

Document for DOCUMENT ROOM 36-412
Research Laboratory of Electronics
Massachusetts Institute of Technology

#

211

THE QUANTUM THEORY OF ANTIFERROMAGNETISM

R. J. HARRISON

LOAN COPY

TECHNICAL REPORT NO. 172

AUGUST 31, 1950

mg

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

The research reported in this document was made possible through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research) and the Air Force (Air Materiel Command), under Signal Corps Contract No. W36-039-sc-32037, Project No. 102B; Department of the Army Project No. 3-99-10-022.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS

Technical Report No. 172

August 31, 1950

THE QUANTUM THEORY OF ANTIFERROMAGNETISM*

R. J. Harrison

Abstract

Recent neutron diffraction experiments on antiferromagnetic crystals at low temperatures indicate the existence of a long range correlation in direction of the spin magnetic moments. Although the Ising model predicts such a correlation, a consideration of the uncertainty relationships for spin shows that in the correct quantum description of an antiferromagnet, the ordering cannot be the simple type given by this semiclassical model. The problem of determining the lowest quantum state of an antiferromagnet is much the same as the quantum chemical problem of determining the state of a large molecule. The short range correlation of spin moments is measured by a quantity called the "bond order" in molecular problems. This assumes its maximum in the lowest eigenstate.

In an attempt to make progress in solving the general quantum mechanical problem, a new method of approximation has been developed and applied to the linear antiferromagnetic chain. The spin eigenfunction is expanded in terms of a set of "valence bond spin functions", grouped according to "degree of excitation". The amplitudes for each excitation are determined by the solution of an infinite set of linear difference equations.

* This report is based on a Doctoral thesis in the department of Physics, Massachusetts Institute of Technology.

THE QUANTUM THEORY OF ANTIFERROMAGNETISM

I. Introduction

In the semiclassical Ising model for ferro- and antiferromagnetic substances, the energy is proportional to the sum of scalar products of the magnetic moments of neighboring pairs of the elementary magnets. The proportionality constant is positive for ferromagnetic, and negative for antiferromagnetic substances. For antiferromagnetism, the lowest state would be reached when, if geometry permitted*, neighboring magnetic moments were always antiparallel; the highest state would correspond to all magnetic moments parallel.

Recent experiments (1) give evidence that the ordering of the spin magnetic moments of an antiferromagnetic crystal below the Curie temperature is, qualitatively at least, like that predicted by the Ising theory. A more detailed analysis of the experimental results may be warranted because of differences between predictions of semiclassical theory and the quantum theory which we shall now discuss.

One of the early developments of the new quantum theory was the explanation of the chemical bond. An atom having an electron with an unpaired spin forms a bond with a similar atom, provided that the spin state of the resulting molecule is one in which the electrons of the two atoms have antiparallel spins. The mechanism of binding was described in terms of a negative exchange integral. It was shown later that positive exchange integrals could arise in certain circumstances, and that a parallel alignment of spins would then be favored. Such positive exchange integrals were characteristic of the situations where ferromagnetism occurred, and provided an explanation of this phenomenon.

Since antiferromagnetism differs from ferromagnetism in having a negative exchange integral, one returns to the theory of the chemical bond when investigating ferromagnetism. The problem of describing the quantum state of an antiferromagnet is like that of describing the state of a complex chemical molecule. Even if we confine our attention to the lowest state, as we shall do, the problem is made extremely complicated by the phenomena of "resonance" (3). Thus one may consider the example of the benzene molecule which has six π electrons with unpaired spins. Here the lowest state is a mixture of the two Kekulé structures and the three Dewar structures. As the number of spins increases, the number of bond structures among which the molecule is in resonance increases very rapidly.

There is just one state which has a relatively simple description. This is the state which corresponds to the lowest state for a ferromagnet and to the highest state for an antiferromagnet. This state, with large total spin, is not very different from its analogue in the semiclassical Ising model. On the other hand, the quantum state, which is

* Instances of geometries where one cannot have ordered antiparallel arrangements are the face-centered cubic lattice, and the two-dimensional triangular net (2).

the lowest for an antiferromagnet and has zero spin, behaves quite differently from the corresponding state in the Ising-model description. This fact makes the analogy of an antiferromagnet with a large molecule more pertinent than the analogy with a ferromagnet.

In previous work (5, 6, 7) it has been recognized that the lowest quantum state of an antiferromagnet does not exhibit perfect ordering of spins. In the discussion that follows, we try to bring out more clearly the physical concept involved in describing the state of spatial ordering of spin magnetic moments. We shall do this by drawing certain qualitative conclusions from the consideration of Dirac's well-known operator (8) for the exchange interaction. This operator, acting on the spin part of the wave function, is the operator for that part of the energy arising from the exchange interaction. Calling this operator H_1 , we have

$$H_1 = -\frac{1}{2} \alpha \sum_{i,j} (\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}) \quad (1.1)$$

where, in the nearest neighbor approximation that we consider, the sum is taken only over neighboring pairs of electrons i and j ; α is the exchange integral, and is negative in the antiferromagnetic case; $\underline{\sigma}^{(i)}$ and $\underline{\sigma}^{(j)}$ are the Pauli spin matrices operating on the spin eigenfunctions of the i 'th and the j 'th electrons.

This expression is useful because it gives the explicit dependence of the energy upon the average value of spin correlation $\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}$ between neighboring spins. If one replaces the operators $\underline{\sigma}^{(i)}$ and $\underline{\sigma}^{(j)}$ by classical vector quantities proportional to magnetic moment or to spin angular momentum, one essentially has the semiclassical expression for the energy in the Ising model of a ferro- or antiferromagnetic material. One may verify immediately for this semiclassical case the statements we made at the beginning regarding the arrangement of spins for the lowest and for the highest energy states.

In quantum mechanics, the situation takes on added complexity. Because of the properties of spin angular momentum, one cannot have a quantum mechanical state in which two electrons can be described as having parallel spin angular momentum vectors. Two neighboring spins are as nearly parallel as they can be if they have equal components in some one particular direction, while the other components are completely uncorrelated. Under these circumstances $\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}$ takes its maximum value of 1. This may be compared with the magnitude of

$$\left[\sigma^{(i)} \right]^2 = \left[\sigma^{(j)} \right]^2 = 3 \quad ,$$

for the individual electrons.

If two electrons are paired to form a singlet state, the spin angular momentum of one will be antiparallel to that of the other, but, at the same time, it will be completely

uncorrelated in direction with that of any other electron. Thus if 1 and 2 are the electrons thus paired or "bonded", $\underline{\sigma}^{(1)} \cdot \underline{\sigma}^{(2)}$ has the value -3; if k is any other electron, $\underline{\sigma}^{(1)} \cdot \underline{\sigma}^{(k)}$ and $\underline{\sigma}^{(2)} \cdot \underline{\sigma}^{(k)}$ have the average value zero.

The state in which all 2N electrons have the same components of spin in a given direction z, e.g. all spins "up", will be an eigenstate of the total spin angular momentum operator corresponding to the eigenvalue $S(S+1)h^2/4\pi^2 = N(N+1)h^2/4\pi^2$, and of the energy operator H_1 corresponding to the eigenvalue $-Nz\alpha$, where z is the number of nearest neighbors. This is the highest state for the antiferromagnet, and the lowest state for the ideal ferromagnet. It very much resembles the parallel spin case of the Ising model; the difference is that instead of all magnetic moments being parallel, the components of all moments in one specific direction are equal, with no correlation among any other components.

Let us examine what would seem to be the natural analogue of the antiparallel case of the semiclassical model. This would be an alternating "spin-up, spin-down, spin-up, spin-down, etc." arrangement in which every other electron is in an eigenstate corresponding to $\sigma_z = +1$, while its neighbor is in an eigenstate corresponding to $\sigma_z = -1$.

$$\overline{\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}}$$

would have the value -1, giving the value zero for the average of H_1 over such an arrangement. However this arrangement is not an eigenfunction of either the total spin or of the energy operator H_1 . It does not, therefore, correspond at all to the lowest state of an ideal antiferromagnet.

From the above example, and from the minimal properties of the correct eigenfunction, we may deduce that the lowest eigenvalue for the antiferromagnet is smaller than zero. While there are Nz terms in the sum (Eq. 1.1), there can be at most N pairs of neighboring electrons in singlet states. Therefore, the energy must be greater than that corresponding to

$$\overline{\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}} = -3 \quad ,$$

giving $Nz\alpha$ as the lower bound for the lowest eigenvalue*.

We shall see later that one can describe the situation in which specified pairs of electrons are in singlet states, by valence bond spin functions in which a bond connects

* Whether the difference in energies between the highest and lowest states is greater, or is smaller, in the quantum case than it is in the semiclassical case depends on how we express the correspondence between the two. If we replace the spin operators by classical angular momentum vectors of magnitude $\frac{1}{2}h/2\pi$, the energy values for the highest states will correspond, while the lowest state will have a lower energy in the quantum case. If we replace the spin operators by classical angular momentum vectors of magnitude $\frac{1}{2}\sqrt{3}h/2\pi$, then the quantum energy is lower in the highest state, and higher in the lowest state than the corresponding semiclassical values.

each such pair. The lowest energy of the system will be reached when as many pairs of neighbors will be in singlet states as much of the time as possible. This situation is realized by the system being in "resonance" among different types of bond structures, as was described for benzene.

The definition of bond order given by Penney (9) for molecular problems, is a measure of the fraction of the time a given pair of electrons spends in the singlet state. The bond order is zero if

$$\overline{\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}}$$

is zero, and is unity, if

$$\overline{\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}}$$

is equal to its value for the singlet state, -3 . Since the energy is also a measure of

$$\overline{\underline{\sigma}^{(i)} \cdot \underline{\sigma}^{(j)}}$$

the bond order may be expressed directly in terms of the energy.

We have found it useful to apply this concept of bond order in discussing antiferromagnetism since it is a parameter that is a direct measure of the short-range order of the spin magnetic moments of the electrons. The lowest state is the one that maximizes the absolute value of the short-range spin order. Thus in Slater's or Hulthén's description of a linear antiferromagnet, which we shall discuss in more detail later, the lowest quantum state is described as a superposition of states, the majority of which have arrangements differing from the regular "alternating" arrangement of spins by having about 20 percent of the spins "mixed up".

This might make it appear that the order is less than that corresponding to the alternating arrangement, but the resonance between states results in the opposite being true. The bond order for the actual lowest state of the linear chain is about 0.6, whereas, as discussed before, the alternating arrangement corresponds to a bond order of 0.33.

Returning to the experimental results on neutron diffraction in antiferromagnetic crystals, we see that it would be desirable to determine from the theoretical analysis of the scattering of neutrons in a magnetic medium exactly which parameters are measurable by neutron diffraction. Presumably these are order parameters analogous to those measurable in X-ray work on alloys. Then one could proceed to compare these measured quantities with the theoretical predictions. We have seen that one certainly expects a high degree of short-range order. However we have no theoretical information regarding other order parameters, and in particular, regarding long-range order. The experimental evidence points to the existence of at least some degree of long-range order in actual materials. It may be pointed out that in these crystals, there is the so-called "superexchange" (10, 11) interaction which is described by a somewhat more general type of operator than in Eq. 1.1. Although the considerations which preclude any

strictly ordered arrangement are independent of the form of the interaction operator, the lowest state under the more general operator will no longer be simply the one which maximizes the short-range spin correlation. It is therefore conceivable, though perhaps not likely, that if one could solve the spin state problem for both cases, one might find that long-range order occurred for one case but not the other.

We have therefore investigated the quantum mechanical solutions for antiferromagnetic models in an attempt to make headway in clarifying this situation. We present details of this investigation, and include a summary of pertinent previous work in this field.

II. Development of Formalism

Finding the lowest quantum state in a model describing an antiferromagnetic material in terms of a localized system of spins having only nearest neighbor interaction poses problems of interest and difficulty, so we shall try to set up our formalism in a way which allows us to confine our attention to such a system, disregarding for the moment the further complications which can arise in actual antiferromagnetic crystals.

We shall first show how the Hamiltonian operator acting on the complete wave function describing both the spatial and the spin coordinates of a system of $2N$ electrons can be replaced by an equivalent operator acting only upon the spin part of the wave function. (With the assumption that only nearest neighbors interact, this operator is the one described in section I.) The problem is then one of determining the eigenvalues and corresponding eigenfunctions of this spin operator. We can do this by expanding the spin eigenfunction in terms of an independent set of spin functions, and determining the coefficients in this expansion by solution of a set of linear equations whose secular equation determines the eigenvalue.

The idealized system which we shall consider consists of a set of $2N$ electrons which are more or less localized at $2N$ sites. The complete wave function Ψ for this system involves both the spatial and the spin coordinates of each of these electrons. We assume that the interaction between the electrons is described by a Hamiltonian operator H , not involving the spin, and whose form we shall leave otherwise unspecified for the moment. The function Ψ is an eigenfunction of H , such that the Schroedinger equation

$$(H - W)\Psi = 0 \tag{2.1}$$

is satisfied.

The Pauli principle requires that Ψ be antisymmetric in the coordinates of all electrons. We may construct functions of this type from a function, the "representative term", which treats the electrons as distinguishable, by forming the sum of all the $(2N)!$ functions which arise from the $(2N)!$ possible permutations of the electrons among themselves in the representative term. Each function includes a factor of plus or minus one, according as the permutation is even or odd. The representative term itself may be written as the product of a spatial part and a spin part.

We assume that Ψ may be written in this way. Thus

$$\Psi = \sum_j (-1)^j P_j \left[u(1, 2, \dots, 2N) \bar{\phi}(1, 2, \dots, 2N) \right] \quad (2.2)$$

where

$$u(1, 2, \dots, 2N) \cdot \bar{\phi}(1, 2, \dots, 2N) \quad (2.3)$$

is the representative term, with u the spatial part, and $\bar{\phi}$ the spin part; P_j is one of the $(2N)!$ operations of permuting the $2N$ electrons among themselves, acting on both u and $\bar{\phi}$; and j is even or odd when P_j is even or odd. If we insert this form of Ψ into the Schroedinger equation (2.1), multiply by

$$u(1, 2, \dots, 2N) \equiv u(x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_{2N}, y_{2N}, z_{2N}),$$

and integrate over the spatial coordinates x_1, \dots, z_{2N} , we get

$$\sum_j (-1)^j (u | H - W | P_j u) \cdot P_j \bar{\phi} = 0$$

or

$$\sum_j G_j P_j \bar{\phi} = 0 \quad (2.4)$$

where we have written

$$G_j = (-1)^j (u | H - W | P_j u) \quad . \quad (2.5)$$

$\bar{\phi}$, the spin part of the representative term of the complete eigenfunction Ψ of the operator $(H - W)$ is thus an eigenfunction of the operator

$$\sum_j G_j P_j \quad .$$

We shall therefore call $\bar{\phi}$ the spin eigenfunction of the system. The operator

$$\sum_j G_j P_j$$

involves the matrix elements G_j , which are pure numbers, and the permutation operators P_j , permuting the coordinates of the electrons in the spin eigenfunction.

So far our discussion has been quite general. We shall now introduce certain simplifying assumptions.

These are, first, that permutations involving the exchange of more than a single pair

of electrons do not give any contribution; and, second, that only those single exchanges which involve the exchange of neighboring pairs of electrons give a contribution. (The last assumption may be relaxed in special cases to include interactions with a finite number of non-nearest neighbors.)

These assumptions give for G_j :

$$\left. \begin{aligned} G_j &= Q - W, \text{ if } P_j \text{ is the identical permutation, } Q \text{ representing} \\ &\quad \text{the ordinary Coulomb energy term;} \\ G_j &= -\alpha, \text{ if } P_j \text{ interchanges a pair of nearest neighbors;} \\ G_j &= 0, \text{ in all other cases.} \end{aligned} \right\} \quad (2.6)$$

Although we have spoken of interchanging neighboring electrons as if they were distinguishable as well as localized, this is only because we are confining ourselves to dealing with the representative term of the entire wave function. When we exchange electrons in the spin part of that term, if the spatial part describes the electrons as localized at definite sites or in given orbitals, this exchange effectively interchanges the spins of the electrons which are localized at these positions. Hence we shall sometimes speak of interchanging the orbitals, or the spins.

We may also note here that the permutation operator which acts on a spin function to interchange the spins of the electrons l and m may be represented by the operator

$$P_{l,m} = -\frac{1}{2}(1 + \underline{\sigma}^{(l)} \cdot \underline{\sigma}^{(m)})$$

where $\underline{\sigma}^{(l)}$ and $\underline{\sigma}^{(m)}$ are the Pauli spin operators for electrons l and m , so that Eq. 2.4, together with the assumptions of (2.6) made for the form of G_j , gives

$$\left\{ Q - W + \frac{\alpha}{2} \sum_{\substack{l, m \\ \text{nearest} \\ \text{neighbors}}} (1 + \underline{\sigma}^{(l)} \cdot \underline{\sigma}^{(m)}) \right\} \cdot \Phi = 0 \quad (2.7)$$

which leads directly to the form of the operator H_1 given in Eq. 1.1.

We have now formulated our eigenvalue problem for the spin eigenfunction Φ , in terms of the operator

$$\sum_j G_j P_j \quad ;$$

so that Φ is determined by Eq. 2.4 together with the relations of (2.6) for G_j . This problem can be solved formally by expanding Φ in terms of a set of independent spin functions ϕ_i ,

$$\Phi = \sum_i a_i \phi_i \quad (2.8)$$

and then, by the procedure which we shall outline, obtaining a set of linear equations for the coefficients a_i .

The eigenfunction Φ is a function describing a many-particle spin state. We must represent such many-spin wave functions in terms of the single electron spin functions.

There are 2^{2N} independent spin functions in a set which is sufficient to describe any possible spin state of $2N$ electrons. However, it is possible to use a smaller subset of spin functions which are eigenfunctions corresponding to a common eigenvalue of any operator which commutes with

$$\sum_j G_j P_j \quad .$$

\underline{S}^2 and \underline{S}_z , the operators for total spin, and for components of spin in the z direction respectively, are such operators.

There are two types of spin eigenfunctions which we shall discuss. The first type, which we shall call the spin product functions, are eigenfunctions of \underline{S}_z only. The second type are the valence bond functions and are eigenfunctions of \underline{S}^2 as well as \underline{S}_z , and therefore have the advantage of effecting a greater reduction in the degree of the secular equation when they are used than do the spin product functions.

There are 2^{2N} independent spin product functions, which can be formed by taking the product over the $2N$ electrons of spin functions α or β for each electron. For example, one of these spin product functions is

$$\alpha(1) \beta(2) \beta(3) \alpha(4) \beta(5) \alpha(6) \alpha(7) \dots \alpha(2N) \quad .$$

α and β are the well known spin eigenfunctions corresponding to the eigenvalues $\frac{1}{2}h/2\pi$, and $-\frac{1}{2}h/2\pi$, respectively, for the z component of spin angular momentum of a single electron.

All of the spin product functions are orthogonal to each other in spin space. They are also eigenfunctions of the operator \underline{S}_z corresponding to the eigenvalue $(n_\alpha - n_\beta)h/4\pi$, where n_α is the number of α 's, and n_β is the number of β 's. The subsets of the set of spin product functions corresponding to the different values of \underline{S}_z fulfill the condition of being noncombining. The subset corresponding to $\underline{S}_z = 0$, in which equal numbers of α 's and β 's occur in the product, is of particular interest. There are $(2N)!/(N!)^2$ independent functions in this subset.

The valence bond functions may often be used to advantage because they are eigenfunctions of \underline{S}^2 as well as of \underline{S}_z . They may be defined in several equivalent ways.

First, they may be constructed out of the spin product functions just defined. A bond function having a bond between two specified orbitals will be formed of all spin product functions which assign opposite spins to the electrons in these two orbitals. The sign is plus or minus, depending on the direction of the bond with respect to the spin. Thus, a bond function with bonds between the M electron pairs a - b , c - d , e - f , \dots , but not between any of the remaining $2(N - M)$ electrons, is

$$\phi_{a-b, c-d, e-f, \dots} = \sum_j \delta_{ab} \cdot \delta_{cd} \cdot \delta_{ef} \cdots \phi_j \quad (2.9)$$

where the sum is over all spin product eigenfunctions ϕ_j with a particular value $(n_\alpha - n_\beta)h/4\pi$ for S_z , and δ_{ab} , etc., is +1 for a ϕ_j which contains $\alpha(a) \beta(b)$ as factor; it is -1 when $\beta(a) \alpha(b)$ occurs in ϕ_j , and it is zero if either $\alpha(a) \alpha(b)$ or $\beta(a) \beta(b)$ occurs.

If we perform the summation we get

$$\phi_{a-b, c-d, e-f, \dots} = [\alpha(a)\beta(b) - \beta(a)\alpha(b)] [\alpha(c)\beta(d) - \beta(d)\alpha(c)] \cdots \times \sum \alpha(1)\beta(m) \cdots \quad (2.10)$$

where the products are taken over the M pairs, and the sum is taken over the $[2(N-M)]! / (n_\alpha - M)! (n_\beta - M)!$ different ways of distributing $(n_\alpha - M)$ α 's and the $(n_\beta - M)$ β 's over the $2(N-M)$ electrons. It can be shown that bond functions containing the same number of bonds are eigenfunctions of \underline{S}^2 corresponding to the same eigenvalue. For the singlet state with $S^2 = S_z = 0$, all electrons are paired, there being N bonds.

Not all the bond functions which one can construct by drawing bonds between pairs of electrons are independent. Rumer (12) has given a rule for constructing a complete linearly independent set of bond functions. In the case of the singlet state, this rule is that one represents the electrons by points distributed around a circle. Bonds are represented by lines connecting pairs of points. If one draws all possible arrangements for the N bonds, such that each lies wholly within the circle and no two bonds intersect, one will obtain $(2N)!/N!(N+1)!$ diagrams, representing a complete, linearly independent set of singlet-state bond eigenfunctions. These are the canonical structures. The reduction of the order of the secular matrix by use of the bond functions is a factor of N+1 greater than one gets by using spin product functions. Thus, for example, with 10 electrons, the secular equations are of order 252 when using spin product functions, and of order 42 when using bond functions.

Any other bond function with the same number of bonds may be expanded in terms of the noncrossing set. If there is only one pair of crossed bonds, the uncrossing may be done by means of the rule, expressed symbolically,

$$\begin{array}{c} a & & b \\ & \searrow & / \\ & c & & d \end{array} = \begin{array}{c} a & \longrightarrow & b \\ & & \\ c & \longleftarrow & d \end{array} - \begin{array}{c} a \\ \downarrow \\ c \end{array} \begin{array}{c} b \\ \uparrow \\ d \end{array} \quad (2.11)$$

This follows immediately from the identity

$$[\alpha(d)\beta(a) - \beta(d)\alpha(a)] [\alpha(c)\beta(b) - \beta(c)\alpha(b)] = [\alpha(a)\beta(b) - \beta(a)\alpha(b)] [\alpha(c)\beta(d) - \beta(c)\alpha(d)] - [\alpha(a)\beta(c) - \beta(a)\alpha(c)] [\alpha(d)\beta(b) - \beta(d)\alpha(b)]$$

Cases involving multiply crossed bonds may be handled by successive uncrossings, using the above rule, or by other equivalent methods (4).

Having made this digression in order to discuss some properties of the kinds of functions ϕ in terms of which we may expand the eigenfunction $\bar{\Phi}$, let us now return to the problem of determining the coefficients a_i in that expansion, which we rewrite

$$\bar{\Phi} = \sum_{i=1}^{p_N} a_i \phi_i \quad (2.12)$$

where p_N is the number of functions ϕ in the subset in the terms of which we expand.

$\bar{\Phi}$ must satisfy Eq. 2.4, which becomes

$$\sum_i \sum_j a_i G_j P_j \phi_i = 0 \quad (2.13)$$

If we multiply this by ϕ_l and perform the summation over spin coordinates, we are led to the following set of linear equations for the a_i 's:

$$\sum_i \sum_j \sum_{\sigma} a_i G_j \phi_l P_j \phi_i = \sum_{i=1}^{p_N} a_i [H - W]_{li} = 0$$

$$l = 1, 2, \dots, p_N \quad (2.14)$$

where

$$[H - W]_{li} = \sum_j G_j (\phi_l P_j \phi_i) \quad (2.15)$$

is the matrix element of the secular equation.

If the functions ϕ_i are the set of spin product functions, because of their orthogonality property these matrix elements are zero unless ϕ_l and ϕ_i are alike except for the exchange of not more than a single pair of neighboring spins. In this case we are led to a secular equation which is nearly diagonal.

If however the functions ϕ_i are the set of bond functions, they no longer are orthogonal, and the secular equation contains a large number of nondiagonal terms. The actual calculation of the matrix elements is expedited by sets of rules which have been developed for the purpose (13, 14).

In this latter case, where we are dealing with nonorthogonal functions ϕ_i , it is advantageous to derive the equations for the a_i in a slightly different form, due to Wheland (15), for which the secular equation has fewer nondiagonal elements. Our derivation differs somewhat from Wheland's in order that we may be able to demonstrate the exact relationship between these equations and the ones we have already derived, which are due originally to Slater (16).

For this purpose, we introduce a set of functions ϕ' which is determined by the condition that it form a biorthogonal set of functions together with the functions ϕ . That is,

$$(\phi'_l | \phi_k) = 0, \quad l \neq k; \quad (\phi'_l | \phi_l) = 1$$

$$l, k = 1, 2, \dots, p_N \quad . \quad (2.16)$$

(If the functions ϕ are orthogonal, the set ϕ' is identical to the set ϕ .)

If we then multiply Eq. 2.13 by the function ϕ'_1 , and again sum over spin coordinates, we get

$$\sum_i \sum_j a_i G_j (\phi'_1 | P_j \phi_i) = 0 \quad . \quad (2.17)$$

If the effect of the permutation operator P_j can be expressed as

$$P_j \phi_i = \sum_k b_{ik}^{(j)} \phi_k \quad (2.18)$$

then, using Eq. 2.16, Eq. 2.17 becomes

$$\sum_i \sum_j a_i G_j b_{i1}^{(j)} = 0$$

or

$$\sum_{i=1}^{p_N} a_i \left(\sum_j G_j b_{i1}^{(j)} \right) = 0 \quad (2.19)$$

$$l = 1, 2, \dots, p_N$$

where

$$\sum_j G_j b_{i1}^{(j)} \quad (2.20)$$

is the il 'th matrix element of Wheland's secular equation, and is not, in general, equal to the li 'th element. For orthogonal functions ϕ_i , these equations are identical with the ones given by Slater. For nonorthogonal functions, we may easily show the equivalence as follows:

Since

$$\phi_1 = \sum_m (\phi_1 | \phi_m) \phi'_m \quad (2.21)$$

as one can verify from Eq. 2.16, then

$$\begin{aligned}
[H - W]_{li} &= \sum_j G_j (\phi_l | P_j \phi_i) \\
&= \sum_j G_j \left(\sum_m (\phi_l | \phi_m) \phi'_m | \sum_k b_{ik}^{(j)} \phi_k \right) \\
&= \sum_j G_j \sum_m (\phi_l | \phi_m) b_{im}^{(j)}
\end{aligned}$$

or

$$[H - W]_{li} = \sum_m (\phi_l | \phi_m) \left(\sum_j G_j b_{im}^{(j)} \right) \quad . \quad (2.22)$$

This is the desired relationship between Eq. 2.15 and Eq. 2.20.

In order to apply this latter "method of spin valence", we must be able to compute the coefficients $b_{ik}^{(j)}$ of Eq. 2.18. When we apply the permutation operator P_j to a bond function ϕ_i , if this operator exchanges the two electrons sharing a bond, the result is just the negative of the original bond function. If the exchange is not of this type, the resulting bond function will be a "crossed" structure, which when resolved into sets of the canonical structures by means of the rules for "uncrossing" bonds we have discussed, determines the coefficients $b_{ik}^{(j)}$.

III. The Infinite Linear Antiferromagnetic Chain – Present Status

If one is interested in obtaining a solution to a three-dimensional problem, one usually starts (and often finishes) by consideration of a corresponding one-dimensional problem. Accordingly, we shall begin by considering the solutions to the problem of the linear chain.

Bethe (5) has given a general solution to the problem of the linear ferromagnetic or antiferromagnetic chain. He describes the spin state of the system in terms of the spin product eigenfunctions, specifying these functions by means of the numbers n_1, n_2, \dots, n_r , which denote the location of the α 's. Thus the correct eigenfunction of the system is given by

$$\Phi = \sum a_{n_1 n_2 \dots n_r} \beta(1) \dots \alpha(n_1) \beta(n_1 + 1) \dots \alpha(n_2) \dots \alpha(n_r) \dots \beta(2N) \quad . \quad (3.1)$$

He obtains a set of linear partial difference equations of the first order for the coefficients $a_{n_1 \dots n_r}$, which have the formal solution

$$a_{n_1 \dots n_r} = \sum_P P \exp \left(\sum_{i=1}^r k_i n_i + \frac{1}{2} \sum_{i < j} \phi_{i,j} \right) \quad (3.2)$$

where the sum $\sum_{\mathbb{P}}$ is taken over the $r!$ permutations of the number 1 through r ; and the permutation operator \mathbb{P} is considered operative on the indices i of the coefficients k_i and on the indices i, j of the coefficients ϕ_{ij} , (but not on the indices i of the numbers n_i); and where the coefficients k_i and ϕ_{ij} may be determined from a set of equations arising from the periodic boundary conditions. There are a number of different sets of these coefficients corresponding to the different eigenvalues of the problem. These eigenvalues are given by $\epsilon = \sum_i (1 - \cos k_i)$, with the energy W related by $W = Q - NaZ + 2 \epsilon a$.

Bethe shows that the highest ferromagnetic ($a > 0$), and lowest antiferromagnetic ($a < 0$) state for $2N$ spins is the singlet state, with the k_i 's determined by the relations

$$2Nk_i = 2\pi(2i-1) + \sum_{j=1}^N \phi_{ij}, \quad i = 1, 2, \dots, N, \quad (3.3)$$

with

$$\cot \phi_{ij}/2 = \frac{1}{2} \cot k_i/2 - \frac{1}{2} \cot k_j/2, \quad -\pi \leq \phi_{ij} \leq \pi.$$

From solutions of these equations, one may in principle obtain complete information as to the eigenfunctions and eigenvalues for any value of N . In practice, it is difficult to carry out the algebra involved.

In the limit for N very large, Bethe (5) and Hulthén (7), have given asymptotic solutions to the above equations. k is given in terms of a distribution function $A(k)$:

$$A(k) = \frac{N}{4 \sin^2 \frac{k}{2} \cosh \left(\frac{\pi}{2} \cot \frac{k}{2} \right)} \quad (3.4)$$

which represents the number of values of k_i within the interval dk . The eigenvalue is given by

$$\epsilon = \sum (1 - \cos k) + \int 2 \sin^2 \frac{k}{2} A(k) dk = 2N \log 2. \quad (3.5)$$

The energy of the lowest state is thus

$$\begin{aligned} W - Q &= -2Na + 2Na(2 \log 2) \\ &= 2N|a| - 2N|a|(1.3863) \\ &= -0.3863(2N)|a| \end{aligned} \quad (3.6)$$

corresponding to a fractional bond order of 0.590.

The problem is still incomplete unless one has the eigenfunction. Although a formal solution would seem to be given by Bethe's method, it is difficult to interpret the result. Accordingly, we have considered other approaches to the problem of the linear

antiferromagnetic chain, which provide fuller information in this respect.

Slater-Hulthén Approximation for Infinite Chain

One such approach is the approximate treatment of the infinite chain developed by Slater (6) and elaborated on by Hulthén (7), involving the use of spin product wave functions.

Slater's approximation started from the assumption that the spin arrangement for the lowest level of the antiferromagnetic chain differed very little from the completely alternating arrangement $\alpha\beta\alpha\beta\alpha\beta\dots$. A given spin product function was characterized by a single parameter p giving the number of exchanges of neighboring spins which would produce this state, starting from the alternating state. All spin product functions with the same value of this parameter were assumed to enter with the same coefficient in the correct wave function. In computing matrix elements of a given spin product function with other functions, only those spin functions were considered in which the sites where neighboring spins were interchanges were far apart from each other. The others were discarded.

With these assumptions, Slater obtained a second order linear difference equation for the amplitude of the p 'th spin product function. He found an approximate solution of this difference equation by the WKB method, obtaining a value of $-0.29 |a| (2N)$ for the energy of the lowest state, and a value of $+0.69 |a| (2N)$ for the energy of the highest state. The energy value for the highest state would not be expected to be good, since the approximation is not valid in this region. We may compare the value $-0.29 |a| (2N)$ for the lowest energy state though with the value $-0.3863 |a| (2N)$ from Bethe's solution.

The spin state having the largest amplitude in Slater's solution was one for which p was about 6 percent of the total number of spins.

Hulthén, in his "first approximation", assumes that all spin product functions having the same number M of pairs of neighbors with opposite components of spin in the z direction enter with the same amplitude in the correct eigenfunction. The value of M ranges from $2N$, corresponding to the completely alternating spin product function, to the value two. The parameter M thus is designed to cover a larger range than Slater's parameter p with which it is approximately equivalent, for small p .

Having made this assumption, that the parameter M is sufficient to describe the eigenfunction, Hulthén computes the matrix elements of the spin function formed by summing all spin product functions having the value M , with the corresponding functions characterized by the values M , $M-2$, and $M+2$. He considers the combinatorial problem in a more detailed and exact fashion than is given in Slater's treatment. In this way he is again led to a linear second order difference equation for the coefficients in the expansion of the eigenfunction. He also solves the equation by the WKB method. He obtains a lowest energy eigenvalue of $-0.31(2N) |a|$, and a value for the highest energy of $W = Q + 2N |a|$. That the highest energy level is in agreement with the correct value in this case comes from the facts that the combinatorial coefficients entering into the

difference equation were exact for all values of M , and that the approximation that all states with a given M have equal weight becomes strictly true when M assumes its minimum value. The approximate eigenfunction that Hulthén obtains is given by

$$\Phi = \sum_{M=2}^{2N} p(M) \left[\frac{1}{W(M)} \sum_{M \text{ constant}}^{W(M)} \phi(M) \right] \quad (3.7)$$

where the sum within brackets is taken over all the $W(M)$ spin product functions consistent with a given M . The coefficient $p(M)$ may be approximated in the neighborhood of its maximum by the Gaussian function

$$p(M) \sim e^{-aN \frac{(M-M_0)^2}{M_0^2}} \quad (3.8)$$

where a is a numerical constant of the order unity, and M_0 is equal to $0.775(2N)$. The maximum value of the function $p(M) \sqrt{W(M)}$ occurs at the value $M = 0.658(2N)$. Hulthén also carries out a "second approximation" in which he introduces a second parameter describing how many "wrong" second neighbors a given spin has. With this second approximation, Hulthén obtains the energy value for the lowest state $-0.347(2N)|a|$, which is somewhat closer to Bethe's value.

Valence Bond Approximations

Because of the advantages inherent in describing a system of spins in terms of the valence bond spin eigenfunctions, it was thought desirable to develop a method of obtaining an approximate eigenfunction for the infinite chain using these types of functions in the expansion. The nature of the approximation tried was one used by Pauling and Wheland (17) in their solution of the secular equation of the naphthalene molecule by the valence bond method. Without approximations, the equations determining the lowest state of naphthalene are of fourteenth order. This order is reduced to four, by the assumption that all canonical structures with the same "degree of excitation" occur with the same coefficient. The degree of excitation is defined as follows.

In a given canonical structure, a bond which is drawn between orbitals that are adjacent in the molecule is called "effective"; a bond between orbitals that are not actually adjacent in the molecule is "ineffective". The degree of excitation, is just the number of "ineffective" bonds in the canonical structure. Or in other words, structures with the same degree of excitation have the same number of double bonds. Since structures with a given number of double bonds have the same energy in a model neglecting resonance, one might expect that their amplitudes when in quantum mechanical "resonance" would not be too different.

The full set of fourteen equations was later solved by Sherman (18), checking the fact that this type of approximation gave good results.

By using this same approximation for the infinite linear chain, namely, making the assumption that all canonical structures having the same degree of excitation enter the eigenfunction of the problem with the same coefficient, we were able to find a solution using a procedure very similar to that of Slater and Hulthén. We had expected that since the valence bond functions are correct singlet spin functions, we might obtain a better approximation, for the same labor, than had been obtained when starting from the spin product functions. We did in fact obtain a value for the energy of the lowest state which was closer to the "exact" value than was obtained by Hulthén, even in his "second approximation".

IV. Approximate Solution of Infinite Linear Chain by Valence Bond Method

We assume a periodic chain with $2N$ electrons per period. The canonical set of bond functions for this case is found by drawing a plane $2N$ -gon, and connecting the $2N$ corners together in pairs such that none of the connecting lines or bonds intersect. The degree of excitation for any given canonical structure is exactly the number of bonds which connect other than nearest neighbors.

We shall assume that the correct spin eigenfunction can be obtained to a good degree of approximation by an expansion in terms of the complete set of canonical singlet bond functions in which all canonical structures having the same degree of excitation enter with the same coefficient. Symbolically, the spin eigenfunction Φ is assumed to be of the form

$$\Phi = \sum_{i=0}^{N-2} a_i \left(\sum_{i \text{ const.}}^{k_i} \phi_i \right) \quad (4.1)$$

where the sum within brackets is over all the k_i functions ϕ_i having the degree of excitation i . The number k_i , of canonical structures of excitation i has been given by Wheland (19).

In order to apply the formalism developed in section II to the determination of the best spin eigenfunction in our problem, we must find out what happens to the functions ϕ_i , which we shall take to represent bond functions of excitation i , when we apply the permutation operators, exchanging nearest neighbors.

The answer is that we either get back the original function with a negative sign, or we get a "crossed" bond function, which when "uncrossed" gives the original function minus a different bond function. This other bond function may have the same degree of excitation, or it may be one degree higher, or one or two degrees lower in excitation. Figure 1 shows the way that these different cases can arise.

In using Eq. 2.13 we must form the sums $\sum_j G_j P_j \phi_i$, where P_j consists of the identical permutation and the $2N$ permutations of neighboring pairs, and where $G_j = Q - W$ for the identical permutation, and $-a$ for the others.

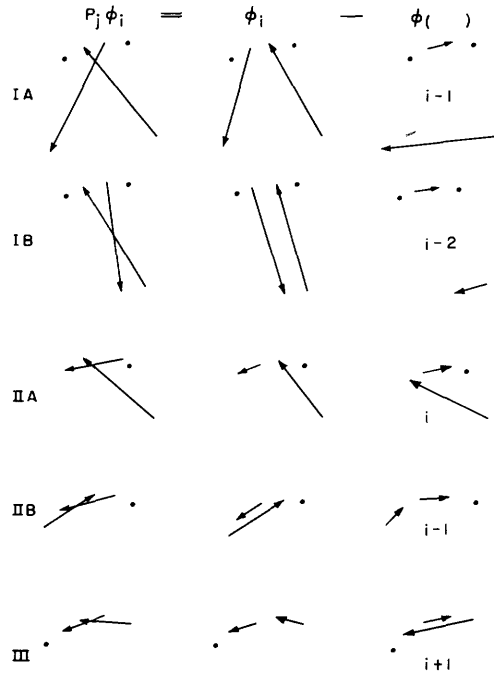


Figure 1

For the $2N$ possible exchanges of neighboring pairs, for a bond function ϕ_i of excitation i :

$N-i$ times, there will be a bond between the pair exchanged, giving a contribution $(N-i)\alpha\phi_i$ to the sum $\sum_j G_j P_j \phi_i$.

An average of $\Pi_A^i \cdot 2N$ times, case IIA will occur, contributing $-\Pi_A^i 2N\alpha(\phi_i - \phi_i'')$ where the ϕ_i'' 's are in the set of which ϕ_i is a representative, though not the original function ϕ_i .

Case IA or IIB will occur an average of $(I_A^i + \Pi_B^i) \cdot 2N$ times contributing $-(I_A^i + \Pi_B^i) \cdot 2N\alpha[\phi_i - \phi_{i-1}]$ where ϕ_{i-1} is of the set having excitation $i-1$.

Case III will occur an average of $\text{III}^i \cdot 2N$ times, contributing $-\text{III}^i \cdot 2N\alpha[\phi_i - \phi_{i+1}]$.

Case IB will occur an average of $I_B^i \cdot 2N$ times, contributing $-I_B^i \cdot 2N\alpha[\phi_i - \phi_{i-2}]$.

Thus, altogether,

$$\begin{aligned} \sum_j G_j P_j \phi_i &= (Q - W)\phi_i + [(N - i)\alpha - (\Pi_A^i + I_A^i + \Pi_B^i + \text{III}^i + I_B^i)2N\alpha]\phi_i \\ &+ \Pi_A^i \cdot 2N\alpha\phi_i'' + \text{III} \cdot 2N\alpha\phi_{i+1} + (I_A^i + \Pi_B^i) \cdot 2N\alpha\phi_{i-1} \\ &+ I_B^i \cdot 2N\alpha\phi_{i-2} \end{aligned} \quad (4.2)$$

Or, since

$$(I_A^i + I_B^i + \Pi_A^i + \Pi_B^i + \text{III}^i) \cdot 2N = N + i,$$

we have, setting $\phi_i'' = \phi_i$

$$\begin{aligned} \sum_j G_j P_j \phi_i &= \text{III}^i \cdot 2N\alpha\phi_{i+1} + \left\{ Q - W - 2i\alpha + \Pi_A^i \cdot 2N\alpha \right\} \phi_i \\ &+ (I_A^i + \Pi_B^i) 2N\alpha\phi_{i-1} + I_B^i \cdot 2N\alpha\phi_{i-2} \end{aligned} \quad (4.3)$$

Defining

$$\begin{aligned} C_{+1}^i &= \text{III}^i \cdot 2N\alpha; & C_0^i &= \left\{ Q - W - 2i\alpha + \Pi_A^i \cdot 2N\alpha \right\} \\ C_{-1}^i &= (I_A^i + \Pi_B^i) \cdot 2N\alpha; & C_{-2}^i &= I_B^i \cdot 2N\alpha \end{aligned} \quad (4.4)$$

we have

$$\sum_j G_j P_j \phi_i = C_{-2}^i \phi_{i-2} + C_{-1}^i \phi_{i-1} + C_0^i \phi_i + C_1^i \phi_{i+1} \quad . \quad (4.5)$$

To determine the coefficients a_i of Eq. 4.1, we use the equations in the form

$$\sum_{i=0}^{N-2} \sum_{i=\text{const}}^{k_i} a_i \sum_j G_j P_j \phi_i = 0 \quad ,$$

or

$$\sum_{i=0}^{N-2} \sum_j (k_i a_i) G_j P_j \phi_i = 0 \quad . \quad (4.6)$$

Since

$$\sum_j G_j P_j \phi_{i-1} = C_{-2}^{i-1} \phi_{i-3} + C_{-1}^{i-1} \phi_{i-2} + C_0^{i-1} \phi_{i-1} + C_1^{i-1} \phi_i$$

$$\sum_j G_j P_j \phi_i = C_{-2}^i \phi_{i-2} + C_{-1}^i \phi_{i-1} + C_0^i \phi_i + C_1^i \phi_{i+1}$$

$$\sum_j G_j P_j \phi_{i+1} = C_{-2}^{i+1} \phi_{i-1} + C_{-1}^{i+1} \phi_i + C_0^{i+1} \phi_{i+1} + C_1^{i+1} \phi_{i+2}$$

$$\sum_j G_j P_j \phi_{i+2} = C_{-2}^{i+2} \phi_i + C_{-1}^{i+2} \phi_{i+1} + C_0^{i+2} \phi_{i+2} + C_1^{i+2} \phi_{i+3}$$

we are led to the set of difference equations,

$$C_1^{i-1} p_{i-1} + C_0^i p_i + C_{-1}^{i+1} p_{i+1} + C_{-2}^{i+2} p_{i+2} = 0$$

$$i = 1, 2, \dots, N-4 \quad (4.7)$$

$$\left. \begin{aligned} C_0^0 p_0 + C_{-1}^1 p_1 + C_{-2}^2 p_2 &= 0 \\ C_1^{N-4} p_{N-4} + C_0^{N-3} p_{N-3} + C_{-1}^{N-2} p_{N-2} &= 0 \\ C_1^{N-3} p_{N-3} + C_0^{N-2} p_{N-2} &= 0 \end{aligned} \right\} \quad (4.8)$$

where $p_i = k_i a_i$.

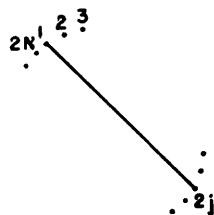
If we can find the coefficients C^i and then solve the difference equations (4.7) subject to the boundary conditions (4.8) we shall have obtained a solution to our problem.

Thus the first step is to evaluate the coefficients C^i . These involve the calculation of the probabilities I_A^i , I_B^i , II_A^i , II_B^i , and III^i . We can compute these by means of a method which is very similar to that used by Wheland (19) in computing the value of k_i . Therefore, we will first repeat verbatim Wheland's computation of the k_i 's.

"The orbitals are arranged in the circle in the same order in which they come in the chain. The polynomial C_N , associated with the chain of $2N$ atoms, is given by the recursion formula

$$C_N = C_{N-1} + z \sum C_{j-1} C_{N-j} \quad (4.9)$$

which is derived in the following manner:



We first number the orbitals in order, from 1 at one end of the chain to $2N$ at the other end. Now we consider the canonical structures which contain a bond between orbital 1 and some other definite orbital, say $2j$. This bond divides the molecule into three parts: (1) the orbitals 1 and $2j$ involved in the bond; (2) the orbitals $2, 3, \dots, 2j-1$ lying on one side of the bond; and (3) the orbitals $2j+1, 2j+2, \dots, 2N$ lying on the other side of the bond. (In the special cases in which $j = 1$, or N , the second, or the third part will be missing. We then simply treat the missing part formally as a chain of 0 orbitals, and associate with it the polynomial $C_0 = 1$.) If these three parts are considered separately, they can be represented by the polynomials J_1 , J_2 , and J_3 , respectively. Then $J_1 = 1$ if $j = 1$ (i. e., if the bond between orbitals 1 and $2j$ is effective) and $J_1 = z$ otherwise (i. e., if the bond is ineffective). Similarly $J_2 = C_{j-1}$, and $J_3 = C_{N-j}$. If the three parts are now considered as a whole, the corresponding polynomial, which represents the totality of structures with a bond between orbitals 1 and $2j$, is the product $J_1 J_2 J_3$. This follows from the fact that there can be no bonds between any two of the three parts, since such bonds would necessarily lead to noncanonical structures. The law for combining such noninteracting systems is then formally identical with that for multiplying polynomials. We now consider the canonical structures in which the orbital 1 is bonded with some further definite orbital, say $2k$. As in the previous case, the associated polynomial is equal to $K_1 K_2 K_3$, with $K_1 = 1$ or z , $K_2 = C_{k-1}$, and $K_3 = C_{N-k}$. Obviously none of the structures of this second group will occur also in the first group, and consequently the totality of structures in which the orbital 1 is bonded to either $2j$ or $2k$ will be represented by the sum $J_1 J_2 J_3 + K_1 K_2 K_3$. This same procedure can now be repeated until all of the possible canonical structures have been taken into

consideration. The resulting polynomial is then

$$C_N = \sum J_1 J_2 J_3$$

which is just the recursion formula given originally."

One can easily find the general term of this series using this recursion formula. The expression for the general term, given by Wheland is

$$C_N = \frac{1}{N} \sum_{j=0}^{N-1} \binom{N}{j} \binom{N}{j+1} z^j \quad . \quad (4.10)$$

The derivation of the expression for the polynomial, R_N , corresponding to the closed ring (or cyclic chain) with $2N$ members is exactly the same, except that $J_1 = 1$ for $j = 1$, or N , and $J_1 = z$ otherwise. This gives

$$R_N = \sum J_1 J_2 J_3 = z \sum_{j=2}^{N-1} C_{j-1} C_{N-j} + 2C_{N-1} \quad , \quad (4.11)$$

or

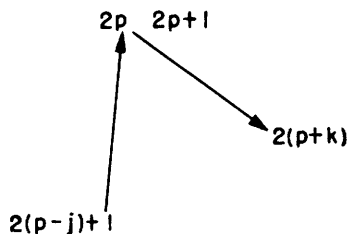
$$R_N = C_N + (1 - z)C_{N-1} \quad . \quad (4.12)$$

The general term of this may also be given explicitly

$$R_N = \frac{2}{N-1} \sum_{j=0}^{N-2} \binom{N}{j} \binom{N-1}{j+1} z^j \quad . \quad (4.13)$$

We now proceed to use these results, and this method, to calculate the probabilities $I \rightarrow III$ that we are interested in.

Let us consider those cases in which we interchange electrons $2p$ and $2p+1$, and in which there is no bond connecting $2p$ and $2p+1$. There is a bond from electron $2(p-j) + 1$ to $2p$, and from $2p+1$ to $2(p+k)$.



These bonds divide the ring into 5 separate chains.

Chain	Polynomial	I	II	III
$2p + 1$ and $2(p + k)$	$J_1 =$	z	1 or z	1
$2(p - j) + 1$ and $2p$	$J_2 =$	z	z or 1	1
$2p + 2, 2p + 3, \dots, 2(p + k) - 1$	$J_3 = C_{k-1}$			
$2(p - j) + 2, 2(p - j) + 3, \dots, \dots, 2p - 1$	$J_4 = C_{j-1}$			
$2(p + k) + 1, 2(p + k) + 2, \dots, \dots, 2N, 1, 2, \dots, 2(p - j)$	$J_5 = C_{N-k-j}$			

Thus to calculate the probabilities we are interested in, we must compute sums of the form $\sum_j \sum_k J_1 J_2 J_3 J_4 J_5$ where the conditions on each sum depend on the case with which we deal. The different cases may be characterized as follows:

I_A	I_B	II_A	II_B	III
$k \neq 1$	$k \neq 1$	$k = 1, j \neq 1$	$k = 1, j \neq 1$	$k = 1$
$j \neq 1$	$j \neq 1$	or $k \neq 1, j = 1$	or $k \neq 1, j = 1$	$j = 1$
$j + k \neq N$	$j + k = N$	$j + k \neq N$	$j + k = N$	

Therefore the sums we compute for each case are:

$$\begin{aligned}
 I_A: & \sum_{j=2}^{N-k-1} \sum_{k=2}^{N-3} J_1 J_2 J_3 J_4 J_5 \\
 I_B: & \sum_{\substack{k=2 \\ (j=N-k)}}^{N-2} J_1 J_2 J_3 J_4 J_5 \\
 II_A: & \sum_{\substack{j=2 \\ (k=1)}}^{N-2} J_1 J_2 J_3 J_4 J_5 + \sum_{\substack{k=2 \\ (j=1)}}^{N-2} J_1 J_2 J_3 J_4 J_5 \\
 II_B: & \sum_{\substack{j=1 \\ (k=N-1)}}^{(k=N-1)} J_1 J_2 J_3 J_4 J_5 + \sum_{\substack{j=N-1 \\ (k=1)}}^{(k=1)} J_1 J_2 J_3 J_4 J_5 \\
 III: & \sum_{\substack{j=1 \\ (k=1)}}^{(j=1)} J_1 J_2 J_3 J_4 J_5
 \end{aligned}$$

These sums are

$$\begin{aligned}
I_A: \quad & z^2 \sum_{j=2}^{N-k-1} \sum_{k=2}^{N-3} C_{k-1} C_{j-1} C_{N-k-j} = z \sum_{k=2}^{N-3} (C_{N-k} - C_{N-k-1} - z C_{N-k-1}) C_{k-1} \\
& = C_N - C_{N-1} - z [C_{N-1} + C_{N-2} + (1+z)C_{N-3}] - (1+z) [C_{N-1} - C_{N-2} - z(C_{N-2} + C_{N-3})] \\
& = C_N - 2(1+z) C_{N-1} + (1+z+z^2) C_{N-2} \tag{4.14}
\end{aligned}$$

$$\begin{aligned}
I_B: \quad & z^2 \sum_{k=2}^{N-2} C_{k-1} C_{N-k-1} C_0 = z(C_{N-1} - C_{N-2}) - z^2 C_{N-2} \\
& = z C_{N-1} - z(1+z) C_{N-2} \tag{4.15}
\end{aligned}$$

$$\begin{aligned}
II_A: \quad & 2z \sum_{j=2}^{N-2} C_0 C_{j-1} C_{N-1-j} = 2 [C_{N-1} - C_{N-2} - z C_{N-2}] \\
& = 2 C_{N-1} - 2(1+z) C_{N-2} \tag{4.16}
\end{aligned}$$

$$II_B: \quad 2z C_{N-2} C_0 C_0 = 2z C_{N-2} \tag{4.17}$$

$$III: \quad C_0 C_0 C_{N-2} = C_{N-2} \tag{4.18}$$

where we have used Eq. 4.9 repeatedly, together with the fact that $C_0 = C_1 = 1$; $C_2 = 1 + z$.

If we add up all these polynomials for cases I through III, together with the polynomial C_{N-1} , representing the case where there is a bond between p and $p+1$, we should have just R_N .

$$\begin{aligned}
\sum & = C_N + C_{N-1} (-2 - 2z + z + 2 + 1) + (1 + z + z^2 - z - z^2 - 2z + 2z + 1) C_{N-2} \\
& = C_N + C_{N-1} (1 - z) = R_N \quad .
\end{aligned}$$

Next we calculate explicitly the polynomials corresponding to the different cases, using Eq. 4.10. These results are:

$$I_A: \quad \frac{3}{N-2} \sum_{j=2}^{N-3} \binom{N-2}{j-2} \binom{N-2}{j+1} z^j \tag{4.19}$$

$$I_B: \quad \frac{2}{N-2} \sum_{j=2}^{N-2} \binom{N-2}{j} \binom{N-2}{j-2} z^j \tag{4.20}$$

$$\text{II}_A: \quad \frac{4}{N-2} \sum_{j=1}^{N-3} \binom{N-2}{j-1} \binom{N-2}{j+1} z^j \quad (4.21)$$

$$\text{II}_B: \quad \frac{2}{N-2} \sum_{j=1}^{N-2} \binom{N-2}{j-1} \binom{N-2}{j} z^j \quad (4.22)$$

$$\text{III:} \quad \frac{1}{N-2} \sum_{j=0}^{N-3} \binom{N-2}{j} \binom{N-2}{j+1} z^j \quad (4.23)$$

We calculate the probabilities for the different cases by taking the coefficient of the i 'th power of z in the polynomial corresponding to a particular case, and dividing it by the total number of structures corresponding to the excitation i . (This defines what we mean, for example, by the probability II_A^i ; namely, that if we consider a given exchange of nearest neighbors, if this is applied to all the k_i structures corresponding to the excitation i , case II_A will occur exactly $\text{II}_A^i \cdot k_i$ times.)

These probabilities are:

$$\text{I}_A^i = \frac{3}{2} \frac{i(i-1)(N-i-2)}{N(N-1)(N-2)} \quad (4.24)$$

$$\text{I}_B^i = \frac{i(i-1)(i+1)}{N(N-1)(N-2)} \quad (4.25)$$

$$\text{II}_A^i = 2 \frac{i(N-i)(N-i-2)}{N(N-1)(N-2)} \quad (4.26)$$

$$\text{II}_B^i = \frac{i(i+1)(N-i)}{N(N-1)(N-2)} \quad (4.27)$$

$$\text{III}^i = \frac{1}{2} \frac{(N-i)(N-i-1)(N-i-2)}{N(N-1)(N-2)} \quad (4.28)$$

$$k_i = \frac{2N!(N-2)!}{i!(i+1)!(N-i)!(N-i-2)!} \quad (4.29)$$

We can again check, and verify that $\text{I}_A^i + \text{I}_B^i + \text{II}_A^i + \text{II}_B^i + \text{III}^i$ is equal to $(N+i)/2N$.

The coefficients C^i may be computed by reference to (4.4).

We find that:

$$C_1^i = \alpha \frac{(N-i)(N-i-1)(N-i-2)}{(N-1)(N-2)}; \quad C_1^{i-1} = \alpha \frac{(N-i+1)(N-i)(N-i-1)}{(N-1)(N-2)} \quad (4.30)$$

$$C_0^i = \left\{ Q - W - 2 \alpha + 4 \alpha \frac{(N-i)(N-i-2)}{(N-1)(N-2)} \right\} \quad (4.31)$$

$$C_{-1}^i = \alpha \frac{i [3(i-1)(N-i-2) + 2(i+1)(N-i)]}{(N-1)(N-2)} ; \quad C_{-1}^{i+1} = \alpha \frac{(i+1) [3i(N-i-3) + 2(i+2)(N-i-1)]}{(N-1)(N-2)} \quad (4.32)$$

$$C_{-2}^i = 2\alpha \frac{(i+1)i(i-1)}{(N-1)(N-2)} ; \quad C_{-2}^{i+2} = 2\alpha \frac{(i+1)(i+2)(i+3)}{(N-1)(N-2)} . \quad (4.33)$$

We are thus now ready to proceed with the solution of the difference equation (4.7).

Since N is assumed large, we can write the coefficients C in terms of the variable $x = i/N$, neglecting terms of the order $1/N$. Then

$$C_1(x) = \alpha N(1-x)^3; \quad (4.30a)$$

$$C_0(x) = Q - W + 2Na x(2x^2 - 4x + 1) \quad (4.31a)$$

$$C_{-1}(x) = \alpha N \cdot 5x^2(1-x) \quad (4.32a)$$

$$C_{-2}(x) = 2\alpha N \cdot x^3 . \quad (4.33a)$$

If we write

$$(Q - W)/Na = 2\lambda \quad (4.34)$$

Eq. 4.7 now becomes

$$(1-x)^3 p_{x-1/N} + [2\lambda + 2x(2x^2 - 4x + 1)] p_x + 5x^3(1-x) p_{x+1/N} + 2x^3 p_{x+2/N} = 0 . \quad (4.35)$$

We assume $p = e^{N \int \beta(x) dx}$. Then, again neglecting quantities of the order of $1/N$, we have

$$p_{i+1} = e^\beta p_i \quad (4.36)$$

and

$$2x^3 e^{3\beta} + 5x^2(1-x) e^{2\beta} + 2[\lambda + x(2x^2 - 4x + 1)] e^\beta + (1-x)^3 = 0 . \quad (4.37)$$

One may thus determine β from a solution of this cubic equation in e^β . So that Eq. 4.36 will correspond to a solution of the difference equations, we must also see that it satisfies the boundary conditions of Eq. 4.8.

If we examine these conditions we find that they require

$$e^\beta \rightarrow \infty \text{ as } x \rightarrow 0, \text{ and } e^\beta \rightarrow 0 \text{ as } x \rightarrow 1 . \quad (4.38)$$

The cubic equation (4.37) will have three roots for every value of x . Their general behavior is sketched for different values of x in Fig. 2.

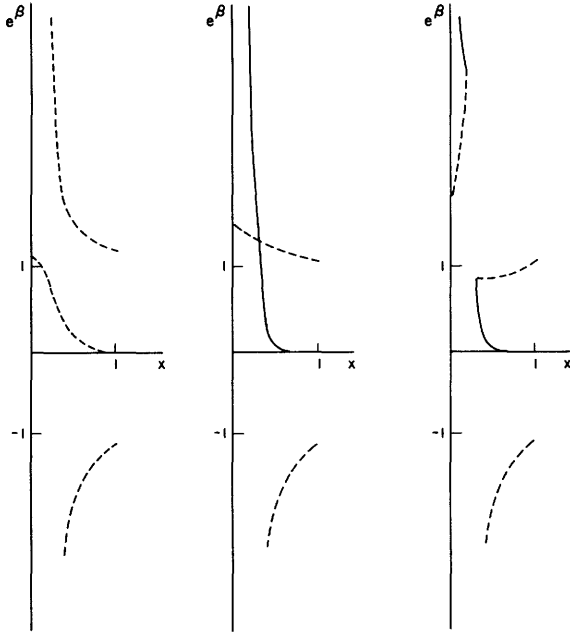


Figure 2

The behavior exhibited by the first sketch is such that no solution of Eq. 4.37 satisfies both boundary conditions. If the value of λ is such as to let two roots of the cubic become equal for at least one value of x , then it is possible to satisfy boundary conditions by switching branches at this point. This is illustrated in the second and third sketches; the second corresponds to the least value of λ for which the equation can have a double root, and hence to the lowest eigenvalue for λ . In the third sketch, imaginary roots exist for a certain region of x . Solutions to the difference equation exist in this case for a set of discrete, but closely spaced, values of λ which must be such as to allow the solutions within the regions having complex roots to be joined on to the solutions where real

roots exist, in such a way that the p 's will be real. We, therefore, compute the value of λ corresponding to the lowest eigenvalue of our problem as follows.

The condition for equal roots of Eq. 4.37 is that the discriminant be equal to zero.

The discriminantal equation can be factored to give the condition:

$$(\lambda - x) \left\{ \lambda + x \left[2x^2 - 4x + 1 + \frac{7 + \sqrt{4913}}{32} (x - 1)^2 \right] \right\} \left\{ \lambda + x \left[2x^2 - 4x + 1 + \frac{7 - \sqrt{4913}}{32} (x - 1)^2 \right] \right\} = 0. \quad (4.39)$$

This has three roots

$$\begin{aligned} \lambda_1 &= x \\ \lambda_2 &= -x \left[2x^2 - 4x + 1 + \frac{7 + \sqrt{4913}}{32} (x - 1)^2 \right] \\ \lambda_3 &= -x \left[2x^2 - 4x + 1 + \frac{7 - \sqrt{4913}}{32} (x - 1)^2 \right] \end{aligned} \quad (4.40)$$

Since the range of x is from zero to one, we see immediately that any value of λ between zero and one will give equal roots for some value of x , and in fact, since the value of 1 is the maximum value of either λ_1 , λ_2 , or λ_3 in this range, we conclude that the highest eigenvalue corresponds to $\lambda = 1$.

The lowest eigenvalue of the problem occurs when λ is equal to the minimum value assumed by either λ_1 , λ_2 , or λ_3 , for any x , which occurs for

$$x = \frac{4 - \sqrt{217 - 3\sqrt{4913}}}{6} = 0.2346 \quad (4.41)$$

at the value

$$\lambda_2 = -0.3707 \quad . \quad (4.42)$$

Since $\lambda = (Q - W)/2Na$, we obtain

$$\text{ENERGY OF HIGHEST STATE: } W = Q - 2Na$$

$$\text{ENERGY OF LOWEST STATE: } W = Q - 2Na + 1.3707(2Na) \quad . \quad (4.43)$$

Thus the value of the energy we obtain for the lowest eigenstate is closer to Bethe's exact value than Hulthén's second approximation, and quite a bit closer than Hulthén's first approximation.

Therefore, we have confirmed our expectation that the use of bond functions would enable us to obtain a good approximation.

The maximum value of $p_i = k_i a_i$ will occur when $dp/dx = 0$, or when $\beta = 0$, or when $e^\beta = 1$. If we substitute this relation in Eq. 4.37 we find that

$$(x)_{p_{\max}} = 1 + 2\lambda = 0.2586 \quad (4.44)$$

or, in other words, the degree of excitation which corresponds to the maximum value of $k_i a_i$, corresponds to an i of $0.2586N$, roughly one-quarter of the maximum possible degree of excitation.

We may approximate p in the neighborhood of this maximum value by the Gaussian function

$$p \cong \text{const} \exp \left[-N(x - x_0)^2 / 2(3x_0^2 + 1) \right] \cong \text{const} e^{-\frac{N(x-x_0)^2}{2.4}} \quad (4.45)$$

where x_0 is the fractional excitation (0.259) corresponding to p_{\max} . This is a very sharp distribution, due to the factor N in the exponent.

Knowing that we have such a sharp maximum of $k_i a_i$, we may derive the relation (4.44) between the eigenvalue, and the maximum excitation, from a result of Wheland (19), namely that

$$W = Q + \sum_i \frac{a_i k_i (N-i)}{\sum_i a_i k_i} a \quad . \quad (4.46)$$

Assuming that only the set of bond functions near $(a_i k_i)_{\max}$ is important, we get $W = Q + (N-i)a$. Since $\lambda = (Q-W)/Na$, from Eq. 4.34, we have $N-i/2N = -\lambda$; or $i/N = 1+2\lambda$, which is just relation (4.44).

We may therefore summarize the results of our valence bond approximation as follows.

We have expanded the spin eigenfunction describing the state of a linear antiferromagnetic chain in terms of the Rumer canonical set of valence bond functions, under the approximation that all bond functions having the same degree of excitation, i , have the same coefficient a_i in that expansion.

The secular equation then contains only matrix elements between the state i , and the states $i + 1$, i , $i - 1$, $i - 2$; and reduces to a set of third order linear difference equations. This set of difference equations is solved, in the limit of large N , by the WKB method. The range of values of energy within which there exist solutions satisfying the boundary conditions determines the highest and lowest energy eigenvalues. In the solution for the lowest state, appreciable amplitudes occur only for bond functions within a narrow range about values of the degree of excitation $0.2586N$. The results for the lowest energy eigenvalues, compared with those obtained by other methods are

Bethe "exact" value	$\frac{W - Q}{2Na} = 0.3863$
Slater approximation	0.29
Hulthén "first approximation"	0.316
"second approximation"	0.347
Our approximation	0.3707

V. Special Solutions

Solutions pertinent to the problem of determining the lowest spin eigenfunction for the linear antiferromagnetic chain with periodic boundary conditions for 2, 4, 6, 8, and 10 spins have been given in the literature. The 6, 8, and 10 spin solutions arise in the treatments of the molecules of benzene, cyclooctatetrane, and cyclodecapentane, respectively. The various types of structures represented in the valence bond diagrams have, in a sense, a physical reality to the chemists, who had empirically developed similar models before the development of quantum mechanics. The molecule was considered to undergo a "resonance" between the different kinds of bond structures with the resonance energy reckoned as the difference between the actual energy, and the energy which would correspond to just one of the "unexcited" bond structures. Resonance among a large number of structures would give a large resonance energy, and hence greater stability.





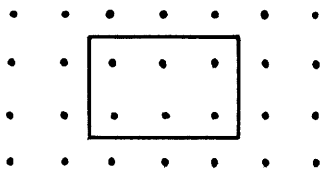
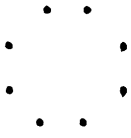
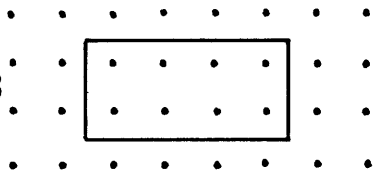
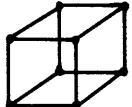
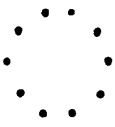

We present a tabulation of the results for energy, amplitude of excitation of each type of bond structure, and bond order for these cyclic structures.

We have also investigated other six- and eight-spin structures in other arrangements, in order to try to get some information about what happens in two and three dimensions. We have also tabulated these results for energy and bond order.

We notice that the fractional bond order is less in these cases than it is in the one-dimensional case, as could have been anticipated from the fact that it is harder for a given spin to stay paired with another in a singlet state when the number of nearest neighbors increases. Even in these cases, though, the bond order is still appreciably

	ϕ_i	$\Phi = \sum a_i \phi_i$
2		
4		$a_A = a_B$
6		$a_A = 1/2$ $a_B = 0.626/3$
8		$a_A = 1/2$ $a_B = 1.45/8$ $a_C = 0.20/4$
10		$a_A = 1/2$ $a_B = 1.71/10$ $a_C = 0.67/5$ $a_D = 0.31/10$ $a_E = 0.50/10$ $a_F = 0.03/5$

Note: A(x 2) means that there are 2 structures of type A equivalent by symmetry operations, etc.

NUMBER OF ELECTRONS	STRUCTURE	ENERGY		BOND ORDER
		Highest	Lowest	
2		$-a$	a	1.0
4		$-4a$	$2a$	0.67
6		$-6a$	$2.6055a$	0.623
6		$-7a$	$2.759a$	0.596
(Periodic boundary conditions)				
6		$-12a$	$4.606a$	0.589
8		$-8a$	$3.3022a$	0.609
(Periodic boundary conditions)				
8		$-16a$	$5.869a$	0.578
8		$-24a$	$7.28a$	0.536
10		$-10a$	$14.031a$	0.602
$2N \rightarrow \infty$		$-2Na$	$0.3863(2N)a$	0.591

larger than the value $1/3$ that it would have if the spin components relative to a fixed axis alternated in sign. This difference may, of course, get smaller if we consider larger numbers of spins in the three-dimensional case. The degree of complication of the computation increases rapidly as the number of spins increases, but it is possible that some approximate type of treatment could produce good results. We might also mention in this connection that secular equations of quite high order have been set up for the valence bond method using punched card methods (20), so that it may be feasible to solve a problem with enough spins to afford a reasonable approximation to the infinite lattice.

Conclusion

In ferromagnetism the lowest state in the semiclassical Ising model is characterized by a simple ordered structure; this remains valid in the quantum mechanical model with the modification that only one component of spin shares in the ordering.

For antiferromagnetism, even in the Ising model, the lowest state may, for certain geometrical configurations, be only partially ordered. In quantum mechanics this state is quite complicated for any geometry. The situation is more analogous to the problem of chemical binding than it is to ferromagnetism. There exists the phenomenon of "resonance" among different bond structures such as occurs in certain molecules.

Both from the neutron diffraction experiments of Shull, and from considerations based on Nernst's law, one is led to believe that some sort of ordering actually exists in antiferromagnetic materials at low temperatures. From simple quantum mechanical considerations, however, this state cannot exhibit perfect ordering in the usual sense. Our main problem was to clarify the nature of this order.

A knowledge of the state of a system comes from both the eigenfunction and from the eigenvalue for the energy. Knowledge of the value of the energy enables us to calculate the "bond order". This concept, originally used by chemists, which we have adapted for our problem, describes the short range order of a spin system. We may thus calculate the bond order for the infinite linear chain, where we have a value of the energy due to Bethe. Energy values and bond orders have also been obtained in two and three dimensions for some special models, but not for the infinite lattices.

We have obtained an approximate eigenfunction in terms of the valence bond spin functions for the problem of the infinite linear chain to supplement the formal solution given by Bethe. We do not feel that we have obtained as complete information from the eigenfunction as might be possible; in particular, because of the simplicity of Bethe's results for the energy of the linear chain, one is led to suspect that the description of the eigenfunction must also contain elements of simplicity which have thus far remained hidden.

We have not entered into any consideration of excited states. It would be necessary to do so if one wished to consider the statistical problem, or the effect of the magnetic field.

Acknowledgment

The author wishes to thank Professor Laszlo Tisza for suggesting this problem, as well as for his guidance of the work and many helpful suggestions.

References

1. C. G. Shull, J. S. Smart: *Phys. Rev.* 76, 1256 (1949).
2. G. H. Wannier: *Phys. Rev.* 78, 341(A) (1950); 79, 357 (1950).
3. H. Eyring, J. Walter, G. E. Kimball: *Quantum Chemistry*, Chap. XIII (Wiley and Sons, N. Y. 1944).
4. H. Eyring, H. Gershinowitz: *J. Chem. Phys.* 3, 224 (1933).
5. H. Bethe: *Zeits. f. Physik* 71, 205 (1931); *Handbuch der Physik* XXIV/2, 604 ff., 618 (J. Springer, Berlin 1933).
6. J. C. Slater: *Phys. Rev.* 35, 509 (1930).
7. L. Hulthen: *Arkiv. f. Mat. Astro. o. Fys.* 26A, 11 (1938).
8. P. A. M. Dirac: *Proc. Roy. Soc. A* 123, 714 (1929). Also see P. A. M. Dirac: *The Principles of Quantum Mechanics*, Chap. IX 3rd. ed. (Oxford, London, 1947); or J. H. Van Vleck: *The Theory of Electric and Magnetic Susceptibilities*, Chap. XII (Oxford, London, 1932).
9. W. G. Penney: *Proc. Roy. Soc. A* 158, 306 (1937).
10. H. A. Kremers: *Physica* 1, 182 (1934).
11. P. W. Anderson: *Phys. Rev.* 79, 350 (1950).
12. G. Rumer: *Nach. Ges. Wiss. Gott., M. P. Klasse*, 337 (1932). Also see G. Rumer, E. Teller, H. Weyl: *ibid.*, 499 (1932); H. Weyl: *ibid.*, 285 (1930); 33 (1931); W. Heitler, G. Rumer: *ibid.*, 277 (1930).
13. R. S. Bear, H. Eyring: *J. Chem. Phys.* 3, 98 (1933).
14. L. Pauling: *J. Chem. Phys.* 1, 280 (1933).
15. G. W. Wheland: *J. Chem. Phys.* 3, 230 (1935).
16. J. C. Slater: *Phys. Rev.* 38, 1109 (1931).
17. L. Pauling, G. W. Wheland: *J. Chem. Phys.* 1, 362 (1933).
18. J. Sherman: *J. Chem. Phys.* 2, 488 (1934).
19. G. H. Wheland: *J. Chem. Phys.* 3, 356 (1935).
20. M. B. Oakley, G. E. Kimball: *J. Chem. Phys.* 17, 706 (1949).

