PHYSICAL REVIEW VOLUME 125, NUMBER 1 JANUARY 1, 1962

Theory of Ferromagnetism and the Ordering of Electronic Energy Levels

ELLIOTT LIEB AND DANIEL MATTIS
Thomas J. Watson Research Center, International Business Machines Corporation, Yorktown Heights, New York

(Received May 25, 1961; revised manuscript received September 11, 1961)

Consider a system of N electrons in one dimension subject to an arbitrary symmetric potential, $V(x_1,\cdots,x_N)$, and let $E(S)$ be the lowest energy belonging to the total spin value S. We have proved the following theorem: $E(S)<E(S')$ if $S<S'$. Hence, the ground state is unmagnetized. The theorem also holds in two or three dimensions (although it is possible to have degeneracies) provided $V(x_1,y_1,z_1;\cdots;x_N,y_N,z_N)$ is separately symmetric in the x_i, y_i, and z_i. The potential need not be separable, however. Our theorem has strong implications in the theory of ferromagnetism because it is generally assumed that for certain repulsive potentials, the ground state is magnetized. If such be the case, it is a very delicate matter, for a plausible theory must not be so general as to give ferromagnetism in one dimension, nor in three dimensions with a separately symmetric potential.

INTRODUCTION

THIS paper consists mainly in the enunciation and proof of a theorem about the ordering of the energy levels of a system of interacting fermions. As such, our primary concern is with mathematics. Without striving to be pedantic, we have endeavored to construct the proof with care and rigor.

We take advantage of the symmetry properties of electron wave functions belonging to various values of total spin-angular-momentum S. In certain cases we are able to order the ground-state energies belonging to the various spins without any explicit numerical calculations. This circumvents the great difficulties of the N-body problem, such as the applicability of perturbation theory, etc.

Our theorem is not without some theoretical consequences. Notably, whenever it is applicable, there can be no ferromagnetism unless one postulates explicitly spin- or velocity-dependent forces. (The theorem does not apply to electrons in a three-dimensional lattice interacting with Coulomb or central forces; but conversely, correct theories of ferromagnetism should not predict ferromagnetism for interactions which *are* covered by the theorem.) However, as the mathematics stand quite independently of such applications, we shall defer further considerations of the physical consequences of this work to the end of the paper (Sec. IV).

In Sec. I we shall state and prove the theorem for a one-dimensional electron system. In order to pass to higher dimensions, it will be necessary to prove further theorems on one-dimensional systems which have no direct relevance to fermions. These will be discussed in Sec. II. Section III will contain the proof of our theorem for certain specialized problems in two or more dimensions. We have added an Appendix on an analogous theorem for certain one-dimensional chains of three-dimensional atoms; the proof uses a different technique from that in Sec. I, insomuch as we switch to the delta function (or lattice gas) representation and use second quantization.

As a preliminary, we should like to recall a well-known theorem on the two-fermion problem, proved many years ago.[1] Consider the general two-particle Hamiltonian[2] (in any number of dimensions)

$$H=\mathbf{p}_1{}^2+\mathbf{p}_2{}^2+V(\mathbf{r}_1,\mathbf{r}_2), \tag{1}$$

where V is any symmetric potential. That is to say, the particles may interact with each other and/or with an external potential, the only proviso being that no spin- or velocity-dependent forces are present. The boundary conditions can be anything so long as they are homogeneous. Since the total spin S is a good quantum number, the ground state is either $S=0$ or $S=1$. The theorem states that the ground state always has $S=0$, a statement borne out by the hydrogen molecule, for example.

Since V is real (hence the necessity for excluding velocity-dependent forces) the eigenfunctions of H are real. An eigenstate ${}_M{}^S\Psi$ with $S=0$ must be of the form[3]

$${}_0{}^0\Psi=\Phi_{Sy}(\mathbf{r}_1,\mathbf{r}_2)[(+-)-(-+)], \tag{2}$$

where Φ_{Sy} is a symmetric real function and where we have used an obvious notation for the spin part of the wave function. Alternatively, a state with $S=1$ (and $M=1$, for example) must be of the form

$${}_1{}^1\Psi=\Phi_A(\mathbf{r}_1,\mathbf{r}_2)[(++)], \tag{3}$$

where Φ_A is antisymmetric. In both cases the symmetry property of Φ is determined by the Pauli principle which states that ${}_M{}^S\Psi$ must be antisymmetric.

Now if the ground-state wave function were $S=1$, consider the trial function obtained by taking the absolute value of Φ_A,

$${}_0{}^0\Psi\equiv|\Phi_A|[(+-)-(-+)], \tag{4}$$

which has $S=0$ and satisfies the Pauli principle. Φ_A is the spatial part of the $S=1$ ground-state function. It is readily verified that the variational energy of ${}_0{}^0\Psi$ is the same as that of the supposed ground state, ${}_1{}^1\Psi$ (we shall return to a proof of this in Sec. I). Thus, by

---

[1] The authors believe this theorem is due to E. P. Wigner.

[2] $\hbar^2/2m=1$.

[3] We shall use the notation ${}_M{}^S\psi$ for a function with a definite S and M value; ${}^S\psi$ if it has only a definite S value; and ${}_M\psi$ if it has only a definite M value.

reductio ad absurdum, it follows that there is always an $S=0$ eigenfunction having an energy at least as low as the lowest $S=1$ function. In fact, as we shall show later, the trial function given above cannot be an eigenfunction unless V is pathologic (e.g., a repulsive core in one dimension). Thus, if $E(S)$ denotes the lowest energy belonging to a given S value, $E(0)<E(1)$ for two particles.

This paper is an extension of the two-particle theorem to an arbitrary number of particles. We are able to do it completely for one dimension, and in certain cases for higher dimensions. The general result is $E(S)\leq E(S+1)$.

I. THE ONE-DIMENSIONAL SYSTEM

We start with the general Hamiltonian[2]

$$H=-\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2}+V(x_1,\cdots,x_N), \tag{5}$$

where V is real and symmetric in the N variables $x_1,\cdots,x_N$. Otherwise, it is completely arbitrary. The boundary conditions may be any of the following:

(i) If the particles are in a "box" (i.e., $0\leq x_i\leq L$), then $\Psi=0$ if any $x_i=0$ or L;

(ii) the same as (i) except that $\partial\Psi/\partial x_i=0$ if any $x_i=0$ or L;

(iii) if $-\infty<x_i<\infty$ or $0<x_i<\infty$, then we restrict ourselves to square integrable (bound state) wave functions, assuming these exist. Periodic boundary conditions are excluded, for they require a slight modification of the theorem and a somewhat lengthier analysis.

Although H does not contain the spins explicitly, every eigenfunction belongs to a definite S value, which may take on any of the values $N/2$, $(N/2)-1$, $\cdots$, 0, or $\frac{1}{2}$. If we denote the lowest or ground-state energy belonging to a given S value by $E(S)$, then the theorem to be proved is

Theorem I. If $S>S'$, *then* $E(S)>E(S')$ *unless* V *is pathologic, in which case* $E(S)\geq E(S')$.

The term pathologic potential will be defined in the sequel.

In order to prove this theorem it is first necessary to characterize the spatial part of a wave function of space and spin. To this end, let ${}_M\Psi$ be a wave function satisfying the Pauli principle and having a definite spin azimuthal quantum number M. (That is, $S_z\,{}_M\Psi=M\,{}_M\Psi$, where $S_z=\sum_{i=1}^N S_z^i$.) Then ${}_M\Psi$ may be expanded in the complete set of spin functions having the M value in question. The coefficients of the expansion will be spatial functions. Thus

$${}_M\Psi=\textstyle\sum_j {}_M\Phi_j(x_1,\cdots,x_N)G_j^M, \tag{6}$$

where G_j^M is a spin function of which a typical one is

$$G_1^M=(\underbrace{--\cdots--}_{p}\underbrace{++\cdots++}_{N-p}),$$

with

$$p=(N/2)-M \tag{8}$$

(i.e., particles $1-p$ have spins down, the rest being up). Since ${}_M\Psi$ is a Pauli function and since the various G_j^M are obtained from each other by a permutation of the spin variables, it follows that the ${}_M\Phi_j$ are related to each other by a permutation of $x_1,\cdots,x_N$. Moreover, the following is easily verified:

(a) If any ${}_M\Phi_j\equiv0$, then all ${}_M\Phi_j\equiv0$.

(b) If Ψ is an eigenfunction of H with energy E, then so is each ${}_M\Phi_j$ separately.

(c) ${}_M\Phi_1$ (henceforth to be denoted simply by ${}_M\Phi$) is nonvanishing and is of the form

$${}_M\Phi=\Phi(x_1,\cdots,x_p|x_{p+1},\cdots,x_N),\quad p=(N/2)-M, \tag{9}$$

where the notation is meant to imply that Φ is antisymmetric in the variables $x_1,\cdots,x_p$ *and* in the variables $x_{p+1},\cdots,x_N$.

(d) Given any spatial function having the symmetry properties of (9) above, it may be used to generate a nonvanishing Pauli function such as ${}_M\Psi\equiv\sum_p(-)^p\times(p\Phi)(pG_1^M)$, where the summation is over all permutations p of the N particles. If Φ is an eigenfunction of H, then so is ${}_M\Psi$.

The next question to consider is what further condition must we impose on ${}_M\Phi$ in order that ${}_M\Psi$ have a definite S value (i.e., $S^2\,{}_M{}^S\Psi=S(S+1)\,{}_M{}^S\Psi$ and $S_z\,{}_M{}^S\Psi=M\,{}_M{}^S\Psi$). For simplicity, let us take $M\geq0$ (i.e., $p\leq N/2$), in which case the S value of Ψ in Eq. (6) could be any of $M, M+1,\cdots,N/2$, or a mixture of these. Suppose we wish $S=M$. Then, a necessary and sufficient condition is $S_+\,{}_M{}^M\Psi=0$ where $S_+=\sum_{i=1}^N S_+^i$. The operator S_+ acts on the G_j^M functions; acting on G_1^M it generates G_1^{M+1} among others. But other G_j^M's also generate G_1^{M+1} and if one sets equal to zero the coefficient of G_1^{M+1} in $S_+\,{}_M{}^M\Psi$, one finds

(e) A necessary and sufficient condition that ${}_M\Psi$ be ${}_M{}^M\Psi$ is that ${}_M{}^M\Phi$ (the coefficient of G_1^M) be of the form (9) and that the bar cannot be moved to the left. By this is meant that ${}_M{}^M\Phi$ cannot be antisymmetrized with respect to the variables $x_p,\cdots,x_N$. In other words,

$$(1-\sum_{j=p+1}^{N}P_{p,j})\,{}_M{}^M\Phi\equiv0, \tag{10}$$

where $P_{p,j}$ is the simple transposition permutation of x_p and x_j.

(f) If the bar can be moved to the left once, but not twice, then Ψ is, in general, a mixture of ${}^{M+1}{}_M\Psi$ and ${}^M{}_M\Psi$, and so forth.

(g) If $M>0$, we can always lower the M value of Ψ by $S_-\,{}_M\Psi=\text{const}\,{}_{M-1}\Psi\neq0$, and hence the bar can always be moved to the right. In other words, if a function is of the form (9), the bar can always be moved to the right if $p<N/2$, but it cannot always be moved to the left. This is a known result of the theory of the

permutation group which we have proved by recourse to the more generally known theory of angular momentum.

(h) Any function satisfying (9) and (10) may be used to generate a nonvanishing ${}_M{}^S\Psi$ as in (d) above. These remarks tell us that the higher the S value of ${}^S\Phi$, the more antisymmetric it must be. For instance, a totally antisymmetric function always belongs to $S=N/2$, but to any $M\leq N/2$.

Now since S_- and S_+ commute with H, $E(S)$ is degenerate with respect to M value; i.e., $H\,{}_M{}^S\Psi=E(S)\,{}_M{}^S\Psi$ implies there exists ${}_j{}^S\Psi$ with the same eigenvalue, where j can take all $2S+1$ values between S and $-S$. We shall therefore prove Theorem I in the following manner: Let $E(M)$ and ${}_M\Psi$ be the lowest eigenvalue and eigenfunction, respectively, of H belonging to a given M value ≥ 0. We shall show that ${}_M\Psi$ is ${}_M{}^M\Psi$ and hence that $E(M)=E(S)$. We shall further show that $E(M)<E(M+1)$ unless V is pathologic. In other words, we shall show that the lowest eigenfunction of H of the form (9) also satisfies Eq. (10).

Let R be the full domain, all $0\leq x_i\leq L$ [assuming we use boundary condition (i) or (ii) for example], and define $R_M\subset R$ by

$$R_M\colon\quad 0\leq x_1\leq\cdots\leq x_p\leq L$$

and

$$0\leq x_{p+1}\leq\cdots\leq x_N\leq L. \tag{11}$$

Consider the Schrödinger equation in R_M;

$$H\varphi=E\varphi, \tag{12}$$

with boundary conditions:

$$\varphi=0 \text{ on the boundary of } R_M. \tag{13}$$

It is well known that the ground-state function φ_0 of Eqs. (12) and (13) satisfies

$$\varphi_0\geq 0 \quad\text{in}\quad R_M. \tag{14}$$

For suppose (14) were not satisfied and consider $\varphi=|\varphi_0|$. Now $\langle\varphi|\varphi\rangle=\langle\varphi_0|\varphi_0\rangle$ and $H\varphi=E\varphi$ everywhere except where φ_0 vanishes, at which points φ is continuous but has discontinuous derivatives. Thus $H\varphi=E\varphi+\delta$ functions, the latter occurring when φ vanishes. Hence

$$\int_{R_M}\varphi(H\varphi)=E\int_{R_M}\varphi\varphi=E\langle\varphi|\varphi\rangle.$$

Therefore φ satisfies the same boundary conditions as φ_0 and has the same variational energy. This implies that among the ground states of (12) and (13) (assuming the possibility of a degeneracy) there is at least one satisfying (14). The following are also true, although the proof is tedious:

(i) If V is bounded, then in fact $\varphi_0\neq 0$ *inside* R_M.

(j) There can be no degeneracy unless $\varphi_0=0$ *inside* R_M.

We can now define the term "pathologic potential." It is a potential with a sufficiently strong infinity to cause φ_0 to vanish *inside* R_M. An infinite repulsive core is an example. Thus there is, in general, no ground-state degeneracy, but even if there is, one of the eigenfunctions satisfies (14).

Let P be a permutation of the variables $x_1,\cdots,x_p$ and Q a permutation of the variables $x_{p+1},\cdots,x_N$, and define $PQ(R_M)$ as the domain defined by the appropriate permutation of the variables in condition (11). All the domains $PQ(R_M)$ are disjoint, except possibly for the boundaries, and together span R. Let φ be any solution to (12) and (13) and define the function Φ everywhere in R by

$$\Phi=(-)^P(-)^Q PQ\varphi \quad\text{in}\quad PQ(R_M). \tag{15}$$

Because of the boundary conditions, (13), it is easily verified that Φ is continuous and has a continuous derivative everywhere in R, and hence satisfies

$$H\Phi=E\Phi \quad\text{in}\quad R. \tag{16}$$

Conversely, any eigenfunction of H in R that satisfies Eq. (9) satisfies (12) and (13) in R_M. Thus if φ_0 is the ground state of (12) and (13), Φ, as defined by (15), is the ground state of H belonging to M.

Now consider

$$\begin{aligned}{}_M{}^M\chi&=\operatorname{Det}\begin{vmatrix}1x_1\cdots x_1{}^{p-1}\\ \vdots\qquad\vdots\\ 1x_p\cdots x_p{}^{p-1}\end{vmatrix}\times\operatorname{Det}\begin{vmatrix}1x_{p+1}\cdots x_{p+1}{}^{N-p-1}\\ \vdots\qquad\vdots\\ 1x_N\ \cdots x_N{}^{N-p-1}\end{vmatrix}\\ &=\prod_{\substack{j,k=1\\ j>k}}^{p}(x_j-x_k)\times\prod_{\substack{j,k=p+1\\ j>k}}^{N}(x_j-x_k).\end{aligned} \tag{17}$$

(Except for a totally symmetric Gaussian factor, this is the solution to the problem of noninteracting one-dimensional electrons in a harmonic oscillator potential.) The function ${}_M{}^M\chi$ clearly satisfies (9), (10), and (14). It is easily verified that if ${}_M{}^Sf$ and ${}_M{}^{S'}g$ are any two functions having different S values but the same M value, then

$$\int_R {}_M{}^Sf\,{}_M{}^{S'}g=P!(N-P)!\int_{R_M}{}_M{}^Sf\,{}_M{}^{S'}g, \qquad (p=\tfrac{1}{2}N-M) \tag{18}$$

and further the right-hand side of (18) vanishes if $S\neq S'$. Since ${}_M{}^M\chi$ and the ground state of H belonging to M, ${}_M\varphi$, are both non-negative in R_M, Eq. (18) implies that ${}_M\varphi$ is not orthogonal to ${}_M{}^M\chi$ in R. If the ground state of (12) and (13) is nondegenerate, then ${}_M\varphi$ can belong to only one S value and this S value must therefore be $S=M$. If one wants $S=M+1$ or higher, it is necessary to go at least to the first-excited state of (12) and (13). In case of degeneracy, at least one of the ground-state functions belongs to $S=M$.

This completes the proof of Theorem I.

II. GENERALIZATION OF THEOREM I, AND THE "POURING PRINCIPLE"

Thus far we have considered eigenfunctions of H which have the symmetry property (9), the only allowed symmetry class for constructing a Pauli function of space and spin. But there are other symmetry classes with their corresponding eigenfunctions—the totally symmetric function, for example. The latter is a Bose function and plays no role for fermions, unless only two particles are involved.

The most widely known classification of symmetry classes is due to Young; but we shall find it convenient to use a slightly modified version of his scheme.[4] It is well known that every function of 2 variables can be written as the sum of an antisymmetric and a symmetric function which *a fortiori* are orthogonal to each other. Thus,

$$\phi(x_1,x_2)=\tfrac{1}{2}(1+p_{12})\phi+\tfrac{1}{2}(1-p_{12})\phi. \qquad (19)$$

The operator $1-p_{12}$ is said to antisymmetrize ϕ, while the operator $1+p_{12}$ symmetrizes it. If ϕ is a function of $x_1,\cdots,x_N$, the operator $1-p_{12}$ antisymmetrizes it with respect to the variables x_1,x_2; the operator $(1-p_{13}-p_{23})(1-p_{12})$ antisymmetrizes it with respect to x_1, x_2, x_3, and so forth. It is quite possible that ϕ may be antisymmetrized with respect to x_1 and x_2, but not with respect to x_1, x_2, x_3; $(x_1-x_2)x_3$ is an example of such a function.

Now consider a function ϕ of the variables $x_1,\cdots,x_N$ which is of the form

$$\phi(x_1,\cdots|\cdots x_{N-n_1-n_2}|x_{N-n_1-n_2+1},\cdots,x_{N-n_1}|\\ x_{N-n_1+1},\cdots,x_N), \qquad (20)$$

or any of its permutations, by which we mean that it is separately antisymmetric in the n_1 variables $x_{N-n_1+1},\cdots,x_N$ and in the n_2 variables $x_{N-n_1-n_2+1},\cdots,x_{N-n_1}$ and so forth, where $n_1\geq n_2\geq\cdots$. The bars are to be regarded as unmovable leftwards[5]; e.g., ϕ cannot be antisymmetrized with respect to $x_{N-n_1},\cdots,x_N$. The function ϕ is said to belong to the symmetry class characterized by the numbers n_1, n_2, etc. For example, a totally symmetric function has N "boxes" with one variable in each, whereas an antisymmetric function has 1 "box" containing all N variables. It is possible to prove the following properties of ϕ:

(a) Any bar may always be moved to the right (i.e., the antisymmetrization procedure which would move the bar to the right gives a nonvanishing result) if the number of variables to the right of the bar is greater than the number to the left.

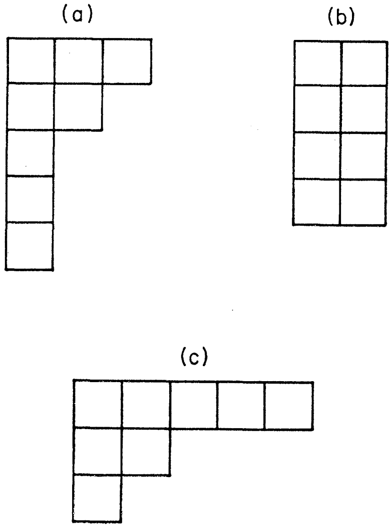

FIG. 1. The tableaux corresponding to several different symmetry classes for eight particles. By the pouring principle $E(a)>E(c)$ and $E(b)>E(c)$, but one cannot say whether (a) or (b) has the lower energy. Note that (c) happens to be the conjugate of (a).

(b) The largest group of variables with respect to which ϕ may be antisymmetrized is n_1. From among the remainder, the largest group would be n_2, and so forth.

(c) If ϕ and ψ belong to two different symmetry classes, they are orthogonal.

(d) For a function of the form (20), the antisymmetrization operator which moves a bar to the right has as its unique inverse the operator which moves the bar to the left. Therefore, if ψ is obtained from ϕ by moving a bar to the right, it is said to belong to the same class as ϕ.

From these remarks we see that by a combination of antisymmetrization and orthogonalization, one can reduce an arbitrary function to a sum of orthogonal functions, each belonging to a different symmetry class. The function in each symmetry class itself is decomposable into a sum of functions of the form (20) and its permutations, although this last decomposition is not unique.

There is a convenient pictorial representation for the symmetry classes called *tableau* illustrated in Fig. 1. One draws a column of n_1 boxes. To the right of it, and starting at the same height, one places a column of n_2 boxes, and so forth. A function of the form (20) is further designated by inserting the appropriate variables in the appropriate column, the order of the variables in any column being immaterial. It will be seen that the lengths of the *rows* decrease from top to bottom; it is therefore possible to define the conjugate to a symmetry class to be the one in which rows are replaced by columns. Thus (a) and (c) in Fig. 1 are conjugate. We shall return to this later, however.

Returning to the problem at hand, it will be appreciated that since H is permutation-invariant, every eigenfunction not only belongs to a definite symmetry class, but is of the form (20). Theorem I, then, tells us about the order of the ground-state energies of functions of the one- and two-column class. Under certain circum-

---

[4] See, for example, D. E. Rutherford, *Substitutional Analysis* (Edinburgh University Press, Edinburgh, 1948). Young's scheme proceeds by symmetrization, instead of antisymmetrization as used here.

[5] We have departed slightly from the notation in Eq. (9). Formerly the bar was regarded as possibly movable leftwards.

stances it is possible to extend this theorem to more than two-column classes. To do this we need one more concept, which we shall call the "pouring principle."

Definition. If α and β are two different symmetry classes, it is possible to *pour α into β* if the bars in the α function can be moved to the right [subject to (a) above] so that one gets a function antisymmetric in the same groups of variables as the β function. More formally, if the α tableau has the columns $n_1 \geq n_2 \geq$, and the β tableau has the columns $n_1' \geq n_2' \geq \cdots$, then we require that $n_1 \geq n_1'$; $(n_1 - n_1') + n_2 \geq n_2'$; $(n_1 - n_1' + n_2 - n_2') + n_3 \geq n_3'$; etc., where any missing columns are to be regarded as having $n=0$. If α can be poured into β, we denote this fact by $\alpha \xrightarrow{p} \beta$.

Thus, in Fig. 1, (a) can be poured into (c) and (b) into (c), but neither (a) nor (b) can be poured into the other. Note that if α can be poured into β, then the conjugate of β can be poured into conjugate of α. We can now state the extension of Theorem I in one dimension.

Theorem II. Let α and β be two symmetry classes and let $E(\alpha)$ and $E(\beta)$ be the respective ground-state energies of eigenfunctions of H in the two classes. If α can be poured in β, then $E(\alpha) > E(\beta)$ unless V is pathologic, in which case $E(\alpha) \geq E(\beta)$.

The proof is exactly the same as for Theorem I. One antisymmetrizes the α function until it matches the bars of the β function (a process which does not change the energy of the α function). Next one defines the fundamental region in analogy with (11) above and shows that the ground-state function in this region is positive, and is therefore not orthogonal to a determinantal function which is positive in this region and is known to belong to the β class.

Corollary: Consider[2]

$$H = -\sum_{i=1}^{N} \frac{\partial^2}{\partial x_i^2} - \sum_{i=1}^{M} \frac{\partial^2}{\partial y_i^2} + V(x_1,\cdots,x_N;\, y_1,\cdots,y_M). \qquad (21)$$

If V is symmetric in the x variables, then every eigenfunction belongs to some symmetry class under permutation of these variables. Theorem II obviously is true for this more general problem. If, moreover, V is also separately symmetric in the y variables, although of no particular symmetry under the interchange of an x with a y, then every eigenfunction falls into some symmetry class α in the x variables and some class α' in the y variables. If two functions are characterized by (α,α') and (β,β'), respectively, then $E(\alpha,\alpha') > E(\beta,\beta')$ if α can be poured into β and α' into β'.

III. SEPARATELY SYMMETRIC POTENTIAL IN HIGHER DIMENSIONS

We shall now turn to the proof of the extension of Theorem I to higher dimensions when the potential is separately symmetric. Only two dimensions will be explicitly considered, for the extension to three dimensions is a corollary. The general fermion Hamiltonian is of the form (21) with $M=N$ and V is a symmetric function of the pairs x_1, y_1, etc. A separately symmetric potential is one for which V is, in addition, separately symmetric in the x variables, and in the y variables. The theorem to be proved is:

Theorem III. If V is separately symmetric and if $S > S'$, then $E(S) \geq E(S')$. (Note that Theorem III is not quite so strong as Theorem I, because the equalities can occur without pathologic potentials.)

Before discussing Theorem III, it is necessary to consider the "Kronecker product" of two tableaux. Suppose we have a function of the x's and the y's which is of a definite symmetry class α in the x's and a definite class α' in the y's. Let us consider this function together with all the functions derived from it by permutations, and let us ask if there is some linear combination of these functions which is of a definite symmetry class, β, in the N *pairs* of variables (x_i,y_i). If so, which β classes can be generated and how many independent β functions are there? [By independent β functions we mean functions which cannot be obtained from each other by permutations of the (x_i,y_i) pairs.] This question is analogous to the problem in the theory of angular momentum of combining J_1 and J_2 to give a resultant J. There the answer is given simply by the Clebsch-Gordan theorem: All $J_1 + J_2 \geq J \geq |J_1 - J_2|$ may be produced once and only once. Unfortunately, there is no such simple rule for tableaux except in two special cases: when β is symmetric, or antisymmetric (i.e., one row, or one column).

Lemma I. To generate a totally antisymmetric function in the pairs $x_i y_i$, it is required that the x and y tableaux α and α' be conjugates. (To generate a totally symmetric function in the pairs, the x tableau must be the same as the y tableau.)

This is a well-known result from the theory of permutation groups.[4]

As we saw in Sec. I, the spatial part of a Pauli (antisymmetric) function of definite S must be of the two-column type, where $n_1 = (N/2) + S$ and $n_2 = (N/2) - S$. By Lemma I we are led to believe that spin functions of a given S are of the conjugate two-row type. This is indeed the case. There are $\binom{N}{p}$ different G_M functions, and these will be seen to generate all two-row tableaux in which the first row is $(N/2)+M$ or longer. Because the ${}_M\varphi_j$ in Eq. (6) are all derivable from each other, they will not all be linearly independent. The G_j^{M}'s therefore appear only in certain definite linear combinations; if φ is of the two-column type, these linear combinations can be shown to be of the conjugate two-row type.

Any Pauli eigenfunction of H of definite S value is thus seen to be the triple Kronecker product of a function belonging to the α and α' class in the x and y variables, respectively, and of a two-row function of the spin variables. The resultant must be a one-column function in the triplets (x_i,y_i,s_i), where s_i is the spin variable. The problem can be viewed in two ways: a Kronecker product of α and α' must be of the appropriate two-column type; or a Kronecker product of α and the two-row spin function must be conjugate to α'.

To prove Theorem III, we take the latter view. Let α and α' be the tableaux of the spatial function, giving the lowest energy for a given S value and suppose α' is (c) of Fig. 1, so that the (x_i,s_i) class, which we shall denote by β, is (a) of Fig. 1. A typical representative, χ, would have each of the pairs (x_i,s_i) in some box of the (a) tableau and each of the y_i in some box of the (c) tableau. Suppose x has (x_1,s_1), $\cdots$, (x_5,s_5) in the first column, (x_6,s_6) and (x_7,s_7) in the second, and (x_8,s_8) in the third. It will be seen that χ may be regarded as a five-particle one-dimensional Pauli function in the first five pairs, or as a two-particle Pauli function in the next two pairs.

Now, to prove Theorem III we need the following lemma:

Lemma II. Let χ be an eigenfunction with the following properties: (1) it is of the α and α' symmetry classes in the x and y variables, respectively, and is in fact formed from the lowest eigenfunction of H having these classes; (2) in the (x,s) pairs, it is antisymmetric in the same sets of variables as a function of class β, the conjugate of α', but it is not necessarily itself of the class β; (3) considered as a function of each of the sets of (x,s) pairs in which it is antisymmetric, it has a definite S value (i.e., each column has a definite S value). These S values we shall call $S_1, S_2, \cdots$. Let S be any total S value for all the particles which can be compounded from $S_1, S_2, \cdots$ by the usual Clebsch-Gordan rule. Then $E(S) \leq E(\alpha,\alpha')$.

Proof. By applying the appropriate S_+ and S_- operators of each column the appropriate number of times to χ, we can generate functions having all possible M values in each column and which still have the same energy as χ. These functions are then to be added together with the appropriate coefficients to generate a new function, ${}^S\chi \neq 0$ having the required total S value. ${}^S\chi$ has the same energy as χ. Since the total S^2 operator commutes with all permutations, ${}^S\chi$ may be written as the sum of functions belonging to definite symmetry classes and all having the given S value. One of these tableaux may be β itself, in which case the lemma is proved. If not, then since ${}^S\chi$ is already antisymmetric in the same sets of variables as the β tableau, at least one of the component tableaux, say γ, must be such that $\gamma \xrightarrow{p} \beta$. To make a Pauli function out of this component we should need a function whose y class is γ^+, the conjugate of γ. But γ^+ satisfies $\alpha' \xrightarrow{p} \gamma^+$, and therefore $E(\alpha,\gamma^+) \leq E(\alpha,\alpha')$. Thus, if we carry out the same procedure again with the ground-state function of the (α,γ^+) class, we shall be able to construct a Pauli function having the S value in question and with an energy lower than $E(\alpha,\alpha')$.

Corollary. The ground-state function of a given S value has the properties of χ above, and, moreover, belongs to the symmetry class β in the (x,s) variables.

Proof. If not, then by the procedure of the above proof, we should either be able to lower the energy by changing the y tableau, or else we should end up with the function ${}^S\chi$ having the properties stated in the corollary.

Now we consider the proof of Theorem III. The lowest eigenfunction having a given $S \neq 0$ value is the sum of permutations of a function ${}^S\chi$ having the properties stated in the corollary and therefore has the S values S_1, S_2, etc., in each column of β, the (x,s) tableau. There are three cases to be considered:

(1) If S is not the minimum that can be compounded of S_1, S_2, etc., then we can construct a function of spin $S-1$ and with at least as low an energy.

(2) If S_1, S_2, etc., are already the lowest possible (i.e., either 0 or $\frac{1}{2}$) then (1) above is transparently true.

(3) If S is indeed the minimum of $S_1 \oplus S_2 \oplus$, etc., then lowering some one of these S_i values, say S_1, would permit us to create a function of total S equal to $S-1$. But there does exist a function having a lower energy than ${}^S\chi$ and having the properties of χ listed in Lemma II, except that the first column has the value S_1-1. To see this, we regard the function having the α tableau in the x's as being the sum of functions each belonging to definite symmetry classes in each of the groups of x variables appearing in each column of β. One of these sets of tableaux must be the respective conjugates of S_1, S_2, etc. Looking at χ from this point of view, it is clear from the results of Secs. I and II that we can lower one or more of the S_i at will and at the same time lower the energy.

This completes the proof of Theorem III for two dimensions; the extension to higher dimensions is obvious: We simply treat the s_i, z_i pairs in the same manner as the spins above. Let us remark, however, on the reason for the lack of strict inequality as we had in Theorem I. Suppose the lowest function of $S=1$ had an (x,s) tableau such as (a) in Fig. 1 with $S_1=S_3=\frac{1}{2}$ and $S_2=0$. Since each column already has its lowest possible S value, the only way in which $S=0$ could have a *lower* energy is by having an (x,s) tableau containing only columns of even length, such as (b) of Fig. 1, with $S=0$ in each. But it is not possible to prove that such a function has, in fact, a lower energy, and therefore it is possible for the ground state as well as for excited states to have a degeneracy in more than one dimension. (This degeneracy can be estimated never to exceed $N^{\frac{2}{3}}$, in three dimensions. It is therefore not an extensive property of the system.)

IV. ON THEORIES OF FERROMAGNETISM

Although we could extend them to other particles obeying various statistics, the results of this paper apply most directly to the problem of interacting electrons, and as such have some bearing on the theory of ferromagnetism.

It is well known that ferromagnetism must be a consequence of the electronic interactions, for a noninteracting electron system always obeys the theorem $E(S) \leq E(S+1)$, regardless of the external potential. This is antiferromagnetism, or at most, paramagnetism. Ferromagnetism is assumed to occur when the ground state belongs to a nonvanishing S whose magnitude is proportional to the size of the system. It is also well known that the direct magnetic spin-spin forces are negligibly weak, so that the spatial forces in conjunction with the Pauli principle are held to be responsible for the phenomenon.

The simplest realistic problem which offers some hope of being soluble is the linear chain of three-dimensional atoms. The atomic states are supposed known when the atoms are infinitely far apart, and the problem is to find the new configurations when overlap becomes important. Peierls[6] considers this very problem in the chapter on ferromagnetism in his book. The approximation which he makes is that there is only one orbital state per atom, and he concludes that the electronic interactions can lead to ferromagnetism. However, recent and more realistic calculations[7–9] on such chains have proved the contrary to be true. Also in Sec. I we showed that under no circumstances can a one-dimensional electron system be ferromagnetic with only space-dependent forces; this includes the special case of a chain of one-dimensional "atoms." In the Appendix we also treat a model applicable to an idealized chain of three-dimensional atoms, with similar results. It therefore seems that a linear chain can be magnetic only if the individual atoms have orbital degeneracy, that is, if the single atom displays a magnetic or truly three-dimensional character; but it is not known whether this is a *sufficient* condition for ferromagnetism to occur in a linear system.

Our theorem has no relevance to atomic magnetism *per se* (Hund's rule) because it does not apply to the central force problem in three dimensions. But if we consider ferromagnetism to be an extensive property of a solid, the theorem does have some relevance. For we shall show that it is not merely sufficient to have (i) a band structure, (ii) strong repulsive interactions, and (iii) three dimensions, to produce ferromagnetism. We shall base ourselves on the results of Sec. III.

Suppose, for example, that highly magnetized states of a noninteracting set of electrons lie rather close in energy to the $S=0$ ground state. If one introduces a repulsive interaction potential into the problem, and treats this by lowest-order perturbation theory, certain terms called the "exchange integral" will favor the magnetized states. To lowest order, one may find that the magnetized states have crossed the $S=0$ state, and the interacting system is supposed to become ferromagnetic. But this conclusion would be fallacious if the effect were cancelled by second-order or higher-order terms in the perturbation series (or if the perturbation expansion did not converge). Indeed, we now give an example based on the previous section, which is a case in point. Consider, for example, the unperturbed Hamiltonian to be

$$H_0 = \sum p_i^2 + \sum_i \left[V(x_i) + V(y_i) + V(z_i)\right]. \tag{22}$$

Let $V(x)$ be a periodic potential so that one-electron eigenfunctions are Bloch functions. The potential can be chosen such that the bands display the usual degeneracies and other features of motion in a three-dimensional cubic lattice.

If now we introduce an interaction, say

$$H_I = g^2 \sum_{ij} \left(\frac{1}{|x_i - x_j| + d} + \frac{1}{|y_i - y_j| + d} + \frac{1}{|z_i - z_j| + d} \right), \tag{23}$$

or some other repulsive, separately symmetric interaction, the total Hamiltonian $H_0 + H_I$ is still subject to our theorem and is *not* ferromagnetic. (The theorem does not exclude paramagnetism, however, for the ground state might be degenerate with states of nonvanishing spin angular momentum.) But what are the conclusions we would reach if we were to apply first-order perturbation theory to H_I? This amounts to calculating the expectation value of $H_0 + H_I$ using the Slater determinants appropriate to the unperturbed problem. The unperturbed functions with the most spatial nodes are better, variationally speaking, than those with fewer nodes for *sufficiently large* g^2, and we might be led to conclude that there exist some values of S such that $E(S) < E(0)$, which is erroneous. It is therefore clear that we cannot always trust perturbation theory to properly order the levels, for when it is carried out only to finite order, it might be more accurate for some values of S than for others, depending on the particular features of the problem. The same can be said of variational calculations.

A notorious example of the above is the low-density electron gas with Coulomb interactions which is in a background of compensating positive charge. The expectation value of the Hamiltonian using the unperturbed plane-wave states is lower for the ferromagnetic configuration than for the nonferromagnetic ones, at sufficiently low density. But perturbation theory diverges for this problem,[10] and this ferromagnetism is indeed fictitious. A recent and accurate calculation[11] by

---

[6] R. E. Peierls, *Quantum Theory of Solids* (Oxford University Press, New York, 1955).

[7] R. K. Nesbet, Phys. Rev. **122**, 1497 (1961).

[8] L. F. Mattheiss, Phys. Rev. **123**, 1209 (1961).

[9] David I. Paul, Phys. Rev. **118**, 92 (1960), and Phys. Rev. **120**, 463 (1960).

[10] M. Gell-Mann and K. A. Brueckner, Phys. Rev. **106**, 364 (1957).

[11] W. J. Carr, Jr., Phys. Rev. **122**, 1437 (1961).

Carr leads that author to conclude that at all but the lowest densities the electronic spins are *antiferromagnetically* aligned.

In concluding, we should recall that our theorem is not valid if there are explicitly spin-dependent forces, or velocity-dependent forces. In the latter case, the eigenfunctions are not real and our method of proof does not apply. Nor does it apply to the Coulomb potential which governs real electrons. But it does serve as a warning that the criterion for ferromagnetism must be rather detailed, and not so broad as to violate the results of this investigation.

We have also found it possible to order many of the energy levels of the Heisenberg Hamiltonian $\sum K_{ij}\mathbf{S}_i\cdot\mathbf{S}_j$ (where the S_i are spins on a lattice in one, two, or three dimensions), by analogy with the calculation in the Appendix. These results will be reported in a subsequent publication.

ACKNOWLEDGMENTS

We should like to thank our colleagues at I.B.M. for many useful discussions, and particularly Dr. D. Jepsen, Dr. T. D. Schultz, and Dr. M. Gutzwiller for their excellent suggestions.

APPENDIX. ONE-DIMENSIONAL LATTICE GAS

We are interested in a theorem analogous to the one in the text, for a chain of *three*-dimensional atoms. This problem cannot be solved in all generality, therefore we are led to consider the following tractable model.

(a) We use a truncated Hamiltonian such that only valence electrons are mobile.

(b) Each atom in the linear chain has only one valence state (capable of double occupancy, however, because of spin degeneracy).

(c) The atoms are at a distance d from their nearest neighbors. This distance is such that only nearest-neighbor overlap is important.

(d) The matrix element for a one-electron hop from site j to $j\pm 1$ is K, a constant. Two-electron hops and exchange effects are neglected. This is equivalent to

(e) assuming that aside from the "hopping" matrix elements the Hamiltonian is diagonal, with an energy calculable by specifying which atoms have empty valence states, which have singly occupied valence state, and which have doubly occupied valence states. These assumptions lead directly to Eq. (A4).

Our model reduces, in a certain limit, to the one-dimensional problem of Sec. I. The Appendix provides therefore an alternate proof for Theorem I.

To see this, let us consider one-dimensional space as consisting of discrete points labeled $i=1,2,\cdots,N$, separated by a distance d. The length of the chain is therefore Nd. Next, introduce the second-quantized Fermi operators c_{is} and $c_{is}^\dagger$, where $s=$ "up" or "down" according to the spin coordinate, and

$$[c_{is},c_{js'}]_+=0,\quad [c_{is}^\dagger,c_{js'}]_+=\delta_{ij}\delta_{ss'}. \tag{A1}$$

Now consider the Hamiltonian

$$H=-K\sum_{i,s}(c_{i+1,s}{}^\dagger c_{is}+c_{i-1,s}{}^\dagger c_{is}-2c_{is}{}^\dagger c_{is})\\ +\sum_{i,j,s,s'}V(|i-j|)c_{is}{}^\dagger c_{is}c_{js'}{}^\dagger c_{js'}. \tag{A2}$$

If we transform to running waves,

$$H\rightarrow 2K\sum(1-\cos kd)c_{ks}{}^\dagger c_{ks}\\ +\sum V_g c_{k+gs}{}^\dagger c_{ks}c_{k'-gs'}{}^\dagger c_{k's'}, \tag{A.3}$$

where V_g is the Fourier transform of $V(|i-j|)$. In the limit $K^{-\frac{1}{2}}=d=N^{-1}=0$ this reduces to the problem of a one-dimensional electron gas with two-body forces. The Hamiltonian (A2) is a special case of the problem we shall now consider,

$$H=-K\sum_{i,s}(c_{i+1,s}{}^\dagger c_{is}+\text{H.c.})+V(\cdots n_j\cdots), \tag{A4}$$

where V is an arbitrary symmetric function of the operators $n_j\equiv(c_{j\uparrow}{}^\dagger c_{j\uparrow}+c_{j\downarrow}{}^\dagger c_{j\downarrow})$. This Hamiltonian satisfies (a)–(e), is identical with the general Hamiltonian of Sec. I in the limit $K^{-\frac{1}{2}}=d=0$, and can be shown to commute with the spin operators which, in our new representation, are

$$S_z=\tfrac{1}{2}\sum_i(c_{i\uparrow}{}^\dagger c_{i\uparrow}-c_{i\downarrow}{}^\dagger c_{i\downarrow}),\quad S_y=\tfrac{1}{2}i\sum(c_{i\uparrow}{}^\dagger c_{i\downarrow}-c_{i\downarrow}{}^\dagger c_{i\uparrow}),$$

and

$$S_x=\tfrac{1}{2}\sum(c_{i\uparrow}{}^\dagger c_{i\downarrow}+\text{H.c.}). \tag{A5}$$

The problem is soluble because there exists a transformation to Pauli (pseudo-spin) variables, in which the Schrödinger equation can be reduced to a series of algebraic equations. We define the Pauli operators as follows:

$$\begin{aligned} b_{i\downarrow}&\equiv c_{i\downarrow}\exp\{\pi i\sum_{j=1}^{i-1}c_{j\downarrow}{}^\dagger c_{j\downarrow}\},\\ b_{i\uparrow}&\equiv c_{i\uparrow}\exp\{\pi i[\sum_{j=1}^{N}c_{j\downarrow}{}^\dagger c_{j\downarrow}+\sum_{j=1}^{i-1}c_{j\uparrow}{}^\dagger c_{j\uparrow}]\}, \end{aligned} \tag{A6}$$

The $b_i{}^\dagger$'s are given by the Hermitean conjugates of these defining equations. All Pauli operators *commute except* b_{is} and $b_{is}{}^\dagger$, (for all i and s), which anticommute:

$$b_{is}{}^\dagger b_{is}+b_{is}b_{is}{}^\dagger=1. \tag{A7}$$

In terms of these new operators,

$$H=-K\sum(b_{i+1,s}{}^\dagger b_{is}+\text{H.c.})+V(\cdots n_j\cdots), \tag{A8}$$

where $n_j=b_{j\uparrow}{}^\dagger b_{j\uparrow}+b_{j\downarrow}{}^\dagger b_{j\downarrow}$. The Hamiltonian remains simple under this transformation only for this very special case of a linear chain, and nearest-neighbor hops. We now assume $K>0$. Otherwise, K can be made positive by a trivial canonical transformation.

Also, for the moment, let us pretend that the number of electrons is even, and is $2p$. Obviously, p cannot exceed N, the number of sites, as each site can accommodate two electrons at most.

Ground State in $M=0$ Subspace

States of all allowed spin angular momentum can be rotated into the $M=0$ subspace with no change in energy. The ground state here is therefore the ground state of the Hamiltonian. The configurations which form a complete set in this subspace have p electrons with spin up and p with spin down. (A configuration specifies which sites are vacant or occupied); for example, if $p=1$ and $N=2$, the complete set of $M=0$ configurations is

$$\phi_1=b_{1\uparrow}{}^\dagger b_{1\downarrow}{}^\dagger|0\rangle, \qquad \phi_2=b_{2\uparrow}{}^\dagger b_{2\downarrow}{}^\dagger|0\rangle,$$
$$\phi_3=b_{1\uparrow}{}^\dagger b_{2\downarrow}{}^\dagger|0\rangle, \quad \text{and} \quad \phi_4=b_{1\downarrow}{}^\dagger b_{2\uparrow}{}^\dagger|0\rangle.$$

The Pauli operators for different particles commute, and therefore, the configurations can all be defined to have the same sign for arbitrary ordering of the operators. The number of distinct configurations is $t\equiv\binom{N}{p}^2$, and we shall label them ϕ_α, where $\alpha=1, 2, \cdots, t$.

Let the ground-state function be

$$\psi_0=\sum f^\alpha\phi_\alpha, \tag{A9}$$

with energy E_0 and real amplitudes f^α. If we denote the eigenvalues of V by V_α,

$$V\phi_\alpha=V_\alpha\phi_\alpha, \quad |V_\alpha|<\infty, \tag{A10}$$

then Schrödinger's equation can be expressed in terms of the amplitudes as

$$-K\sum\nolimits_{\beta(\alpha)} f^{\beta(\alpha)}=(E_0-V_\alpha)f^\alpha. \tag{A11}$$

The index $\beta(\alpha)$ runs over those configurations for which

$$\langle\beta(\alpha)|H-V|\alpha\rangle\neq 0. \tag{A12}$$

We parenthetically observe that a variational function,

$$\psi=\sum g^\alpha\phi_\alpha, \tag{A13}$$

has variational energy W,

$$W=\frac{-K\sum\sum g^\alpha g^{\beta(\alpha)}+\sum V_\alpha(g^\alpha)^2}{\sum(g^\alpha)^2}\geq E_0. \tag{A14}$$

Clearly, all nonzero amplitudes f^α can be chosen positive in the ground state. For if they oscillate in sign, define a trial function by $g^\alpha\equiv|f^\alpha|$. Then by inspection of (A14), $W\leq E_0$. This is a contradiction unless

$$\psi'\equiv\sum|f^\alpha|\phi_\alpha,$$

is also a ground-state eigenfunction. Therefore,

$$-K\sum\nolimits_{\beta(\alpha)}|f^{\beta(\alpha)}|=(E_0-V_\alpha)|f^\alpha|. \tag{A15}$$

If any $f^\alpha=0$, then by Eq. (A15) all $f^{\beta(\alpha)}$ also vanish. Since by repeated application of H to any given state, all other states are eventually reached, we can conclude that *all* f^α vanish if some one f^α vanishes. Therefore

$$f^\alpha\neq 0 \text{ for all } \alpha. \tag{A16}$$

(We note that $E_0-V_\alpha<0$ for all α, since the ground-state energy must lie lower than the most favorable potential energy.)

Equations (A11) and (A15) are *incompatible unless* all f^α have the same sign, for if we combine them, we obtain

$$\sum|f^{\beta(\alpha)}|=|\sum f^{\beta(\alpha)}|, \text{ for all } \alpha. \tag{A17}$$

Hence $\psi'=\pm\psi_0$. That the ground state is nondegenerate follows from the observation that all other eigenfunctions of H must be orthogonal to ψ_0 and therefore must have a change of sign, and therefore cannot obey Eq. (A17).

The spin of ψ_0 is found by noting that ψ_0 is *not* orthogonal to the ground state for $V=0$, because they both contain all configurations of Pauli operators with no changes of sign in the amplitudes. The ground state for $V=0$ can be found by inspection of (A3), and belongs to $S=0$. Therefore, the ground state belongs to $S=0$ in general.

Ground State for $M>0$

By a similar procedure, the ground state in any $M>0$ subspace is found to belong to $S=M$. Since each such subspace contains all states of $S\geq M$, this automatically orders the ground states belonging to the various values of S, whether the number of electrons is even or odd. Denoting by $E(S)$ the lowest energy belonging to total spin S, we have therefore proved the following:

Theorem. $E(S+1)>E(S)$.

Note that the restrictions (a)–(e) preclude "double hops," which is related to so-called "exchange." It is therefore reasonable to assume that if ferromagnetism is possible in a linear chain of the sort we have considered, that these neglected exchange matrix elements would be responsible. However, very recent and accurate calculations[7–9] have shown this exchange mechanism to be rather weak. Nesbet[7] finds that the direct "exchange is small compared with the sum of the various antiferromagnetic effects." Mattheiss[8] finds that the true energy levels of such a chain are accurately approximated by the states of the Heisenberg antiferromagnet with nearest-neighbor interactions. Finally, Paul[9] also concludes that linear chains of atoms in s states are nonferromagnetic. It would be interesting to investigate whether an orbital degeneracy on each atom could lead to ferromagnetism for the linear chain in the same way as it appears to be responsible for the magnetic moment of the O_2 molecule.