Antiferromagnetism

W. Marshall

References	Article cited in: **http://rspa.royalsocietypublishing.org/content/232/1188/48#related-urls**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Proc. R. Soc. Lond. A* go to:
http://rspa.royalsocietypublishing.org/subscriptions

Bacq, Z. M. & Alexander, P. 1955 *Fundamentals of radiobiology*. London: Butterworths.
Bellamy, . 1954 *The infra-red spectra of complex molecules*. London: Methuen.
Charlesby, A. 1952 *Proc. Roy. Soc.* A, **215**, 187.
Charlesby, A. 1954a *Proc. Roy. Soc.* A, **217**, 122.
Charlesby, A. 1954b *Proc. Roy. Soc.* A, **224**, 120.
Fox, M. & Flory, P. J. 1949 *J. Phys. Chem.* **53**, 197.
Lawton, E. J., Bueche, A. M. & Balwit, J. S. 1953 *Nature, Lond.*, **172**, 76.
Setlow, R. B. 1952 *Proc. Nat. Acad. Sci., Wash.*, **38**, 166.
Setlow, R. B. & Doyle, B. 1953 *Arch. Biochem. Biophys.* **46**, 46.
Sheppard, N. & Simpson, D. M. 1952 *Quart. Rev. Chem. Soc.* **6**, 1.

Antiferromagnetism

By W. Marshall†

Theoretical Physical Department, University of Birmingham

The ground state of a lattice with one electron per atom and antiferromagnetic interactions between nearest neighbours only is examined by a variational method similar in principle to the treatments by Hulthén (1938) and Kasteleijn (1952) of the linear–chain problem. The calculation involves a statistical problem which is shown to be exactly equivalent to the Ising ferromagnetic problem. This cannot be solved exactly, except in the one-dimensional case, and so the Bethe–Peierls method is used to solve it approximately. In complete contradiction to the Kubo (1953) variational calculation it is concluded that all lattices have disordered ground states. The energy differences between ordered and disordered states are small and so the nature of the ground state is likely to be sensitive to small additional ordering forces.

1. Introduction

The theoretical position with regard to antiferromagnetism is far from satisfactory, although qualitatively the simple Van Vleck (1941) theory is very successful. As it is hoped to show in a later paper (Marshall 1955), the spin-wave theory of Anderson (1952), Kubo (1952) and Ziman (1952a, b, 1953) is open to many criticisms and, in the opinion of the author, conclusions drawn from it, particularly as to the conditions required for antiferromagnetism are not to be relied upon. Neither is the Kubo (1953) variational calculation completely satisfactory, although it does avoid most of the objections to the usual spin-wave theory, because certain approximations are made which may not be valid and because, as Kubo himself points out, it fails to predict the ground state of the linear chain to be disordered.

It therefore seems worth while to examine the ground states of various lattices by a variational method similar in principle to the treatments of the linear chain problem by Hulthén (1938) and Kasteleijn (1952). By doing this it is hoped to obtain some idea of the conditions required for the ground state to have anti-ferromagnetic ordering. In particular it is hoped to determine if nearest neighbour

† Present address: Theoretical Physics Division, A.E.R.E., Harwell, Berks.

antiferromagnetic exchange interactions alone are sufficient to give an ordered ground state or whether some additional ordering force is required. Of course, the conditions required for the ground state to have antiferromagnetic ordering are only necessary conditions for antiferromagnetism because a lattice can show anti-ferromagnetic ordering in the ground state without being antiferromagnetic, just as a linear chain with ferromagnetic interactions has a ground state showing ferro-magnetic ordering, although it is not ferromagnetic. This is because the statistical weights of the ordered and disordered states near the ground state become important at any temperature greater than zero. Hence, only a theory capable of examining the excited states of the system would tell us if these conditions were also sufficient for antiferromagnetism.

2. The variational calculation

We assume that:

(*a*) The Heitler–London model of electrons localized upon each magnetic atom is valid.

(*b*) The magnetic atoms lie on a lattice the points of which can be divided into two sublattices, A and B, such that all the nearest neighbours of a point on one sub-lattice lie on the other. This excludes the face-centred cubic and hexagonal plane lattices but includes the linear chain, square plane, simple cubic and body-centred cubic lattices.

(*c*) Only nearest neighbour exchange interactions are important and these are all equal.

(*d*) There is just one electron, apart from those in closed shells, per magnetic atom, i.e. with each magnetic atom is associated a spin of $S = \frac{1}{2}$.

Because of (*a*) this work cannot be applied to antiferromagnetic metals since the Heitler–London model cannot give electrical conductivity. However, most anti-ferromagnets are compounds in which the magnetic atoms are kept far apart by non-magnetic atoms and for these the Heitler–London model is probably good. The fact that the Heitler–London model can be used to describe the $3d$ electrons in a ferromagnet with some success (it gives the correct law for the approach to satura-tion at low temperatures) but not the $3d$ electrons of an antiferromagnet is not surprising. For in a ferromagnet at low temperatures most of the spins of these electrons are parallel and so the Pauli principle keeps them apart and this helps to keep them localized. In an antiferromagnet where there are equal numbers of parallel and antiparallel spins this effect cannot be important.

Assumption (*c*) is a very serious approximation and these calculations indicate that next nearest neighbour interactions or anisotropy forces, as well as nearest neighbour interactions, play a vital part in forming antiferromagnetic ordering. We shall return to this point in §5. Also the assumption that all the nearest neighbour interactions are equal is unlikely to be valid in general since the magnetic electrons are in $3d$ orbitals. But probably this asymmetric electron distribution is important only because it produces anisotropy forces through the residual spin-orbit coupling. (*d*) is probably not a serious limitation.

The Hamiltonian, with the addition of a constant term to that usually empolyed, can be written (Van Vleck 1945),

$$H = 2J \sum_{\langle j, k \rangle} (\mathbf{S}_j . \mathbf{S}_k - \tfrac{1}{4}) \tag{1}$$

$$= 2J \sum_{\langle j, k \rangle} (\mathbf{S}_j^z \mathbf{S}_k^z - \tfrac{1}{4}) + J \sum_{\langle j, k \rangle} (\mathbf{S}_j^+ \mathbf{S}_k^- + \mathbf{S}_j^- \mathbf{S}_k^+), \tag{2}$$

where $\mathbf{S}_j$ and $\mathbf{S}_k$ are spin operators expressed in units of $\hbar$, $\mathbf{S}^{\pm} = \mathbf{S}^x \pm \mathbf{S}^y$, and the symbol $\langle j, k \rangle$ denotes that the sum goes over all nearest neighbour pairs j, k.

The solution of the antiferromagnetic problem using the complete Hamiltonian (2) is always very much more difficult to obtain than that of the ferromagnetic case. To illustrate this consider the linear chain. Then in (2) if J is negative, as it is in the ferromagnetic case, the ground state is immediately obvious, namely

$$\ldots + + + + + + \ldots.$$

But in the antiferromagnetic case J is positive and the lowest diagonal element of (2), $-JN$, is given by the perfectly ordered chain

$$\ldots + - + - + - + - \ldots.$$

But this is not a stationary state eigenfunction for it has non-diagonal matrix elements, each of magnitude J, with N other arrangements. Although these non-diagonal matrix elements are much smaller than the diagonal elements there are so many of them that we expect their disordering effect to be comparable to the ordering effect of the diagonal terms. It is therefore not at all obvious if the ground state of (2) is ordered or not.

Physically we can regard the problem as follows. Two adjacent electrons must constantly be exchanging their orientations. This does not matter if the orientations are the same, which is so if the spins are parallel ($S_1^z = S_2^z$), and so a nearest neighbour pair of parallel spins is a 'stable' arrangement. But it does matter if the spins are antiparallel ($S_1^z = -S_2^z$) and hence the effect of this 'spin exchanging' is to mix up and destroy those arrangements which have a large number of antiparallel nearest neighbour pairs. On the other hand, it is just these arrangements which, if we took only the diagonal part of (2), would give the lowest energies.

Because of (d) each spin has two possible orientations and, since there are N magnetic atoms in the crystal, a complete orthogonal set of 2^N wave functions can be constructed. Let any one of these be $\{\mu\}$ and suppose that, for this arrangement, m_μ is the total z-component of spin, q_μ is the number of antiparallel (i.e. $+-$) nearest neighbour pairs and p_μ is the number of plus spins on the A sublattice.

Let an eigenfunction of the system be

$$\psi = \sum_\mu c_\mu \{\mu\}. \tag{3}$$

Now the operator $2J(S_j^z S_k^z - \tfrac{1}{4})$ gives zero if the spins on j, k are prallel and gives $-J$ if they are antiparallel. The operator $(\mathbf{S}_j^+ \mathbf{S}_k^- + \mathbf{S}_j^- \mathbf{S}_k^+)$ also gives zero if the spins of j, k are parallel and if they are antiparallel simply interchanges the spins, i.e. $+- \rightarrow -+$. Hence

$$H\psi = J \sum_\mu c_\mu [-q_\mu \{\mu\} + \sum_{\mu'} \{\mu'\}] = J \sum_\mu \{\mu\} \sum_{\{\mu'\}} (c_{\mu'} - c_\mu), \tag{4}$$

where $\{\mu'\}$ differs from $\{\mu\}$ by the interchange of one antiparallel neighbouring pair and so the sum over μ' contains q_μ terms. The energy eigenvalue of this state is

$$E_c = [\psi^* \psi]^{-1} [\psi^* H \psi] = J [\sum_\mu c_\mu^* c_\mu]^{-1} [\sum_\mu c_\mu^* \sum_{\mu'} (c_{\mu'} - c_\mu)]. \tag{5}$$

The ground state is given by that set of coefficients c_μ which makes E_c a minimum. This set has the form

$$c_\mu = (-1)^{p_\mu} a_\mu, \tag{6}$$

where a_μ is real and positive. To prove this suppose that it were not so, i.e. that E_c were the ground state eigenvalue and the a_μ were not all real and positive. Then $p_{\mu'} = p_\mu \pm 1$ by definition of $\{\mu'\}$ and so $c_\mu^* c_{\mu'} = -a_\mu^* a_{\mu'}$. Hence

$$E_c = -J [\sum_\mu a_\mu^* a_\mu]^{-1} [\sum_\mu a_\mu^* \sum_{\mu'} (a_{\mu'} + a_\mu)]. \tag{7}$$

Clearly we could get a lower eigenvalue than E_c by using coefficients b_μ defined by $b_\mu = (-1)^{p_\mu} |a_\mu|$, this would be a contradiction of the statement that E_c is the ground state eigenvalue. Hence (6) is proved. The author is indebted to Professor R. E. Peierls for the above proof. It is easy to verify that (6) remains true if an additional ferromagnetic next nearest neighbour interaction is present. It is not necessarily true if the next nearest neighbour interactions are antiferromagnetic.

The total z-component of spin commutes with H and so the eigenvalues of the former operator may be selected one at a time and the corresponding energy eigenvalues and eigenfunctions determined for each of them. It follows from the fact that each component of the total spin commutes with H that $\mathbf{S}^2$ does so also. It can now be shown that the ground state has a total spin equal to zero whenever (6) is true. This proof is given in appendix A where is it also shown that the theorem corresponding to (6) in the case of $S > \frac{1}{2}$, i.e. more than one electron per magnetic atom, is easy to find. Hence the proof that the ground state is a singlet holds for any integer number of electrons per magnetic atom. This is a result which has been assumed intuitively by Anderson (1952) but it has not been proved before except for small groups of atoms.

We now wish to find some approximate form for the a_μ of (6) in order to construct a trial function. The approximation we make is that a_μ is a function only of p_μ, the number of plus spins on the A sublattice, and q_μ, the number of nearest neighbour antiparallel pairs. Naturally we need consider only those arrangements $\{\mu\}$ for which m_μ, the total z-component of spin, is zero. Hence

$$\psi_{\text{tr.}} = \sum_{\substack{\mu \\ [m_\mu = 0]}} (-1)^{p_\mu} [W(p_\mu, q_\mu)]^{-\frac{1}{2}} a(p_\mu, q_\mu) \{\mu\}, \tag{8}$$

where $W(p_\mu, q_\mu)$ is the number of arrangements $\{\mu\}$ with the values p_μ, q_μ. The factor $[W(p_\mu, q_\mu)]^{-\frac{1}{2}}$ which could have been absorbed into $a(p_\mu, q_\mu)$ is introduced for convenience. (8) is essentially the same trial function as used by Kasteleijn (1952) in his treatment of the linear chain problem. Hulthén (1938), who also considered the latter problem, used a simpler trial function in which a_μ depended only upon q_μ.

The particular form assumed for a_μ needs some justification. $-Jq_\mu$ is the eigenvalue of the operator $2J \sum_{\langle j,k \rangle} (\mathbf{S}_j^z \mathbf{S}_k^z - \frac{1}{4})$ which is the diagonal part of (2). q_μ is therefore directly related to the Ising energy and short range order existing in the arrangement $\{\mu\}$. p_μ directly measures the long range order in $\{\mu\}$. It therefore seems certain that p_μ and q_μ are the most important parameters to use but it is a very serious approximation to assume that they are the only ones necessary and the method stands or falls on the question of whether or not it is a good enough approximation to make in constructing $\psi_{\mathrm{tr.}}$.

One criticism of (8) is that it is not necessarily isotropic because the z-axis is now a preferred direction in space and not merely an axis of quantization. It is therefore of interest to calculate the expectation value of the square of the total spin for (8). This is done in appendix B where it is shown that

$$\langle \mathbf{S}^2 \rangle \lesssim N, \quad \text{i.e.} \quad \langle \, | \, \mathbf{S} \, | \, \rangle \lesssim \sqrt{N}. \tag{9}$$

Although $\langle \mathbf{S}^2 \rangle \approx N$ is a large number it is negligible compared to the possible range of values of $\langle \mathbf{S}^2 \rangle$ which is $\approx N^2$. This does not mean that (8) is necessarily a good approximation to the ground state because most of the true eigenstates of the system have $\mathbf{S}^2 \approx N$ and therefore any arbitrarily chosen linear combination of them would be expected to give $\langle \mathbf{S}^2 \rangle \approx N$ also. But (9) does show that the anisotropy introduced into (8) by using p_μ and q_μ as variation parameters is not too large and this gives us more confidence in the approximation. This situation is similar to that occurring in the spin-wave theory (Anderson 1952), where the approximations introduced also give an anisotropic ground state with $\langle \mathbf{S}^2 \rangle \approx N$.

We rewrite (8) in the form

$$\psi_{\mathrm{tr.}} = \sum_{p,q} a(p,q)\, \psi(p,q), \tag{10}$$

where

$$\psi(p,q) = \{W(p,q)\}^{-\frac{1}{2}} \sum_\mu (-1)^p\, \delta_{(m_\mu,\,0)}\, \delta_{(p_\mu,\,p)}\, \delta_{(q_\mu,\,q)}\, \{\mu\}. \tag{11}$$

Clearly the functions $\psi(p,q)$ are orthonormal. The matrix elements

$$\psi^*(p',q')\, H \psi(p,q)$$

are calculated in the following way

$$H\psi(p,q)$$

$$= -qJ\psi(p,q) + J(-1)^p \{W(p,q)\}^{-\frac{1}{2}} \sum_\mu \delta_{(m_\mu,\,0)}\, \delta_{(p_\mu,\,p)}\, \delta_{(q_\mu,\,q)} \sum_{\langle j,k \rangle} (\mathbf{S}_j^+ \mathbf{S}_k^- + \mathbf{S}_j^- \mathbf{S}_k^+) \{\mu\}. \tag{12}$$

The last expression of (12) contains $qW(p,q)$ terms since the inner sum contains q non-zero terms.

Define $f_{p,q}^{p';q'}$ as the probability that, if two adjacent antiparallel spins in an arrangement with given p and q are interchanged, the p value will change to p' and the q value will change to q'. Then the $qW(p,q)$ terms of the last expression in (12) are divided up amongst the different p and q values in such a way that there is a fraction $f_{p,q}^{p';q'}$ of them with the values p', q'. Each term also has the opposite sign

to the one it would have in (11) because $p' = p \pm 1$ for all non-zero $f_{p,q}^{p',q'}$. Therefore, multiplying on the left of (12) with $\psi^*(p', q')$ we get

$$(\psi^*(p', q') H \psi(p, q)) = -qJ\delta_{p,p'}\delta_{qq'} - Jq[W(p,q)]^{\frac{1}{2}}[W(p',q')]^{-\frac{1}{2}}f_{p,q}^{p',q'}. \qquad (13)$$

Since H is Hermitian,

$$q[W(p,q)]^{\frac{1}{2}}[W(p',q')]^{-\frac{1}{2}}f_{p,q}^{p',q'} = q'[W(p',q')]^{\frac{1}{2}}[W(p,q)]^{-\frac{1}{2}}f_{p',q'}^{p,q}. \qquad (14)$$

Hence, eliminating $[W(p,q)]^{\frac{1}{2}}[W(p',q')]^{-\frac{1}{2}}$ from (13) gives

$$(\psi^*(p', q') H \psi(p, q)) = -qJ\delta_{pp'}\delta_{qq'} - J\{qq' f_{p,q}^{p',q'} f_{p',q'}^{p,q}\}^{\frac{1}{2}}. \qquad (15)$$

Operating on (10) with H and multiplying on the left by $\psi^*(p, q)$ gives

$$-(E/J)\,a(p,q) = qa(p,q) + \sum_{p',q'} a(p',q')\{qq' f_{p,q}^{p',q'} f_{p',q'}^{p,q}\}^{\frac{1}{2}}. \qquad (16)$$

It is now convenient to define variables x and y by

$$q = \tfrac{1}{4}Nz(1+x) \quad \text{and} \quad p = \tfrac{1}{4}N(1+y), \qquad (17)$$

where z is the number of nearest neighbours to each atom. Since $0 \leqslant p \leqslant \frac{1}{2}N$ then $-1 \leqslant y \leqslant 1$. q is not strictly independent of p. In fact it can be shown that

$$\left| 2zp - \tfrac{1}{2}zN \right| \leqslant q \leqslant \tfrac{1}{2}zN$$

and hence $2\,|\,y\,| - 1 \leqslant x \leqslant 1$. Subject to this condition x and y, i.e. p and q, may be considered as independent variables. Now define

$$a(p,q) = \exp \tfrac{1}{4}Nz\chi(x,y). \qquad (18)$$

Then
$$a(p',q') = a(p,q)\exp \tfrac{1}{4}Nz[\chi(x',y') - \chi(x,y)]. \qquad (19)$$

Now $f_{p,q}^{p',q'}$ is non-zero only for $|\,q'-q\,| \leqslant 2(z-1)$ and $p' = p \pm 1$ so for the non-zero terms of (2·16), $(x'-x) \approx 1/N$ and $(y'-y) \approx 1/N$. Hence assuming $\chi(x,y)$ is a smoothly varying function and ignoring terms $\approx 1/N$, (16) gives

$$-E/J = q + q\sum_{p'\,q'}\{f_{p,q}^{p',q'} f_{p',q'}^{p,q}\}^{\frac{1}{2}}\exp \tfrac{1}{4}Nz\left[(x'-x)\frac{\partial\chi}{\partial x} + (y'-y)\frac{\partial\chi}{\partial y}\right]. \qquad (20)$$

Let $q' = q + 2r$ and $p' = p + v$. Then neglecting terms $\approx v/p$ and xr/q, i.e. $\approx 1/N$

$$f_{p+v,q+r}^{p,q} = f_{p,q}^{p-v,q-r}. \qquad (21)$$

Thus (20) reduces eventually to

$$-E/J = q\left[1 + 2\sum_{r=-(z-1)}^{(z-1)}\{f_{p,q}^{p+1,q+2r} f_{p,q}^{p-1,q-2r}\}^{\frac{1}{2}}\cosh\left\{2r\frac{\partial\chi}{\partial x} + z\frac{\partial\chi}{\partial y}\right\}\right]. \qquad (22)$$

We now notice that $a(p,q)$ cannot have a maximum value at any of the extreme limits of the ranges of p and q. To show this substitute $q = q_{\max}$ into (16) and then the sum over q' in the second term of (16) must run only over values of $q' \leqslant q_{\max}$, i.e. over values of $r \leqslant 0$. Therefore as $q \to q_{\max}$, (22) must tend asymptotically to an equation with the sum over all r replaced by a sum over $r \leqslant 0$ only. Hence as $q \to q_{\max}$, $\partial\chi/\partial x \to -\infty$. Therefore $a(p,q)$ does not have a maximum at $q = q_{\max}$ and in fact initially it increases as q decreases from $q_{\max}$. $a(p,q)$ can be shown to have

a similar behaviour at all the extreme limits of p and q, and since it always initially increases as p, q move away from these extreme limits it follows that $a(p,q)$ must have at least one maximum at some point inside the allowed ranges for p and q. At a maximum point, say p_0, q_0, $\dfrac{\partial \chi}{\partial x} = \dfrac{\partial \chi}{\partial y} = 0$ and so

$$-E/J = q_0\left[1 + 2\sum_{r=-z+1}^{z-1}\{f_{p_0,q_0}^{p_0+1,q_0+2r}\, f_{p_0,q_0}^{p_0-1,q_0-2r}\}^{\frac{1}{2}}\right]. \qquad (23)$$

Once $f_{p,q}^{p',q'}$ are known as functions of p, q then (23) gives E as a function of two parameters p_0, q_0 and we can find the values of these which give the minimum value for E. Near (p_0,q_0)

$$a(p,q) \approx a(p_0,q_0)\exp\left\{\approx \frac{1}{N}(p-p_0)^2 + \approx \frac{1}{N}(q-q_0)^2\right\}. \qquad (24)$$

Hence $a(p,q)$ is extremely small except for values of p and q negligibly, $\approx N^{\frac{1}{2}}$, different from p_0, q_0. It is therefore possible to think of the system as actually being in the 'state' specified by the quantum numbers p_0, q_0.

If there is only one point, p_0, q_0, which makes E a minimum in (23) then, because the physical problem is not changed if $p \to \frac{1}{2}N - p$, we must have $p_0 = \frac{1}{4}N$. This corresponds to a disordered ground state and antiferromagnetism is impossible. If there are two points which make E a minimum in (24) then they must be (p_0,q_0) and $(\frac{1}{2}N - p_0, q_0)$ and since at least one of them must give a maximum of $a(p,q)$ it follows from symmetry that they both give maxima of $a(p,q)$. In this situation there is an overwhelming probability for $p \approx p_0$ or $p \approx \frac{1}{2}N - p_0$ and it is still true that p has equal chances of being greater or less than $\frac{1}{4}N$. However, if it is determined at some instance of time that $p \approx p_0$ then, because the interaction between the parts $p \approx p_0$ and $p \approx \frac{1}{2}N - p_0$ is so weak the system will take a very long time to rotate into the state $p \approx \frac{1}{2}N - p_0$. Such a long time in fact—Anderson (1952) estimates ≈ 3 years—that the possibility can be neglected. This is the same as saying that the wave function represented by just one-half of (10), the half with $p \approx p_0$, while not a true stationary state, is stationary for all practical purposes. Thus this 'state' shows ordering and antiferromagnetism is possible. If there were three or more points giving a minimum value for E in (23) then we would have to be very careful in deciding which were maxima of $a(p,q)$, but in practise this case does not occur and so we shall not consider it.

For the particular case of the linear chain problem, (23) reduces to the equation derived by Kasteleijn (1952) by a different method. The original treatment by Hulthén (1938) involves a simpler trial function and, like Kasteleijn's paper, is applicable only to the linear chain problem. Consequently the equation derived by Hulthén is similar to (23) but is simpler and contains less information.

3. The transition amplitudes

In this section the quantities $f_{p,q}^{p',q'}$ which Kasteleijn (1952) calls 'transition amplitudes' will be calculated as functions of p and q. The corresponding quantities which Hulthén needed in his simpler treatment were determined by exact counting.

This is very complicated even in the linear-chain case and an extension of his method to two or three dimensions is impossible. Kasteleijn uses a much simpler method to calculate the transition amplitudes but here again the extension of his method to two and three dimensions is difficult. We shall look at the problem in a rather different way.

First of all it is convenient to change the notation of calling spins 'plus' or 'minus' to that of calling them 'right' or 'wrong' in such a way that a plus spin on the A sublattice is 'right' and a minus spin on A is 'wrong' while a plus spin on the B sublattice is 'wrong' and a minus spin on B is 'right'. The words 'right' and 'wrong' will be abbreviated as R and W respectively. With this notation the perfectly ordered state $\ldots + - + - + - \ldots$ is written $\ldots RRRRRR \ldots$. An antiparallel $(+ -)$ neighbouring pair of spins must have both spins right (RR) or both spins wrong (WW). A parallel neighbouring pair must be RW or WR. p, the number of plus spins on the A sublattice is also the number of R spins on A. The number of R spins in the whole lattice is therefore $2p$.

Recalling the definition of $f_{p,q}^{p',q'}$, we see that we want the probability of finding certain arrangements of spins surrounding a $(+ -)$ nearest neighbour pair averaged over all those distributions with fixed values of p and q. For example, with the simple plane lattice we want the probability of various arrangements of spins over the cluster of eight atoms shown in figure 1 when we are given that the centre pair is $+ -$, i.e. either RR or WW, and that there are $2p$ R spins and $q(+ -)$ pairs in the whole distribution. Thus, for instance, $f_{p,q}^{p-1,q+6}$ is the probability, given 12 is $+ -$, that the cluster has the arrangement

$$W\,W \qquad\qquad + - $$
$$W\,R\,R\,W, \text{ i.e. } + + - -.$$
$$W\,W \qquad\qquad + - $$

For, on interchanging the central $+ -$ pair, we get

$$W\,W \qquad\qquad + - $$
$$W\,W\,W\,W, \text{ i.e. } + - + -.$$
$$W\,W \qquad\qquad + - $$

Comparing these arrangements we see that the number of R spins has decreased by two, i.e. $p \to p-1$, and the number of $+ -$ pairs has increased by 6, i.e. $q \to q+6$.

FIGURE 1. The cluster for the simple plane lattice.

The problem is now transformed in a way exactly analogous to the transformation from a canonical to a grand canonical ensemble in statistical mechanics. Instead of averaging over just those distributions with fixed values of p and q, the averaging process is extended over all possible distributions with every possible value of p and q but the distributions are weighted with a factor $t^{-q}s^{-p}$. t and s are

chosen to make the mean values of p and q in this weighted ensemble of distributions equal to the values which we wish to examine. Then by analogy with the statistical mechanics problem, the standard deviation of p and q from their mean values will be $\approx \sqrt{N}$ and the value of any smoothly varying function of p, q, in particular the function $f_{p,q}^{p',q'}$, averaged over the ensemble will be equal to value of the function at the mean values apart from terms $\approx N^{-1}$.

In this ensemble, if the effects of spins outside the cluster could be ignored then the probability of finding any small cluster of spins would depend solely upon the number of R spins and the number of RR and WW pairs in the cluster. Or, what comes to the same thing, to construct the relative probability of a cluster we would write a factor s for every W spin and a factor t for every RW nearest neighbour pair in the cluster. But the effect of spins outside the cluster can only be ignored if there is no preference for them to be either R or W or to form RR or WW pairs and this is only so for $t = s = 1$.

We now notice that there is an exact equivalence between the factors t and s and factors which would be introduced in the construction of the partition function for an Ising ferromagnetic lattice. In the latter problem we define a plus sign as R and a minus spin as W so that the saturated state is $\ldots R R R R R \ldots$. Then, if a neighbouring RW pair has an energy V greater than a neighbouring RR or WW pair, each RW pair in a distribution gives a factor $\exp(-V/kT)$ in constructing the partition function. In addition, if a magnetic field H is applied each W spin has an energy $2\mu H$ greater than a R spin and so each W spin gives a factor $\exp(-2\mu H/kT)$. It is therefore clear that the probability of finding various distributions of spins over clusters in this Ising lattice is exactly the same as in our problem if $\exp(-V/kT)$ is replaced by t and $\exp(-2\mu H/kT)$ is replaced by s because the effect of the spins outside the cluster must be the same in each case. Now this Ising problem has been studied a great deal and though as yet it has been solved exactly only in a few special cases, numerous approximation methods have been developed to consider it and these we can take over to use in our problem.

Leaving aside for the moment which method is going to be applied to this problem let us consider in more detail what it is that we wish to calculate. Let $(RR)_n$, which is a complicated function of s and t, be the relative probability of finding in the ensemble of distributions a RR pair surrounded by a 'shell' containing n W spins, summed over all possible positions of these n W spins in the shell. Then $0 \leqslant n \leqslant 2\gamma$, where $\gamma = z - 1$ and hence 2γ is the number of atoms in the shell which in the simple plane case consists of atoms 3, 4, 5, 6, 7 and 8 in figure 1. Then the relative probability of finding a RR pair summed over all possible shells is

$$(RR) = \sum_{n=0}^{2\gamma} (RR)_n. \tag{25}$$

Define $(WW)_n$, (WW), (RW), etc., similarly.

The equations determining t and s in terms of p and q are

$$\{(RR) + (WW)\}\{(RR) + (WW) + 2(RW)\}^{-1} = q/\tfrac{1}{2}Nz, \tag{26}$$

$$\{(RR) + (RW)\}\{(RR) + (WW) + 2(RW)\}^{-1} = p/\tfrac{1}{2}N. \tag{27}$$

These equations are obtained by asking for the probability of finding a $(+-)$ pair or a R spin respectively in the ensemble.

We now evaluate the sum in (23). Remembering the definitions of $f_{p,q}^{p',q'}$, etc. we see that

$$f_{p,q}^{p+1,\,q+2r} = \{(RR)+(WW)\}^{-1}(WW)_{\gamma-r} \quad \text{and} \quad f_{p,q}^{p-1,\,q-2r} = \{(RR)+(WW)\}^{-1}(RR)_{\gamma-r}. \tag{28}$$

Hence
$$\sum_{r=-\gamma}^{\gamma} \{f_{p,q}^{p+1,\,q+2r} f_{p,q}^{p-1,\,q-2r}\}^{\frac{1}{2}} = \{(RR)+(WW)\}^{-1} \sum_{n=0}^{2\gamma} \{(WW)_n (RR)_n\}^{\frac{1}{2}}. \tag{29}$$

Because they have the same shells $(RR)_n$ and $(WW)_n$ are simply related to one another and it is easily verified that

$$(WW)_n = s^2 t^{2\gamma-2n}(RR)_n. \tag{30}$$

Hence (29) becomes
$$st^{\gamma}\{(RR)+(WW)\}^{-1} \sum_{n=0}^{2\gamma} t^{-n}(RR)_n = st^{\gamma}\{(RR)+(WW)\}^{-1}(NN), \tag{31}$$

where this equation defines (NN).

Substituting from (26) and (31) into (23) gives finally

$$E = -\tfrac{1}{2}NJz \times \text{maximum of } G(s,t), \tag{32}$$

$$G(s,t) = 1 + 2\{st^{\gamma}(NN)-(RW)\}\{(RR)+(WW)+2(RW)\}^{-1}. \tag{33}$$

This is the expression to be evaluated.

The function $G(s,t)$ has a very important property which will be needed later on. In the ensemble W spins have been weighted with a factor s and this is exactly equivalent to weighting R spins with a factor s^{-1}. Hence since the physical situation is unchanged by interchanging the words R and W it follows that

$$G(s,t) = G(1/s,t) \tag{34}$$

and also all derivatives of $G(s,t)$ with respect to t must obey similar equations. Of course when s is replaced by s^{-1} in (27) then p goes to $\tfrac{1}{2}N-p$.

Only for the linear chain can (33) be evaluated exactly and for other lattices some approximate solution of the statistical problem must be used. We shall use the solution most easily adapted to this problem, that first developed by Bethe (1935) and Peierls (1936) for the Ising model of ferromagnetism. For the latter problem it is known to give a good approximation to the correct solution. In §5 the accuracy of this method with regard to our problem will be discussed.

In the Bethe–Peierls method when constructing the relative probability of any distribution of spins over a cluster we write a factor s for every W spin and a factor t for every RW pair in the cluster and also, to take account approximately of the effect of spins outside the cluster, we write an additional factor λ_1, for every W spin in the shell. Thus in the simple plane lattice with the cluster drawn in figure 1, the relative probability of any arrangement of spins over the cluster is obtained by writing t for every RW nearest neighbour pair, s for every W spin on a 'centre'

atom like 1 or 2 and a factor $\lambda = s\lambda_1$ for every W spin on a shell atom like 3, 4, 5, 6, 7 or 8. In this way we can write down all the quantities $(RR)_n$, (RR), etc., as simple functions of s, t and λ. We now determine λ in terms of s and t by the consistency equation

$$(WW) + (RW) = \frac{\lambda}{2\gamma} \frac{\partial}{\partial \lambda} \{(RR) + (WW) + 2(RW)\}. \tag{35}$$

The right-hand side of (35) is a convenient way of evaluating the relative probability of a shell spin being W while the left-hand side is the relative probability of a centre spin being W. Thus (35) ensures that the probability of a spin being W is the same whether it is considered as a 'shell' or a 'centre' spin.

So far p and q have been treated as the independent variables while t, s and λ have been regarded as dependent upon them through (26), (27) and (35). From now on, however, t and λ will be treated as the independent variables and (35) will be regarded as an equation to determine s while (26) and (27) give p and q in terms of t and λ. Hence (33) is now regarded as a function of t and λ, say $F(\lambda, t)$. Clearly the property corresponding to (34) is

$$F(\lambda, t) = F(1/\lambda, t). \tag{36}$$

Apart from the linear chain case the method used here will not give as good an approximation as the standard Bethe–Peierls theory because in the latter the clusters considered have only one central atom and all the shell atoms are equivalent to one another. In our problem shell atoms like 3 and 8 in figure 1 are not equivalent and the effect of outside atoms on these will be different. We should therefore use one long range parameter for W spins on 8 or 5 and another parameter for W spins on 3, 4, 6 or 7. This has been done for the simple cubic case, although the calculation will not be given here, and it was found that the effect of the non-symmetrical shell was unimportant. It therefore seems reasonable to assume it unimportant for other lattices too. Of course these remarks do not apply to the linear chain problem because the two shell atoms are equivalent in this case. In fact as will be shown in §4·1, for the linear chain the Bethe–Peierls method is exact.

In looking for the maxima of $F(\lambda, t)$ only points for which $\lambda > 0$ and $t > 0$ need be considered because λ or t negative gives probabilities which are meaningless. Also, values of $t > 1$ correspond to situations in which RW pairs, i.e. $++$ and $--$ pairs, are more common than RR or WW pairs. Such situations are of no importance for the antiferromagnetic problem so attention need be confined to the range $0 \leqslant t \leqslant 1$. In addition, values of $F(\lambda, t)$ for $\lambda > 1$ will be given in terms of values for $\lambda < 1$ by (36) and we need only consider the range $0 \leqslant \lambda \leqslant 1$.

4. Results for various lattices

4·1. *The linear chain*

The linear-chain problem, though not physically interesting, provides a useful check on the accuracy of the trial function (8) because the correct ground state eigenvalue has been determined exactly by Bethe (1931). The cluster for the linear

chain is particularly simple and has just four atoms, 4 1 2 3, and we can immediately write down

$$(RR) = (1+\lambda t)^2, \quad (RW) = st(1+\lambda t)(\lambda + t),$$
$$(WW) = s^2(\lambda + t)^2, \quad (NN) = (1+\lambda)^2.$$

$$(37)$$

To obtain (RR) we must make R both 1 and 2 and sum over all possible shells, i.e. over all possible arrangements of spins on 3 and 4. Now instead of calculating $(RR)_0$, $(RR)_1$ and $(RR)_2$ separately and summing we can sum over the contributions from each shell atom separately. Thus if 3 is R it gives a term 1, and if it is W it gives a term λt because 2 and 3 then form a RW pair and 3 is a W shell atom. Summing over the possibilities for 3 therefore gives a factor $1 + \lambda t$. 4 gives a similar factor. Hence we obtain (RR) as given by (37). The other expressions in (37) are derived similarly.

Equation (35), which is regarded as an equation to determine s, has the simple root

$$s = \lambda(1+\lambda t)(\lambda + t)^{-1}. \tag{38}$$

It is interesting to note at this point that (37) and (38) give an exact solution for the Ising ferromagnetic problem. Thus the magnetic moment, M, of the latter problem is given by

$$M/N\mu = \{(RR) - (WW)\}\{(RR) + (WW) + 2(RW)\}^{-1} = \frac{1-\lambda^2}{1+\lambda^2+2t\lambda}. \tag{39}$$

After substituting $t = \exp(-V/kT)$ and $s = \exp(-2\mu H/kT)$ and using (38) to solve for λ, (39) gives the exact solution derived by Ising (1925). The reason for this is easy to see; in the linear chain the parameter λ takes account of the effect of atoms outside the cluster exactly and so there is no approximation involved in the Bethe–Peierls method. It follows that in our problem the only approximation for the linear chain case is that involved in the trial function (8).

Equations (33), (37) and (38) give eventually

$$F(\lambda, t) = 1 + \frac{2\lambda t(1-t)(1+\lambda+\lambda^2+\lambda t)}{(1+\lambda t)(\lambda + t)(1+\lambda^2+2\lambda t)}. \tag{40}$$

We now wish to find the maxima of $F(\lambda, t)$. After some tedious algebra it is possible to show that the only solution of $\frac{\partial}{\partial \lambda}\{F(\lambda, t)\} = 0$, which is a necessary condition for a maximum of $F(\lambda, t)$, is $\lambda = 1$, apart from trivial solutions such as $\lambda = 0$, $t = 0$, etc. which do not correspond to maxima. Substituting $\lambda = 1 - \epsilon$ into (40) and expanding to second order in ϵ gives

$$F(1-\epsilon, t) = 1 + \frac{t(1-t)(3+t)}{(1+t)^3} - \epsilon^2 \frac{t(1-t)(1+6t+t^2)}{2(1+t)^5}. \tag{41}$$

Hence, $\frac{\partial^2}{\partial \lambda^2}\{F(\lambda, t)\} < 0$ for all t in the range $0 < t < 1$ along $\lambda = 1$. $F(1, t)$ has one maximum, of value $1 \cdot 3156$, which occurs at $t = -5 + 2\sqrt{7} = 0 \cdot 2916$. This is therefore the only maximum point of $F(\lambda, t)$.

Substituting $\lambda = s = 1$ into (27) gives $p = \frac{1}{4}N$, i.e. the lattice is disordered in the ground state and hence antiferromagnetism is impossible. (26) gives

$$q = \tfrac{1}{2}Nz \times 0 \cdot 7742,$$

i.e. 77 % of the nearest neighbour pairs in the chain one antiparallel. The ground state eigenvalue obtained from (32) is the same as that obtained by Hulthén (1938) and Kasteleijn (1952).

$$E_0 = -JN \times 1{\cdot}3156. \tag{42}$$

This compares quite well with Bethe's (1931) rigorous result of $E_0 = -JN \times 1{\cdot}3863$ and, remembering the change in the zero of energy in (1), lies within the rigorous limits set upon E_0 by Anderson (1951). The trial function (8) therefore seems to be quite a good one.

4·2. *The square plane and simple cubic lattices*

In principle the calculations for these lattices are the same as for the linear chain. However, the algebra involved is more tedious chiefly because (35) no longer has a simple root like (38). The cluster for the simple plane lattice is shown in figure 1. The cluster for the simple cubic lattice has four additional shell atoms, two vertically above and two vertically below 1 and 2. Because of the similarity of these clusters it is convenient to treat both lattices together.

The first task is to write down (RR), (WW), etc. To illustrate the method of doing this consider (RR). We suppose that in figure 1 both 1 and 2 are R. Then, as in the linear chain case, we can sum over the contributions of 5 and 8 individually and these give together a factor $(1+\lambda t)^2$ in (RR). However, the contributions from 3, 4, 6, 7, etc., cannot be summed individually like those from 5 and 8 because 34, 67, etc., are nearest neighbour pairs and so the contribution from, say 4, will depend upon whether 3 is R or W. We must therefore sum over the arrangements considering the pairs 34, 67, etc., as units. Thus if 43 is RR then the group

$$4 \quad 3$$
$$1 \quad 2$$

contributes a term 1. If 43 is RW we get λt^2 for there is now a W spin in the shell and two, 43 and 32, RW pairs. If 43 is WR we get another term λt^2 while if 43 is WW we get $\lambda^2 t^2$ for then there are two W spins in the shell and two, 41 and 32, RW pairs. The total contribution from 43 is therefore $(1 + 2\lambda t^2 + \lambda^2 t^2)$. Each pair like 43 gives a similar factor and there are $(\gamma - 1)$ such pairs, where $\gamma - 1 = z - 2$ is 2 or 4 for the plane or simple cubic lattice respectively. Hence for both lattices

$$(RR) = (1+\lambda t)^2 (1 + 2\lambda t^2 + \lambda^2 t^2)^{\gamma-1}. \tag{43}$$

Similarly
$$\left.\begin{aligned}
(WW) &= s^2 (\lambda + t)^2 (t^2 + 2\lambda t^2 + \lambda^2)^{\gamma-1}, \\
(RW) &= st^\gamma (t + \lambda)(1 + \lambda t)(1 + \lambda + \lambda t^2 + \lambda^2)^{\gamma-1}, \\
(NN) &= (1+\lambda)^2 (1 + 2t\lambda + \lambda^2)^{\gamma-1}.
\end{aligned}\right\} \tag{44}$$

The consistency equation (35) is now a quadratic equation for s with complicated coefficients and its solution does not have the simple form that it has in the linear chain case. Because of this, to try to solve the simultaneous equations

$$\frac{\partial F}{\partial \lambda} = \frac{\partial F}{\partial t} = 0$$

rigorously is too complicated and so we must plot numerically the two loci

$$(a) \quad \frac{\partial F}{\partial \lambda} = 0 \quad \text{and} \quad \frac{\partial^2 F}{\partial \lambda^2} < 0,$$

$$(b) \quad \frac{\partial F}{\partial t} = 0 \quad \text{and} \quad \frac{\partial^2 F}{\partial t^2} < 0. \tag{45}$$

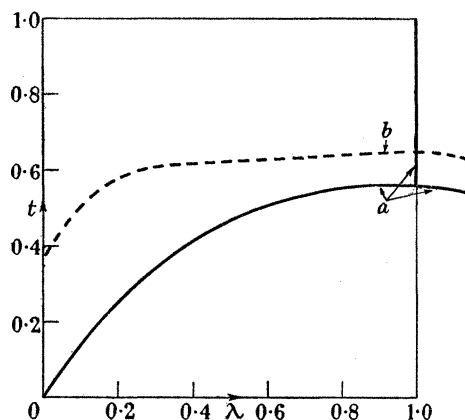The result of doing this for the simple cubic lattice is shown in figure 2.

FIGURE 2. The loci (*a*) and (*b*) for the simple cubic lattice.

The behaviour of the loci on and near $\lambda = 1$ is of special interest and can be established analytically by expanding (43), (44), the solution of (35), and hence (33) in powers of $(1 - \lambda) = \epsilon$. This gives

$$F(1 - \epsilon, t) = 1 + \frac{2^{\gamma+1} t^{\gamma} (1 + t)^{\gamma-3} - t^{\gamma} (3 + t^2)^{\gamma-1}}{(1 + 3t^2)^{\gamma-1} + t^{\gamma} (3 + t^2)^{\gamma-1}} + \tfrac{1}{2} \epsilon^2 \theta(t), \tag{46}$$

where $\theta(t)$ is a complicated function of t which will not be given here. The fact that in the expansion (46) the coefficient of ϵ vanishes shows that $\partial F/\partial \lambda = 0$ for all points on $\lambda = 1$. Indeed this can be shown from the general relation (36) without knowing the explicit form of $F(\lambda, t)$. A consideration of the sign of $\theta(t)$ shows which part of the line $\lambda = 1$ lies on locus (*a*). As figure 2 shows, this part is that for which $1 > t > 0.56$ in the simple cubic case. At the point upon $\lambda = 1$, where $\partial^2 F/\partial \lambda^2$ changes sign, i.e. $\theta(t) = 0$, which in the simple cubic case occurs at $t = 0.56$ (and in the linear chain case does not occur at all), the locus (*a*) breaks away from the line $\lambda = 1$. When it does so it is initially perpendicular to $\lambda = 1$, as shown in figure 2, because from (36) it is possible to show, no matter what the explicit form of $F(\lambda, t)$ may be, that if the loci (*a*) and (*b*) cross the line $\lambda = 1$ then they must do so perpendicularly.

$F(1, t)$, obtained from (46) by putting $\epsilon = 0$, has one maximum which for the simple cubic case occurs at $t = 0.6505$. As mentioned above, the locus (*b*) is perpendicular to $\lambda = 1$ when it passes through this point.

The behaviour of the loci near $\lambda = 0$ can also be determined analytically by expanding in powers of λ. Other points of the loci have been determined numerically.

It is seen from figure 2 that the only point of intersection of the two loci is $\lambda = 1$, $t = 0\cdot6505$. Since from (27), (43) and (44) all points on the line $\lambda = 1$ correspond to 'states' with $p = \frac{1}{4}N$. This means that the ground state of the system is disordered and antiferromagnetism is impossible. The energy eigenvalue of the ground state, given by (32) and (46) by putting $\lambda = 1$, $t = 0\cdot6505$ in $F(\lambda, t)$ for the simple cubic case, is

$$E_0 = -\tfrac{1}{2}NJz1\cdot1015 \quad \text{(for } z=6\text{).} \tag{47}$$

The behaviour for $F(\lambda, t)$ for the simple plane lattice is qualitatively similar to that for the simple cubic lattice and will not be discussed in detail here. The maximum of $F(\lambda, t)$ for this case occurs at $\lambda = 1$, $t = 0\cdot5280$, i.e. at a disordered 'state' and so antiferromagnetism is impossible in this case also. The energy eigenvalue is

$$E_0 = -\tfrac{1}{2}NJz1\cdot1560 \quad \text{(for } z=4\text{).} \tag{48}$$

4·3. *The body-centred cubic lattice*

The determination of (RR), (WW), etc., for this lattice is difficult and the final expressions are very complicated. This is because the shells have a complex structure; some shell atoms are nearest neighbours of three other shell atoms, some are nearest neighbours of only one and two are nearest neighbours of none. It can be shown that if we define

$$f\{A; B; C; D\} = A^3(C+Dt^2)^3 + 3A^2Bt^2(C+D)^2(C+Dt^2)$$
$$+ B^3(D+Ct^2)^3 + 3B^2At^2(C+D)^2(D+Ct^2), \quad (49)$$

then

$$\left.\begin{aligned}
(RR) &= (1+\lambda t)^2 f\{1+\lambda t^2; \lambda t^2(1+\lambda); 1+\lambda t^2; \lambda t^2(1+\lambda)\}, \\
(WW) &= s^2(\lambda+t)^2 f\{t^2(1+\lambda); \lambda(\lambda+t^2); t^2(1+\lambda); \lambda(\lambda+t^2)\}, \\
(RW) &= st^7(1+\lambda t)(\lambda+t)f\{1+\lambda; \lambda(\lambda+t^2); 1+\lambda t^2; \lambda(1+\lambda)\}, \\
(NN) &= (1+\lambda)^2 f\{1+\lambda t; \lambda(t+\lambda); 1+\lambda t; \lambda(t+\lambda)\}.
\end{aligned}\right\} \tag{50}$$

Because of the complicated nature of these expressions, $F(\lambda, t)$ for the body-centred cubic lattice has not been investigated as thoroughly as for the simple cubic case but it has been determined that the behaviour near $\lambda = 0$ and $\lambda = 1$ is qualitatively similar to that in the simple cubic case. Thus a maximum of $F(\lambda, t)$ occurs on the line $\lambda = 1$ at $t = 0\cdot7175$ and of value $1\cdot0785$ in this case.

The behaviour of the loci in regions away from $\lambda = 0$ or $\lambda = 1$ has not been investigated and therefore we cannot be certain that they do not intersect at any other point. However, in view of their behaviour near $\lambda = 0$ and $\lambda = 1$ it seems unlikely that they will intersect at any point for which $\lambda < 1$ and so it will be assumed that the point $\lambda = 1$, $t = 0\cdot7175$ is the only maximum of $F(\lambda, t)$ for this case. In principle this assumption is easily verified but in practice a great deal of numerical calculation would be required and in view of arguments which will be presented in § 5 this would probably not be worth while.

Subject to this assumption then, the ground state of this lattice is disordered also and has an energy

$$E_0 = -\tfrac{1}{2}JNz1\cdot0785 \quad \text{(for } z=8\text{).} \tag{51}$$

5. DISCUSSION OF THE RESULTS

The results of § 4 show that, subject to the assumptions (a), (b), (c) and (d) given in § 2, all lattices have disordered ground states and hence antiferromagnetism is impossible under these conditions. The theory leading to this conclusion involves two approximations. The first of these is the form assumed for the trial function (8). This is a very natural form to choose but it is difficult to assess how good an approximation it is. As pointed out in § 4·1, for the linear chain case where no further approximation is required, an eigenvalue (42) is obtained which agrees quite well with Bethe's rigorous result and this indicates that (8) is probably a good trial function. But the agreement with Bethe's result is not exceptionally good and so we cannot be sure that this last conclusion is correct.

For two- and three-dimensional lattices a further approximation is made because the Bethe–Peierls method is used to evaluate the transition amplitudes of § 3. It is therefore no longer true that the eigenvalues obtained are necessarily upper bounds to the true eigenvalues for these lattices. Nevertheless, it is interesting to compare these results with those of Kubo's (1953) variational calculations which, in complete contrast to these conclusions, predict all lattices to have ordered ground states with the same assumptions (a), (b), (c) and (d). Remembering the change of origin in (1) these are:

		this calculation	Kubo's calculation	
linear chain	$2E/JNz =$	$-1{\cdot}3156$	$-1{\cdot}198$	
simple plane	$2E/JNz =$	$-1{\cdot}1560$	$-1{\cdot}147$	
simple cubic	$2E/JNz =$	$-1{\cdot}1015$	—	(52)
body-centred cubic	$2E/JNz =$	$-1{\cdot}0785$	$-1{\cdot}069$	

Because the eigenvalue for the linear chain is so very much better than Kubo's we may safely conclude that the trial function is better than his. This is satisfying because it is believed that the ground state of an antiferromagnetic linear chain is certainly disordered (Anderson 1952). The other eigenvalues obtained are a little lower than Kubo's but, for the reason stated above, it cannot be inferred that the trial function (8) is better than his for these lattices.

It is not easy to see if the 'first shell' Bethe–Peierls approximation is accurate enough for us to have confidence in these conclusions. The critical factor determining whether the ground state is ordered or not is the sign of $\frac{\partial^2}{\partial\lambda^2} F(\lambda, t)$ evaluated at the maximum of the disordered ($\lambda = 1$) state. For the simple cubic lattice this quantity is $-0{\cdot}0296$. Now the better calculation which was mentioned in § 3, that using two long-range parameters to account for the asymmetry of the 'shell', shows no qualitative change from the simpler calculation presented here and the same eigenvalue (3·11) is obtained but $\frac{\partial^2}{\partial\lambda^2} F(\lambda, t)$ at the maximum becomes $-0{\cdot}0158$.

This difference is partly due to a change of 'scale' in λ, i.e. a given point (λ, t) does not represent the same state in both calculations. But even when allowance is made

W. Marshall

for this fact, the better calculation still shows a smaller, by a factor $\approx \frac{2}{3}$, energy difference between the best disordered state and a given state which is slightly ordered than the original calculation. A 'second-shell' Bethe–Peierls calculation would be expected to produce effects of the same order of magnitude as these corrections. Thus it may well be that such a calculation would reverse the sign of $\frac{\partial^2}{\partial \lambda^2} F(\lambda, t)$ at the best disordered state and so give an ordered ground state although, as far as we can judge, it appears more probable that the sign of this quantity will be unchanged and our conclusions confirmed.

It would therefore be particularly interesting if (33) were investigated using a better approximation than the 'first shell' Bethe–Peierls method. Until this is done the question of whether the ground states of these lattices are ordered or not must remain open. Despite this, from these results we can draw the important conclusion
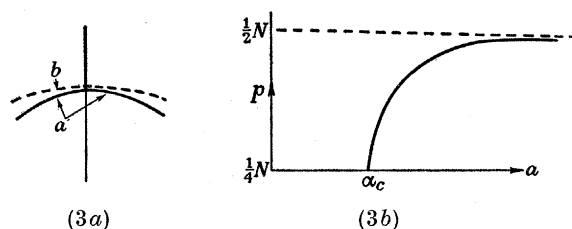

FIGURE 3. The influence of additional forces upon the order existing in the ground state. α measures the strength of the additional forces. p is the number of 'right' spins on each sub-lattice. (3a) shows the loci (a) and (b) at the critical point $\alpha = \alpha_c$.

that the energy differences between the disordered and some ordered states are not large. Indeed this is clear because the results of the Kubo variational calculation (52) show that there are some highly ordered states which have energies close to those obtained here for the best disordered states. Hence the nature of the ground state is likely to be sensitive to even small anisotropy forces or small next-nearest neighbour interactions.

In connexion with this last point, assuming the lattice to be disordered with only nearest-neighbour exchange interactions present, some qualitative conclusions can be drawn regarding the effect of small additional forces. For example, suppose an additional ferromagnetic next-nearest neighbour interaction of strength $J\alpha$ were added to the Hamiltonian (1). Then this term would increase the tendency towards order and thus the two loci (a) and (b) of figure 2 would move closer together. At some critical value of α, say α_c, the loci would just touch at some point on $\lambda = 1$ and the situation near the point of intersection would be as shown in figure (3a).

If α is now increased slightly, (b) moves downwards and (a) moves upwards so that the point of intersection moves rapidly away from the line $\lambda = 1$. In fact, since the loci are perpendicular to $\lambda = 1$ at the point of crossing, the tangent to a curve of order, i.e. the number of R spins in the lattice ($2p$), plotted against α would be vertical at α_c. Over a short range of α just greater than α_c, p will continue to change rapidly because the loci still intersect at a narrow angle and the p-α curve will therefore look like figure (3b). Because the first shell Bethe–Peierls method is not

quite accurate enough for our purposes it is not worth while investigating if it gives a reasonable value for α_c. However, for the linear chain case where this method is exact, it is possible to show that α_c is infinite, i.e. the ground state of the linear chain remains disordered no matter how strong the next-nearest neighbour interactions may be. This behaviour is probably not typical of other lattices.

A more efficient method of increasing the tendency towards order is through anisotropic forces. Clearly, if in some sense α measures the strength of these forces then the p-α curve must look like figure $(3b)$ for the previous analysis follows through as before.

The experimental evidence for the presence of strong additional forces is good. Thus for many antiferromagnets the anisotropic and exchange interactions are comparable (Lidiard 1954). Also all antiferromagnets of the 'two sublattice' type (except the metal chromium, to which this theory does not apply anyway) have rectangular unit cells in which two of the atoms we call 'next nearest neighbours' are actually the closest to the one we are considering and their effects would therefore be expected to be comparable to the 'nearest neighbour' interactions.

6. Conclusion

It has certainly not been proved rigorously that lattices are disordered when only nearest neighbour antiferromagnetic exchange interactions are present. However, a method has been established by which this question may be examined more closely and this would be an interesting calculation to perform especially in view of the complete contradiction between these results and those of the Kubo (1952) variational calculation. The latter, although it uses the Holstein–Primakoff (1940) notation, avoids all the weaknesses of the normal 'spin-wave' theory of antiferromagnetism. Kubo does, however, make some approximations which may not be valid and presumably it is because of these that his method breaks down for the linear chain. For other lattices it seems probable that his method gives the better approximation to the energies of highly ordered states while this calculation gives the better approximation for disordered and slightly ordered states. In this sense therefore the theories are complementary and it is a delicate question as to which theory gives the correct conclusion regarding the order existing in the ground state.

Appendix A

Here it is proved that the ground state is a singlet whenever (6) is true. To simplify the notation only the linear chain problem with $S = \frac{1}{2}$ will be considered. The generalization to two and three dimensions is obvious and it will be shown later how the proof can be extended to cases when $S \geqslant 1$.

Suppose the ground state eigenfunction with zero total z component of spin is

$$\psi_0 = \sum_\mu c_\mu \{\mu\}. \tag{A 1}$$

If $\mathbf{S}^+ = \sum_{e=1}^{N} \mathbf{S}_e^+$, then $\psi_1 = \mathbf{S}^+ \psi_0$ is identically zero if $\mathbf{S}^2 \psi_0 = 0$, and otherwise is a state degenerate with ψ_0 and with total z component of spin equal to one. Now

$$\psi_1 = \sum_\mu c_\mu \sum_e \mathbf{S}_e^+ \{\mu\} = \sum_{\mu_1} \{\mu_1\} \sum_\mu c_\mu, \tag{A 2}$$

where in this last expression the sum over μ_1 goes over all arrangements with $m_{\mu_1} = 1$ and $\{\mu\}$ is an arrangement obtained from $\{\mu_1\}$ by replacing a '$+$ spin' by a '$-$ spin'. The inner sum therefore contains $\frac{1}{2}N+1$ terms.

In (A 2) consider the coefficients of the arrangements $\{\mu_1\}$ and $\{\mu_1^t\}$, where $\{\mu_1^t\}$ is obtained from $\{\mu_1\}$ by translating it one step in some direction, i.e.

$$\text{if} \qquad \{\mu_1\} = \ldots + - + + - - + - \ldots$$

$$\text{then} \quad \{\mu_1^t\} = \quad \ldots + - + + - - + - \ldots.$$

Suppose $\{\mu_1\}$ has plus spins on the atoms $n_1, n_2, \ldots, n_{\frac{1}{2}N+1}$, and $\{\mu_1^t\}$ has plus spins on the atoms $n_1+1, n_2+1, \ldots, n_{\frac{1}{2}N+1}+1$. Then

$$\sum_\mu c_\mu = c(n_2, n_3, \ldots, n_{\frac{1}{2}N+1}) + c(n_1, n_3, \ldots, n_{\frac{1}{2}N+1}) + \ldots + c(n_1, n_2, \ldots, n_{\frac{1}{2}N}), \tag{A 3}$$

and

$$\sum_\mu c_\mu^t = c(n_2+1, n_3+1, \ldots, n_{\frac{1}{2}N+1}+1) + c(n_1+1, n_3+1, \ldots, n_{\frac{1}{2}N+1}+1)$$
$$+ \ldots + c(n_1+1, n_2+1, \ldots, n_{\frac{1}{2}N}+1). \tag{A 4}$$

Now $c(n_2, n_3, \ldots, n_{\frac{1}{2}N+1})$ and $c(n_2+1, n_3+1, \ldots, n_{\frac{1}{2}N+1}+1)$ must have the same magnitude because they are coefficients of two arrangements occurring in ψ_0 which differ from one another only by a translation. Because ψ_0 is the ground state we know by (6) that they must both be real and since in the translation $p \to \frac{1}{2}N - p$ it follows that

$$c(n_2, n_3, \ldots, n_{\frac{1}{2}N+1}) = (-1)^{\frac{1}{2}N} c(n_2+1, n_3+1, \ldots, n_{\frac{1}{2}N+1}+1). \tag{A 5}$$

Similarly, we can examine the other terms of (A 3) and (A 4) in pairs and so

$$\sum_\mu c_\mu = (-1)^{\frac{1}{2}N} \sum_{\mu^t} c_\mu^t. \tag{A 6}$$

But ψ_1, if it is not identically zero, is also a ground state of the system and must therefore also have coefficients satisfying (6). But when $\{\mu_1\} \to \{\mu_1^t\}$ in ψ_1 then $p \to \frac{1}{2}N - p \pm 1$. Therefore

$$\sum_\mu c_\mu = (-1)^{\frac{1}{2}N \pm 1} \sum_{\mu^t} c_\mu^t. \tag{A 7}$$

Hence from (A 6) and (A 7), $\sum_\mu c_\mu = \sum_{\mu^t} c_\mu^t = 0$ and ψ_1 is identically zero. Hence the ground state has a total spin equal to zero whenever (6) is true.

The extension of this proof to the case of $S > \frac{1}{2}$, i.e. more than one electron outside the closed shells per magnetic atom, follows immediately since a theorem corresponding to (6) can always be found. For example, if $S = 1$ the ground-state coefficients have the form

$$c_\mu = (-1)^{n_\mu} a_\mu, \tag{A 8}$$

where a_μ is real and positive and n_μ is the number of spins having $\mathbf{S}^z = 0$ on the A sublattice.

APPENDIX B

The expectation value of the square of the total spin for the trial function (10) is

$$\langle \mathbf{S}^2 \rangle = \{ \sum_{p,q} a^2(p,q) \}^{-1} \{ \sum_{p,q,p',q'} a(p',q') a(p,q) (\psi^*(p',q') \mathbf{S}^2 \psi(p,q)) \}. \tag{B 1}$$

Using (18) and (24), (B 1) reduces eventually to

$$\langle \mathbf{S}^2 \rangle = \frac{1}{2} \sum_{p',q'} \{ (\psi^*(p',q') \mathbf{S}^2 \psi(p_0,q_0)) + (\psi^*(p',q') \mathbf{S}^2 \psi(\tfrac{1}{2}N - p_0, q_0)) \} + \approx N, \tag{B 2}$$

where (p_0, q_0) is a maximum point of $a(p,q)$.

The matrix elements $\psi^*(p',q') \mathbf{S}^2 \psi(p,q)$ are calculated as follows. We note that

$$\mathbf{S}^2 \{\mu\} = \tfrac{1}{2} N \{\mu\} + \tfrac{1}{2} \sum_{n,n'} (\mathbf{S}_n^+ \mathbf{S}_{n'}^- + \mathbf{S}_n^- \mathbf{S}_{n'}^+) \{\mu\} = \tfrac{1}{2} N \{\mu\} + \sum_{\mu^+} \{\mu^+\}, \tag{B 3}$$

where $\{\mu^+\}$ differs from $\{\mu\}$ by an interchange of two opposite spins. Therefore

$$\mathbf{S}^2 \psi(p,q) = \tfrac{1}{2} N \psi(p,q) + \{ W(p,q) \}^{-\frac{1}{2}} \sum_\mu \delta_{p,p_\mu} \delta_{q,q_\mu} (-1)^{p_\mu} \sum_{\mu^+} \{\mu^+\}. \tag{B 4}$$

The second expression on the right-hand side of (B4) contains $\frac{1}{4} N^2 W(p,q)$ terms. Suppose a fraction $F_{p,q}^{p';q'}$ of them belong to $\psi(p',q')$. Then $F_{p,q}^{p';q'}$ is the probability, given a distribution with fixed p, q that these values will change to p', q' when two opposite spins are interchanged.

Using the property that $\mathbf{S}^2$ is Hermitian and remembering that

$$F_{p,q\pm 2r}^{p,q} = F_{p,q}^{p,q\mp 2r} + \approx \frac{1}{N}, \tag{B 5}$$

we can derive, in a similar manner to the derivation of (23) that

$$\sum_{p',q'} (\psi(p',q') \mathbf{S}^2 \psi(p,q)) = \tfrac{1}{4} N^2 \left\{ \begin{array}{l} F_{p,q}^{p,q} - 2\{ F_{p,q}^{p+1,q} F_{p,q}^{p-1,q} \}^{\frac{1}{2}} \\[2mm] + 2 \sum_{r=1}^{z} [\{ F_{p,q}^{p,q+2r} F_{p,q}^{p,q-2r} \}^{\frac{1}{2}} - \{ F_{p,q}^{p+1,q+2r} F_{p,q}^{p-1,q-2r} \}^{\frac{1}{2}} \\[2mm] \qquad\qquad - \{ F_{p,q}^{p+1,q-2r} F_{p,q}^{p-1,q+2r} \}^{\frac{1}{2}}] \end{array} \right\} + \approx N. \tag{B 6}$$

Of all the $\frac{1}{4} N^2$ opposite $(+ -)$ pairs in the distribution let $[R_n R_m]$ be that fraction of them which are RR and also have the property that one spin of the pair is surrounded by a shell consisting of its z nearest neighbours, containing n W and $(z-n)$

R spins while the other spin is surrounded by a shell containing $m\,W$ and $(z-m)\,R$ spins. Define $[R_n W_m]$, $[W_n W_m]$ similarly. Then

$$F_{p,q}^{p-1,q+2r} = \sum_{n=0}^{z-r} [R_{r+n}R_{z-n}], \quad F_{p,q}^{p+1,q-2r} = \sum_{n=0}^{z-r} [R_n R_{z-r-n}],$$

$$F_{p,q}^{p+1,q+2r} = \sum_{n=0}^{z-r} [W_n W_{z-r-n}], \quad F_{p,q}^{p+1,q-2r} = \sum_{n=0}^{z-r} [W_{r+n}W_{z-n}],$$ (B 7)

$$F_{p,q}^{p,q+2r} = 2\sum_{n=0}^{z-r} [R_{r+n}W_n], \qquad F_{p,q}^{p,q-2r} = 2\sum_{n=0}^{z-r} [R_n W_{n+r}].$$

Because they have the same shells $[R_n R_m]$ and $[R_n W_m]$, etc., are simply related. In fact

$$[R_n W_m] = s^{z-2m}[R_n R_m], \quad [W_n W_m] = s^2 t^{2z-2n-2m}[R_n R_m].$$ (B 8)

Substituting (B 7) and (B 8) into (B 6) gives

$$\sum_{p',q'} (\psi^*(p',q')\,\mathbf{S}^2\psi(p,q)) = \tfrac{1}{2}N^2 s\left[\sum_{n=0}^{z} t^{z-2n}[R_n R_n] + 2t^z \sum_{r=1}^{z} t^{-r} \sum_{n=0}^{z-r} t^{-2n}[R_n R_{r+n}]\right.$$

$$- \sum_{n=0}^{z} [R_n R_{z-n}] - \sum_{r=1}^{z} t^r \sum_{n=0}^{z-r} [R_n R_{z-r-n}]$$

$$\left. - \sum_{r=1}^{z} t^{-r} \sum_{n=0}^{z-r} [R_{r+n}R_{z-n}] + \approx\frac{1}{N}\right].$$ (B 9)

We can now show that the coefficient of the term $\approx N^2$ vanishes identically. For example consider terms like $[R_m R_m]$. The first term gives $t^{z-2m}[R_m R_m]$ the second term gives no contribution like this, the third term gives $-[R_m R_m]$ if $m = \tfrac{1}{2}z$, the fourth term gives $-t^{z-2m}[R_m R_m]$ if $m \leqslant \tfrac{1}{2}z - 1$ and the fifth term gives $-t^{z-2m}[R_m R_m]$ if $m \geqslant \tfrac{1}{2}z + 1$. Hence the net result is that terms like $[R_m R_m]$ vanish identically. Similarly terms like $[R_m R_p]$ vanish. Hence substituting (B 9) into (B 2),

$$\langle \mathbf{S}^2 \rangle \approx N,$$

which is (9).

References

Anderson, P. W. 1951 *Phys. Rev.* **83**, 1260.
Anderson, P. W. 1952 *Phys. Rev.* **86**, 694.
Bethe, H. A. 1931 *Z. Phys.* **71**, 694.
Bethe, H. A. 1935 *Proc. Roy. Soc.* A, **150**, 552.
Holstein, I. & Primakoff, H. 1940 *Phys. Rev.* **58**, 1098.
Hulthén, L. 1938 *Ark. Mat. Astr. Fys.* **26** A, no. 1.
Ising, E. 1925 *Z. Phys.* **31**, 253.
Kasteleijn, P. W. 1952 *Physica*, **18**, 104.
Kubo, R. 1952 *Phys. Rev.* **87**, 568.
Kubo, R. 1953 *Rev. Mod. Phys.* **25**, 344.
Lidiard, A. 1954 *Rep. Progr. Phys.* **17**, 201.
Marshall, W. 1955 *Proc. Roy. Soc.* A, **232**, 69.
Peierls, R. E. 1936 *Proc. Roy. Soc.* A, **154**, 207.
Van Vleck, J. H. 1941 *J. Chem. Phys.* **9**, 85.
Van Vleck, J. H. 1945 *Rev. Mod. Phys.* **17**, 25.
Ziman, J. M. 1952 *a Proc. Phys. Soc. Lond.* A, **65**, 540.
Ziman, J. M. 1952 *b Proc. Phys. Soc. Lond..* A, **65**, 548.
Ziman, J. M. 1953 *Proc. Phys. Soc. Lond.* A, **66**, 89.