AN IMPROVED MANY-ELECTRON THEORY
FOR ATOMS AND MOLECULES WHICH
USES EIGENFUNCTIONS OF TOTAL SPIN

Thesis by

William Andrew Goddard, III

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1965
(Submitted September 18, 1964)
ACKNOWLEDGMENTS:

I wish to thank my research advisor, Professor Pol Duwez, for his encouragement and patience which were important to the eventual completion of this project. I wish to express my sincere appreciation to Professor C. A. Barnes, Dr. S. I. Chan, and Dr. R. M. Pitzer for reading my thesis and providing important discussions and criticisms of it. I also wish to express my appreciation to the National Science Foundation for the grant of Graduate Fellowships.

I wish to dedicate this thesis to my teacher, Dr. Paul Pietrokowsky, whose inspiration while I was an undergraduate led directly to my interest in symmetry and solids which in turn led me to the research discussed herein.
AN IMPROVED MANY-ELECTRON THEORY FOR ATOMS
AND MOLECULES USING EIGENFUNCTIONS OF TOTAL SPIN

THESIS BY William Andrew Goddard, III

ABSTRACT

The object is to obtain good approximations for the ground state wave function and energy for atoms and simple molecules (e.g., \( \text{H}_2 \), \( \text{HF} \), \( \text{H}_2\text{O} \), \( \text{CH}_4 \)). We neglect relativistic effects including all spin couplings and we fix the nuclear positions; thus, the Hamiltonian for a molecule with \( N \) electrons is
\[
H = \sum_{i=1}^{N} -\frac{\hbar^2}{2m} \nabla_i^2 + \sum_{i<j} V(r_{ij}) + \sum_{i} \frac{\varepsilon_i^2}{r_i^2},
\]
where \( V(r_{ij}) \) is the electrostatic interaction between the electrons and the nuclear framework. Since the Hamiltonian does not contain spin interactions, then the many electron wave function is an eigenfunction of \( S^2 \); in addition, the many-electron wave function must satisfy Pauli's principle.

A method has been developed to obtain explicitly (for any \( N \)) many-electron wave functions which simultaneously are eigenfunctions of \( S^2 \) and satisfy Pauli's principle. The method is simple and elegant and lends itself readily to applications. Given any function of the spatial coordinates of \( N \) particles, \( \Phi \), and any function of the spin coordinates of \( N \) particles, \( \chi \), then \( G_{i}^{\alpha} \Phi \chi \) is an eigenfunction of \( S^2 \) and satisfies Pauli's principle. We will be particularly interested in the best description of the ground state of the many-electron system by a single \( G_{i}^{\alpha} \Phi \chi \). The primary reason for this is that such a description is readily interpretable and, in addition, the energy promises to be rather accurate.

With no further restrictions (two different sets of orthonormal one-electron functions are used as the basis for the spatial space, i.e.,
different orbitals for different spins, the components of $\Phi$ are any $N$ of these one-electron functions) the best $C^N_{F}$ approximation to the many-electron wave function is found for any $N$ (number of electrons), any $S$ (total spin), and any nuclear configuration. For a given compound calculations can be made for the various nuclear configurations to determine the molecular structure for each possible value of spin. The optimum set of orbitals are each the solution of a one-electron Hamiltonian and thus can be interpreted as the state of an electron moving in the potential due to the other electrons. In addition, these orbitals are not required to be basis functions of the irreducible representations of the spatial symmetry group (as are the Hartree-Fock orbitals) thus, they may be somewhat localized. These optimum orbitals may be of chemical significance.

The very much more restrictive case is considered where only one set of orthonormal basis functions spanning spatial space is used from which to select the $N$ components of $\Phi$. Due to the presence of doubly occupied orbitals this method leads to rather large correlation errors. Using the $C^N_{1}$ method the Hartree-Fock equations and the first-order perturbed wave functions thereof are derived

The VO$_2$ distorted rutile crystal structure is explained and the (uninvestigated) magnetic structure predicted.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>A: A Method of Constructing Many-Electron Wave Functions Which Satisfy</td>
<td>8</td>
</tr>
<tr>
<td>Pauli's Principle and are Eigenfunctions of Total Spin</td>
<td></td>
</tr>
<tr>
<td>a. The Symmetric Group and $G_l^\alpha$</td>
<td>9</td>
</tr>
<tr>
<td>b. The Spatial and Spin Tensor Spaces</td>
<td>15</td>
</tr>
<tr>
<td>c. The General Function Which Satisfies Pauli's Principle and is an</td>
<td>29</td>
</tr>
<tr>
<td>Eigenfunction of $S^2$.</td>
<td></td>
</tr>
<tr>
<td>B: Same Set of Orbitals For Different Spin</td>
<td>33</td>
</tr>
<tr>
<td>a. Orthogonality and Normalization of the $G_l^\alpha \Phi \chi$</td>
<td>35</td>
</tr>
<tr>
<td>b. First- and Second-Order Density Matrices for $G_l^\alpha \Phi \chi$</td>
<td>40</td>
</tr>
<tr>
<td>c. The Set of $G_l^\alpha \Phi \chi$ With the Same $\Phi$; Hund's</td>
<td>46</td>
</tr>
<tr>
<td>Rule and The Heisenberg Hamiltonian</td>
<td></td>
</tr>
<tr>
<td>d. The Hartree-Fock Equations for the $G_l^\alpha \Phi \chi$</td>
<td>52</td>
</tr>
<tr>
<td>e. Perturbation of the Hartree-Fock Equations for $G_l^\alpha \Phi \chi$</td>
<td>62</td>
</tr>
<tr>
<td>f. Construction From the $G_l^\alpha \Phi \chi$ of Basis Functions</td>
<td>75</td>
</tr>
<tr>
<td>of Irreducible Representations of the Spatial Symmetry Group</td>
<td></td>
</tr>
<tr>
<td>C: Different Orbitals for Different Spins</td>
<td>79</td>
</tr>
<tr>
<td>a. The First- and Second-Order Density Matrices</td>
<td>83</td>
</tr>
<tr>
<td>b. The GF Equations</td>
<td>89</td>
</tr>
<tr>
<td>c. The Significance of the GF Orbitals</td>
<td>98</td>
</tr>
<tr>
<td>d. The Defect in The GF Method</td>
<td>103</td>
</tr>
<tr>
<td>e. Restrictions on the GF Orbitals Due to Spatial Symmetry</td>
<td>107</td>
</tr>
<tr>
<td>f. The Alternate Orbital Method</td>
<td>109</td>
</tr>
<tr>
<td>D: Spin Considerations</td>
<td>112</td>
</tr>
<tr>
<td>a. First- and Second-Order Density Matrices</td>
<td>113</td>
</tr>
<tr>
<td>with Unsummed Spin</td>
<td></td>
</tr>
<tr>
<td>b. Concerning the Lowest States of a One-Dimensional System of</td>
<td>115</td>
</tr>
<tr>
<td>Electrons</td>
<td></td>
</tr>
</tbody>
</table>
E: Explanation of the VO$_2$ Distorted Rutile Crystal Structure
and Prediction of the Magnetic Structure .................. 110
a. Survey of the Rutile-Structure Compounds ............... 121
b. The Stable Magnetic Structures For the Rutile
Crystal Structure ........................................... 125
c. The Distorted Rutile Structure .......................... 141

Appendix .......................................................... 150

References ....................................................... 175
INTRODUCTION

We are interested in obtaining good approximations for the ground state wave function and energy for atoms and simple molecules. We neglect relativistic effects including all spin couplings and we fix the nuclear positions; thus, the Hamiltonian for a molecule with \( N \) electrons is

\[
\hat{H} = \sum_{i=1}^{N} \frac{-\hbar^2}{2m} \nabla_i^2 + \sum_{i=1}^{N} V(r_i) + \sum_{\sigma \neq \pi} \frac{\phi^2}{\psi_\sigma^n} \tag{A-1}
\]

where \( V(r_i) \) is the electrostatic interaction between the electrons and the nuclear framework.

A good first approximation to the ground state solution of A-1 is the Hartree-Fock approximation. In the Hartree-Fock scheme one tries to find the lowest energy solution for which the wave function is approximated as a single Slater determinant. For a typical molecular ground state for a molecule with \( N \)-electrons (\( N \) even), the Slater determinant contains \( \frac{N}{2} \) orthogonal molecular orbitals \( \phi_i \) associated with an \( \alpha \) spin and the same \( \frac{N}{2} \) orthogonal molecular orbitals associated with a \( \beta \) spin. The orbitals, \( \phi_i \), are varied in order to minimize the energy. The minimum in the energy is called the Hartree-Fock energy, \( E_{\text{HF}} \), and the orbitals which minimize the energy are called the Hartree-Fock orbitals, \( \phi_i^{\text{HF}} \). The Hartree-Fock orbitals are solutions of a one-electron Hamiltonian (which includes an average electron-electron interaction term) called the Hartree-Fock Hamiltonian, \( \hat{H}_{\text{HF}} \phi_i^{\text{HF}} = \epsilon_i \phi_i^{\text{HF}} \).

The difference between \( E_{\text{HF}} \) and the exact energy of the ground state of A-1 is called the correlation energy. The Hartree-Fock method yields accurate energies as compared to the total energy; however, the correlation energy although small compared to the total energy is of the
order of magnitude of the binding energy of the molecule (SLATER (1960) p.32; SINANOĞLU (1961ab); LÖWDIN (1959) 242,246). For this reason it is desirable to improve upon the Hartree-Fock method in order to obtain better energies and wave functions. Three major methods of improvement have been proposed (LÖWDIN (1959) p.259) (i) superposition of configurations, (ii) wave functions with explicit interelectron coordinates, and (iii) different orbitals for different spin.

Before discussing these methods we consider some possible expansions of the exact wave function \( \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \) (where \( \mathbf{r}_i \) is the spatial \((x_i, y_i, z_i)\) and spin \((\sigma_i)\) coordinate of particle \(i\)). Let \( \psi_i(\mathbf{r}_i) \) be a function of the coordinates of one electron and let the set of \( \psi_i \) for all \( i \) form a complete orthonormal set; then (LÖWDIN (1955a)) the set of \( \psi_i(\mathbf{r}_i) \psi_j(\mathbf{r}_j) \psi_k(\mathbf{r}_k) \) for all \( i, j, k \) forms a complete orthonormal set for the \( N \)-electron space and the exact wave function of an \( N \)-electron system may be expanded as

\[
\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \sum_{i,j,k} C_{ij} \psi_i(\mathbf{r}_i) \psi_j(\mathbf{r}_j) \psi_k(\mathbf{r}_k)
\]

(A-2)

In order to satisfy Pauli's principle it is sufficient to take the Slater determinant of each different set of functions of A-2. If two functions in the determinant are the same, then the determinant is zero; thus, if the subscripts of \( \psi_i \) are ordered, then

\[
\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \sum_{i<j<k} C_{ijk} \psi_i(\mathbf{r}_i) \psi_j(\mathbf{r}_j) \psi_k(\mathbf{r}_k) \text{ DET} \sqrt[3]{\psi_i(\mathbf{r}_i) \psi_j(\mathbf{r}_j) \psi_k(\mathbf{r}_k)}
\]

(A-3)

Since the Hamiltonian, (A-1) is independent of spin, each one-electron function \( \psi_i(\mathbf{r}) \) can be factored into a spatial part \( \phi_i(\mathbf{r}) \) and a spin part \( \sigma_i(\alpha) \), i.e.,

\[
\psi_i(\mathbf{r}) = \phi_i(\mathbf{r}) \sigma_i(\alpha)
\]

We will take \( \sigma_i \) to be an eigenfunction of \( S_z \), viz., \( \sigma_i(S_z \alpha = \frac{\lambda}{2}) \) or
The complete set of one-electron functions, \( \psi'(r') \), in terms of which the actual wave function of the N-electron system is expanded in A-2 is taken to be orthonormal. The one-electron spin functions \( \alpha \) and \( \beta \) form an orthonormal set. What about the one-electron spatial functions? If \( \psi'(r') = \phi'(r'_1) \sigma(r'_2) \) and \( \psi_m = \phi_m \alpha \), then \( \phi \) and \( \phi_m \) are orthogonal when \( \psi \) and \( \psi_m \) are. However, if \( \psi_m = \phi_m \beta \) then we can say nothing about the relation between \( \phi \) and \( \phi_m \). Therefore, we can, at best, consider two complete orthonormal bases spanning the spatial vector space (Löwdin (1955a)) -- one complete orthonormal set for expansion of all \( \psi \) with \( \alpha \) spin and the other orthonormal set for expansion of all \( \psi \) with \( \beta \) spin. We call a one-electron spatial function an orbital. So, we have one complete set of orbitals \( \phi_{ia} \) for \( \alpha \) spin and another complete set of orbitals \( \phi_{ib} \) (over the same space) for \( \beta \) spin. It has been extremely common in the past to require that the set of orbitals be the same for \( \alpha \) and \( \beta \) spin, i.e., \( \phi_{ia} = \phi_{ib} \); in this case we have only one complete set of functions to deal with and an enormous simplification results. Of course, it makes no difference if we really use the infinite expansion of A-3, but we usually only consider a few terms and want to know the best choice for these few terms. This type of an assumption is involved in most discussions of electron shells in atoms, molecular orbitals in molecules, bonding pairs of electrons, lone pairs, etc. The restriction to a single set of orbitals I denote as SODS (same orbitals for different spin), and if different orbitals are allowed for different spins, I denote it as DODS (different orbitals for different spins). We see that the Hartree-Fock wave function is the best approximation to the exact wave function using only one term from A-3 with the
restriction that SODS be used.

In configuration interaction the SODS restriction is used but several terms from A-3 are used to approximate the wave function and the coefficients of the A-3 expansion are varied in order to minimize the energy. According to SINANOĞLU (1962a) the best configuration interaction calculations on small molecules so far have given only one-third to one-half of the binding energy. Löwdin (1959, p.296) states that if the results for helium apply to molecules, then orbitals of sufficiently high angular momentum, in order to get the small absolute errors in the binding energy which are desired, have definitely not been included in configuration interaction calculations. But the inclusion of orbitals of sufficiently high angular momentum would result in an enormous computational problem because of the number of different configurations involved. Besides problems of accuracy, the configuration interaction wave function is quite difficult to interpret (Löwdin (1955a) p.1483).

Wave functions involving interelectron correlations have allowed accurate calculations in He and H₂ and other two-electron systems, but, according to Löwdin (1959, p.304), no strict foundation for a generalization of the theory has been developed. Sinanoğlu has made important contributions by developing an approach for closed shell systems where the correlations on the HF states are solved for one pair at a time (Sinanoğlu (1961a, 1962ab, 1963)). The Sinanoğlu approach is valid only for closed-shell systems, and even for them a wave function involving interelectronic coordinates is hard to interpret (see the difficulty Slater (1960, p.37ff) has when considering the simplest case with correlation, viz., helium).
In the method of different orbitals for different spins (DODS) one allows two different sets of orthogonal orbitals, one for the spin-orbitals with $\alpha$ spin and one for the spin-orbitals with $\beta$ spin (this much restriction can be made without approximation). Since the correlation error is primarily due to two electrons being forced into the same orbital (Löwdin (1959) p.307), then one can hope for a big improvement in accuracy with DODS. Several applications have been made on two-electron problems (see Löwdin (1959) p.307 for references) but the extension to larger numbers of electrons has been prevented by problems in handling the spin problem (Löwdin (1959) p.307).

The Hamiltonian $A$-1 does not involve spin; thus, the wave function must be an eigenfunction of total spin, i.e., of $S^2$. This is no problem to satisfy for molecular ground states in the Hartree-Fock method because the SODS restriction is used and (usually) every molecular orbital is doubly-occupied, once with an $\alpha$ spin and once with a $\beta$ spin. However, when DODS is used no such simplification occurs even for molecular ground state wave functions; thus, we have the spin problem Löwdin refers to. Forgetting the spin problem for the moment, the DODS method is attractive because it may allow quite accurate energies for a single "configuration" (by a single configuration we mean a state well approximated by just $N$ different orbitals and which is an eigenfunction of $S^2$) and because the wave function for such a state admits of some physical interpretation. To be more specific we approximate the exact wave function by a single DODS configuration and vary the orbitals in order to minimize the energy just as in the Hartree-Fock method. The result is that the orbitals which minimize the energy are eigenfunctions of a one-electron Hamiltonian and
thus, each orbital can be interpreted as the state of an electron moving in the average field of all other electrons.

To facilitate the development of the DODS approach, it is important that we construct a general method of writing wave functions (for N-electrons) which satisfy Pauli's principle and which are also eigenfunctions of $S^2$. This is accomplished in section A. The result is an operator $G'_{\alpha}$ which upon operating on a function $\mathbf{F}$ of the spatial coordinates of N-electrons and a function $\mathbf{\lambda}$ of the spin coordinates of N-electrons yields a function $G'_{\alpha}\mathbf{F}\mathbf{\lambda}$ which satisfies Pauli's principle and is an eigenfunction of total spin. Fortunately, the $G'_{\alpha}$ operator has properties adapted to the later applications.

The $G'_{\alpha}$ operator is used in conjunction with the SODS restriction in section B culminating in the Hartree-Fock equations and the first-order perturbed wave functions.

In section C the $G'_{\alpha}$ operator is used in conjunction with DODS in the manner described in the second paragraph before the last. The result is the best possible $G'_{\alpha}\mathbf{F}\mathbf{\lambda}$ function. The resulting one-electron equations for the best orbitals (the Roothaan technique of solving for the best orbitals is used) are called the GF equations and the best orbitals are called the GF orbitals. For any given $N$ this method can be used for any possible $S$, not just the ground state $S$. Because DODS is allowed the GF method can be carried out for any nuclear configuration (as distances go to $\infty$, the orbitals automatically change to atomic orbitals). As distinct from Hartree-Fock orbitals for nondegenerate, $S = 0$ molecular ground states, the GF orbitals are not required to be basis functions of the irreducible representations of the spatial symmetry group. Thus, the
GF orbitals are not required to be spread over the molecule or solid as are the Hartree-Fock orbitals. Thus, the GF orbitals may be somewhat localized and may even be interpretable in terms of such concepts as bonding and non-bonding.
Section A

A Method of Constructing Many-Electron Wave Functions Which Satisfy Pauli's Principle and Are Eigenfunctions of $S^2$

Introduction

In order to develop the DODS approach of solving for the ground state energy and wave function of a molecule, we must construct a general method of writing wave functions (for N-electrons) which satisfy Pauli's principle and which are also eigenfunctions of $S^2$. Methods of accomplishing this have been obtained by KOTANI et al (1955) and by LÖWDIN (1955c). Löwdin's method is not well suited for our purposes because generalizations to arbitrary N are too difficult and because of problems of linear dependency of the resulting wave functions (see p. Ac-3). Kotani's method (see p. Ac-2) could actually be used for the DODS approach; however, we shall find it convenient for our later development to construct a third method from a somewhat different standpoint than Kotani's. In section Ac it is shown that the new method constructs many-electron functions which are equivalent to those constructed with the Kotani method. Sections Aa and Ab are preliminaries discussing the symmetry properties of the symmetric group (the group of permutations) and the general transformations of tensors; the results of sections Aa and Ab are combined in section Ac to obtain a general function which has the desired properties.
The symmetric group on \( N \) objects, \( S_N \), is the group of \( N! \) different permutations of \( N \) identical objects among \( N \) distinguishable positions. We write the permutation \( \gamma = (135)(2678) \) to mean that the object in position 1 is moved to position 3, the object in position 3 is moved to position 5, the object in position 5 is moved to position 1, the object in position 2 is moved to position 6, the object in position 6 is moved to position 7, etc.; note that the object in position 4 is not moved. We do not interpret the above permutation to mean that the object in position 1 is replaced by the object in position 3, etc.; this interpretation is the inverse of the previous one. If the objects are numbered and the positions are not, we interpret the above permutation as object 1 is replaced by object 3, etc., (i.e., the position on object 1 is moved to the position on object 3, etc.); in this case we do not interpret the permutation as object 1 moves to where object 3 is, etc..

There is a pictorial method of discussing the different irreducible representations of \( S_N \) which uses Young's tableaux, e.g., for \( S_4 \) there are five irreducible representations which are characterized by the five Young shapes (or patterns or diagrams),

\[
\begin{array}{cccc}
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array} & \\
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array} & \\
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array} & \\
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array}
\end{array}
\]

If we place numbers in the boxes of a shape, we get a Young tableau, e.g., \( \begin{array}{c}
\begin{array}{c}
2 & 1 & 3 \\
4
\end{array}
\end{array} \). If the numbers 1 through \( N \) are placed such that they always increase down each column and increase to the right in

\[
\begin{array}{cccccccccccccccc}
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array} & \\
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array} & \\
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array} & \\
\begin{array}{c}
\begin{array}{c}
\end{array}
\end{array}
\end{array}
\]

* The first four paragraphs are based on RUTHERFORD (1948),
each row, we have a standard tableau, e.g., \( \begin{array}{ccc} 1 & 2 & 4 \\ \hline 3 \\ 2 \end{array} \) and \( \begin{array}{ccc} 1 & 3 & 4 \\ \hline 2 & 1 & 4 \end{array} \). The number of different standard tableaux of a particular shape is the dimension of the irreducible representation denoted by the shape. We denote a shape by the number of elements in each row, e.g., the shape \( \alpha = \left[ \begin{array}{ccc} 2 & q & r, s^3 \end{array} \right] \) has two rows of length \( p \), one row of length \( q \), one row of length \( r \), and three rows of length \( s \); the convention is that the longer rows be mentioned first, thus \( p > q > r > s \). Remember that \( N = 2p+q+r+3s \), for this reason \( \left[ \begin{array}{ccc} 2 & q & r, s^3 \end{array} \right] \) is called a partition of \( N \).

Some convenient representations of the irreducible representations can be related to the standard tableaux. In particular, the orthogonal representation is of interest. Let \( U_{ij}^{\alpha} \) be the \( ij \) component of the matrix representing the permutation \( \pi \) in the \( \alpha \) irreducible representation of \( S_N \). Consider the \( N-1 \) different elementary transpositions \((k,k+1)\), where \( k \) varies from 1 to \( N-1 \); this set of transpositions is the set of generators of \( S_N \) since any element of \( S_N \) may be expressed in terms of elementary transpositions. We must now define the axial distance in a tableau. Consider two elements, say \( p \) and \( q \), in a standard tableau, let the row and column of the position of \( p \) and \( q \) in the standard tableau be \( r_p \) and \( c_p \), and \( r_q \) and \( c_q \) respectively, e.g., for

\[
\begin{array}{ccc}
1 & 2 & 9 \\ \hline 3 & 5 & 10 \\ 4 & 7 & 12 \\ 6 & 8 \\
\end{array}
\]

\( r_p = 1 \) and \( c_p = 3 \), and \( r_q = 3 \) and \( c_q = 2 \). Define the axial distance \( d_{pq}^{\alpha} \) between \( p \) and \( q \) in tableau \( i \) as

\[
d_{pq}^{\alpha} = (r_p - c_p) - (r_q - c_q),
\]

for our example \( d_{12}^{\alpha} = 3 \) (we see that the axial distance is the number of horizontal and vertical jumps in moving from \( p \) to \( q \) where left and down are positive).

We denote the standard tableau \( j \) of the irreducible representation
\( \alpha \) by \( S_j^\alpha \), where \( j \) ranges from 1 to \( f^\alpha \) and \( f^\alpha \) is the dimension of the irreducible representation \( \alpha \). The different tableaux of the same shape (i.e., same irreducible representation) are obtained from each other by permutations of \( S_N \); we let the permutation taking \( S_j^\alpha \) into \( S_i^\alpha \) be \( \sigma_{ij} \) (i.e., \( S_i = \sigma_{ij} S_j \), the subscript \( \alpha \) is often suppressed below).

Now

\[
U_{ij}^\alpha(k, k+1) = \frac{1}{d_{k+1, \mu}^{\alpha}} \quad (Aa-1)
\]

\[
U_{ij}^\alpha(k, k+1) = 0 \quad \text{if} \quad S_i^\alpha \neq (k, k+1) S_j^\alpha
\]

\[
= \sqrt{1 - \left( \frac{1}{d_{k+1, \mu}^{\alpha}} \right)^2} \quad \text{if} \quad S_i^\alpha = (k, k+1) S_j^\alpha
\]

Therefore, we have off-diagonal elements in \( U_{(k, k+1)} \) only for positions whose tableaux differ only by the transposition of \( k \) and \( k+1 \); also, if \( k \) and \( k+1 \) are in the same column of \( S_i \), then \( U_{ij}^\alpha(k, k+1) = -\delta_{ij} \) and if \( k \) and \( k+1 \) are in the same row of \( S_i \) then \( U_{ij}^\alpha(k, k+1) = +\delta_{ij} \). The matrices for \( N - 2 \) and \( N - 4 \) are in appendix \( F \). Since \( (k, k+1)^2 \) is the identity and \( U_{ij} = U_{ji} \) in \( Aa-1 \), then the matrices for the elementary transpositions are orthogonal. But the product of orthogonal matrices is also orthogonal; so, all the matrices in this representation are orthogonal.

i.e.,

\[
U_{rs}^\alpha U_{sr}^{\alpha^{-1}} = U_{rs}^{\alpha^{-1}} \quad (Aa-2)
\]

We mention the orthogonal units, \( O_{rs}^\alpha \), these are linear combinations of the elements of \( S_N^\alpha \); the set of these for the different \( \alpha \) and \( r \) and \( s \) varying from 1 to \( f^\alpha \) form a basis of the group algebra.

For any shapes \( \alpha \) and \( \beta \) of \( S_N \) and any standard tableaux \( r, s, p, \) and \( q \)

\[
O_{rs}^\alpha O_{pq}^\beta = \delta_{rs}^{sp} \delta_{pq}^{\alpha \beta} \quad (Aa-3)
\]

For any \( \chi \) of \( S_N \), (where \( \sigma^\chi = N! / f^\alpha = \) an integer)
\( \Sigma = \sum_{\alpha, r, s} U_{rs}^{\alpha} 0_{rs}^{\alpha} \)  \hspace{1cm} (Aa-4)

\( 0_{rs}^{\alpha} = \frac{1}{\sigma} \sum_{\alpha} U_{rs}^{\alpha} \)  \hspace{1cm} (Aa-5)

Therefore, \( \Sigma_1 \Sigma_2 = \sum_{\alpha} \sum_{\beta} U_{rs_1}^{\alpha} U_{pq}^{\beta} 0_{rs_2}^{\beta} 0_{pq}^{\alpha} = \sum_{\alpha} U_{rs_2}^{\alpha} U_{sq_1}^{\alpha} 0_{rq_2}^{\alpha} 0_{rq}^{\alpha} \)

We denote the identity or unity permutation by \( \varepsilon \) (stands for einheit); note that \( U_{pq}^{\varepsilon} = \delta_{pq} \).

Some examples:

\( N=2 \)

\[
\begin{align*}
0_{11} &= \frac{1}{2} (\varepsilon + (1,2)) ; \\
S_1 &= [12] \\
\begin{bmatrix} [2] \\
[1] \\
0_{11} &= \frac{1}{2} (\varepsilon - (1,2)) ; \\
S_1 &= \frac{1}{2}
\end{bmatrix}
\end{align*}
\]

\( N=3 \)

\[
\begin{align*}
0_{11} &= \frac{1}{6} \left[ \varepsilon + (1,2,3) + (1,3,2) + (1,2) + (1,3) + (2,3) \right] ; \\
S_1 &= \begin{bmatrix} [3] \\
[2] \\
\begin{bmatrix} [2,1] \\
0_{11} &= \frac{1}{6} \left( 2\varepsilon - (1,2,3) + (1,2,3) + (1,2) + (1,3) + (2,3) \right) ; \\
S_1 &= \frac{1}{2}
\end{bmatrix}
\end{align*}
\]

It is a general result that \( \sum_{\alpha} \sum_{\beta} O_{\alpha\beta}^{\alpha} = \varepsilon \)

Any permutation can be written in a number of different ways as a product of transpositions, but the number of transpositions in such an expansion is either always an odd number or always an even number.
The parity of a permutation, \( \sigma \), is defined as 1 or -1 for permutations expressible as an even or odd number of transpositions respectively. For any number \( N \geq 2 \) there are two irreducible representations of \( S_N \) with dimension one, they are \( \alpha = [N] \) (the trivial representation) and \( \alpha = [1^N] \) (the alternating representation).

\[
\begin{align*}
0^{[N]}_{11} &= \frac{1}{N!} \sum_{\sigma \in S_N} \sigma \\
0^{[1^N]}_{11} &= \frac{1}{N!} \sum_{\sigma \in S_N} \sigma
\end{align*}
\] (Aa-8)

A shape which is obtained from a second shape by interchanging rows and columns is called the associate of the second shape, the same is true for tableaux and the associate is denoted by a bar, e.g., if \( S_j^{\alpha} = \begin{array}{ccc}
1 & 2 & 3 \\
4 & 5 & 7
\end{array} \) then \( S_j^{\alpha} = \begin{array}{ccc}
1 & 2 & 3 \\
4 & 5 & 7
\end{array} \). Note that the associate of a standard tableau is a standard tableau and thus \( \alpha = \overline{\alpha} \). Since

\[
d(k+1,k) = -d(k,k+1)
\]

then

\[
\begin{align*}
U_{i}^{\alpha} \overline{U^{\alpha}}(i,k) &= -U_{i}^{\alpha} \\
U_{i}^{\alpha} \overline{U^{\alpha}}(i,k+1) &= U_{j}^{\alpha} \overline{U^{\alpha}}(j,k+1)
\end{align*}
\] (Aa-9)

We wish to consider linear combinations of permutations acting on two different spaces. We will let \( O_{rs}^{\alpha} \) act only on the first space and \( \omega_{pq}^{\beta} \) act only on the second space; of course, \( O_{rs}^{\alpha} \) and \( \omega_{pq}^{\beta} \) commute. Consider the expression:

\[
U_{i}^{\alpha} O_{r}^{\alpha} \omega_{l}^{\alpha} \omega_{r}^{\alpha} \] (Aa-10)

Remember that \( S_{r}^{\alpha} = S_{r}^{\alpha} S_{i}^{\alpha} \). Let the transposition \( \sigma \) act on \( \sigma_{i}^{\alpha} \); it then acts on both \( O_{rs}^{\alpha} \) and \( \omega_{pq}^{\beta} \). From (Aa-4)

\[
\begin{align*}
\sigma = \sum_{\beta \sigma} \omega_{\beta \sigma}^{\alpha} &- \sum_{\beta \sigma} \omega_{\beta \sigma}^{\alpha} \\
\sigma = \sum_{\beta \sigma} \omega_{\beta \sigma}^{\alpha} &- \sum_{\beta \sigma} \omega_{\beta \sigma}^{\alpha}
\end{align*}
\]

* The symbol \( \equiv \) is read as "is defined as"
Thus, \( g^\alpha_i = \sum_p \sum_{pq} \sum_{rt} \sum_{ri} \int \sigma^p_{rt} \sigma^q_{ri} u^\alpha_{st} u^\beta_{pq} \omega^\alpha_{rt} \omega^\beta_{ri} \) 

\[
= \sum_{opr} \sum_{orp} \int \sigma^p_{rt} \sigma^q_{ri} u^\alpha_{st} u^\beta_{pq} \omega^\alpha_{rt} \omega^\beta_{ri} \quad (Aa-11)
\]

Where we have used \( \sigma_{ri} = \sigma_{rp} \sigma_{pi} \) and \( \int \sigma_{rs} \omega_{sq} = \int \sigma_{sr} \omega_{qs} \); also, note that \( \frac{\partial}{\partial t} = \frac{\partial}{\partial t} = \frac{\partial}{\partial t} \).

If \( \gamma \) is an elementary transposition, then from appendix I

\[
\frac{\partial}{\partial r} u^\alpha_{pr} = -u^\alpha_{qr} .
\]

Thus,

\[
\sum_r \int \sigma^p_{qr} u^\alpha_{sr} u^\beta_{qr} \omega^\alpha_{sr} \omega^\beta_{qr} = -\gamma_{sq} \quad (Aa-12)
\]

Using (Aa-12) in (Aa-11) we obtain

\[
(k,k+1)G^\alpha_i = -\sum_s \int \sigma^p_{si} \sigma^q_{si} u^\alpha_{si} u^\beta_{si} = -G^\alpha_i .
\]

But any permutation \( \gamma \) may be expressed in terms of elementary transpositions and \( \gamma \) will involve an even or odd number of elementary transpositions depending on whether \( \gamma \) is an even or odd permutation, hence:

\[
\gamma^\alpha_i = \frac{\partial}{\partial \gamma} G^\alpha_i \quad (Aa-13)
\]

When \( G^\alpha_i \) operates on the \( N \)-electron space, the resulting function satisfies the Pauli principle.
Consider a group $G$ of linear transformations in an $n$-dimensional space $R_n$. If we choose a basis for $R_n$, then any vector $\mathbf{x}$ in $R_n$ is described by its components, $x_i$, on the basis vectors. If $a_i$, a transformation of the group $G$, transforms $\mathbf{x}$ into $\mathbf{x}'$ then we write (where we use the summation convention, and $\varepsilon$ is read as "is an element of")

$$\mathbf{x}' = a_i x_i = a_{ij} x_j,$$  \hspace{1cm} (Ab-1)

Consider the $n^2$ quantities $x_i y_j$ formed from products of the components of the vectors $\mathbf{x}$ and $\mathbf{y}$ in $R_n$. When the transformation (Ab-1) is applied to the vectors of $R_n$, $x_i y_j$ is transformed as $x_i' y_j' = a_{ik} a_{jl} x_k y_l$, i.e., the $n^2$ quantities $x_i y_j$ transform according to $a \times a$, the Kronecker square of the transformation $a$. A set of $n^2$ quantities $F_{ij}$ whose law of transformation is $F_{ij} = a_{ik} a_{jl} F_{kl}$ forms a tensor of rank two with respect to the group $G$. Similarly, a set of $n^r$ quantities $F_{ij\ldots k}$ whose law of transformation is

$$F_{ij\ldots k} = a_{il} a_{jm} \ldots a_{kp} a_{lm} \ldots a$$ \hspace{1cm} (Ab-2)

forms an $r$th-rank tensor with respect to the group $G$.

Reconsider the case of second rank tensors, say $F_{i_1i_2}$, where $i \in \{1, 2, \ldots, n\}$. Under the transformation of (Ab-1) the set of tensors $(F_{i_1i_2} + F_{i_1i_2})$ transforms into itself and the set of tensors $(F_{i_1i_2} - F_{i_1i_2})$ transforms into itself. For example,

$$a_{j_1i_1} a_{j_2i_2} (F_{i_1i_2} + F_{i_1i_2}) = a_{j_1i_1} a_{j_2i_2} F_{i_1i_2} + a_{j_2i_2} a_{j_1i_1} F_{i_1i_2} = a_{j_1i_1} a_{j_2i_2} F_{i_1i_2} + a_{j_2i_2} a_{j_1i_1} F_{i_1i_2} = -F_{i_1i_2} - F_{i_1i_2}$$

* The first two paragraphs are based on HAMERMESH (1962) Chapter 10.
Also, \( a_{j_1 j_2} (F_{i_1 i_2} \pm F_{i_2 i_1}) = \pm (F_{j_1 j_2} \pm F_{j_2 j_1}) \). Thus,

\[
(F_{j_1 j_2} \pm F_{j_2 j_1}) = \frac{1}{2} (a_{j_1 j_2} a_{j_2 j_1} \pm a_{j_1 j_2} a_{j_2 j_1}) (F_{i_1 i_2} \pm F_{i_2 i_1}) \text{ and,}
\]

\[
(a_{j_1 i_1} a_{j_2 i_2} \mp a_{j_1 i_2} a_{j_2 i_1}) (F_{i_1 i_2} \mp F_{i_2 i_1}) = 0 \quad (Ab-3)
\]

Thus, the transformation matrix is reduced and the tensor subspaces are invariant. Hence, by taking linear combinations of the \( F_{i_1 i_2} \) and permuting the indices, we are able to decompose the space of second rank tensors into two invariant subspaces. Let the permutation operator \( p \) operate on equation (Ab-2),

\[
(pF')_{i_1 i_2 \cdots i_r} = F_{\rho(1) \rho(2) \cdots \rho(r)} = a_{\rho(1) \rho(2)} a_{\rho(2) \rho(3)} \cdots a_{\rho(r) \rho(1)} F_{\gamma_1 \gamma_2 \cdots \gamma_r}
\]

Now rename \( j_1 \) as \( j_p(1) \), \( j_2 \) as \( j_p(2) \), etc.; then,

\[
(pF')_{i_1 i_2 \cdots i_r} = a_{j_p(1) j_p(2) \cdots j_p(r)} (pF)_{j_1 j_2 \cdots j_r}
\]

But the order of the \( a_{ij} \) here is not important; therefore,

\[
(pF')_{i_1 i_2 \cdots i_r} = a_{i_1 i_2 \cdots i_r} (pF)_{j_1 j_2 \cdots j_r} \quad (Ab-4)
\]

We see that the operator \( p \) commutes with the transformation (Ab-2) in the tensor space; therefore, the tensor space can be decomposed by operating on the \( F_{i_1 \cdots i_r} \) with the \( O^\alpha_{pq} \) for \( S_r \) (the symmetric group on \( r \) objects). Then, in place of the \( n^r \) independent quantities \( F_{i_1 \cdots i_r} \), we deal with a linearly independent set of \( n^r \) quantities selected from the \( O^\alpha_{pq} F_{i_1 \cdots i_r} \). From (Ab-4)

\[
(O^\alpha_{pq} F_{i_1 \cdots i_r}) = a_{i_1 j_1 i_2 j_2 \cdots i_r j_r} (O^\alpha_{pq} F_{j_1 \cdots j_r}) \quad (Ab-5)
\]
Since $j_1, j_2,$ etc., independently go through all values from 1 to $n$, then different terms have the same components of $F$ but in a different order. We can take $F = \alpha' F'$, where $F'$ has a special order; then, $O_{pq}^\alpha \Rightarrow$ yields a linear combination of $O_{pr}^\alpha$ involving various $r$. So, the space of $O_{pq}^\alpha$ with the same $\alpha$ and $p$ is invariant, and the total tensor space has been decomposed into a number of invariant subspaces.

Consider the case of $N=2$, there are two irreducible representations of $S_N$, both of dimension one (cf. Ab-3).

\[
[2]_{11}^{[2]} F_{i_1i_2} = \frac{1}{2} (F_{i_1i_2} + F_{i_2i_1}), \text{ this is a symmetric tensor.}
\]

\[
[2]_{11}^{[2]} F_{i_1i_2} = \frac{1}{2} (F_{i_1i_2} - F_{i_2i_1}), \text{ this is an antisymmetric tensor.}
\]

\[
[2]_{11}^{[2]} a_{i_1i_2} = \frac{1}{2} (a_{i_1i_2} + a_{i_2i_1}) + a_{i_1i_2} a_{i_2i_1}
\]

\[
[2]_{11}^{[2]} a_{i_1i_2} = \frac{1}{2} (a_{i_1i_2} - a_{i_1i_2}) - a_{i_1i_2} a_{i_2i_1}
\]

The 2nd-rank tensor space can be decomposed into two invariant subspaces corresponding to the two irreducible representations of $S_2$.

Consider the case of $N=3$, there are three irreducible representations of $S_N$, two of dimension one and one of dimension two. But now, use a two-dimensional space, i.e., $i = 1$ or 2. Then,

\[
[3]_{11}^{[3]} F_{i_1i_2i_3} = \frac{1}{6} \left[ F_{i_1i_2i_3} + F_{i_2i_3i_1} + F_{i_3i_1i_2} - F_{i_2i_1i_3} - F_{i_1i_3i_2} - F_{i_1i_2i_3} \right]
\]

But, this is zero for any choice of the $i_1, i_2,$ and $i_3$ since $n=2$. Therefore, we can have non-empty subspaces for $N=3$ and $n=2$ only for the other two irreducible representations of $S_N$, i.e., $[3]$ and $[2,1]$ (they do occur). The general result (see Thm. Ab-1 iii) is that for an $n$-dimensional
transformation space we get non-zero tensors after operating with
\( O_{pq}^\alpha \) only for shapes with no more than \( n \) rows. In our above example \( \begin{bmatrix} 1 & 3 \end{bmatrix} \)
has three rows; consequently, we obtained zero when operating on a ten-
sor based on a two-dimensional space.

We now derive some theorems which will be useful in the further
development. The \( n \)-dimensional space transformed by the group \( G \) (with
elements \( a \)) is called the transformation space. The \( n \)-di-
men-sional vector space spanned by the \( r \)-th rank tensors \( \mathcal{L} \) is called the tensor space.
A component of \( \mathcal{L} \) is denoted by \( F_{i_1 i_2 \ldots i_r} \), the \( i_j \) are called indices of
the tensor component and are often suppressed.

Before going on to the next theorem, it will be necessary to go
back to the definition of the \( O_{rq}^\alpha \) (which may be found in the reference
Rutherford (1948)). It will be convenient to discuss the semi-normal
units \( e_{rq}^\alpha \) rather than the orthogonal units \( O_{rq}^\alpha \). Since \( e_{rq}^\alpha \) is propor-
tional to \( O_{rq}^\alpha \), where the proportionality constant is not zero, then
\( O_{rq} F = 0 \) iff \( e_{rq} F = 0 \). Let \( \sigma_r^\alpha \) be the group of all permutations for
which no element of one column of \( S_r^\alpha \) is put into a different column.
Let \( \sigma_r^\alpha \) be the group of all permutations for which no element of one row
of \( S_r^\alpha \) is put into a different row. Let
\[
N_r^\alpha \equiv \sum_{\sigma_r^\alpha} \sigma_r^\alpha \quad \text{and} \quad P_r^\alpha \equiv \sum_{\sigma_r^\alpha} \sigma_r^\alpha
\]
define \( E_{rs}^\alpha \equiv \sum_{\sigma_r^\alpha} \sigma_r^\alpha \sigma_s^\alpha \). Denote as \( S^\alpha \) the tableau obtained from \( S \) by re-
moving \( n \), denote as \( S^{\ast \ast} = S^{2\ast} \) the tableau obtained from \( S^\ast \) by removing
\( n-1 \), etc. Let
\[
e^{(m-1)*} =
e^{(m-1)*}
e^{(m-2)*}
e^{(m-1)*}
e^{(m-2)*}
e^{(m-1)*}
\]
\[
e^{(m-2)*} = \frac{1}{\delta^{(m-2)*}} e_{rr}^{(m-2)*} E_{r \alpha} \delta_x^{(m-1)*}
\]
\[
e^{(m-1)*} = \frac{1}{\delta^{(m-1)*}} e_{rr}^{(m-1)*} E_{r \alpha} \delta_x^{(m-1)*}
\]
\[
e^{(m-2)*} = \frac{1}{\delta^{(m-2)*}} e_{rr}^{(m-2)*} E_{r \alpha} \delta_x^{(m-2)*}
\]
Theorem Ab-1:

(i) If for a given $F$ it is not possible to arrange the first $m$ indices of $F$ to get $F'$ such that when $F'$ is placed in $S_r^\alpha$ to give $S_r^\alpha(F')$ no two identical indices among the first $m+1$ indices of $F'$ occur in the same column of $S_r^\alpha(F')$, then $0_{sr}^\alpha F = 0$ for all $s$.

(ii) If $S_r^\alpha$ has more rows than $F$ has different indices, then $0_{sr}^\alpha F = 0$.

(iii) In particular, for an $n$ dimension transformation space, if $\alpha$ has more than $n$ rows, then $0_{sr}^\alpha F = 0$ for all possible $F$.

Proof:

(i) Let $e_{sr}^{j*}$ operate on $F$. But, $E_{sr}^{j*} = F_{sr}^{j*} N_{sr}^{j*}$ and $N_{sr}^{j*}$ antisymmetrizes each column of $S_r^{j*}$; so, if $E_{sr}^{j*}$ operates on $F'$ and if two identical indices of the first $n-j$ indices of $F'$ are in the same column of $S_r^{j*}$, then $N_{sr}^{j*} F' = 0$. Now, $e_{sr}^{(j-1)*}$ rearranges the first $n-j-1$ indices of $F$; so, if there is no way to rearrange the first $(n-j-1)$ indices of $F$ such that the first $n-j$ indices of $F'$ have all identical indices in different columns of $S_r^{j*}$, then $e_{sr}^{(j-1)*} F = 0$ (for all $s$) and, thus, $e_{sr}^{j*} F = 0$ (for all $s$).

(ii) If $S_r^\alpha$ has more rows than $F$ has different indices, then no rearrangement of $F$ (from applying $e^\alpha$) can avoid having an index appear more than once in the first column of $S_r^\alpha$. Thus, $e_{sr}^\alpha F = 0$ (for all $s$).

(iii) In particular, an $n$-dimensional space $F$ can have at most $n$ different indices; so, $e_{rs}^\alpha F = 0$ for all $\alpha$ with more than $n$ rows. QED

This theorem is not a necessary condition for $0_{rs} F = 0$; the reason is that we are limited to studying the continuous collections of indices of $F$.

* It is convenient in discussing such theorems to place the indices of $F$ in the tableau (say $S_p^\alpha$) in place of the numbers $1...N$; such an array I call an array and denote by $S_p^\alpha(F)$. 
F starting at the first index, so, we might miss some arrangement of
indices of F farther along which in themselves can give rise to $O_{rs}F = 0$
and, indeed, see theorem (Ab-2) for one case. It is possible that
theorems Ab-1 and Ab-2 could provide a necessary condition for $O_{rs}F = 0$;
but this seems likely only for $n = 2$.

We want to be able to select a linearly independent set of

$$0_{rp}^{F} ijk \ldots$$

Theorem Ab-2: If F is written with all identical indices adjacent*,
then (i) is the arrays $S_p(F)$ and $S_q(F)$ are identical, then

$$0_{rp}^{F} = \left\{ \begin{array}{l}
\text{constant (independent of } r) \quad 0_{rq}^{F} \end{array} \right\}$$

and (ii) if the array

$$S_p(F)$$

has identical elements in the same column, then $0_{rp}^{F} = 0$.

Proof:

(i) If $S_p(F) - S_q(F)$ and if $\sigma_{pq}$ is an elementary transposition, say,

$$(k,k+1),$$

then $0_{rp}^{F} = 0_{rp}^{\sigma_{pq}(k,k+1)F} = \sum_{t} U_{\alpha t}^{B_{\sigma_{pq}(k,k+1)}} \sigma_{rp}^{\alpha} \sigma_{tq}^{\beta} F = \sum_{t} U_{\alpha t}^{B_{\sigma_{pq}(k,k+1)}} \sigma_{rp}^{\alpha} F$

But, $U_{pt(k,k+1)} = 0$ unless $t = p$ or $t = q$ since $(k,k+1)$ is an elementary transposition and since $\sigma_{pq} = (k,k+1)$. Thus,

$$0_{rp}^{F} = U_{pp(k,k+1)}^{F} 0_{rp}^{F} + U_{pq(k,k+1)}^{F} 0_{rq}^{F}.$$ But, $U_{pp(k,k+1)} \neq 1$ since if it was then $k$ and $k+1$ would be in the same row of $S_p$ in which case

$S_q = (k,k+1)S_p$ would not be a standard tableau which is contrary to our assumptions. Thus, $0_{rp}^{F} = \text{constant } 0_{rq}^{F}$. Since $F$ has identical

indices adjacent, then the set of identical arrays is such that each

$S_q$ of the set differs from at least one other by an elementary transposition and the chain of tableaux so related is unbroken and connects

* Note the requirement on F; this important theorem is not true for just any F; consequently, tensor components will usually be written with the convention of theorem Ab-2.
to all \( q \) of the set. Thus, the \( \alpha_{pq}^F \) are proportional as \( p \) goes through the elements of the set of identical tableaux.

(ii) If \( S_p \) has a pair of adjacent letters in the same column, say \( k \) and \( k+1 \), and if the \( k \)th and \((k+1)\)th indices of \( F \) are identical, then

\[
\alpha_{pq}^F = \alpha_{pq}^{(k,k+1)F} = \sum_{st \in \text{pt}(k,k+1)} \alpha_{pq}^F = \sum_{pt(k,k+1)} \alpha_{pq}^F . \quad \text{But,}
\]

\[
\alpha_{pt(k,k+1)}^F = - \sum_{pt} \alpha_{pt}^F , \quad \text{Thus, } \alpha_{pq}^F = - \alpha_{pq}^F , \quad \text{and hence, } \alpha_{pq}^F = 0 . \quad \text{If}
\]

\( S_q \) (\( F \)) has identical indices, in the same column, but these positions in \( S_q \) are not occupied by adjacent letters (e.g., \( \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \) for \( F = F_{11234} \)) then among the set of identical arrays there is an \( S_q \) (\( F \)) with adjacent letters in these positions*, viz., \( \begin{array}{c} 1 \\ 3 \\ 2 \\ 4 \end{array} \) in the above example. Thus, using part (i) and first part of (ii) we obtain

\[
\alpha_{pq}^F = 0 .
\]

An example for \( F_{11234} \) and \( \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \)

\[
S_1 = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} ; \quad S_1(F) = \begin{array}{c} 1 \\ 1 \\ 2 \\ 3 \\ 2 \\ 3 \end{array} ; \quad \alpha_{pq}^F \neq 0
\]

\[
S_2 = \begin{array}{c} 1 \\ 2 \\ 4 \\ 3 \\ 5 \end{array} ; \quad S_2(F) = \begin{array}{c} 1 \\ 1 \\ 2 \\ 3 \\ 2 \\ 3 \end{array} ; \quad \alpha_{pq}^F = \text{const } \alpha_{pq}^F ; \quad \text{constant independent of } p
\]

\[
S_3 = \begin{array}{c} 1 \\ 3 \\ 4 \\ 2 \\ 5 \end{array} ; \quad S_3(F) = \begin{array}{c} 1 \\ 2 \\ 2 \\ 3 \\ 1 \\ 3 \end{array} ; \quad \alpha_{pq}^F = 0
\]

\[
S_4 = \begin{array}{c} 1 \\ 2 \\ 5 \\ 3 \\ 4 \end{array} ; \quad S_4(F) = \begin{array}{c} 1 \\ 1 \\ 3 \\ 2 \\ 2 \\ 2 \end{array} ; \quad \alpha_{pq}^F \neq 0
\]

\[
S_5 = \begin{array}{c} 1 \\ 3 \\ 5 \\ 2 \\ 4 \end{array} ; \quad S_5(F) = \begin{array}{c} 1 \\ 2 \\ 3 \\ 1 \\ 2 \end{array} ; \quad \alpha_{pq}^F = 0
\]

* The reason being that all the arrays obtained from \( S_q(F) \) by permuting letters which have identical components in \( F \) must be identical. We select out of all these arrays the set of those which are based on standard tableaux; this set includes all standard tableaux obtained by permuting the above letters. Thus, one tableau has adjacent elements in the same column.
Some theorems on transformations of symmetry tensors follow.

**Theorem Ab-3.** The tensor of symmetry \([i^n]\) transforms as \(\det a\)  

Proof:  
\[
\begin{align*}
\alpha(0,1_{F_{11}} F_1 n) &= \sum_{\alpha, \beta, \gamma} \partial_{\alpha(0)} \partial_{\beta(1)} \partial_{\gamma(n)} \alpha_{ll} \beta_{l1} \gamma_{l1} \varepsilon_{l1} F_{11} F_{l1} F_{l1} \varepsilon_{l1} F_1 n \\
0_{11} F_{ij} \ldots k &= 0 \text{ unless } i,j, \text{ etc. are all different}; \text{ thus,} \\
\xi_{i,j, \ldots k} &= \xi_{1,2, \ldots n} \\
\text{Let } F_{ij \ldots k} &= \gamma F_{12 \ldots n} ; \\
\text{then, } 0_{11} F_{12 \ldots n} &= \beta_{12} 0_{11} F_{12 \ldots n} ; \\
\text{thus, } a(0_{11} F_{12 \ldots n}) &= (\sum_{\alpha(0)} \partial_{\alpha(0)} \partial_{\beta(1)} \partial_{\gamma(n)} \beta_{12} F_{12 \ldots n} (\det a) F_{12 \ldots n} = (\det a) a(0_{11} F_{12 \ldots n}) .
\end{align*}
\]

**Q.E.D.**  

**Theorem Ab-4:** The tensor of symmetry \([2^n]\) transforms as \((\det a)^2\)  

Proof:  
\[
\begin{align*}
\alpha(0_{rf} F_{11} n) &= \sum_{\xi_{ij} \gamma_{mn}} \partial_{\xi_{ij}} \partial_{\gamma_{mn}} \gamma_{mn} (\det a)^2 F_{11} n \\
0_{rf} F &= 0 \text{ unless the first } n \text{ indices are all different and} \\
\text{the last } n \text{ indices are all different.} \\
\text{Let } F_{ij \ldots m} &= \gamma \gamma^{''} F_{11} n , \text{ where } \gamma^{''} \text{ and } \gamma^{'''} \text{ operate on the first} \\
\text{and last } n \text{ components of } F \text{ respectively. Then,} \\
0_{rf} F_{ij \ldots m} &= \beta_{12} \beta_{12} 0_{rf} F_{11} n \\
\sum_{\xi_{ij} \gamma_{mn}} \partial_{\xi_{ij}} \partial_{\gamma_{mn}} \gamma_{mn} (\det a)^2 F_{11} n &= (\det a)^2 ; \text{ thus,} \\
a(0_{rf} F_{11} n) &= (\det a)^2 (0_{rf} F_{11} n) . \text{ Let} \\
F_{11} n &= \gamma F_{11} n , \text{ then} \\
0_{rf} F_{11} n &= \sum_{j} U_{fj} \gamma r_{rf} F_{11} n \gamma - U_{fl} \gamma r_{r1} F_{11} n \gamma \\
\text{Hence, } a(0_{rf} F_{11} n) &= (\det a)^2 (0_{rf} F_{11} n) . \\
\text{Since } U_{fl} \gamma \neq 0, \text{ any non-zero } 0_{rf} F_{ij \ldots k} \text{ can be expressed in terms} \\
of 0_{rf} F_{11} n . \text{ Q.E.D.}
\end{align*}
\]
Theorem Ab-5: The tensors of symmetry \([I^{m-n}]\) transform as \((a^*)\text{ det } \overline{a}\).

(HAMERMESH (1962) p. 391 states that \([I^{m-n}]\) is equivalent to \([I]\) if \(\text{det } a = 1\); this statement is true only if \(\text{det } a = 1\) and \(a\) is equivalent to \(a^t\); if \(a\) is orthogonal, i.e., for \(\text{SO}(m)\), the statement is true; it is not true in general, e.g., for \(\text{SU}(3)\)).

Proof: \(a(0_{11} F_{i+1,i+1,...,i-1}) = \sum_{gjlm} \partial g_{i+l+1} \partial k_{i+2} \cdots \partial m_{i-1} O_{11} F_{gjlm}\).

But, \(0_{11} F_{jk...m} = 0\) unless \(jk...m\) are all different; therefore**

\([I^{m-n}]\)

\(0_{11} F_{jk...m} = \epsilon_{pjk...m}0_{p+1,...,p-1}\) if \(p\) is the letter missing from \(jk...m\). So,

\(a(0_{11} F_{i+1,...,i-1}) = \{\sum_{gjlm} \epsilon_{pjkm} \partial g_{i+l+1} \partial k_{i+2} \cdots \partial m_{i-1}\} O_{11} F_{p+1,...,p-1}\).

Let \(A_{qi} = \sum_{gjlm} \epsilon_{pjkm} \partial g_{i+l+1} \partial k_{i+2} \cdots \partial m_{i-1}\) (this is the cofactor of \(\text{det } a\)).

Then, \(a_{qi} A_{pi} = \epsilon_{pjkm} a_q a_i a_{i+1} \cdots a_{m-1} = S_{mp} \text{ det } a\) since if \(q \neq p, q\) is one of the other letters \(j, k, ..., m\) and the determinant has two equal rows. Thus \(A_{qi} = a_{qi}\) and \(A_{pq} = a_{pq}(\text{ det } a)\)

QED

Note that for \(\text{SO}(n)***\) the tensors \([I^{m-n}]\) and \([I]\) transform in the same way. For \(\text{SU}(n)\), \([I^{m-n}]\) transforms as \(a^*\); therefore, for \(\text{SU}(2)\), \(a\) and \(a^*\) must be equivalent (this certainly provides insight into the reason for an important difference between \(\text{SU}(2)\) and \(\text{SU}(3)\), namely, that \(a\) is equivalent to \(a^*\) for \(\text{SU}(2)\) but \(a\) is not equivalent

---

* \(a^t, a^\dagger, a^*\), and \(a^\dagger\) denote the transpose, complex conjugate, and Hermitian conjugate of \(a\), respectively.

** \(\epsilon_{pq...m}\) is zero unless all subscripts are different, it changes sign under transposition, and it is positive when all subscripts are in order.

*** \(\text{O}(n)\) is the group of orthogonal transformations in \(n\) dimensions (i.e., if \(a\) \(\in \text{O}(n)\) then \(a^t = a^{-1}\)). The \(S\) in \(\text{SO}(n)\) indicates that we take the subgroup of \(\text{O}(n)\) with determinant equal to \(+1\). \(\text{U}(n)\) is the group of unitary transformation in \(n\) dimensions (i.e., if \(a\) \(\in \text{U}(n)\) then \(a^\dagger = a^{-1}\)).
Theorem Ab-6: The tensor of symmetry $[1^{m-2}]$ transforms under $a$ as $a^{-1}$ times the way $[1^2]$ transforms under $a^{-1}$. (Similarly we can prove the corresponding theorem for $[1^{m-3}]$ and $[1^3]$, etc.).

Proof: Proceeding as in Thm Ab-5 we obtain

$$a(0^{[m-2]}_{11} F_{i+1,...,j-1,j+1,...,i-1}) = \sum_{p,q} A_{pqij} 0^{[m-2]}_{ij} F_{p+1,...,q-1,q+1,...,p-1}$$

where $A_{pqij} = \sum_{m,m',m'\prime} \partial_{m_i m_j} \partial_{m_{p-1} m_{q-1}} \partial_{m_{i+1} m_{j+1}}$.

Then, $a_i a_j A_{pqij} = (\det a)(S_{sp} S_{tq} - S_{sq} S_{tp})$

Thus, $A_{pquv} = a^{-1}_m a^{-1}_v a^{-1}_u a^{-1}_p$.

Theorem Ab-7:* $0^{[m_1,m_2,...,m_8]}_{rf}$ transforms as $(\det a)$ times the way

$$0^{[m_1-m_8,m_2-m_8,...,m_8-m_8]}_{rf}$$

transforms.

$$0^{[m_1-m_8,m_2-m_8,...,m_8-m_8]}_{rf}$$

under a the same way that

$$0^{[m_1-m_8,m_2-m_8,...,m_8-m_8]}_{rf}$$

transforms under $a^{-1}$.

Proof:

$$a(0^{[m_1,m_2,...,m_8]}_{rf} F_{i_1,i_2,...,i_8}) = \sum_{p,q} \partial_{m_i} \partial_{m_j} \partial_{m_{p-1}} \partial_{m_{q-1}} 0^{[m_1,m_2,...,m_8]}_{rf} F_{i_1,i_2,...,i_8}$$

But, the first $m_8$ sets of $\mathcal{S}$ indices in $F_{i_1,...,i_8}$ must each contain each letters $1,...,\mathcal{S}$ once; the next $m_{8-1} - m_8$ sets must each contain all but one of the letters $1,...,\mathcal{S}$ once; etc. Next, we let

$F_{i_1,...,i_8} = \mathcal{V} F_{i_1,...,i_8}$, where the first $m_8$ sets of $\mathcal{S}$ indices of $F'$ are in order $1,...,\mathcal{S}$; the next $m_{8-1} - m_8$ sets of $\mathcal{S} - 1$ letters are in order starting from the missing letter (e.g., if $p$ is missing, then $p+1, p+2, ... , 1, ... p-1$); the next $m_8 - m_{8-1}$ sets of $\mathcal{S} - 2$ $\mathcal{S}$-1 letters.

* Here I let the dimension of the transformation space be $\mathcal{S}$.
letters start with the first letter above the lower of the two
missing letters and the letters are in order (e.g., if \( p \) and \( q \) are
missing and \( p < q \), then \( p+1, p+2, \ldots, q-1, q+1, \ldots, \Xi, 1, \ldots, p-1 \)),
etc. Then, \( 0_{rr}F = 0_{rr}FF' = \sum_{1,4,7} F_{1,4,7} \). Hence, each of the \( F \).

of the first \( m_s \) terms separates into a

\[
\sum_{1,4,7} a_1 a_2 \cdots a_7 = \det a \text{ just as in THM Ab-3. Thus,}
\]

\[
0_{rr} \left[ m_1, m_2, \ldots, m_s \right] F \text{ transforms the same way as } (\det a)^{m_s} \text{ times}
\]

\[
0_{rr} \left[ m_1, m_2, \ldots, m_s, \ldots, m_{s-1}, m_s \right] F. \text{ Now, consider } 0_{rr} \left[ m_1, m_2, \ldots, m_s, \ldots, m_{s-1}, m_s \right] F,
\]

each column transforms separately and the column of length \( [1^{s-1}] \)
transforms like \( [1] \) if \( a \) is replaced by \( a^{-1} \), and the column of
length \( [1^{s-2}] \) transforms like \( [1^2] \) if \( a \) is replaced by \( a^{-1} \), etc.,
as in THM Ab-6. Thus, \( 0_{rr} \left[ m_1, m_2, \ldots, m_s, \ldots, m_{s-1}, m_s \right] F \) transforms like
\[
0_{rr} \left[ m_1, m_2, \ldots, m_s, \ldots, m_{s-1}, m_s \right] F \text{ if } a \text{ is replaced by } a^{-1}. \]

**Theorem Ab-8:** If there is only one linearly independent tensor \( 0_{rj} \alpha \)
(r fixed and \( j \) varying) of symmetry \( \alpha = [m_1, m_2, \ldots, m_s] \),
then THM Ab-7 applies to any of the non-zero \( 0_{rj} \alpha \).

**Proof:** Trivial.

It may be true that THM Ab-7 applies to any non-zero \( 0_{rj} \alpha \) even if
there is more than one linearly independent \( 0_{rj} \alpha \) but I have not been
able to prove it and have not needed it (and I doubt it).

We have not yet put any requirements on the group of transforma-
tions \( G \) of the \( n \)-dimensional transformation space. If the group is
\( \text{GL}(n) \), the group of all linear transformation of an \( n \)-dimensional space,
then the invariant subspaces of the \( n \)-dimensional tensor space obtained
by applying the \( 0_{rj} \alpha \) are indecomposable for fixed \( \alpha \) and \( \rho \) with \( q \) varying.
For some subgroups of \(\text{GL}(n)\), e.g., \(\text{SU}(n)\), (see NAPIKETSHIP (1962), p. 390) the invariant subspaces are still indecomposable, whereas for others, e.g., \(\text{SO}(n)\), the invariant subspaces are in general decomposable. For example, if \(G = \text{SO}(3)\) (the three-dimensional proper rotation group) the six-dimensional space of symmetric tensors can be further decomposed into a one-dimensional space which transforms like a scalar (i.e., invariant) and a five-dimensional indecomposable space which transforms like the spherical harmonics with \(L = 2\); the three-dimensional space of the antisymmetric tensors is indecomposable.

We are particularly interested in (complex) two-dimensional spaces with \(\text{SU}(2)\) as the transformation group. We choose as the basis vectors \(\alpha\) and \(\beta\), the eigenfunctions of \(S_z\). A typical tensor component (for an \(N\)-electron problem) might be \(\gamma(1...N) = \alpha(1)\beta(2)\alpha(3)\beta(4)\alpha(5)\beta(6)\alpha(7)\beta(8)\ldots\alpha(N)\).

From THM Ab-1 iii we can consider only one and two rowed shapes since \(O_{ri}^\alpha \gamma = 0\) if \(\alpha\) has more than two rows. From THM Ab-2 we see that the \(O_{ri}^\alpha \gamma\) for different \(i\) and the same \(\alpha\) and \(r\) are all proportional. So, although for general \(\gamma\) we might have \(f^\alpha\) linearly independent sets of \(O_{ri}^\alpha \gamma\) (with \(r\) varying from 1 to \(f^\alpha\)), there is at most one. Thus, we choose the convention of considering only the components \(\gamma\) with all of the \(\alpha\) written first and all of the \(\beta\) written last; we write \(O_{ri}^\alpha \gamma\) picking \(i = 1\) since \(O_{ri}^\alpha \gamma\) is always non-zero if \(O_{ri}^\alpha \gamma\) is non-zero for any \(i\). The shape has two rows; thus, we write the partition \(\alpha\) as \(\alpha = [\frac{N}{2} + S, \frac{N}{2} - S]\), where \(\frac{N}{2} + S\) and \(\frac{N}{2} - S\) are integers. If \(N\) is even, there is one partition \(\alpha\) for each integral value of \(S\) from 0 to \(\frac{N}{2}\); if \(N\) is odd, there is one partition \(\alpha\) for each half-integral \(*\) value.

\(*\) We use the physicists' usual convention that half-integral implies what might be better termed half-odd-integral.
of $S$ from $\frac{1}{2}$ to $\frac{N}{2}$. The number of linearly independent $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}$ for the different $\mathcal{K}$ (r and $\sigma$ fixed) is 2$S$+1 (use THM Ab-2). Because of our convention on $\mathcal{K}$ these different $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}$ correspond to different $S_z$; and of the 2$S$+1 different $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}$ there is one for each $S_z$ from $-S$ to $+S$ spaced at integral intervals. Also, $S^+_0 \mathcal{O}_{r_1} \mathcal{K}^S = 0$ and $S^-_0 \mathcal{O}_{r_1} \mathcal{K}^-_S = 0$ (use THM Ab-2). The total number of linearly independent $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}$ for different $\mathcal{K}$, $\sigma$, and $r$ is $Z^N$; and $r^{\sigma}$ is just the number obtained from the branching diagram for the number of different states of the same spin. Since the spaces of different $\sigma$ and $r$ are each indecomposable, we have shown that the set of states $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}$ for $\sigma - \left[ \frac{-N}{2}, \frac{N}{2} - \epsilon \right]$ span the same space as the set of states of spin $S$ for $N$ particles each of spin $\frac{1}{2}$. Since there is a one-to-one correspondence between $S$ and $\sigma$, we use them interchangeably; when I write $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}^S$, I mean that $\mathcal{K}$ has $S_z = S$ for the $S$ corresponding to $\sigma$. There is only one $\mathcal{O}_{rj}^{\sigma} \mathcal{K}^S \neq 0$ for varying $j$ and fixed $r$ and $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}^S$ is an eigenfunction of $S^2$ with $S_z = S$ (since $S^+_0 \mathcal{O}_{r_1} \mathcal{K}^S = 0$). $S^-_0 \mathcal{O}_{r_1} \mathcal{K}^S$ gives a linear combination of $\mathcal{O}_{rj}^{\sigma} \mathcal{K}^S_{-1}$ for varying $j$ and fixed $r$ (i.e., $S^-_0 \mathcal{K}^S = \mathcal{K}^S_{-1} + \mathcal{K}^S_{-1} + \cdots + \mathcal{K}^S_{-1}$, where only $\mathcal{K}^S_{-1}$ has the standard order; thus, $S^-_0 \mathcal{K}^S = (\epsilon + \epsilon' + \cdots + \epsilon^n) \mathcal{K}^S_{-1}$ where, e.g., $\epsilon^{n-1} \mathcal{K}^S_{-1} = \mathcal{K}^S_{-1}$, and, e.g., $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}^S_{-1} = \frac{1}{Z} \sum \mathcal{O}_{\sigma}^{\sigma} \mathcal{K}^S_{-1}$) But, all of these are proportional; so, by merely multiplying by the appropriate (normalization) constant the $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}^S_z$ are all connected by $S^\pm$. Each $\mathcal{O}_{r_1}^{\sigma} \mathcal{K}^S_z$ is an eigenfunction of $S^2$. 
Some examples:

For $N=5$, $\chi=\alpha\beta\beta\beta\beta\beta$ ($S_z=-\frac{1}{2}$), and 

\[
S_1 = \alpha \beta \beta \beta \\
S_2 = \alpha \beta \beta \\
S_3 = \alpha \beta \beta \\
S_4 = \alpha \beta \beta \\
S_5 = \alpha \beta \beta
\]

$p_1 F = \text{const}$, $p_2 F = \text{const}$, $p_4 F = 0$.

For $N=6$

\[
S = 3 \\
S = 2 \\
S = 1 \\
S = 0
\]

For $N=2$

\[
S = 1 \\
S = 0
\]

We started with a $2^2 = 4$ dimensional tensor space ($\alpha\alpha$, $\alpha\beta$, $\beta\alpha$, $\beta\beta$ being a set of independent quantities), and we have transformed to a new set of basis vectors $0_{11}^{[2]} \alpha\alpha$, $0_{11}^{[2]} \alpha\beta$, $0_{11}^{[2]} \beta\alpha$, and $0_{11}^{[2]} \beta\beta$ which are eigenfunctions of $S^2$ and $S_z$. 
Ac: The General Function Which Satisfies Pauli's Principle and Is an Eigenfunction of $S^2$

Consider a product of $N$ one-electron wave functions

$$\Psi = \Psi_i(r_1) \Psi_j(r_2) \cdots \Psi_m(r_N),$$

and let each $\Psi_i(r) = \phi_i(r) \sigma_z \phi_i(r)$, where $\sigma_z = \alpha$ or $\beta$. Then $\Psi_i(l_1, \cdots, N) = \phi_i(r_1) \phi_i(r_2) \cdots \phi_i(r_N) \sigma_z \phi_i(r_1) \cdots \sigma_z \phi_i(r_N)$. Let

$$\Phi(l_1, \cdots, N; \lambda_1, \lambda_2, \cdots, \lambda_N) \equiv \phi_i(r_1) \phi_i(r_2) \cdots \phi_i(r_N) \sigma_z \phi_i(r_1) \cdots \sigma_z \phi_i(r_N)$$

Then $\Phi$ involves only spatial coordinates, and $\Psi$ involves only spin coordinates. Let the operator $0^\alpha_{pq}$ operate only on the spatial coordinates and $\omega^\alpha_{pq}$ operate only on the spin coordinates. Using Aa-10:

$$\left< \alpha_\Phi \Phi \right> \Phi = \sum_{r, \lambda} \left< \alpha_{\Phi r \lambda} \right> \left( \omega^\alpha_{\Phi \lambda} \Psi \right)$$

(Ac-1)

But, from section Ab $\omega^\alpha_{\Phi \lambda} \Psi$ is an eigenfunction of $S^2$. Therefore, since $S^2$ does not operate on $\left< \alpha_{\Phi r \lambda} \right>$, then $\alpha^\alpha_{\Phi \lambda} \Phi$ is an eigenfunction of $S^2$.

From Aa-13 $\alpha^\alpha_{\Phi \lambda} \Phi = \sum_{\beta} \alpha^\lambda_{\Phi \lambda} \Phi$; thus,

$$\beta \left< \alpha^\lambda_{\Phi \lambda} \Phi \right> = \sum_{\beta} \left< \alpha^\lambda_{\Phi \lambda} \Phi \right> \left( \beta \alpha^\lambda_{\Phi \lambda} \Phi \right)$$

(Ac-2)

Therefore, $\alpha^\alpha_{\Phi \lambda} \Phi$ satisfies Pauli's principle and is an eigenfunction of $S^2$. Note that since $\alpha$ is limited to two rows then $\lambda$ is limited to two columns.

A useful theorem follows.

Theorem Ac-1:

1. $\left< \alpha^\alpha_{\Phi \lambda} \Phi \middle| \Phi \right> = \left< \Phi \middle| \alpha^\alpha_{\Phi \lambda} \Phi \right>$

2. $\left< \alpha^\alpha_{\Phi \lambda} \Phi \middle| 0^\beta_{\Phi \rho} \Phi \right> = \sum_{\beta} \left< \alpha^\lambda_{\Phi \lambda} \Phi \middle| 0^\rho_{\Phi \rho} \Phi \right>$

* For Bose particles, $\alpha^\lambda_{\Phi \lambda} \Phi$; thus, we would use $\left< \alpha_{\Phi \lambda} \Phi \middle| 0^\rho_{\Phi \rho} \Phi \right>$ to get an eigenfunction of $S^2$ if the spin of the particles is $\frac{1}{2}$. It is believed that such particles do not exist.
Proof:

(i) let \( (dx) = dx_1 dx_2 \ldots dx_N \):

\[
\langle \Omega \Phi_1 \Phi_2 \rangle = \int (\Omega^{\dagger} \Phi_2 \Phi_2^\dagger (dx) = \frac{1}{\delta^N} \sum_{\lambda} U_{\lambda \lambda} \langle \phi_\lambda \Phi_2 \Phi_2 (dx) \rangle
\]

But the integral is some number and is independent of the spatial coordinates since they have been integrated over; therefore, the integral is unchanged by a permutation of the spatial coordinates.

\[
\langle \Phi_1 \Phi_2 \rangle = \int (\Phi_2 \Phi_2^\dagger (dx) = \int (\Phi_2 \Phi_2^\dagger (dx) = \int (\Phi_2 \Phi_2^\dagger (dx) = \int (\Phi_2 \Phi_2^\dagger (dx)
\]

If \( \Phi = \Phi \), we obtain

\[
\langle \Omega \Phi_1 \Phi_2 \rangle = \langle \Phi_1 \Phi_2 \rangle
\]

\[
\langle \Omega \Phi_1 \Phi_2 \rangle = \frac{1}{\delta^N} \sum_{\lambda} U_{\lambda \lambda} \langle \phi_\lambda \Phi_2 \Phi_2 (dx) \rangle = \langle \Phi_1 \Phi_2 \rangle
\]

(ii) Using part i

\[
\langle \Omega \Phi_1 \Phi_2 \rangle = \langle \Phi_1 \phi_\lambda \Phi_2 \rangle = \sum_{\lambda} \langle \Phi_1 \phi_\lambda \Phi_2 \rangle = \sum_{\lambda} \langle \Phi_1 \phi_\lambda \Phi_2 \rangle = \langle \Phi_1 \Phi_2 \rangle
\]

\[
\langle \Omega \Phi_1 \Phi_2 \rangle = \frac{1}{\delta^N} \sum_{\lambda} U_{\lambda \lambda} \langle \phi_\lambda \Phi_2 \Phi_2 (dx) \rangle = \langle \Phi_1 \Phi_2 \rangle
\]

\[
\langle \Omega \Phi_1 \Phi_2 \rangle = \frac{1}{\delta^N} \sum_{\lambda} U_{\lambda \lambda} \langle \phi_\lambda \Phi_2 \Phi_2 (dx) \rangle
\]

We cannot go too much further without becoming becoming a little more specific about the \( \Phi_1 \).

Another method of obtaining N-electron wave functions which are eigenfunctions of \( S^2 \) and simultaneously satisfy Pauli's principle has been obtained independently by Kotani et al (1955). By a quite different approach than that used here Kotani obtains (K p.17)

\[
\Phi_{S,M}^{(m)} = \frac{1}{\sqrt{S}} \sum_{\lambda} \psi_{S,M,i}^{(m)} \Omega_{S,M,i} \Phi_{S,M,i}\]

where \( \psi_{S,M,i}^{(m)} = \sqrt{\frac{1}{\delta^N}} \sum_{\lambda} U_{\lambda \lambda}^{(m)} (P) \cdot \psi^{\lambda} \)

( \( \psi^\lambda \) is a product of N orbitals) and \( \Omega_{S,M,i} \) is one of \( f_S \) orthonormal spin functions which is an eigenfunction of \( S^2 \) with eigenvalue \( S(S+1) \)

(the \( \Omega \) are obtained by vector coupling the N spins in the \( f_S \) different ways given by the branching diagram). The \( \Omega \) transform as (K p.7)
\( p_{a, \alpha, \beta} \) in order that they satisfy Pauli's principle (K p. 14). We must compare the \( \Phi_{\alpha}^{(1)} \) to
\[
\Phi_{\alpha} = \sum \psi_{\alpha}(\rho) \Phi_{\alpha}(\rho') \psi_{\alpha}(\rho'')
\]
Here
\[
\frac{\psi_{\alpha}(\rho)}{\rho} = \frac{1}{2} \sum_{\beta, \gamma} \psi_{\alpha}^\beta \psi_{\alpha}^\gamma \psi_{\alpha}^{\beta' \gamma'} \psi_{\alpha}^{\beta' \gamma'}
\]
We may identify \( \psi_{\alpha}(\rho) \) and
\[
\sum_{\beta, \gamma} \psi_{\alpha}(\rho) \Phi_{\alpha}(\rho') \psi_{\alpha}(\rho'')
\]
in which case \( \psi_{\alpha}(\rho) = \psi_{\alpha}(\rho') \). Also (K p. 14)
\[
\psi_{\alpha}(\rho) = \sum \psi_{\alpha}(\rho) \psi_{\alpha}(\rho') \psi_{\alpha}(\rho'')
\]
However,
\[
\frac{\psi_{\alpha}(\rho)}{\rho} = \sum \psi_{\alpha}(\rho) \psi_{\alpha}(\rho') \psi_{\alpha}(\rho'')
\]
But from appendix I \( \sum \psi_{\alpha}(\rho) \psi_{\alpha}(\rho') = \frac{\rho}{\rho} \). Thus,
\[
\psi_{\alpha}(\rho) = \sum \psi_{\alpha}(\rho) \psi_{\alpha}(\rho')
\]
Thus, we see that the Kotani function and the \( \Phi_{\alpha}^{(1)} \)
are equivalent. In fact Kotani derives the equivalent of Theorem
Ba-1 and Ba-2 and appendix A (K p. 22 and K p. 23-26); he also obtains
the equivalent of equations Be-3 and 4 (K p. 27). In general Kotani's
proofs proceed somewhat differently than those discussed in this report.

An advantage of the \( \Phi_{\alpha}^{(1)} \) form is the close relationship to the Young
tableaux permitting a pictorial visualization of the relationship be-
tween and the properties of the \( \Phi_{\alpha}^{(1)} \); however, the \( \Phi_{\alpha}^{(1)} \)
can be visualized using the branching diagrams and, indeed, there is a direct
relation between the branching diagrams and the standard Young tableaux.

Another method of obtaining N-electron wave functions which are
eigenfunctions of \( S^2 \) and simultaneously satisfy Pauli's principle is
due to L"ODIN (1955c). L"od"din starts with a Slater determinant and
then operates with a projection operator to select the component of a
specified total spin. The projection operator is
\[
\sum_{k=0}^{N} \frac{\rho_k}{\rho} \left( \frac{\rho_k}{\rho} \right)
\]
where the product is over all \( k \) from 0 to \( \frac{N}{2} \) (except \( \frac{N}{2} \)) and
\[
\frac{\rho_k}{\rho} = \frac{\rho_{k}^{\alpha}}{\rho^{\alpha}} \frac{\rho_{k}^{\beta}}{\rho^{\beta}}
\]
The problem is that the result of operating with \( \rho_{k=0}^{N} \) on a Slater determinant includes, in general, a
sum over very many Slater determinants many of which are in general, linearly dependent and, in addition, the result of $2^F$ operating on different Slater determinants with the same orbitals results in very many linearly dependent and nonorthogonal terms. Thus, the use of this scheme is quite complex; however, Löwdin has succeeded in obtaining a general expression for what might be the most important of the many different functions with a given total spin.
Section B

Same Orbitals for Different Spin -- the Hartree-Fock Method

Introduction

Before proceeding to the DOWS approach of solving for the ground state energy and wave function of a molecule, we will develop the Hartree-Fock method using the $\mathcal{G}_i^\alpha$ operators. Throughout section B the SODS restriction is employed since the Hartree-Fock method requires SODS. Although it is true that the best Slater determinant approximation to the ground wave function of a molecule is usually an eigenfunction of $S^2$ if the SODS restriction is used, it is not true that every Slater determinant is an eigenfunction of $S^2$ even if the SODS restriction is applied. Thus, even for the SODS restricted case it is necessary to use the $\mathcal{G}_i^\alpha$ operators in general treatments. For this reason we carry out a more complete development than is required for constructing the Hartree-Fock method. Sections Ba and Bb are preliminaries for the later sections of B. The orthogonality properties of the $\mathcal{G}_i^\alpha \mathcal{P}$ are developed in section Ba; the first- and second-order density matrices which are used for determining expectation values are determined in section Bb. Section Bc contains a simple application of the results of section Bb to obtain a generalized Hund's rule. The Hartree-Fock method is derived and discussed in section Bd. The first-order perturbed wave function (with the Hartree-Fock wave function as the zero-order state) is obtained in section Be; with the first-order wave function, the second- and third-order corrections to the energy can be determined. In section Bf the procedure involved in combining the $\mathcal{G}_i^\alpha \mathcal{P}$ to obtain proper symmetry functions of the spatial symmetry group are discussed.
for the proper rotation group, $SO(3)$. 
Ba: Orthogonality and Normalization of the $G^\alpha_1 \Phi \chi$

First some useful theorems.

Theorem Ba-1: If $F_{1,2} = \Phi = \Phi_1^{(1)} \Phi_2^{(2)} \Phi_3^{(3)} \cdots$ (where the set of $\Phi_i$ is orthonormal) is written with all identical indices adjacent and if $S_p(\Phi) \neq S_q(\Phi)$, then $0_{rp}^\alpha$ and $0_{rq}^\alpha$ are orthogonal.

Proof: Using THM Ac-1

$$\langle 0_{rp}^\alpha | 0_{rq}^\alpha \rangle = \langle \Phi | 0_{pq}^\alpha | \Phi \rangle = \sum_{\gamma} \alpha^\gamma \sum_{pq}^\gamma \langle \Phi | \gamma | \Phi \rangle,$$

but $\langle \Phi | \gamma | \Phi \rangle \neq 0$ implies that $\gamma$ only permutes sets of identical indices, and if for convenience we consider that only one set is permuted, then $\gamma$ can be expressed in terms of elementary transpositions involving only the letters corresponding to the pertinent set of indices of $\Phi$.

Hence, $\gamma S_p(\Phi) = S_q(\Phi)$; thus, $U_{pq}^\alpha \neq 0$ only if $S_p(\Phi) = S_q(\Phi)$,

but we have assumed $S_p(\Phi) \neq S_q(\Phi)$. If more than one set of identical indices is permuted then each one is treated the same way. QED

From THM Ah-7 if $S_p(\Phi) = S_q(\Phi)$ then $0_{rp}^\alpha$ and $0_{rq}^\alpha$ are proportional so that $G_p^\alpha \Phi \chi$ and $G_q^\alpha \Phi \chi$ are proportional. We want to select the set of all of the different $G_p^\alpha \Phi \chi$; so, a method of deciding which ones are proportional is necessary.

Theorem Ba-2: If $\Phi$ is written with all $\Phi_i$ which occur twice (I call these doubly-occupied orbitals) placed together and first, and the $\Phi_i$ occurring once (I call these singly-occupied orbitals) placed last (i.e.,

$$\Phi = \Phi_1^{(1)} \Phi_2^{(2)} \Phi_3^{(3)} \cdots \Phi_m^{(2m-1)} \Phi_m^{(2m)} \Phi_{2m+1}^{(2m+1)} \Phi_{2m+2}^{(2m+2)} \cdots \Phi_N^{(N)}$$

where the set $i_1, i_2, \ldots, i_n$ is ordered and the set $i_{2n+1}, \ldots$ is ordered, and $n$ is the number of doubly-occupied orbitals) the set of non-zero $G_p^\alpha \Phi \chi$ is linearly independent and, a fortiori, orthogonal. (This theorem depends on $\alpha$ being one- or two-columned.)
PROOF: Let $T_{\Phi}$ be the group of permutations leaving $\Phi$ invariant (it is an Abelian group of $2^n$ elements); then,

$$<\Phi|\Phi> = \begin{cases} 0 & \text{if } \varepsilon \notin T_{\Phi} \\ 1 & \text{if } \varepsilon \in T_{\Phi} \end{cases}$$

Let $U_{pq} = \sum_{p}^{q} q$ since both $S_{P}$ and $S_{Q}$ are of the form $\varepsilon$. We have

$$<\omega_{p}^{\alpha} \Phi|\omega_{q}^{\beta} \Phi> = <\Phi|\omega_{p}^{\alpha} \Phi> - \sum_{p}^{2n} S_{pq}$$

(Ba-1)

But, using THM Ab-1

$$<\omega_{p}^{\alpha} \Phi, X, \omega_{q}^{\beta} \Phi> = \sum_{p}^{2n} \sum_{q}^{2n} S_{pq} <\Phi|\omega_{p}^{\alpha} \Phi> <X|\omega_{q}^{\beta} \Phi> = \sum_{p}^{2n} S_{pq} \sum_{q}^{2n} S_{pq}$$

and therefore, if $\Phi_{\alpha}$ and $\Phi_{\beta}$ are in the standard order,

$$<\omega_{p}^{\alpha} \Phi, X, \omega_{q}^{\beta} \Phi> = \sum_{p}^{2n} \sum_{q}^{2n} S_{pq} S_{pq}$$

(Ba-2)

QED

When considering the spin tensor space or the spatial tensor space by itself, it is permissible to select specific types of tensor components (i.e., ordered as in THM Ab-2) on which the $\omega_{pq}^{\alpha}$ must operate since we are interested in the number of different non-empty spaces, and we have found a convention which does not miss any spaces. We must be careful in making a convention for $\Phi_{\alpha} \Phi_{\beta} \chi$ that we do not miss any non-zero $G_{ij}^{\alpha} \chi$. In particular we cannot arbitrarily select the convention for $\Phi_{\alpha}$ and $\chi$ separately (e.g., if $\alpha = [23]$, $\Phi = \Phi_{1} \Phi_{2} \Phi_{3}$, and $\chi = \alpha \beta \gamma \rho \sigma \tau$ then $\omega_{pq}^{\alpha} \Phi \neq 0$ only for $q = 1$ and $\omega_{pq}^{\alpha} \chi = 0$ whereas if $\chi = \alpha \beta \gamma \rho \sigma \tau \beta$ then $\omega_{pq}^{\alpha} \chi \neq 0$). For SDS it is most convenient to use the convention of THM Ab-2 for $\Phi_{\alpha}$ and then to use a convention for $\chi$ which will give a non-zero $\omega_{pq}^{\alpha} \chi$ for every $q$ which gives a non-zero $\omega_{pq}^{\alpha} \Phi$ regardless of what $\Phi$ is. The convention for $\chi$ is that the first $N$ $2S$ positions of $\chi$ have alternating $\alpha$ and $\beta$ then the remaining $2S$ positions have the remaining
Consider the case of $M_S = S$; then (by THM Ab-2)

$$\omega_{pq} \mathcal{N}_f = S_{pq} \omega_{pq} \mathcal{N}_f.$$

Let $\mathcal{N} = \mathcal{N}_f$, then (using Aa-4)

$$\omega_{pq} \mathcal{N} = U_{pq} \mathcal{N}_f$$

Thus,

$$\frac{\langle X_{pq} \mathcal{N}_f \rangle}{\langle X_{pq} \mathcal{N}_f \rangle} = \frac{\langle X_{pq} \mathcal{N}_f \rangle}{\langle X_{pq} \mathcal{N}_f \rangle} = \sqrt{\langle X_{pq} \mathcal{N}_f \rangle} \langle X_{pq} \mathcal{N}_f \rangle}$$

But, $\langle X_{pq} \mathcal{N}_f \rangle \neq 0$ only for $e \in T\mathcal{N}_f$ (i.e., $e$ which do not move elements from one row to a different row). If $e \in T\mathcal{N}_f$ then $U_{pq} e = 1$; thus,

$$\langle X_{pq} \mathcal{N}_f \rangle = \frac{N}{2} + S) \cdot (\frac{N}{2} - S)! / \mathcal{O}^\alpha$$

The $G_I^\alpha \mathcal{N}_f$ used so far have not been normalized, denote the normalized function as $G_I^\alpha \mathcal{N}_f$. From Ba-2, Ba-3 and Ba-4 ($M_S = S$ has been assumed, * There is a general convention for numbering the standard tableaux (see RUTHERFORD (1948) p.24). All that will be necessary here is to know that $S_1$ is the standard tableau obtained by filling successive rows one at a time and $S_2$ is the standard tableau obtained by filling successive columns one at a time
but it is shown below that Ba-7 and Ba-8 are valid for any \( M_S \)

\[
\mathcal{C}_i^\alpha \Phi^\alpha = C_n^\alpha (M_S) \sum \mathcal{F}_{\tau_{\Phi_i}} (0^\alpha_{\tau_{\Phi_i}}) (\omega^\alpha_{\tau_{\Phi_i}} \chi^\alpha_i)
\]  
(Ba-7)

where

\[
C_n^\alpha (M_S) = \sqrt{\frac{\mathcal{S}^\alpha}{\mathcal{N}_S}} |k^\alpha_1 \omega_{\tau_1} \chi^\alpha_1\rangle
\]  
(Ba-8)

For any \( M_S \) we let \( C_1^\alpha \Phi^\alpha = C_1^\alpha (\Phi^\alpha) = \sum \mathcal{F}_{\tau_{\Phi_i}} (0^\alpha_{\tau_{\Phi_i}}) (\omega^\alpha_{\tau_{\Phi_i}} \chi^\alpha_i) \)

where \( C_1^\alpha (\Phi^\alpha) = \sqrt{\frac{\mathcal{S}^\alpha}{\mathcal{N}_S}} |k^\alpha_1 \omega_{\tau_1} \chi^\alpha_1\rangle \)

Thus,

\[
\langle \mathcal{C}_2^\alpha \Phi^\alpha | \mathcal{C}_2^\beta \Phi^\beta \rangle = |S^\alpha \beta| S_\gamma
\]  
(Ba-9)

For arbitrary \( M_S \) the non-zero \( \omega_{\tau_1} \chi^\alpha_i \) for different \( i \) are all proportional but I have not discovered a simple way to find the coefficient for the general case (see Appendix E for an example). If \( \omega_{\tau_1} \chi \neq 0 \), let

\[
\omega_{\tau_2} \chi = \eta_{\tau_2} \omega_{\tau_1} \chi
\]  
(Ba-10)

Then equations Ba-7, and 8 are valid for any \( M_S \) and

\[
\langle k^\alpha_1 \omega_{\tau_1} \chi^\alpha_1 | k^\alpha_2 \omega_{\tau_1} \chi^\alpha_2 \rangle = \eta_{\tau_2} \eta_{\tau_1}^2 \langle k^\alpha_1 \omega_{\tau_1} \chi^\alpha_1 | k^\alpha_1 \omega_{\tau_1} \chi^\alpha_1 \rangle = \sqrt{\langle k^\alpha_1 \omega_{\tau_1} \chi^\alpha_1 | k^\alpha_2 \omega_{\tau_2} \chi^\alpha_2 \rangle}
\]  
(Ba-11)

I define

\[
\mathcal{C}_2^\alpha \Phi^\alpha = \sum \mathcal{F}_{\tau_{\Phi_i}} (C_n^\alpha \Phi_i) (\omega^\alpha_{\tau_2} \chi^\alpha_i)
\]  
(Ba-12)

We expand the exact wave function for the system in terms of the \( \mathcal{C}_2^\alpha \Phi(\text{the exact wave function is required to be an eigenfunction of } S^2 \text{ and } S_z) \)

\[
\Phi_{\alpha'} (\tau_1, \tau_2, ..., \tau_n) = \sum \sum \mathcal{C}_2^\alpha \mathcal{C}_2^\beta \Phi^\alpha \Phi^\beta
\]  
(Ba-13)
where the sum over $\Phi$ is over all different ordered $\Phi$

$$\Phi = \phi_1 \phi_2 \phi_3 \cdots \phi_{2n-1} \phi_2 \phi_3 \cdots \phi_{2n+1} \phi_{n+1} \phi_n \phi_{n+1} \cdots \phi_{2n} \phi_{2n+1} \cdots \phi_{2n+1} \phi_n (N)$$

Thus,

$$\sum_{\Phi} = \sum_{m=0}^{\infty} \sum_{\Phi_{M}} \sum_{\Phi_{M-2m}}$$

Remember that although for a given value of $S$ there are restrictions on the possible $n$, in general there are several possible $n$ for the same $S$. Our convention for $M_n^S$ is such that it also depends on $n$. Thus,

$$M_n^S (\vec{r}_1, \cdots, \vec{r}_N) = \sum_{M^n} \sum_{M_{M-2n}} \sum_{C_{M^n} (\vec{r}_1, \cdots, \vec{r}_n, \vec{r}_{n+1}, \cdots, \vec{r}_{2n-2m})} \sum_{\phi_1 \phi_2 \phi_3 \cdots \phi_{2n-1} \phi_2 \phi_3 \cdots \phi_{2n+1} \phi_{n+1} \phi_n \phi_{n+1} \cdots \phi_{2n} \phi_{2n+1} \cdots \phi_{2n+1} \phi_n (N)$$

By Ba-4 the terms in this sum are all orthogonal.
Bb*: First- and Second-Order Density Matrices for $e_i^a \Phi^b$

Our Hamiltonian, A-1, contains only one- and two-electron operators; so, for evaluating energies it will only be necessary to deal with the electron densities and pair densities discussed below.

Consider the one-electron operator $H^1 = \sum_i h^1_i$, where $h^1_i$ operates on electron $i$ ($H^1$ is invariant under permutations of the particles). If $\Psi$ is an N-electron wave function which satisfies Pauli's principle, then $\langle \Psi | H^1 | \Psi \rangle = \sum_i \langle \Psi | h^1_i | \Psi \rangle = \frac{1}{N!} \sum_i \langle \Psi | h^1_i | \Psi \rangle$, but since all coordinates in $\langle \Psi | h^1_i | \Psi \rangle$ are integrated over, then

$$\langle \Psi | H^1 | \Psi \rangle = \sum_i \langle \Psi | h^1_i | \Psi \rangle = \langle \Psi | H^1 \Psi \rangle = N \langle \Psi | H^1 | \Psi \rangle.$$ But $\Psi \equiv \Phi^a$, so, let $\xi = (1,i)$, then $\langle \Psi | h^1_1 | \Psi \rangle = \langle \Psi | h^1_i | \Psi \rangle$. Thus,

$$\langle \Psi | H^1 | \Psi \rangle = \sum_i \langle \Psi | h^1_i | \Psi \rangle = N \langle \Psi | H^1 | \Psi \rangle.$$ Let

$$\rho(1,i') = N \sum \psi^\ast (1,\ldots,n) \psi (i',\ldots,n) (d\xi_1/\Lambda) \quad (Bb-1)$$

where $(dx_1') = dx_2 dx_3 \ldots dx_N$ and, $(ds) = ds_1 ds_2 \ldots ds_N$ (i.e., integrate (actually sum) over all spins) then,

$$\langle \Psi | H^1 | \Psi \rangle = \sum_i \langle \Psi | h^1_i | \Psi \rangle = \sum_i \rho(1,i') \sum_j \rho(1',j) \langle \Psi | dx_1 \quad (Bb-2)$$

where the $\delta_{i,j}$ ensures that the operator $h^1_i$ operates on $\Psi$ and not on $\Psi^\ast$ (I may abbreviate it as $h(1) \rho(1))$. $\rho(1,1')$ is called the first-order density matrix, and $\rho(1)$ (i.e., $1 = 1'$) is called the electron density since $\rho(1) \ dx_1$ is the probability of there being an electron in the volume $dx_1$.

Consider the two electron operator $H^2 = \sum_{p<q} \delta_{pq}$, then, as

* See Löwdin (1955) for a general discussion of the density matrices. The notation is based on that of Ruedenberg (1962).
above,
\[
\langle \psi | H^2 | \psi \rangle = \sum_{\mu \nu} \langle \psi | \hat{g}_{\mu \nu} | \psi \rangle = \frac{1}{2} \sum_{\mu \nu} \langle \psi | \hat{g}_{\mu \nu} | \psi \rangle = \frac{1}{2} \sum_{\mu \nu} \langle \psi | \hat{g}_{\mu \nu} | \psi \rangle = \frac{N(N-1)}{2} \langle \psi | \hat{g}_{12} | \psi \rangle \quad (Bb-3)
\]

Let \( \Pi(1,2) = N(N-1) \int \psi^*(1,..N) \psi(1,..N) (dx''_{12}) (ds) \), where 
\( (dx''_{12}) = dx_3 dx_4 .. dx_N \). \( \Pi(1,2) \) is called the electron pair density since \( \Pi(1,2) dx_1 dx_2 \) is the probability of there being an electron in \( dx_1 \) and an electron in \( dx_2 \) simultaneously. Then,
\[
\langle \psi | H^2 | \psi \rangle = \frac{1}{2} \sum_{\mu \nu} \Pi(1,2) dx_1 dx_2 \quad (Bb-4)
\]

If the wave function of the system is approximated by a Slater determinant of SODS molecular orbitals (a common approximation), then
\[
\rho(1,1') = \sum_{i=1}^{N} \phi_i^*(1) \phi_i(1') \quad (Bb-5)
\]
\[
\Pi(1,2) = \sum_{i,j} \left[ \rho_i(1) \rho_i(2) \rho_j(1) \rho_j(2) - \rho_i(1) \rho_j^*(2) \rho_i^*(1) \rho_j(2) \right] \quad (Bb-6)
\]

Ruedenberg (1962) lets,
\[
\rho(1,2) = \rho(1) \rho(2) - \rho_x(1,2) \quad (Bb-7)
\]

where \( \rho_x \) is called the exchange pair density. Ruedenberg suggests that \( \rho_x \) does two things, first it corrects for the excess contribution in \( \rho(1) \rho(2) \) due to the same electron being at positions 1 and 2 (the integral (with factor \( \frac{1}{2} \)) of \( \rho(1) \rho(2) \) yields \( N^2/2 \) whereas there are only \( N(N-1)/2 \) pairs, the integral of \( \rho_x \) provides the correction for the \( N/2 \) self pairs), and second, it includes the correlations.
between the electrons*. Ruedenberg expands the (SODS) molecular orbitals in a basis of atomic orbitals, and by carefully partitioning the resulting expansions of \( \rho \), \( \Pi \), and \( \Sigma \), he is able to give a truly lucid and interesting discussion of the nature of the chemical bond based on some Hartree-Fock calculations on simple molecules. Relations similar to Bb-1 through 4 follow for \( \langle \Psi_1 | H | \Psi_2 \rangle \) where \( \Psi_1 \neq \Psi_2 \); so, we consider \( \Psi_1 = g_1 \Phi_1 \chi_1 \) and \( \Psi_2 = g_2 \Phi_2 \chi_2 \). The Nth-order density matrix is

\[
\gamma_{ \Phi \Phi }^{ \alpha \beta } (x_1 \ldots x_n) = N! \int \left( g_1 \Phi_1 \chi_1 \right)^\alpha \left( g_2 \Phi_2 \chi_2 \right)^\beta dx_1 \ldots dx_n
\]

where Ba-8 has been used. \( \leq \sum_{\tau} (g_{\alpha \Phi} \Phi)^\tau (g_{\beta \Phi} \Phi)^* \) = \( \leq \sum_{\tau} \sum_{\tau'} \left( U_{\alpha \tau} \Phi_{\tau'} \right) \left( U_{\beta \tau} \Phi_{\tau'}^* \right) \)

where we have used Aa-2 and \( \leq \sum_{\alpha} U_{\alpha \tau} \Phi_{\tau'} = U_{\alpha \tau} \Phi_{\tau'} \).

Thus, using \( \gamma_{ \Phi \Phi }^{ \alpha \beta } = \sum_{\tau} U_{\alpha \tau} \Phi_{\tau'} \left( U_{\beta \tau} \Phi_{\tau'}^* \right) \)

\[
\gamma_{ \Phi \Phi }^{ \alpha \beta } (x_1 \ldots x_n) = \sum_{\tau} U_{\alpha \tau} \Phi_{\tau'} \left( U_{\beta \tau} \Phi_{\tau'}^* \right) \]

The second-order density matrix is defined as

\[
\gamma_{ \Phi \Phi }^{ \alpha \beta } (x_1 \ldots x_n) = \sum_{\tau} U_{\alpha \tau} \Phi_{\tau'} \left( U_{\beta \tau} \Phi_{\tau'}^* \right) \]

where \( \langle dx_1 \ldots dx_n \rangle = dx_1 dx_2 \ldots dx_n \). The first-order density matrix is defined as

\[
\gamma_{ \Phi \Phi }^{ \alpha \beta } (x_1 \ldots x_n) = \sum_{\tau} U_{\alpha \tau} \Phi_{\tau'} \left( U_{\beta \tau} \Phi_{\tau'}^* \right) (dx_1) = \frac{1}{N} \sum_{\tau} (g_{\alpha \Phi} \Phi)^\tau (g_{\beta \Phi} \Phi)^* (dx_1) \]

Consider \( \langle \Phi_{\alpha} \Phi_{\beta} \rangle \) this is independent of electrons 2 through \( N \) (do not confuse electrons with states; the integral, of course, does

* It should be realized that compared to the case where no correlation is included (Ruedenberg says this is the case with the SODS HF wave function) the introduction of correlation changes both \( \rho \) and \( \Pi \) (as is easily seen since the Virial then must be satisfied in both cases and only \( \rho \) leads to kinetic terms).
depend on the states which in the term \( \Phi \) happen to be associated with specific electrons, the application of \( G_i^\alpha \) makes the electrons indistinguishable so the integral is invariant if we operate with \((i, \tau''(1))\).

In particular, we choose \( \beta = \gamma^{-1} \). So,

\[
\sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') = (1, \tau''(1)) \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx) \]

Thus,

\[
N \subseteq \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') = N \subseteq \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx) \]

Similarly,

\[
\sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx''') = (1, \tau''(2)) \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx) \]

and

\[
N(N-1) \subseteq \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx''') = (1, \tau''(2)) \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx) \]

From THM Ab-1

\[
\sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx) = \Phi_1^\alpha \Phi_2^\beta \]

For the general case

\[
N(N-1) \cdots (N-r+1) \subseteq \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') = \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') \]

Using the preceding steps

\[
\Phi_1^\alpha \Phi_2^\beta = \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') \]

(\text{Bb-11})

\[
\Phi_1^\alpha \Phi_2^\beta = \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') \]

(\text{Bb-12})

Consider \( \Phi_1^\alpha = \Phi_2^\beta \) (remember we use the convention of THM Ab-2).

From Ba-1 \( \langle \Phi|0_{ij} \Phi \rangle = \frac{2^n}{\theta^n} S_{ij} \); also,

\[
\sum \Phi_1^\alpha \Phi_2^\beta (dx') = \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') \]

And,

\[
\sum \Phi_1^\alpha \Phi_2^\beta (dx') = \Phi_1^\alpha (p) \Phi_2^\beta (p) \]

So,

\[
\sum \Phi_1^\alpha \Phi_2^\beta (dx') = \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') \]

Also,

\[
\sum \Phi_1^\alpha \Phi_2^\beta (dx') = \sum \left( \Phi_1^\alpha \Phi_2^\beta \right) (dx') \]

If \( \tau' \in T \), \( \tau' \in T \) for \( (p, q) \) \( \tau' \) where \( \tau' \in T \), this case \( U_{ij} \tau = U_{ij} (p, q) \)

\[
- 0 \text{ otherwise} \]

The second case is different from the first if and only if \( (p, q) \notin T \).
So, \( \Phi^*(q_0^* \Phi) (dx'_{pq}) = \sum_{e} \left[ \Phi^*_e (q_0^*) \Phi_e (p) \Phi_e (q) + U_{ij}^e (p_q) \Phi^*_i (q) \Phi^*_j (p) \Phi_j (q) \Phi_i (p) \right] \),

where \( U_{ij}^e (p_q) \triangleq U_{ij} (p_q) \) if \( (p, q) \notin T_\Phi \)

\[ = 0 \quad \text{if} \quad (p, q) \in T_\Phi \]

Thus,

\[ \delta_{\alpha \beta} (1, \Phi \chi_M, \Phi \chi_M) = \delta_{\alpha \beta} \delta_{M, M} \sum_{p} \Phi^*_p \Phi_p (1) \]

\[ \subseteq \delta_{\alpha \beta} \delta_{M, M} \sum_{p} \left[ \Phi^*_p \Phi_p (1) + U_{ij}^e (p) \Phi^*_i (q) \Phi^*_j (p) \Phi_j (q) \Phi_i (p) \right] \] (Bb-13)

\[ \subseteq \delta_{\alpha \beta} \delta_{M, M} \sum_{p} \left[ \Phi^*_p \Phi_p (1) + U_{ij}^e (p) \Phi^*_i (q) \Phi^*_j (p) \Phi_j (q) \Phi_i (p) \right] \] (Bb-14)

If \( i = j \) and \( \alpha = \beta \), then Bb-13 and -14 are the \( \mathcal{R} \) and \( \mathcal{P} \) for the state \( \chi_i \Phi \chi_M \), compare with Bb-5, 6. It should be noted that in general the second-order density matrix is not expressible in terms of the first-order density matrix as is the case for Slater determinants. Even for the same space state, \( \chi_i \Phi \chi_M \) and \( \chi_j \Phi \chi_M \) have off-diagonal matrix elements (in the 'exchange terms') if more than one linearly independent \( \chi_i \Phi \) is possible for the same \( \alpha \) and for the given \( \Phi \). The same thing, of course, occurs when one uses Slater determinants except that there are interactions between more Slater determinants than there are between \( \chi_i \Phi \) (the \( \chi_i \Phi \) to interact must be of the same \( S \) and \( S_z \), the Slater determinants need only be of the same \( S_z \). Using THM Ab-2 there is only one linearly independent \( \chi_i \Phi \) (and thus, no off-diagonal matrix elements) for the same \( \alpha \) when \( 2S = N - 2n \), this is the case of maximum spin for a given number of doubly-occupied orbitals; the same statement is true for Slater determinants if, in addition, \( M = \pm S \).

Now, consider \( \Phi_1 \neq \Phi_2 \). \( \int \Phi_1^* \Phi_2 (dx') = \frac{1}{2} \sum_{p} U_{ij}^e (p) \Phi^*_i (q) \Phi^*_j (p) \Phi_j (q) \Phi_i (p) \)

For this to be non-zero \( \Phi_1 \) and \( \Phi_2 \) can differ by at most one state; however, the differing states will generally be at different positions (i.e., not match). Let \( \Phi_2' = \chi_1 \Phi_2 \) match \( \Phi_1 \) except for one state...
(note $\Phi'_2$ does not necessarily have all doubly-occupied states adjacent). If $\varepsilon' = \varepsilon_1', \varepsilon_2'$ then $\sum_{\pi'} \phi^* \Phi^*_{\pi'}(dx') \neq 0$ only if $\varepsilon_1', \varepsilon_2' \in T_{\Phi'_2}$ and $p$ is the position of the non-matching states. If $\Phi'_2$ and $\Phi'_1$ differ by having one more $c$ state and one less $d$ state ($d \neq c$), write $\Phi'_2 = \Phi'_1(c;d)$. Then (where $n' = n_1$)

$$\sum_{\Phi'_1} (0, \Phi^*_1(c;d))(dx'_p) = \sum_{\Phi'_2} \frac{2}{\pi_1} U_{ij} \varepsilon_1 \Phi^*_d(p) \Phi_c(p).$$

For $\sum_{\Phi'_1} \Phi'_2(dx''_{pq}) \neq 0$, $\Phi'_1$ and $\Phi'_2$ can differ in at most two states. If $\Phi'_2$ differs by two states from $\Phi'_1$ and contains one more of $c_1$ and $c_2$ states and one less of $d_1$ and $d_2$ states than $\Phi'_1$, let $\Phi'_2 = \Phi'_1(c_1'c_2';d_1'd_2')$. Let $\Phi'_2 = \Phi'_1$ match $\Phi'_1$ except in two positions and let $d_1$ and $c_1$ correspond.

$$\sum_{\Phi'_1} \Phi'_2(c_1'c_2';d_1'd_2')(dx''_{pq}) = \sum_{\Phi'_2} \left[ U_{ij} \varepsilon_1 \left( S_{d_1'} S_{d_2'} \phi^*_d(p) \phi^{*}_c(p) \phi_c(p) \phi^{*}_{c_1}(p) + \phi^*_d(p) \phi^{*}_c(p) \phi_c(p) \phi^{*}_{c_2}(p) \right) + U_{ij} \varepsilon_2 \left( S_{d_1'} S_{d_2'} \phi^*_d(p) \phi^{*}_c(p) \phi_c(p) \phi^{*}_{c_1}(p) + \phi^*_d(p) \phi^{*}_c(p) \phi_c(p) \phi^{*}_{c_2}(p) \right) \right]$$

If $\Phi'_2$ and $\Phi'_1$ differ only by one state, set $d_2 = c_2$ equal to whatever the $S$-function implies. Hence, $\Phi_1^\alpha$.

$$S_{\alpha} \Phi_1(1;\infty) \Phi_2(1,2;\infty) = \sum_{\Phi'_1} \Phi_1(1) \Phi_2(1,2;\infty)$$

Using Bb-13 through 17 any operator involving one- and two-electron operators can be evaluated for the expansion in terms of $g_1^\alpha$ of the exact wave function of Ba-13.
Bc: The Set of $\Phi_1^\alpha$ With the Same $\Phi$: Hund's Rule and the Heisenberg Hamiltonian

Consider the matrix elements of the Hamiltonian between the various $\Phi_i$ for the same space state $\Phi$. $H = \sum_x \lambda_x + \sum_{\Phi_i} g_{ij}$

Using Bb-2, 4, 13, 14

$$E_{ij} = \sum_{\alpha}^{\Phi} \sum_{i,j} \left[ \frac{1}{2} \sum_{q} \phi_q^\dagger \phi_q \right] + \frac{1}{2} \sum_{\Phi_i} \sum_{q} \phi_q^\dagger \phi_q$$

(Bc-1)

It is common to define

$$J_{pq} = \left< \phi_{ip}^\dagger \phi_{iq}^\dagger \phi_{q} \phi_{ip} \right> = "Coulomb integral"$$

$$K_{pq} = \left< \phi_{ip}^\dagger \phi_{iq}^\dagger \phi_{q} \phi_{iq} \right> = "Exchange integral"$$

(Bc-2)

We would like to compare the energies for the different $\alpha$, $i$, and $j$; the important quantity is $U_{ij}(p,q)$. Because of our choice of standard form for $\Phi$ (THM Ab-2) all tableaux $S_i$ for which $C_i^\alpha \Phi \neq 0$ must be of the form (remember $n$ is the number of doubly-occupied orbitals)

$$\begin{array}{c|c}
1 & 2 \\
3 & 4 \\
5 & 6 \\
\vdots & \vdots \\
2n-1 & 2n \\
\end{array}$$

I call a tableau of this form the lowest-in-2n. So the possible $\alpha$ and $i$ for the above case are in a one-to-one correspondence with the possible $\alpha'$ and $i'$ for N-2n electrons and general $\Phi'(1,\ldots,N-2n)$ (i.e., each state singly-occupied). In appendix A it is shown that

$$U_{ij}(p,q) = -\frac{1}{2} \sum_{ij}$$

If $p \leq 2n$ and $q$ is any letter not in the same row as $p$. Thus,

$$E_{ij} = A S_{ij} + \frac{1}{2} \sum_{p,q \geq 2n} U_{ij}(p,q) K_{pq}$$

(Bc-3)
where \( A \) (which is independent of \( \alpha, \ i, \) and \( j \)) is
\[
A = \sum_{p,q} \langle p \mid \lambda + 1/2 \rangle \langle p \mid \lambda \rangle + \frac{1}{2} \sum_{p,q} \langle p \mid \lambda + 1/2 \rangle \langle p \mid \lambda \rangle + \frac{1}{2} \sum_{p,q} \langle p \mid \lambda + 1/2 \rangle \langle p \mid \lambda \rangle + \frac{1}{2} \sum_{p,q} \langle p \mid \lambda + 1/2 \rangle \langle p \mid \lambda \rangle,
\]
\[\text{(Bc-4)}\]

where we used \( J_{pp} = K_{pp} \). We now delete the A and obtain
\[
E_{i,j}^\alpha = \sum_{p,q > 2n} U_{ij}(p,q) K_{pq}^\alpha
\]
\[\text{(Bc-5)}\]

For \( \alpha \) with more than one non-zero \( G_{i,j}^\alpha \), we must diagonalize \( E_{i,j}^\alpha \).

In so doing we take a unitary transformation on the \( G_{i,j}^\alpha \) of the same \( \alpha \) which subjects each matrix \( U_{ij}(p,q) \) in the above equation to the same unitary transformation, but since \( U_{ij} \) is unitary it is left unitary. The diagonal elements of a unitary matrix have magnitude less than or equal to one (\( |U_{ij}|^2 \sum_{ij} U_{ij} U_{ij}^\dagger = 1 \)); thus, if \( V \) is the transformed \( U \),
\[
E_{i}^\alpha = \sum_{p,q > 2n} V_{ii}(p,q) K_{pq}^\alpha
\]
\[\text{and } |V_{ii}| \leq 1.\]

The \( K_{pq}^\alpha \) are all positive since each is the self-energy of a charge distribution (ROOTHAAN (1951)); thus, the lowest possible state would be one (say \( i \)) for which \( V_{ii}(p,q) = -1 \) for all \( p \) and \( q \) (we do not expect to achieve this usually since there are \((N-2n)(N-2n-1)\) different matrices being transformed by the same unitary transformation). Consider the \( \alpha = [2^n, 1^{N-2n}] \), there is only one tableau lowest-in-2n and, hence, only one non-zero \( G_{i,j}^\alpha \). If \( p > 2n \), \( U_{ij}^\alpha(p,p+1) = -S_{ij} \); thus,
\[
U_{ij}^\alpha(p,q) = -S_{ij} \quad \text{for } p > q > 2n.
\]
So the \( E_{i,j}^\alpha = S_{ij} \sum_{p,q > 2n} K_{pq}^\alpha \) is already diagonal and this state has the lowest possible energy. We have found, then, that the lowest possible state* among all possible states

* We have not shown that there will not be a tie but such a case would not be due to symmetry but rather to an accidental relation between the \( K_{pq}^\alpha \) and would thus be quite unlikely.
with the same orbital occupations is the state of highest spin. We have not stated how the \( \phi_{ip} \) are to be determined. If they are eigenfunctions of the one-electron operator \( h \) (i.e., neglecting electron-electron interactions), then equation Bc-1 is the first-order perturbation energy. If we include the electron-electron interactions in the self-consistent field sense as in section Bd and choose the \( \phi_i \) which minimize the energy, then the best \( \phi_i \) are actually different for different \( j \) and same \( \alpha \) of \( C^\alpha_j \). The point is that regardless of which \( \phi_i \) are used the \( \alpha \) of the maximum spin for which \( C^\alpha_j \phi_i \neq 0 \) is the lowest energy state. This is a general Hund's rule being applicable to any system of electrons (within the restrictions of SODS and no spin interactions). It should be remembered that the Hund's rule which is applicable to solids might be stated as: Given that the ground state of an atom has electrons occupying certain closed shells and a certain number of electrons in certain incomplete shells; then among all possible sets of orbital occupation numbers giving the same number of electrons in each shell, the actual set of occupation numbers for the various states of the incomplete shell is that allowing the maximum possible spin and the maximum \( L \) consistent with that spin; and the ground state is an eigenfunction with the above values of \( S \) and \( L \) and \( J \) equal to the maximum or minimum amount (depending on whether the shell is less than or more than half filled) consistent with the \( S \) and \( L \) eigenvalues. In considering the different sets of occupation numbers of the orbitals of an incomplete shell, we must consider the case (if allowed) of no doubly-occupied orbitals, all of the occupied orbitals being singly-occupied; one doubly-filled orbital, the rest singly-
occupied; two doubly-occupied orbitals; etc.. We have shown that among the possible spin states with no doubly-occupied orbitals the lowest energy state is the one with the maximum spin, among the states with one doubly-occupied orbital the lowest energy state is the state of maximum spin, the same for two doubly-occupied orbitals, etc.. To continue on to proving Hund's rule we must show that (at least for atoms) the state with the fewest doubly-occupied orbitals (and highest S for this n) is the lowest energy state. We might first try showing that the highest spin state for no doubly-occupied orbitals has a lower energy than the highest spin state for one doubly-occupied orbital and a lower energy than the highest spin state for two doubly-occupied orbital, etc., (I presumed here for definiteness that the shell is not more than half-filled, if it is more than half-filled then we modify accordingly). Actually, even if proved we would still not have even part of Hund's rule because the states must be eigenstates of L. So, even if the lowest energy $G_1$ has the highest possible spin, it is possible that when combinations are taken to get eigenfunctions of L, some states of lower spin will split more and give a lower energy. I have not been able to prove a complete Hund's rule even for the case of an atom (every proof eventually involves considering each possible number of electrons in the incomplete shell and diagonalizing for L but this tedious procedure has been done before and does not generalize to any N). KOSTER (1953) in an unpublished work considered a set of singly-occupied states and started from an equation similar to $\text{Be 5}$ (except that the coefficients of $K_{pq}$ were components of some undetermined unitary matrix) which he obtained from Dirac's vector model and found that among the
set of states (no closed shells) with no doubly-occupied orbitals, the
state of spin = \( \frac{N}{2} \) is the lowest. Since we can discuss cases where
there are some doubly-occupied orbitals and thus can discuss systems
having some closed shells, we have a much more general result than
Koster's. As Koster did (and for lack of a better term) I will call my
Hund's-type rule the extended Hund's rule since it applies to any sys-
tem of electrons and despite the fact that for atoms it is more restrict-
ed than Hund's rule.

Consider \( \sum_i E_i^\alpha = \sum_{p>q>2n} K_{pq} \sum_i U_{ii}(p,q) \) . Pretend that the
closed shells are not there; then we have \( r = N-2n \) electrons and \( r \) dif-
ferent one-electron states; so, all \( \alpha \) of the \( G_{ik}^\alpha \) are allowed. Then

\[
\sum_i U_{ii}(p,q) = \delta_{\alpha}^{\alpha} = \text{character of transposition (same for all transposi-
tions)}.
\]

\[
E_{ave} = \frac{1}{r} \sum_i E_i^\alpha = \frac{1}{r} \sum_{p>q>2n} K_{pq} .
\]

From appendix B

\[
\frac{\delta_{\alpha}^{\alpha}}{\delta_{\alpha}^{\alpha}} = -\frac{2}{r(r-1)} \left[ \frac{r(r-\Delta)}{4} + S(S+1) \right] ; \quad \text{Thus,}
\]

\[
E_{ave} = -\frac{2}{r(r-1)} \left[ \frac{r(r-\Delta)}{4} + S(S+1) \right] \sum_{p>q>2n} K_{pq} . \tag{Bc-6}
\]

\[
S(S+1)G_{\alpha}^\alpha = \sum_{\beta=1}^{2} \sum_{\rho=1}^{2} \frac{\gamma_{\alpha}^\beta \gamma_{\alpha}^\rho}{r} = \left( \sum_{\beta=1}^{2} \sum_{\rho=1}^{2} \frac{\gamma_{\alpha}^\beta \gamma_{\alpha}^\rho}{r} \right) G_{\alpha}^\alpha = \left( \sum_{\beta=1}^{2} \sum_{\rho=1}^{2} \frac{\gamma_{\alpha}^\beta \gamma_{\alpha}^\rho}{r} \right) \tag{Bc-7}
\]

thus,

\[
F_{ave} = -\left[ \frac{1}{r} + \frac{2}{r(r-1)} \sum_{\frac{\gamma_{\alpha}^\beta \gamma_{\alpha}^\rho}{r}} \right] \left( \sum_{p>q>2n} K_{pq} \right) . \tag{Bc-8}
\]

If we now set all the \( K_{pq} \) equal

\[
E_{ave} = -\left[ \frac{r(r-1)}{4} + \sum_{k,l} \frac{1}{4} \gamma_{k} \cdot \gamma_{l} \right] K = -\sum_{k,l} \left( \frac{1}{4} + \gamma_{k} \cdot \gamma_{l} \right) K . \tag{Bc-9}
\]

and if we

now let \( K = K_{kl} \) for the \( kl \) term, etc.,

\[
E_{ave} = -\sum_{k,l} \left( \frac{1}{4} + \gamma_{k} \cdot \gamma_{l} \right) K_{kl} \tag{Bc-9}
\]

Equation Bc-9 is the Heisenberg Hamiltonian used for parameterizing
magnetic interactions among electrons in solids. It must be emphasized
that unless all of the $K_{pq}$ are equal then the $f^\alpha$ different states of the same $\alpha$ (same total spin) are at $f^\alpha$ different energies; so, the differences between $E_{\text{ave}}$ for different $\alpha$ have little to do with the actual spectrum of energies. The pretense that the closed shells are not there is valid since the $U_{ij}(p,q)$ for $p, q \geq 2n$ are the same as for the case where there are only $r = N-2n$ electrons (e.g., see the lemma in appendix A). The transition from Bc-8 to Bc-9 is, of course, utterly ridiculous. It should be remembered that all of the $K_{pq}$ are positive.

A restriction common to this whole section is that we use the same set of $\phi_i$ for different $G^\alpha_i$ and then we require that the $G^\alpha_i \Phi^\alpha_j$ or linear combinations (with the same $\Phi^\alpha_j$) form eigenstates of the $N$-electron system. It is all right to form $G^\alpha_i \Phi^\alpha_j$ with some set of $\phi_i$ and then to vary the $\phi_i$ to try to obtain a good approximation to an eigenstate, but we must do this for each $\alpha$ separately. For the same basic set of $\phi_i$, we would vary, among other parameters, the scale of the different $\phi_i$, (we do not change the type of symmetry of the $\phi_i$) to minimize the energy and find the ground state of each, but the scale changes of the same $\phi_i$ for different $\alpha$ would in general be different. Hence, the integrals in (Bc-1) are different for different $\alpha$. But we have assumed them to be the same; thus, the energy differences obtained from equation Bc-2 and, hence, from all other equations in this section, are quite without physical meaning. The only purpose for which they can be used is to select the ground state as justified on p. Be-3. The Heisenberg Hamiltonian is often used in a way which is not justifiable on theoretical grounds; we will do so ourselves in section E.
Let us try to find the lowest energy $E_1^{\Phi}$ for a given type of \( \Phi \). Each $\phi_1$ is required to be a basis function of an irreducible representation of the spatial symmetry group of the Hamiltonian. Let the dimension of the $K$ irreducible representation be $d_K$. If there is among the $\phi_i$ in $\Phi$ a set of $d_K$ doubly-occupied orbitals transforming as the $d_K$ different basis functions of $K$, we say that the set of $d_K$ orbitals comprises a closed shell. If there are several shells for the same $K$, then we number them $1K$, $2K$, $3K$ in order of increasing one-electron energy (as from a Hartree-Fock calculation). A set of functions which accounts for less than $2d_K$ electrons (but at least one) and which transforms as basis functions of an irreducible representation under the operations of the spatial symmetry group comprises an open shell. A set of $d_K$ singly occupied orbitals transforming as the $d_K$ different basis functions of $K$ is termed to comprise a half-filled shell (note it is not sufficient that there be $d_K$ electrons in the shell). Two $\Phi$'s are defined as being of the same type if they contain the same number of one-electron functions of each symmetry. The variations in each one-electron function must leave the number of functions of each shell the same, and all of the basis functions of the same shell must change in the same way under the variation so that they remain basis functions for the irreducible representation. We require that $\delta E_{11}^{\alpha} = 0$ under the constraint that the $N-n$ occupied one-electron functions remain orthogonal (this also preserves the norm of the many-electron function). Thus,

$$\delta (E_{11}^{\alpha} - \sum_{\rho,\eta} \epsilon_{p\rho} \langle \phi_\rho | \phi_\eta \rangle ) = 0$$

(Bd-1)
Taking the complex conjugate of this equation, we see that $\zeta_{pq}$ is Hermitian. Using Bc-1

$$\gamma [\sum_{p=1}^{N} \langle \phi_{p} | l_{p} | \phi_{p} \rangle + \frac{1}{2} \sum_{p,q} \left( \langle \phi_{p} | g_{pq} | \phi_{q} \rangle + \langle \phi_{q} | g_{pq} | \phi_{p} \rangle \right) - \frac{\pi}{\rho} \langle \phi_{p} | l_{p} | \phi_{p} \rangle] = 0$$

One separates the variation of $\delta \phi_{p}^{*}$ from that of $\delta \phi_{p}$ by replacing $\delta \phi_{p}$ by $i \delta \phi_{p}$ then multiplying the equation by $i$ and adding the result to the above equation (see MESSIAH (1962) p. 764). Thus,

$$\sum_{p=1}^{N} \left[ \langle \delta \phi_{p} | h | \phi_{p} \rangle + \sum_{g,p} \left( \langle \phi_{p} | g_{pq} | \phi_{q} \rangle + \frac{\rho}{\rho} \langle \phi_{q} | g_{pq} | \phi_{p} \rangle - \frac{\rho}{\rho} \langle \phi_{p} | h | \phi_{p} \rangle \right)\right] = 0 \quad \text{(Bd-2)}$$

If all of the $\delta \phi_{p}$ are linearly independent, then the coefficient of $\delta \phi_{p}^{*}$ is $0$, that is,

$$\sum_{p=1}^{N} \left[ \langle \delta \phi_{p} | h | \phi_{p} \rangle + \sum_{g,p} \left( \langle \phi_{p} | g_{pq} | \phi_{q} \rangle + \frac{\rho}{\rho} \langle \phi_{q} | g_{pq} | \phi_{p} \rangle \right) - \frac{\rho}{\rho} \langle \phi_{p} | h | \phi_{p} \rangle \right] = 0 \quad \text{(Bd-3)}$$

If the coefficient of $\delta \phi_{p}$ in the equation corresponding to Bd-2 is set equal to zero, then the resulting differential equation is the complex conjugate of Bd-3; thus, we obtain nothing new. Now, although it is true that all of the orbitals in a shell have to vary together so that they remain a basis of an irreducible representation, the $\delta \phi_{p}$ for two states of the same shell are orthogonal because the $\delta \phi_{p}$ also transform as different basis functions of the irreducible representation.

If $\phi_{p}$ is doubly occupied then the same equation occurs twice. Thus, equation Bd-3 is valid. Equation Bd-3 is the Hartree-Fock equation and can in principle be solved iteratively by integrating for $\phi_{p}$ after solving for the integrals from the previous approximation (this is impractical except for atoms). Consider $\langle \phi_{p} | \phi_{q} \rangle$ and $\langle \phi_{p} | \phi_{q} \rangle$ on equation Bd-3 (and use $\langle \phi_{p} | \phi_{q} \rangle = \delta_{pq}$)
\begin{align}
\langle \psi | h | \psi \rangle + \sum_{q 
eq p} \langle \psi | g | q \rangle \langle q | \psi \rangle = \varepsilon_{p} \tag{Bd-4}
\end{align}

\begin{align}
\langle \psi | h | \psi \rangle + \sum_{q 
eq p} \langle \psi | g | q \rangle \langle q | \psi \rangle = \varepsilon_{p} \tag{Bd-5}
\end{align}

Since \( h \) and \( g \) are invariant under the spatial symmetry group, \( \varepsilon_{pr} = 0 \) unless \( \phi_{p}^{i} \) and \( \phi_{q}^{r} \) transform as the same basis function of the same irreducible representation (we are using the discussion following Bd-9 and Bd-10 here). Actually we should have expected \( \varepsilon_{pr} = 0 \) for different symmetry functions since \( \langle \phi_{p}^{i} | \phi_{q}^{r} \rangle = 0 \) by symmetry. A transformation among the occupied states which leaves \( G_{1}^{\Phi} \) invariant (except for a phase factor) leaves \( \varepsilon_{ii} \) unchanged. If \( \phi_{i} = \sum_{j} S_{ij} \phi_{j} \) (where the sum is over all different states), then

\begin{align}
\sum_{p, q} \varepsilon_{pq} \langle \phi_{p} | \phi_{q} \rangle = \sum_{i, j} \langle \phi_{i} | \phi_{j} \rangle \sum_{p, q} S_{pi}^{*} \varepsilon_{pq} S_{qj} \tag{S is unitary}
\end{align}

(since we require that orthonormality of the \( \phi_{i} \) be preserved) and \( \varepsilon_{pq} \) is Hermitian; thus, an \( S \) can be chosen which will diagonalize \( \varepsilon \).

We wish to preserve the symmetry properties of the \( \phi_{i} \) (which is possible since the only non-diagonal \( \varepsilon_{pq} \) are between states of the same symmetry). Hence, it is now necessary to see in which cases the transformation \( S \) leaves \( G_{1}^{\Phi} \) invariant.

Consider a set of functions which are transformed by \( S \) among themselves; then

\begin{align}
\Phi = \phi_{1}^{(1)} \phi_{2}^{(2)} \cdots \phi_{N}^{(N)} = \sum_{g_{1} \cdots g_{N}} \phi_{1}^{g_{1}} \phi_{2}^{g_{2}} \cdots \phi_{N}^{g_{N}} \phi_{1}^{(1)} \phi_{2}^{(2)} \cdots \phi_{N}^{(N)} \tag{Bd-6}
\end{align}

and consider \( G_{1}^{\Phi + \lambda} \). Assume that all of the functions in the transformed set are doubly-occupied orbitals. If the functions \( \phi_{i}^{p} \) and \( \phi_{i}^{q} \) are transposed, then we must use for \( \lambda \) in Bd-6 say
say \( \mathcal{G} = (2p-1,2p-1)(2q,2q) \) or \( (2p-1,2p)(2q,2q) \) \( \text{(Bd-7)} \)

The reason is that the \( \mathcal{G} \) in equation Bd-6 must refer to permutations of the (numbered)* electrons, this \( \mathcal{G} \) takes the two electrons in \( i_p \) states and puts them in \( i_q \) states and vice versa.

\[
G_i^\alpha (\varepsilon \Phi) \chi = \sum_r b_{\alpha r} (0_{\mathcal{G}_r} \varepsilon \Phi) (0_{\mathcal{G}_r} \chi)
\]

\[0_{\mathcal{G}_r} \varepsilon \Phi = \sum_j U_{ijr} 0_{\mathcal{G}_r} \Phi \text{ (from Aa-4), but by THM Ab-2}\]

\[0_{\mathcal{G}_r} \chi = S_{ij} 0_{\mathcal{G}_r} \chi \text{ (since } \Phi \text{ is ordered and } S_i \text{ is lowest-in-2n). Thus, }\]

\[0_{\mathcal{G}_r} \varepsilon \Phi = U_{ijr} 0_{\mathcal{G}_r} \Phi \text{. From appendix C } U_{ij}(2p-1,2q-1)(2p,2q) = S_{ij}, \text{ and thus for any permutation } \varepsilon \text{ composed of transpositions of type} \]

\[Bd-7 \quad U_{ij} \varepsilon = S_{ij} \text{ thus, } G_i^\alpha (\varepsilon \Phi) \chi = G_i^\alpha \Phi \chi, \text{ and the } G_i^\alpha \Phi \chi \text{ is invariant under all transformations of the space of doubly-occupied orbitals (we might say that if the pair of electrons in a doubly-occupied orbital are considered as a single entity then this entity is a Bose particle).} \]

It follows then that the \( G_{pq} \) between doubly-occupied orbitals are zero**. The \( G_{pq} \) between doubly-occupied orbitals and singly-occupied orbitals is invariant under all transformations of the space of doubly-occupied orbitals (we might say that if the pair of electrons in a doubly-occupied orbital are considered as a single entity then this entity is a Bose particle).

** In case anyone is worrying about numbering the electrons when they are indistinguishable, the particles in a function such as \( \Phi \) are distinguishable (if they are in different states); only when we operate with \( G_i \) or with the Slater antisymmetrizers do we get a wave function which describes indistinguishable particles.

** Actually the different basis functions of the same irreducible representation must be transformed by the same \( S \) so that they still form a basis for the irreducible representation, so, one might think that the \( G_{pq} \) between functions transforming as different basis functions of \( D_{pq} \) the irreducible representation could not all be simultaneously diagonalized; however, by applying the group operations on Bd-5 we see that these different matrices are the same and thus are diagonalized by the same \( S \).
orbitals of the same symmetry type need not be zero. However, if all of
the singly-occupied orbitals are basis functions of the same irre-
ducible representation and if this irreducible representation occurs
only once\(^*\) then the \( \varepsilon \) can be completely diagonalized\(^**\) (e.g., for
the (argon)\( d^n \) and (argon)\( d^{n-2} s^2 \) configurations of the first row trans-
ition elements). A case such as lithium would involve off-diagonal
\( \varepsilon_{pq} \). These same problems occur for the same reasons for Slater deter-
minants despite the common derivation (e.g., MESSIAH (1962) p. 777)
showing that the \( \varepsilon_{pq} \) can be diagonalized. The reason is that we have
restricted our wave function by requiring that the same set of orbitals
be used for different spins\(^***\) (SODS); this is an artificial restric-
tion which is justified because of the vast simplification in calcula-
tions (making them feasible) and because the restriction usually fits
with whatever experimental information is available. If no restrictions
are made on the spin-orbitals other than their orthonormality, then in-
deed a transformation can always be performed to diagonalize the \( \varepsilon_{pq} \);

\(^*\) This is true even if some of the states of this irreducible repre-
sentation are doubly-occupied, the reason is that \( \varepsilon_{pq} \) is non-zero
only between states which transform as the same basis function of
the same irreducible representation, thus an irreducible represen-
tation must occur twice in order to have off-diagonal \( \varepsilon_{pq} \).

\(^**\) A similar result for atoms with Slater determinant wave functions
is given by WATSON (1959) p. 7

\(^***\) Actually we required that the \( \Phi \) be basis functions of the irre-
ducible representations of the spatial symmetry group; however, even
without this additional assumption the unitary transformations of \( \Phi \n\) in order to leave \( \varepsilon_{ij} \) invariant must be restricted to transformations
among the doubly-occupied orbitals (see p. Bd-3) and transformations
among the singly-occupied orbitals (this latter is allowed only if
\( \alpha = [2^m_j ] \), see p. Cb-1,2) and thus, the \( \varepsilon_{ij} \) will, in general,
have off-diagonal elements for \( i \leq 2n < j \).
the "no restriction" means that we cannot require the SODS approximation.

We assume in the remainder that conditions prevail which allow

the \( \xi_{pq} \) to be diagonalized. From Bd-3

\[
\hbar l_2 \phi_p \Phi_p + \sum_{i \neq j} (\phi_i \phi_i^* \phi_j \Phi_p \Phi_p) + \hbar l_2 \text{Coulomb term} = \xi_{pq} \phi_p \Phi_p
\]

where \( \xi_{ii}(p,q) = \xi_{ii}(p,q) \) if \( i \neq j \)

\[
\xi_{ii}(p,q) = \begin{cases} 
0 & \text{if } i = j \\
- \frac{1}{2} & \text{if } i = j \text{ and } p,q \leq 2n \text{(to correct for the extra Coulomb term)} \\
-1 & \text{if } i = j \text{ and } p,q > 2n
\end{cases}
\]

Using the results of appendix A for \( \xi_{ij}(p,q) \) as in section Be., if \( p < 2n \)

\[
\hbar l_2 \phi_p \Phi_p + \sum_{i \neq j} (\phi_i \phi_i^* \phi_j \Phi_p \Phi_p - \frac{1}{2} \phi_i \phi_i^* \phi_j \Phi_p \Phi_p) = \xi_{pq} \phi_p \Phi_p
\]

if \( p > 2n \)

\[
\hbar l_2 \phi_p \Phi_p + \sum_{i \neq j} (\phi_i \phi_i^* \phi_j \Phi_p \Phi_p - \frac{1}{2} \phi_i \phi_i^* \phi_j \Phi_p \Phi_p) + \sum_{n \text{ even}} \phi_i \phi_i^* \phi_j \Phi_p \Phi_p + \hbar l_2 \text{Coulomb term} = \xi_{pq} \phi_p \Phi_p
\]

Now, in order for equations Bd-9 and 10 to be consistent with the restrictions on the \( \phi_i \), the equations above must transform as \( \phi_i \).

so, the term operating on \( \phi_i \) must be invariant under transformations of the spatial symmetry group. The sum \( \sum_1 \phi_i^* \phi_1 \phi_2 \) is invariant under all operations of a group iff the sum includes all basis functions of every irreducible representation of which it contains even one basis function. Therefore, our equations are consistent only if all shells are either full or half-full (remember the definition on p. Bd-1).

\( R \equiv \Phi_1(\alpha_1) \phi_1(\alpha_2) = \sum_{\alpha_2} \phi_1^*(\delta_1) \phi_1^*(\delta_2) \equiv R_1 \Phi_1 \Phi_2 \); \( R \) is unitary since the orthogonality is preserved. The set of functions is left invariant since the functions in each irreducible representation must transform among themselves; so, the sum over \( i \) is over all columns; thus, \( \sum_1 \phi_i^* \phi_1 \Phi_2 \Phi_2 = \sum_1 \phi_2 \Phi_1 \Phi_2 \). By the definition of irreducibility if some but not all of the basis functions of an irreducible representation are in the sum then the sum is not invariant.
and in addition for \( p, q > 2n \) (i.e., among the states of the half-full shells) then \( V_{ii}(p,q) \) must be the same for all \( q \) of the same shell (e.g., if there is only one unfilled shell then \( V_{ii}(p,q) \) must be independent of \( p \) and \( q \) for \( p, q > 2n \); the only case is \( V_{ii}(p,q) = -1 \) (maximum spin)). In cases other than these we cannot require the orbitals to be basis functions of the irreducible representations of the spatial symmetry group of the \( N \)-electron system. The only good solution is to relax the SODS restriction which of course complicates the problem considerably; this will be discussed in section C. For the remainder of this section we presume that all shells are either full or half-full. However, Hartree-Fock calculations are often made on systems for which they can't be self-consistent (open-but not half-shell); the results are not especially bad (WATSON, 1959) so I will leave my equations with the \( V_{ii} \) rather than -1 in Bd-10 to allow for these cases. The only cases where Bd-9 and Bd-10 are consistent with the \( \phi_{ip} \) all being basis functions of irreducible representations are just the cases of maximum spin of closed shell and half-filled shell systems, in these cases the Slater determinant is an eigenfunction of spin and gives the same Hartree-Fock equations.

* In the usual derivations of the Hartree-Fock equations for Slater determinants (this is redundant in the usual terminology since the Fock usually implies that the wave function is approximated by a Slater determinant) a spin (\( \alpha \) or \( \beta \)) is associated with each orbital. One gets an equation similar to Bd-8 except that \( V_{ii}'(p,q) \) is replaced by -1, each orbital is multiplied by a spin, and the integral (sum) over these spins is not completed. When the integral over spins is taken, the exchange terms between "orbitals with different spins" are deleted. It should be realized that for a system with no spin interactions, it is not possible to associate a definite spin with a specific orbital; only the total spin and total spin component can be fixed. Of course, if SODS is used and if an orbital is doubly-occupied, then one must have the orbital once with an \( \alpha \) spin and once with a \( \beta \) spin in every non-cancelled term of the expansion of the wave function. So, between
Note that if \( N \neq 2n \) then even for half-filled orbitals the operator on \( \Phi_{ip} \) in equation Bd-10 is not necessarily the same for different \( p \) and is not the same as the operator for Bd-9. Thus, the eigenfunctions would not automatically be orthogonal if it was not that they are different basis functions of an irreducible representation which does not occur in the doubly-occupied set of orbitals; thus, we see why an off-diagonal \( \varepsilon_{pq} \) is needed when the irreducible representation used in the singly-occupied set is repeated.

If the \( \Phi_{ip} \) in Bd-2 are expanded in terms of a fixed set of basis functions* (ROOTHAAN 1951) then the variations of \( \Phi_{ip} \) are variations in the coefficients of the expansion. In this case we obtain

\[
[H_{ij} + \sum_{ij} \frac{1}{2} \varepsilon_{ij} \Phi_{ij}(\xi_{ij} \psi_{ij})] C_{ij} = \varepsilon_{ij} \sum \Phi_{ij} C_{ij}, \quad \text{(Bd-11)}
\]

where \( \Phi_{ij} = \sum_{ij} \Phi_{ij} \), \( \xi_{ij} = N-M \), \( \lambda \) is a fixed basis function

\[
H_{ij} = \langle \chi_{ij} | H | \chi_{ij} \rangle, \quad \Phi_{ij} = \langle \chi_{ij} | \chi_{ij} \rangle
\]

Substituting for the \( U_{ii}^{(p,q)} \), if \( p \leq 2n \)

\[
[H_{ij} + \sum_{ij} \frac{1}{2} \varepsilon_{ij} \Phi_{ij}(\xi_{ij} \psi_{ij} - \frac{1}{2} \xi_{ij} \psi_{ij})] C_{ij} = \varepsilon_{ij} \sum \Phi_{ij} C_{ij}, \quad \text{(Bd-12)}
\]

* We use more basis functions than the number of different orbitals.

(continued from previous page)

two doubly-occupied orbitals there are four Coulomb terms and two exchange terms, from Bd-0 we see that the same situation prevails here since each orbital appears twice and has one Coulomb and one-half exchange term for each occasion that the other orbital appears. In the maximum spin case for a configuration, Bd-0 has one exchange term between singly-occupied orbitals and one-half exchange term between a singly-occupied orbital and each occurrence of the doubly-occupied orbital. So, the result for Slater determinants corresponds to the correct result for the case of maximum spin of a configuration (this includes closed shell systems, i.e., \( S = 0 \)); this case includes all those for which the Hartree-Fock equations can be self-consistent within the restriction that the orbitals be basis functions of the spatial symmetry group.
if $p > 2n$

$$\left[ H^{\mu \nu} + \sum_{\xi \gamma \nu} C^*_{\xi \gamma} C_{\xi \gamma} \left( q_{\mu \xi \gamma} \nu \lambda - \frac{1}{2} q_{\mu \xi \gamma} \lambda \nu \right) + \sum_{\xi \gamma \nu} C^*_{\xi \gamma} C_{\xi \gamma} \left( q_{\mu \xi \gamma} \nu \lambda \right) + \right. \left. + U^{\mu \nu}(p, q) q_{\mu \xi \gamma} \nu \lambda \right] C_{\rho \nu} = \delta_{\rho \nu} S_{\mu \nu} C_{\rho \nu}$$  \hspace{1cm} (Bd-13)

The advantage of the Roothaan procedure of solving Bd-12 over the Hartree procedure of integrating equation Bd-9 is that in the former it is merely necessary to diagonalize an $M \times M$ matrix whereas the latter involves a set of coupled partial differential equations which cannot in general be solved (in the case of $SO(3)$ the variables separate and a set of coupled ordinary differential equations result, even in this case the Roothaan procedure is apparently more convenient). As before, the eigenvectors of Bd-13 are orthogonal to the others because they are required to be by symmetry restrictions and not because they are different eigenvectors of a Hermitian matrix. As noted before Bd-13 is consistent with $\phi_i$ being a basis function of an irreducible representation only if $\alpha - \left[ \frac{p_R}{2n}, \frac{1}{2} (N-2m) \right] \nu$ in which case $U^{\mu \nu}(p, q) = -1$, and the operator in Bd-13 is independent of $p$ for $p > 2n$. This is an important case since it has the maximum possible spin consistent with the number of doubly-occupied orbitals.

If unitary transformations are taken on the set of Hartree-Fock orbitals by having the doubly-occupied orbitals transform among themselves, then $\phi_i$ is left invariant (see discussion after Bd-7). Under these transformations the Hartree-Fock Hamiltonian (Bc-1), the density (Bb-5), and the pair density (Bb-6) are left invariant. Often such transformations are used to obtain orbitals which are expected to be more consistent than the Hartree-Fock orbitals with the correlation terms of the complete Hamiltonian which are missing from the Hartree-Fock Hamiltonian. The disadvantage with such orbitals (e.g., Wannier
orbitals, hybridized orbitals, etc.) is that the off-diagonal \( <\rho \nu> \) (Bd-5) are non-zero. So, instead of Bd-8 the equation for the new orbitals is

\[
\chi_1 |\phi_\rho> + \sum_\delta ( <\phi_\delta |\delta |\phi_\rho> |\delta \rho> + U_{\nu\delta}^{\tau\tau} <\phi_\delta |\delta |\phi_\rho> |\phi_\nu>) = \\
= \sum_\delta [ <\phi_\delta |\delta |\phi_\rho> + \sum_\nu ( <\phi_\nu |\delta |\phi_\rho> + U_{\nu\delta}^{\tau\tau} <\phi_\nu |\delta |\phi_\rho> |\phi_\nu>) |\phi_\rho> \] \quad (Bd-14)
\]

which is extremely inconvenient to use.
In the Hartree-Fock equations each electron moves in the averaged field of the other particles. Actually, the motion of the real electrons is correlated so that they can reduce their mutual repulsions. Correlation effects are apparently important enough so that the Hartree-Fock equations are often of insufficient accuracy when discussing chemical aspects of molecules (Löwdin (1959), Sinanoğlu (1961)). In this section we find the first-order perturbed wave function; using the Hartree-Fock wave function as the zero-order function; with the first-order wave function, the second- and third-order perturbed energies can be found. The treatment is modeled after that of Sinanoğlu (1961a) who considered only closed shells, for which case the Slater determinant is an eigenfunction of $S^2$. The basic idea of using the Hartree-Fock wave function as the zero-order wave function is due to Møller and Plesset (1934).

Since the SODS Hartree-Fock equations are used as the zero-order functions, this approach can only be used where the SODS Hartree-Fock equations are valid (i.e., self-consistent) which is for closed-shell and half-shell systems. If we allow DODS, then the treatment is very much more difficult, and in that case the self-consistent field equations (see section C) are probably so much better that the perturbation is not necessary in order to get accurate results.
SUMMARY: The first-order perturbation wave function is obtained for the two types of cases for which the SODS Hartree-Fock equations are self-consistent (hence, this is limited to systems with only full and half-full shells). This first-order wave function is a sum of terms each of which is obtained from the zero-order wave function by replacing a pair of orbitals $\phi_i^p$ and $\phi_i^q$ with a two-electron function $\Theta_{i\bar{i}}^{p\bar{p} \bar{q}}$. If $i = \bar{i}$ then the $\Theta_{i\bar{i}}^{p\bar{p} \bar{q}}$ can be found by solving an inhomogeneous differential equation. If $i \neq \bar{i}$ then $\Theta_{i\bar{i}}^{p\bar{p} \bar{q}}$ is expressed as a sum over the excited states of the Hartree-Fock Hamiltonian. This allows us to accurately determine the $\Theta_{i\bar{i}}^{p\bar{p} \bar{q}}$ (which should be large since the states were artificially forced to be the same) and to determine $\Theta_{i\bar{i}}^{p\bar{p} \bar{q}}$ by a simpler less accurate method more in keeping with the lesser importance of $\Theta_{i\bar{i}}^{p\bar{p} \bar{q}}$. It should be stressed that we do include systems with half-filled shells and maximum spin.
From Bd-8, the equation for the Hartree-Fock molecular orbitals is

\[ [\hbar + \sum_q (\langle i_q | g | i_q \rangle + U''_{ii}(p,q) \langle i_q | g | i_p \rangle P_{pq}) | i_p \rangle = \epsilon_{pq} | i_p \rangle, \]

where \( U''_{ii}(p,q) \triangleq U_{ii}(pq) \) if \( i_p \neq i_q \)

\[ \Delta = -\frac{1}{2} \text{ if } \frac{i_p}{i_q} = 1 \text{ and } p,q \leq 2n \]

\[ \Delta = -1 \text{ if } \frac{i_p}{i_q} = 1 \text{ and } p,q > 2n \]

Let \( V_p = \sum_q (\langle i_q | g | i_q \rangle + U''_{ii}(pq) \langle i_q | g | i_p \rangle P_{pq}) \), then

\( (\hbar + V_p) | i_p \rangle = \epsilon_{pq} | i_p \rangle \).

Take \( | i_p \rangle \) and obtain

\( \epsilon_{pq} = [\langle p | h_{1p} | p \rangle + \sum (\langle p \phi | g | \phi \rangle + U''_{ii}(pq) \langle p \phi | g | i_p \rangle P_{pq})] \)

Let \( H_0 = \sum_{p=1}^{N} (\hbar + V_p) \); so, \( H_0 \phi_1 \ldots \phi_N = (\sum_{p=1}^{N} \epsilon_{pp} \phi_1 \ldots \phi_N) \).

Then multiply by \( X \) and operate with \( \bar{G}_i \); thus,

\( H_0 \overline{X} = (\sum_{p=1}^{N} \epsilon_{pp}) \bar{G}_i \overline{X} \) \text{ and } \( E_0 = \sum_{p=1}^{N} \epsilon_{pp} \) \hspace{1cm} (Be-1)

The perturbation is \( H_1 = \sum_{p \neq q} \frac{N}{p} - \sum \frac{V_p}{p} \). The first-order energy correction is

\( E_1 = -\frac{1}{2} \sum_{p \neq q} \left( \langle p | g | q \rangle + U''_{ii}(pq) \langle p | g | q \rangle \right) \) \hspace{1cm} (Be-2)

This first-order correction merely corrects for the electron-electron interaction energies being counted twice in \( E_0 \). Let \( \langle w | w \rangle = 1 \), then (e.g., SINANOGLU (1961a) \( (H_0 - E_0)X_1 = (E_1 - H_1) \psi_0 \)), where \( X_1 \) is the first-order correction to \( \psi_0 \).

From above

\( E_1 - H_1 = \sum_{p \neq q} \left( J_{pq} + U''_{ii}(pq) K_{pq} - g_{pq} \right) + \sum_{p \neq q} \left[ \frac{N}{2} - \sum_{p \neq q} \left( J_{pq} + U''_{ii}(pq) K_{pq} \right) \right] \)

let \( X_1 \triangleq Y - W \); \( m_{pq} \triangleq J_{pq} + U''_{ii}(pq) K_{pq} - g_{pq} \).
we have
\[ (H_0 - E_0) \psi_0 = \left( \sum_{\mathbf{p} \mathbf{q} \mathbf{p}' \mathbf{q}'} M_{\mathbf{p} \mathbf{q} \mathbf{p}' \mathbf{q}'} \right) \psi_0 \]
\[ = (\sum_{\mathbf{p} \mathbf{q}} M_{\mathbf{p} \mathbf{q}}) \psi_0 \]
\[ = (\sum_{\mathbf{p} \mathbf{q}} M_{\mathbf{p} \mathbf{q}}) \psi_0 \]
\[ \text{(Be-3)} \]

\[ \psi_0 \] is non-degenerate (except for the $2S + 1$ different spin states); so, \( \langle \psi_1 | M_{\mathbf{p} \mathbf{q}} | \psi_0 \rangle = 0 \) for all solutions (i.e., \( \psi_0 \)) of the homogeneous equation, and thus solutions to the inhomogeneous equation for \( \psi \) exist. We have \( \psi = \frac{1}{H_0 - E_0} \left( \sum_{\mathbf{p} \mathbf{q}} M_{\mathbf{p} \mathbf{q}} \right) \psi_0 \)

As discussed in section Bd the Hartree-Fock scheme as restricted here is valid only for states involving only closed- and half-filled (maximum spin) shells; so, only \( c_{\alpha}^\psi \chi \neq 0 \) if \( \bar{\Phi} \) is written in standard order. Thus, \( \bar{\Phi} = \sum_{\alpha} \bar{c}_{\alpha} \bar{\Phi}^\psi \chi^\psi \).

Since the one-electron orbitals and spin functions form a complete set of functions on the one-electron space, then the product functions \( \psi = \bar{\Phi}_{\alpha} \bar{\Phi}_{\beta} \cdots \bar{\Phi}_{\mu} \bar{S}_{\nu} \bar{S}_{\kappa} \) form a complete set of functions on the \( N \)-electron space. Thus, since the \( G_{\alpha}^{\psi} \chi \) is obtained by a non-singular transformation from the \( \psi \), then the set of \( G_{\alpha}^{\psi} \chi \) for all \( \bar{\Phi}, \bar{\Phi}, \bar{\Phi} \), and ordered \( \bar{\Phi} \) and \( \chi \) is complete. Thus, \( 1 = \sum_{\bar{\Phi}_{\alpha}, \bar{\Phi}_{\beta}, \bar{\Phi}_{\mu}} | \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \langle \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \bar{c}_{\beta}^{\psi} \chi_{\nu} | \rangle \)

and \( Y = \sum_{\bar{\Phi}_{\alpha}, \bar{\Phi}_{\beta}, \bar{\Phi}_{\mu}} | \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \langle \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \bar{c}_{\beta}^{\psi} \chi_{\nu} | \rangle \)

\[ = \sum_{\bar{\Phi}_{\alpha}, \bar{\Phi}_{\beta}, \bar{\Phi}_{\mu}} | \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \langle \bar{c}_{\alpha}^{\psi} \chi_{\mu} | M_{\mathbf{p} \mathbf{q}} | \bar{c}_{\beta}^{\psi} \chi_{\nu} | \rangle \tag{Be-4} \]

From before if \( \bar{\Phi} = \bar{\Phi}_0 \), then \( \langle \psi | \sum_{\mathbf{p} \mathbf{q}} M_{\mathbf{p} \mathbf{q}} | \psi_0 \rangle = 0 \). From Bb-12 (or by just expanding) \( \langle \bar{c}_{\alpha}^{\psi} \chi_5 | \sum_{\mathbf{p} \mathbf{q}} M_{\mathbf{p} \mathbf{q}} | \bar{c}_{\beta}^{\psi} \chi_0 | \rangle = \sum_{\mathbf{p} \mathbf{q}} \frac{\delta_{\alpha \beta}}{2 \sqrt{2 \mathbf{p} \mathbf{q}}} \langle \phi | \sum_{\mathbf{p} \mathbf{q}} M_{\mathbf{p} \mathbf{q}} | \psi_0 \rangle \bar{c}_{\alpha}^{\psi} \bar{\Phi}_0 \rangle \)

Use Ba-12 for \( \bar{c}_{\alpha}^{\psi} \chi \) and Ba-8 for \( c_{\alpha}^{\psi} \chi \), then

\[ Y = \sum_{\bar{\Phi}_{\alpha}, \bar{\Phi}_{\beta}, \bar{\Phi}_{\mu}} \frac{\delta_{\alpha \beta}}{2 \sqrt{2 \mathbf{p} \mathbf{q}}} \langle \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \bar{c}_{\beta}^{\psi} \chi_{\nu} | \rangle \bar{u}_{\mu \nu} \langle \bar{c}_{\alpha}^{\psi} \chi_{\mu} | \sum_{\mathbf{p} \mathbf{q}} M_{\mathbf{p} \mathbf{q}} | \bar{c}_{\beta}^{\psi} \chi_{\nu} | \rangle \]

\[ \bar{\Phi} = \sum_{\alpha} \bar{c}_{\alpha} \bar{\Phi}^\psi \chi^\psi \]

Let
$$G_{\rho}^{\alpha} \Phi^I = \sum_{\rho} \sum_{\rho} \left( O_{\rho} \Phi^I \right) \left( O_{\rho} \Phi^I \right) = \sum_{\rho} \sum_{\rho} \left( O_{\rho} \Phi^I \right) \left( O_{\rho} \Phi^I \right)$$

Thus,

$$\sum_{\rho} \sum_{\rho} \left( O_{\rho} \Phi^I \right) \left( O_{\rho} \Phi^I \right) = \sum_{\rho} \sum_{\rho} \left( O_{\rho} \Phi^I \right) \left( O_{\rho} \Phi^I \right)$$

So,

$$Y = \sum_{\rho} \sum_{\rho} \left( O_{\rho} \Phi^I \right) \left( O_{\rho} \Phi^I \right)$$

But, for $$\left( \frac{\Phi^I}{M_{pq}} \right) \neq 0$$, $$\Phi^I$$ must match $$\Phi^I$$ in all positions excepting $$p$$ and $$q$$. Let $$\Phi^I = \sum_{\rho} \Phi^I \left( \rho \Phi^I \right) \left( \rho \Phi^I \right)$$ differ from $$\Phi^I$$ only by having $$\Phi^I$$ rather than $$\Phi^I$$ in position $$p$$ and $$\Phi^I$$ rather than $$\Phi^I$$ in position $$q$$ (of course $$\Phi^I$$ may equal $$\Phi^I$$, etc.); then

$$Y = \sum_{\rho} \sum_{\rho} \left( \frac{\Phi^I}{M_{pq}} \right) \left( \frac{\Phi^I}{M_{pq}} \right)$$

Define

$$E_{\rho}^{\rho} (\rho \rho) = \sum_{\rho} \Phi^I \left( \rho \Phi^I \right) \left( \rho \Phi^I \right) \left( \rho \Phi^I \right)$$

the sum over $$k$$ and $$m$$ only goes over the $$\Phi^I$$ and $$\Phi^I$$ such that

$$G_{\rho}^I \Phi^I \neq 0$$

and also for which the set $$\Phi^I$$ and $$\Phi^I$$ is not equal to the set $$\Phi^I$$ and $$\Phi^I$$ (see the remark after Be-4). Then

$$Y = \sum_{\rho} \sum_{\rho} \left( \frac{\Phi^I}{M_{pq}} \right) \left( \frac{\Phi^I}{M_{pq}} \right)$$

There are three possibilities for $$E_{\rho}^{\rho} (\rho \rho)$$:

- both $$i$$ and $$i$$ are doubly-occupied, one is doubly-occupied, or none is doubly-occupied; I use FF, FM and HH to denote these three possibilities (F is for full and H is for half-full).

Let: $$k > i_N$$ denote that $$k$$ can be any orbital not contained in $$\Phi^I$$

$$k > i_N$$ denote that $$k$$ can be any orbital not doubly-occupied in $$\Phi^I$$

(i.e., $$k$$ can be a singly-occupied orbital in $$\Phi^I$$ or else an orbital not contained in $$\Phi^I$$)
Neither \( k \) nor \( m \) can be a doubly-occupied orbital of \( \Phi \) since \( \Phi \) would have the same orbital three times and thus \( G_1 \Phi \chi = 0 \) by THM Ab-1, and \( k \) and \( m \) cannot be singly-occupied orbitals of \( \Phi_0 \) other than \( i_p \) or \( i_q \) since we would have \( n = n_0 + 1 \) (\( \Phi \) has \( n_0 \) + 1 doubly-occupied orbitals) and \( G_1 \Phi \chi = 0 \) by THM Ab-2. Also \( k \neq m \) since in all cases other than above get \( n = n_0 + 1 \).

\[
\Theta_{p}^{H \Phi} (r, \theta) = \sum_{\Phi_{m}^{+}} \Phi_{p}^{(n)} \Phi_{m}^{(n)} < \Phi_{p}^{(n)} \Phi_{m}^{(n)} | M_{1} \Phi_{m}^{+} >
\]

\[
\Theta_{p}^{H \Phi} (r, \theta) = \sum_{\Phi_{m}^{+}} \Phi_{p}^{(n)} \Phi_{m}^{(n)} < \Phi_{p}^{(n)} \Phi_{m}^{(n)} | M_{1} \Phi_{m}^{+} >
\]

\[
\Theta_{p}^{I \Phi} (r, \theta) = \sum_{\Phi_{m}^{+}} \Phi_{p}^{(n)} \Phi_{m}^{(n)} < \Phi_{p}^{(n)} \Phi_{m}^{(n)} | M_{1} \Phi_{m}^{+} >
\]

The other possibilities yield either some orbital triply-occupied or else result in \( n > n_0 \) (note that \( n < n_0 \) allows \( G_1 \Phi \chi = 0 \) and is thus acceptable).
\[ \Theta_{\lambda_p \lambda_q}^{FF}(p, q) = \sum_{m, \lambda_m} \bar{\phi}_m(p) \bar{\phi}_m(q) \langle \psi_m | \lambda_m \rangle \langle \lambda_m | \mu \rangle \]

\[ \begin{cases} \bar{\psi} > N_j \quad m > N_j \\
\bar{\psi} > N_j \quad \bar{2} m < m \leq N_j \\
\bar{2} m < m \leq N_j \quad m > N_j \\
\bar{2} p > \lambda_m \quad m = \lambda_p \\
\bar{2} p > \lambda_{2 m_0} \quad \bar{2} m_0 \quad \bar{2} m_0 < m < N_j \quad m > N_j \end{cases} \]

**DOUBLY EXCITED**

**SINGLY EXCITED**

Now we partition \( W \) (from Be-3) in the same way

\[ W = \frac{1}{\hbar_0 - E_0} \sum_{\bar{p} \bar{q}} \langle \phi_{\bar{p}} + \phi_{\bar{q}} | \psi_\xi \rangle \langle \xi | \phi_{\bar{p}} + \phi_{\bar{q}} \rangle = \langle \phi_{\bar{p}} + \phi_{\bar{q}} | \psi_\xi \rangle \langle \psi_\xi | \phi_{\bar{p}} + \phi_{\bar{q}} \rangle \]

where \( \frac{\phi_{\bar{p}} + \phi_{\bar{q}}}{\hbar_0 - E_0} \). The development proceeds as for \( Y \) except that since \( \tilde{K} \) is a one-electron operator, the excited \( \Phi' \) in Be-5 can differ from \( \Phi_0 \) in at most one place. \( \Phi' \) must differ from \( \Phi_0 \) in at least one place since

\[ \langle \tilde{G}_\xi \Phi_0 \rangle \langle \tilde{K}_\xi \Phi_0 \rangle = 0 \]

(remember \( \tilde{G}_\xi \Phi_0 = \sum_{i} \tilde{G}_\xi \Phi_i \)). Instead of Be-6 we get

\[ W = \sum_{\bar{p} \bar{q}} \langle \phi_{\bar{p}} + \phi_{\bar{q}} | \tilde{G}_\xi \Phi_0 \rangle \langle \tilde{G}_\xi \Phi_0 | \phi_{\bar{p}} + \phi_{\bar{q}} \rangle \]

We obtain

\[ W = \sum_{\bar{p} \bar{q}} \langle \phi_{\bar{p}} + \phi_{\bar{q}} | \phi_{\bar{p}} \phi_{\bar{q}} \rangle \langle \phi_{\bar{p}} \phi_{\bar{q}} | \Phi_0 \rangle \langle \Phi_0 | \phi_{\bar{p}} + \phi_{\bar{q}} \rangle \]

where

\[ \Theta_{\lambda_p \lambda_q}^{HH}(p, q) = \sum_{m, \lambda_m} \bar{\phi}_m(p) \bar{\phi}_m(q) \langle \psi_m | \lambda_m \rangle \langle \lambda_m | \mu \rangle \]

Analyzing as before

\[ \Theta_{\lambda_p \lambda_q}^{HH}(p, q) = \sum_{m, \lambda_m} \bar{\phi}_m(p) \bar{\phi}_m(q) \langle \psi_m | \lambda_m \rangle \langle \lambda_m | \mu \rangle \]

(De-11)

\[ \Theta_{\lambda_p \lambda_q}^{HH}(p, q) = \sum_{m, \lambda_m} \bar{\phi}_m(p) \bar{\phi}_m(q) \langle \psi_m | \lambda_m \rangle \langle \lambda_m | \mu \rangle \]

(De-12)

\[ \Theta_{\lambda_p \lambda_q}^{HH}(p, q) = \sum_{m, \lambda_m} \bar{\phi}_m(p) \bar{\phi}_m(q) \langle \psi_m | \lambda_m \rangle \langle \lambda_m | \mu \rangle \]

(De-13)

Next we need the relation between \( \langle km | \psi_i \rangle \) and \( \langle km | \tilde{K} \psi_i \rangle \).
69

\[ \begin{align*}
\text{for } k \neq 1 \quad p_j & \quad \langle p_g | \alpha | p_j \rangle = -\langle \phi_1 g | \alpha | p_j \rangle \\
\langle p_g | \alpha | 1 p_j \rangle & = -\langle \phi_1 g | \alpha | 1 p_j \rangle \\
\langle \alpha | \beta | p_j \rangle & = -\langle \phi_1 g | \beta | p_j \rangle - U_{\beta}(p_j) \langle \phi_1 g | 1 p_j \rangle
\end{align*} \] (Be-14)

Let \( \Theta = \Theta' - \Theta'' \), then \( X_1 = \sum_{n_0} \frac{C_{\beta'}^{\gamma}}{C_{\alpha'}^{\gamma}} \langle \phi_1 | \beta | \gamma \rangle \).

Comparing Be-11, 12, 13 and Be-7, 8, 9 and using Be-14, we see that if

\[ | C_{\alpha'}^{\gamma} \frac{\phi_1}{\beta} \gamma \rangle = U_{\beta}(p_j) \langle C_{\alpha'}^{\gamma} \frac{\phi_1}{\beta} \gamma \rangle \] (Be-15)

then the singly-excited states of \( \Theta'' \) would cancel those of \( \Theta' \) leaving \( \Theta'' \) with no singly-excited states. Let \( \sum_{n_0} \frac{C_{\beta'}^{\gamma}}{C_{\alpha'}^{\gamma}} \langle \phi_1 | \beta | \gamma \rangle = \langle \phi_1 | \beta | \gamma \rangle \). Let \( \frac{\phi_1}{\beta} \gamma \) be the part of \( \Theta'' \) containing only doubly-excited states and let \( \Theta' = \Theta' + \Theta'' \). Then,

\[ X_1 = \sum_{n_0} \frac{C_{\beta'}^{\gamma}}{C_{\alpha'}^{\gamma}} \langle \phi_1 | \beta | \gamma \rangle + \]

\[ + \sum_{n_0 < 2n_0} \sum_{n_0 < 2n_0} \left[ \sum_{n_0 < 2n_0} \langle \phi_1 | \beta | \gamma \rangle \right] \sum_{\gamma > 2n_0} \langle \phi_1 | \beta | \gamma \rangle \langle \phi_1 | \beta | \gamma \rangle \left[ \sum_{\gamma > 2n_0} \langle \phi_1 | \beta | \gamma \rangle \right]
\]

We are considering the case of \( \alpha = \sum_{J^2} J^2 \) (maximum spin for the configuration); for this case \( G_i^0 \Omega \chi = \sum_{J^2} G_i^0 \Omega \chi \). Consider the singly-excited states in the above equation; if \( \phi_1 \) is excited then

\[ \sum_{n_0} \frac{C_{\beta'}^{\gamma}}{C_{\alpha'}^{\gamma}} \langle \phi_1 | \beta | \gamma \rangle = 0 \quad \text{if } J \neq 1 \] for the same reason as above.
Thus,

$$X_1 = C_{\alpha_0} \left[ \sum_{p \sigma} G_{\alpha_0 \beta_0 \alpha \beta_0 \phi_{p\sigma} \psi_{p\sigma}} + \sum_{p < m_0} \sum_{q > m_0} \sum_{\gamma \delta} \langle \varepsilon_{\gamma \delta} \mid \gamma \delta \phi_{p\sigma} \psi_{p\sigma} \rangle \right]^{1/2} \sum_{\alpha_1 \beta_1} \langle \alpha_1 \beta_1 \phi_{p\sigma} \psi_{p\sigma} \rangle \langle \alpha_1 \beta_1 \mid \varepsilon_{\gamma \delta} \phi_{p\sigma} \psi_{p\sigma} \rangle \langle \alpha_1 \beta_1 \mid \varepsilon_{\gamma \delta} \phi_{p\sigma} \psi_{p\sigma} \rangle$$

(Be-16)

if only closed shells are present then

$$X_1 = C_{\alpha_0} \left[ \sum_{p \sigma} G_{\alpha_0 \beta_0 \alpha \beta_0 \phi_{p\sigma} \psi_{p\sigma}} \right]$$

(Be-17)

Using Be-7, 8, 9 and remembering that \(\Theta\) contains only double excitations,

---

* The Brillouin theorem (Brillouin (1933), Møller and Plesset (1934), Löwdin (1959) p. 283) states that the matrix elements of total energy and thus of the perturbing energy between the Hartree-Fock configuration and the singly-excited (Slater determinant) configurations vanish identically. One might wonder how this is consistent with the appearance of non-zero terms in Be-16 for singly-excited states. The answer is that the Brillouin theorem discusses spin-orbitals; so, the situation where the orbital stays the same and the spin changes is considered to be an excitation. Thus, if \(\phi_{\alpha}\) is a singly-occupied orbital and associated with the spin \(\alpha\) in the Slater determinant and if \(\phi_{\beta}\) is associated with spin \(\beta\), then the excited state obtained by replacing \(\phi_{\alpha}\) with \(\phi_{\beta}\) and \(\phi_{\beta}\) with \(\phi_{\alpha}\) is considered (in the Brillouin theorem) as a double excitation although only one orbital is changed and the total spin may remain unchanged (e.g., for \(\alpha = \{^{2-1}_0, \ldots, e\}^{1/2}\) (maximum spin) and \(M = S\) the excited state is a mixture of \(S\) and \(S + 1\)). Thus, the Brillouin theorem is somewhat emptied of significance than is usually thought. A contributor to this common misunderstanding is the common error in assigning a definite spin to a definite orbital as discussed in the footnote on p. 8d-8. Note that for the situation with all orbitals doubly-occupied then no cases as in the example can occur and, in fact, only matrix elements to excited states differing in two orbitals can occur. This latter result coupled with the fact that most calculations are done on systems with all orbitals doubly-occupied is probably why the above defect in the Brillouin theorem has not been noticed earlier.
\[
\langle \Phi_\text{H}(p) | \Theta_{\text{HH}}^{\text{FH}}(p, q) \rangle = 0, \text{ i.e., } \Theta_{\text{HH}}^{\text{FH}} \text{ is orthogonal to all occupied states}
\]

\[
\langle \Phi_\text{H}(p) | \Theta_{\text{FF}}^{\text{FH}}(p, q) \rangle = 0 \text{ otherwise}
\]

\[
\langle \Phi_\text{H}(p) | \Theta_{\text{HH}}^{\text{FF}}(p, q) \rangle = 0 \text{ otherwise}
\]

\[
\langle \Phi_\text{H}(p) | \Theta_{\text{FF}}^{\text{FF}}(p, q) \rangle = 0 \text{ otherwise}
\]

If only closed shells are present, then

\[
\langle \Phi_\text{H}(p) | \Theta_{\text{FF}}^{\text{FF}}(p, q) \rangle = 0 \text{ i.e., } \Theta_{\text{FF}} \text{ is orthogonal to all occupied states}
\]

\[
\langle \Phi_\text{H}(p) | \Theta_{\text{FF}}^{\text{FF}}(p, q) \rangle = 0
\]

One cannot arbitrarily require equation Be-18 to be true, but it is true for closed shell systems (because of the Pauli principle). This is the reason that STANNOCH (1962b, p. 3200) is able to require it (he deals with closed shell). Unfortunately, such a relation is not true for non-closed shell systems. For closed-shell systems

\[
\Theta_{\text{FF}}^{\text{FF}} = \sum_{m, n > N} \sum_{m, n > N} \langle \phi_m | \mu_{p} + \nu_{q} - \phi_n \rangle
\]

where \( \mu_p \) is the one-electron Hartree-Fock operator and \( \phi_n \) is one-electron Hartree-Fock energy. This is equivalent to (if the equation has a solution)
since with the second requirement the sum in Be-19 can be taken over the complete set. The inhomogeneous equation has a solution only if the inhomogeneity is orthogonal to the solutions of the homogeneous equation. Since

$$\langle i; i \mid m_{ij} \mid i; i \rangle_{pq} = J_{pq} + U_{i(p,q)}^{K} - J_{pq} = U_{i(p,q)}^{K} \neq 0 \quad \text{and} \quad \langle i; i \mid m_{ij} \mid i; i \rangle_{pq} = \delta_{i(p,q)}^{K} - J_{pq} \neq 0,$$

then the inhomogeneous equation for $\Theta_{i; i}$ has a solution, but for $\Theta_{i; i}$ where $i \neq i$, it does not. The reason for wanting a differential equation for $\Theta_{i; i}$ is that the expansions of the form Be-19 are slowly convergent with a significant contribution from the continuum (e.g., SCHULL and LÖWDIN (1955)); whereas, for two-particle differential equations quite accurate results can be obtained. However, the $\Theta_{i; i}$ terms are likely to be very much more important than the terms for $i \neq i$ due to the two electrons being forced to be in exactly the same spatial state in the former (LÖWDIN (1959)). Even if we had a differential equation for $\Theta_{i; i}$, it probably would be worth solving only for $\Theta_{i; i}$, and the other less important terms, if included at all, would be estimated by a "short" expansion. If the Roothaan procedure is used to solve the Hartree-Fock equations, the left-over molecular orbitals (we have $N-n$ molecular orbitals but use $N>N-n$ basis functions, after diagonalizing the matrix we select the $N-n$ ones with lowest $\epsilon_{pp}$ as the Hartree-Fock orbitals the others are the left-over or virtual orbitals) can be used for the expansion of $\Theta_{i; i}$ for $i \neq i$ (see LÖWDIN (1955b) p. 1504).
For systems with half-closed shells then similarly to the above case

\[
(h_p^1 + h_p^2 - c_p^r \phi_g^r) \phi_p^q(\phi_{p^q}) = m_p^q \tau_p^r(\phi_{p^q}), \quad \text{now } \tau_p^r = \phi_g^r \quad (\text{Be-21})
\]

where we require that \( \langle \phi_g^r | \phi_p^q \rangle = 0 \) for all doubly-occupied \( i \) and \( \langle \phi_g^r | \phi_p^q \rangle = 0 \) for all occupied states.

The second and third order energies can be obtained from \( X_1 \) (SINANOĞLU (1961a)).

\[
E_2 = \langle \phi_1 | h_{12} | X_1 \rangle
\]

\[
E_3 = \langle \phi_1 | h_{12} | X_1 \rangle
\]

Let \( X_1 \) involve only singly-excited states and \( X_0 \) involve only doubly-excited states (see Be-16). \( X_1 \) is orthogonal to \( \phi_1 \), so \( \langle \phi_1 | X_1 \rangle = 0 \).

Also, \( \langle \phi_1 \phi_{p^q} \rangle = 0 \) since \( V_p \) is a one-electron operator.

\[
E_2^V = \langle \phi_1 | \sum \phi_i \phi_i \phi_{p^q} \phi_{p^q} \rangle = \sum \frac{\sigma_n^m \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle}{r_{\phi P}^2} \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle
\]

\[
= \sum \frac{\sigma_n^m \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle}{r_{\phi P}^2} \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle
\]

where \( \sum \) goes over all \( \varphi \) giving different \( \varphi_0 \).

\[
E_2^{SV} = \langle \phi_1 | \sum \varphi_i \phi_i \phi_{p^q} \phi_{p^q} \rangle = \sum \frac{\sigma_n^m \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle}{r_{\phi P}^2} \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle
\]

if \( k > i, \) then get zero since \( \sigma \) must be \( \leq \) and \( U_{j i} = 0 \)

if \( i, k \leq i \), let \( k = i, \) then \( r = p \) or \( r = t \);

if \( r = p, \) we must have \( \sigma = \epsilon, \) if \( r = t \) we must have \( \sigma = (p, t) \).

Thus,

\[
E_2^{SV} = \sum \frac{\sigma_n^m \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle}{r_{\phi P}^2} \langle \phi_i | \varphi_0 \rangle \langle \phi_i | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle \langle \phi_{p^q} | \varphi_0 \rangle
\]
Also,
\[ E_{2g} = \langle X^k \rangle \langle \delta_{p\perp} \rangle \]
\[ = \sum_{p, q \perp \Delta} \sum_{\phi_{pq} \in \mathcal{G}_{pq}} \sum_{\phi_{pq} \in \mathcal{G}_{pq}} \langle \delta_{p\perp} \rangle \langle \delta_{p\perp} \rangle \]
\[ \times \text{if } k > i_N \text{ let } p = r \text{ then } \mathcal{E} = \mathcal{E} \text{ or } (p, t); \]
\[ \quad \text{if } k = i_m \text{ where } 2n < m \leq N \text{ and } p = r \text{ then } \mathcal{E} = \mathcal{E}, (p, t), (m, p), \text{ or } \]
\[ (m, p)(p, t) \]

Thus,
\[ E_{2g} = \frac{1}{2} \sum_{p \leq 2n} \sum_{\phi_{pq} \in \mathcal{G}_{pq}} \sum_{\phi_{pq} \in \mathcal{G}_{pq}} \langle \delta_{p\perp} \rangle \langle \delta_{p\perp} \rangle \times \]
\[ \sum_{\phi_{pq} \in \mathcal{G}_{pq}} \langle \delta_{p\perp} \rangle \langle \delta_{p\perp} \rangle \]
\[ + \sum_{2n < m \leq N} \left( \langle \delta_{p\perp} \rangle \langle \delta_{p\perp} \rangle \right) \]
\[ \text{The analysis for } E_3 \text{ is similar with more terms.} \]
BF: Construction from the $G_i^j$ of Basis Functions of Irreducible Representations of the Spatial Symmetry Group

The $G_i^j$ are eigenfunctions of $S^2$ but are not basis functions of the irreducible representations of the spatial symmetry group. Let us consider a case of taking linear combinations of the $G_i^j$ to get wave functions which are basis functions of the irreducible representation of the spatial symmetry group. For convenience we take the spatial symmetry group to be SO(3) (the proper rotation group in three dimensions).

We consider a set of one-electron states all of which belong to the same irreducible representation or shell of O(3) (e.g., L = 2 or d functions). The usual procedure starts with a Slater determinant which happens to be an eigenfunction of $L^2$ and $S^2$. It is automatically an eigenfunction of $L_z$ and $S_z$ since the one electron functions are eigenfunctions of $L_z$ and $S_z$. Consider the maximum possible $S_z$ consistent with the given number of occupied states of the shell, then the Slater determinant of this maximum $S_z$ and the maximum possible $L_z$ for this $S_z$ is an eigenfunction of $S^2$ and $L^2$ (if we apply $S^+$ or $L^+$ we get zero). Also, if the next lower value of $S_z$ allows a higher maximum value of $L_z$, then the Slater determinant for this next lower $S_z$ and highest $L_z$ is also an eigenfunction of $S^2$ and $L^2$, etc. for still lower $S_z$ (no need to consider $S_z$ less than zero). With each of these eigenfunctions of $S^2$
and $L^2$, we can get the other states of the same $S$ and $L$ by applying $S^-$ and $L^-$ (the lowering operators). The $S$ and $L$ for which the Slater determinant of maximum $S_z$ and $L_z$ is not an eigenfunction of $S^2$ and $L^2$ can be found by orthogonalizing to the previously determined states. (A problem occurs when more than one state with the same $S$ and $L$ occurs).

We are really constructing zero-order eigenfunctions (neglecting electron-electron interactions) for a degenerate problem, but since the

\[ \sum_{L_1} \frac{\varepsilon^2}{L_1^2} \]

is invariant under spatial rotation and spin transformation, then only states which have the same values of $S$ and $L$ can interact.

Neglecting eigenfunctions coming from other configurations (i.e., a different set of one-electron functions, e.g., not all 3d functions if the above set is for 3d functions) the perturbing Hamiltonian is diagonal for cases where only one state has the same values of $L$ and $S$, but we still must diagonalize the submatrix for cases where there is more than one state with same $L$ and $S$. Thus, in the latter case the wave function is not determined by symmetry. I have no quarrel with the usual procedure except that the wave functions are linear combinations of Slater determinants which are unnecessarily complicated to work with.

The use of the $G_1^\alpha$ allows the ultimate possible simplification of the wave functions; the simplification is small in some cases but significant in others. I will find some of the eigenfunctions of the $d^3$ configuration to demonstrate the method and compare it with the usual method. For $S = \frac{3}{2}$ the new method is no better than the usual approach since the Slater determinants automatically factor into a space and spin part and the $O^{[137]}_{11}$ is just the determinant of space states.

Consider $S = \frac{1}{2}$ then $\alpha = [2,1] = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. We wish the maximum $L_z$;
if \( \overline{\Phi} = \phi_2 \phi_2 \phi_1 \), then \( G_1^{\overline{\Phi}} = O(\text{THM Ab-1}) \), but for \( \overline{\Phi} = \phi_2 \phi_2 \phi_1 \)

\( G_1^{\overline{\Phi}} \neq O \). Thus we will consider the \( L = 5 \) eigenstates. Recall that

\[
G_1^{\overline{\Phi}}X = \sum \frac{1}{r} \left( G_0^{\overline{\Phi}} \right) (n_{\overline{t} }) X. (54)
\]

Since the \( L^+ \) operates only on the \( O_{rl} \overline{\Phi} \), we need only consider the \( O_{rl} \overline{\Phi} \) (remember that states of different \( M_s \) are obtained by simply using a different number of \( \alpha \) and \( \beta \) in the ordered \( X \)).

\( L = 5 \),

\( L_z = 5 \): (55)

\[
O_{rl} \phi_2 \phi_2 \phi_1 = 0_{rl} (13) \phi_1 \phi_1 \phi_2 = U_{11} (13) \phi_1 \phi_1 \phi_2
\]

\( L_z = 4 \): (54)

\[
O_{rl} \phi_2 \phi_2 \phi_1 = \sqrt{\frac{1}{10}} O_{rl} \phi_1 \phi_2 \phi_1 = \frac{1}{\sqrt{10}} (2 \phi_2 \phi_1 + 2 \phi_1 \phi_2 + \sqrt{5} \phi_2 \phi_2)
\]

but we want the standard order \( \phi_1 \phi_1 \phi_2 \), thus

\[
O_{rl} \phi_2 \phi_2 \phi_1 = 0_{rl} (13) \phi_1 \phi_1 \phi_2 = U_{11} (13) 0_{rl} \phi_1 \phi_1 \phi_2
\]

\( L_z = 3 \): (53)

\[
O_{rl} \phi_2 \phi_2 \phi_1 = \frac{1}{\sqrt{10}} \left[ -2 \phi_2 \phi_2 \phi_2 - 2 \phi_2 \phi_1 \phi_1 + \sqrt{5} \phi_2 \phi_1 \phi_1 + \sqrt{5} \phi_2 \phi_2 \phi_2 \right]
\]

\[
- \frac{1}{\sqrt{10}} \left[ -2 \phi_2 \phi_2 \phi_2 - 2 \phi_2 \phi_1 \phi_1 - 2 \phi_1 \phi_2 \phi_1 + \phi_1 \phi_1 \phi_1 \right]
\]

\( L = 3 \),

\[
O_{rl} \phi_2 \phi_2 \phi_1 = \frac{1}{\sqrt{10}} \left[ -2 \phi_2 \phi_2 \phi_2 - 2 \phi_2 \phi_1 \phi_1 + \sqrt{5} \phi_2 \phi_1 \phi_1 + \sqrt{5} \phi_2 \phi_2 \phi_2 \right]
\]

\( \text{it should be noted that the norm of } O_{rl} \phi_2 \phi_2 \phi_1 \]

where \( k = 1 \neq m \) is twice as much as it is if \( k > 1 > m \); thus,

the norm of (53) is the same as that of (54) and (55) (see Ba-1)
\[ L_z = 2: \quad (52)_{r1} = \frac{1}{\sqrt{2+\sqrt{5}}} \left[ v_1 c_{11}(2 \phi_1 \phi_1 + \phi_1 \phi_2 + \phi_1 \phi_3 + \sqrt{2} \phi_1 \phi_4) + v_2 c_{12}(2 \phi_2 \phi_2 + \phi_3 \phi_3 + \sqrt{2} \phi_3 \phi_4) \right] \]

\[ = \frac{1}{\sqrt{\sqrt{20}}} \left[ 2 c_{11}(2 \phi_1 \phi_1 + 6 \phi_1 \phi_3 + \sqrt{2} \phi_1 \phi_4) + 2 c_{12}(2 \phi_2 \phi_2 + \sqrt{2} \phi_3 \phi_4) \right] \]

This is certainly enough to demonstrate the technique. The advantages of the new method result from the linear independence and, a fortiori, the orthogonality of the various terms in the expansion of \((L L_z)_{r1}\) and the \((L L_z)_{r1} \) terms occurring in \( C_{11}^{pr} \). This is in contrast to linear dependencies and non-orthogonalities and non-separated spin and space states in the usual sum of Slater determinants (cf. CONDON and SHORTLEY (1963) p. 227).
Section C

Different Orbitals For Different Spin -- The GF Method

Introduction

We will now develop the DODS approach of obtaining accurate approximations to the ground state energy and wave function of the Hamiltonian A-1. The $G_i^\alpha$ operator will be used in order to consider $N$-electron wave functions which satisfy Pauli's principle and are eigenfunctions of $S^2$. The best possible $C^\alpha_f \Phi \chi$ wave function is found for each spin and number of electrons by using the variational principle $(\Phi = \phi_{i_a} \phi_{i_b} \cdots \phi_{i_m} \phi_{i_n})$, where the $\phi_{i_a}$ are mutually orthonormal orbitals and the $\phi_{i_b}$ are mutually orthonormal orbitals. The Roothaan technique or expanding the orbitals in terms of a large fixed basis set of functions is used to obtain one-electron equations for the $\phi_{i_a}$ (the $\phi_{i_a}$ are eigenfunctions of the one-electron Hamiltonian $H^a$ and the $\phi_{i_b}$ are eigenfunctions of $H^b$). The $\phi_{i_a}$ and $\phi_{i_b}$ which lead to the minimum total energy are called the GF orbitals and the one-electron equations for the GF orbitals (e.g., $H^a \phi_{i_a} = \varepsilon_{i_a} \phi_{i_a}$) are called the GF equations. In section Ca the first- and second-order density matrices for $G_f^\alpha \Phi \chi$ are derived; these are used in obtaining the expectation value of the energy in section Cb. In section Cb the variational principle is applied to obtain the GF equations, which yield the best possible $G_f^\alpha \Phi \chi$. In section Cc the significance of the GF orbitals is discussed. In section Cd the defect of the GF method is pointed out and discussed. In section Ce the restrictions placed on the GF orbitals by the spatial symmetry is discussed for the two particle case.

The DODS scheme was used by Harris and Taylor (1963) for $H_2$. The
single configuration used by Harris and Taylor did not use general orbitals; however, they chose very good (elliptic) parametric orbitals and were able to remove over half the correlation energy for $\frac{1}{2}$ with the Hartree-Fock and exact energies given in PRAGA and RANSIL (1961). We would expect the GF method to allow even more improvement.

Some other approaches have been developed which eliminate part or all of the SODS restriction (but which involve other restrictions). The alternate orbital, unrestricted Hartree-Fock and extended valence bond methods will be briefly described. The alternate orbital method is sketched in section GF.

The unrestricted Hartree-Fock method (we call this UHF) (e.g., PRATT (1956), WATSON (1959), WATSON and FREEMAN (1960)): In the UHF a single Slater determinant is used for the wave function but the set of orbitals with $\alpha$ spin are varied independently of the orbitals of $\beta$ spin. Such a wave function is not an eigenfunction of $S^2$ with the result that each orbital has a definite spin associated with it (see footnote on page 6b-7). Apparently, most of the UHF calculations have dealt with atoms and all have used additional approximations. Apparently, most have assumed that the atomic orbitals are basis functions of the spatial symmetry group (i.e., transform as spherical harmonics) although some have discussed the possibility of having different radial functions for orbitals of the same $L$ but different $m_L$ (this means that the set of functions of different $m_L$ and same $L$ no longer form an irreducible representation of $SO(3)$). LÖWDIN (1959, p.314) has pointed out that the UHF method may not allow as much splitting of the doubly-occupied orbitals as would the general DODS solution. In addition, LÖWDIN (1959, p.315)
points out that since the varied UHF state is not an eigenfunction of $s^2$
then even if a pure spin component is projected from the UHF state, one
must be very careful in making physical interpretations of the UHF re-
sults.

Another approach which does not use the SODS restriction is the
method of solving for an approximate ground state of $H_2$ is the Heitler-
Londin method. In this method one approximates the wave function by

$$(\alpha(1)\beta(2) + \beta(1)\alpha(2))(\alpha(1)\beta(2) - \alpha(2)\beta(1))$$

where $\alpha(1)$ is an atomic orbital on atom $a$ and $\beta(1)$ is an atomic orbital on atom $b$. Then one chooses
some parametric form for the atomic orbitals (say an exponential) and
varies the parameters to minimize the energy. The extended valence bond
(EVB) method is an extension of Heitler-Londin method to systems more
complicated than $H_2$. In the EVB method one has one atomic orbital for
each bonding electron and atomic orbital for each pair of nonbonding
electrons. The nonbonding atomic orbitals are assumed to be orthogonal
to all other orbitals, each bonding atomic orbital is assumed to be or-
thogonal to all other atomic orbitals except one other (with which it
bonds). A particular combination of the orbitals and spins is used which
results in a total wave function of spin equal to zero (one of the $f^\alpha$
such functions). With these assumptions a fairly simple energy expres-
sion is obtained. A parametric form is assumed for the various atomic
orbitals, and the parameters are varied in order to minimize the energy.
Due to increased flexibility of the wave function, the EVB method yields
lower energies than methods involving parametric molecular orbitals.

* Dr. R. M. Pitzer pointed out this method to me.
(SLATER (1963) p.196). A configuration interaction based on the EVB method accounted for 80% of the binding energy of \( \text{H}_2\text{O} \) (MERRIFIELD (1961)). Note that since the orbitals are fixed to be the atomic type, they cannot vary as generally as in the GF method. The orthogonality assumptions provide another restriction in the EVB method as compared to the GF method; moreover, these orthogonality assumptions should lead to a rather complicated Lagrange multiplier matrix.
Ca: The First- and Second-Order Density Matrices

Let the wave function be $G_1^{\alpha} \Phi X$, where $\alpha = \left[ \frac{N}{2} + S, \frac{N}{2} - S \right]$. Define $n = \frac{N}{2} + S$. We let $\Phi$ contain $n = \frac{N}{2} + S$ orthogonal orbitals $\phi_{ia}$ and $N - n = m$ orthogonal orbitals $\phi_{ib}$. Let $S(ia;jb) = \int \phi_{ia}^* \phi_{jb} \, d\alpha$ (for SODS $S(ia;jb) = \delta_{ij}$, for alternate orbitals $S(ia;jb) = \lambda \delta_{ij}$).

We wish to determine the $\phi_{ia}$ and $\phi_{ib}$ such that the energy of the system is a minimum.

Let $\overline{\Phi} = \phi_{i1}^{(1)} \phi_{i2}^{(2)} \cdots \phi_{n}^{(N)} \bar{\phi}_{b1}^{(1)} \bar{\phi}_{b2}^{(2)} \cdots \bar{\phi}_{mb}^{(N)}$ \hspace{1cm} (Ca-1)

and $X = \alpha(1) \alpha(2) \cdots \alpha(N) \beta(m+1) \beta(m+2) \cdots \beta(N)$ \hspace{1cm} (Ca-2)

(we are here considering $M_s = S$). Then by THM Ab-2

$G_1^{\alpha} \Phi X = \sum_{i} G_f^{\alpha} \phi_{i} X$. We will consider $G_f^{\alpha} \phi_i X$ in the following.

The energy, $E = \frac{\langle G_f^{\alpha} \phi_i X | H | G_f^{\alpha} \phi_i X \rangle}{\langle G_f^{\alpha} \phi_i X | G_f^{\alpha} \phi_i X \rangle}$ \hspace{1cm} (Ca-3)

where $H = \sum_{\alpha=1}^{N} \sum_{\gamma=1}^{N} \frac{q_{\alpha\gamma}}{2} \phi_{\alpha} \phi_{\gamma}^* + \frac{1}{2} \sum_{\alpha=1}^{N} \sum_{\gamma=1}^{N} \sum_{\beta=1}^{N} \sum_{\gamma=1}^{N} \frac{q_{\alpha\beta} q_{\gamma\beta}}{2} \phi_{\alpha} \phi_{\gamma}^* \phi_{\beta} \phi_{\beta}^*$

The next steps are to evaluate the normalization integral and the first- and second-order density matrices for $G_f^{\alpha} \phi_i X$.

The normalization integral:

From (Ac-3) $\langle G_f^{\alpha} \phi_i X | H | G_f^{\alpha} \phi_i X \rangle = \alpha^* \langle \phi_i | \phi_i \rangle \langle X | \phi_i \rangle \langle \phi_i \phi_i^* | X \rangle$

From Eq. Da-4 $\langle \phi_i | \phi_i \rangle = \frac{(N+S)(N-S)}{N}$

$\Theta^* \langle \phi_i | \phi_i \rangle = \frac{1}{N} \sum_{\alpha=1}^{N} \sum_{\gamma=1}^{N} \frac{q_{\alpha\gamma}}{2} \phi_{\alpha} \phi_{\gamma}^*$

Due to the orthogonality among the $a$ states and among the $b$ states, then either $\phi_{ia}$ is in the $i$ position of $\overline{\Phi} = \Phi^\prime$ or else it is in a position $n$; similarly either $\phi_{ib}$ is in the $n$ position of $\Phi^\prime$ or $\overline{\Phi}$.

* But see section Ce.
else it is in a position \( \leq n \). If \( \tau \) takes no element \( \leq n \) to a position \( > n \), then \( \langle \mathcal{F}(\tau) \rangle \neq 0 \) only for \( \tau \in \). If \( \tau \) takes one element, say \( i \), from a position \( m \) to a position \( > n \), then it must take one element say \( j \) from a position \( > n \) to a position \( \leq n \); \( \langle \mathcal{F}(\tau) \rangle \neq 0 \) only for \( \tau = (i,j) \). If \( \tau \) takes two elements, \( i \) and \( j \), from positions \( \leq n \) to positions \( > n \); then it must take two elements, say \( p \) and \( q \), from positions \( > n \) to positions \( \leq n \); \( \langle \mathcal{F}(\tau) \rangle \neq 0 \) only for \( \tau = (i,p)(j,q) ; (i,j)(i,p)(j,q) ; (p,q)(i,j)(p,q) ; (i,j)(p,q)(i) \). If \( \tau \) takes \( r \) elements \( (p_1, p_2, \ldots, p_r) \) from positions \( \leq n \) to positions \( > n \), then it must take \( r \) elements (say, \( q_1, q_2, \ldots, q_r \)) from positions \( > n \) to positions \( \leq n \); \( \langle \mathcal{F}(\tau) \rangle \neq 0 \) only if \( \tau = \tau_a \tau_b \tau_r \) where

\[
\tau_r = (p_1, q_1)(p_2, q_2) \cdots (p_r, q_r)
\]

and \( \tau_a \) is any permutation of \( \{p_1, p_2, \ldots, p_r\} \) and \( \tau_b \) is any permutation of \( \{q_1, q_2, \ldots, q_r\} \) (the choice of the specific set of \( r \) disjoint transpositions for \( \tau_r \) is arbitrary, the permutations \( \tau_a \) and \( \tau_b \) take \( \tau_r \) through \( (r!)^2 \) different permutations for which \( \langle \mathcal{F}(\tau) \rangle \neq 0 \)). Now we evaluate

\[
U_{ff} \tau_a \tau_b \tau_r.
\]

From (Aa-1), \( U_{ff} \tau_a = S_{\tau_a} b_{\tau_a} \) and \( U_{ff} \tau_b = S_{\tau_b} b_{\tau_b} \); thus, \( U_{ff} \tau_a \tau_b \tau_r = b_{\tau_a} b_{\tau_b} U_{ff} \tau_r \), but from appendix II*

\[
U_{ff} \tau_r = \lambda. \quad \text{Thus},
\]

\[
\langle \mathcal{F}(\tau) \rangle = \sum_{\tau_a \tau_b} \sum_{\tau_r} \sum_{b_{\tau_a} b_{\tau_b} \tau_r} \langle \mathcal{F}(\tau) \rangle = \sum_{b_{\tau_a} b_{\tau_b}} \sum_{\tau_r} \langle \mathcal{F}(\tau) \rangle.
\]

But the integral factors into two integrals, one of which is

\[
\langle \mathcal{F}(\tau) \rangle = \langle \mathcal{F}(\tau_a \tau_b \tau_r) \rangle = S(p, a; q, b) \cdots S(p, a; q, b).
\]

The sum \( \sum_{b_{\tau_a} b_{\tau_b}} \) times the above is recognized as the determinant of the \( r \)-dimensional matrix \( S(p, a; q, b) \) which from (Aa-8) is written as

* You must note that the proof in appendix II is incomplete.
Thus,

$$\Theta <\Phi | \Omega_{\eta} | \Phi> = \sum_{r=0}^{\infty} \frac{(\eta)^{2}}{(r)!} \sum_{P_{r}} \left[ O_{i}^{(r)} S(P_{r}; g_{a}; b) \cdots S(P_{r}; g_{a}; b) \right] x \left[ O_{i}^{(r)} S(g_{a}; b) \cdots S(g_{a}; b) \right]$$

(A-4)

A standard order of the $p_i$ and of the $q_i$ is presumed to have been defined. Since $S(ia; jb) = S^\ast (jb; ia)$ each term in the sum of Ca-4 is real and positive. The $P_r$ is an abbreviation for the sum of the $p_i$ from 1 to n such that $p_1 < p_2 < \cdots < p_r$.

The first-order density matrix: From section 8b

$$\rho(1,1) \equiv \sum_{\eta} \left[ \Phi_{\eta}^{\ast} (1) \Phi_{\eta} (1) \right] = \sum_{\eta} \sum_{\eta'} \frac{(\eta')^{2}}{(\eta)^{2}} \Phi_{\eta'} (1) \Phi_{\eta}^{\ast} (1)$$

Let

$$\sum_{\eta} \Phi_{\eta}^{\ast} (1) \Phi_{\eta} (1) \sum_{\eta} \frac{(\eta')^{2}}{(\eta)^{2}} \Phi_{\eta'} (1) \Phi_{\eta}^{\ast} (1) D_{\eta \eta'}^{\eta \eta'}$$

(Ca-5)

where $\eta_i$ is the ith state in $\Phi$, I shall often omit the $\eta$. Using the same reasoning as in deriving the normalization integral

$$D_{\eta \eta'}^{\eta \eta'} = \sum_{r=0}^{\infty} \frac{(\eta)^{2}}{(\eta')^{2}} \sum_{P_{r}} \left[ O_{i}^{(r)} S(P_{r}; g_{a}; b) \cdots S(P_{r}; g_{a}; b) \right] x \left[ O_{i}^{(r)} S(g_{a}; b) \cdots S(g_{a}; b) \right]$$

(Ca-6)

I will always write $n$ as the upper limit of $r$, the limit is set by the smaller of the maximum number of elements in $P_r$ and $Q_r$. In the case of $D_{ia}$ this number is the smaller of $m$ and $n-1$ since $i \in P_r$. For $D_{ja}$,

$$\lambda \neq \mu \lambda \eta \mu$$

where $\lambda_i = (ia, ja)$ and $\lambda_2 = (ja, ja)$

$$D_{ia} = -\sum_{r=0}^{n-1} \frac{(\eta)^{2}}{(\eta')^{2}} \sum_{P_{r}} \left[ O_{i}^{(r)} S(P_{r}; g_{a}; b) \cdots S(P_{r}; g_{a}; b) \right] x \left[ O_{i}^{(r)} S(g_{a}; b) \cdots S(g_{a}; b) \right]$$

where the -1 is due to $\eta_{2} = -1$ and $(ia, ja)$ transposes $ia$ and $ja$. For $D_{jb}$,

$$\lambda \neq \mu \lambda \eta \mu$$

where $\lambda_i = (ia, jb)$

$$D_{ja} = -\sum_{r=0}^{n-1} \frac{(\eta)^{2}}{(\eta')^{2}} \sum_{P_{r}} \left[ O_{i}^{(r)} S(P_{r}; g_{a}; b) \cdots S(P_{r}; g_{a}; b) \right] x \left[ O_{i}^{(r)} S(g_{a}; b) \cdots S(g_{a}; b) \right]$$

where the -1 is due to $\eta_{2} = -1$ and $(ia, ja)$ transposes $ia$ and $ja$. For $D_{jb}$,
Thus,
\[ D_{ia, ja}^{ab} = \sum_{r=0}^{m} \frac{\binom{r+1}{m}}{\binom{r}{m}} \sum_{\substack{P_r, Q_r \in Pr, Qr \in Qr \in}} \left[ Q_i^{r+1} S^{ab} S^{ab} \ldots \right] \left[ Q_i^{1+2} S^{ab} S^{ab} \ldots \right] \]  
(Ca-8)

where the \( \binom{m}{m} \) occurs above because of the \( r+1 \) transpositions between sets \( a \) and \( b \). Equations Ca-5, 6, 7, 8 determine the first-order density matrix. Note that \( D_{ij} \) is Hermitian.

The second-order density matrix: From section Bb
\[ \varphi(1, 2; 1', 2') \equiv N(\text{m}) \left( S_{g, g} X \right) \left( S_{g, g} X \right) (dx_r, dx'_{r}) (dx_{r'}')(dx_{r''}) \]
\[ = \delta^{\mu} \left( x_{rg} x_{g} \right) \sum_{i, g} \left( \alpha_{i}, 1, 2, 2' \right) S_{g}^{\mu} S_{g}^{\mu} (dx_{r}) \]
\[ S_{g}^{\mu} S_{g}^{\mu} (dx_{r}) \equiv \sum_{\alpha, r}^{\mu} \left( \alpha_{g}, 2 \right) \phi_{\alpha}^{(2)}(g) \phi_{\alpha}^{(2)}(g) D_{\alpha, r} \]

The same approach as for the first-order density matrices is used. ( \( \varphi \) is shown below the \( D \), in some cases). Indices are equal only if the same letter is used (i.e., \( i \neq j \neq k \neq t \).

\[ D_{ia, ja}^{ab} = \sum_{r=0}^{m} \frac{\binom{r+1}{m}}{\binom{r}{m}} \sum_{\substack{P_r, Q_r \in Pr, Qr \in Qr \in}} \left[ Q_i^{r+1} S^{ab} S^{ab} \ldots \right] \left[ Q_i^{1+2} S^{ab} S^{ab} \ldots \right] \]

\[ D_{ia, ja}^{ab} = -\sum_{r=1}^{m} \frac{\binom{r+1}{m}}{\binom{r}{m}} \sum_{\substack{P_r, Q_r \in Pr, Qr \in Qr \in}} \left[ Q_i^{r+1} S^{ab} S^{ab} \ldots \right] \left[ Q_i^{1+2} S^{ab} S^{ab} \ldots \right] \]

\[ D_{ia, ja}^{ab} = \sum_{r=2}^{m} \frac{\binom{r+1}{m}}{\binom{r}{m}} \sum_{\substack{P_r, Q_r \in Pr, Qr \in Qr \in}} \left[ Q_i^{r+1} S^{ab} S^{ab} \ldots \right] \left[ Q_i^{1+2} S^{ab} S^{ab} \ldots \right] \]

\[ D_{ia, kb}^{ab} = \sum_{r=0}^{m} \frac{\binom{r+1}{m}}{\binom{r}{m}} \sum_{\substack{P_r, Q_r \in Pr, Qr \in Qr \in}} \left[ Q_i^{r+1} S^{ab} S^{ab} \ldots \right] \left[ Q_i^{1+2} S^{ab} S^{ab} \ldots \right] \]

\[ D_{ia, ja}^{ab} = -\sum_{r=0}^{m} \frac{\binom{r+1}{m}}{\binom{r}{m}} \sum_{\substack{P_r, Q_r \in Pr, Qr \in Qr \in}} \left[ Q_i^{r+1} S^{ab} S^{ab} \ldots \right] \left[ Q_i^{1+2} S^{ab} S^{ab} \ldots \right] \]

\[ D_{ia, ja}^{ab} = \sum_{r=0}^{m} \frac{\binom{r+1}{m}}{\binom{r}{m}} \sum_{\substack{P_r, Q_r \in Pr, Qr \in Qr \in}} \left[ Q_i^{r+1} S^{ab} S^{ab} \ldots \right] \left[ Q_i^{1+2} S^{ab} S^{ab} \ldots \right] \]
\[ D_{ia,kb}^{ia, kb} = \sum_{r=0}^{m} \left( \frac{r!}{r+1} \right)^2 \sum_{p, q} \left| C_{ij} \right|^2 S(p, q; g, h) \cdots \right|^2 \]

\[ D_{ia,kb}^{kb, ia} = \sum_{r=0}^{m} \left( \frac{r!}{r+1} \right)^2 \sum_{p, q} \left| C_{ij} \right|^2 S(p, q; g, h) \cdots \right|^2 \]

\[ D_{ia,jb}^{ia, jb} = -\sum_{r=0}^{m} \left( \frac{r!}{r+1} \right)^2 \sum_{p, q} \left[ (kb, gb) C_{il}^{il} S(p, q; g, h) \cdots \right] \left[ (gb, gb) C_{ij}^{ij} S(p, q; g, h) \cdots \right] \]

\[ D_{tb, ia}^{ia, tb} = -\sum_{r=0}^{m} \left( \frac{r!}{r+1} \right)^2 \sum_{p, q} \left[ (gb, gb) C_{il}^{il} S(p, q; g, h) \cdots \right] \left[ (gb, gb) C_{ij}^{ij} S(p, q; g, h) \cdots \right] \]

where we used \( \zeta = \zeta_2 \zeta_3 \zeta_4 \zeta_5 \zeta_1 \) and \( \zeta_2 = (jb, kb) \)

\[ D_{ja,kb}^{ia, kb} = \sum_{r=0}^{m} \left( \frac{r!}{r+1} \right)^2 \sum_{p, q} \left[ C_{il}^{il} S(p, q; gb) S(p, q; gb) \cdots \right] \]

where we used \( \zeta = \zeta_2 \zeta_3 \zeta_4 \zeta_5 \zeta_1 \)

\[ D_{tb, ja}^{ia, sb} = \sum_{r=0}^{m} \left( \frac{r!}{r+2} \right)^2 \sum_{p, q} \left[ C_{ij}^{ij} S(p, q; gb) S(p, q; gb) \cdots \right] \]

where we used \( \zeta = \zeta_2 \zeta_3 \zeta_4 \zeta_5 \zeta_6 \zeta_7 \zeta_8 \zeta_9 \zeta_10 \)
Note that $D_{kS_3, \bar{S}_4}^{\bar{S}_1, \bar{S}_4} = (D_{kS_3, \bar{S}_4}^{\bar{S}_1, \bar{S}_4})^* = D_{kS_3, \bar{S}_4}^{\bar{S}_1, \bar{S}_4}$ and

$D_{kS_3, \bar{S}_4}^{\bar{S}_1, \bar{S}_4} = -D_{kS_3, \bar{S}_4}^{\bar{S}_1, \bar{S}_4}$; using these relations all other $D$ are obtained from the above ones.

Using the preceding work

$$\frac{\langle \phi_{i \lambda} \phi_{k \bar{\lambda}} | \phi_{j \bar{\lambda}} \phi_{t \lambda} \rangle}{\beta^d \langle \phi_{k \bar{\lambda}} | \phi_{t \lambda} \rangle} = \sum_{\gamma=1}^{N} \left[ \delta_{\gamma} \langle \phi_{i \lambda} | \phi_{\gamma \bar{\lambda}} \rangle U_{\gamma}^{1/2} \right] \sum_{\delta=1}^{N} \langle \phi_{j \bar{\lambda}} | \phi_{\delta \lambda} \rangle D_{\bar{\lambda}, \gamma}^{\lambda, \bar{\lambda}}$$

(Ca-9)

where the $i, j, k, t$ denote the position of the state in $\Phi$. But,

$D_{k, t}^{\bar{\lambda}, \lambda} = -D_{k, t}^{\bar{\lambda}, \lambda}$; so, regardless of what we define the, so far undefined, $D_{k, t}^{\bar{\lambda}, \lambda}$ to be, then $D_{k, t}^{\bar{\lambda}, \lambda} = -D_{k, t}^{\bar{\lambda}, \lambda}$. Also,

$$\langle \phi_{i \lambda} | \phi_{k \bar{\lambda}} \rangle = \langle \phi_{i \lambda} | \phi_{k \bar{\lambda}} \rangle$$

; thus, we can write

$$\frac{\langle \phi_{i \lambda} | \phi_{k \bar{\lambda}} \rangle}{\beta^d \langle \phi_{k \bar{\lambda}} | \phi_{t \lambda} \rangle} = \sum_{\gamma=1}^{N} \langle \phi_{i \lambda} | \phi_{\gamma \bar{\lambda}} \rangle U_{\gamma}^{1/2} \sum_{\delta=1}^{N} \langle \phi_{i \lambda} | \phi_{\delta \lambda} \rangle D_{\bar{\lambda}, \gamma}^{\lambda, \bar{\lambda}}$$

(Ca-10)

In the next section the density matrices derived here will be utilized in solving for the best DODS $\Phi^\alpha$ approximation to the ground state of an $N$-electron molecule.
Cb: The GF Equations

We wish to minimize the energy $E = \frac{\langle \Phi | H | \Phi \rangle}{\langle \Phi | \Phi \rangle}$ with the constraint that the $n$ $\Phi_{ia}$ remain orthonormal and the $m$ $\Phi_{ib}$ remain orthonormal. Thus, $\delta I = 0$ where*

$$I = \sum_i E - \sum_i \xi_d \langle \Phi | \alpha_i | \Phi \rangle + \sum_i \xi_d \langle \Phi | \beta_i | \Phi \rangle$$

(Cb-1)

Consider $\sum_i \langle \Phi | (H-E) \alpha_i | \Phi \rangle = \langle \Phi | (H-E) \alpha_i | \Phi \rangle + \langle \Phi | (H-E) \alpha_i | \Phi \rangle$

$$= \langle \Phi | (H-E) \alpha_i | \Phi \rangle + \langle \Phi | (H-E) \alpha_i | \Phi \rangle$$

hence, it is real. Thus $\delta I$ of the other term must be real since $\delta I = 0$.

Consider the variation of $\Phi_{ia}$

$\delta (\text{other term}) = \sum_i \left[ \xi_d \langle \Phi | \alpha_i | \Phi \rangle + \xi_d \langle \Phi | \beta_i | \Phi \rangle \right] + \xi_d \langle \Phi | \alpha_i | \Phi \rangle$

Since this is real, by taking the complex conjugate we obtain $\xi_d = \xi_d^*$, i.e., $\xi^{\text{mu}} = \xi^{\text{nu}}$ are Hermitian. Take some unitary transformation of the $\Phi_{ia}$, e.g., $\Phi_{ia} = S_i \Phi_{ia}$ then $\Phi = \sum_i S_i \Phi_{ia}$, where $S_i = \Phi_i \Phi_i^*$

Also, $\xi_d \Phi = \sum_{\Phi_i} C_{\Phi_{ia}} (\xi_d \Phi_{ia}) = \sum_{\Phi_i} C_{\Phi_{ia}} \xi_{\Phi_i} (\Phi_{ia})$. But, $\langle \Phi | \alpha_i | \Phi \rangle = \delta \xi_{\Phi_i}$ and $\langle \Phi | \beta_i | \Phi \rangle = \delta \xi_{\Phi_i}$ (since the unitary transformation preserves the orthonormality of the $\Phi_{ia}$). Thus,

$$\sum_{\Phi_i} C_{\Phi_{ia}} C_{\Phi_{ia}}^* = \delta \xi_{\Phi_i}$$

(Cb-2)

Moreover, $0_{rf} \Phi = U_{rf} \Phi = \sum_{\Phi_i} C_{\Phi_{ia}} \Phi_{ia}$ if $U$ involves only elements $\leq n$ (or $> n$); so, $0_{rf} \Phi = \left( \sum_{\Phi_i} C_{\Phi_{ia}} \Phi_{ia} \right) 0_{rf} \Phi$. Hence,

$$\langle \Phi | (H-E) | \Phi \rangle = \langle \Phi | (H-E) | \Phi \rangle = \sum_{\Phi_i} C_{\Phi_{ia}} C_{\Phi_{ia}}^*$$

Recall that $b_{\Phi_a} b_{\Phi_a} = b_{\Phi_a} b_{\Phi_a}^* - b_{\Phi_a}^* b_{\Phi_a} = b_{\Phi_a}^*$, where $b_{\Phi_a}^* = b_{\Phi_a}$. Thus, $\sum_{\Phi_a} C_{\Phi_{ia}} C_{\Phi_{ia}}^* b_{\Phi_a} b_{\Phi_a} = \sum_{\Phi_a} b_{\Phi_a} b_{\Phi_a} = c_{\Phi_a} c_{\Phi_a} = 1$ using Cb-2.

* The $\xi$, a constant, is used to simplify later formulae.
Therefore, $\langle \Phi | (\text{H-E}) \text{O}_{\text{ff}} | \Phi \rangle = \langle \Phi | (\text{H-E}) \text{O}_{\text{ff}} | \Phi \rangle$. Thus,

$$I = \langle \Phi | (\text{H-E}) \text{O}_{\text{ff}} | \Phi \rangle - \sum_{\substack{p, q, r, t \in \mathbb{P} \\& \\text{distinct}}} \sum_{q, r, t \in \mathbb{P}} \sum_{a, b \in \mathbb{A}} \sum_{\substack{\xi_a, \xi_b \in \mathbb{P}}} \mathcal{S}(\xi_a; \xi_b) \Phi_{\text{ab}} \Phi_{\text{ab}}^*.$$ 

Hence, the transformed equation is the same as the original except that $\varepsilon_{ij}$ is replaced by $\varepsilon_{ij}^a \sum_{q \in \mathbb{P}} \xi_q^a \varepsilon_{qj}$. Since $\varepsilon^a$ is Hermitian and $\mathcal{S}^a$ is unitary, an $\mathcal{S}^a$ exists which diagonalizes $\varepsilon^a$. Similarly for $\varepsilon^b$.

In the following we use the $\Phi_i$ defined by these transformations, thus

$$I = \mathcal{S} \langle \Phi | (\text{H-E}) \text{O}_{\text{ff}} | \Phi \rangle - \sum_{\Phi_i \in \Phi} \langle \Phi_i | \Phi_i \rangle$$

(Cb-3)

Using the same analysis as on p. Bd-2, we see that $\Phi^a$ and $\Phi^b$ can be varied independently; we vary $\Phi^a$. We are now interested in the coefficient of $\Phi^a$ in $I$.

$$\mathcal{S} \langle \Phi | \text{O}_{\text{ff}} | \Phi \rangle (\text{ka}) \equiv \text{[Sum of terms in } \mathcal{S} \langle \Phi | \text{O}_{\text{ff}} | \Phi \rangle \text{ involving ka]} = \sum_{r=0}^{\infty} \frac{(r)!}{(r)!} \prod_{p=0}^{r} \left| C_{\omega r} \mathcal{S}(\alpha_1, \omega) \cdot \ldots \right|^2.$$ 

Now relax the restriction that the $p_i$ be ordered (i.e., $p_1 < p_2 < \ldots < p_r$); as a consequence we must divide by $r!$. Let $p_1 = k$ and multiply by $r$. Expand the determinant

$$r! \left[ C_{\omega r} \mathcal{S}(\alpha_1, \omega) \cdot \ldots \right] = \sum_{\xi_1, \ldots, \xi_r} \mathcal{S}(\xi_1, \ldots, \xi_r) \mathcal{S}(\xi_1, \ldots, \xi_r) \cdot \ldots = \sum_{\xi_1, \ldots, \xi_r} \mathcal{S}(\xi_1, \ldots, \xi_r) \mathcal{S}(\xi_1, \ldots, \xi_r) \cdot \ldots = \sum_{\xi_1, \ldots, \xi_r} \mathcal{S}(\xi_1, \ldots, \xi_r) \mathcal{S}(\xi_1, \ldots, \xi_r) \cdot \ldots = \mathcal{S}(\xi_1, \ldots, \xi_r) \mathcal{S}(\xi_1, \ldots, \xi_r) \cdot \ldots$$

In the second equation we set $\varepsilon = (\text{ Mafia } \varepsilon')$, where $\varepsilon'$ involves only elements 2 to $r$. Let

$$\mathcal{S} \langle \Phi | \text{O}_{\text{ff}} | \Phi \rangle (\text{ka}) = \sum_{\omega} \mathcal{S}(\alpha_1, \omega) \mathcal{F}(\omega, \mu),$$

where

$$\mathcal{F}(\omega, \mu) \equiv \sum_{r=1}^{\infty} \frac{(r)!}{(r)!} \prod_{p=0}^{r} \left[ \mathcal{S}(\alpha_1, \omega) \cdot \ldots \right],$$

and

$$\mathcal{S}(\alpha_1, \omega) \cdot \ldots$$
\( P(r) \) denotes the unordered set which is ordered in \( P_r \), and \( P'(r) \) has \( p_l \) deleted. Similarly, \( D_{ia}^{ka}(\omega) = \sum_{\mu_p} \delta_{\omega_i} S(\omega_i; \mu_b) F_{\mu_p}^{\mu_\omega}(\omega_i; \omega) \) if \( i \neq k \), where (for \( i \neq k \))

\[
D_{ia}^{ka}(\omega; \mu) = \sum_{r=1}^{\infty} \frac{r!}{(r-1)!} \frac{1}{P(r), Q(\mu)} \left[ \frac{1}{(a_r, \mu)} \left( \sum_{\mu' \in P'} \frac{1}{(a_r, \mu')} \left[ \frac{[O_1^{\mu'}]}{S^{(a_r; \omega)} S(\omega_i; \omega)} \right] \right) \right] \times \frac{1}{(a_r, \mu)} \left[ \frac{[O_1^{\mu}]}{S^{(a_r; \omega)} S(\omega_i; \omega)} \right] \times \frac{1}{(a_r, \mu)} \left[ \frac{[O_1^{\mu}]}{S^{(a_r; \omega)} S(\omega_i; \omega)} \right] \frac{1}{(a_r, \mu)} \left[ \frac{[O_1^{\mu}]}{S^{(a_r; \omega)} S(\omega_i; \omega)} \right]
\]

Continuing in the same manner let \( D_{ja}^{ka}(\omega; \mu) = \sum_{\mu_p} \delta_{\omega_j} S(\omega_j; \mu_b) F_{\mu_p}^{\mu_\omega}(\omega_j; \omega) \) if \( i \neq k \).

\[
D_{ja}^{ka}(\omega; \mu) = \sum_{r=1}^{\infty} \frac{r!}{(r-1)!} \frac{1}{P(r), Q(\mu)} \left[ \frac{1}{(a_r, \mu)} \left( \sum_{\mu' \in P'} \frac{1}{(a_r, \mu')} \left[ \frac{[O_1^{\mu'}]}{S^{(a_r; \omega)} S(\omega_j; \omega)} \right] \right) \right] \times \frac{1}{(a_r, \mu)} \left[ \frac{[O_1^{\mu}]}{S^{(a_r; \omega)} S(\omega_j; \omega)} \right] \times \frac{1}{(a_r, \mu)} \left[ \frac{[O_1^{\mu}]}{S^{(a_r; \omega)} S(\omega_j; \omega)} \right] \frac{1}{(a_r, \mu)} \left[ \frac{[O_1^{\mu}]}{S^{(a_r; \omega)} S(\omega_j; \omega)} \right]
\]

Defining \( F_{i,j}^{s,t}(\omega; \mu) \) in a similar way we obtain \( D_{i,j}^{s,t}(\omega; \mu) = \sum_{\mu_p} \delta_{\omega_i} S(\omega_i; \mu_b) F_{\mu_p}^{\mu_\omega}(\omega_i; \omega) \) if \( i, j \neq k \). Thus, the coefficient of \( S_{\omega}^{\mu} \) in \( \mathcal{E} \) is (after multiplying by \( \delta_{\omega}^{\mu} \))

\[
\mathcal{E} \{ \omega \} = \sum_{\omega_i} \left\{ \frac{1}{\omega} \delta_{\omega_i} S(\omega_i; \mu_b) D_{\omega}^{\mu_\omega} + \sum_{i, j=1}^{N} \frac{1}{\omega_i} \delta_{\omega_i} S(\omega_i; \mu_b) D_{\omega}^{\mu_\omega} + \sum_{i=1}^{N} \left[ \frac{1}{\omega_i} \delta_{\omega_i} S(\omega_i; \mu_b) F_{\omega}^{\mu_\omega} \right] + \sum_{j=1}^{N} \left[ \frac{1}{\omega_j} \delta_{\omega_j} S(\omega_j; \mu_b) F_{\omega}^{\mu_\omega} \right] + \sum_{i, j=1}^{N} \left[ \frac{1}{\omega_i} \delta_{\omega_i} S(\omega_i; \mu_b) F_{\omega}^{\mu_\omega} \right] + \sum_{j=1}^{N} \left[ \frac{1}{\omega_j} \delta_{\omega_j} S(\omega_j; \mu_b) F_{\omega}^{\mu_\omega} \right] \right\} \mathcal{E} = 0
\]

(Cb-4)

The differential-integral equation which is the coefficient of \( S_{\omega}^{\mu} \) must be equal to zero since the \( S_{\omega}^{\mu} \) are linearly independent. This equation is analogous to the Hartree-Fock equation which results from using a Slater determinant wave function. I will call the coefficient
of $\Phi_{ka}^*$ in Eq. (4.1) the GF equation for $\Phi_{ka}$ since it results from using $\Phi_{ka}$ as the wave function. We will write the GF equation as

$$H(\Phi_{ka}^*|\Phi_{ka} > = \varepsilon_{ka} \langle \Phi_{ka} |\Phi_{ka} >$$

(Cb-5)

where

$$H(ka) = \sum_{q \in Z} N q_{ka}^2 \hbar P_{ka} + \sum_{q \in 2.Z} \langle \Phi_{ka} |\Phi_{ka} > \sum_{p \in 2.Z} D_{ka,p}^2 P_{ka} +$$

$$+ \sum_{a \in Z} \left[ \sum_{q \in Z} \langle \Phi_{ka} |\Phi_{ka} > F_0^a(\Phi_{ka}; \mu) + \sum_{q \in 2.Z} \langle \Phi_{ka} |\Phi_{ka} > F_0^a(\Phi_{ka}; \mu) +$$

$$- EF(\Phi_{ka}; \mu) \right] P_{ka}$$

and $P_{ka,j}$ changes $\Phi_{ka}$ into $\Phi_j$. The biggest trouble with the GF equations is that they are a set of coupled integral partial differential equations, and, thus, are essentially impossible to solve. In order to get a set of equations which can be solved, we expand each $\Phi_j$ in terms of a fixed set of $M \times N$ basis functions, $\chi_j$. Then in the variation of $\Phi_j$ the coefficients of the expansion are varied; ROOTHAAN (1951) described this approach for the SODS closed-shell approximation of molecules. If the basis set $\chi_j$ were complete, then the exact solution of Cb-5 would be obtained. The basis sets will be finite, so some approximation is involved, the limit on the number of basis functions is set by the computers available. Let $\Phi_{1a} = \sum_{\chi_j} c_{a_j} \chi_j$ and $\Phi_{1b} = \sum_{\mu} c_{a_{1b}} \chi_{\mu}$. Let

$$\langle \chi_j | h | \chi_j \rangle = h_{jbj} \; j \langle \chi_j | \chi_j \rangle = \sum_{\mu} c_{a_{1b}}^{*} \chi_{\mu}$$

Thus, all of the integrals are evaluated once. From Cb-4

$$\sum_{a \in 2.Z} \sum_{b \in 2.Z} \sum_{a \in 2.Z} D_{ka,p}^2 \sum_{a \in 2.Z} + \sum_{a \in 2.Z} \sum_{b \in 2.Z} \sum_{a \in 2.Z} D_{ka,p}^2 \sum_{a \in 2.Z} +$$

$$+ \sum_{a \in 2.Z} \sum_{b \in 2.Z} \sum_{a \in 2.Z} \left[ \langle \chi_j | h | \chi_{j} \rangle F_0^a(\Phi_{ka}; \mu) + \sum_{a \in 2.Z} \langle \chi_j | \chi_j \rangle F_0^a(\Phi_{ka}; \mu) - EF(\Phi_{ka}; \mu) \right] =$$

$$= \varepsilon_{ka} \sum_{a \in 2.Z} c_{a}^{2}$$

Every term on the left side contains a factor of $c_{a}^{2}$, thus we can

* The $E$ appearing here is just $\langle \Phi | H | \Phi \rangle / \langle \Phi | \Phi \rangle$ and is evaluated for the guessed wave function just like all the other terms in $H(ka)$. $E$ is not used as a Lagrange multiplier in the SODS HF equations since in that case the constraint that the orbitals be orthonormal is sufficient to guarantee that the normalization of the $N$-electron wave function be unchanged.
write $\text{Cb-7}$ as

$$H_{\mu \nu}(a)C_{\mu \nu}^{a} = \epsilon_{\mu \nu}^{a}S_{\mu \nu}C_{\mu \nu}^{a}$$  \hspace{1cm} (Cb-8)

I call this equation the CFR equation since the Roothaan approach is used to solve for the GF orbitals $\Phi$. The method of solving Cb-8 is to guess the coefficients $C_{\mu \nu}^{a}$, then evaluate $H_{\mu \nu}^{a}$, then diagonalize $H_{\mu \nu}^{a}$ using the $S_{\mu \nu}$ metric (i.e., solve $|H_{\mu \nu}^{a} - \epsilon_{\mu \nu}^{a}|S_{\mu \nu} = 0$). The $C_{\mu \nu}^{a}$, which are solutions of Cb-8 are used to re-evaluate $H_{\mu \nu}^{a}$, then the new $H_{\mu \nu}^{a}$ is diagonalized etc.. The same is done for $\Phi$ to find the $\Phi_{ib}$. The only problem left is that from Cb-7 it would appear that $H_{\mu \nu}^{a}(a)$ depends on $k$; we must now show that this is not the case.

$$\phi(k \alpha_{a}; \mu) \equiv \sum_{\mu} \sum_{\pi} \sum_{\rho(\mu)} \sum_{\sigma(\mu)} \frac{\rho(\mu)}{P(k \alpha_{a}; \mu)} \left[ \frac{\partial r_{\mu}}{\partial \sigma(\mu)} O_{\mu \nu}^{\alpha} S(\mu \nu) \ldots \right] \left[ \frac{\partial r_{\nu}}{\partial \rho(\mu)} S(\mu \nu) \ldots \right] = 0$$ since the second determinant has the same column twice

Then $F(a; \mu) = F(k \alpha_{a}; \mu) + f(k \alpha_{a}; \mu)$ is independent of $k$. Thus,

$$\langle \Phi(k \alpha_{a}; \mu) | \Phi(k \alpha_{a}; \mu) F(\alpha; \mu) \rangle = \sum_{\alpha} S(k \alpha_{a}; \mu) F(\alpha; \mu) \right.$$

Also, $f_{ij}^{a}(k \alpha_{a}; \mu)$ and $f_{ij}^{a}(k \alpha_{a}; \mu)$ are defined similarly and are zero for the same reason. Thus, $F_{ij}^{a}(k \alpha_{a}; \mu)$, $F_{ij}^{a}(k \alpha_{a}; \mu)$, and $F(\alpha; \mu)$ in Cb-4 and Cb-5 can be replaced with

$$F_{ij}^{a}(k \alpha_{a}; \mu), F_{ij}^{a}(k \alpha_{a}; \mu), \text{ and } F(a; \mu).$$

Let

$$D_{ij}^{a}(k \alpha_{a}; \mu) = \sum_{\alpha} S(\mu \nu; k \alpha_{a}) F_{ij}^{a}(k \alpha_{a}; \mu)$$

where

$$F_{ij}^{a}(k \alpha_{a}; \mu) = -\sum_{\alpha} \sum_{\beta} \sum_{\gamma} \sum_{\delta} \sum_{\epsilon} \frac{\beta(\mu)}{P(\mu \nu; k \alpha_{a})} \left[ \frac{\partial r_{\beta}}{\partial \gamma(\mu)} O_{\beta \gamma}^{\alpha} S(\beta \gamma) \ldots \right] \left[ \frac{\partial r_{\gamma}}{\partial \delta(\mu)} \ldots \right]$$

Then $\sum_{\alpha} S(\mu \nu; k \alpha_{a}) f_{ij}^{a}(k \alpha_{a}; \mu) = 0$ since this corresponds to the det $S(\alpha \beta; \mu)$.
\[ D_{ka}^{ja} \text{ having the ka column twice; thus, } \eta_{ja}^{ka} = \sum_{j=1}^{m} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) \]

where \( F_{ja}^{\nu}(a; \nu) = F_{ja}^{\nu}(ka; \nu) \div F_{ja}^{\nu}(ka; \nu) \). The other \( D_{j}^{ka} \) and \( D_{j}^{ka,i} \) with \( j, s \neq k \) are treated similarly. First consider,

\[
\sum_{\theta_{2}} \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle \nu_{\theta_{1}}^{\nu_{\theta_{2}}} = \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle \nu_{\theta_{1}}^{\nu_{\theta_{2}}} D_{\theta_{1}}^{\theta_{2}} \theta_{1}^{\theta_{2}} + \sum_{\theta_{2}} \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle \nu_{\theta_{1}}^{\nu_{\theta_{2}}} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu)
\]

\[
= \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle \left[ D_{\theta_{1}}^{\theta_{2}} - \sum_{\theta_{1}} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) \right] + \sum_{\theta_{2}} \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu)
\]

But,

\[
\eta_{ka}^{k} \sum_{\nu} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) = \sum_{\nu} \left( \frac{1}{\nu} \sum_{\theta_{1}} |\langle \Phi_{\theta_{1}} | 1_{l'} | \Phi_{\theta_{1}} \rangle \nu_{\theta_{1}}^{\nu_{\theta_{2}}} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) \right)^{2}
\]

\[
= \sum_{\nu} \left( \frac{1}{\nu} \sum_{\theta_{1}} |\langle \Phi_{\theta_{1}} | 1_{l'} | \Phi_{\theta_{1}} \rangle \nu_{\theta_{1}}^{\nu_{\theta_{2}}} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) \right)^{2}
\]

Similarly,

\[
\sum_{\nu} \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle D_{\theta_{1}}^{\theta_{2}} = \sum_{\nu} \left[ \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle D_{\theta_{1}}^{\theta_{2}} + \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle D_{\theta_{1}}^{\theta_{2}} \right]
\]

\[
= \sum_{\nu} \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle \sum_{\nu} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) + \sum_{\nu} \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle \left[ D_{\theta_{1}}^{\theta_{2}} - \sum_{\theta_{1}} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) \right] + \langle \Phi_{\theta_{2}} | 1_{l'} | \Phi_{\theta_{2}} \rangle \sum_{\nu} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu)
\]

But,

\[
D_{ka, j}^{k, a} = \sum_{\nu} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) = \sum_{\nu} \left( \frac{1}{\nu} \sum_{\theta_{1}} |\langle \Phi_{\theta_{1}} | 1_{l'} | \Phi_{\theta_{1}} \rangle \nu_{\theta_{1}}^{\nu_{\theta_{2}}} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) \right)^{2}
\]

\[
= \sum_{\nu} \left( \frac{1}{\nu} \sum_{\theta_{1}} |\langle \Phi_{\theta_{1}} | 1_{l'} | \Phi_{\theta_{1}} \rangle \nu_{\theta_{1}}^{\nu_{\theta_{2}}} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) \right)^{2}
\]

Similarly, for the others we obtain

\[
D_{\theta_{1}}^{\theta_{2}} = \sum_{\nu} S(\nu; b, \theta_{2}) F_{\theta_{2}}^{\nu}(\theta; \nu) = D_{\theta_{1}}^{\theta_{2}}
\]
Also,
\[
D_{tb,kb}^{ja,jb} = \sum_{\mathcal{N}} S(\nu b; k\omega) F_{tb,kb}^{a,i} S(\omega a; \nu r) = \sum_{r=1}^{m} \sum_{(j+1)}^{(m+1)} \sum_{l=1}^{m} \sum_{(j+1)}^{(m+1)} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)] x_{l} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)] + \\
+ \sum_{r=0}^{m} \frac{(r+1)(r+1)}{(r+1)!} \sum_{(j+1)}^{(m+1)} \sum_{l=1}^{m} \sum_{(j+1)}^{(m+1)} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)] x_{l} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)]
\]

The second term is
\[
= + \sum_{r=1}^{m} \sum_{(j+1)}^{(m+1)} \sum_{l=1}^{m} \sum_{(j+1)}^{(m+1)} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)] x_{l} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)]
\]

Thus, \(D_{tb,kb}^{ja,jb} = \sum_{\mathcal{N}} S(\nu b; k\omega) F_{tb,kb}^{a,i} (\omega ; \nu r) = -D_{tb}^{ja,jb}\) Moreover, for the other cases where \(jb\) or \(tb\) is replaced by a term with \(a\), this result follows immediately from Cb-9 since \(D_{ta,ka}^{ja,jb} = -D_{ka,ta}^{ja,jb}\). Hence,

\[
D_{ta,ka}^{ja,jb} = \sum_{\mathcal{N}} S(\nu b; k\omega) F_{ta,ka}^{a,i} (\omega ; \nu r) = -D_{ta,ka}^{ja,jb}
\]

Now consider the coefficient of \(\langle \phi_{g} \mid \phi_{h} \rangle\) in Cb-4. Since \(F_{s,t}^{ja,jb}(ka;u) = 0\) and \(F_{s,t}^{ka,j}(ka;u) = 0\), then as before, we define \(F_{s,t}^{i,j}(ka;u)\) (which is zero) and \(F_{s,t}^{i,j}(a;u)\) as

\[
= F_{s,t}^{i,j}(ka;u) + (ka;u). \text{ Thus, Cb-4 becomes}
\]

\[
\langle \phi_{g} \mid \phi_{h} \rangle + \sum_{\mathcal{N}} S(\nu b; k\omega) F_{s,t}^{a,i} (\omega ; \nu r) + \\
+ \sum_{(j+1)}^{(m+1)} \sum_{l=1}^{m} \sum_{(j+1)}^{(m+1)} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)] x_{l} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)] + \\
+ \sum_{r=0}^{m} \frac{(r+1)(r+1)}{(r+1)!} \sum_{(j+1)}^{(m+1)} \sum_{l=1}^{m} \sum_{(j+1)}^{(m+1)} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)] x_{l} [c_{ij} S(\nu b; k\omega) S(\omega a; \nu r)]
\]

\[
= \epsilon_{a,b} \langle \phi_{g} \mid \phi_{h} \rangle
\]
The coeff of \( \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \) is

\[
\sum \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] \sum \frac{m}{2} \sum \sum \frac{m}{2} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] + \sum \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] \sum \frac{m}{2} \sum \sum \frac{m}{2} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] + \sum \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] \sum \frac{m}{2} \sum \sum \frac{m}{2} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right]
\]

Here \( F_{a;\mu}^{i} > = 0 \) and \( F_{a;\mu}^{ij} > = 0 \) as before. We let

\[
F_{a;\mu}^{i} > = \sum_{\lambda} S(\nu; \lambda) F_{a;\mu}^{i} > \left( \Phi_{a; \mu}, \nu \right), \quad \text{where}
\]

\[
F_{a;\mu}^{i} > = - \sum_{\lambda} \left[ \Phi_{a;\mu}^{i} > \left( \Phi_{a; \mu}, \nu \right) \sum_{\lambda} \sum_{\mu} \left[ \Phi_{a;\mu}^{i} > \left( \Phi_{a; \mu}, \nu \right) \right] \right]
\]

Next define \( F_{a;\mu}^{i} > \) which is zero, and \( F_{a;\mu}^{i} > \). Do the same for the other \( F_{a;\mu}^{i} > \) and \( F_{a;\mu}^{i} > \). Thus, Cb-4 becomes

\[
\sum_{\nu} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] + \sum_{\nu} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] + \sum_{\nu} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right]
\]

Hence, (see Cb-8)

\[
H_{a;\mu}^{a} = \sum_{\nu} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] + \sum_{\nu} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right] + \sum_{\nu} \left[ \langle \Phi_{a;\mu} | F_{\Phi_{a;\mu}} > \right]
\]

And we see that \( H_{a;\mu}^{a} \) is, indeed, independent of \( k \). Thus, the \( a \) \Phi_{a}^{i} \) are eigenfunctions of \( H_{a;\mu}^{a} \), and the \( b \) \Phi_{b}^{i} \) are eigenfunctions of \( H_{b;\mu}^{b} \).

\( H_{b;\mu}^{b} \) is obtained from Cb-11 by letting \( a \leftrightarrow b \) and the limits on \( \nu \) and
\[ H^{a}_{\mu \nu} \] is the GF Hamiltonian and the \( \Phi_{ia} \) are the GF orbitals.

Notice that by using the Roothaan procedure we avoid the coupling of equations present in Ch-5 because of exchange and nonorthogonality. Since \( H \) is Hermitian and independent of \( k \), then the \( \Phi_{ia} \) must be orthogonal. When we diagonalize \( H^{a}_{\mu \nu} \) and \( H^{b}_{\mu \nu} \) we get two independent sets of \( M \) eigenvectors: we select \( n \) from the first set and \( m \) from the second set for the \( \Phi_{ia} \) and \( \Phi_{ib} \) to use for the next iteration. The ones selected must be those which give the lowest value of \( E \) when evaluated for the new \( \Phi_{ia} \) and \( \Phi_{ib} \). It probably will result that the \( \Phi_{ia} \) selected will usually be those with the lowest \( \epsilon_{ii} \) as apparently occurs in the SODS HF approach.

---

I should mention that although everyone takes it for granted in the usual SODS HF approach (which is only applicable for closed shell systems, let \( N = 2n \)) that the \( n \) states of lowest \( \epsilon_{ii} \) are the occupied ones (e.g., ROOTHAN (1951) p. 77), I believe that that is not necessarily true. From Be-1 and Be-2 the energy, \( E' \), of the SODS HF state is (for a closed shell system)

\[
E' = \sum_{\Phi_i} \epsilon_{ii} - \frac{1}{2} \sum_{\Phi_i \Phi_j} \left( \sum_{q} (J_{pq} - \frac{1}{2} K_{pq}) \right) = \sum_{\Phi_i} \epsilon_{ii} + \frac{1}{2} \sum_{\Phi_i \Phi_j} \left( J_{pq} - \frac{1}{2} K_{pq} \right)
\]

Thus, the criterion for choosing the occupied states among the eigenstates of the HF Hamiltonian should be: The \( n \) states of lowest energy \( \epsilon_{ii} \) are occupied. We see then that if \( i = n \) is one of the \( n \) states of lowest \( \epsilon_{ii} \), and if \( p \) is some state with higher energy \( \epsilon_{pp} \) than \( \epsilon_{ii} \), and state \( p \) is occupied and state \( n \) is empty if

\[
\frac{1}{2} \sum_{q} (J_{pq} - \frac{1}{2} K_{pq}) > \epsilon_{pp} - \epsilon_{nn}
\]

This is true only when \( p \) is more favorable than \( n \) as regards the one-electron operator but less favorable as regards the electron-electron repulsion. For those those who would try to use the Virial theorem to dispute the above, remember that the electron-nucleus interactions are included in \( h \) and the nucleus-nucleus interactions have been omitted.
Cc: The Significance of the GR Orbitals

We now have a method of solving for the energy and the wave function of the molecular ground state (and the lowest state of each spin) without the restriction (SODS) of doubly-occupied orbitals and without the restriction (BFO) that the orbitals be basis functions of the irreducible representations of the total spatial symmetry group. The relaxation of the SODS restriction should yield more accurate energies and many-electron wave functions. The BFO restriction* forces Hartree-Fock molecular orbitals to be spread out through the entire molecular or solid**. The relaxation of the BFO restriction allows the variationally obtained orbitals (the GF orbitals) a great deal more freedom to vary in whichever manner minimizes the total energy; thus, again the many-electron wave function is improved and, furthermore, the GF orbitals -- which do not vary under the very constraining BFO restriction -- have very much more opportunity to become physically significant than do the Hartree-Fock orbitals.

There has been a prevalent feeling that the Hartree-Fock orbitals lack chemical significance and that chemically significant orbitals should be rather localized. The result is that a number of transformations of the Hartree-Fock orbitals into more localized orbitals under the restriction that the many-electron wave function be invariant have been proposed (e.g., WANNIER (1937), LENNARD-JONES (1949), DAUDEL (1955), and EDMISTON & RUEDENBERG (1963); others are listed in the last reference).

* ROOTHAAN (1951, p.80) has shown that for a nondegenerate molecular state with spin equal to zero the SODS restriction is sufficient to imply that the Hartree-Fock orbitals be BFOs. Dr. R. M. Pitzer pointed out this proof to me.

** The electron density and pair density are, of course, spread out through the molecule but we are here discussing the one-electron states from which the many-electron wave function is to be constructed.
All of these attempts have in common the disadvantage that the Hartree-Fock Lagrange-multiplier matrix is nondiagonal (see equation Bd-14) and thus, the localized orbitals are not eigenfunctions of some one-electron Hamiltonian. The Hartree-Fock orbitals are eigenfunctions of a one-electron Hamiltonian and can each be interpreted as a state in the average electrostatic field due to the other electrons, but the BFO restriction keeps the Hartree-Fock orbitals from being chemically significant. The GF orbitals are eigenfunctions of a one-electron Hamiltonian and can each be interpreted as a state in the average electrostatic field due to the other electrons; however, the GF orbitals are not required to be spread over the molecule and, thus, are free to become localized if such behavior would decrease the total molecular energy. Since in the GF method all artificial restrictions (SODS & BFO) have been removed except the one that no interelectron coordinates appear, then the GF orbitals are the best possible ones from which to form the many-electron wave function, and if any chemically significant interpretation of the many-electron wave function in terms of one-electron states exists, then these one-electron states are the GF orbitals*. Thus, if some common intuitive concepts of bonding are correct then we might expect some GF orbitals to be mainly localized near two (or a few) adjacent nuclei and perhaps others mainly localized near one nucleus. Remember that the total energy is minimized; the kinetic terms must prevent the orbitals from becoming com-

* Just as it is possible that Hartree-Fock solutions might predict equilibrium molecular structures in spite of the high correlation energy, it is not impossible that some method of localizing the Hartree-Fock molecular orbitals might approximate the GF orbitals; if so, the usefulness of the Hartree-Fock method would be greatly enhanced (and the method of localization might provide some physical insight).
pletely localized; thus, we must expect at least small subsidiary maxima near several nuclei. Since the $\Phi_{ia}$ and $\Phi_{jb}$ orbitals are (allowed to be) nonorthogonal, we might expect the bonding orbitals to pair together with perhaps some sort of extremal value for $S(ia;jb)$.

It should be emphasized that the lack of progress in understanding the structures of molecules and solids is partly due to the lack of a physical picture of the electron states of the system which corresponds adequately with reality. The importance of the physical picture has been stressed by Ruedenberg (1962) and Löwdin (1955a). With the GF method we can calculate the ground states for each spin and for all values of interatomic distance and all arrangements of the nuclei. Recall that the HF method causes large correlation errors at very large distances). Using the GF method one can, thus, find the stable structure and obtain information needed for studying the effect of nuclei motions. We can calculate the GF orbitals for various nuclear configurations and study the characteristics of the GF orbitals for the ground state.

For a given $N$ (number of electrons) and a given $S$ (total spin) there is just one set of GF equations. For the various compounds with the same total number of electrons and the same total spin we merely substitute different values for $\langle \mathbf{N}_{p}h_{b}\rangle$ and $\langle \mathbf{N}_{p}\mathbf{N}_{b}/g_{p}\rangle_{h_{b}-h_{p}}$.

My reasoning for expecting that the GF orbitals are localized.

Consider first the situation where there are no interactions between the electrons. The Hamiltonian $H = \sum_{1}^{N} \hat{h}_{i}$ is separable so the eigenfunction can be written as a product of functions each of which depends on a different set of one-electron variables. Thus, $H \Phi = E \Phi$, $\Phi = \phi_{1}(1) \phi_{2}(2) \cdots \phi_{N}(N)$, $\hat{h}_{i} \phi = \epsilon_{i} \phi$, $E = \sum_{i}^{N} \epsilon_{i}$. Since $H$ is invariant
under permutations, we can write $\chi_1^{\alpha_1} = \chi_1^{\alpha_1}$ just as before. The point is that since the $\phi_i$ are all eigenfunctions of the Hermitian operator $\mathbf{h}$, then they can be taken as orthonormal. Therefore, the non-orthogonality of the orbitals in the many-electron problem is due to the interactions of the electrons. Now consider the case where a portion of the electron interactions has been included (e.g., by the HF or GF method); the remaining perturbation represents the (average) difference between the instantaneous potential and the average potential due to all of the HF or GF orbitals; this perturbation we call the fluctuation potential. Since the Coulomb interaction is long range, then in order to minimize the average fluctuation potential (i.e., minimize the energy) the orbitals should be localized* (since then for large distances the average potential due to an electron in the orbital $i$ is essentially the same as the potential due to an electron at any of the positions for which $\phi_j$ is large and thus, the fluctuation is zero). Hence, the contributions to the fluctuation error from a particular orbital would be due only to the interactions of this orbital with a few others. In the process of minimizing the energy the orbitals would be localized and then varied to minimize the interactions with even the neighboring orbitals. Bear in mind that through all this varying of orbitals to minimize electron interaction the changes have to be such as to not increase the kinetic and electron-nuclear terms excessively (the total energy is minimized). It should be pointed out in the above reasoning that we are not really speaking of electrons in specific orbitals, but rather the discus-

* In the limit of only electron-electron interactions (no kinetic or electron-nuclear term) we would expect the electrons to be uniformly distributed and kept at maximum distances from each other, i.e., localized.
sion presupposes a partition of the perturbation interaction terms in such a way that the contributions can be interpreted as if the electrons were in fact in specific orbitals (in section Be the term \( \frac{-k}{pq} \frac{-k}{pq} \frac{q}{qp} \) could be interpreted as the fluctuation interaction between an electron in orbital p and an electron in orbital q). Such a partition should involve only pairs of electrons; because of nonorthogonality of the GF orbitals each term involves many orbitals and thus, I have not been able to convincingly show that the above interpretation is really valid. Although intuitively reasonable, the conjecture concerning the localization properties of the GF functions must remain a conjecture until some actual calculations are performed.

It would seem that the above reasoning follows in the same way for the case of an attractive-short-range interaction of particles. Thus, if it were valid to neglect spin interactions, it would seem that the many-nucleon wave function should be built out of localized orbitals. The use of the localized orbitals would allow much more advantage to be taken of the attractive nuclear forces than, say, orbitals which are eigenfunctions of the total symmetry group (i.e., \( \chi_{k}^{m} \) for the finite nuclei, \( e^{ikr} \) for infinite nuclear matter).

It should be emphasized that the use of GF orbitals is primarily to try to get a wave function which is sufficiently accurate to describe the system of particles but sufficiently simple to allow physical interpretation. It is not necessary that such a wave function should exist; however, the success of the chemical concepts of bonding indicate that such a wave function must exist.
Cd: The Defect in the GF Method

Now that we are aware of the genuine promise of important results which the GF method holds out, we must go back to a point which was completely glossed over at the very beginning of the development. Although, it is true that for \( \Phi \) and \( \mathcal{X} \) as in Ca-1 and Ca-2, we obtain
\[
G_i^a \Phi \mathcal{X} = \delta_{if} G_f^a \Phi \mathcal{X},
\]
this is not true if any other permutation of the states in \( \mathcal{X} \) is used. Since these new \( G_i^a \Phi \mathcal{X} \) are linearly independent of \( G_f^a \Phi \mathcal{X} \), then we should have considered the Hamiltonian matrix over this complete space and diagonalized it. Then we might have tried all possible sets of \( \phi_4 \) and \( \phi_6 \) and chosen the one that results in the lowest energy; this is a horribly complicated variational problem since the Hamiltonian matrix should be kept diagonal throughout the variation.

An alternate and simpler procedure would be to consider each different \( G_i^a \Phi \mathcal{X} \) separately and vary the \( \phi_4 \) and \( \phi_6 \) in order to minimize the energy. Those \( G_i^a \Phi \mathcal{X} \) giving the lowest energies would be examined and a compromise set of \( \phi_4 \) and \( \phi_6 \) chosen and a limited configuration interaction performed. This too is much too complicated. There are other possible schemes, none of which are very inviting.

Practically all discussions of DODS (vide infra) have presupposed that the total wave function is antisymmetric under transformations among the \( \phi_4 \) and among the \( \phi_6 \) (I will call this the anti-anti case). This requirement is sufficient to restrict the possible \( G_i^a \Phi \mathcal{X} \) to just the \( G_f^a \Phi \mathcal{X} \) we have considered in the previous sections of C; so, if such a requirement could be justified rigorously, then the foundation for the GF method would be absolutely sound and rigorous.

All of the work on alternate orbitals by Löwdin, de Heer, and
PAUNCZ has presumed the anti-anti case. This assumption is explicitly stated in one paper PAUNCZ et al. (1962a). YOSHIZUMI and ITOH (1955) actually considered the anti-anti case and one other case; the former allowed an energy stabilization of 2.35 ev whereas the other allowed an energy stabilization of only 0.3 ev. Yoshizumi and Itoh argue that the reason for the stabilization of the anti-anti case is that the correlation between the electrons in states of the same type (i.e., a or b) is automatically included by the Pauli principle and that the correlation between electrons in different types of states are taken into account by the semi-localization of the alternate orbitals. LIEB and MATTIS (1967) prove* the theorem that for a one-dimensional system of electrons the lowest energy state of spin S has a lower energy than the lowest energy state of spin S' if S < S'. This would be an important theorem except for one thing -- throughout their entire analysis they assumed that every state (or at least the lowest of a given spin) is of anti-anti type. Their proof would be correct if, in fact, the anti-anti state always has the lowest energy**. WIGNER (1934) treated the correlation energy for a many-electron system by using for the wave functions the product of a determinant of orbitals for spin times the determinant of the orbitals for spin, i.e., Wigner assumed the anti-anti case. The "unrestricted HF" method (see p. C-3) also assumes the anti-anti case. In other words in every case (with which I am familiar) where DODS is considered it is implicitly or explicitly assumed that the anti-anti case prevails.

* But see section Db, the theorem and the proof are incorrect.

** Which cannot be true for the one-dimensional case, see section Db.
The $G_f^*$ state for $\vec{r} = |n,m| \phi_{ia}$ is antisymmetrical for interchanges of
the $n \phi_{ia}$ orbitals and also for interchanges of the $m \phi_{ib}$ orbitals;
thus, the order of the subgroup of $S_n$ which transforms only antisymmetry-
cally related (and thus correlated) orbitals into each other is $n!m!$
This is the maximum possible antisymmetry for any state $G_i$ of the
same spin. The next best state is $S_{nb} = \frac{1}{(n-1)!(m-1)!}$. It is clear that the
states of higher antisymmetry have
correlations automatically included
between more orbitals and thus, can be expected to achieve lower energies.
If the total wave function changes sign when two orbitals are transposed
then the electrons in these orbitals cannot be at the same point at the
same time; whereas, if this condition does not hold then there is a
finite possibility of their coinciding. Since the electron interaction
contains a positive singularity, such close distances between electrons
are highly unfavorable; the anti-anti state has (by a factor of at least
$nm$ e.g., 9 for Benzene) the fewest possible close encounters by elec-
trons and thus, the variation can concentrate more on varying the orbitals
to minimize these few close encounters. On the other hand the non-anti-
anti cases must waste much of the flexibility in the varying orbitals
to try to make-up for the much larger number of close encounters. In
this light it is quite easy to believe that the anti-anti state is far superior to the others. Since we do make a variational calculation, then even if the actual ground state uses some \( G_i \phi \mathcal{N} \) different from \( G_f \mathcal{N} \), we can vary the \( \phi_i \) to get the best possible \( G_f \mathcal{N} \), which the preceding arguments imply should be quite close to the actual solution.

A real disadvantage to making variational calculations on non-anti-anti states is that the proof that the Lagrange multiplier matrix is diagonal (p. Cb-1, 2) only holds for \( G_f \mathcal{N} \) thus, the variational calculations are immensely more complicated than the GF equations. As mentioned on p. Bd-5 one can show that for Slater determinant wave functions with general spin orbitals, a transformation may be taken on the spin orbitals which diagonalizes the Lagrange multiplier matrix. However, such a wave function is not necessarily an eigenfunction of \( S^2 \) and it is not clear to me that one can necessarily be sure that the resulting spin orbitals can each be factored into either \( \phi_i \alpha \) or \( \phi_i \beta \). Thus, it is not (proved to be) true that the Lagrange multiplier matrix is diagonal for even the general case where the spin orbitals are factored into an orbital times \( \alpha \) or \( \beta \); thus, the first sentence of this paragraph does not contradict the general proof of the Lagrange multiplier matrix.
Ce: Restrictions on the GF Orbitals Due to Spatial Symmetry

The total wave function must, besides satisfying Pauli's principle and being an eigenfunction of \( S^2 \), be a basis function for an irreducible representation of the spatial symmetry group. As an example of how this may restrict the GF orbitals we consider the case of \( N = 2 \) electrons, \( S = 0 \), and the total wave function is nondegenerate. The wave function is \( \mathbf{G}_f \mathbf{a}_a \mathbf{b}_b \) and \( \mathbf{K}_f \mathbf{a}_a \mathbf{b}_b = \mathbf{C}_R \mathbf{G}_f \mathbf{a}_a \mathbf{b}_b \), where \( |C_R| = 1 \). Take \( \mathbf{a}_a \) and \( \mathbf{b}_b \) to be real and note that \( \mathbf{G}_f \mathbf{a}_a \mathbf{b}_b \mathbf{b}_b = \mathbf{C}_R \mathbf{G}_f \mathbf{a}_a \mathbf{b}_b \). Define

\[
\psi_1 = \frac{1}{\sqrt{2(1+S)}} (\mathbf{a}_a + \mathbf{b}_b) \quad \text{and} \quad \psi_2 = \frac{1}{\sqrt{2(1-S)}} (\mathbf{a}_a - \mathbf{b}_b)
\]

where \( S = \langle \mathbf{a}_a | \mathbf{b}_b \rangle \); then \( \psi_1 \) and \( \psi_2 \) are orthonormal. Thus,

\[
\mathbf{a}_a = \sqrt{\frac{1+S}{2}} \psi_1 + \sqrt{\frac{1-S}{2}} \psi_2 \quad \text{and} \quad \mathbf{b}_b = \sqrt{\frac{1-S}{2}} \psi_1 - \sqrt{\frac{1+S}{2}} \psi_2
\]

Hence, \( \mathbf{G}_f \mathbf{a}_a \mathbf{b}_b \mathbf{b}_b = (\frac{1+S}{2}) \mathbf{G}_f \psi_1 \psi_1 \mathbf{b}_b - (\frac{1-S}{2}) \mathbf{G}_f \psi_2 \psi_2 \mathbf{b}_b \). Define the functions \( \psi_3, \psi_4, \ldots \) such that the set of all \( \psi_i \) is complete and orthonormal.

Expand \( R \psi_1 \) and \( R \psi_2 \) as \( R \psi_1 = \sum \alpha \beta \psi_3 \) and \( R \psi_2 = \sum \alpha \beta \psi_4 \). Therefore,

\[
\mathbf{G}_f \mathbf{a}_a \mathbf{b}_b \mathbf{b}_b = \sum \frac{1}{2} \left[ (1+S) \alpha \beta \psi_3 \psi_3 - (1-S) \alpha \beta \psi_4 \psi_4 \right] + \left( (1+S) \alpha \beta \psi_3 \psi_3 - (1-S) \alpha \beta \psi_4 \psi_4 \right) \mathbf{G}_f \psi_1 \psi_2 \mathbf{b}_b
\]

The various \( \mathbf{G}_f \psi_1 \psi_2 \mathbf{b}_b \) are orthogonal and, hence, linearly independent; thus, we may equate coefficients.

From \( i = j > 2 \), \( \sqrt{1+S} \alpha_{1i} = \pm \sqrt{1-S} \alpha_{2i} \)

From \( j > i > 2 \), either all of the signs in the above equation are + or all are -.

From \( j > 2, i = 1 \), \( (1+S) \alpha_{1j} \alpha_{1i} = (1-S) \alpha_{2j} \alpha_{2i} = \pm \sqrt{1-S} \sqrt{1+S} \alpha_{1j} \alpha_{2i} \), thus, either all \( \alpha_{1j} = 0 \) for \( j > 2 \) or \( \sqrt{1+S} \alpha_{11} = \pm \sqrt{1-S} \alpha_{21} \)

From \( i = j = 1 \), \( (1+S) \alpha_{11}^2 - (1-S) \alpha_{21}^2 = C_R (1+S) \) thus, we cannot have the second of the above choices and \( \alpha_{1j} = 0 \) for \( j > 2 \)

* Dr. R. M. Pitzer suggested that this case be done explicitly.
Similarly we obtain $A_{2j} = 0$ for $j > 2$.

The $A_{11} A_{12} A_{21} A_{22}$ must satisfy

Ce-1: $(1+S)A_{11}^2 - (1-S)A_{21}^2 = C_R (1+S)$

Ce-2: $(1+S)A_{12}^2 - (1-S)A_{22}^2 = -C_R (1-S)$

Ce-3: $(1+S)A_{11} A_{12} = (1-S)A_{21} A_{22}$

Ce-4: $|A_{11}|^2 + |A_{12}|^2 = 1$; $|A_{21}|^2 + |A_{22}|^2 = 1$

Ce-5: $A_{11}^* A_{21} + A_{12}^* A_{22} = 0$

where Ce-4,5 are obtained from the requirement that $\langle \mathbf{r}_j | \mathbf{r}_j \rangle = S_{jj}$.

We let $A_{ij} = \alpha_{ij} e^{i \Theta_{ij}}$, where $\alpha_{ij}$ is real and positive. We let $C_R = 1$ and find that $\alpha_{11} = \alpha_{22} = 1$ and $\alpha_{12} = \alpha_{21} = 0$ and $\Theta_{11} = 0, \pm \pi$ and $\Theta_{22} = 0, \pm \pi$. Thus, $A_{11} = \pm 1$, $A_{22} = \pm 1$, and $A_{12} = A_{21} = 0$. But

$$R \phi_a = \sqrt{\frac{1-S}{2}} A_{11} \psi_1 + \sqrt{\frac{1+S}{2}} A_{22} \psi_2$$

$$R \phi_b = \sqrt{\frac{1-S}{2}} A_{11} \psi_1 - \sqrt{\frac{1+S}{2}} A_{22} \psi_2$$

Hence, either $R \phi_a = \pm \psi_a$ and $R \phi_b = \pm \psi_b$  \hspace{1cm} (Ce-6)

or  \hspace{1cm} $\kappa \phi_a = \pm \psi_b$ and $\kappa \phi_b = \pm \psi_a$  \hspace{1cm} (Ce-7)

We note that the $\phi_a$ function may transform into the $\phi_b$ function under some symmetry transformations. For $H_2$ we would expect Ce-6 (with + signs) to prevail for rotations about the axis and Ce-7 to prevail when the atoms are interchanged by $R$. For $He$ we would expect Ce-6 with plus signs to prevail for all elements of $SO(3)$. Thus, the GF orbitals would be symmetric with respect to the axis for the ground state of $H_2$ and would be $S$ functions for the ground state of $He$ (cf. LÖWDIN (1955c), p.1517).
Löwdin suggested a simple way to remove much of the correlation error in the \( \pi \) electron M.O. of alternate systems (e.g., benzene). In these systems there is a natural set of excited states which can be mixed with the \( n \) filled (\( \pi \) electron) states (mixing each filled state with one excited state and using different excited states for different filled states) to split the doubly-occupied orbitals. By using one mixing parameter for all the states, a set of relatively simple equations are obtained. Solutions have been carried out for \( C_2H_2 \) and \( C_6H_6 \) (Yoshizumi and Troh (1955), Troh and Yoshizumi (1955)) and for cyclic systems of \( 2n \) atoms (where \( n \) is odd) (Pauncz et al. (1962)). The general theory is presented in Pauncz et al. (1962a). The extension to the case where different mixing parameters are used for different orbitals is considered by de Heer (1962). All the preceding studies have neglected the inner shell electrons. All of the formulae of the preceding works can be derived from GF theory by simply letting \( S(ia;jb) = \lambda \delta_{ij} \) in the \( D_{ij}^* \) and \( D_{ij}^{kt} \) of equation Ca-10 (or to allow different mixing parameters we simply let \( S(ia;jb) = \lambda \delta_{ij} \)). In all cases even if different parameters are used \( D_{ia}^{ja} = 0 \) if \( i \neq j \), \( D_{ka,ra}^{ia,ja} = 0 \) unless \( \xi_{i,j}^2 = \xi_{k,t}^2 \), \( D_{ia,ja} = 0 \) if \( k = i \) or \( j \), \( D_{ia,kb}^{ia,ka,ta} = 0 \) and \( D_{ia,ka}^{ia,kb} = 0 \) if \( k \neq t \).

If all \( \lambda \) are equal we obtain (\( N \) must be equal to \( 2n \) for the alternate orbital method).
\[ D_{ia} = \sum_{r=0}^{m-1} \left( \frac{m-r}{m} \right)^{1.1} \frac{1}{1-\lambda^2} \left[ 1 - \frac{\lambda^2 (1-\lambda^2)^m}{m (1-\lambda^2)^3} \right] \xrightarrow{\lambda \to 1} \frac{m+1}{2} \]

\[ D_{ib} = \sum_{r=0}^{m-1} \frac{2^{r+1}}{(r+1)} \left( \frac{r}{r+1} \right) = \frac{1}{1-\lambda^2} \left[ -\lambda^2 + \frac{(1-\lambda^{2m})}{m (1-\lambda^2)^2} \right] \xrightarrow{\lambda \to 1} \frac{m+1}{2} \]

\[ D_{ia, ja} = \sum_{r=0}^{m-2} \frac{2^r}{(m+r)} \left( \frac{r}{r+1} \right) = A = \frac{1}{1-\lambda^2} \left[ -\lambda^2 + \frac{2\lambda^{2m}}{(m-1)(1-\lambda^2)^2} + \frac{2\lambda^2 (1-\lambda^{2m})}{m (m-1)(1-\lambda^2)^2} \right] \xrightarrow{\lambda \to 1} \frac{m+1}{3} \]

\[ D_{ia, ja} = \sum_{r=0}^{m-2} \frac{2^r}{(m+r)} \left( \frac{r}{r+1} \right) \equiv B = \lambda^2 \left[ \frac{2^{m-2}}{(m-1)} - \frac{2\lambda^{2m}}{(m-1)(1-\lambda^2)^2} + \frac{2(1-\lambda^{2m})}{m (m-1)(1-\lambda^2)^2} \right] \xrightarrow{\lambda \to 1} \frac{m+1}{3} \]

\[ D_{ib, ja} = \sum_{r=0}^{m-2} \frac{2^r+1}{(m+r+1)} \left( \frac{r}{r+1} \right) = \lambda^2 C \quad IF \quad r \neq 0 \quad \xrightarrow{\lambda \to 1} \frac{m+1}{6} \]

\[ D_{ia, ib} = \sum_{r=0}^{m-2} \lambda^2 \left( \frac{r}{r+1} \right) = \lambda^2 D \quad IF \quad r \neq 0 \quad \xrightarrow{\lambda \to 1} \frac{m+1}{6} \]

\[ D_{ia, ja} = \sum_{r=0}^{m-2} \lambda^2 \left( \frac{r}{r+1} \right) = \lambda^2 E \quad \xrightarrow{\lambda \to 1} \frac{m+1}{3} \]

\[ D_{ib, ia} = \sum_{r=0}^{m-1} \lambda^2 \left( \frac{r}{r+1} \right) = \lambda^2 F = \frac{-\lambda^{2m}}{1-\lambda^2} + \frac{(1-\lambda^{2m})}{m (1-\lambda^2)^2} \xrightarrow{\lambda \to 1} \frac{m+1}{2} \]

\[ D_{ib, ja} = \sum_{r=0}^{m-2} \lambda^2 \left( \frac{r}{r+1} \right) = \lambda^2 G \quad IF \quad r \neq 0 \quad \xrightarrow{\lambda \to 1} \frac{m+1}{6} \]

\[ D_{ib, ja} = \sum_{r=0}^{m-2} \frac{2^r+1}{(m+r+1)} \left( \frac{r}{r+1} \right) = \lambda^2 H \quad IF \quad r \neq 0 \quad \xrightarrow{\lambda \to 1} \frac{m+1}{3} \]

all other cases are zero.

From here one next evaluates the \( \langle \ln \rangle \) and \( \langle |y| \rangle \) integrals in equation Ca-10 then minimizes the energy with respect to \( \lambda \). The limits shown at the right for each quantity are for \( \lambda \to 1 \), i.e., for the SODS case, and in fact when the limiting quantities are placed in Ca-10 the result


LÖWDIN (1962b, p.270) states that if DODS are used, then for an
S = 0 state a transformation on the \( \phi_{ia} \) and another on the \( \phi_{ib} \) can be
found for which \( S(ia;jb) = \lambda S_{ij} \), and therefore, it is not necessary to
go beyond the general alternate orbital scheme. It should be realized
that the transformation to \( S(ia;jb) = \lambda S_{ij} \) introduces off-diagonal
elements in the Lagrange multiplier matrix which would make the GF
equations unsolvable and also would remove the wonderful properties
promised for the \( \phi_{ia} \) in section Cb. Thus, one is left with no way to
solve the energy problem since the orbitals are not known and the matrix
elements can't be evaluated.

By its nature the alternate orbital method is wedded to SODS-BFO
type molecular orbitals; so it is possible that the orbitals obtained
will never be significantly more useful than the SODS-BFO M. O. The
energies obtained should be more accurate, of course.
Section D

Spin Considerations

Introduction

In all previous discussions the spins have been completely summed over in all matrix elements since only operators independent of spin have been used. There are electron systems with non-zero spin density and/or spin pair-density (e.g., for crystals, antiferromagnets such as MnF$_2$ and ferromagnets such as CrO$_2$ and for molecules, O$_2$). Thus, measurements depending on these spin densities and pair-densities (e.g., neutron diffraction for crystals) yield information about the electron system complementary to that obtained from measurements not involving spin. The interactions resulting in the various magnetic structures are no doubt much the same ones as are involved in chemical bonding so the explanation of one awaits the elucidation of the other. So, the variety of new information involving spin measurements should be helpful in developing the theory of the chemical bond. In section Db the first- and second-order density matrices are obtained; these would be used in obtaining the expectation values for operators involving spin. In section Db a theorem concerning the lowest energy state among the states of different spin for a one-dimensional system is discussed.
Da: First- and Second-Order Density Matrices With Unsummed Spin

Before, we have always summed over the spin when evaluating matrix elements. Now we will find the first- and second-order density matrices, $\gamma$ and $\gamma^2$ (which have unsummed spins), for $\Phi^2$

\[
\Phi^2 = \frac{1}{\sqrt{\mu^2}} \langle \Phi | \Phi \rangle \langle \Phi | \Phi \rangle \Phi \Phi
\]

\[
\Sigma (\Phi^2 \Phi^2) = \frac{1}{\mu^2} \sum_{\mu \nu} U_{\mu \nu} U_{\nu \mu}^* \Sigma (\Phi \Phi) \Sigma (\Phi \Phi)
\]

\[
\Sigma (\Phi \Phi) = \frac{1}{\mu^2} \sum_{\mu \nu} U_{\mu \nu} U_{\nu \mu}^* \Sigma (\Phi \Phi) \Sigma (\Phi \Phi)
\]

But, $\Sigma (\Phi \Phi) = 0$ unless $\nu = \sigma = \sigma'$, where $\sigma$ involves only elements $= n$

\[
\text{and } \sigma' \text{ only elements } > n
\]

For convenience let $M = S$ in the following

\[
\Sigma (\Phi \Phi) = \frac{(\frac{N_s}{2} + 1)! (\frac{N_s}{2} - 1)!}{\mu^{2N_s} / \mu^2} \sum_{\nu = \sigma = \sigma'} U_{\nu \nu} U_{\nu \nu}^* \Sigma (\Phi \Phi) \Sigma (\Phi \Phi)
\]

\[
\Sigma (\Phi \Phi) = \frac{(\frac{N_s}{2} + 1)! (\frac{N_s}{2} - 1)!}{\mu^{2N_s} / \mu^2} \sum_{\nu = \sigma = \sigma'} U_{\nu \nu} U_{\nu \nu}^* \Sigma (\Phi \Phi) \Sigma (\Phi \Phi)
\]

\[
\gamma = N \Sigma (\Phi^2 \Phi^2) (\Phi^2 \Phi^2) (\Phi^2 \Phi^2) \Phi \Phi
\]

\[
= \frac{(\frac{N_s}{2} + 1)! (\frac{N_s}{2} - 1)!}{\mu^{2N_s} / \mu^2} \sum_{\nu = \sigma = \sigma'} U_{\nu \nu} U_{\nu \nu}^* \Sigma (\Phi \Phi) \Sigma (\Phi \Phi)
\]

But, $U_{\nu \nu} = \sum_{\nu = \sigma = \sigma'} U_{\nu \nu}^*$ from appendix I.

Thus, $\sum_{\nu = \sigma = \sigma'} U_{\nu \nu} U_{\nu \nu}^* = \sum_{\nu = \sigma = \sigma'} U_{\nu \nu}^*$

\[
\gamma = \frac{(\frac{N_s}{2} + 1)! (\frac{N_s}{2} - 1)!}{\mu^{2N_s} / \mu^2} \sum_{\nu = \sigma = \sigma'} U_{\nu \nu} U_{\nu \nu}^* \Sigma (\Phi \Phi) \Sigma (\Phi \Phi)
\]

where $\gamma$ is the normalized $\gamma$, $\gamma(\sigma, \sigma') = [\gamma(\sigma, \sigma') \rho_\sigma(1) \rho_\sigma(1) \rho_\sigma(1)]$, where $\rho_\sigma(1) \equiv \frac{N}{\mu^{2N_s} / \mu^2} \Sigma (\Phi \Phi) (\Phi \Phi) (\Phi \Phi) \Phi \Phi$

For $\rho_\sigma(1)$ the sum over $\nu$ is for $\nu(\sigma) > \sigma$. $\rho_\sigma(1) + \rho_\sigma(1) = \rho(1)$ since

\[
\sum_{\nu = \sigma = \sigma'} U_{\nu \nu} U_{\nu \nu}^* = \delta_{\sigma \sigma'}
\]
Recall that \( \theta = \frac{N!}{\cosh(N-1)} \).

In the simple cases \((N = 3, 4)\) which I have tried, the sums simplify somewhat, but I have not found a general way to simplify these equations.

Once the \( \Phi \) is known \( \chi \) and \( \rho^2 \) can be evaluated to predict such things as the Fermi contact term of NMR, the chemical isomeric shift of Mossbauer studies, and the elastic spin scattering of neutrons. Thus, it is by careful study of the \( \chi \) and \( \rho^2 \) that we must hope to try to extract the reasons for the stability of the various magnetic structures.
Concerning the Lowest State of a One Dimensional System
Of Electrons

LIEB and MATTIS (1962) prove the theorem that for a one-dimensional system of electrons if \( E(S) \) is the energy of the lowest energy state of spin \( S \) and \( E(S') \) is the energy of the lowest state spin \( S' \), then \( E(S) \leq E(S') \) if \( S \leq S' \). The theorem, although dealing with a thought problem, could be useful; unfortunately, the theorem is quite wrong. The error in the theorem results from the lack of an adequate method of constructing an eigenstate of \( S^2 \) which also satisfies Pauli's principle. We will use the \( G_1 \) method for this purpose.

The basic ideas in the proof are: (i) If the many-electron wave function for a one dimensional system (one-dimensional electrons) is antisymmetric under interchange of some spatial coordinates (say 1 and 2), then each point in configuration space for which \( x_1 = x_2 \) is a node of the wave function and this node divides the space into two regions (\( x_1 < x_2 \) and \( x_1 > x_2 \)). The knowledge of the wave function in either region is sufficient to fix the wave function over the entire space. Thus, if the wave function is antisymmetric under interchanges of \( \alpha_1, \alpha_2, \alpha_3 \) and also of \( x_4, x_5, x_6, x_7 \), then (if the system is finite and the wave is required to be zero on the boundaries, say \( x = 0 \) and \( L \)) consideration may be restricted to the domain \( 0 < x_1 < x_2 < x_5 < L \) and \( 0 < x_4 < x_5 < x_6 < x_7 < L \) where the boundaries are nodes. (ii) the ground state of \( H \phi = E\phi \) has no nodes; so, if two eigenfunctions of the many-electron problem have the same restricted domain and one has further nodes but the other does not, then the latter has the lower
energy*. Consider, for example, $\alpha = [2, 1, 2]$, $S_2 = \begin{bmatrix} 2 & 2 \\ 3 & 4 \end{bmatrix}$, then consider $0 < \gamma < \xi < L$. $G_2^\alpha \Phi \gamma$ has no required nodes in this region but $G_3^\alpha \Phi \gamma$ has additional required nodes for $\gamma_1 = \gamma_3$ and $\gamma_2 = \gamma_3$ thus $P_{22}^\alpha \leq P_{33}^\alpha$.

First, we consider LIEB and MATTIS' error. All of their proof (excepting the basic idea paraphrased above) is predicated upon an implicit misunderstanding, and thus, is irrelevant. The erroneous assumption which is made explicit in the last paragraph on p. 166 is that every many-electron eigenfunction with $S_z = M$ is antisymmetric in one set of $\frac{N}{2} + M$ variables and in another set of $\frac{N}{2} - M$ variables. The fact is that only for $S_z^\alpha$ where $\alpha = [\frac{N}{2}, \frac{N}{2} - M]$ (i.e., $M_s = S$) is that assumption true, and $G_1^\alpha \Phi \gamma$ is never (vide infra) the lowest energy state of spin for a one-dimensional problem.

Theorem Db-1: Of all the states $G_i^\alpha \Phi \gamma$ for a system of $N$ one-dimensional electrons, $G_1^\alpha \Phi \gamma$ has the lowest energy, where $\alpha = [\frac{N}{2}, \frac{N}{2} - M]$ and $M = 0$ or 1.

Proof: $S_1 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ or $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ and, thus, has no necessarily antisymmetric variables; and, hence, can be nodless through all of configuration space. Every other $S_1$ for all $i$ and $\alpha$ has at least one occasion where adjacent letters are in the same column and thus has at least one required node.

Thus, the lowest state for three-dimensional Bose electrons (i.e., ones identical with real electrons except for the symmetry of the wave function under permutation) would always be the totally symmetric state $F_i^\alpha \Phi \gamma = F_i^\alpha \Phi \gamma(\omega, \chi)$ with $\alpha = [N]$ (thus, the total spin = $N/2$). This is because $F_i^\alpha \Phi \gamma = \text{const} \langle \alpha | \Phi_{H} | \alpha \rangle$ and any other $\alpha$ would have some nodes while $\alpha = [N]$ has no nodes. It is believed that such particles do not exist.
It should be noted that Lieb and Mattis tried to prove that the lowest state is

\[
S_f^{[\alpha]} = \begin{array}{c}
1 \\
2 \\
3 \\
\vdots \\
n \\
N
\end{array}
\]

(for an even number of electrons)

We see that in fact there are many states of lower energy. For example,

\[
S_1^{[\alpha_1, \alpha_2]} = \begin{array}{c}
1 \\
2 \\
3 \\
2n-3 \\
2n-2 \\
2n-1 \\
2m
\end{array}
\]

In spite of ThM Db-1 it is not possible to prove (at least not by using the ideas in the second paragraph of this section) that the lowest energy state of a one-dimensional system of electrons is a state of lowest spin. An eigenstate of spin \( \alpha \) (i.e., \( \alpha = \left\lfloor \frac{n}{2} + \frac{N}{2} - \frac{S}{2} \right\rfloor \)) can be constructed from any linear combination of the \( f^\alpha \) different \( G_{\alpha}^\epsilon \); since there is no pair of coordinates which is antisymmetric in every tableau of shape \( \alpha \), then the wave function for any spin can be made nodeless. It is apparent that the use of the \( G_{\alpha}^\epsilon \) provides a clear insight into many-electron problem, at least for the Hamiltonian A-1.
Explanation of the $\text{VO}_2$-Distorted Rutile Crystal

Structure and Prediction of the Magnetic Structure

Introduction

In the previous sections all discussions have been at a fairly rigorous level in that the few approximations made could be stated and the resulting errors approximated (e.g., the errors due to neglecting nuclear motion and relativistic effects (at least for light atoms) have been treated by others, the error in using the Roothaan approach can be estimated by using larger or smaller basis sets); however, although important results may be foreseen in the future, actual calculations must perforce deal with systems of only about a dozen electrons or less. In this section we will make a major assumption which has only the most tenuous of theoretical justification -- but an aura of empirical basis -- and proceed from there with a few approximations to obtain some quite interesting results. The results are of such a nature as to place some more a posteriori evidence in favor of the original assumptions.

The kind of system that we deal with is a nonconducting oxide or fluoride of a transitional metal element. The physical ideas involved are: (1) the compounds are ionic; this is used to mean that the predominant energy lowering mechanism is related to a transfer of electrons (as compared to the free atom) from the transition element atom to the oxygen atom, each oxygen atom is presumed to take two electrons and each fluorine
atom one: (2) any electrons left on the transition element after the charge transfer and in addition to the number of the heaviest lighter inert gas atom are not involved significantly in binding and are considered localized on the atom; I call these the extravalence electrons; (3) the extravalence electrons are all in one-electron states belonging to the \( L = 2 \) irreducible representation of \( SO(3) \) (we would use \( L = 3 \) for rare earth elements); (4) Hund's rule applies to the extravalence electrons of each atom, so that each atom has the maximum possible spin consistent with the number of extravalence electrons and with the requirement that they be in \( L = 2 \) states (we forget about the net orbital angular momentum, which would occur in the free atom, pretending that it is zero, i.e., quenched). In section 5a a survey of the pertinent properties of the compounds exhibiting the rutile structure is presented; the conclusion is that the above ideas apply to these compounds.

At this point we make the big assumption: we allow the set of extravalence electrons from different atoms to interact and assume that the relative energy of the possible states is given by (the Heisenberg Hamiltonian)

\[
H = 2 \sum_{i,j} J_{ij} \mathbf{s}_i \cdot \mathbf{s}_j
\]  

(E-1)

where \( i \) and \( j \) refer to different atoms and the sum is over the different pairs. In analogy to expressions such as 3c-9 the \( J_{ij} \) is sometimes referred to as an "exchange integral", note that I use the opposite sign. The \( S_i \) used here is considered to be a vector with its magnitude being

* It should be clear that since the electrons are indistinguishable, I am not really discussing electrons but rather the one-electron states to be later used in constructing the many-electron wave function. Convention and expediency make the present terminology convenient.
the spin of atom i and its direction being the direction of the spin of atom i (as given by neutron diffraction experiments). To really be analogous to, say, Bc-9 the equation should involve $S_i$ rather than $S_i^\uparrow$, one might want to consider this as a second assumption. Using the Hamiltonian E-1 the possible stable magnetic structures for compounds exhibiting the rutile structure are determined in section Eb.

In section Eb one of the stable magnetic structures, $\gamma 3$, is analyzed and it is determined that for the $\gamma 3$ magnetic structure the rutile crystal structure is not stable. The necessary distortions of the rutile crystal structure result in the VO$_2$ crystal structure; thus, it is predicted that the $\gamma 3$ magnetic structure is probably exhibited by VO$_2$ and the distortion of the VO$_2$ is explained as the result of the interactions responsible for the magnetic structure.
Ea: Survey of the Rutile-Structure Compounds

The rutile structure is exhibited by the following compounds: MgF₂(1), MnF₂(2)(3), FeF₂(2)(3), CoF₂(2)(3), NiF₂(2)(3), ZnF₂(2)(3), PdF₂(6), TiO₂(1)(7), CrO₂(4)(8), MnO₂(17), TaO₂(5), GeO₂(1), and SnO₂(1). The following six substances have a structure which is essentially the same as rutile but with a slightly lower symmetry and slight (∼0.05 to 0.3 Å) displacements of the atoms: VO₂(12), NbO₂(13)(14), MoO₂(9)(10), TeO₂(11), WO₂(9)(10) and ReO₂(11). The following two compounds have structures directly related to the rutile structure but with a significant distortion; CrF₂(16) and CuF₂(15). The following compounds are reported to have the rutile structure (but the work is rather old, Goldschmidt 1926, ref. 18 PbO₂, RuO₂, OsO₂, IrO₂, and TeO₂ (from the Jahn-Teller theorem we would expect RuO₂ and OsO₂ to be strongly distorted), TeO₂ seems to me to be an inappropriate compound.

---

18. Goldschmidt (1926) Geochemische Verteilungsgesetze VI, measurements by W. H. Zachariasen; quoted in Strukturbericht 1, 158,211, 213 (1931)
for the rutile structure (the extra electrons are not "d-type" electrons), IrO$_2$ seems a little out of place when compared to the compounds for which the rutile structure is well verified). There are several cases of the rutile structure with more than one cation (19); however, these have not been investigated very thoroughly.

The compounds MnF$_2$(20), FeF$_2$(20), CoF$_2$(20), and NiF$_2$(20) are all antiferromagnetic with the same arrangement of localized magnetic moments; MnO$_2$(21)(22) has a unique spiral antiferromagnetic structure; and CrO$_2$(4)(8) is ferromagnetic. The compound VO$_2$ is reported to have an antiferromagnetic Neel point at about 340°K ((23) measuring magnetic susceptibility reports 341°K with most of the change between 338°K and 343°K, (25) by the same means reports between 335°K and 355°K, (24) measuring specific heat reports 340°K); this is about the same temperature as the conductivity transition ((26) gets 340-325°K when cooling and 335-350°K when heating, the low temperature form is a semiconductor

\[ \rho = \frac{\rho}{\rho} \text{ at the transition}. \]

On the other hand MnO$_2$, MnF$_2$, FeF$_2$, (19) K. Brandt, Arkiv Kemi, Min. Geol. 1/A, #15 (1943).
(20) R. A. Erickson, Phys. Rev. 90, 779-785 (1953).
(21) R. A. Erickson, Phys. Rev. 85, 745 (1952). This is an abstract. The work was on powder specimens; when single crystals were later used, it was found that the structure given here is incorrect. See reference 22. Erickson's single crystal data (and thus, his method of preparation, etc.) on which reference 22 is based have not (as far as I know) been published.
(25) N. Perakis and J. Wucher, Compt. Rend. 235, 354-356 (1952); reference 27 gives the wrong volume and page number for this reference.
CoO₂ and NiF₂ have Neel points of about 90°K, 70°K, 00°K, 40°K, and 70°K respectively ((27) p. 14 ff, I have averaged and rounded off the experimental values). Ferromagnetic CrO₂ has a Curie point at 116°C (28). Erickson found that the temperature dependence of the magnetic structure factor of the (100) diffraction for MnF₂, FeF₂, CoF₂, and NiF₂ are consistent with localized spins on the transition element of magnitude 5/2, 3/2, 3/2, and 1 respectively (20), such a measurement is really not too sensitive but the results are consistent with what people think is reasonable. The saturation magnetization of CrO₂ is consistent with S = 1 (29); this is a reliable indication of the magnitude of the localized spin. Erickson’s measurements (20) indicate that the orbital angular momentum is quenched in MnF₂, CoF₂, and NiF₂ but not quite quenched in FeF₂.

The fact that every case of this structure involves oxygen or fluorine as the anion suggests that the stability of the rutile structure is not due to covalent (i.e., electron interference) effects but rather to electron transfer effects. The occurrence of the structure for many transition element cations and many non-transition element cations suggests that the extravalence electrons on the transition elements are not important in the stability of the rutile structure. The above evidence and the fact that all the rutile compounds are apparently insulators (except for VO₂ which shows a transition from a semiconductor to a poor conductor (26) and CrO₂ which has a resistivity

(28) A. Michel and J. Bénard, Compt. Rend. 200, 1316-1318 (1935)
(29) C. Guillaud, A. Michel, J. Bénard, & M. Fallot, Compt. Rend. 219, 50-60 (1944)
of $10^{-2}$ to $10^{-3} \angle$-cm at room temperature (30)) suggests that the set of all extravalence electrons of all the transition element ions may be considered as somewhat independent of the other electrons (31) in the solid (the ones which are responsible for the stability or the structure) and, in addition, that these extravalence electrons may be considered as localized. The localized spins on the transition elements (for those compounds on which measurements have been made) are consistent with the Hund's rule assumption and with the usual change transfer values; the angular momentum seems to be fairly well quenched. Thus, the compounds having the rutile structure fit the requirements for using the Heisenberg Hamiltonian quite well (probably better than any other structure).

(31) See footnote on page E-2
Eb: The Stable Magnetic Structures for the Rutile Crystal Structure

We use \( E-1 \) and assume that \( J_{ij} \) is equal to zero except between nearest neighbors (two at \( \pm c e_2 \), \( J = J_2 \)), next-nearest neighbors (eight \( 8 + \frac{2}{2} a e_2 + \frac{2}{2} a e_2 + \frac{2}{2} a e_2 \), \( J = J_1 \)), and the favorable (for superexchange because of the intervening oxygen atoms) pair of next-next-nearest neighbors (two at \( \pm ae_2 \) \( \pm ae_2 \) and \( \pm ae_2 \) \( \pm ae_2 \) for one sublattice and at \( \pm ae_2 \) \( \pm ae_2 \) and \( \pm ae_2 \) \( \pm ae_2 \) for the other sublattice, \( J = J_4 \)).

1. The spin is treated as a vector and expanded in a Fourier series (the periodicity is due to the assumption of periodic boundary conditions over the microcrystal, as usual) which effects a partial diagonalization of the Hamiltonian. Thus,

\[
H = 2J_1 \sum_{\langle i,j \rangle} S_i \cdot S_j + 2J_2 \left[ \sum_{\langle i,j \rangle} S_i \cdot S_j + \sum_{\langle i,j \rangle} S_i \cdot S_j \right] + 2J_4 \left[ \sum_{\langle i,j \rangle} S_i \cdot S_j + \sum_{\langle i,j \rangle} S_i \cdot S_j \right]
\]

where, \( \langle i,j \rangle \) indicates all pairs of a certain type, e.g., for \( J_2 \)

have the sum over all nearest neighbor pairs.

For \( J_1 \):

\[
i' = (n_1 \pm \frac{1}{2}) a_1 + (n_2 \pm \frac{1}{2}) a_2 + (n_3 \pm \frac{1}{2}) c
\]

For \( J_2 \):

\[
i' = n_1 a_1 + n_2 a_2 + n_3 c; j = n_1 a_1 + n_2 a_2 + (n_3 \pm 1)c
\]

For \( J_4 \):

\[
i' = n_1 a_1 + n_2 a_2 + n_3 c; j = (n_1 \pm 1)a_1 + (n_2 \pm 1)a_2 + n_3 c
\]

for one sublattice

\[
i' = n_1 a_1 + n_2 a_2 + (n_3 \pm \frac{1}{2}) c; j' = (n_1 \pm 1)a_1 + (n_2 \pm 1)a_2 + (n_3 \pm \frac{1}{2}) c
\]

for other sublattice

2. \( S_i = \sum_k S(k) e^{\frac{2\pi i k}{a} \cdot r_i} \); \( S_i' = \sum_k S(k) e^{\frac{2\pi i k}{a} \cdot r_i} \).
where the sum is over the \( k \) in the first Brillouin zone, there are \( N \) such states and thus, \( N \) unit cells in the microcrystal (2N transition element cations).

3. Substituting Eb-2 in Eb-1 we obtain

\[
(31) \quad 2J_1 \sum_{<\mathbf{k},\mathbf{k}'>} \mathbf{S}_i \cdot \mathbf{S}_i' = 2J_1 \sum_{<\mathbf{k},\mathbf{k}'>} \mathbf{S}(\mathbf{k}) \cdot \mathbf{S}(\mathbf{k}') \sum_{\mathbf{h}} e^{i(\mathbf{h} \cdot \mathbf{k} + \mathbf{h} \cdot \mathbf{k}')} \mathbf{g}_{\mathbf{h}} \mathbf{y}_{\mathbf{k}}
\]

where

\[
\mathbf{g}_{\mathbf{h}} \mathbf{y}_{\mathbf{k}} = \left[ e^{\frac{i}{2}(\mathbf{h} \cdot \mathbf{x} + \mathbf{h} \cdot \mathbf{y} + \mathbf{h} \cdot \mathbf{z})} + e^{\frac{i}{2}(\mathbf{h} \cdot \mathbf{x} - \mathbf{h} \cdot \mathbf{y} - \mathbf{h} \cdot \mathbf{z})} \right] \mathbf{g}_0
\]

\[
= \mathbf{g}_0 \left( e^{\frac{i}{2} \mathbf{y} \cdot \mathbf{r}} - e^{-\frac{i}{2} \mathbf{y} \cdot \mathbf{r}} \right),
\]

(3ii) But (see e.g., ZIMAN (1960) p. 10)

\[
\mathbf{y}_{\mathbf{k}} = \mathbf{e}^{i \mathbf{k} \cdot \mathbf{r}} N \sum_{\mathbf{h}} \mathbf{g}_0 \mathbf{h} \cdot \mathbf{g}_0
\]

The unit cell is simple tetragonal; thus, so is the first Brillouin zone; so, the only way for \( \mathbf{k} + \mathbf{k}' = \mathbf{g} = 0 \) is for \( \mathbf{k} \) to be on the surface of the Brillouin zone. Note that both \( \mathbf{k} \) and \( \mathbf{g} - \mathbf{k} \) must be in the first Brillouin zone. The term \( \mathbf{e}^{i \mathbf{g} \cdot \mathbf{r}} \) is due to not requiring that the coordinate system origin be placed on a lattice site of the sublattice under consideration. \( \mathbf{r}_0 \) denotes the vector from the origin to some lattice site of the primed sublattice. It would seem that there is an inconsistency here since we could get either \( \mathbf{e}^{i \mathbf{g} \cdot \mathbf{r}_0} \) or \( \mathbf{e}^{-i \mathbf{g} \cdot \mathbf{r}_0} \) depending on which sublattice we sum over. However, it turns out (vide infra) that the only occasions when \( \mathbf{g} \neq 0 \) have \( \mathbf{y}_{\mathbf{k}} = 0 \), and if \( \mathbf{g} = 0 \) then \( \mathbf{e}^{i \mathbf{g} \cdot \mathbf{r}_0} = \mathbf{e}^{-i \mathbf{g} \cdot \mathbf{r}_0} \). Similarly (remember to count each pair only once)

\[
(3iii) \quad 2J_2 \sum_{<\mathbf{k},\mathbf{k}'>} \mathbf{S}_i \cdot \mathbf{S}_i' = \sum_{<\mathbf{k},\mathbf{k}'>} \mathbf{S}(\mathbf{k}) \cdot \mathbf{S}(\mathbf{k}') \sum_{\mathbf{h}} e^{i(\mathbf{h} \cdot \mathbf{k} + \mathbf{h} \cdot \mathbf{k}')} 2 \mathbf{g}_{\mathbf{h}} \mathbf{y}_{\mathbf{k}}
\]

where,

\[
2 \mathbf{g}_{\mathbf{h}} \mathbf{y}_{\mathbf{k}} = \mathbf{e}^{i \mathbf{h} \cdot \mathbf{c}} + \mathbf{e}^{-i \mathbf{h} \cdot \mathbf{c}} = 2 \mathbf{e}^{i \mathbf{h} \cdot \mathbf{k} \cdot \mathbf{c}}
\]
(3iv) \[ 2 J_+ \sum_{\nu, \delta} S_\nu \cdot S_\delta = 2 J_+ \sum_{\nu, \delta} S(\nu) \cdot S(\delta) e^{i (\nu + \delta) \cdot \mathbf{K}_N} 2 \mathbf{K}_N \]

\[ 2 \mathbf{K}_N = e^{i (\mathbf{K}_N - \mathbf{K}_N) \cdot \mathbf{a}} + e^{-i (\mathbf{K}_N - \mathbf{K}_N) \cdot \mathbf{a}} = 2 \cos (\mathbf{K}_N \cdot \mathbf{a}) \mathbf{a} \]

\[ 2 \mathbf{K}_N = 2 \cos (\mathbf{K}_N \cdot \mathbf{a}) \mathbf{a} \]

(3v) thus,

\[ H = N \sum_{\nu, \delta} \sum_{\mathbf{K}_N} \chi_{\nu, \delta} \left[ S(\nu) \cdot S(\delta) e^{i \mathbf{K}_N \cdot \mathbf{a}} + S(\nu) \cdot S(\delta) e^{-i \mathbf{K}_N \cdot \mathbf{a}} \right] + 2 J_+ \sum_{\nu, \delta} \sum_{\mathbf{K}_N} \left[ \chi_{\nu, \delta} \left( S(\nu) \cdot S(\delta) e^{i \mathbf{K}_N \cdot \mathbf{a}} + S(\nu) \cdot S(\delta) e^{-i \mathbf{K}_N \cdot \mathbf{a}} \right) \right] + 2 J_+ \sum_{\nu, \delta} \left[ \chi_{\nu, \delta} \left( S(\nu) \cdot S(\delta) e^{i \mathbf{K}_N \cdot \mathbf{a}} + S(\nu) \cdot S(\delta) e^{-i \mathbf{K}_N \cdot \mathbf{a}} \right) \right] + 2 J_+ \sum_{\nu, \delta} \left[ \chi_{\nu, \delta} \left( S(\nu) \cdot S(\delta) e^{i \mathbf{K}_N \cdot \mathbf{a}} + S(\nu) \cdot S(\delta) e^{-i \mathbf{K}_N \cdot \mathbf{a}} \right) \right] \]

\[ \chi_{\nu, \delta} = \cos (\mathbf{K}_N \cdot \mathbf{a}) \]

(4) However, there are some conditions on the possible values of \( S(\nu) \)

(4i) \( S_\nu \) is real, thus \( S(-\nu) = S^*(\nu) \)

(4ii) \( S_\nu' S_\nu = S_\nu^2 = S^2 \) is a scalar and thus is left invariant by lattice translation; \( S_\nu^2 = \sum_{\mathbf{e}_k} S(\nu) \cdot S(\nu) e^{i (\nu + \nu') \mathbf{e}_k} \)

\[ = \sum_{\mathbf{e}_k} \left[ S(\nu) \cdot S(\nu) + \sum_{\mathbf{e}_k} S(\nu) \cdot S(\nu) e^{i (\nu + \nu') \mathbf{e}_k} \right] \]

Hence, \( S_\nu^2 (\nu, \nu') \) is invariant under lattice translations.

(4iii) Consider first the terms with \( \nu' = \nu \) (i.e., assume \( S(\nu) S(\nu') - 0 \) if \( \nu' = \nu \) or \( -\nu \) then \( S(\nu) - S(\nu) e^{i 2 \mathbf{e}_k \cdot \mathbf{K}_N} \) is invariant under lattice translations; thus,

\[ \sum_{\mathbf{e}_k} \left[ \left( S(\nu) S(\nu) + S(\nu) S(\nu) \right) e^{i 2 \mathbf{e}_k \cdot \mathbf{K}_N} + \left( S(\nu) S(\nu) - S(\nu) S(\nu) \right) e^{-i 2 \mathbf{e}_k \cdot \mathbf{K}_N} \right] \]

is invariant under lattice translations.
Thus, either

(4iii⃗{9}) $S(\mathbf{k}) \cdot S(\mathbf{k}) = 0$ (and thus $S(-\mathbf{k}) \cdot S(-\mathbf{k}) = 0$); no restriction on $\mathbf{k}$

(4iii⃗{9}) $2k \cdot r_{i} = n\pi$, thus, $k_{i} = 0, \pi$; no restriction on $S(\mathbf{k}) \cdot S(\mathbf{k})$

(4iii⃗{9}) $S(\mathbf{k}) \cdot S(\mathbf{k}) = \text{pure imaginary and } r_{i} = 0$ for some $i$; then

\[ \sin 2k \cdot r_{j} = 0, \text{ hence, } 2k \cdot r_{j} = n\pi \text{ or } k_{j} = 0, \pm \frac{\pi}{2}, \pi \Rightarrow r_{j} = 0; \]

we must include at least one $k_{j} = \pm \frac{\pi}{2}$ to get cases different from $\beta$

(4iv) condition (iii) is too restrictive, we should only require invariance for

\[ \sum_{\mathbf{k}} S(\mathbf{k}) \cdot S(\mathbf{k}) e^{ik \cdot r_{i}} \text{ where } K = \mathbf{k} + \mathbf{k}', \text{ for all } K \]

\[ \sum_{\mathbf{k}} S(\mathbf{k}) \cdot S(\mathbf{k}) e^{ik \cdot r_{i}} + S^{*}(\mathbf{k}) \cdot S^{*}(\mathbf{k}) e^{-ik \cdot r_{i}}, \text{ for each } K; K \text{ goes all possible values of } k + k'. \]

Thus, either,

(4iv⃗{9}) \[ \sum_{\mathbf{k}} S(\mathbf{k}) \cdot S(\mathbf{k}) = 0 ; \text{ no restriction on } K \]

(4iv⃗{9}) $K \cdot r_{j} = 2\pi n$, thus $K_{i} = 0, 2\pi = k_{i} + k'_{i}$

(4iv⃗{9}) \[ \sum_{\mathbf{k}} S(\mathbf{k}) \cdot S(\mathbf{k}) = \text{imag} \text{ and } r_{i} = 0 \text{ for some } i \text{ then } \]

\[ \sin(K \cdot r_{j}) = 0; \text{ hence, } K \cdot r_{j} = n\pi \; ; k_{i} = 0, \pi, \pi = k_{i} + k'_{i}. \]

(4v) in all cases we get \[ \sum_{\mathbf{k}} S_{i}^{2} = \sum_{\mathbf{k}} \left[ S(\mathbf{k}) \cdot S(\mathbf{k}) + \sum_{q \geq 0} S(\mathbf{q}) \cdot S(\mathbf{q}) e^{i\mathbf{q} \cdot \mathbf{r}_{i}} \right] \]

\[ G = \sum_{\mathbf{k}} \left[ \cos \left( \frac{1}{2} \mathbf{q}_{1} \cdot \mathbf{q}_{1} \right) \cos \left( \frac{1}{2} \mathbf{q}_{2} \cdot \mathbf{q}_{2} \right) \cos \left( \frac{1}{2} \mathbf{q}_{3} \cdot \mathbf{q}_{3} \right) \cos \left( \frac{1}{2} \mathbf{q}_{4} \cdot \mathbf{q}_{4} \right) \cos \left( \frac{1}{2} \mathbf{q}_{5} \cdot \mathbf{q}_{5} \right) \cos \left( \frac{1}{2} \mathbf{q}_{6} \cdot \mathbf{q}_{6} \right) \cos \left( \frac{1}{2} \mathbf{q}_{7} \cdot \mathbf{q}_{7} \right) \right] \]

For $k \neq \frac{q}{2}$ at most one term of $\frac{q}{2}$ is allowed and even then $k$ must lie on the surface of the Brillouin zone.

(5i) Now minimize $H$ with respect to $S(\mathbf{k})$ and $S'(\mathbf{k})$ holding all

\[ s_{1}^{2} = s^{2}; \text{ thus, consider } H - \sum_{\mathbf{k}} \lambda_{k} S_{i}^{2} - \sum_{\mathbf{q}} \lambda_{k} S_{i}^{2} \text{ (see 3v and 4v)}. \]

\[ \frac{\partial}{\partial S(\mathbf{q})} : \sum_{\mathbf{k}} \left[ \frac{1}{6} \lambda_{k} S(\mathbf{q} - \mathbf{k}) + \sum_{q \geq 0} S(\mathbf{q} - \mathbf{k}) e^{i\mathbf{q} \cdot \mathbf{r}_{i}} + 2 \sum_{q \geq 0} S(\mathbf{q} - \mathbf{k}) e^{i\mathbf{q} \cdot \mathbf{r}_{i}} \right] \]

\[ = \sum_{\mathbf{k}} \lambda_{k} S(\mathbf{q} - \mathbf{k}) + 2 \sum_{q \geq 0} S(\mathbf{q} - \mathbf{k}) e^{i\mathbf{q} \cdot \mathbf{r}_{i}} \]

\[ \frac{\partial}{\partial S'(\mathbf{q})} : \sum_{\mathbf{k}} \left[ \frac{1}{6} \lambda_{k} S(\mathbf{q} - \mathbf{k}) + \sum_{q \geq 0} S(\mathbf{q} - \mathbf{k}) e^{i\mathbf{q} \cdot \mathbf{r}_{i}} + 2 \sum_{q \geq 0} S(\mathbf{q} - \mathbf{k}) e^{i\mathbf{q} \cdot \mathbf{r}_{i}} \right] + 
\]
(5ii): Note that \( \sum \lambda' \left( e^{i\mathbf{g} \cdot \mathbf{r}} e^{i\mathbf{g} \cdot \mathbf{r}_0} \right) \), where \( \mathbf{r}_{1,\mathbf{r}} \) is the vector to some atom position of the unprimed sublattice, and
\[
\sum \lambda' e^{i\mathbf{g} \cdot \mathbf{r}'} e^{i\mathbf{g} \cdot \mathbf{r}_{10}} \leq \lambda',
\]
where \( \mathbf{r}'_{10} \) is some atom position of the primed sublattice. But, \( e^{i\mathbf{g} \cdot \mathbf{r}} e^{i\mathbf{g} \cdot \mathbf{r}_0} e^{i\frac{1}{2} (\mathbf{g} \cdot \mathbf{a}_1 + \mathbf{g} \cdot \mathbf{a}_2 + \mathbf{g} \cdot \mathbf{a}_3)} = (-1)^n e^{i\mathbf{g} \cdot \mathbf{r}_0} \).

where \( n \) is the number of the \( S(-\mathbf{k}) = 2\pi \) (0 to 3). Cases where \( n \neq 1 \) can occur only for \( k = \frac{\pi}{2} \). Let \( N\lambda' \leq \lambda \), \( N\lambda' \leq \lambda' \).

(5iii): Consider \( k \) not on surface of the Brillouin zone
\[
16 J_1 \lambda' \left( S(\mathbf{k}) + (4 J_2 \lambda_{k2} + 4 J_3 \lambda_{k3}) S(-\mathbf{k}) \right) = 2 \lambda S(\mathbf{k})
\]

If we multiply the first equation by \( S(k) \) and the second equation by \( S'(k) \) and sum over \( k \), we get \( \frac{2H}{N} = 2(S(k)+S'(k))S^2 \). Thus, \( H = N(S(k)+S'(k))S^2 \). Hence, \( H \) is proportional to \( S(k)+S'(k) \). But, if \( S(k) \) and \( S'(k) \) are non-zero, then the determinant equals zero or \( S(k) = 0 \) and \( S'(k) = 0 \).

\[
\begin{vmatrix}
A - B - \lambda & C \\
C & A - D - \lambda
\end{vmatrix} = 0
\]

\( A = 2J_2 \lambda_{k2} \); \( B = 2J_4 \lambda_{k4} \); \( C = 8J_1 \lambda_{k1} \); \( D = 2J_4 \lambda'_{k4} \).

Thus, \( \lambda' - (\lambda + \lambda') \) - \( (A + D) + (A + D) \) - \( C^2 \) = 0. Thus, we have \( \lambda' \) as function of \( \lambda \), but \( H \) is proportional to \( \lambda + \lambda' \) and we want \( H \) to be a minimum; thus, we take \( \frac{dH}{d\lambda} = NS^2 \left( 1 + \frac{dA}{d\lambda} \right) = 0 \); thus, \( \frac{dA}{d\lambda} = -1 \). Substituting above we get \( \lambda' + \lambda \frac{dA}{d\lambda} - \frac{dA}{d\lambda} (A + D) = 0 \). Hence,

\( \lambda' - \lambda = (A + D) - (A + D) = D - B \). Substituting in the secular equation and solving, we obtain \( \lambda = (A + D) \pm C \) and \( \lambda' = (A + D) \pm C \). Let,

\( \mu \leq \lambda + \lambda' = 2A + (B + D)\pm 2C = 4J_2 \lambda_{k2} \pm 2J_4 \lambda_{k4} \). \( \pm 16J_1 \lambda_{k1} \); then

\( H = NS^2 \); also, \( \left( \frac{S(-\mathbf{k})}{S(\mathbf{k})} \right) \left( \frac{S(-\mathbf{k})}{S(\mathbf{k})} \right) = 0 \); thus, \( S(-\mathbf{k}) = \pm S'(\mathbf{k}) \) or \( S(\mathbf{k}) = \pm S'(\mathbf{k}) \). Hence, \( S(\mathbf{k}) = \pm S'(\mathbf{k}) \). Now note that the only non-zero \( S(k) \) and \( S'(k) \) are those for which \( \mu \) has the same numerical
value as the above one. For this to be true for general values of $J_1$, $J_2$, and $J_4$ we must have \( \tau_{K1}' = \tau_{K1} \), thus \( \tau_2' = \pm \tau_2 \) or \( \pm \tau_2 \); 
\( \tau_{K2}' = \tau_{K2} \), thus \( \tau_2' = \pm \tau_2 \) or \( \pm \tau_2 \); and \( \tau_4' = \pm \tau_4 \) or \( \pm \tau_4 \). These conditions taken together imply that if $k$ is taken into $k'$ by one of the symmetry operations in the group $D_{4h}$ then \( \mu_k = \mu_{k'} \). Since the rutile space group is homomorphous onto the $D_{4h}$ point group, we expect the $k$ states related by the operations of the $D_{4h}$ point group to be equivalent. If one of the interactions $J_1, J_2,$ and $J_4$ had been neglected, however, then the group of operations relating the $k$ states of equal energy would have been some group containing $D_{4h}$ as a subgroup. In the sense that there is no degeneracy (excepting the accidental degeneracy which results for specific relations between the $J_1, J_2,$ and $J_4$ values) other than that required by the crystal symmetry, we have included enough types of interactions to eliminate all falsely degenerate states.

Regardless of the particular values of the $S(k)$ for the various degenerate modes which are combined for a particular spin state, \( H = N \mu S^2 \) and is thus, the same. Consequently, it is sufficient when determining the lowest energy state to pick one mode out of each degenerate set.

(5iv) Consider $k$ on surface of Brillouin zone but $k \neq g/2$. Now note that the argument $k$ of $S(k)$ must be in first Brillouin zone; therefore, we can get only one $g$, say $g_1$, such that $g_1 - k$ is in first Brillouin zone if $k \neq g/2$. But $k$ on the Brillouin
zone surface implies that \(-k\) is not in the first Brillouin zone, \(g_1\cdot k\) on the B. Z. surface implies that \(k-g_1\) is not in the first B. Z.; thus, we can have no \(S(-k)\) or \(S(k-g_1)\) terms. For every possible \(g_1\), \(n_{g_1} = 1\).

From 5i
\[
16 J_1 \gamma_{k_1} S'(g_1, \epsilon) e^{\frac{\epsilon}{k_1}} \langle 4J_2 \gamma_{k_2} + 4J_4 \gamma_{k_4} \rangle S(g_1, \epsilon) e^{\frac{\epsilon}{k_1}} = 2\lambda S'(g_1, \epsilon) e^{\frac{\epsilon}{k_1}}
\]
\[
16 J_1 \gamma_{k_1} S(g_1, \epsilon) e^{\frac{\epsilon}{k_1}} \langle 4J_2 \gamma_{k_2} + 4J_4 \gamma_{k_4} \rangle S'(g_1, \epsilon) e^{\frac{\epsilon}{k_1}} = 2\lambda' S'(g_1, \epsilon) e^{\frac{\epsilon}{k_1}}
\]
\[
\langle \gamma_{k_1} \rangle = -\langle \gamma_{-k_1} \rangle; \text{ thus, } \begin{vmatrix} A-B-\lambda' & C \\ C & -A-D+\lambda' \end{vmatrix} = 0
\]

\[
A = 2J_2 \gamma_{k_2}; \quad B = 2J_4 \gamma_{k_4}; \quad C = 8J_1 \gamma_{k_1}; \quad D = 2J_4 \gamma_{k_4}
\]

Solving as in (iii) \(\lambda = \lambda' + (D-B)\), \(\lambda = (A+B) \pm C\),

thus, there is no real solution unless \(C = 0\). But if \(k\) is on the B. Z. surface, then at least one \(k_1 = \pi\), and since \(\gamma_{k_1} = \cos \frac{k_1}{2} \cos \frac{k_2}{2} \cos \frac{k_4}{2}\), then \(\gamma_{k_1} = 0\) and, thus, \(C = 0\); thus, \(\mu = \lambda + \lambda' - 2A + B + D = 4J_2 \gamma_{k_2} + 2J_4 (\gamma_{k_4} + \gamma_{k_4}')\).

No relation between \(S(k)\) and \(S'(k)\) is required since the minimization conditions separate into a term involving only \(S(k)\) and a term involving only \(S'(k)\).

(5v) Consider \(k\) on the surface of Brillouin zone and such that \(k = g/2\) for some \(g \neq 0\). Thus, we consider
\[
k \in \langle \frac{\pi}{2}, \frac{\pi}{2}, 0 \rangle, \langle \frac{\pi}{2}, 0, \frac{\pi}{2} \rangle, \langle 0, \frac{\pi}{2}, \frac{\pi}{2} \rangle, \langle \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2} \rangle, \langle 0, 0, \frac{\pi}{2} \rangle, \langle \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2} \rangle.
\]

In each case the only possible \(g-k\) is \(g-k = k\) or \(g = 2k\) (using 5i).

Note that if \(k = g/2\) then \(-k\) is not in the first B. Z. Also note that \(\gamma_{k_1} = 0\). Thus, from 5i we obtain
\[
\langle 4J_2 \gamma_{k_2} + 4J_4 \gamma_{k_4} \rangle S(k) e^{\frac{\epsilon}{k_1}} = 2\lambda S(k) e^{\frac{\epsilon}{k_1}}
\]
\[
\langle 4J_2 \gamma_{k_2} + 4J_4 \gamma_{k_4} \rangle S'(k) e^{\frac{\epsilon}{k_1}} = 2\lambda' S'(k) e^{\frac{\epsilon}{k_1}}
\]
Thus, \( \mu = \lambda + \lambda' = 4J_2 \chi_{K_2} + 2J_4 (x_{K_4} + y_{K_4}) \) and there is no required relation between \( S(k) \) and \( S'(k) \).

(vi): We summarize the preceding. A necessary condition for \( H \) to be a minimum with respect to variations in \( S(k) \) and \( S'(k) \) is that:

\[
\mu = 4J_2 \chi_{K_2} + 2J_4 (x_{K_4} + y_{K_4}) \pm 16J_4 \chi_{K_1}\]

where \( H = N \mu S^2 \).

If \( \chi_{K_1} \neq 0 \) then \( S'(k) = \pm S(k) \).

If \( \chi_{K_1} = 0 \) then no required relation exists between \( S'(k) \) and \( S(k) \).

The above \( \pm \) cases are not independent, the pluses correspond and the minuses correspond.

(vii): In any case \( S(k) \) and \( S'(k) \) must satisfy one of the conditions 4iv\( \alpha \), 4iv\( \beta \), 4iv\( \gamma \). The cases \( \beta \) and \( \gamma \) involve only a few discrete \( k \) values; however, for case \( \alpha \) there is no restriction on the possible values of \( k \). If the state has \( S(k) \neq 0 \) only for one value of \( |k| \), then I call it a single mode state; we have shown in 5iii that it is sufficient to consider only such states when determining the ground state. Thus, conditions 4iii\( \alpha \), 4iii\( \beta \), and 4iii\( \gamma \) are used.

(viii\( \alpha \)): Since for case \( \alpha \) there is no restriction on the possible \( k \) values, we shall find the minimum of \( \mu \) with respect to variations of \( k \).

\[
\mu = 4J_2 \cos k \sin z + 4J_4 (\cos k \alpha)(\cos \beta \alpha)(\cos \gamma \alpha) \pm 16J_4 (\cos \alpha \gamma)(\cos \beta \gamma)(\cos \alpha \beta) \]

\[
\frac{\partial \mu}{\partial z} = -4c \left[ J_2 \sin k \sin \frac{z}{2} + \frac{1}{2} J_1 \cos \frac{2\alpha \gamma}{2} \cos \frac{2\beta \gamma}{2} \sin \frac{2\alpha \beta}{2} \right] = 0
\]

A: \sin k \cos \frac{2\alpha \gamma}{2} = J_1 \cos \frac{2\alpha \gamma}{2} \cos \frac{2\beta \gamma}{2} - 0

\[
\frac{\partial \mu}{\partial \alpha} = -4a \left[ J_4 \sin \alpha \alpha \cos \beta \gamma \pm \frac{1}{2} J_1 \sin \frac{2\alpha \gamma}{2} \cos \frac{2\beta \gamma}{2} \cos \frac{2\alpha \beta}{2} \right] = 0
\]

\[
\frac{\partial \mu}{\partial \beta} = -4a \left[ J_4 \cos \alpha \alpha \sin \beta \gamma \pm \frac{1}{2} J_1 \cos \frac{2\alpha \gamma}{2} \sin \frac{2\alpha \beta}{2} \cos \frac{2\beta \gamma}{2} \cos \frac{2\alpha \beta}{2} \right] = 0
\]

Adding or subtracting these we get
\[ J_4 \sin \left( \frac{k_x + k_y}{2} \right) a + 2 J_2 \sin \left( \frac{a x + b y}{2} \right) \cos \left( \frac{a x^2}{2} + \frac{b y^2}{2} \right) = 0 \]

Thus,

\[ \sin \left( \frac{k_x + k_y}{2} \right) \frac{a}{2} \left[ \frac{J_4}{2} \cos (k_x + k_y) a + J_4 \cos \left( \frac{a x^2}{2} \right) \right] = 0 \]

A, B, and C are all necessary conditions for \( \mu \) to be a minimum with respect to variations in \( k \).

(5vii) To satisfy A we must have either

I: \( k_a c = 0 \) or II: \( J_2 \cos \frac{a x^2}{2} = \frac{1}{2} J_4 \cos \frac{a x^2}{2} + J_2 \cos \frac{a y^2}{2} \). To satisfy B we must have either \( R: \frac{a x^2}{2} = \frac{a y^2}{2} \) or \( k_a a = \frac{a y a}{2} = \pi \), or

C: \( J_4 \cos (k_x + k_y) \frac{a}{2} = \pm J_1 \cos \frac{a y^2}{2} \).

To satisfy C we must have either

r: \( k_a a = k_a \frac{a}{2} \) or s: \( J_4 \cos (k_x - k_y) \frac{a}{2} = \pm J_1 \cos \frac{a y^2}{2} \). We consider all possible combinations of these conditions

\( \alpha 1 = I \rho r: k_x c = 0, k_a a = \frac{1}{2} k_y a, \) and so, \( k_c - k_a a = 0 \)

\( \alpha 5 = I \rho s: k_x a = 0, k_a a = k_a y, \) and so, \( J_4 \cos k_a a = \pm J_1 \)

\( \alpha 3 = I \sigma r: k_x c = 0, k_a a = k_a y \)

\( \sigma s: \) No solution

\( \alpha 6 = I \sigma r: k_x c = 0, k_a a = k_a y, \) and so, \( J_4 \cos k_a a = \pm J_1 \)

\( \alpha 7 = I \sigma s: k_x c = 0, k_a a = \frac{1}{2} k_y a, \) and so, \( J_4 \cos \frac{k_x}{2} \cos \frac{k_y}{2} = \pm J_1 \)

\( \alpha 8 = II \rho r: k_a a = k_a a = 0, \) and so, \( J_2 \cos \frac{a x^2}{2} = \pm J_1 \)

\( \alpha 10 = II \rho s: k_a a = -k_a y, \) and so, \( J_2 \cos \frac{a x^2}{2} = \pm J_1 \left( \cos \frac{a y^2}{2} \right) \) and

\( J_4 \cos \frac{a x^2}{2} = \pm J_1 \left( \cos \frac{a x^2}{2} \right)^2 \) and thus,

\( J_1 \left( \cos \frac{a x^2}{2} \right)^2 \) when \( \frac{k_x}{J_1} \) and \( \frac{k_y}{J_1} \) are interchanged.
\[
\cos \frac{a \alpha}{2} = -\frac{1}{2} \left( \frac{J_2 J_4}{J_1^2} \right) \left( 1 - 2 \left( \frac{J_2 J_4}{J_1^2} \right)^2 \right) = \frac{1}{2} \left( 1 - \frac{J_2^2}{J_1 J_4} \right),
\]
thus, \[
J_1 \cos \frac{a \alpha}{2} = \pm J_4 \left[ 2 \left( \cos \frac{a \alpha}{2} \right)^2 - 1 \right] = \pm J_4 \left[ \left( \frac{J_2^2}{J_1 J_4} \right) \right] \left( 2 - \frac{J_2^2}{J_1 J_4} \right).
\]

\[\alpha = 2 \text{II} \sigma : k \alpha = k \alpha = \pi; k \beta = \pi\]

\[\alpha = 3 \text{II} \sigma ^* : \text{No solution}\]

\[\alpha = 9 \text{II} \sigma : k \alpha = k \alpha, J_2 \cos \frac{a \alpha}{2} = \pm J_4 \left( \cos \frac{a \alpha}{2} \right)^2, \text{ and} \]

\[J_4 = \pm J_4 \cos \frac{a \alpha}{2}; \text{ thus, } J_4 \left( \cos \frac{a \alpha}{2} \right)^2 = \frac{J_2}{J_4^2};\]

\[\left( \cos \frac{a \alpha}{2} \right)^2 = \frac{J_2}{J_4^2} \]

\[\alpha = 2 \text{II} \epsilon ^* : k \alpha = \cos \frac{a \alpha}{2} = J_4 \cos \left( \frac{a \alpha}{2} \right) = \pm J_4 \cos \left( \frac{a \alpha}{2} \right) \text{ and } J_2 \cos \frac{a \alpha}{2} = \pm J_4 \cos \frac{a \alpha}{2}; \text{ thus, } \cos \frac{a \alpha}{2} = 0 \]

Now for the energies:

\[\mu = 4J_2 \cos k \alpha + 4J_4 \cos k \beta \left( \cos \frac{a \alpha}{2} \right) \cos \frac{a \alpha}{2} \]

\[\alpha = 1 \text{I} \rho : \mu = 4J_2 + 4J_4 \pm 16J_1 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\alpha = 5 \text{I} \rho ^* : \mu = 4J_2 + 4J_4 \left( \frac{J_2}{J_4} \right)^2 + 16J_1 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\alpha = 3 \text{I} \sigma : \mu = 4J_2 + 4J_4 \]

\[\alpha = 6 \text{I} \tau : \mu = 4J_2 + 4J_4 \left( \frac{J_2}{J_4} \right)^2 \pm 8J_1 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\alpha = 7 \text{I} \psi : \mu = 4J_2 + 4J_4 \left( 2 \left( \cos \frac{a \alpha}{2} \right)^2 - 1 \right) \pm 16J_1 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\alpha = 8 \text{II} \rho : \mu = 4J_2 \left( \frac{J_2}{J_4} \right)^2 - 1 \right) + 4J_4 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\alpha = 9 \text{II} \rho ^* : \mu = 4J_2 - 4J_4 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\alpha = 2 \text{II} \epsilon : \mu = -4J_2 - 4J_4 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\alpha = 10 \text{II} \psi : \mu = 4J_2 \left( \frac{J_2}{J_4} \right)^2 - 1 \right) + 4J_4 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]

\[\mu = -4J_2 - 4J_4 \cos \frac{a \alpha}{2} \cos \frac{a \alpha}{2} \]
(6i): The conditions 4iii \( \alpha \), 4iii \( \beta \), 4iii \( \gamma \), 5vi, and 5viii allow no other states than those on page Eb-12 to be the ground state (leaving out states degenerate because of symmetry reasons).

(6ii): It is appropriate here to consider a point which was omitted in section 4iii. For cases where \( \gamma_0 \neq 0 \) then from 5vi we have \( S(k) = \pm S(k) \) so that the spins on the two sublattices are related. In section 4 we considered the requirement that \( S_1^2 = S \) be invariant under lattice translations; for cases \( \alpha \) and \( \beta \) \( S'(k) = \pm S(k) \) implies that \( S_1 = S^2 \); however, case \( \gamma \) is not sufficient to guarantee that \( S_1 = S_1 \) = \( S^2 \). In order for case \( \gamma \) to be satisfied on both sublattices when \( S'(k) = \pm S(k) \) we must have \( \sin 2k \cdot r_j = 0 \) for \( r_j = \) translation vector between two cations on the same or on different sublattices, therefore, \( 2(k \cdot \frac{a}{x} \pm k \cdot \frac{b}{y} \pm k \cdot \frac{c}{z}) = 2n \bar{r} \) and we have the additional requirement that \( k_1 = \pm \frac{\pi}{2} \) must occur in pairs (e.g., \( k = (\frac{\pi}{2}, \frac{\pi}{2}, \bar{r}), (\frac{\pi}{2}, -\frac{\pi}{2}, 0), \) etc.). This means that states \( \sigma_1, \sigma_2, \sigma_3, \) and \( \sigma_0 \) are not allowed as ground states.

(7i): We must now find the stable states for the various values of the exchange parameters. For every case where \( \mu \) contains a term in \( J_1 \), there is a state of the same total energy for both \( J_1 = + |J_1| \) and \( J_1 = - |J_1| \); therefore, it is sufficient to consider \( J_1 = + |J_1| \). Since we are here interested in the stable states and not the total energy, we may divide all energies by \( |J_1| \) (if \( J_1 \neq 0 \)), and therefore we may take \( j_2 = J_2/|J_1| \) and \( j_4 = J_4/|J_1| \) as the two independent variables. We will then partition the \( j_2-j_4 \) plane into regions where the various structures are stable. Note that the analysis in 5viii yields all structures which are stable over an area in the \( j_2-j_4 \) plane of measure greater than zero.*

* \( 810 \) is degenerate with \( \beta 4 \) for a finite area but that is for a special reason, see the footnote on the bottom of page Eb-12.
### Table Eb-1: Energies of the Magnetic States

<table>
<thead>
<tr>
<th>$k_1$</th>
<th>$k_2$</th>
<th>$k_3$</th>
<th>$j_{K1}$</th>
<th>$j_{K2}$</th>
<th>$j_{M}$</th>
<th>$j_{N}$</th>
<th>$j_{P}$</th>
<th>$j_{Q}$</th>
<th>$H$ *</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4$J_2$ + 4$J_4$ + 16$J_1$</td>
<td></td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-4$J_2$ + 4$J_4$</td>
<td></td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>4$J_2$ - 4$J_4$</td>
<td></td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-4$J_2$ - 4$J_4$</td>
<td></td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4$J_2$ + 4$J_4$</td>
<td></td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-4$J_2$ + 4$J_4$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0</td>
<td>0</td>
<td>$\pi$/2</td>
<td>$\pi$/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4$J_1$ + $\pi$2$J_1$</td>
<td>NOT ALLOWED *</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>$\pi$/2</td>
<td>0</td>
<td>$\pi$/2</td>
<td>1</td>
<td>0</td>
<td>4$J_1$ + $\pi$2$J_1$</td>
<td>NOT ALLOWED *</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>$\pi$/2</td>
<td>$\pi$/2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$\pm 2J_1$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_4$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>$\pi$/2</td>
<td>0</td>
<td>$\pi$/2</td>
<td>1</td>
<td>0</td>
<td>-4$J_2$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_5$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4$J_1$ + 8$J_1$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_6$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>$\pi$/2</td>
<td>0</td>
<td>$\pi$/2</td>
<td>1</td>
<td>-1</td>
<td>4$J_1$ + 8$J_1$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_7$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>1</td>
<td>$\pi$/2</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>±4$\sqrt{2}J_1$</td>
<td>NOT ALLOWED *</td>
</tr>
<tr>
<td>$\gamma_8$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>1</td>
<td>$\pi$/2</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>±4$\sqrt{2}J_1$</td>
<td>NOT ALLOWED *</td>
</tr>
<tr>
<td>$\gamma_9$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>1</td>
<td>$\pi$/2</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>-4$J_2$ + 4$J_4$ ***</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{10}$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>1</td>
<td>$\pi$/2</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>-4$J_2$ + 4$J_4$ ***</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{11}$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4$J_2$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4$J_2$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{14}$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4$J_2$</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{15}$</td>
<td>$\pi$/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4$J_4$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>SAME AS $\beta_1$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>SAME AS $\beta_4$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>SAME AS $\alpha_5$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>SAME AS $\alpha_6$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4$J_2$ - 4($J_2$, $J_4$, $J_1$) ± 8$J_1$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_6$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4$J_2$ + 4$J_4$ - 8($J_2$, $J_4$, $J_1$)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_7$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4$J_2$ + 4$J_4$ - 8($J_2$, $J_4$, $J_1$)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_8$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4$J_2$ + 4$J_4$ - 8($J_2$, $J_4$, $J_1$)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_9$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4$J_2$ + 4$J_4$ - 8($J_2$, $J_4$, $J_1$)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_{10}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4$J_2$ + 4$J_4$ - 8($J_2$, $J_4$, $J_1$)</td>
<td></td>
</tr>
</tbody>
</table>

* See 611
** $S'(k) = \pm S(k)$ if $\mathcal{R}_n \neq 0$
*** $\gamma_9$ and $\gamma_{10}$ have $\mathcal{R}_n = 0$, therefore, there is no required relation between $S'(k)$ and $S(k)$; this allows combinations of the two to be used to maximize the $J_4$ interaction which I have done.
Therefore, neither $\beta_2, \beta_3$ or any of the $\gamma$ states can be stable except at isolated points or lines. However, this conclusion is true only for the exact rutile structure, it will result that at least one structure which is stable only at a point for the exact rutile structures becomes stable over a region of the $j_2 - j_4$ plane if small distortions in the rutile structure are allowed. For this reason we will carry these states along in the analysis.

(7ii): Comparing the various $\beta$ and $\gamma$ states we obtain the lowest states for the various $j_2$ and $j_4$ values as shown in Figure Eb-1. Comparing the various $\gamma$ states we arrive at the lowest states for the various $j_2$ and $j_4$ values as shown in Figure Eb-2 (remember that for some states the possible $j_2$ and $j_4$ are restricted in 5vii in order to keep the magnitude of the argument of each cosine term from becoming greater than one). Figures Eb-1 and 2 are for the convenience of the reader who wishes to check the calculations. Considering $\alpha$, $\beta$, and $\gamma$ states together the lowest states are those shown on Figure Eb-3. The stable magnetic structure $\beta 1$ is exhibited by MnF$_2$, FeF$_2$, CoF$_2$, and NiF$_2$ (ERICKSON (1953)) and the stable magnetic structure $\alpha 8$ is exhibited by MnO$_2$ (YOSHIMORI (1959)). The $\beta 1$ structure yields a ferromagnetic arrangement of spins (all parallel) if $J_1$ is negative; this is the structure of CrO$_2$ (GUILLAUD, et al (1944)). It should be pointed out that Yoshimori, using Erickson's data for single crystal MnO$_2$ (which Erickson could not interpret), used an analysis similar to the preceding one to find the $\alpha 8$-type stable magnetic structure which is consistent with the experimental data. Yoshimori used the same $J_1$ and $J_2$ interactions but he used a physically unreasonable third type of exchange inter-
Figure 1: The Lowest Allowed States Among the $\beta$ and $\gamma$ Cases

Figure 2: The Lowest States Among $\alpha_{5}$ Through $\alpha_{10}$
Figure 1: The Lowest Allowed States Among the $\beta$ and $\gamma$ Cases

Figure 2: The Lowest States Among $\alpha_5$ Through $\alpha_{10}$
Figure 3: The Stable Magnetic Arrangements For the Rutile Structure

\( \alpha 3 \) is stable at the point \( j_2 = j_4 = 1 \) and is degenerate there with \( \beta 1, \beta 4, \gamma 10, \alpha 7, \) and \( \alpha 8 \).

\( \alpha 5 \) and \( \alpha 6 \) are stable on the line \( j_4 = +1, j_2 \leq +1 \) and are degenerate there with \( \alpha 7 \) and \( \beta 1 \); on this line \( \alpha 5 \) and \( \alpha 6 \) are identical with \( \beta 1 \).

\( \alpha 9 \) and \( \alpha 10 \) are stable on the line \( j_2 j_4 = 1, 0 < j_4 \leq 1 \) and are degenerate there with \( \alpha 8 \) and \( \beta 4 \); on this line \( \alpha 9 \) and \( \alpha 10 \) are identical with \( \alpha 8 \).

\( \alpha 1 \) is identical with \( \beta 1 \).

\( \alpha 2 \) is identical with \( \beta 4 \).
Figure 3: The Stable Magnetic Arrangements For the Rutile Structure

\( \gamma 3 \) is stable at the point \( j_2 = j_4 = 1 \) and is degenerate there with \( \beta 1, \beta 4, \gamma 10, \alpha 7, \) and \( \alpha 8 \)

\( \alpha 5 \) and \( \alpha 6 \) are stable on the line \( j_2 = \pm 1, j_2 \leq \pm 1 \) and are degenerate there with \( \alpha 7 \) and \( \beta 1 \); on this line \( \alpha 5 \) and \( \alpha 6 \) are identical with \( \beta 1 \)

\( \alpha 9 \) and \( \alpha 10 \) are stable on the line \( j_2 j_4 = 1, 0 \leq j_4 \leq 1 \) and are degenerate there with \( \alpha 8 \) and \( \beta 4 \); on this line \( \alpha 9 \) and \( \alpha 10 \) are identical with \( \alpha 8 \)

\( \alpha 1 \) is identical with \( \beta 1 \)

\( \alpha 2 \) is identical with \( \beta 4 \)
action. From the start he used the same Lagrange multipliers for each sublattice. He also failed to discuss the $\mathbf{g} \neq 0$ terms in 3ii. But most important he completely failed to consider the $\gamma_3$ method of satisfying the spin invariance (see part 4); we will see in section Ec that one of these $\gamma$ states ($\gamma_3$) explains the distorted rutile structure exhibited by VO$_2$, MoO$_2$, WO$_2$, ReO$_2$, and TcO$_2$. Since the type of analysis is not restricted to the rutile structure, we can hope that some other cases of distorted structures may be similarly explained as a result of a $\gamma$-type of magnetic order.
Ec: The Distorted Rutile Structure

From Figure Eb-3 the $\gamma 3$ structure is stable if $j_2 = 1$ and $j_4 = 1$. In this structure each cation has neighbors at $\pm c$; one with the same spin and one with the opposite spin. (This results since

$$S_i = \sum_k S(k)e^{ikr} = 2u_4 \cos(k \cdot r + \frac{\pi}{2})$$

and $k \cdot c = \frac{\pi}{2}$; thus, as $r$ increases in the $c$ direction, we get

$+s_i +s_i -s_i -s_i +s_i +s_i -s_i -s_i$ etc.). Therefore, if we let the neighboring atoms in the $c$ direction with opposite spin come closer together (and, therefore, the neighboring atoms with the same spin get farther apart), then the exchange integral for the favorable pair increases while that of the unfavorable pair decreases (assuming $J_1 > 0$, trivial changes if $J_1 < 0$) and we get a change in $\mu$ of $(-4J_4' + 4J_4'')$. Thus, for $\gamma 3$

$$\mu = \pm 8J_1 - 4(J_4'' - J_4') = \pm 8J_1 - 4\alpha J_1$$

where $\alpha = \frac{J_4'' - J_4'}{J_1} > 0$.

(Ec-1)

So, if the movements of the various cations are consistent with each other (they are, see below), we see that the rutile structure is not stable for the $\gamma 3$ magnetic arrangement, but, in fact, the $\gamma 3$ magnetic arrangement implies a distortion of the crystal structure which then makes the $\gamma 3$ magnetic arrangement stable with respect to $\beta 1, \beta 4, \alpha 7,$ and $\alpha 8$ for a region of the exchange integrals. The region of stability of $\gamma 3$ for $x = \frac{1}{2}$ is shown on Figure Ec-1.

It is now necessary to see what restrictions on the $\gamma 3$-stabilizing movements of the cations are imposed by the crystal structure. Each cation is surrounded by six anions which are located approximately as if they were at the vertices of an octahedron centered at the cation.
Figure Ec-1: The Stability Region For the \( \gamma^3 \) Magnetic Arrangement For \( x = 1/2 \)
Figure Ec-1: The Stability Region For the $\gamma 3$ Magnetic Arrangement For $x = 1/2$
(e.g., anions 1, 2, 3, 4, 5, and 6 "surround" cation A). If A and B displace toward each other along the line between them, they conflict with anions 5 and 6 which are already very close together. The extravalence electrons on the cations have been previously assumed to be in states non-interacting with the states which contribute to the binding of the solid. For this reason the spatial regions of high density of extravalence electrons are in regions away from the anions (i.e., "point" toward the midpoint of the plane formed by anions 134, 234, 123, 124, 356, 456, 345, or 346 for atom A). Consequently, if cation A moves from the center of the octahedron, it will move toward one of the faces of the octahedron. If it is to move toward B, then it must approach face 356 or 456; similarly, if B is to move toward A, it must move toward face 567 or face 568. But note that if A moves toward face 356 then A approaches C and if A moves toward face 456 then A moves toward D. If D and C have different spins, say +S and -S, respectively, and if A has spin +S then A will move toward atom C rather than atom D (if \( J_{1} > 0 \)), etc. The above conclusions will be used to see which, if any, distorted structures can result from the \( Y3 \) magnetic arrangement.
From before, $S_1 = \gamma_4 \cos(k \cdot r_1 + \frac{\pi}{2})$, $S'_1 = -2\gamma_4 \cos(k \cdot r_1 + \frac{\pi}{4})$ and $k_x a_c = \frac{\pi}{2}$, $k_y a = 0$, $k_z c = \frac{\pi}{2}$. This information is used to determine the orientation of the magnetic moments throughout the crystal as shown on Figure Ec-2 (left hand column, periodicity is 4a, a, 4c in the x, y and z directions). Letting the ions displace we get the distorted structure shown in the right hand column. It is important to note that this distorted structure is completely determined by the above considerations (it was not obvious that the movements of the ions would all cooperate to allow the distortion, but they do). The arrow indicates the displacement in the x-y plane and the U or D indicates the direction of displacement in the z direction (the total displacement is in the direction perpendicular to the face of the octahedron). The crystallographic structure is determined by the pattern on the right. We see that the size of the unit cell (orthorhombic) is 2a, a, 2c in the x, y and z directions. Actually, closer examination shows that the true primitive cell is half the size of the orthorhombic cell and is monoclinic (same y and z primitive vectors, but a new x primitive vector such that $\alpha'_x = \frac{1}{2} \alpha_x - \frac{1}{2} \alpha_z$. In fact this structure is identical with the VO$_2$ structure!! So, the VO$_2$ structure which is exhibited by VO$_2$, MoO$_2$, TcO$_2$, ReO$_2$, and WO$_2$ is derived here as the only stable structure for a certain range of "exchange integrals" (if the cations are such that they may be displaced from the center of the anion octahedron). Note that the required distortion also results in an additional (not included in Ec 1) energy lowering for $\gamma_3$ due to the $J_4$ interaction. For L - 2 electron states (i.e., d electrons) the fivefold degeneracy is partially lifted for cubic symmetry; in which case, there are three states at
The magnetic structure is on the left. The crystallographic structure is on the right.

0 indicates an anion in the same plane.

Layers separated by $\Delta z = 2c$ have reversed spin.

The orthorhombic magnetic unit, cell (not primitive) $a' = 4a$, $b' = a$, $c' = 4c$.

The orthorhombic crystallographic unit cell (not primitive) has $a' = 2a$, $b' = a$, $c' = 2c$.

FIGURE Ec-2: The $\delta 3$ Magnetic Arrangement and the Resulting (VO$_2$) Distorted Rutile Crystal Structure
The magnetic structure is on the left. The crystallographic structure is on the right.

0 indicates an anion in the same plane.

Layers separated by $\Delta z = 2c$ have reversed spin.

The orthorhombic magnetic unit cell (not primitive) $a' = 4a, b' = a, c' = 4c$.

The orthorhombic crystallographic unit cell (not primitive) has $a' = 2a, b' = a, c' = 2c$.

FIGURE Ec-2: The $\delta 3$ Magnetic Arrangement and the Resulting $(\text{VO}_2)$ Distorted Rutile Crystal Structure
one energy (we call these $t_2$ states) and two states at another energy (we call these $e$ states). If the cubic symmetry is due to an octahedron of anions, then the three-fold level has a lower energy (this is assuming that these electrons are not used for binding, which is appropriate here). For the three fold states the electrons avoid the anions; for the two fold states the electrons are concentrated near the anions. Therefore, if the atom has any $e$ electrons, it will be forced to stay in the center of the octahedron, since any distortions push the $e$ electrons into very unfavorable surroundings; so, if extravalence electrons are present, (which is appropriate for transition element cations of the V through Ga columns for oxides and the Sc through Cu columns for fluorides then the $\gamma 3$ magnetic arrangement can occur only for cations which have one to three extravalence electrons, viz., V, Cr, Mn, Nb, Mo, Tc, Ta, W, Re for the oxides (the fluorides with one to three valence electrons do not have the rutile structure). Five of these, as mentioned, do, in fact, have the VO$_2$ structure which, I say, implies that they have the $\gamma 3$ magnetic arrangement (this prediction can be checked by experiment), one case MnO$_2$ has the $\alpha R$ magnetic arrangement, one case CrO$_2$ is ferromagnetic, one case TaO$_2$ has been studied only fairly carefully using x-ray diffraction and is reported to be undistorted rutile (actually it was shown to have a period of length c in z direction), and one case NbO$_2$ has a distorted structure different from VO$_2$. According to my analysis the NbO$_2$ distorted structure (which using my analysis should correspond to a rather complicated magnetic arrangement involving four independent magnetic substructures) is always unstable with respect to VO$_2$; I have no expla-
ation for NbO$_2$ which is simultaneously consistent with the previous
analysis and with the reported x-ray diffraction data. If TaO$_2$ has
exchange integrals such that it is in the stability region of $\beta$ ($\beta$
(quite possible noting the region in which VO$_2$ is stable), then the
$\gamma$10 structure would result ($\beta$ and $\gamma$10 have same energy for the un-
distorted case, but $\gamma$10 allows an energy-lowering distortion) and the
distortion in this case results in an x and y period of $\sqrt{2}a$ but no
change in the z period -- this structure is consistent with the litera-
ture; neutron diffraction experiments if they verify the predicted
magnetic structure would imply, then, the larger crystallographic unit

cell. The above deductions serve to indicate approximate values of ex-
change integrals for the various compounds; it should be noted that
there is no known relation between these exchange integrals and the ac-
tual state of affairs of the electrons and nuclei (e.g., wave functions,
perturbing potentials, and the like) and this prevents us from making
any meaningful estimates of the changes of these exchange integrals
under distortion.

One more point should be mentioned. All of the regular rutile
structure antiferromagnets have Néel points below 100$^\circ$K, but VO$_2$ has
a Néel point of 340$^\circ$K. This is consistent with the above explanation
of VO$_2$ structure since the distortion allows a much greater energy
stabilization of the magnetic structure.

Using the known Néel points of VO$_2$ and MnO$_2$ and assuming that the
$J_1$ exchange integral is the same in the two cases (this may be a poor
assumption since the Mn has three $t_2$ electrons while the V has one;
thus, they interact differently with the oxygens involved in the super-
exchange type $J_1$ interaction) we may calculate for VO$_2$ the $J_\mu - J_\nu$, i.e.,
the change in exchange parameter with distance times the distance. We
use $\mathcal{E}TN = \frac{\mu (\omega_2 - \omega_3)}{3}$. YOSHIDORI (1959, p.811) estimates for MnO$_2$, $J_2 = 1.60$
and $j_4 = \frac{1}{2}$ (he uses $j_3$ but the energy expressions for $\alpha \delta$ differ by my
$j_4$ being replaced by his $2j_3$). Thus, for MnO$_2$
$$\frac{\mathcal{E}TN}{J_1} = \frac{5}{4} \left[ x y_2 + x y_4 + \frac{y_8}{2} \right] = 11.75; T_N = 90^\circK, \text{ thus, } z + x = 46 \text{ and }$$
x = 44 = (J_\mu - J_\nu)/J_1; \text{ thus, } J_\mu - J_\nu = \frac{44}{1600} = \frac{1}{36} \text{ ev. This value seems un-
reasonably high. If a value of 10 times the previous } J_1 \text{ is used, the}
new } x = 2.6 \text{ which is much more reasonable and } J_\mu - J_\nu = \frac{2.6}{160} = \frac{1}{60} \text{ ev.}$$
The displacement of the V from the center of symmetry is 0.24 \AA (ANDERSSON
(1956)); thus, (averaging) $J_2$ changes as $\frac{1}{30}$ ev per \AA; note that the
effect of changes in $J_4$ has been neglected; inclusion of $J_4$ would de-
crease the calculated change in $J_2$.

The point to be made about the results of section E is that al-
though the crystallographic structure was known beforehand, the magnetic
structure was not (and is still not known experimentally). Thus, the
theory besides predicting some results which were known has predicted a
magnetic structure which is unique and is certainly complicated enough
to have never been suspected a priori. As a matter of fact, if VO$_2$ has,
in fact, the magnetic structure $\gamma_3$, then I conjecture that it would
have been very difficult to successfully interpret the experimental neu-
tron diffraction data even if it were available (this may explain why no
one has reported studies of this obvious candidate for neutron diffraction
study). If the predicted magnetic structure actually occurs, then we
have some justification in trying the same method on some other ionic

* Even the $++-+-++++$ alternation of spins along the nearest-
neighbor directions is, I believe, unique.
oxides. On the other hand, if the predicted magnetic structure for \( \nu_2 \)
does not occur (and similarly for the one for \( \text{TaO}_2 \)) then we may simply
discard the theory.
APPENDIX A:

Summary: We prove that for tableau $S^x_i$ and $S^y_j$ like this

$$
\begin{array}{c|c}
1 & 2 \\
3 & 4 \\
\vdots & \vdots \\
2n-1 & 2n \\
2n+1 & \\
\end{array}
$$

if $p \leq 2n$ and if $q$ is not in the same row as $p$ then

$$U^x_{ij}(p,q) = \frac{1}{2} S_{ij}$$

Lemma: Let $S_i = \sigma^y_{ij} S_j$, in order for $U^x_{ij}$ to be non-zero,

then for any expansion of $\gamma$ in elementary transpositions it must be possible to get some expansion of $\sigma^y_{ij}$ in elementary transpositions by deleting elementary transpositions from the expansion of $\gamma$.

Proof: Remember from Aa-1 that all diagonal elements of $U^x_{ij}(k,k+1)$ are non-zero but the off-diagonal elements are non-zero iff

$$S_i = (k,k+1) S_j \quad U^x_{ij} = \sum_{m, r} \sum_{p, q} U^x_{i s m} U^x_{m r} \cdots U^x_{p q r} U^x_{q i k}$$

where:

$$\gamma = \gamma_1 \cdots \gamma_r$$

is an expansion of $\gamma$ in elementary transpositions.

But $U^x_{ij} \neq 0$ iff $q - j \in \gamma$ and $q - \gamma_j \in \gamma$. If $q - j$ then

$$U^{pq}_{\gamma_1} \neq 0$$

iff $p = j$ or $p = \gamma_j$, $j \in (r-1)j$. If $q = rj$ then

$$U^{pq}_{\gamma_r} \neq 0$$

iff $p = rj$ or $p = (r-1)rj$. We carry this on until we arrive at $\gamma_1$, by which time we have $2^{n-1}$ (not all different) possibilities for $s$; then we find that either $s = i$ or $s = \gamma_i \in \gamma$ for non-zero contributions to $U^x_{ij} \gamma$. Thus, all non-zero contributions to $U^x_{ij} \gamma$ result from ordered selections of elementary transpositions from $\gamma$ such that $S_i = \gamma_1 \cdots \gamma_c S_j$. We note that $\sigma^y_{ij} = \gamma_1 \cdots \gamma_c$ so $\gamma_1 \cdots \gamma_c$ is an expansion of $\sigma^y_{ij}$ in elementary transpositions.

If no such set of $\gamma_1 \cdots \gamma_c = \sigma^y_{ij}$ exists then $U^x_{ij} \gamma = 0$. Note that this result is independent of the particular choice of the expansion of $\gamma$ in elementary transpositions.

QED

THEOREM a: If $d_{p,p+1} = \xi_i$ then $U^{d}_{ij}(p,q) = U^{d}_{ij}(p+1,q)$
Proof:  \((p,q) = (p,p+1)(p+1,q)(p+1,p)\)

\[
U_{ii}(p,q) = \sum_{jk} U_{ij}(p,p+1) U_{jk}(p+1,q) U_{ki}(p,p+1) = \\
= (U_{ii}(p,p+1))^2 U_{ii}(p+1,q) = U_{ii}(p+1,q)
\]

Define a tableau lowest-in-2n as one like

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
| ... | ...
| 2n-1 | 2n |
| 2n-1 | ...

Theorem 8.1: If \(S_i^q\) and \(S_j^q\) are lowest-in-\(2n\) and \(p \leq 2n \leq q\), then

\[U_{ij}(p,q) = U_{ij}(2n,q)\]

Proof:  \((p,q) = (p,p+1)(p+1,q)(p+1,p);\)  \(U_{ij}(p,q) =

\[
= \sum_{ks} U_{ik}(p,p+1) U_{ks}(p+1,q) U_{sj}(p+1,p).
\]

From equation Aa-1

\[U_{ik}(p,p+1) \neq 0\] only for \(s_k - s_i\) and \(s_k = (p,p+1)s_i\), denote the latter by \(k = pi\). \(S_i\) and \(S_j\) are the same in \(p+1\) and \(p\). Thus,

\[U_{mj}(p+1,p) = U_{mi}(p+1,p);\] but, from the lemma if \(S_i = (p,p+1)s_k\), then

\[S_m = (p,p+1)s_j\]. Thus, if \(S_j = \neq S_i\), then \(S_m = \neq S_k\); denote \(j\) by \(\neq i\) and \(m\) by \(\neq k\). \(S_k\) and \(S_m\) are the same in \(p\) and \(p+1\); so,

\[U_{mj}(p+1,p) = U_{zk}(p+1,p) = U_{ki}(p+1,p);\]

\[U_{ij}(p,q) = \left( (U_{ii}(p,p+1))^2 U_{ij}(p+1,q) + (U_{i,pi}(p,p+1))^2 U_{pi,pj}(p+1,q) \right)^2 \]

If \(p\) is odd, then \(U_{ij}(p,p+1) = \delta_{ij}\) and thus \(U_{ij}(p,q) = U_{ij}(p+1,q)\); so we can take \(p\) to be even. If \(p = 2n\) the theorem is proved; thus, assume \(p < 2n\). The \(p+1\) is odd so \(U_{ij}(p+1,q) = U_{ij}(p+2,q)\) and

\[U_{pi,pj}(p+1,q) = U_{pi,pj}(p+2,q) = U_{ij}(p+2,q)\]

\[(U_{i,pi}(p,p+1)) = 1 - (U_{ii}(p,p+1))\); so, \(U_{ij}(p,q) = U_{ij}(p+1,q)\). We continue this until \(U_{ij}(p,q) = U_{ij}(2n,q)\)

QED
Theorem bii: If $p < q < 2n$ then $U_{ij}(p,q) = -\frac{1}{2} \delta_{ij}$

Proof: Take $q$ odd since if $q$ is even, $U_{ij}(p,q) = U_{ij}(p,q-1)$. Then stop iteration in Theorem bi when $p = q - 1$; so,

$U_{ij}(p,q) = U_{ij}(q-1,q)$. If $i = j$, $U_{ii}(q-1,q) = -\frac{1}{2}$ (from equation Aa-1). If $i \neq j$ then by the lemma $U_{ij}(q-1,q) = U$.

Theorem c: For a tableau like $S_i$, $U_{ii}(p,q) = -\frac{1}{2}$

$S_i = \begin{array}{|c|}
\hline
p \\
p+1 \\
p+2 \\
\vdots \\
q \\
\hline
\end{array}$

Proof: This follows immediately from repeated application of part a.

Theorem di: For a tableau like $S_i$,

$U_{ii}(p,p+2) = -\frac{1}{2}$

Proof: Directly from Theorem a

Theorem dii: For a tableau like $S_i$,

$U_{ii}(p,p+3) = -\frac{1}{2}$

Proof: $U_{ii}(p,p+3) = [(U_{ii}(p+3,p+2))^2 + (U_{ii}(p+2,p+3))^2 U_{ii}(p+3,p+2)]$

since $U_{ii}(p+3) = 0$. But from Theorem di

$U_{ii}(p+2,p+2) = -\frac{1}{2}$, and from Theorem c $U_{ii}(p,p+2) = -\frac{1}{2}$.

From equation Aa-1 $(U_{ii}(p+2,p+2))^2 = 1 - (U_{ii}(p+3,p+2))^2$;

thus, $U_{ii}(p,p+3) = -\frac{1}{2}$

Theorem diii: For a tableau like $S_i$,

$U_{ii}(p,q) = -\frac{1}{2}$

Proof: All the steps in Theorem dii are the same in this case (pre-
cumbing the cases for small $q$ to be already done).
Theorem div: For a tableau like $S_i$, $U_{ii}(p, r) = -\frac{1}{2}$. $S_i = \begin{align*} p & \quad \text{p} \\ p+1 & \quad \text{q} \\ p+2 & \quad q+1 \\ \vdots & \quad \vdots \\ q-1 & \quad r \\ q & \quad r+1 \end{align*}$

Proof: This follows from Theorem dill and repeated application of Theorem A.

Theorem eii: For a tableau like $S_i$, $S_i = \begin{align*} p & \quad \text{p} \\ q & \quad q+1 \\ \vdots & \quad \vdots \\ q-1 & \quad q+1 \end{align*}$

Proof: $U_{ii}(p, q+1) = \left[U_{ii}(q, q+1)\right]^2 U_{ii}(p, q) + (U_{ii}(q, q+1))^2 U_{qi, qi}(p, q)$;

$U_{ii}(p, q) = -\frac{1}{2}$ from Theorem d; $U_{qi, qi}(p, q) = -\frac{1}{2}$ from Theorem c;

thus, $U_{ii}(p, q+1) = -\frac{1}{2}$

Theorem eiii: For a tableau like $S_i$, $S_i = \begin{align*} p & \quad \text{p} \\ q & \quad q+1 \\ \vdots & \quad \vdots \\ q-2 & \quad q+1 \\ q-1 & \quad q+1 \end{align*}$

Proof: $U_{ii}(p, q) = \left[U_{ii}(q, q+1)\right]^2 U_{ii}(p, q) + (U_{ii}(q, q+1))^2 U_{qi, qi}(p, q)$;

$U_{ii}(p, q) = -\frac{1}{2}$ from div; $U_{qi, qi}(p, q) = -\frac{1}{2}$ from Theorem eii; thus,

$U_{ii}(p, q) = -\frac{1}{2}$

Theorem eiv: For a tableau like $S_i$, $S_i = \begin{align*} p & \quad \text{p} \\ p+1 & \quad \text{q} \\ p+2 & \quad q+1 \\ \vdots & \quad \vdots \\ q-1 & \quad r+1 \\ q & \quad r+1 \end{align*}$

Proof: All the steps are the same as in Theorem eii (presuming the cases for lower $r$ to be already due).

Theorem f: Further additions of strings of elements successively in one column and then the other follow in the same way
as in theorem d, c, and e. Any tableau of the form
\[
\begin{array}{c}
  \vdots \\
  p+1 \hline \\
  p \\
  \vdots
\end{array}
\]

can be constructed from the general case by selecting various lengths for the different strings. The general result is: For a tableau of the form
\[
\begin{array}{c}
  \vdots \\
  p+1 \\
  \vdots
\end{array}
\]

if \( q > p \) then
\[ U_{ii}(p,q) = - \frac{1}{2} \]

Now consider \( U_{ij}(p,q) \) and require \( i \neq j \).

**Theorem g1:** For the tableau of Theorem c, \( U_{ij}(p,q) = 0 \) since there is only one tableau lowest-in-2n for that shape.

**Theorem g2:** For the tableau of Theorem d, \( U_{ij}(p,q) = 0 \); same reason as for Theorem g1.

**Theorem g3:** For the tableau of Theorem d, \( S_j = (p+2, p+3)S_i \)
\[
U_{ij}(p,q) = [U_{ii}(p+2, p+3)U_{ii}(p+3, p+2) + U_{ii}(p+3, p+2)U_{ii}(p+2, p)]
\]

But, from parts c and d \( U_{jj}(p+2, p) = U_{ii}(p+2, p) = - \frac{1}{2} \) and from equation A-1 \( U_{ii}(p+3, p+2) = - U_{jj}(p+3, p+2) \). Thus, \( U_{ij}(p, p+3) = 0 \).

**Theorem g4:** Continuing as in Theorem d, \( U_{ij}(p,q) = 0 \) since all steps in above case are still valid.

**Theorem g5:** The equivalent of div, c and e for \( U_{ij}(p,q) \) follow just as in the above cases where we always take
\[
U_{ij}(p,q) = [U_{ii}(p-1, p)U_{ii}(p-1, p)U_{ii}(p-1, p) + U_{ii}(p-1, p)U_{ii}(p-1, p)]
\]

which results since \( U_{ik}(q-1, p) \) has been previously shown to be diagonal; thus, \( S_i = (q, q-1)S_j \). So, \( U_{ii}(q, q-1) = - U_{jj}(q-1, q) \); then we use \( U_{ii}(q-1, p) = U_{jj}(q-1, p) \) which results from Theorems c, d, e, and f; so, we obtain \( U_{ij}(p,q) = 0 \) if \( i \neq j \).
Appendix B:

Consider \( H = \text{sum of all elements in a class } C \) of the group \( G \). By the definition of a class, if \( g \in G \) the \( gh = Hg \); so, by Schur's lemma (LOMONT (1959) p. 7) for every irreducible representation of the group, \( U_H = a E \) where \( U_H \) represents \( H \), \( a \) is a constant, and \( E \) represents \( C \).

If \( G = S_N \) and \( C \) is the class of transpositions, then we obtain

\[
\sum_{\tau \in \Sigma} U_{\tau}^{\alpha} = a \chi_{\alpha}^{\tau}.
\]

Taking the trace \( \alpha \chi_{\alpha}^{\tau} = f_{\alpha}^{\tau} \), where \( \chi_{\alpha}^{\tau} \) is the character (trace) of a transposition (it is the same for all transpositions), \( t \) is the number of transpositions in \( S_N \), and \( f_{\alpha}^{\tau} \) is the degree of the representation \( \alpha \). \( t = N(N-1)/2 \). From LOMONT (1959) p. 263:

\[
f_{\alpha}^{\tau} = \binom{N}{2} - \binom{N}{1}; \text{ p. 264}: \chi_{\alpha}^{\tau} = -\left( N^2 - 2N \right) \left( \frac{N}{2} + s \right)^2 + \frac{N}{2} \left( \frac{N}{2} - 1 \right)
\]

\[
N = (N-k) + k = \left[ \frac{N^2}{2} + s \right] + \left[ \frac{N}{2} - s \right]
\]

\[
\chi_{\alpha}^{\tau} = -\frac{(N-2)! \left( \frac{N}{2} - s \right) \left( \frac{N}{2} + s \right) \left( \frac{N}{2} - s - 1 \right) \left( \frac{N}{2} + s - 1 \right)}{\left( \frac{N}{2} - s \right) \left( \frac{N}{2} + s \right) \left( \frac{N}{2} - s - 1 \right) \left( \frac{N}{2} + s - 1 \right)}\frac{A}{2N + 2S(S+1)}
\]

\[
A = \left( \frac{N}{2} - s \right) \left( \frac{N}{2} + s \right) \left( \frac{N}{2} + s + 1 \right) + \left( \frac{N}{2} - s \right) \left( \frac{N}{2} - s - 1 \right) \left( \frac{N}{2} + s - 1 \right) \left( \frac{N}{2} + s - 1 \right) - \left( \frac{N}{2} - s \right) \left( \frac{N}{2} + s \right) \left( \frac{N}{2} + s + 1 \right) - \left( \frac{N}{2} - s \right) \left( \frac{N}{2} + s \right) \left( \frac{N}{2} + s + 1 \right)
\]

\[
= 2S + 1 \left[ \frac{N^2}{2} - 2N + 2S(S+1) \right]
\]

\[
\chi_{\alpha}^{\tau} = \frac{N(N-1) \left[ \frac{N^2}{2} - 2N + 2S(S+1) \right]}{\left( \frac{N}{2} - s \right) \left( \frac{N}{2} + s \right) \left( \frac{N}{2} - s - 1 \right) \left( \frac{N}{2} + s - 1 \right)}
\]

\[
f_{\alpha}^{\tau} = \frac{N! \left( \frac{N}{2} + s + 1 \right) - \left( \frac{N}{2} - s \right)}{\left( \frac{N}{2} + s + 1 \right) \left( \frac{N}{2} - s \right)} = \frac{N! (2S+1)}{\left( \frac{N}{2} + s + 1 \right) \left( \frac{N}{2} - s \right)}
\]

\[
t = N(N-1)/2
\]

Thus, \( \chi_{\alpha}^{\tau} = -\frac{\frac{N^2}{2} - 2N + 2S(S+1)}{2} = -\frac{N(N-1)}{4} - S(S+1) \). Thus,

\[
\sum_{\alpha \in C} U_{\alpha}^{\tau} = -a \chi_{\alpha}^{\tau} \left[ \frac{N(N-1)}{4} + S(S+1) \right]; \text{ (this above formula is given by LOWDIN (1955a) p. 1476). Also, } X_{\tau} = -\frac{\chi_{\alpha}^{\tau}}{4} \left[ \frac{N(N-1)}{4} + S(S+1) \right].
\]
APPENDIX C:

Theorem: If $S_i$ or $S_j$ is lowest-in-2n and if $p < q < n$, then

$$U_{ij}(2p-1,2q-1)(2p,2q) - S_{ij}.$$ (Note that the proof of this theorem is not complete.)

Proof: Since $U_{ij}$ is unitary, it is sufficient to prove that

$$U_{ii}(2p-1,2q-1)(2p,2q) = 1 \cdot (2p-1,2p)(2p-1,2q-1)(2p,2q) = (2p,2q-1)(2p-1,2q)(2p-1,2p)$$
and $U_{ik}(2p-1,2p) = S_{ik}$ thus,

$$U_{ii}(2p-1,2q-1)(2p,2q) = U_{ii}(2p,2q-1)(2p-1,2q).$$ The elements in
the rows before $p$ and the rows after $q$ are irrelevant; so, for
$p-q = 1$ it is sufficient to let $p = 1$ and $q = 2$, for $p-q = 2$ it is
sufficient to let $p = 1$ and $q = 3$ etc. The only value for $i$ is 1.

(i) Consider $p-q = 1$, then $S_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

$$(23)(14) = \frac{(23)(123)(34)(23)(12)}{(13)}; \text{ hence,}$$

$$U_{11}(23)(14) = \sum_{m} U_{1m}(1,3) U_{mm}(3,4) U_{m1}(2,3).$$ Denote $(k,k+1)S_1$ as

$S_k$; denote $(k,k+1)(m,m+1)S_1$ as $S_{km}$, etc.

$$U_{11}(23)(14) = (-\frac{1}{2}) U_{11}(1,3) + (-\frac{\sqrt{3}}{2}) U_{12}(1,3) = \frac{1}{4} + \frac{3}{4} = 1$$

(ii) Consider $p-q = 2$, then $S_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

$$(25)(16) = (2,3,4)(5,4,3,2)(1,2,3,4,5)(6,5,4)(4,3,2)(1,2)$$

$$U_{11}(25)(16) = \sum_{i} (-\frac{1}{2})^3 U_{1(i,5)} + (-\frac{1}{2})\left[ U_{21}(1,5) + U_{22}(1,5) + U_{12}(1,5) \right] +$$

$$+(-\frac{1}{2})(-\frac{\sqrt{3}}{2}) \left[ U_{22}(1,5) + U_{22}(1,5) + U_{12}(1,5) \right] - (-\frac{\sqrt{3}}{2})^2 U_{12}(1,5).$$

From appendix A, $U_{11}(1,5) = -\frac{1}{2}$, I call terms like this no-jump
terms. By direct evaluation, $U_{11}(1,5) = U_{12}(1,5) = U_{12}(1,5) = 0$;
I call these one-jump terms.
\[ U_{2,2}(1,5) = U_{2,4}(1,5) = U_{1,24}(1,5) = -\frac{1}{2} \] I call these two-jump terms. \[ U_{2,24}(1,5) = -\frac{\sqrt{3}}{3} \] I call this a three-jump term. Thus,

\[ U_{11}(25)(16) = 1 \]

(iii) Consider \( p-q = 3 \), then \( S_1 = \)

\[
\begin{array}{cccc}
1 & 2 \\
3 & 4 \\
5 & 6 \\
7 & 8 \\
\end{array}
\]

\( (27)(1,8) = (2,3,4)(4,5,6)(7,1)(8,7,6)(6,5,4)(4,3,2)(1,2) \) The one- and three-jump terms all are zero by direct evaluation. The no-jump term is \( \frac{1}{2} \) by appendix A. By direct evaluation,

\[ U_{2,2}(1,7) = U_{2,4}(1,7) = U_{2,6}(1,7) = U_{4,2}(1,7) = U_{4,2}(1,7) = U_{4,6}(1,7) = -\frac{1}{2} \]

\[ U_{24,24}(1,7) = U_{1,24}(1,7) = U_{1,24}(1,7) = U_{1,24}(1,7) = 0 \]

\[ U_{24,24}(1,7) = -\frac{1}{2} \]

\[ U_{24,24}(1,7) = U_{24,24}(1,7) = U_{24,24}(1,7) = U_{24,24}(1,7) = -\frac{1}{3} \]

\[ U_{24,24}(1,7) = -\frac{2}{3} \sqrt{3} \]. Thus, \( U_{11}(2,7)(1,8) = \)

\[ (-\frac{1}{2})(\frac{5}{\pi}) + (-\frac{1}{2})(\frac{5}{\pi})(\frac{1}{2}) + (-\frac{1}{2})(\frac{1}{2}) + (-\frac{1}{2})(\frac{1}{2})(\frac{2}{3}) + (-\frac{1}{2})(\frac{2}{3})(\frac{5}{\pi}) + (-\frac{1}{2})(\frac{5}{\pi})(\frac{2}{3}) = 1 \]

(iv) Trends in the above terms can be recognized from which some of the general terms may be guessed. Unfortunately, I have not been able, yet, to complete the inductive step. Parts (ii) and (iii) relied on no special properties of the \( S_1 \) for \( p-q = 2 \) and \( 3 \) thus, it is clear that the theorem will be true for the general case.
Appendix D: A Theorem Used in the Perturbation of the Code

Hartree-Fock State.

Theorem: Consider $\alpha = [2^n, 1^N 2^n]$. Let $\Phi$ have the standard order and obtain $\Phi_q$ from $\Phi$ by replacing the state $\phi_{iq}$ with some state $\phi_k$ for which $C_{1q}^{iq} \neq 0$. In the following cases

$$0_{rl}(p,q) \Phi_q = U_{ll}(p,q)0_{rl} \Phi_q$$

(i) $p$ and $q > 2n$.

PROOF (i): $k$ is not a doubly-occupied state of $\Phi_o$ since then $\Phi_q$ would have a triply occupied state; $k$ is not a singly occupied state of $\Phi_o$ since then $\Phi_f$ would have $n' = n + 1$ (THM Ab-1). Thus, $k$ is not occupied in $\Phi_o$; so, $0_{rl} \Phi = S_{jl}0_{rl} \Phi$;

$$0_{rl}(p,q) \Phi_q = \sum_j U_{lj}(p,q)0_{rl} \Phi_q = U_{ll}(p,q)0_{rl} \Phi_q$$

(ii) $p$ and $q < 2n_o$, $l_j < k$, and $p$ and $q$ not in the same row of $\Phi_o$.

PROOF (ii): Let $k = i_m$ and let $i_q = i_{q-1}$, then $(m, q-1) \Phi = \Phi'$ has the standard order and $n = n_o$. $0_{rl} \Phi = U_{ll}(m, q-1)0_{rl} \Phi'$;

$$0_{rl}(p,q) \Phi'_q = U_{ll}(p,q)(m, q-1)0_{rl} \Phi';$$

$(p, q)(m, q-1) = (q, q-1)(p, q-1)(m, q)(q, q-1)$. Thus,

$$U_{ll}(p, q-1)(p, q-1)(m, q)(q, q-1) = U_{ll}(p, q-1)(m, q).$$

If $p < q-1$, then

$$U_{ll}(p, q-1)(m, q) = U_{ll}(p, q-1)U_{ll}(m, q)$$

from the Lemma of appendix A.

If $p > q-1$, use $(p, q-1)(m, q-1)(p-1, q)(p-1, q) = (m, p)(p-1, q)$; then,

$$U_{ll}(p, q-1)(m, q) = U_{ll}(p, q-1)(m, q-1)(p-1, q)(p-1, q) = U_{ll}(m, p)(p-1, q) = U_{ll}(m, p)U_{ll}(p-1, q)$$

where appendix C and the Lemma of appendix A have been used. From appendix A, $U_{ll}(m, p) = U_{ll}(m, q)$ and $U_{ll}(p-1, q) = U_{ll}(p, q)$. Thus,
\( 0_{r_1(p,q)} \Phi_q = \frac{U_{i_1}(p,q)}{U_{i_1}(m,q)} \alpha_i \Phi_q = U_{i_1}(m,q) \alpha_i \Phi_q \)

(iii) \( p, q \leq 2n \) and \( k > n \).

**PROOF (iii):** Let \( i_{q-1} = i \) then \((q-1, 2n-1)(q, 2n) \Phi_q = \Phi'\) has the standard order. Using appendix C, \(0_{r_1} \Phi_q = \sum_j U_{l_j}(q, 2n)(q-1, 2n-1) 0_{r_j} \Phi' = U_{l_1}(q, 2n)(q-1, 2n-1) 0_{r_1} \Phi' \)

\( 0_{r_1(p,q)} \Phi_q = \sum_j U_{l_i(p,q)}(q-1, 2n-1) 0_{r_j} \Phi' \)

\((q, q-1)(p, q)(q, 2n)(q-1, 2n-1) = (p, q-1)(q-1, 2n)(q, 2n-1)(q, q-1).\)

Using the Lemma of appendix A and the appendix C

\( U_{l_j}(p, q)(q, 2n)(q-1, 2n-1) = U_{l_j}(p, q-1)(q-1, 2n)(q, 2n-1) = U_{l_1}(p, q-1) U_{l_1}(q-1, 2n)(q, 2n-1) \delta_{j_1} \cdot \) Thus,

\( 0_{r_1(p,q)} \Phi_q = U_{l_1}(p, q) 0_{r_1} \Phi_q \)
APPENDIX E:

I have not developed general simple formulae relating the proportional \( w_{rj} \) \( \chi \) (for different \( j \)) to each other. We know from THM Ab-2 that the proportional \( w_{rj} \) \( \chi \) all have the same array if the convention for \( \chi \) is used. As an example of how the proportionality constants can be found for any particular case we consider the partition \([4,2]\) and \( \chi = \alpha \beta \beta \beta \). Recall that \( w_{rj} \) \( \chi = \sum w_{rj} \) \( \chi = \sum u_{rj} w_{rj} \) \( \chi \)

\[
S_1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \quad S_2 = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 4 & 6 & 7 & 8 \end{bmatrix} \quad S_3 = \begin{bmatrix} 1 & 2 & 4 & 5 \\ 3 & 6 \end{bmatrix}
\]

\[
S_4 = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 4 & 5 & 7 & 8 \end{bmatrix} \quad S_5 = \begin{bmatrix} 1 & 2 & 4 & 6 \\ 3 & 5 \end{bmatrix} \quad S_6 = \begin{bmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 \end{bmatrix}
\]

\[
(\sum u_{rj}^{-1}) w_{rj} \chi + \sum u_{rj} w_{rj} \chi = 0
\]

\[
(1 - u_{12(45)}) w_{12} \chi = u_{12(45)} w_{12} \chi
\]

\[
(1 - u_{23(34)}) w_{23} \chi = u_{23(34)} w_{23} \chi ; \quad (1 - u_{22(56)}) w_{22} \chi = u_{22(56)} w_{22} \chi
\]

\[
(1 - u_{33(56)}) w_{33} \chi = u_{33(56)} w_{33} \chi
\]

BUT IF \( S_2 = (8, 8, 8, 1) S_1 \), THEN \( u_{12(45)} \) \( \chi \) = \( \sqrt{1 - u_{12(45)}^2} \)

\[
\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}
\]

Thus, \( \sqrt{1 - u_{11(45)}} w_{11} \chi = \sqrt{1 + u_{11(45)}} w_{12} \chi \). Note that \( u_{11(45)} = - u_{12(45)} \) ; Thus, \( \sqrt{1 + u_{11(45)}} w_{11} \chi = \sqrt{1 - u_{12(45)}} w_{12} \chi \).
\[ \omega_{42} x = \frac{\sqrt{1 - U_{42}(45)}}{\sqrt{1 + U_{42}(34)}} \omega_{41} x \]

\[ \omega_{43} x = \sqrt{1 - U_{43}(50)} \omega_{42} x = \frac{\sqrt{(1 - U_{22}(34))(1 - U_{42}(45))}}{\sqrt{(1 + U_{22}(34))(1 + U_{42}(45))}} \omega_{41} x \]

\[ \omega_{44} x = \frac{\sqrt{1 - U_{44}(34)}}{\sqrt{1 + U_{44}(34)}} \omega_{44} x \]

\[ \omega_{45} x = \sqrt{1 - U_{45}(45)} \omega_{44} x = \frac{\sqrt{(1 - U_{44}(45))(1 - U_{45}(45))}}{\sqrt{(1 + U_{44}(45))(1 + U_{45}(45))}} \omega_{41} x \]

Thus,

\[ \omega_{42} x = \sqrt{\frac{5}{3}} \omega_{41} x \]

\[ \omega_{43} x = \sqrt{\frac{4 - \frac{5}{3}}{2}} \omega_{41} x \]

\[ \omega_{44} x = \sqrt{\frac{4 - \frac{5}{3}}{2 - \frac{5}{3}}} \omega_{41} x \]

\[ \omega_{45} x = \sqrt{\frac{4 - \frac{5}{3}}{2 - \frac{5}{3}}} \omega_{41} x \]

\[ \omega_{46} x = \sqrt{\frac{3 + \frac{4 - \frac{5}{3}}{2}}{2 - \frac{5}{3}}} \omega_{41} x \]
APPENDIX F: The Matrices for the Orthogonal Representation
for N = 3 and 4

N = 3, \( \alpha = [2, 1] \)

\[ U_{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad U_{1,2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \]

\[ U_{(2,3)} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}; \quad U_{(1,3)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \]

\[ U_{(1,2,3)} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}; \quad U_{(1,3,2)} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \]

N = 4, \( \alpha = [2, 2] \)

\[ U_{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad U_{1,2,3} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad U_{1,3,4} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \]

\[ U_{(3,4)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad U_{(1,3,4)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \]

\[ U_{(1,4)}(2,3) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad U_{(1,3,2)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \]

\[ U_{(1,4,3,2)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad U_{(1,3,2,4)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \]

\[ U_{a}^{\alpha} = U_{a}^{\alpha}; \quad (1, 2, 3)(1, 3, 2) = \mathbb{1}, \quad (1, 3, 4)(2, 4, 3) = \mathbb{1} \]

Of all the irreducible representations of all the symmetric groups, the only irreducible representation of dimension greater than one which is not a faithful representation of the group is this one, \( \alpha = [2, 2] \) of \( S_{4} \).
\[ N = 4, \ \alpha = [3, 1] : \quad S_1 = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \quad S_2 = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \quad S_3 = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \]

\[ \alpha = [2, 1^2] : \quad S_1 = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \quad S_2 = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \quad S_3 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \]

The upper sign applies to \( \alpha = [3, 1] \) the lower sign to \( \alpha = [2, 1^2] \)

\[
\begin{align*}
U_{\Omega} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ; \quad U_{(1,2)} = \begin{pmatrix} \pm i & 0 & 0 \\ 0 & \pm i & 0 \\ 0 & 0 & \mp i \end{pmatrix} ; \quad U_{(2,3)} = \begin{pmatrix} \pm i & 0 & 0 \\ 0 & \frac{-i}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{-i}{2} \end{pmatrix} \\
U_{(3,4)} &= \begin{pmatrix} \frac{1}{3} & \frac{\sqrt{3}}{3} + \frac{i}{3} & 0 \\ \frac{\sqrt{3}}{3} - \frac{i}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} ; \quad U_{(1,2)(3,4)} = \begin{pmatrix} \frac{1}{3} & \frac{\sqrt{3}}{3} + \frac{i}{3} & 0 \\ \frac{\sqrt{3}}{3} - \frac{i}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} ; \quad U_{(2,3)(1,4)} = \begin{pmatrix} \frac{1}{3} & \frac{\sqrt{3}}{3} + \frac{i}{3} & 0 \\ \frac{\sqrt{3}}{3} - \frac{i}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{align*}
\]
Appendix G. A Little Bit About Group Theory

A representation: Let $G$ be a group of transformations on some vector space $S$. Let $a$ be some vector in the space and consider the quantities $G_1 a, G_2 a, ..., G_n a$ where $n$ is the order of the group. Then $G_j G_k a = G_{jk} a$ if $G_j G_k$, and we see that the set of quantities $G_1 a, G_2 a, ..., G_n a$ are transformed into each other by the operations of the group. We obtain a set of orthonormal basis vectors $b_1, b_2, ..., b_m$ by, say, using the Gram-Schmidt procedure on the previous set $(m \leq n)$. $G_i b_j$ equals a linear combination of $b_k$, say, $G_i b_j = D^i_{k j} b_k$; if $G_i G_j = G_r$; then $G_i G_j b_j = D^p_{k j} G_r b_r = D^p_{k j} G^q_{r j} b_r = G_r b_r = D^i_{r j} b_r$; thus, $D^i_{ij} = D^i_{kr} D^k_{rj}$ (since the $b_i$ are linearly independent). Therefore, the set of matrices $D^1, D^2, ..., D^n$ combine with each other in the same way that the group elements combine. Such a set of matrices is called a representation of the group $G$ (appendix F contains several representations of the symmetric group; note that the set of matrices $D^1 = 1, D^2 = 1, ..., D^n = 1$, where 1 is the real number, forms a representation of any group of order $n$, this is called the trivial representation).

An irreducible representation: If there exists some vector $c$ in the vector space $S$ (continuing with the above example) such that when the set $G_1 c, G_2 c, ..., G_n c$ is orthonormalized to obtain $f_1, f_2, ..., f_\mu$, where $\mu = m$, then the previous representation is said to have been reduced. If the set $f_1, f_2, ..., f_\mu$ is completed to give the space of $b_1, ..., b_m$ and if the new matrices representing the group elements are determined, then it will be found that: for every matrix, $D^1_{pq} = 0$ if $p < \mu$ or $q > \mu$ or if $p > \mu$ and $q < \mu$. The original space has been decomposed into two subspaces each of which is left invariant by the group.
operations. If the space spanned by the basis vectors of some representation cannot be decomposed into two non-null spaces, then the representation is said to be irreducible.

Consider the transformation group of order two operating on three-dimensional space where the elements are the identity (ε) and the inversion (i). If \( f(\overrightarrow{r}) \) is some function of \( \overrightarrow{r} \), then \( \varepsilon f(\overrightarrow{r}) = f(\overrightarrow{r}) \)
\( i f(\overrightarrow{r}) = f(-\overrightarrow{r}) \). Thus, \( f(\overrightarrow{r}) \) and \( f(-\overrightarrow{r}) \) correspond to the \( b_1, b_2, \ldots \).

\( \varepsilon f(\overrightarrow{r}) = f(\overrightarrow{r}), \varepsilon f(-\overrightarrow{r}) = f(-\overrightarrow{r}), i f(\overrightarrow{r}) = f(-\overrightarrow{r}), i f(-\overrightarrow{r}) = f(\overrightarrow{r}) \); thus, if \( G_1 = \varepsilon \) and \( G_2 = i, D_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \) and \( D_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \). The group relations are \( \varepsilon \varepsilon = \varepsilon \), \( i \varepsilon = \varepsilon \), \( \varepsilon i = i \varepsilon = i \), and we see that \( D_1 D_1 = D_1 \),
\( D_2 D_2 = D_1, D_1 D_2 = D_2 D_1 = D_2 \). But \([f(\overrightarrow{r}) - f(-\overrightarrow{r})]\) is contained in the space spanned by \( f(\overrightarrow{r}) \) and \( f(-\overrightarrow{r}) \); \( \varepsilon [f(\overrightarrow{r}) - f(-\overrightarrow{r})] = [f(\overrightarrow{r}) - f(-\overrightarrow{r})] \),
\( i [f(\overrightarrow{r}) - f(-\overrightarrow{r})] = -[f(\overrightarrow{r}) - f(-\overrightarrow{r})] \). Thus, the one dimensional space \([f(\overrightarrow{r}) - f(-\overrightarrow{r})]\) is invariant under the group and the previous representation has been reduced. The complementary space is \([f(\overrightarrow{r}) + f(-\overrightarrow{r})]\) for which \( \varepsilon [f(\overrightarrow{r}) + f(-\overrightarrow{r})] = i [f(\overrightarrow{r}) + f(-\overrightarrow{r})] = [f(\overrightarrow{r}) + f(-\overrightarrow{r})] \).

Consider the homogeneous differential equation \( H \psi_i = E_i \psi_i \); consider transformations of the space of eigenfunctions, \( \psi_i \). The set of transformations (say \( h \), leaving \( H \) invariant forms a group. (If \( G_k \in h \) and \( G_j \in h \), then \( G_k H = H G_k \) and \( G_k H = H G_k \)), thus,
\( G_k G_j H = G_k H G_j = H G_k G_j \); hence, \( i i = i \), \( iG_k G_j = G_k G_j \) then \( G_k \in h \); the unit element is in the set since it does not change \( H \) and \( \varepsilon H = H \varepsilon \); if \( G_k \in h \)
then consider \( H G_k^{-1} \), we have \( G_k H G_k^{-1} = H G_k^{-1} G_k = H \), thus \( G_k^{-1} (G_k H G_k^{-1}) = G_k^{-1} H \)
or \( G_k^{-1} H = H G_k^{-1} \), hence \( G_k^{-1} \in h \); thus, the set is a group). Let \( G_j \in h \)
operate on \( \psi_i = E_i \psi_i \); then \( E_j G_j \psi_i = G_j H \psi_i = H G_j \psi_i \); thus, \( G_j \psi_i \) is an eigenfunction of \( H \) degenerate with the eigenfunction \( \psi_i \). Continuing
in this way the functions of the set \( c_1 \psi_1, c_2 \psi_2, \ldots, c_n \psi_1 \) are all degenerate. As before we construct an orthonormal set \( \phi_1, \phi_2, \ldots, \phi_m \) (m \( \leq \) n) from these; the group \( h \) transforms the \( \phi_i \) among themselves and the transformation matrices form a representation of the group \( h \). All of the \( \phi_i \) are necessarily degenerate. Approaching the other direction if \( \kappa_1, \kappa_2, \ldots, \kappa_\mu \) are basis functions of an irreducible representation of \( h \) and if \( \kappa_1 \) is an eigenfunction of \( H \) then \( \kappa_2, \ldots, \kappa_\mu \) must all be eigenfunctions of \( H \) and degenerate with \( \kappa_1 \). For real physical systems it is almost everywhere true that each set of degenerate eigenfunctions of the Hamiltonian forms a basis for an irreducible representation of the symmetry group of the Hamiltonian. (Exceptions are called accidental degeneracies.) Note, however, that if the assumed symmetry group (say \( h' \)) of \( H \) is smaller than the real symmetry group (say \( h \)) of \( H \), then since \( h' \) is a subgroup of \( h \) the irreducible representations of \( h \) may be reducible for \( h' \); thus, some symmetry required degeneracies might appear to be accidental.

The different irreducible representations of \( \text{SO}(3) \) are labeled by numbers called the angular momentum, \( L \). The degeneracy of the irreducible representation \( L \) is \( 2L+1 \) and a set of basis functions for irreducible representation \( L \) is the set of spherical harmonics of degree \( L \), \( Y_{LM}(\theta, \phi) \). The eigenfunctions of the Hamiltonian can be required to transform as spherical harmonics.

Let \( C_1 \) be the group containing the identity and the inversion. This group has two irreducible representations both of dimension one. For one, the trivial representation, \( D^e = D^i = 1 \) and the basis function is invariant under \( i \); such a function is said to have even parity.
For the other irreducible representation \( \eta = 1 \), \( \eta^\dagger = -1 \) and the basis function changes sign under inversion; such a function is said to have odd parity.

For the Abelian group of order \( n \) (all the elements commute) all of the irreducible representations are one-dimensional. If \( t \) is the generator of the group then the \( n \) elements are \( \epsilon, t, t^2, \ldots, t^{n-1} \) (note \( t^n = \epsilon \)). The irreducible representations are labeled by the numbers \( k \), and the basis function of irreducible representation \( k \) changes as

\[
t^m \phi^k = C^{ink} \phi^k \quad (\text{note } nk = 1).
\]

The translation group in one dimension, \( T(1) \), is Abelian; so, let \( t = t(a) \) be the primitive translation (of magnitude \( a \)); then \( t^m = t(ma) \) and \( t(ma) \phi^k(x) = e^{ink} \phi^k(x) \).

So far, the \( m \) and \( k \) are dimensionless, it is convenient to let

\[
ma = M \quad \text{and} \quad \frac{k}{a} = K
\]

then \( MK = mk \). The three-dimensional translation group, \( T(3) \), is the direct product of three Abelian groups. The typical element is \( t(m_1a_1)t(m_2a_2)t(m_3a_3) \). This group is still Abelian; so, the irreducible representations are one-dimensional and

\[
t(m_1a_1, m_2a_2, m_3a_3) \phi^k(x_1, x_2, x_3) = e^{i(k_1m_1 + k_2m_2 + k_3m_3)} \phi^k(x_1, x_2, x_3).
\]

We let \( M = m_1a_1 + m_2a_2 + m_3a_3 \) (\( a_1, a_2, a_3 \) not required to be orthogonal) and \( K = k_1a_1 + k_2a_2 + k_3a_3 \) where \( a^i \cdot a^j = \delta_{ij} \) then,

\[
M \cdot K = m_1k_1 + m_2k_2 + m_3k_3 \quad \text{and} \quad t(M) \phi^k(\vec{x}) = \phi^k(\vec{x} + \vec{M}) = e^{i \vec{K} \cdot \vec{M}} \phi^k(\vec{X}).
\]

Such fundamental principles as conservation of momentum and energy arise from considerations of symmetry. Similarly, the selection rules are group theoretical. Consider diffraction by a solid: \( H \) is invariant under the translation group \( T(3) \); let \( \phi_\mathbf{R}(\vec{R}, \vec{R}_1, \ldots, \vec{R}_n) \) be the ground state eigentfunction of \( H \) (\( n \) is the number of electrons and nuclei in the solid). The initial situation is the solid in the ground state and a
free X-ray or neutron (it is a basis function of the Euclidean group
(all three-dimensional translations and all three-dimensional rotations)
and thus, transforms as $e^{ik \cdot x}$ under translation) with wave number $k$;
the final state is the solid in the ground state and a free X-ray or
neutron with wave number $k'$. Let $H'$ be the interaction between the
solid and the X-ray or neutron. The transition matrix element is
\[
I = \langle \psi_{k'}(\mathbf{r}_1', r_2', \ldots, \mathbf{r}_N') \mid H' \mid \psi_k(\mathbf{r}_1, r_2, \ldots, \mathbf{r}_N) \rangle.
\]
This quantity is invariant under lattice translation ($H'$ has the symmetry of the lattice);
so, we transform with $T(M)$, $\mathbf{L} = T(M) \mathbf{L} = e^{i\mathbf{k} \cdot \mathbf{r}} e^{i\mathbf{k'} \cdot \mathbf{r'}}$.
Thus, $e^{i(\mathbf{k} - \mathbf{k'}) \cdot \mathbf{r}} = 1$ or $(\mathbf{k} - \mathbf{k'}) \cdot \mathbf{r} = 2\pi n (n$ is an integer). Thus, if
\[
(\mathbf{k} - \mathbf{k'}) \cdot \mathbf{r} = 2\pi g_1 \hat{\mathbf{a}}^1 + g_2 \hat{\mathbf{a}}^2 + g_3 \hat{\mathbf{a}}^3
\]
then the $g_i$ are integers; such vectors as $\mathbf{g}$ are called reciprocal lattice vectors, and we see that the
$2\pi \mathbf{q}$ include the possible scattering vectors for all allowed diffractions. The other selection rules (e.g., absorption of light) follow in
a similar manner.

For a set of $N$ identical particles $H$ is invariant under the permuta-
tions of the particles; thus, $\psi_i$ is a basis function of an irreduc-
ible representation of the symmetric group, $S_N$ (see section A3 for some
discussion of $S_N$). If we consider some perturbation of the problem, $H'$,
then the transition matrix element is $\langle \psi_j' \mid H' \mid \psi_i \rangle$; but if the particles
are still identical, then $H'$ is also invariant under the operations of
$S_N$. Thus, $\langle \psi_j' \mid H' \mid \psi_i \rangle \neq 0$ only when the $\psi_i$ and $\psi_j$ transform as the
same basis function of the same irreducible representation. Thus, if
the system is in a state which belongs to an irreducible representation
of $S_N$ dimension greater than one, then the system is degenerate but the
degeneracy can never be detected as long as the particles are
identical. It is unsatisfactory to speak of things which are in principle undetectable: so, the whole problem is avoided by postulating that physical systems of identical particles must belong to one of the one-dimensional irreducible representations of $S_N$ (this postulate is called the Pauli principle). There are only two one-dimensional irreducible representations of $S_N$; these are the trivial representation and the alternating representation. For the trivial representation all permutations are represented by +1; for the alternating representation even permutations are represented by +1, and odd permutations are represented by -1. Particles whose many-particle systems have wave functions which are basis functions of the trivial representation (alternating representation) are called Bosons (Fermions). Electrons, protons and neutrons are Fermions (given that the electron has spin $1/2$, the footnote on page Db-1 implies that electrons are Fermions since, experimentally, the ground state of many-electron systems is not totally symmetrical and the spin is not $N/2$).

Probably the best place to begin the study of group theory is the first three chapters of NAMERMESH (1962).
APPENDIX H:

Summary:

(i)* Let \( S_{1}^{\alpha} = \begin{array}{cccc}
1 & 2 & \cdots & n \\
\frac{n}{n+1} & \frac{n+2}{n+1} & \cdots & \frac{n}{n+1}
\end{array} \) where \( \alpha =\begin{cases} n, & N-n \end{cases} \)

Then if \( \tau \) is the product of \( r \) disjoint transpositions each of which involves one element in the first row and one element in the second row of \( S_{1}^{\alpha} \), \( U_{1l}^{\alpha} = \binom{\alpha}{r} \binom{\nu}{r} \)

(ii)* if \( S_{1}^{\alpha} = \begin{array}{cccc}
1 & n+1 & \cdots & n \\
2 & n+2 & \cdots & n \\
\vdots & \vdots & \ddots & \vdots \\
n & n & \cdots & n
\end{array} \) where \( \alpha =\begin{cases} 2^{n}, & i^{N-2n} \end{cases} \)

then \( U_{lf}^{\alpha} = \binom{i}{r} \binom{\nu}{r} \). Two permutations are defined to be disjoint if no element of one appears in the other (e.g., the cycles of a permutation are disjoint).

Consider \( r = 1 \), let \( \tau = (ij) \), where \( i < n < j \) and \( \alpha =\begin{cases} n, & N-n \end{cases} \).

We note that \( U_{kl}(n,n-1,...,i) = \begin{cases} \tau_{kl}, & U_{kl}(n+1,n+2,...,j) = \tau_{kl} \end{cases} \)

\( U_{l1}(i,j) = U_{l1}(i,i+1,...,n)(n,j)(n,n-1,...,i) = U_{r1}(i,i+1,...,n)(n,j) = \)

\( = U_{l1}(i,i+1,...,n)(j,j-1,...,n+1)(n,n+1)(n,n+2,...,j) = \)

\( = U_{l1}(i,i+1,...,n)(j,j-1,...,n+1)(n,n+1) \).

Consider \( r = 2 \), let \( \tau = (k,m)(i,j) \), where \( k < i < n < j < m \)

\( U_{l1}(k,m)(i,j) = U_{l1}(k,m)(i,i+1,...,n)(j,j-1,...,n+1)(n,n+1) \); consider

\( U_{l1}(n,n-1,...,k)(i,i+1,...,n), k < i \) thus, by the lemma of appendix A the elementary transpositions involving letters \( i-1 \) cannot have off-diagonal contributions. Thus, the position of \( i-1 \) cannot be changed under any of the transpositions giving off-diagonal contributions.

* The proof of this theorem is not complete.
Hence, all permutations involving only letters $\leq i-1$ give unity contributions to $U_{\alpha e}$, for this reason I replace these transpositions by $\leq$ or neglect; thus, $U_{ij}(n,n-1,\ldots,k)(i,i+1,\ldots,n) = U_{ij}(n,i-1)$

$U_{tl}(k,m)_{(i,j)} = U_{tl}(k,k+1,\ldots,n)m(n,i-1)(j,j-1,\ldots,n+1)(n,n+1)$

Thus, $U_{ll}(1,1)(n,m)_{(i,j)} = U_{ii}(n-1,n+1)$

**Theorem App. H-1:** If the $2r$ inequalities $p < k < \ldots < i < n < j < \ldots < m < q$

are satisfied and $\alpha = (n, N-n)$, then

(i) $U_{tl}(p,q)_{(k,m)}(i,j) = \frac{(-1)^{r}}{C_{m}^{r}}$ where $A_{r} = (n+m)_{(m,m-1,\ldots,m+2)}(m+1, m+3, \ldots) \cdots (m+2r, m+2r+1, m+2r+2, \ldots) \times C_{m-r}^{r}$. 

(ii) $U_{ll}(p,q)_{(k,m)}(i,j) = (-1)^{r}/C_{m}^{r}$

**Proof (i):**

(i) It is true for $r = 1$ and 2. Assume it is true for $r-1$ $\Rightarrow$ s.

$U_{tl}(l_{r}q_{r})_{(k,m)}_{(i,j)} = \frac{(-1)^{r}}{C_{m}^{r}}$ where $A_{r} = (n+m)_{(m,m-1,\ldots,m+2)}(m+1, m+3, \ldots) \cdots (m+2r, m+2r+1, m+2r+2, \ldots) \times C_{m-r}^{r}$.
Proof (ii): \( U_{11A_1} = U_{11(n,n+1)} = -\frac{1}{n} \) (from \( A_{1-1} \))

\[
U_{11A_2} = U_{11(n,n+1)(n,n-1)(n+1,n+2)(n,n+1)} =
\]

\[
= \left( U_{i(i,m+1)} \right)^2 U_{i(i,m-1)(m+1,m+2)} (U_{h(k,m+1)})^2 U_{k(k,m-1)(m+1,m+2)}
\]

\[
= \left( \frac{1}{m} \right)^2 \left( 1 - \frac{1}{m^2} \right) \left( \frac{1}{m^2} \right)^2 = \frac{1}{m^4} + \frac{2m}{m^2(m-1)} = \frac{1}{m^2} \left( \frac{2}{m} \right)
\]

where \( S_{h} = (n,n+1)S_{1} \), \( S_{n} = \left( \begin{array}{cccc}
1 & 2 & \cdots & n-1 & n+1 \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
\end{array} \right) \frac{n}{n+2} \)

\[
U_{11A_3} = U_{11(n,n+1)(n,n-1,n-2)(n+1,n+2,n+3)(n,n+1)(n,n-1,n+1,n+2)(n,n+1)}
\]

We can go on calculating this for each \( r \) but for the proof we need the induction which I have not accomplished yet. What is needed is a way to generate expansions of \( \left( \frac{1}{m} \right)^r \). Thus, the proof is incomplete.

**Theorem App. H-2**: If \( p < k < \ldots < i < n < j < \ldots < m < g \) and \( \alpha = \lceil 2^n, 1^{N-2n} \rceil \), then

(i) \( U_{tf}(p,q)(k,m)..(i,j) = (-1)^{j-i} \cdot \cdots \cdot (-1)^{j-i} \cdot \left( \frac{1}{m} \right)^{\sum_{l=k}^{m} \alpha_{l}} \)

(ii) \( U_{ff}(p,q)(k,m) = (\frac{1}{m})^{\sum_{l=k}^{m} \alpha_{l}} \)
Proof (i): the proof is the same as for THM li where the elementary transpositions which were \( +1 \) before are \(-1\) which causes the factor of the parity of the \((p,\ldots,n)\) etc. to appear.

Proof (ii): the parity factor is squared for \( t = f \) so \( U_{ff}(\tau) = U_{ff}(\tau) \). just as in THM liii. We follow the same steps as before except that all terms are positive.

Theorem App. H-3: If \( p, k, \ldots, i \) are all different and \( s_{n}, \) and \( j, \ldots, m, q \) are all different and \( > n \), then

(i) if \( \alpha = [n, N-n] \), \( U_{ll}(p,q)(k,m)\ldots(i,j) = (-1)^{r}/(\gamma) \)

(ii) if \( \alpha = [2^{n}, 1^{N-2n}] \), \( U_{ff}(p,q)(k,m)\ldots(i,j) = 1/(\gamma) \)

Proof: \( U_{ll}(p,q)(k,m)\ldots(i,j) \) if \( \tau \) involves letters \( < n \) or if \( \tau \) involves letters \( > n \). But the similarity transformation \( \tau \) rearranges the \( p, q, \ldots, i \) among the transpositions. Thus, the \( p, k, \ldots, i \) can be put in the order of THM-1. Similarly, if \( \tau \) involves only elements \( > n \) the \( q, m, \ldots, j \) can be put in the order of THM-1. The proof for ii is exactly the same.

QED

Theorem App. H-4: If \( \alpha = [2^{n}, 1^{N-2n}] \) and for \( \tau_{r} = (p,q)(k,m)\ldots(i,j) \) of THM 2 and if \( \tau \) involves only elements \( < n \) or else elements \( > n \) then

\( U_{ff}(\tau_{r}) = U_{ff}(\tau) \)

Proof:

\( U_{ff}(\tau) = U_{ff}U_{ff}(\tau) = U_{ff}(\tau) = U_{ff}(\tau) \)

QED
APPENDIX 1:

Theorem: $\overline{U_{RS}} = \overline{U_{RS}} U_{RS}$

Proof: Consider $\mathcal{E} = (k, k+1)$, then $U_{RS}(k, k+1) = (\delta_{1k} + (1-k) \delta_{2k}) U_{RS}$

from (Aa-9). We write this as $U_{RS}(k, k+1) = \overline{U_{RS}} U_{RS}$

since $\overline{U_{RS}} = 1$ if $r \neq c$ and $U_{RS}(k, k+1)$ if $r = c$.

where $\xi, \xi', \ldots, \xi_m$ are elementary transpositions. But $\overline{U_{RS}} \overline{U_{RS}} \cdots \overline{U_{RS}} = \overline{U_{RS}}$

and $\overline{U_{RS}} \overline{U_{RS}} \cdots \overline{U_{RS}} = \overline{U_{RS}}$. QED
REFERENCES:


Brillouin, L., 1933, La Méthode du Champ Self-Consistent, Actualités Scientifiques et Industrielles, Herman and Co., Paris. This is issued in the form of a pamphlet and filed in the Cal Tech library as a book.


Koster, G. F., 1953, Quarterly Progress Reports, MIT Solid State and Molecular Theory Group, July 15, 1953, p. 37-38


Löwdin, P. O., 1962a, Rev. Mod. Phys. 34, 80-86.


Møller, C., and M. S. Plesset, 1934, Phys. Rev. 46, 618-622.


Pratt, Jr., G. W., 1956, Phys. Rev. 102, 1303-1307.


Ruedenberg, K., 1962, Rev. Mod. Phys. 34, 326-376.

Rutherford, D. E., 1948, Substitutional Analysis, Edinburgh University Press. It is a shame that such an important and useful book is out of print.


