{22} - 6x^{24} + 0x^{30}\right]$$

$$R = 1 - 2x^{12} - 14x^{22} + 26x^{24} - 396x^{32} + 864x^{34} - 470x^{36} + 0x^{38}$$

$$u = 1 + 2x^{10} - 2x^{12} + 22x^{20} - 48x^{22} + 26x^{24} + 352x^{30} - 1168x^{32} + 1286x^{34} - 470x^{36}$$

- equations (5,30)

We shall discuss these series for λ in the next chapter.

§ 4 Kirkwood's Method.

This method starts from equation (5,15). We have there for the macroscopic partition function

$$\mathcal{Z} = W(R)\, e^{-\frac{E'(R)}{kT}} = \sum_{\langle R\rangle} e^{\kappa s} \qquad \text{(compare (2,2))} \qquad (5,31)$$

where the summation is over all configurations with the same R and where $s = \sum_{\langle i,j \rangle} \mu_i \mu_j$.

From this we have

$$\frac{E'(R)}{kT} = -\ln \frac{\sum_{\langle R\rangle} e^{\kappa s}}{W(R)} \qquad (5,32)$$

If we denote the average of s^n, over all configurations for which R has a particular value, by $\langle s^n \rangle$, we have, from the usual definition of an average

$$\langle s^n \rangle = \frac{1}{W(R)} \sum_{\langle R\rangle} s^n \qquad (5,33)$$

and we can expand the exponential in (5,32)

$$\frac{E'(R)}{kT} = -\ln\left\{1 + \kappa\langle s\rangle + \frac{\kappa^2}{2!}\langle s^2\rangle + \frac{\kappa^3}{3!}\langle s^3\rangle + \cdots\right\} \qquad (5,34)$$

$$= -\sum_{n=1}^{\infty} \frac{\kappa^n}{n!} M_n \quad \text{on expanding the logarithm.} \qquad (5,35)$$

where $M_1 = \langle s\rangle$ $\qquad M_3 = \langle s^3\rangle - 3\langle s^2\rangle\langle s\rangle + 2\langle s\rangle^3$

$M_2 = \langle s^2\rangle - \langle s\rangle^2$ $\qquad$ etc.

We now require to evaluate the quantities M_n as functions of R - this forms the essence of Kirkwood's Method. However, in order to avoid needless repitition we shall omit these evaluations, since pages 274 - 279 of reference 18 give a very clear account of them. We shall merely quote the results. These are

$$\begin{aligned} M_1 &= Q R^2 \\ M_2 &= Q (1-R^2)^2 \\ M_3 &= 4Q R^2 (1-R^2)^2 \\ M_4 &= 2Q (1-R^2)^2 \left[(3z-9)(1-R^2)^2 + 2(1-3R^2)^2 \right] \end{aligned} \tag{5,36}$$

of which the first two are general, while those for M_3 and M_4 are valid only for the two dimensional square net and the simple cubic case.

Referring back to (5,14), we find, using (5,36) and (5,35), for the free energy

$$\begin{aligned} \frac{F(R)}{NkT} &= \tfrac{1}{2}(1+R)\ln(1+R) + \tfrac{1}{2}(1-R)\ln(1-R) - \ln 2 \\ &\quad + \frac{z}{2}\left[-KR^2 - \frac{K^2}{2}(1-R^2)^2 \right] \end{aligned} \tag{5,37}$$

where we have taken only the general terms in (5,36)

To find the equilibrium value of R, we again minimise the free energy and find

$$\ln \frac{1+R}{1-R} = zR\left[2K - 2K^2(1-R^2) \right] \tag{5,38}$$

These two equations are sufficient to enable us to solve for $e^{-F/NkT}$, the partition function per spin.

§ 5. High and Low Temperature Series for λ on Kirkwood's Method

At high temperatures, where the exponential and logarithmic series expansions are accurately valid, Kirkwood's method is effectively exact and we can, in principle, find as many terms of the exact series for λ as we wish, provided we take a sufficient number of the M_i into account in equation (5,37). The calculation of the M_i becomes extremely difficult, however, after the first few terms and as a means of finding the exact series, Kirkwood's method is not really practicable. For purposes of comparison with our variational method series we shall, therefore, consider only Kirkwood's first approximation, which involves using (5, 37) as it stands.

Putting R = 0 gives us, from (5,37), the high temperature series

$$\frac{\lambda}{2} = \frac{1}{2} e^{-F/NkT} = e^{\gamma/2 \cdot \frac{K^2}{2}}$$

$$= 1 + \frac{\gamma}{4} K^2 + \frac{\gamma^2}{16} K^4 + \cdots \qquad (5,39)$$

At low temperatures we use (5,38) to obtain R in terms of K. To do this we assume that $R = 1 + R_1$ and substitute

into (5,38)

$$\ln \frac{2+R_1}{-R_1} = z(1+R_1)\left[2K + 4K^2 R_1\right]$$

If we neglect R_1 as compared to unity - which is a good approximation at very low temperatures where $R \sim 1$, we have

$$R_1 = -2e^{-2zK}$$

Similarly, assuming $R = 1 - 2e^{-2zK}(1 + R_2)$, we find a second approximation giving

$$R = 1 - 2e^{-2zK} + 2e^{-4zK} f(K) \qquad (5,40)$$

where $f(K)$ is a polynomial in K.

We can now substitute this expression for R into equation (5,37) and find the first few terms of the series for λ. For these, the second exponential in (5,40) is unnecessary and we find from the first two terms

$$\lambda = e^{\frac{zK}{2}}\left[1 + e^{-2zK} + 2z\, e^{-4zK}(K + 2K^2)\right] \qquad (5,41)$$

This is valid for all lattices, since, in (5,37), we have used only M_1 and M_2.

§ 6. Zernike's Method.

For the last of our approximate methods we shall discuss a rather interesting one due to Zernike. In his paper, Zernike sets himself the task of calculating the probability of finding an A atom, say, at any point in the lattice if it

is known that an A atom definitely occupies the lattice point at the origin O. We are again considering an AB substitutional solid solution. Using the fact that once the nearest neighbours of a particular site are known, the probability for the occupation of the site no longer depends upon the distribution of the other atoms in the lattice, Zernike arrives at a recurrence relation for this probability.

At the end of paragraph 2, chapter 1, it was pointed out that the important energy unit in the case of an AB lattice is the quantity v defined in equation (1,7). It may be shown that, in calculating probabilities for the occupation of lattice sites, the probability for equal neighbours differs from that for unequal neighbours only by a Boltzmann factor $e^{-v/kT} = x$. Thus, if we treat the simple cubic case ($z = 6$), the ratios of the probabilities for an A atom and for a B atom at the central site are as

x^6 to 1 - nearest neighbours all A atoms

x^5 to x - 5A nearest neighbours and one B

x^4 to x^2 - 4A nearest neighbours and two B's

etc.

From these ratios we calculate the actual (normalised) probabilities for A at the central site to be

$$\omega_6 = \frac{x^6}{1+x^6} \; ; \; \omega_5 = \frac{x^4}{1+x^4} \; ; \; \omega_4 = \frac{x^2}{1+x^2} \; ; \; \omega_3 = \frac{1}{2} \text{ etc.} \qquad (5,42)$$

We can now take the arbitrary site at (i,j,k) with nearest neighbours designated by 1, 2, 3, ····· 6, and write down the probability p_{ijk} that this site is occupied by an A atom. If we make the approximation that the probabilities p_1, p_2, etc are independent, where p_i is the probability that atom i is an A atom, we have

$$P_{ijk} = P_1P_2P_3P_4P_5P_6\cdot\omega_6 + (q_1P_2P_3P_4P_5P_6 + P_1q_2P_3P_4P_5P_6 + \cdots)\cdot\omega_5$$

$$+ (q_1q_2P_3P_4P_5P_6 + \cdots)\,\omega_4 + \text{etc}. \qquad (5,43)$$

where $q = 1 - p$.

To simplify (5,43) we substitute for each probability twice its excess over unity i.e.

$$2p_i - 1 = \tau_i \quad \text{or} \quad p_i = \frac{1+\tau_i}{2} \;;\; q_i = \frac{1-\tau_i}{2} \;;\; \omega_i = \frac{1+u_i}{2} \qquad (5,44)$$

and find eventually

$$2^5\tau_{ijk} = (u_6 + 4u_5 + 5u_4)\sum\tau_i + (u_6 - 3u_4)\sum\tau_i\tau_j\tau_k$$

$$+ (u_6 - 4u_5 + 5u_4)\sum\tau_i\tau_j\tau_k\tau_l\tau_m. \qquad (5,45)$$

The summations include all possible combinations of the various products under the summation signs. From (5,42) and (5,44) we can write the u's in terms of x and then substitute into (5,45). This gives for the simple cubic lattice

$$-2^5 r_{ijk} = (\xi_6 + 4\xi_4 + 5\xi_2)\sum r_i + (\xi_6 - 3\xi_2)\sum r_i r_j r_k + (\xi_6 - 4\xi_4 + 5\xi_2)\sum r_i r_j r_k r_\ell r_m \qquad (5,46a)$$

where $$\xi_m = \frac{1-x^m}{1+x^m} \qquad (5,47)$$

For the other two cases we have

$$-2^7 r_{ijk} = (\xi_8 + 6\xi_6 + 14\xi_4 + 14\xi_2)\sum r_i + (\xi_8 + 2\xi_6 - 2\xi_4 - 6\xi_2)\sum r_i r_j r_k + (\xi_8 - 2\xi_6 - 2\xi_4 + 6\xi_2)\sum r_i r_j r_k r_\ell r_m + (\xi_8 - 6\xi_6 + 14\xi_4 - 14\xi_2)\sum 7r\text{'s} \qquad (5,46b)$$

Body Centred Cubic (z = 8).

$$-2^{11} r_{ijk} = (\xi_{12} + 10\xi_{10} + 44\xi_8 + 110\xi_6 + 165\xi_4 + 132\xi_2)\sum r_i + (\xi_{12} + 6\xi_{10} + 12\xi_8 + 2\xi_6 - 27\xi_4 - 36\xi_2)\sum r_i r_j r_k + (\xi_{12} + 2\xi_{10} - 4\xi_8 - 10\xi_6 + 5\xi_4 + 20\xi_2)\sum r_i r_j r_k r_\ell r_m + (\xi_{12} - 2\xi_{10} - 4\xi_8 + 10\xi_6 + 5\xi_4 - 20\xi_2)\sum 7r\text{'s} + (\xi_{12} - 6\xi_{10} + 12\xi_8 - 2\xi_6 - 27\xi_4 + 36\xi_2)\sum 9r\text{'s} + (\xi_{12} - 10\xi_{10} + 44\xi_8 - 110\xi_6 + 165\xi_4 - 132\xi_2)\sum 11r\text{'s} \qquad (5,46c)$$

Face Centred Cubic (z = 12).

(These equations - non-linear partial finite difference equations - have been checked through and two errors which occur in (5,46b) in Zernike's paper have been corrected.)

§ 7. High Temperature Series for λ on Zernike's Method.

Here we have to deal only with the short range order σ ; in Zernike's words, 'the influence of the fixed A atom at the origin of co-ordinates will soon become insensible when we consider sites farther and farther away from it, so that the probability of these sites being occupied by A atoms will tend to 1/2, or the excess, r_{ijk}, to zero. We therefore try to find solutions of our fundamental equations showing this property. Evidently the quantities r_{ijk} found in this way will express the tendency to local order'. It remains to relate the r_{ijk} to some definition of σ which will be equivalent to the definition we have hitherto used.

Zernike, following Bethe, defines

$$\sigma = \frac{n_B - n_A}{(n_B - n_A)_{\text{complete order}}} \qquad (5,48)$$

(which is a suitable definition since $\sigma_{T=0} = 1$ and $\sigma_{T=\infty} = 0$.) where n_A and n_B are respectively the average numbers of nearest neighbours of any A atom which are A atoms and which are B atoms. These numbers are proportional to the probabilities of finding an A atom and a B atom respectively at a particular site and so, using (5,44), we may write

$$\sigma = \frac{r_n}{(r_n)_{\text{complete order}}} \qquad (5,49)$$

n refers to the site nearest to the origin. For the simple cubic and body centred cubic cases $(r_n)_{\text{complete order}} = -1$ since, for these cases at T = 0, each A atom is entirely surrounded by B atoms. For these cases Zernike alternately changes the signs of the r's making the probability for a B neighbour equal to $1/2(1 + r_n)$. We thus have

$$\sigma_{s.c} = r_{1,\infty} \quad ; \quad \sigma_{B.c.c} = r_{111} \tag{5,50a,b}$$

For the face centred cubic case, however, we have the difficulty discussed in paragraph 1 of this chapter. To make Zernike's method suitable for a discussion of the face centred cubic ferromagnet we must assume preference in our AB lattice, for like nearest neighbours. This will give us

$$\sigma_{F.c.c} = r_{110} \tag{5,50c}$$

as opposed to $\sigma_{F.c.c} = -3r_{110}$ for the unlike neighbours case. We see then, that in order to reach an expression for the short range order, we need merely solve our appropriate partial difference equation for the excess probability r for the occupation of one of the sites nearest to the origin for the particular lattice structure in which we are interested. In order to do this we approximate equations (5,46) by taking in each only the part linear in r_i. In general we have

$$n r_{ijk} = \sum r_i \tag{5,51}$$

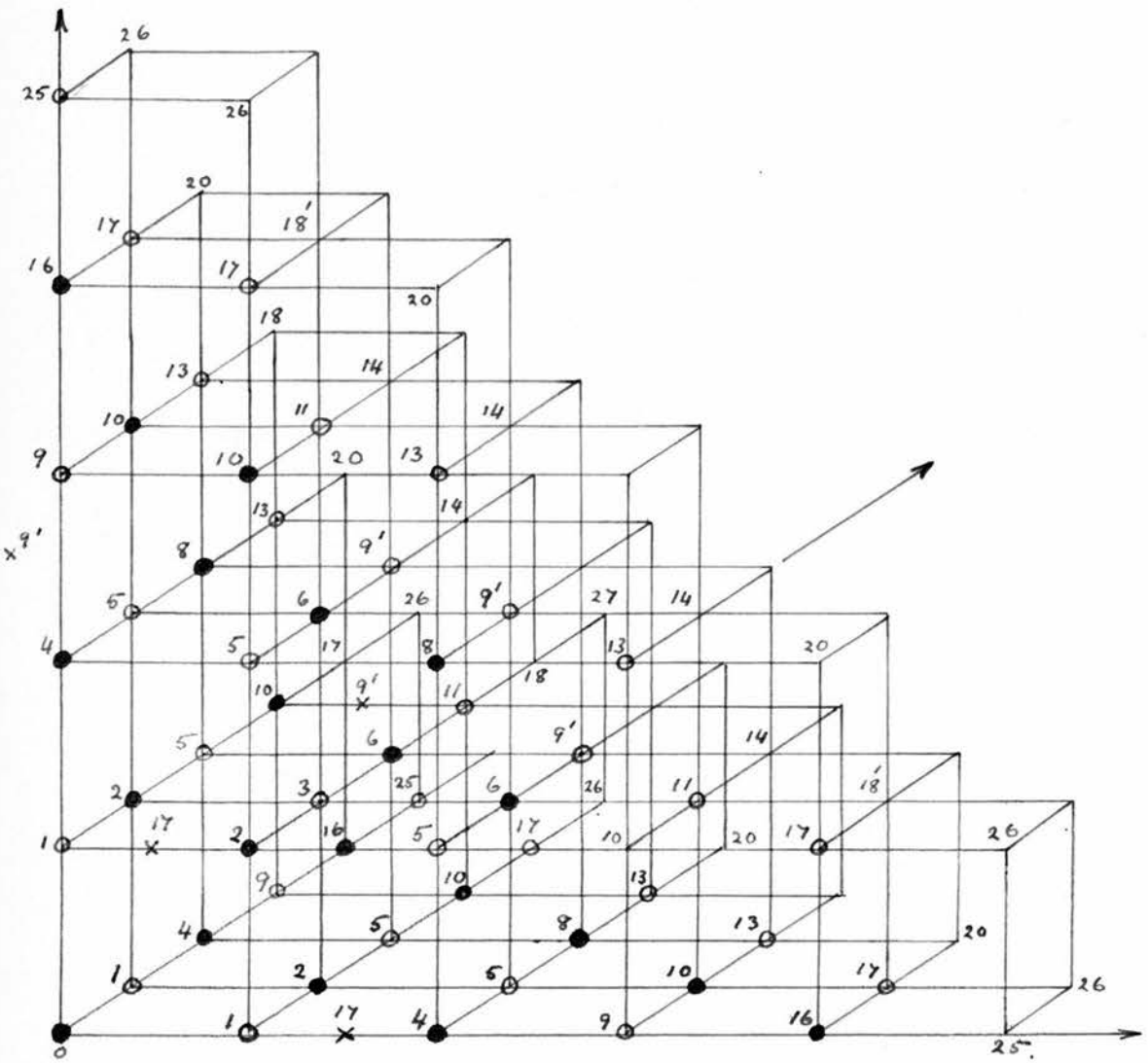


Figure 14.

where the summation includes all nearest neighbours of r_{ijk}. We then apply (5,51) to the succession of lattice points round the origin.

As an example of the method of solution let us consider, once again, the simple cubic lattice, and refer to figure 14. In this sketch, and in the work that follows, we have used, instead of the usual Cartesian co-ordinate system for specifying the lattice points, the square of the distance of each point from the origin, i.e. instead of (x,y,z) we take $x^2 + y^2 + z^2$. This permits us to take full advantage of the symmetry of our lattices. A dash indicates where the same sum of squares results from different combinations of x, y and z. For the simple cubic lattice we find

$$
\begin{aligned}
& nr_1 = 1 + 4r_2 + r_4 && \text{I} \\
& nr_2 = 2r_1 + 2r_3 + 2r_5 \\
& nr_4 = r_1 + 4r_5 + r_9 && \text{II} \\
& nr_3 = 3r_2 + 3r_6 \\
& nr_5 = r_2 + r_4 + 2r_6 + r_8 + r_{10} && \text{III} \\
& nr_9 = r_4 + 4r_{10} + r_{16} \\
& nr_6 = r_3 + 2r_5 + 2r_{9'} + r_{11} \\
& nr_8 = 2r_{9'} + 2r_{13} + 2r_5 \\
& nr_{10} = r_5 + r_9 + 2r_{11} + r_{13} + r_{17} && \text{IV} \\
& nr_{16} = r_9 + 4r_{17} + r_{25}
\end{aligned}
\qquad (5,52)
$$

where the Roman numerals indicate the shells round the origin.

For the atom at the origin we take $r_0 = 1$. If we now neglect in these equations, all but the first term on the right hand side of each, we find first approximations for the r's in terms of the n's

$r_1 = 1/n : r_2 = 2/n^2 : r_4 = 1/n^2 : r_3 = 6/n^3$ etc.

On putting these values into the equations we find second approximations, and so on. Going as far as the fifth shell, Zernike finds for the simple cubic lattice

$$\sigma = r_1 = \frac{1}{n} + \frac{9}{n^3} + \frac{166}{n^5} + \frac{3789}{n^7} + \frac{96282}{n^9} \qquad (5,53a)$$

Similarly, for the other two cases, he gives

$$\sigma = r_3 = \frac{1}{n} + \frac{19}{n^3} + \frac{632}{n^5} + \frac{25715}{n^7} + \frac{1168004}{n^9} \qquad (5,53b)$$

Body Centred Cubic.

$$\sigma = r_2 = \frac{1}{m} + \frac{17}{m^3} - \frac{4}{m^4} + \frac{568}{m^5} - \frac{256}{m^6} + \frac{23377}{m^7} - \frac{14300}{m^8} \qquad (5,53c)$$

where $m + 4 = n$. Face Centred Cubic.

To find the high temperature series for λ we refer once again to equation (5,17) which may be rewritten

$$\ln \lambda = -\frac{F(R)}{NkT} = \ln 2 + \frac{z}{2}\int_0^K \sigma \, dK \qquad (5,54)$$

where we have used the facts that $E = -[1/2]\mathcal{J}Q\sigma$ ($\mathcal{J} = v$), $Q = [1/2]zN$, $\lambda_{T\to\infty} = 2$ (see page 82.), and $K = (-\ln x)/2 = \mathcal{J}/2kT$. For σ we use (5,53), n can be found from (5,51) and (5,46) in terms of the ξ_i, while from (5,47) we see

that $\xi_m = \tanh mK = mK - \frac{(mK)^3}{3} + \frac{2}{15}(mK)^5 - \cdots$

The solution is straightforward and we find for our two cases, the series

$$\lambda = 2\left[1 + 2K^2 + 13.667K^4 + 232.756K^6\right] \qquad (5,55)$$

Body Centred Cubic Lattice.

$$\lambda = 2\left[1 + 3K^2 + 8K^3 + 37K^4 + 232K^5\right] \qquad (5,56)$$

Face Centred Cubic Lattice.

§ 8 Low Temperature Series for λ on Zernike's Method.

If a degree of long range order exists, the solutions of our fundamental equations (5,46) will each tend to a constant limit at great distances from the origin and this limit, say s, will be our long range order parameter. Putting this limiting value of r into our simple cubic equation gives

$$6a_5 s^4 + 20a_3 s^2 + 6a_1 - 32 = 0 \qquad (5,57a)$$

where we have taken out the root s = 0, (corresponding to the high temperature case) and where we have substituted a_1 for the co-efficient of $\sum r_i$, a_3 for the coefficient of $\sum r_i r_j r_k$, etc. The numerical factor in each term is, of course, equal to the number of terms under each summation in (5,46a).

For the other two cases we have

$$8b_7 s^6 + 56 b_5 s^4 + 56 b_3 s^2 + 8b_1 - 128 = 0 \qquad (5,57b)$$

Body Centred Case

$$12 c_{11} s^{10} + 220 c_9 s^8 + 792 c_7 s^6 + 792 c_5 s^4$$

$$+ 220 c_3 s^2 + 12 c_1 - 2048 = 0 \qquad (5,57c)$$

Face Centred Case - preference for like neighbours.

Now if we have the long range order s, the sites must be alternately α and β sites, with more A atoms than B atoms on the α sites and more B atoms than A atoms on the β sites. We can easily see that the probability of finding an A atom on an α site is $(1 + s)/2$ while the probability for an A atom on a β site is $(1 - s)/2$. Alternatively, if we select an A atom as our origin, the probability that it occupies an α site will be $(1 + s)/2$ and that it occupies a β site, $(1 - s)/2$. In the first case the limit of r_{ijk} will be +s and in the second case it will be -s. In order to find the average value of σ we must solve the difference equation for one or other of these limiting values. To facilitate calculation we make the substitution (for the +s case - change the sign of s for the -s case)

$$r_{ijk} = s + (1 - s)t_{ijk} \qquad (5,58)$$

This changes our variable from r to t, but has the advantage that it preserves our boundary condition ($t_{000} = 1$)

and consequently equations (5,53) hold also in the low temperature region.

Substituting (5,58) into equations (5,46) and neglecting non-linear terms, the terms independent of t cancel according to equations (5,57) leaving

$$32\, t_{ijk} = (a_1 + 10 a_2 s^2 + 5 a_5 s^4) \sum t_i \qquad (5,59a)$$

$$128\, t_{ijk} = (b_1 + 21 b_3 s^2 + 35 b_5 s^4 + 7 b_7 s^6) \sum t_i \qquad (5,59b)$$

$$2048\, t_{ijk} = (c_1 + 55 c_3 s^2 + 330 c_5 s^4 + 462 c_7 s^6 + 165 c_9 s^8 + 11 c_{11} s^{10}) \sum t_i \qquad (5,59c)$$

These are linear difference equations of the same form as (5,51). To make their solutions a little easier we can eliminate the term in s^2 in each, using equations (5,57). We find

$$\frac{1}{n_{s.c}} = \frac{1}{32} (16 - 2a_1 + 2a_5 s^4) \qquad (5,60a)$$

$$\frac{1}{n_{b.c.c}} = \frac{1}{128} (48 - 2b_1 + 14 b_5 s^4 + 4 b_7 s^6) \qquad (5,60b)$$

$$\frac{1}{n_{f.c.c.}} = \frac{1}{2048} (512 - 2c_1 + 132 c_5 s^4 + 264 c_7 s^6 + 110 c_9 s^8 + 8 c_{11} s^{10}) \qquad (5,60c)$$

Now, for the low temperature series for λ, we have

$$\ln \lambda = -\frac{\gamma}{4} \ln x + \frac{\gamma}{4} \int_0^x \frac{(1-\sigma)}{x} dx \qquad (5,61)$$

In order to complete the solution for λ we need an expression relating σ to t_i. This we find by calculating

the average probability for the nearest neighbours round the origin. We have

$$\sigma = \tfrac{1}{2}(1+s)\{s+(1-s)t_i\} + \tfrac{1}{2}(1-s)\{-s+(1+s)t_i\}$$

$$= s^2 + (1-s^2)t_i$$

or

$$1-\sigma = (1-s^2)(1-t_i) \qquad (5,62)$$

Since we have to develop a series expansion for λ, we require series for s^2 (since only powers of s^2 occur in our formulae) and for the t_i.

For s^2 we assume series of the form $s^2 = 1 + \sum s_i x^{2i}$ and, after substituting into equations (5,57), equate the coefficients in the resulting expressions. Actually, in order to simplify the working, these equations were cast into a form in powers of $(1 - s^2)$ instead of powers of s^2. On subsequently substituting for the a_i, b_i and c_i, the following alternatives to (5,57) were found

$$[8\xi_8 - 48\xi_6 + 112\xi_4 - 112\xi_2](1-s^2)^3 - [80\xi_8 - 256\xi_6 + 224\xi_4](1-s^2)^2$$

$$+ [192\xi_8 - 256\xi_6](1-s^2) - 128\xi_8 + 128 = 0 \qquad (5,63b)$$

Body Centred Lattice.

$$[12\xi_{12} - 120\xi_{10} + 528\xi_8 - 1320\xi_6 + 1980\xi_4 - 1584\xi_2](1-s^2)^5$$

$$- [280\xi_{12} - 1920\xi_{10} + 5280\xi_8 - 7040\xi_6 + 3960\xi_4](1-s^2)^4$$

$$+ [1792\xi_{12} - 8064\xi_{10} + 12672\xi_8 - 7040\xi_6](1-s^2)^3 - [4608\xi_{12} - 12288\xi_{10} + 8448\xi_8](1-s^2)^2$$

$$+ [5120\xi_{12} - 6144\xi_{10}](1-s^2) - 2048\xi_{12} + 2048 = 0 \qquad (5,63c)$$

for the Face Centred Lattice.

Since we are not immediately concerned with the simple cubic lattice equations we shall neglect these from now on. On substituting the assumed series for s^2 into these equations we find the following values for the coefficients

$$s^2_{b.c.c} = 1 - 4x^8 - 32x^{14} + 40x^{16} - 336x^{20} + 832x^{22} - 396x^{24} \qquad (5,64b)$$

$$s^2_{f.c.c} = 1 - 4x^{12} - 48x^{22} + 56x^{24} + 0x^{30} - 792x^{32} \qquad (5,64c)$$

where for ξ_m we have used the expansion

$$\xi_m = 1 - 2x^m + 2x^{2m} - 2x^{3m} + 2x^{4m} - \cdots \qquad (5,65)$$

In order to find an expansion for the t_i we use the expressions for 1/n given in equations (5,60). On substituting for the a_i, b_i and c_i in terms of the ξ_i and then for the ξ_i in terms of x, using the expansion in (5,65), we find

$$t_3 = x^6 - x^8 + 6x^{12} + 14x^{14} + x^{16} \qquad \text{(Body Centred)} \qquad (5,66b)$$

$$t_2 = x^{10} - x^{12} + 14x^{20} \qquad \text{(Face Centred)} \qquad (5,66c)$$

where we have used equations (5,53).

These expressions for the t_i and for the s^2 yield, on substituting them into (5,62) and (5,61), the following

series for our partition functions per spin

$$\lambda_{B.C.C.} = x^{-2}\left[1 + x^{8} + 4x^{10} - 4x^{16} + 28x^{20} - 70.18x^{22} + 25x^{24}\right] \quad (5,67b)$$

$$\lambda_{F.C.C} = x^{-3}\left[1 + x^{12} + 6x^{22} - 6x^{26} + 0x^{30} + 64.5x^{32}\right] \quad (5,67c)$$

§ 9. In conclusion we must briefly mention the exact series for λ calculated by Trefftz. We shall quote these series in the next chapter for comparison with the various approximate series we have derived.

Trefftz's series are derived according to methods given by Kramers and Wannier [19]) for the high temperature region and by van der Waerden [20]) for the low temperature region. We shall not go into these methods, but merely remark that they consist essentially in counting the various configurations corresponding to particular states of our crystals - the sort of procedure that is used in Kirkwood's method. We have used a similar procedure in connection with the derivation of the low temperature series for the partition function of the face centred cubic ferromagnet. (See paragraph 7, chapter 4.)

VI

THE VARIATIONAL METHOD COMPARED TO THE OTHER APPROXIMATE METHODS

§1. The Body Centred Cubic Lattice.

In chapter IV we had an indication that the body centred case might represent a limiting case of the variational method as we have used it. This was based on the fact that, for this case, all our maximisation parameters went out, the treatment of the problem being, as a consequence, so much simplified that we managed to find the high temperature solution for λ in a closed form. That this result should not be a particularly good one was also implied by the fact that for three of our four subsidiary partition functions we had to use the very crude approximations

$$\omega = \nu\mathcal{g} = 4$$

$$\mathcal{f} = 2$$

(A partition function should, by definition, vary with temperature.)

For the low temperature body centred cubic case we had reasonable expressions for all of our subsidiary partition functions - as a result of the existence of our low temperature (magnetisation) maximisation parameters - and consequently we expect that the low temperature series should be rather better than the high temperature one.

In order now to see just how good our worst case (on the variational treatment) is, we shall collect into tables our various series for comparison with each other. We shall use the abbreviations

E - exact series: V - variational method series:

B - Bethe's method (quasi-chemical) series:

Z - Zernike's method series: K - Kirkwood's method series.

Table 1.

Body Centred Lattice - High Temperature Series.

$\lambda_E = 2[1 + 2K^2 + 13\cdot667K^4 + 156\cdot756K^6]$

$\lambda_Z = 2[1 + 2K^2 + 13\cdot667K^4 + 232\cdot756K^6]$

$\lambda_K = 2[1 + 2K^2 + 4K^4]$

$\lambda_V = 2[1 + 2K^2 + 1\cdot667K^4] \quad = 2(\cosh K)^4$

$\lambda_B = 2[1 + 2K^2 + 1\cdot667K^4] \quad = 2(\cosh K)^4$

From this table it seems that the variational method and Bethe's method share the distinction of being the worst of our approximate methods.

Table 2.

Body Centred Lattice - Low Temperature Series.

$\lambda_E = x^{-2}[1 + x^8 + 4x^{14} - 4x^{16} + 28x^{20} - 60x^{22} + 44x^{24}]$

$\lambda_V = x^{-2}[1 + x^8 + 4x^{14} - 4x^{16} + 28x^{20} - 60x^{22} + 32x^{24}]$

$\lambda_B = x^{-2}[1 + x^8 + 4x^{14} - 4x^{16} + 28x^{20} - 60x^{22} + 32x^{24}]$

$\lambda_Z = x^{-2}[1 + x^8 + 4x^{14} - 4x^{16} + 28x^{20} - 70\cdot18x^{22}]$

$\lambda_K = x^{-2}[1 + x^8 + 0x^{14} + 16x^{16}(K + 2K^2)] \quad (x = e^{-2K}.)$

Here we see that the variational method and Bethe's method are both better than any of the other approximate methods. In fact, comparing the two expansions in both tables, we have the interesting result that our variational treatment of the body centred lattice turns out to give exactly the same series as does Bethe's method. In other words, Bethe's method would appear to be a lower limiting case of the variational method. To date, we have not had time to go into the equivalence of the two methods, but we feel that it should be a simple matter to demonstrate it without having to go to the extreme of actually calculating the series.

The fact that Kirkwood's method is so poor at low temperatures (as we shall see, the above series is typical) is not surprising, since the logarithmic and exponential expansions used in the derivation of Kirkwood's equations are only valid for high temperatures.

It may also be seen from this table, and from table 4, that, at low temperatures, the series derived on Zernike's method is only slightly poorer than that derived on Bethe's method. This implies that the results given by Kramers and Wannier for the low temperature square net series on Zernike's method, are wrong, since they show this series to fail at the second term. ter Haar (unpublished result) has recalculated this series and finds that the correct series is considerably better than that of Kramers and Wannier.

§ 2. The Face Centred Cubic Lattice.

For the face centred case, on our treatment, our conditions are slightly better than they are for the body centred case, since we now have an extra maximisation parameter to work with in our two temperature regions. The effect that this has on our results is seen in the following two tables.

Table 3.

Face Centred Lattice - High Temperature Series.

$$\lambda_E = 2[1 + 3k^2 + 8k^3 + 37k^4 + 185k^5]$$

$$\lambda_z = 2[1 + 3k^2 + 8k^3 + 37k^4 + 232k^5]$$

$$\lambda_v = 2[1 + 3k^2 + 8k^3 + 34k^4]$$

$$\lambda_K = 2[1 + 3k^2 + 0k^3 + 9k^4]$$

$$\lambda_B = 2[1 + 3k^2 + 0k^3 + 4k^4] = z(\cosh K)^6$$

We see that the extra parameter has improved the variational method considerably. Although it is still not the best of our approximate methods at high temperatures, the variational method series compares quite well with the corresponding series derived on Zernike's method - which is the best. With the exception of the variational series which has gone up one place in the table, the different series have the same relative positions that they occupy in table 1; in particular, Bethe's method remains the poorest of the approximations to the exact series.

Table 4.

Face Centred Lattice - Low Temperature Series.

$$\lambda_E = x^{-3}[1 + x^{12} + 6x^{22} - 6x^{24} + 8x^{30} + 42x^{32}]$$

$$\lambda_V = x^{-3}[1 + x^{12} + 6x^{22} - 6x^{24} + 8x^{30} + 72x^{32}]$$

$$\lambda_B = x^{-3}[1 + x^{12} + 6x^{22} - 6x^{24} + 0x^{30}]$$

$$\lambda_Z = x^{-3}[1 + x^{12} + 6x^{22} - 6x^{24} + 0x^{30}]$$

$$\lambda_K = x^{-3}[1 + x^{12} + 0x^{22} + 24x^{24}(K + 2K^2)]$$

In this table the various methods retain their places according to table 2. (With a degree of uncertainty as regards the Bethe and Zernike method series; they are identical so far as we have gone.) The variational series holds good up to the term in x^{32} while the Bethe and Zernike series fail somewhere between x^{24} and x^{30}. Kirkwood's series is again the poorest.

§ 3. The Simple Cubic Lattice.

Finally, in order to emphasize the way in which the addition of further maximisation parameters improves the variational method, we shall quote the results obtained by ter Haar (reference 15) for the high temperature simple cubic series. (See the discussion in paragraph 9 in chapter IV.) From table 5, we see that the different methods again have the relative ordering that they have in table 1, with the exception of the variational method series which has moved up to the top place.

Table 5.

Simple Cubic Lattice - High Temperature Series.

$$\lambda_E = 2[1 + \tfrac{3}{2}K^2 + \tfrac{31}{8}K^4 + \tfrac{5461}{240}K^6 + \tfrac{2436211}{13440}K^8]$$

$$\lambda_V = 2[1 + \tfrac{3}{2}K^2 + \tfrac{31}{8}K^4 + \tfrac{5461}{240}K^6 + \tfrac{2180857}{13440}K^8]$$

$$\lambda_2 = 2[1 + \tfrac{3}{2}K^2 + \tfrac{31}{8}K^4 + \tfrac{8039}{80}K^6]$$

$$\lambda_K = 2[1 + \tfrac{3}{2}K^2 + \tfrac{9}{8}K^4]$$

$$\lambda_B = 2[1 + \tfrac{3}{2}K^2 + \tfrac{7}{8}K^4]$$

§ 4. It is hardly necessary to point out the significance of our results, since the tables we have given do this sufficiently well. We have found that the variational method as we have used it , yields equations which increase in complexity and difficulty of solution as we go through the cubic structures in the order, body centred cubic, face centred cubic, simple cubic. Also, for any particular structure, the high temperature equations are much simpler than are the corresponding low temperature ones. All this, of course, is the result of the variation in the number of our maximisation parameters as we treat the different cases and as we treat the two different temperature regions. In fact we might say that complexity of equations is proportional to the number of undetermined parameters we have to

deal with. We have found also that as the number of our unknown parameters increases in the various problems we have discussed, our solutions correspondingly improve. This, of course, is more or less as we should expect.

It is fair to say, then, that the variational method is, when used to its full advantage, an extremely powerful means of calculating the partition functions for the types of problem we have discussed. Since the simple cubic lattice affords the best application of the method to three dimensional problems, it seems a reasonable extrapolation of our results to state that when the low temperature simple cubic series has been obtained, it will be found to be better than the series found by the other methods.

Altogether, we can say that our results and their implications agree with those of Kramers and Wannier, in that a direct application of the variational method seems to be the most powerful of any of the approximate methods used in this work to calculate the partition functions of ferromagnetic structures.

(We have not carried out our original intention to use equations (1,2) to calculate series for the configurational energies and specific heats of our models, since we felt that no useful purpose would be achieved by doing so. Tables 1. - 5. are quite adequate for a comparison of the various approximate methods.)

APPENDIX

Note on the Direct Application of the Variational Method to the Face Centred Cubic Lattice.

In our attempts to apply the variational method directly to the face centred cubic structure, it soon became evident that, to avoid effectively repeating the discussion given in chapter III, we should have to consider the lattice as being built up by adding two plates (one layer of unit cells) at a time to the existing structure. If we do this, we are performing a step which can be repeated exactly at each stage.

After our first reduction, our lattice will be reduced to two essentially two dimensional ones consisting of four plates of spins and two plates of spins. The next two reductions will leave us with lattices consisting of sixteen strips, four strips, four strips and two strips of spins. The four final reductions will reduce the problem to the consideration of lattices containing respectively thirty-two, eight, eight and four spins. By comparing these lattices with the previous subsidiary ones we have had to deal with, we see that the matrices corresponding to them will be of orders 2^{16}, 2^4, 2^4 and 2^2 respectively. The largest of these is obviously far too large to handle.

In the case of the body centred cubic lattice, any direct approach using the variational method must also be carried out on the above lines. We find a matrix of order 2^8.

REFERENCES

1) Domb, C. and Potts, R.B., Proc.Roy.Soc., A210 (1951) 125

2) Fowler, R.H. and Guggenheim, E.A., Statistical Thermodynamics, Cambridge, (1939)

3) Ising, E., Zeit.Phys., 31 (1925) 253

4) Rushbrooke, G.S., Nuovo Cim., 6 (1949) Suppl. 2, 251

5) Tammann, G., Lehrb.d.Metallographie, Leipzig, (1921)

6) Johansson, C.H. and Linde, J.O., Ann.Physik, 78 (1925) 439

7) Bragg, W.L. and Williams, E.J., Proc.Roy.Soc., A145 (1934) 699
A151 (1935) 540
A152 (1935) 231

8) Bethe, H.A., Proc.Roy.Soc., A150 (1935) 552

9) Fowler, R.H. and Guggenheim, E.A., Proc.Roy.Soc., A174 (1940) 186

10) Chang, T.S., Proc.Camb.Phil.Soc., 35 (1939) 256

11) Kirkwood, J.G., J.Chem.Phys., 6 (1938) 70

12) Zernike, F., Physica, 7 (1940) 565

13) Kramers, H.A. and Wannier, G.H., Phys.Rev., 60 (1941) 252, 263

14) Martin, B. and ter Haar, D., Physica, 18 (1952) 569

15) ter Haar, D., Physica, 18 (1952) 836

16) Trefftz, E., Zeit.Phys., 127 (1950) 371

17) Martin, B., Statistics of the Three Dimensional Ferromagnet, Thesis, Purdue (1951)

18) ter Haar, D., Elements of Statistical Mechanics, Rinehart, New York (1952) ch. 12

19) Kramers, H.A. and Wannier, G.H., Phys.Rev., 64 (1943) 159

20) van der Waerden, B., Zeit.Phys., 118 (1941) 473