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., H_2, HF, H_2O, 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 $H = \sum_{i=1}^{N} \frac{-\hbar^2}{2m} \nabla_i^2 + \sum_{i=1}^{N} V(r_i) + \sum_{i>j=1}^{N} \frac{e^2}{r_{ij}}$, where $V(r_i)$ 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, Φ, and any function of the spin coordinates of N particles, χ, 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 Φ are any N of these one-electron functions) the best $G_f^{\alpha}\Phi\chi$ 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 Φ . Due to the presence of doubly-occupied orbitals this method leads to rather large correlation errors. Using the G_i^{α} 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

Page

INTRODUCTION

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

$$H = \sum_{i=1}^{N} \frac{-\hbar^2}{2m} \nabla_i^2 + \sum_{i=1}^{N} V(r_i) + \sum_{j>i=1}^{N} \frac{e^2}{r_{ij}} \tag{A-1}$$

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 (ϕ_i) associated with an α spin and the same $\frac{N}{2}$ orthogonal molecular orbitals associated with a β spin. The orbitals, ϕ_i, are varied in order to minimize the energy. The minimum in the energy is called the Hartree-Fock energy, E_{HF}, and the orbitals which minimize the energy are called the Hartree-Fock orbitals, ϕ_i^{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, $H^{HF} \phi_i^{HF} = \epsilon_i \phi_i^{HF}$. The difference between E_{HF} and the exact energy of the ground state of A-1 is called the <u>correlation energy</u>. 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; SINANOǦLU (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 $\Psi(\tilde{r}_1,\tilde{r}_2,\cdots,\tilde{r}_N)$ (where $\tilde{r}_i$ is the spatial (r_i) and spin (s_i) coordinate of particle i). Let $\psi_i(\tilde{r}_j)$ be a function of the coordinates of one electron and let the set of ψ_i for all i form a complete orthonormal set; then (LÖWDIN (1955a)) the set of $\psi_i(\tilde{r}_1)\psi_j(\tilde{r}_2)\cdots\psi_k(\tilde{r}_N)$ 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

$$\Psi(\tilde{r}_1,\tilde{r}_2,\cdots,\tilde{r}_N)=\sum_{ij\cdots k} C_{ij\cdots k}\,\psi_i(\tilde{r}_1)\psi_j(\tilde{r}_2)\cdots\psi_k(\tilde{r}_N) \tag{A-2}$$

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 ψ_i are ordered, then

$$\Psi(\tilde{r}_1,\tilde{r}_2,\cdots,\tilde{r}_N)=\sum_{i<j<\cdots<k} C_{ij\cdots k}\ \mathrm{DET}\left\{\psi_i(1)\psi_j(2)\cdots\psi_k(N)\right\} \tag{A-3}$$

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

$$\psi_i(\tilde{r}_j)=\phi_i(r_j)\,\sigma_i(s_j) \tag{A-4}$$

We will take σ_i to be an eigenfunction of S_Z, viz., $\alpha(s_z\alpha=\frac{1}{2}\alpha)$ or

$\beta (S_z \beta = -\frac{1}{2} \beta)$. The complete set of one-electron functions, $\psi_i(\tilde{r}_j)$, 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 α and β form an orthonormal set. What about the one-electron spatial functions? If $\psi_i(\tilde{r}_j) = \phi_i(r_j)\alpha(j)$ and $\psi_m = \phi_m \alpha$, then ϕ_i and ϕ_m are orthogonal when ψ_i and ψ_m are. However, if $\psi_m = \phi_m \beta$ then we can say nothing about the relation between ϕ_i and ϕ_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 ψ_i with α spin and the other orthonormal set for expansion of all ψ_i with β spin. We call a one-electron spatial function an orbital. So, we have one complete set of orbitals ϕ_{ia} for α spin and another complete set of orbitals ϕ_{ib} (over the same space) for β spin. It has been extremely common in the past to require that the set of orbitals be the same for α and β 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 SINANOǦLU (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 H_2 and other two-electron systems, but, according to LÖWDIN (1959, p.304), no strict foundation for a generalization of the theory has been developed. Sinanoǧlu 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 (SINANOǦLU (1961a, 1962ab, 1963)). The Sinanoǧlu 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 α spin and one for the spin-orbitals with β 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 α spin and once with a β 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 <u>may</u> 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 <u>is</u> 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_i^α which upon operating on a function Φ of the spatial coordinates of N-electrons and a function χ of the spin coordinates of N-electrons yields a function $G_i^\alpha \Phi \chi$ which satisfies Pauli's principle and is an eigenfunction of total spin. Fortunately, the G_i^α operator has properties adapted to the later applications.

The G_i^α 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_i^α operator is used in conjunction with DODS in the manner described in the second paragraph before the last. The result is the best possible $G_f^\alpha \Phi \chi$ 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 ∞, the orbitals automatically change to atomic orbitals). As distinct from Hartree-Fock orbitals for nondegenerate, S = 0 molecular ground states, the GF orbitals are <u>not</u> required to be basis functions of the irreducible representations of the spatial symmetry group. Thus, the

GF orbitals are <u>not</u> 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

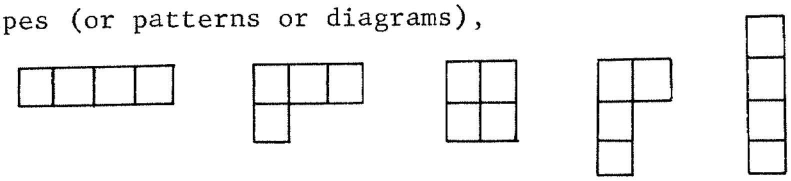
A Method of Constructing Many-Electron Wave Functions Which Satisfy Pauli's Principle and Are Eigenfunctions of S^2

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.

Aa*: The Symmetric Group and G_i^α

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 γ= (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),



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

---

* The first four paragraphs are based on RUTHERFORD (1948).

each row, we have a standard tableau, e.g., $\begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3\\\cline{1-1}\end{array}$ and $\begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2\\\cline{1-1}\end{array}$ but not $\begin{array}{|c|c|c|}\hline 1&4&3\\\hline 2\\\cline{1-1}\end{array}$. 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 $\alpha = [p^2,q,r,s^3]$ 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 $[p^2,q,r,s^3]$ 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 $U^{\alpha}_{ij\gamma}$ be the ij component of the matrix representing the permutation γ in the α 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 $\begin{array}{|c|c|c|c|}\hline 1&2&9&11\\\hline 3&5&10\\\cline{1-3} 4&7&12\\\cline{1-3} 6&8\\\cline{1-2}\end{array}$

$r_9=1$ and $c_9=3$, and $r_7=3$ and $c_7=2$. Define the axial distance d^i_{pq} between p and q in tableau i as $d^i_{pq} = (r_q-c_q)-(r_p-c_p)$, for our example $d_{97}=3$ (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

α by S_j^α, where j ranges from 1 to f^α and f^α is the dimension of the irreducible representation α. 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^α into S_i^α be σ_{ij}^α (i.e., $S_i = \sigma_{ij} S_j$, the subscript α is often suppressed below).

Now

$$\begin{aligned} U^\alpha_{ii(k,k+1)} &= \frac{1}{d^\alpha_{k+1,k}} && \text{(Aa-1)} \\ U_{ij(k,k+1)} &= 0 \quad \text{if} \quad S_i^\alpha \neq (k,k+1)\, S_j^\alpha \\ &= \sqrt{1-\left(\frac{1}{d^\alpha_{k+1,k}}\right)^2} = \sqrt{1-\left(\frac{1}{d^\gamma_{k+1,k}}\right)^2} \quad \text{IF} \quad S_i^\alpha = (k,k+1)\, S_j^\alpha \end{aligned}$$

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 same column of S_i, then $U_{ij(k,k+1)} = -\delta_{ij}$ and if k and k+1 are in same row of S_i then $U_{ij(k,k+1)} = +\delta_{ij}$. The matrices for N = 3 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.,

$$U^\alpha_{rs\tau} = U^\alpha_{sr\tau^{-1}} \qquad \text{(Aa-2)}$$

We mention the orthogonal units, O^α_{rs}, these are linear combinations of the elements of S_N; the set of these for the different α and r and s varying from 1 to f^α form a basis of the group algebra. For any shapes α and β of S_N and any standard tableaux r, s, p, and q

$$O^\alpha_{rs} O^\beta_{pq} = \delta^{\alpha\beta} \delta_{sp} O^\alpha_{rq} \qquad \text{(Aa-3)}$$

For any γ of S_N, (where $\theta^\alpha = N!/f^\alpha$ = an integer)

$$\tau = \sum_{\alpha,r,s} U^{\alpha}_{rs\tau}\, O^{\alpha}_{rs} \qquad \text{(Aa-4)}$$

$$O^{\alpha}_{rs} = \frac{1}{\theta^{\alpha}} \sum_{\tau\in S_N} U^{\alpha}_{rs\tau}\, \tau \qquad \text{(Aa-5)}$$

Therefore, $\tau_1 \tau_2 = \sum_{\alpha rs} \sum_{\beta pq} U^{\alpha}_{rs\tau_1} U^{\beta}_{pq\tau_2} O^{\alpha}_{rs} O^{\beta}_{pq} =$

$$= \sum_{\alpha rsq} U^{\alpha}_{rs\tau_1} U^{\alpha}_{sq\tau_2} O^{\alpha}_{rq} = \sum_{\alpha rq} U^{\alpha}_{rq\tau_1\tau_2} O^{\alpha}_{rq}$$

We denote the identity or unity permutation by $\in$ (stands for einheit); note that $U^{\alpha}_{pq\in} = \delta_{pq}$.

Some examples:

N=2
$$O^{[2]}_{11} = \frac{1}{2}(\in + (1,2)) \;; \qquad S^{[2]}_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array}$$
$$O^{[1^2]}_{11} = \frac{1}{2}(\in - (1,2)) \;; \qquad S^{[1^2]}_1 = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array} \qquad \text{(Aa-6)}$$

N=3
$$O^{[3]}_{11} = \frac{1}{6}\left[\in + (1,2,3) + (1,3,2) + (1,2) + (1,3) + (2,3)\right] \qquad S^{[3]}_1 = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline\end{array}$$

$$O^{[2,1]}_{11} = \frac{1}{6}(2\in - (1,2,3) - 1,3,2) + 2(1,2) - (1,3) - (2,3)) \;;$$

$$O^{[2,1]}_{12} = \frac{\sqrt{3}}{6}((1,2,3) - (1,3,2) - (1,3) + (2,3)) \;; \qquad S^{[2,1]}_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

$$O^{[2,1]}_{21} = \frac{\sqrt{3}}{6}(-(1,2,3) + (1,3,2) - (1,3) + (2,3)); \qquad S^{[2,1]}_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

$$O^{[2,1]}_{22} = \frac{1}{6}(2\in - (1,2,3) - (1,3,2) - 2(1,2) + (1,3) + (2,3))$$

$$O^{[1^3]}_{11} = \frac{1}{6}(\in + (1,2,3) + (1,3,2) - (1,2) - (1,3) - (2,3)); \qquad S^{[1^3]}_1 = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array}$$

It is general result that $\sum_{\alpha} \sum_{\lambda=1}^{f^{\alpha}} O^{\alpha}_{\lambda\lambda} = \in$

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, ζ_τ, 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).

$$\begin{aligned} O_{11}^{[N]} &= \frac{1}{N!} \sum_{\tau \in S_N} \tau \\ O_{11}^{[1^N]} &= \frac{1}{N!} \sum_{\tau \in S_N} \zeta_\tau \tau \end{aligned} \qquad \text{(Aa-8)}$$

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 $S_j^\alpha = \begin{array}{|c|c|c|} \hline 1 & 2 & 6 \\ \hline 3 & 5 & 7 \\ \hline 4 \\ \cline{1-1} \end{array}$ then $S_{\bar{j}}^{\bar{\alpha}} = \begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2} 6 & 7 \\ \cline{1-2} \end{array}$. Note that the associate of a standard tableau is a standard tableau and thus: $f^\alpha = f^{\bar{\alpha}}$. Since

$d_{(k+1,k)} = -d_{(k,k+1)}$ then

$$\begin{aligned} U_{\bar{i}\bar{i}(k+1,k)}^{\bar{\alpha}} &= -U_{ii(k+1,k)}^{\alpha} \\ U_{\bar{i}\bar{j}(k+1,k)}^{\bar{\alpha}} &= U_{ji(k+1,k)}^{\alpha} \qquad \text{if } j \neq i \end{aligned} \qquad \text{(Aa-9)}$$

We wish to consider linear combinations of permutations acting on two different spaces. We will let O_{rs}^α act only on the first space and ω_{pq}^β act only on the second space; of course, O_{rs}^α and ω_{pq}^β commute. Consider the expression*

$$G_i^\alpha \triangleq \sum_r \zeta_{\sigma_{ri}} O_{ri}^\alpha \omega_{\bar{r}\bar{i}}^{\bar{\alpha}} \qquad \text{(Aa-10)}$$

Remember that $S_r^\alpha = \sigma_{ri}^\alpha S_i^\alpha$. Let the transposition τ act on G_i^α; it then acts on both O^α and ω^α. From (Aa-4)

$$\tau = \sum_{\beta s t} U_{st\tau}^\beta O_{st}^\beta = \sum_{\gamma p q} U_{\bar{p}\bar{q}\tau}^\gamma \omega_{\bar{p}\bar{q}}^\gamma$$

---

* The symbol $\triangleq$ is read as "is defined as"

Thus, $$\tau G_i^{\alpha} = \sum_{\beta st}\ \sum_{\gamma \bar{p}\bar{q}}\ \sum_{r} \zeta_{\sigma_{ri}}\, U^{\beta}_{st\tau}\, U^{\gamma}_{\bar{p}\bar{q}\tau}\, 0^{\beta}_{st}\, 0^{\alpha}_{ri}\, \omega^{\gamma}_{\bar{p}\bar{q}}\, \omega^{\bar{\alpha}}_{\bar{r}\bar{i}} =$$

$$= \sum_{spr} \zeta_{\sigma_{rp}}\, \zeta_{\sigma_{pi}}\, U^{\alpha}_{sr\tau}\, U^{\bar{\alpha}}_{\bar{p}\bar{r}\tau}\, 0^{\alpha}_{si}\, \omega^{\bar{\alpha}}_{\bar{p}\bar{i}} \qquad \text{(Aa-11)}$$

Where we have used $\sigma_{ri} = \sigma_{rp}\sigma_{pi}$ and $\zeta_{\tau}\zeta_{\sigma} = \zeta_{\tau\sigma}$; also, note that $\zeta_{\tau} = \zeta_{\tau^{-1}}$

If τ is an elementary transposition, then from appendix I

$\zeta_{\sigma_{qr}} U^{\bar{\alpha}}_{\bar{q}\bar{r}\tau} = -U^{\alpha}_{qr}$. Thus,

$$\sum_{r} \zeta_{\sigma_{qr}}\, U^{\alpha}_{sr\tau}\, U^{\bar{\alpha}}_{\bar{q}\bar{r}\tau} = -\delta_{sq} \qquad \text{(Aa-12)}$$

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

$$(k,k+1) G_i^{\alpha} = -\sum_{s} \zeta_{\sigma_{si}}\, 0^{\alpha}_{si}\, \omega^{\bar{\alpha}}_{\bar{s}\bar{i}} = -G_i^{\alpha} .$$

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

$$\boxed{\tau G_i^{\alpha} = \zeta_{\tau} G_i^{\alpha}} \qquad \text{(Aa-13)}$$

When G_i^{α} operates on the N-electron space, the resulting function satisfies the Pauli principle.

Ab*: The Spatial and Spin Tensor Spaces

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 $\underline{x}$ in R_n is described by its components, x_i, on the basis vectors. If $\underline{\underline{a}}$, a transformation of the group G, transforms $\underline{x}$ into $\underline{x}$ then we write (where we use the summation convention, and $\in$ is read as "is an element of"),

$$\underline{x}' = \underline{\underline{a}}\underline{x} \quad \text{or} \quad x'_i = a_{ij}x_j \text{ , where } i \in \{1,..,n\} \qquad \text{(Ab-1)}$$

Consider the n^2 quantities x_iy_j formed from products of the components of the vectors $\underline{x}$ and $\underline{y}$ in R_n. When the transformation (Ab-1) is applied to the vectors of R_n, x_iy_j is transformed as $x'_iy'_j = a_{ik}a_{jl}x_ky_l$, i.e., the n^2 quantities x_iy_j transform according to $\underline{\underline{a}} \times \underline{\underline{a}}$, the Kronecker square of the transformation $\underline{\underline{a}}$. 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 $F_{ij...k}$ whose law of transformation is

$$F_{ij...k} = a_{il}a_{jm}\cdots a_{kp}F_{lm...p} \qquad \text{(Ab-2)}$$

forms an rth-rank tensor with respect to the group G.

Reconsider the case of second rank tensors, say $F_{i_1i_2}$, where $i \in \{1,2,..,n\}$. Under the transformation of (Ab-1) the set of tensors $(F_{i_1i_2} + F_{i_2i_1})$ transforms into itself and the set of tensors $(F_{i_1i_2} - F_{i_2i_1})$ transforms into itself. For example,

$$a_{j_ii_1}a_{j_2i_2}(F_{i_1i_2} \pm F_{i_2i_1}) = a_{j_1i_1}a_{j_2i_2}F_{i_1i_2} \pm a_{j_2i_2}a_{j_1i_1}F_{i_2i_1} =$$

$$= F_{j_1j_2} \pm F_{j_2j_1}$$

---

* The first two paragraphs are based on HAMERMESH (1962) Chapter 10.

Also, $a_{j_1 i_2} a_{j_2 i_1} (F_{i_1 i_2} \pm F_{i_2 i_1}) = \pm (F_{j_1 j_2} \pm F_{j_2 j_1})$. Thus,

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

$$(a_{j_1 i_1} a_{j_2 i_2} \mp a_{j_1 i_2} a_{j_2 i_1})(F_{i_1 i_2} \pm F_{i_2 i_1}) = 0 \tag{Ab-3}$$

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),

$$(pF')_{i_1 i_2 .. i_r} = F'_{i_{p(1)} i_{p(2)} \cdots i_{p(r)}} = a_{i_{p(1)} j_1} a_{i_{p(2)} j_2} \cdots a_{i_{p(r)} j_r} F_{j_1 j_2 \cdots j_r}$$

Now rename j_1 as $j_{p(1)}$, j_2 as $j_{p(2)}$, etc.; then,

$$(pF')_{i_1 i_2 \ldots i_r} = a_{i_{p(1)} j_{p(1)}} a_{i_{p(2)} j_{p(2)}} \cdots a_{i_{p(r)} j_{p(r)}} \quad (pF)_{j_1 j_2 \ldots j_r}$$

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

$$(pF')_{i_1 i_2 .. i_r} = a_{i_1 j_1} a_{i_2 j_2} \ldots a_{i_r j_r} (pF)_{j_1 j_2 .. j_r} \tag{Ab-4}$$

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 $F_{i_1 .. i_r}$ with the O^{α}_{pq} for S_r (the symmetric group on r objects). Then, in place of the n^r independent quantities $F_{i_1 \ldots i_r}$, we deal with a linearly independent set of n^r quantities selected from the $O^{\alpha}_{pq} F_{i_1 \ldots i_r}$. From (Ab-4)

$$(O^{\alpha}_{pq} F_{i_1 \ldots i_r}) = a_{i_1 j_1} a_{i_2 j_2} \ldots a_{i_r j_r} (O^{\alpha}_{p_q} F_{j_1 \ldots j_r}) \tag{Ab-5}$$

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 $F = \gamma F'$, where F' has a special order; then, $O^{\alpha}_{pq}\gamma$ yields a linear combination of O^{α}_{pr} involving various r. So, the space of O^{α}_{pq} with the same α 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).

$$O^{[2]}_{11} F_{i_1 i_2} = \frac{1}{2}(F_{i_1 i_2} + F_{i_2 i_1}) \text{ , this is a symmetric tensor.}$$

$$O^{[1^2]}_{11} F_{i_1 i_2} = \frac{1}{2}(F_{i_1 i_2} - F_{i_2 i_1}) \text{ , this is an antisymmetric tensor.}$$

$$O^{[2]}_{11} a_{i_1 j_1} a_{i_2 j_2} = \frac{1}{2}(a_{i_1 j_1} a_{i_2 j_2} + a_{i_1 j_2} a_{i_2 j_1})$$

$$O^{[1^2]}_{11} a_{i_1 j_1} a_{i_2 j_2} = \frac{1}{2}(a_{i_1 j_1} a_{i_2 j_2} - a_{i_1 j_2} a_{i_2 j_1})$$

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,

$$O^{[1^3]}_{11} F_{i_1 i_2 i_3} = \frac{1}{6}\left[F_{i_1 i_2 i_3} + F_{i_2 i_3 i_1} + F_{i_3 i_1 i_2} - F_{i_2 i_1 i_3} - F_{i_3 i_2 i_1} - F_{i_1 i_3 i_2}\right]$$

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^{α}_{pq} only for shapes with no more than n rows. In our above example $[1^{3}]$ has three rows; consequently, we obtained zero when operating on a tensor 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^r-dimensional vector space spanned by the rth-rank tensors $\mathfrak{F}$ is called the tensor space. A component of $\mathfrak{F}$ is denoted by $F_{i_1 i_2 \cdots i_r}$, 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^{α}_{rq} (which may be found in the reference RUTHERFORD (1948)). It will be convenient to discuss the semi-normal units e^{α}_{rq} rather than the orthogonal units O^{α}_{rq}. Since e^{α}_{rq} is proportional to O^{α}_{rq}, where the proportionality constant is not zero, then $O_{rq}F = 0$ iff $e_{rq}F = 0$. Let $\mathfrak{N}^{\alpha}_{r}$ be the group of all permutations for which no element of one column of S^{α}_{r} is put into a different column. Let $\mathcal{P}^{\alpha}_{r}$ be the group of all permutations for which no element of one row of S^{α}_{r} is put into a different row. Let

$$N^{\alpha}_{r} \triangleq \sum_{\tau \in \mathfrak{N}^{\alpha}_{r}} \zeta_{\tau}\, \tau \quad ; \quad P^{\alpha}_{r} \triangleq \sum_{\tau \in \mathcal{P}^{\alpha}_{r}} \tau$$

Define $E^{\alpha}_{rs} \triangleq P^{\alpha}_{r} \sigma^{\alpha}_{rs} N^{\alpha}_{s}$. Denote as S^{*} the tableau obtained from S by removing n, denote as $S^{**} = S^{2*}$ the tableau obtained from S^{*} by removing n-1, etc. Let $e^{(m-1)*} = \epsilon$

$$e^{(m-2)*}_{rs} = \frac{1}{\Theta^{(m-2)*}}\, e^{(m-1)*}_{rr}\, E^{(m-2)*}_{rs}\, e^{(m-1)*}_{ss}$$

$$e^{\gamma *}_{rs} = \frac{1}{\Theta^{\gamma *}}\, e^{(\gamma+1)*}_{rr}\, E^{\gamma *}_{rs}\, e^{(\gamma+1)*}_{ss}$$

$$e_{rs} = \frac{1}{\Theta}\, e^{*}_{rr}\, E_{rs}\, e^{*}_{ss}$$

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^{α}_{r} to give* $S^{\alpha}_{r}(F')$ no two identical indices among the first m+1 indices of F' occur in the same column of $S^{\alpha}_{r}(F')$, then $0^{\alpha}_{sr}F = 0$ for all s.

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

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

Proof:

(i) Let e^{j*}_{sr} operate on F. But, $E^{j*}_{sr} = P^{j*}_{s}\sigma_{sr}N^{j*}_{r}$ and N^{j*}_{r} antisymmetrizes each column of S^{j*}_{r} ; so, if E^{j*}_{sr} operates on F' and if two identical indices of the first n-j indices of F' are in the same column of S^{j*}_{r}, then N^{j*}_{r} F' = 0. Now, $e^{(\gamma+1)*}_{rr}F$ 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^{j*}_{r}, then $e^{(\gamma-1)*}_{sr}F = 0$ (for all s) and, thus, $e_{sr}F = 0$ (for all s).

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

(iii) In particular, an n-dimensional space F can have at most n different indices; so, $e_{rs}F = 0$ for all α 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

---

* It is convenient in discussing such theorems to place the indices of F in the tableau (say S^{α}_{p}) in place of the numbers 1...N; such an array I call an array and denote by $S^{\alpha}_{p}(F)$.

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 $0_{rs}F = 0$ 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 $0_{rs}F = 0$; 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 $0_{rp}F = \lceil$constant (independent of r) $0_{rq}F\rceil$ 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 σ_{pq} is an elementary transposition, say, (k,k+1), then $0^{\alpha}_{rp}F = 0^{\alpha}_{rp}(k,k+1)F = \sum_{\beta s t} U^{\beta}_{st(k,k+1)} O^{\alpha}_{rp} O^{\beta}_{st} F = \sum_{t} U^{\alpha}_{pt(k,k+1)} O^{\alpha}_{rt} F$

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, $0_{rp}F = U_{pp(k,k+1)}0_{rp}F + U_{pq(k,k+1)}0_{rq}F$. 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, $0_{rp}F$ = 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 $O_{rp}F$ 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 kth and (k+1)th indices of F are identical, then $O^{\alpha}_{rp}F = O^{\alpha}_{rp}(k,k+1)F = \sum_{\beta st} U^{\beta}_{st(k,k+1)} O^{\alpha}_{rp} O^{\beta}_{st} F = \sum_{t} U^{\alpha}_{pt(k,k+1)} O^{\alpha}_{rt} F$. But, $U^{\alpha}_{pt(k,k+1)} = -\delta_{pt}$, Thus, $O^{\alpha}_{rp}F = -O^{\alpha}_{rp}F$, and hence, $O^{\alpha}_{rp}F = 0$. If $S_p(F)$ has identical indices, in the same column, but these positions in S_p are not occupied by adjacent letters (e.g., $\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$ for $F = F_{6666}$) then among the set of identical arrays there is an $S_q(F)$ with adjacent letters in these positions*, viz., $\begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$ in the above example. Thus, using part (i) and first part of (ii) we obtain $O^{\alpha}_{rp}F = 0$.

An example for F_{11223} and $\begin{array}{|c|c|c|}\hline \ &\ &\ \\\hline \ &\ &\multicolumn{1}{c}{}\\\cline{1-2}\end{array}$

$S_1 = \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5&\multicolumn{1}{c}{}\\\cline{1-2}\end{array}$; $S_1(F) - \begin{matrix}1&1&2\\2&3&\end{matrix}$; $O^{\alpha}_{p1}F \neq 0$

$S_2 = \begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3&5&\multicolumn{1}{c}{}\\\cline{1-2}\end{array}$; $S_2(F) = \begin{matrix}1&1&2\\2&3&\end{matrix}$; $O^{\alpha}_{p2}F = \text{const}\ O^{\alpha}_{p1}F$; constant independent of p

$S_3 = \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&5&\multicolumn{1}{c}{}\\\cline{1-2}\end{array}$; $S_3(F) = \begin{matrix}1&2&2\\1&3&\end{matrix}$; $O^{\alpha}_{p3}F = 0$

$S_4 = \begin{array}{|c|c|c|}\hline 1&2&5\\\hline 3&4&\multicolumn{1}{c}{}\\\cline{1-2}\end{array}$; $S_4(F) = \begin{matrix}1&1&3\\2&2&\end{matrix}$; $O^{\alpha}_{p4}F \neq 0$

$O^{\alpha}_{p4}F \neq \text{const}\ O^{\alpha}_{p2}F$

$S_5 = \begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4&\multicolumn{1}{c}{}\\\cline{1-2}\end{array}$; $S_5(F) = \begin{matrix}1&2&3\\1&2&\end{matrix}$; $O^{\alpha}_{p5}F = 0$

---

* The reason being that all the arrays obtained from $S_p(F)$ 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 $[1^n]$ transforms as det a

Proof: $a(O_{11}F_{1..n}) = \sum_{i,j,..,k} a_{i1} a_{j2} \cdots a_{km} O_{11}^{[1^m]} F_{ij..k}$ But,

$O_{11}F_{ij...k} = 0$ unless i,j, etc., are all different; thus,

$\{i,j,...k\} = \{1,2,..n\}$.

Let $F_{ij..k} = \gamma F_{12..n}$;

then, $O_{11}^{[1^m]} \gamma F_{12..n} = \delta_\gamma O_{11}^{[1^m]} F_{12..n}$;

thus, $a(O_{11}^{[1^m]} F_{12..n}) = (\sum_\gamma a_{\gamma(1),1} a_{\gamma(2),2} \cdots a_{\gamma(m),m} \delta_\gamma) O_{11}^{[1^m]} F_{12..m}$

$= (\mathrm{DET}\ a)\ O_{11}^{[1^m]} F_{12..m}$ QED

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

Proof: $a(O_{rf}^{[2^m]} F_{1..n1..n}) = \sum_{i,j,..,k,..,m} a_{i1} a_{j2} \cdots a_{km} \cdots a_{mm} (O_{rf}^{[2^m]} F_{ij...k..m})$

But, $O_{rf}^{[2^m]} F = 0$ unless the first n indices are all different and the last n indices are all different.

Let $F_{ij..m} = \gamma' \gamma'' F_{1..n1..n}$, where γ' and γ'' operate on the first and last m components of F respectively. Then,

$O_{rf}^{[2^m]} F_{i..m} = \delta_{\gamma'} \delta_{\gamma''} O_{rf}^{[2^m]} F_{1..n1..n}$. But,

$\sum_{ij..k..m} a_{i1} a_{j2} \cdots a_{km} \cdots a_{mm} \delta_{\gamma'} \delta_{\gamma''} = (\det a)^2$; thus,

$a(O_{rf}^{[2^m]} F_{1..n1..n}) = (\det a)^2 (O_{rf}^{[2^m]} F_{1..n1..n})$. Let

$F_{1..n1..n} = \gamma F_{1122..nn}$, then

$O_{rf}^{[2^m]} F_{1..n1..n} = \sum_j U_{fj\gamma} O_{rj}^{[2^m]} F_{1122..nn} = U_{f1\gamma} O_{r1}^{[2^m]} F_{1122..nn}$

Hence, $a(O_{r1}^{[2^m]} F_{1122..nn}) = (\det a)^2 (O_{r1}^{[2^m]} F_{1122..nn})$.

Since $U_{f1\gamma} \neq 0$, any non-zero $O_{rp}^{[2^m]} F_{ij..k}$ can be expressed in terms of $O_{r1}^{[2^m]} F_{1122...nn}$ QED

Theorem Ab-5: The tensors of symmetry $[1^{m-1}]$ transform* as $(a^{t^{-1}} \det a)$

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

Proof: $a(O_{11}^{[1^{m-1}]} F_{i+1,i+2,\ldots,i-1}) = \sum_{jk\ldots m} a_{j,i+1} a_{k,i+2} \cdots a_{m,i-1} O_{11}^{[1^{m-1}]} F_{jk\ldots m}$

But, $O_{11}^{[1^{m-1}]} F_{jk..m} = 0$ unless jk..m are all different; therefore**

$O_{11}^{[1^{m-1}]} F_{jk..m} = \epsilon_{pjk..m} O_{11}^{[1^{m-1}]} F_{p+1,..,p-1}$ if p is the letter missing from jk..m. So,

$a(O_{11}^{[1^{m-1}]} F_{i+1,..,i-1}) = \left(\sum_{p} \sum_{jk\ldots m} \epsilon_{pjk\ldots m} a_{j,i+1} a_{k,i+2} \cdots a_{m,i-1}\right) O_{11}^{[1^{m-1}]} F_{p+1,\ldots,p-1}$

Let $A_{pi} \triangleq \sum_{jk\ldots m} \epsilon_{pjk\ldots m} a_{j,i+1} a_{k,i+2} \cdots a_{m,i-1}$ (this is the cofactor of det a).

Then, $a_{qi} A_{pi} = \epsilon_{pjk..m} a_{q,i} a_{j,i+1} \ldots a_{m,i-1} = \delta_{mp} \det a$ since if $q \neq p$, q is one of the other letters j, k,.., m and the determinant has two equal rows. Thus $A_{pk} = a^{-1}_{km} a_{mi} A_{ip} = a^{-1}_{kp} (\det a)$ QED

Note that for SO(n)*** the tensors $[1^{m-1}]$ and $[1]$ transform in the same way. For SU(n), $[1^{m-1}]$ transforms as a*; therefore, for SU(2) a and a* must be equivalent (this certainly provides insight into the reason for an important difference between SU(2) and SU(3), namely, that a is equivalent to a* for SU(2) but a is not equivalent to a* for SU(3).

---

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

** $\epsilon_{pj\ldots m}$ is zero unless all subscripts are different, it changes sign under transposition, and it is positive when all subscripts are in order

*** O(n) is the group of orthogonal transformations in n dimensions (i.e., if a O(n) then $a^t = a^{-1}$). The S in SO(n) indicates that we take the subgroup of O(n) with determinant equal to +1. U(n) is the group of unitary transformation in n dimensions (i.e., if a U(n) then $a^\dagger = a^{-1}$).

Theorem Ab-6: The tensor of symmetry $[1^{m-2}]$ transforms under a as det a times the way $[1^2]$ transforms under $a^{t^{-1}}$. (Similarly we can prove the corresponding theorem for $[1^{m-3}]$ and $[1^3]$, etc.).

Proof: Proceeding as in Thm Ab-5 we obtain

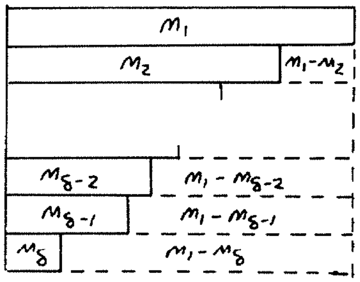
$$a(0_{11}^{[1^{m-2}]} F_{i+1,..,j-1,j+1,..i-1}) = \sum_{p,q} A_{pqij} 0_{11}^{[1^{m-2}]} F_{p+1,..,q-1,q+1,..,p-1}$$

where $A_{pqij} = \sum_{k,\mu,..,m,\nu,..} \epsilon_{pk\mu..mq\nu..} \partial_{k,\lambda+1} \partial_{\mu,\lambda+2} \cdots$

Then, $a_{si} a_{tj} A_{pqij} = (\det a)(\delta_{sp} \delta_{tq} - \delta_{sq} \delta_{tp})$

Thus, $A_{pquv} = a^{-1}_{\mu s} a^{-1}_{\nu t} a_{s\lambda} a_{tj} A_{pqij} = (\det a)(a^{-1}_{\mu p} a^{-1}_{\nu q} - a^{-1}_{\mu q} a^{-1}_{\nu p})$ QED

Theorem Ab-7:* $0_{rf}^{[m_1, m_2, .., m_\delta]} F$ transforms as $(\det a)^{m_\delta}$ times the way $0_{rf}^{[m_1-m_\delta, m_2-m_\delta, .., m_{\delta-1}-m_\delta]} F$ transforms. $0_{rf}^{[m_1-m_\delta, m_2-m_\delta, \ldots, m_{\delta-1}-m_\delta]}$ F transforms under a the same way that $0_{rf}^{[m_1-m_\delta, m_1-m_{\delta-1}, \ldots, m_1-m_2]}$ F transforms under $a^{t^{-1}}$.



Proof:

$$a(0_{rf}^{[m_1, m_2, \ldots, m_\delta]} F_{1..\delta 1..\delta \ldots \nu}) = \sum_{ij..k} \partial_{i1} \partial_{j2} \cdots \partial_{k\nu} O_{rf}^{[m_1, m_2, .., m_\delta]} F_{ij..k}$$

But, the first m_δ sets of δ indices in $F_{ij..k}$ must each contain each letters $1,\ldots,\delta$ once; the next $m_{\delta-1}-m_\delta$ sets must each contain all but one of the letters $1,\ldots,\delta$ once; etc.. Next, we let $F_{ij..k} = \tau F_{1..\delta 1..\delta..}$, where the first m_δ sets of δ indices of F' are in order $1,\ldots,\delta$; the next $m_{\delta-1}-m_\delta$ sets of δ -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 $m_{\delta-2}-m_{\delta-1}$ sets of δ -2

---

* Here I let the dimension of the transformation space be δ .

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 $p+1, p+2, .., q-1, q+1, ..., \delta, 1, ... p-1$), etc. Then, $0_{rf}F = 0_{rf}\gamma F' = \zeta^{\rho}_{\tau_1^{\delta}}\zeta^{\rho}_{\tau_2^{\delta}}\cdots\zeta^{\rho}_{\tau^{\delta-1}_{(m_{\delta-1}-m_\delta)}}\cdots\zeta^{\rho}_{\tau^{1}_{m_1-m_2}}\,O_{r\gamma}F'$

Hence, each of the first m_δ terms separates into a $\sum_{\lambda\gamma\cdots\beta} a_{\lambda 1}a_{\gamma 2}\cdots a_{\beta\delta}\,\zeta^{\rho}_{\gamma^{\delta}} =$ det a just as in THM Ab-3. Thus, $0_{rf}^{[m_1, m_2, .., m_\delta]}F$ transforms the same way as $(\det a)^{m_\delta}$ times $0_{rf}^{[m_1-m_\delta, m_2-m_\delta, \cdots, m_{\delta-1}-m_\delta]}F$. Now, consider $0_{rf}^{[m_1-m_\delta, m_2-m_\delta, \cdots, m_{\delta-1}-m_\delta]}F$, each column transforms separately and the column of length $[1^{\delta-1}]$ transforms like $[1]$ if a is replaced by $a^{t^{-1}}$, and the column of length $[1^{\delta-2}]$ transforms like $[1^2]$ if a is replaced by $a^{t^{-1}}$, etc., as in THM Ab-6. Thus, $0_{rf}^{[m_1-m_\delta, m_2-m_\delta, \cdots, m_{\delta-1}-m_\delta]}F$ transforms like $0_{rf}^{[m_1-m_\delta, m_1-m_{\delta-1}, \cdots, m_1-m_2]}F$ if a is replaced by $a^{t^{-1}}$.

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

Proof: Trivial.

It may be true that THM Ab-7 applies to any non-zero 0^{α}_{rj} even if there is more than one linearly independent 0^{α}_{rj} 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 transformations G of the n-dimensional transformation space. If the group is GL(n), the group of all linear transformation of an n-dimensional space, then the invariant subspaces of the n^r-dimensional tensor space obtained by applying the $0^{\alpha}_{\rho\gamma}$ are indecomposable for fixed α and ρ with q varying.

For some subgroups of GL(n), e.g., SU(n), (see HAMERMESH (1962), p. 390) the invariant subspaces are still indecomposable, whereas for others, e.g., SO(n), the invariant subspaces are in general decomposable. For example, if G = 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.

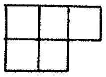
We are particularly interested in (complex) two-dimensional spaces with SU(2) as the transformation group. We choose as the basis vectors α and β, the eigenfunctions of S_z. A typical tensor component (for an N-electron problem) might be $\chi(1\ldots N) = \alpha(1)\beta(2)\alpha(3)\alpha(4)\beta(5)\cdots\alpha(N)$ From THM Ab-1 iii we can consider only one and two rowed shapes since $O^{\alpha}_{rf}\chi = 0$ if α has more than two rows. From THM Ab-2 we see that the $O^{\alpha}_{ri}\chi$ for different i and the same α and r are all proportional. So, although for general χ we might have f^{α} linearly independent sets of $O^{\alpha}_{ri}\chi$ (with r varying from 1 to f^{α}), there is at most one. Thus, we choose the convention of considering only the components χ with all of the α written first and all of the β written last; we write $O^{\alpha}_{ri}\chi$ picking i = 1 since $O^{\alpha}_{r1}\chi$ is always non-zero if $O_{ri}\chi$ is non-zero for any i. The shape has two rows; thus, we write the partition α 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 α for each integral value of S from 0 to $\frac{N}{2}$; if N is odd, there is one partition α 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 $O^{\alpha}_{r1}\chi$ for the different χ (r and α fixed) is 2S+1 (use THM Ab-2). Because of our convention on χ these different $O^{\alpha}_{r1}\chi$ correspond to different S_z; and of the 2S+1 different $O^{\alpha}_{r1}\chi$ there is one for each S_z from -S to +S spaced at integral intervals. Also, $S^{+}O_{r1}\chi_S = 0$ and $S^{-}O_{r1}\chi_{-S} = 0$ (use THM Ab-2). The total number of linearly independent $O^{\alpha}_{r1}\chi$ for different χ, α, and r is 2^N; and f^{α} is just the number obtained from the branching diagram for the number of different states of the same spin. Since the spaces of different α and r are each indecomposable, we have shown that the set of states $O^{\alpha}_{r1}\chi$ for $\alpha = [\frac{N}{2}+S, \frac{N}{2}-S]$ 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 α, we use them interchangeably; when I write $O^{\alpha}_{ri}\chi_S$, I mean that χ has $S_z = S$ for the S corresponding to α. There is only one $O_{rj}\chi_S \neq 0$ for varying j and fixed r and $O_{r1}\chi_S$ is an eigenfunction of S^2 with $S_z = S$ (since $S^{+}O_{r1}\chi_S = 0$). $S^{-}O_{r1}\chi_S$ gives a linear combination of $O_{rj}\chi_{S-1}$ for varying j and fixed r (i.e., $S^{-}\chi_S = \chi_{S-1} + \chi'_{S-1} + \cdots + \chi''_{S-1}$ where only χ_{S-1} has the standard order; thus, $S^{-}\chi_S = (\epsilon + \tau' + \cdots + \tau'')\chi_{S-1}$ where, e.g., $\tau''\chi_{S-1} = \chi''_{S-1}$ and, e.g., $O^{\alpha}_{r1}\tau'\chi_{S-1} = \sum_j U_{1j\tau'}O^{\alpha}_{rj}\chi_{S-1})$ But, all of these are proportional; so, by merely multiplying by the appropriate (normalization) constant the $O^{\alpha}_{r1}\chi_{S_z}$ are all connected by $S^{\pm}$. Each $O^{\alpha}_{r1}\chi_{S_z}$ is an eigenfunction of S^2.

Some examples:

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

$S_1 = \begin{matrix}\alpha & \alpha & \beta \\ \beta & \beta & \end{matrix}$

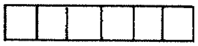
$S_2 = \begin{matrix}\alpha & \alpha & \beta \\ \beta & \beta & \end{matrix}$ $\qquad O_{p1}F = \text{const } O_{p2}F = \text{const } O_{p4}F$

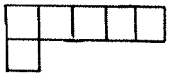
$S_3 = \begin{matrix}\alpha & \beta & \beta \\ \alpha & \beta & \end{matrix}$ $\qquad O_{p3}F = 0$

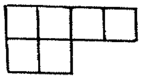
$S_4 = \begin{matrix}\alpha & \alpha & \beta \\ \beta & \beta & \end{matrix}$ $\qquad O_{p5}F = 0$

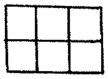
$S_5 = \begin{matrix}\alpha & \beta & \beta \\ \alpha & \beta & \end{matrix}$

For N = 6

S = 3 

S = 2 

S = 1 

S = 0 

For N = 2

S = 1  $S_z = 1 : O_{11}^{[2]}\alpha\alpha = \frac{1}{2}(\alpha\alpha + \alpha\alpha) = \alpha\alpha$

$S_z = 0 : O_{11}^{[2]}\alpha\beta = \frac{1}{2}(\alpha\beta + \beta\alpha)$

$O_{11}^{[2]}\beta\alpha = \frac{1}{2}(\beta\alpha + \alpha\beta) = O_{11}^{[2]}\alpha\beta$

$S_z = -1 : O_{11}^{[2]}\beta\beta = \beta\beta$

S = 0  $S_z = 1 : O_{11}^{[1^2]}\alpha\alpha = 0$

$S_z = 0 : O_{11}^{[1^2]}\alpha\beta = \frac{1}{2}(\alpha\beta - \beta\alpha) = -O_{11}^{[1^2]}\beta\alpha$

$S_z = -1 : O_{11}^{[1^2]}\beta\beta = 0$

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 $O_{11}^{[2]}\alpha\alpha$, $O_{11}^{[2]}\alpha\beta$, $O_{11}^{[2]}\beta\beta$, and $O_{11}^{[1^2]}\alpha\beta$ 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 $\Psi = \psi_i(\tilde{r}_1)\psi_j(\tilde{r}_2)\cdots\psi_k(\tilde{r}_N)$, and let each $\psi_i(\tilde{r}_j) = \phi_i(r_j)\sigma_i(j)$, where $S_i = \alpha$ or β. Then $\Psi(1,\cdot\cdot,N) = \phi_i(r_1)\phi_j(r_2)\cdot\cdot\phi_k(r_N)\sigma_i(1)\sigma_j(2)\cdot\cdot\sigma_k(N)$. Let

$\Phi(1,\cdot\cdot,N; i,j,\cdot\cdot,k) \hat{=} \phi_i(r_1)\phi_j(r_2)\cdot\cdot\phi_k(r_N)$; $\chi(1,\cdot\cdot,N; i,j,\cdot\cdot,k) \hat{=} \sigma_i(1)\sigma_j(2)\cdot\cdot\sigma_k(N)$

Then Φ involves only spatial coordinates, and χ involves only spin coordinates. Let the operator O^{α}_{pq} operate only on the spatial coordinates and ω_{pq} operate only on the spin coordinates. Using Aa-10*

$$\boxed{G_i^{\alpha}\Phi\chi = \sum_r \zeta_{\sigma_{ri}} (O_{ri}^{\alpha}\Phi)(\omega_{ri}^{\bar{\alpha}}\chi)} \qquad \text{(Ac-1)}$$

But, from section Ab $\omega_{ri}^{\bar{\alpha}}\chi$ is an eigenfunction of S^2. Therefore, since S^2 does not operate on $(O^{\alpha}_{ri}\Phi)$, then $G^{\alpha}_i\Phi\chi$ is an eigenfunction of S^2. From Aa-13 $\tau G_i^{\alpha} = \zeta_{\tau} G_i^{\alpha}$; thus,

$$\boxed{\tau(G_i^{\alpha}\Phi\chi) = \zeta_{\tau}(G_i^{\alpha}\Phi\chi)} \qquad \text{(Ac-2)}$$

<u>Therefore, $G^{\alpha}_i\Phi\chi$ satisfies Pauli's principle and is an eigenfunction of S^2!</u> Note that since $\bar{\alpha}$ is limited to two rows then α is limited to two columns.

A useful theorem follows.

Theorem Ac-1:

(i) $\langle O^{\alpha}_{ij}\Phi_1 | \Phi_2\rangle = \langle \Phi_1 | O^{\alpha}_{ji}\Phi_2\rangle$

(ii) $\langle O^{\alpha}_{ij}\Phi_1 | O^{\beta}_{pq}\Phi_2\rangle = \delta^{\alpha\beta}\delta_{ip}\langle \Phi_1 | O^{\alpha}_{jq}\Phi_2\rangle$

---

* For Bose particles, $\tau\psi=\psi$; thus, we would use $\sum_r (O^{\alpha}_{ri}\Phi)(\omega^{\alpha}_{ri}\chi)$ 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) = $dx_1 dx_2 \ldots dx_N$;

$$\langle O^{\alpha}_{ij}\Phi_1 | \Phi_2\rangle = \int (O^{\alpha}_{ij}\Phi_1)^* \Phi_2 (dx) = \frac{1}{\Theta^{\alpha}} \sum_{\tau} U^{\alpha}_{ij\tau} \int (\tau\Phi_1)^* \Phi_2 (dx)$$

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.

$$\int (\tau\Phi_1)^* \Phi_2 (dx) = \tau'' \int (\tau\Phi_1)^* \Phi_2 (dx) = \int (\tau''\tau\Phi_1)^* (\tau''\Phi_2)(dx)$$

If $\tau'' = \tau^{-1}$ we obtain $\int (\tau\Phi_1)^* \Phi_2 (dx) = \int \Phi_1^* (\tau^{-1}\Phi_2)(dx)$;

so, using $U_{ij\tau} = U_{ji\tau^{-1}}$

$$\langle O^{\alpha}_{ij}\Phi_1 | \Phi_2\rangle = \frac{1}{\Theta^{\alpha}} \sum_{\tau} U^{\alpha}_{ji\tau^{-1}} \int \Phi_1 \tau^{-1}\Phi_2 (dx) = \langle \Phi_1 | O^{\alpha}_{ji}\Phi_2\rangle$$

(ii) Using part i

$$\langle O^{\alpha}_{ij}\Phi_1 | O^{\beta}_{pq}\Phi_2\rangle = \langle \Phi_1 | O^{\alpha}_{ji} O^{\beta}_{pq} | \Phi_2\rangle = \delta^{\alpha\beta}\delta_{ip} \langle \Phi_1 | O^{\alpha}_{jq}\Phi_2\rangle \qquad QED$$

Using Thm. Ac-1

$$\langle G^{\alpha}_{i}\psi_1 | G^{\beta}_{j}\psi_2\rangle = \delta^{\alpha\beta}\rho_{\sigma_{ij}} f^{\alpha} \langle \Phi_1 | O^{\alpha}_{ij}\Phi_2\rangle \langle \chi_1 | \omega^{\bar{\alpha}}_{\bar{i}\bar{j}}\chi_2\rangle$$

We cannot go too much further without becoming becoming a little more specific about the Φ_i .

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)

$$\Phi^{(m)}_{S,M} = \frac{1}{\sqrt{f_S}} \sum_{k} \psi^{(m)}_{S,k}\, \Theta_{S,M;k} \qquad \text{, where} \quad \psi^{(m)}_{S,k} = \sqrt{\frac{f}{N!}} \sum_{P} \bar{U}_{km}(P)\, P\psi^{0}$$

(ψ^{0} is a product of N orbitals) and $\Theta_{S,M;k}$ is one of f_S orthonormal spin functions which is an eigenfunction of S^2 with eigenvalue S(S+1) (the Θ are obtained by vector coupling the N spins in the f_S different ways given by the branching diagram). The Θ transform as (K p.7)

$P\Theta_{S,M;k} = \sum_h V_{hk}(P)\Theta_{S,M;h}$ and $U_{hk}(P) = \epsilon_P V_{kh}(P^{-1})$ in order that $\Phi_{S,M}$ satisfy Pauli's principle (K p.14). We must compare the $\Phi_{S,M}^{(m)}$ to $G_i^\alpha \Phi\chi = \sum_r \zeta_{\sigma_{ri}} (O_{ri}^\alpha \Phi)(\omega_{\bar{r}\bar{i}}^{\bar{\alpha}}\chi)$. Here $P\omega_{\bar{r}\bar{i}}^{\bar{\alpha}}\chi = \sum_{\beta\gamma s} U_{\gamma s P}^{\beta}\, \omega_{\gamma s}^{\beta}\omega_{\bar{r}\bar{i}}^{\bar{\alpha}}\chi = \sum_\gamma U_{\bar{\gamma}\bar{r}P}^{\bar{\alpha}}\, \omega_{\bar{\gamma}\bar{i}}\chi$. We may identify $\omega_{\bar{r}\bar{i}}^{\bar{\alpha}}\chi$ and $\Theta_{S,M;r}$ in which case $U_{\bar{\gamma}\bar{r}P}^{\bar{\alpha}} = V_{\gamma r}(P)$. Also (K p. 14) $P\Psi_{S,k}^{(m)} = \sum_h U_{hk}(P)\Psi_{S,h}^{(m)}$. However, $P(\zeta_{\sigma_{ri}} O_{ri}^\alpha \Phi) = \sum_\gamma U_{\gamma r}\, \zeta_{\sigma_{ri}} O_{\gamma i}\Phi = \sum_\gamma \zeta_{\sigma_{\gamma r}} U_{\gamma r P}(\zeta_{\sigma_{\gamma i}} O_{\gamma i}\Phi)$; thus, $U_{\gamma r}(P) = \zeta_{\sigma_{\gamma r}} U_{\gamma r P}$. But from appendix I $\zeta_{\sigma_{\gamma r}} U_{\gamma r P} = \zeta_P U_{\bar{\gamma}\bar{r}P}$; thus, $U_{\gamma r}(P) = \zeta_P U_{\bar{\gamma}\bar{r}P}$. Thus, we see that the Kotani function and the $G_i^\alpha \Phi\chi$ are equivalent. In fact Kotani derives the equivalent of Theorems Ba-1 and Ba-2 and appendix A (K p.22 and K p.23-26); he also obtains the equivalent of equations Bc-3 and 4 (K p.27). In general Kotani's proofs proceed somewhat differently than those discussed in this report. An advantage of the G_i^α form is the close relationship to the Young tableaux permitting a pictorial visualization of the relationship between and the properties of the G_i^α ; however, the $\Theta_{S,M;k}$ 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 LÖWDIN (1955c). Löwdin 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 $\prod_{k \neq \ell} \frac{\hat{S}^2 - k(k+1)}{\ell(\ell+1) - k(k+1)}$ where the product is over all k from 0 or $\frac{1}{2}$ to $\frac{N}{2}$ (except ℓ) and $\hat{S}^2 = (\sum_{\lambda=1}^{N} \hat{s}_\lambda^+)(\sum_{\gamma=1}^{N} \hat{s}_\gamma^-) + (\sum \hat{s}_z)^2 - (\sum \hat{s}_z)$. The problem is that the result of operating with ${}^{2\ell+1}O$ 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 ${}^{2\ell+1}O$ 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 DODS approach of solving for the ground state energy and wave function of a molecule, we will develop the Hartree-Fock method using the G_i^α 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 G_i^α 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 $G_i^\alpha \Phi \chi$ 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 $G_i^\alpha \Phi \chi$ 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 $G_i^\alpha \Phi \chi$

First some useful theorems.

Theorem Ba-1: If $F_{i_1 i_2} = \Phi = \phi_{i_1}(1)\phi_{i_2}(2)\cdots$ (where the set of ϕ_i is orthonormal) is written with all identical indices adjacent and if $S_p(\Phi) \neq S_q(\Phi)$, then $O_{rp}^\alpha \Phi$ and $O_{rq}^\alpha \Phi$ are orthogonal.

Proof: Using THM Ac-1 $\langle O_{rp}^\alpha \Phi | O_{rq} \Phi \rangle = \langle \Phi | O_{pq}^\alpha | \Phi \rangle = \frac{1}{\theta^\alpha} \sum_\tau U_{pq\tau}^\alpha \langle \Phi | \tau | \Phi \rangle$, but $\langle \Phi | \tau \Phi \rangle \neq 0$ implies that τ only permutes sets of identical indices, and if for convenience we consider that only one set is permuted, then τ can be expressed in terms of elementary transpositions involving only the letters corresponding to the pertinent set of indices of Φ. Hence, $\tau S_q(\Phi) = S_q(\Phi)$; thus, $U_{pq\tau} \neq 0$ 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 Ab-2 if $S_p(\Phi) = S_q(\Phi)$ then $O_{rp}^\alpha \Phi$ and $O_{rq}^\alpha \Phi$ 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 $G_i^\alpha \Phi \chi$; so, a method of deciding which ones are proportional is necessary.

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

$$\Phi = \phi_{i_1}(1)\phi_{i_1}(2)\phi_{i_2}(3)\phi_{i_2}(4)\cdots\phi_{i_M}(2M-1)\phi_{i_M}(2M)\phi_{i_{2M+1}}(2M+1)\phi_{i_{2M+2}}(2M+2)\cdots\phi_{i_N}(N)$$

where the set $i_1, i_2, \ldots, i_n$ is ordered and the set $i_{2n+1} \cdots$ is ordered, and n is the number of doubly-occupied orbitals) the set of non-zero $G_i^\alpha \Phi \chi$ is linearly independent and, a fortiori, orthogonal. (This theorem depends on the α being one- or two-columned.)

Proof: Let T_Φ be the group of permutations leaving Φ invariant (it is an Abelian group of 2^n elements); then,

$$\begin{aligned} \langle \Phi | \tau \Phi \rangle &= 0 \text{ if } \tau \notin T_\Phi \\ &= 1 \text{ if } \tau \in T_\Phi \end{aligned}$$

$U_{pq\tau} = \delta_{pq}$ if $\tau \in T_\Phi$ since both S_p and S_q are of the form

1	2
3	4
⋮	⋮
2n-1	2n
⋮	⋮

We have

$$\langle O^{\alpha}_{rp} \Phi | O_{rq} \Phi \rangle = \langle \Phi | O^{\alpha}_{pq} \Phi \rangle = \frac{2^{M}}{\theta^{\alpha}} \delta_{pq} \qquad \text{(Ba-1)}$$

But, using THM Ac-1

$$\begin{aligned} \langle G^{\alpha}_{p} \Phi_1 \chi_1 | G^{\beta}_{q} \Phi_2 \chi_2 \rangle &= \sum_{s} \sum_{r} \rho_{\sigma_{sp}} \rho_{\sigma_{rq}} \langle \Phi_1 | O^{\alpha}_{pq} \Phi_2 \rangle \langle \chi_1 | \omega^{\bar{\alpha}}_{\bar{p}\bar{q}} \chi_2 \rangle \, \delta^{\alpha\beta} \delta_{rs} \\ &= \delta^{\alpha\beta} \rho_{\sigma_{pq}} \, \rho^{\alpha} \langle \Phi_1 | O^{\alpha}_{pq} \Phi_2 \rangle \langle \chi_1 | \omega^{\bar{\alpha}}_{\bar{p}\bar{q}} \chi_2 \rangle \end{aligned}$$

and therefore, if Φ_λ and Φ_γ are in the standard order,

$$\langle G^{\alpha}_{p} \Phi_\lambda \chi_\lambda | G^{\beta}_{q} \Phi_\gamma \chi_\gamma \rangle = \delta^{\alpha\beta} \delta_{pq} \, \delta_{\lambda\gamma} \frac{2^{M} \rho^{\alpha}}{\theta^{\alpha}} \langle \chi_\lambda | \omega^{\bar{\alpha}}_{\bar{p}\bar{p}} \chi_\gamma \rangle \qquad \text{(Ba-2)}$$

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 O^{α}_{ij} 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 $\Psi = \Phi\chi$ that we do not miss any non-zero $G^{\alpha}_{i} \Psi$. In particular we cannot arbitrarily select the convention for Φ and χ separately (e.g., if $\alpha = [2^3]$, $\Phi = \phi_1 \phi_1 \phi_2 \phi_2 \phi_3 \phi_3$, and $\chi = \alpha\alpha\alpha\beta\beta\beta$ then $O^{\alpha}_{pq} \Phi \neq 0$ only for $q = 1$ and $\omega^{\bar{\alpha}}_{\bar{p}\bar{1}} \chi = 0$ whereas if $\chi = \alpha\beta\alpha\beta\alpha\beta$ then $\omega^{\bar{\alpha}}_{\bar{p}\bar{1}} \chi \neq 0$). For SODS it is most convenient to use the convention of THM Ab-2 for Φ and then to use a convention for χ which will give a non-zero $\omega^{\bar{\alpha}}_{\bar{p}\bar{q}} \chi$ for every q which gives a non-zero $O^{\alpha}_{pq} \Phi$ regardless of what Φ is. The convention for χ is that the first N-2S positions of χ have alternating α and β then the remaining 2S positions have the remaining

α first and then the remaining β, e.g., for S = 2, $S_z = 1$, and N = 10, $\chi = \alpha\beta\alpha\beta\alpha\beta\alpha\alpha\alpha\beta$ (so $S_q^{\bar\alpha}(\chi) = \begin{smallmatrix}\alpha\alpha\alpha\,\alpha\alpha\alpha\,\beta\\ \beta\beta\beta\end{smallmatrix}$). From theorems Ab-1 and Ab-2 we see that there is indeed no case where $\omega_{pq}\chi = 0$ from either of these theorems; for a two-dimensional transformation space theorems Ab-1 and Ab-2 probably form a necessary condition for $\omega_{pq}\chi = 0$. There are some cases in which we will use the convention of THM Ab-2 for χ; in those cases we will write* $\chi_{\bar f}$ rather than just χ.

Consider the case of $M_S = S$; then (by THM Ab-2) $\omega_{\bar p\bar q}\chi_{\bar f} = \delta_{\bar q\bar f}\,\omega_{\bar p\bar f}\chi_{\bar f}$. Let $\chi = \sigma\chi_{\bar f}$ then (using Aa-4)

$$\omega_{\bar q\bar p}\,\chi = U_{\bar p\bar f\sigma}\,\omega_{\bar q\bar f}\,\chi_{\bar f} \tag{Ba-3}$$

$$\langle\chi|\omega_{\bar p\bar p}|\chi\rangle = (U_{\bar p\bar f\sigma})^2\langle\chi_{\bar f}|\omega_{\bar f\bar f}|\chi_{\bar f}\rangle \tag{Ba-4}$$

Thus,

$$\langle\chi|\omega_{\bar p\bar f}|\chi\rangle = U_{\bar p\bar f\sigma}\,U_{\bar f\bar f\sigma}\langle\chi_{\bar f}|\omega_{\bar f\bar f}|\chi_{\bar f}\rangle = \sqrt{\langle\chi|\omega_{\bar p\bar p}|\chi\rangle\langle\chi|\omega_{\bar f\bar f}|\chi\rangle} \tag{Ba-5}$$

But, $\langle\chi_{\bar f}|\omega_{\bar f\bar f}|\chi_{\bar f}\rangle = \frac{1}{\Theta^\alpha}\sum_\tau U_{\bar f\bar f\tau}\langle\chi_{\bar f}|\tau|\chi_{\bar f}\rangle$ and $\langle\chi_{\bar f}|\tau\,\chi_{\bar f}\rangle \neq 0$ only for $\tau \in T_{\chi_{\bar f}}$ (i.e., τ which do not move elements from one row to a different row). If $\tau \in T_{\chi_{\bar f}}$ then $U_{\bar f\bar f\tau} = 1$; thus,

$$\langle\chi_{\bar f}|\,\omega_{\bar f\bar f}|\chi_{\bar f}\rangle = (\tfrac{N}{2} + S)!\;(\tfrac{N}{2} - S)!/\,\Theta^\alpha$$

The $G_i^\alpha\Phi\chi$ used so far have not been normalized, denote the normalized function as $\bar G_i^\alpha\Phi\chi$. 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_f is the standard tableau obtained by filling successive columns one at a time

$$S_1 = \begin{array}{|c|c|}\hline 1 & 2\\ \hline 3 & 4\\ \hline 5 & 6\\ \hline 7 & \\ \cline{1-1}\end{array} \qquad S_f = \begin{array}{|c|c|}\hline 1 & 5\\ \hline 2 & 6\\ \hline 3 & 7\\ \hline 4 & \\ \cline{1-1}\end{array}$$

but it is shown below that Ba-7 and Ba-8 are valid for any M_S)

$$\bar{G}_i^\alpha \Phi \chi = C_n^\alpha(M_S) \sum_r \int_{\sigma_{r1}}^{\rho} (O_{ri}^\alpha \Phi)(\omega_{\bar{r}\bar{f}}^{\bar{\alpha}} \chi_{\bar{f}}) \qquad \text{(Ba-7)}$$

where

$$C_n^\alpha(M_s) = \sqrt{\frac{\Theta^\alpha}{2^n f^\alpha \langle \chi_{\bar{f}} | \omega_{\bar{f}\bar{f}} | \chi_{\bar{f}} \rangle}} \qquad \text{(Ba-8)}$$

For any M_s we let $G_i^\alpha \Phi\chi = C_i^\alpha(\Phi\chi) \sum_r \int_{\sigma_{r1}}^{\rho} (O_{ri}^\alpha \Phi)(\omega_{\bar{r}\bar{i}}^{\bar{\alpha}} \chi)$

where $C_i^\alpha(\Phi\chi) = \sqrt{\dfrac{\Theta^\alpha}{2^n f^\alpha \langle \chi_1 | \omega_{\bar{p}\bar{p}}^{\bar{\alpha}} | \chi_1 \rangle}}$

Thus,

$$\langle \bar{G}_i^\alpha \Phi\chi | \bar{G}_j^\beta \Phi\chi \rangle = \delta^{\alpha\beta} \delta_{ij} \qquad \text{(Ba-9)}$$

For arbitrary M_s the non-zero $\omega_{\bar{r}\bar{i}}\chi$ 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 $\omega_{\bar{r}\bar{i}}\chi \neq 0$, let

$$\omega_{\bar{r}\bar{i}} \chi = \eta_{\bar{i}\bar{f}}\, \omega_{\bar{r}\bar{f}} \chi_{\bar{f}} \qquad \text{(Ba-10)}$$

Then equations Ba-7, and 8 are valid for <u>any</u> M_s and

$$\langle \chi_1 | \omega_{\bar{i}\bar{j}} | \chi_2 \rangle = \eta^1_{\bar{i}\bar{f}}\, \eta^2_{\bar{j}\bar{f}} \langle \chi^1_{\bar{f}} | \omega_{\bar{f}\bar{f}} | \chi^2_{\bar{f}} \rangle = \sqrt{\langle \chi_1 | \omega_{\bar{i}\bar{i}} | \chi_1 \rangle \langle \chi_2 | \omega_{\bar{j}\bar{j}} | \chi_2 \rangle} \qquad \text{(Ba-11)}$$

I define

$$G_{\underline{i}}^\alpha \Phi\chi \hat{=} \sum_r \int_{\sigma_{r1}}^{\rho} (O_{r1}^\alpha \Phi)(\omega_{\bar{r}\bar{f}}^{\bar{\alpha}} \chi_{\bar{f}}) \qquad \text{(Ba-12)}$$

We expand the exact wave function for the system in terms of the $G_i^\alpha \Phi\chi$ (the exact wave function is required to be an eigenfunction of S^2 and S_z)

$$\Psi_M^\alpha(\tilde{r}_1, \tilde{r}_2, \cdots, \tilde{r}_N) = \sum_i \sum_\Phi C_i^\alpha(\Phi)\, \bar{G}_i^\alpha \Phi\chi \qquad \text{(Ba-13)}$$

where the sum over $\bar{\Phi}$ is over all different ordered Φ

$$\Phi = \phi_{\lambda_1}(1)\phi_{\lambda_1}(2)\cdots\phi_{\lambda_M}(2M-1)\phi_{\lambda_M}(2M)\phi_{\lambda_{2M+1}}(2M+1)\cdots\phi_{\lambda_N}(N)$$

Thus, $\sum\limits_{\Phi} = \sum\limits_{M=0}^{N/2}\ \sum\limits_{p_1 p_2 \cdots p_M}\ \sum\limits_{g_1 g_2 \cdots g_{N-2M}}$

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 χ_M is such that it also depends on n. Thus,

(Ba-14)

$$\Psi_M^{\alpha}(\tilde{r}_1,\ldots,\tilde{r}_N) = \sum_{\lambda}\sum_{M}\sum_{p_1\cdots p_M}\ \sum_{g_1\cdots g_{N-2M}} C_{\lambda}^{\alpha}(M;p_1,\cdots,p_M;g_1,\cdots,g_{N-2M})\,\bar{G}_{\lambda}^{\alpha}(\phi_{p_1}(1)\phi_{p_1}(2)\cdots\phi_{p_M}(2M)\,\phi_{g_1}(2M-1)\cdots\phi_{g_{N-2M}}(N)\,\chi_M$$

By Ba-4 the terms in this sum are all orthogonal.

Bb*: First- and Second-Order Density Matrices for $G_i^{\alpha}\Phi\chi$

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 $H^1 = \sum_i h_i$, where h_i operates on electron i (H^1 is invariant under permutations of the particles). If Ψ is an N-electron wave function which satisfies Pauli's principle, then $\langle\Psi|H^1|\Psi\rangle = \langle\Psi|\sum_i h_i|\Psi\rangle = \sum_i\langle\Psi|h_i|\Psi\rangle$, but since all coordinates in $\langle\Psi|h_i|\Psi\rangle$ are integrated over, then $\langle\Psi|h_i|\Psi\rangle = \tau\langle\Psi|h_i|\Psi\rangle = \langle\tau\Psi|h_{\tau(i)}|\tau\Psi\rangle$. But $\tau\Psi = \delta_\tau\Psi$, so, let $\tau = (1,i)$, then $\langle\Psi|h_i|\Psi\rangle = \langle\Psi|h_1|\Psi\rangle$. Thus, $\langle\Psi|H^1|\Psi\rangle = \sum_i\langle\Psi|h_1|\Psi\rangle = N\langle\Psi|h_1|\Psi\rangle$. Let

$$\rho(1,1') \triangleq N\int\Psi^*(1,2,\cdots,N)\,\Psi(1',2,\cdots,N)\,(dx_1')(ds) \tag{Bb-1}$$

where $(dx_1') = dx_2dx_3\ldots dx_N$ and, $(ds) = ds_1ds_2\ldots ds_N$(i.e., integrate (actually sum) over all spins) then,

$$\langle\Psi|H^1|\Psi\rangle = \int\{h_1\cdot\rho(1,1')\}_{1'=1}\,dx_1 \tag{Bb-2}$$

where the $\{h_1\cdot\rho(1,1')\}_{1'=1}$ ensures that the operator h_i operates on Ψ and not on $\Psi^*\Psi$ (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 $H^2 = \sum_{p<q} g_{pq}$, then, as

---

* See LÖWDIN (1955) for a general discussion of the density matrices. The notation is based on that of RUEDENBERG (1962).

above,

$$\langle \Psi | H^2 | \Psi \rangle = \langle \Psi | \sum_{p<q} g_{pq} | \Psi \rangle = \frac{1}{2} \sum_{p,q}{}' \langle \Psi | g_{pq} | \Psi \rangle = \frac{1}{2} \sum_{p,q}{}' \langle \Psi | g_{12} | \Psi \rangle$$
$$= \frac{N(N-1)}{2} \langle \Psi | g_{12} | \Psi \rangle \qquad \text{(Bb-3)}$$

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,

$$\langle \Psi | H^2 | \Psi \rangle = \frac{1}{2} \int g_{12}\, \pi(1,2)\, dx_1\, dx_2 \qquad \text{(Bb-4)}$$

If the wave function of the system is approximated by a Slater determinant of SODS molecular orbitals (a common approximation), then

$$\rho(1,1') = \sum_{i=1}^{N} \phi_i^*(1)\, \phi_i(1') \qquad \text{(Bb-5)}$$

$$\pi(1,2) = \sum_{i,j=1}^{N} \left[\phi_i^*(1)\phi_j^*(2)\phi_i(1)\phi_j(2) - \delta_{s_i s_j}\, \phi_i^*(1)\phi_j^*(2)\phi_j(1)\phi_i(2) \right] \qquad \text{(Bb-6)}$$

RUEDENBERG (1962) lets,

$$\pi(1,2) = \rho(1)\rho(2) - \pi_x(1,2) \qquad \text{(Bb-7)}$$

where π_x is called the exchange pair density. Ruedenberg suggests that π_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 π_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 ρ, π, and π_χ, 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 $\langle\varphi_1|H|\varphi_2\rangle$ where $\varphi_1 \neq \varphi_2$; so, we consider $\varphi_1 = \bar{G}_i^\alpha \Phi_1 \chi_1$ and $\varphi_2 = \bar{G}_j^\beta \Phi_2 \chi_2$. The Nth-order density matrix is

$$g_{ij}^{\alpha\beta}(1,2,\cdots,N;\Phi_1\chi_1,\Phi_2\chi_2) \hat{=} N! \int (\bar{G}_i^\alpha \Phi_1 \chi_1)^* (\bar{G}_j^\beta \Phi_2 \chi_2)(ds) =$$

$$= \delta^{\alpha\beta} \frac{N!\,\Theta^\alpha}{f^\alpha\sqrt{2^{M_1}2^{M_2}}} \delta_{\sigma_{ij}} \sum_r (O_{ri}\Phi_1)^*(O_{rj}\Phi_2)$$

where Ba-8 has been used. $\sum_r (O_{ri}^\alpha \Phi_1)^*(O_{rj}^\alpha \Phi_2) = \frac{1}{(\Theta^\alpha)^2} \sum_\tau \sum_{\tau'} U_{ij\tau^{-1}\tau'}^\alpha (\tau\Phi_1)^*(\tau'\Phi_2)$

where we have used Aa-2 and $\sum_r U_{ir\tau^{-1}} U_{rj\tau'} = U_{ij\tau^{-1}\tau'}$

Thus, using $N! = \Theta^\alpha f^\alpha$

$$g_{ij}^{\alpha\beta}(1,2,\cdots,N;\Phi_1\chi_1,\Phi_2\chi_2) = \delta^{\alpha\beta} \delta_{\sigma_{ij}} \frac{1}{\sqrt{2^{M_1}2^{M_2}}} \sum_\tau \sum_{\tau'} U_{ij\tau^{-1}\tau'}^\alpha (\tau\Phi_1)^*(\tau'\Phi_2)$$

The second-order density matrix is defined as

$$g_{ij}^{\alpha\beta}(1,2;\Phi_1\chi_1,\Phi_2\chi_2) \hat{=} \frac{1}{(N-2)!} \int g_{ij}^{\alpha\beta}(1,2,\cdots,N;\Phi_1\chi_1,\Phi_2\chi_2)(dx_{12}'')$$

$$= \delta^{\alpha\beta} \frac{\Theta^\alpha}{f^\alpha\sqrt{2^{M_1}2^{M_2}}} N(N-1)\, \delta_{\sigma_{ij}} \sum_r \int (O_{ri}^\alpha \Phi_1)^*(O_{rj}^\alpha \Phi_2)(dx_{12}'') \qquad \text{(Bb-9)}$$

where $(dx_{12}'') = dx_3 dx_4 \ldots dx_n$. The first-order density matrix is defined as

$$g_{ij}^{\alpha\beta}(1;\Phi_1\chi_1,\Phi_2\chi_2) = \frac{1}{N-1} \int g_{ij}^{\alpha\beta}(1,2;\Phi_1\chi_1,\Phi_2\chi_2)\, dx_2$$

$$= \delta^{\alpha\beta} \frac{\Theta^\alpha N}{f^\alpha\sqrt{2^{M_1}2^{M_2}}} \delta_{\sigma_{ij}} \sum_r \int (O_{ri}^\alpha \Phi_1)^*(O_{rj}^\alpha \Phi_2)(dx_1') \qquad \text{(Bb-10)}$$

Consider $\int(\tau\Phi_1)^*(\tau'\Phi_2)(dx_1')$ 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 ρ and π (as is easily seen since the Virial then must be satisfied in both cases and only ρ leads to kinetic terms).

depend on the states which in the term Φ happen to be associated with specific electrons, the application of G_i^α makes the electrons indistinguishable) so the integral is invariant if we operate with $(i, \tau''(1))\tau''$. In particular, we choose $\tau'' = \tau^{-1}$. So,

$\int (\tau\Phi_1)^*(\tau'\Phi_2)(dx_1') = (1,\tau^{-1}(1)) \int \Phi_1^*(\tau^{-1}\tau'\Phi_2)(dx'_{\tau^{-1}(1)})$. Thus,

$N \sum_r \int (O_{ri}^\alpha \Phi_1)(O_{rj}^\alpha \Phi_2)(dx_1') = \frac{N!}{\Theta^\alpha} \sum_p (1,p) \int \Phi_1^*(O_{ij}^\alpha \Phi_2)(dx_p')$. Similarly,

$\int (\tau\Phi_1)^*(\tau'\Phi_2)(dx_{12}'') = (1,\tau^{-1}(1))(2,\tau^{-1}(2)) \int \Phi_1^*(\tau^{-1}\tau'\Phi_2)(dx''_{\tau^{-1}(1),\tau^{-1}(2)})$ and

$N(N-1) \sum_r \int (O_{ri}^\alpha \Phi_1)(O_{rj}^\alpha \Phi_2)(dx_{12}'') = \frac{N!}{\Theta^\alpha} \sum_{p,q}{}' (1,p)(2,q) \int \Phi_1^*(O_{ij}^\alpha \Phi_2)(dx_{pq}'')$

From THM Ac-1 $\sum_r \int (O_{ri}^\alpha \Phi_1)(O_{rj}^\alpha \Phi_2)(dx) = \rho^\alpha \int \Phi_1^*(O_{ij}^\alpha \Phi_2)(dx)$

For the general case

$$N(N-1)\cdots(N-r+1) \sum_r \int (O_{ri}^\alpha \Phi_1)(O_{rj}^\alpha \Phi_2)(dx_{12\cdots r}''^{\cdots\prime}) =$$
$$= \rho^\alpha \sum_{p_1,p_2,\cdots,p_r}{}' (1,p_1)(2,p_2)\cdots(r,p_r) \int \Phi_1 (O_{ij}^\alpha \Phi_2)(dx_{p_1p_2\cdots p_r}''^{\cdots\prime})$$

Using the preceding steps

$$g_{ij}^{\alpha\beta}(1;\Phi_1\chi_1,\Phi_2\chi_2) = \delta^{\alpha\beta}\delta_{\sigma_{ij}} \frac{\Theta^\alpha}{\sqrt{2^{m_1}2^{m_2}}} \sum_p (1,p) \int \Phi_1^*(O_{ij}^\alpha \Phi_2)(dx_p') \qquad \text{(Bb-11)}$$

$$g_{ij}^{\alpha\beta}(1,2;\Phi_1\chi_1,\Phi_2\chi_2) = \delta^{\alpha\beta}\delta_{\sigma_{ij}} \frac{\Theta^\alpha}{\sqrt{2^{m_1}2^{m_2}}} \sum_{p,q}{}' (1,p)(2,q) \int \Phi_1^*(O_{ij}^\alpha \Phi_2)(dx_{pq}'') \qquad \text{(Bb-12)}$$

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

From Ba-1 $\langle \Phi | O_{ij} | \Phi \rangle = \frac{2^n}{\Theta^\alpha} \delta_{ij}$; also,

$\int \Phi^*(O_{ij}^\alpha \Phi)(dx_p') = \frac{1}{\Theta^\alpha} \sum_\tau U_{ij\tau} \int \Phi^* \tau \Phi (dx_p')$ And,

$$\int \Phi^* \tau \Phi (dx_p') = \phi_{ip}^*(p)\phi_{ip}(p) \quad \text{if} \quad \tau \in T_\Phi$$
$$= 0 \quad \text{if} \quad \tau \notin T_\Phi$$

So,

$\int \Phi^*(O_{ij}^\alpha \Phi)(dx_p') = \frac{2^m}{\Theta^\alpha} \delta_{ij} \phi_{ip}^*(p)\phi_{ip}(p)$ Also,

$\int \Phi^* \tau \Phi (dx_{pq}) = \phi_{ip}^*(p)\phi_{iq}^*(q)\phi_{ip}(p)\phi_{iq}(q)$ If $\tau \in T_\Phi$ (in this case $U_{ij\tau} = \delta_{ij}$)

$= \phi_{ip}^*(p)\phi_{iq}^*(q)\phi_{iq}(p)\phi_{ip}(q)$ If $\tau = (p\ q)\tau'$ where $\tau' \in T_\Phi$ (in this case $U_{ij\tau} = U_{ij(p\ q)}$)

$= 0$ otherwise

The second case is different from the first if and only if $(p,q) \notin T_\Phi$

So, $\int \Phi^*(O_{ij}^{\alpha}\Phi)(dx''_{pq}) = \frac{2^m}{\Theta^\alpha}\left[\delta_{ij}\,\phi^*_{ip}(p)\phi^*_{iq}(q)\phi_{ip}(p)\phi_{iq}(q) + U'_{ij(p,q)}\phi^*_{ip}(p)\phi^*_{iq}(q)\phi_{iq}(p)\phi_{ip}(q)\right]$,

where $U'_{ij(p\,q)} \triangleq U_{ij(p\,q)}$ If $(p,q) \notin T_\Phi$

$= 0$ If $(p,q) \in T_\Phi$

Thus,

$$g^{\alpha\beta}_{ij}(1;\Phi\chi_{M_1},\Phi\chi_{M_2}) = \delta^{\alpha\beta}\delta_{ij}\,\delta_{M_1M_2}\sum_{p=1}^{N}\phi^*_{ip}(1)\phi_{ip}(1') \tag{Bb-13}$$

$$g^{\alpha\beta}_{ij}(1,2;\Phi\chi_{M_1},\Phi\chi_{M_2}) = \delta^{\alpha\beta}\delta_{M_1M_2}\sum_{p,q}{}'\left[\delta_{ij}\phi^{(1)}_{ip}\phi^{(2)}_{iq}\phi^{(1)}_{ip}\phi^{(2)}_{iq} + U'_{ij(p,q)}\phi^{(1)}_{ip}\phi^{(2)}_{iq}\phi^{(1)}_{iq}\phi^{(2)}_{ip}\right] \tag{Bb-14}$$

If i=j and $\alpha=\beta$, then Bb=13 and -14 are the ρ and π for the state $G_i^\alpha\Phi\chi_M$, 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, $G_i^\alpha\Phi\chi$ and $G_j^\alpha\Phi\chi$ have off-diagonal matrix elements (in the 'exchange terms') if more than one linearly independent G_i^α is possible for the same α and for the given Φ . The same thing, of course, occurs when one uses Slater determinants except that there are interactions between more Slater determinants than there are between G_i^α (the G_i^α 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 G_i^α (and thus, no off-diagonal matrix elements) for the same α 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, $M_s = \pm S$.

Now, consider $\Phi_1 \neq \Phi_2$. $\int \Phi_1^* O_{ij}\Phi_2(dx_p') = \frac{1}{\Theta^\alpha}\sum_\tau U_{ij\tau}\Phi_1^*\tau\Phi_2(dx_p')$ For this to be non-zero Φ_1 and Φ_2 can differ by at most one state; however, the differing states will generally be at different positions (i.e., not match). Let $\Phi_2' = \tau_1\Phi_2$ match Φ_1 except for one state (note Φ'_2 does not necessarily have all doubly-occupied states adjacent). If $\tau = \tau_1 \tau'$ then $\int \Phi_1^* \tau \Phi_2 (dx'_p) \neq 0$ only if $\tau' \in T_{\Phi_2}$ and p is the position of the non-matching states. If Φ_2 and Φ_1 differ by Φ_2 having one more c state and one less d state ($d \neq c$), write $\Phi_2 \triangleq \Phi_1(c;d)$. Then (where $n' = n_2$)

$$\int \Phi_1^* (O_{ij} \Phi_1(c;d))(dx'_p) = \delta_{di_p} \frac{2^{M'}}{\Theta} U_{ij\tau_1} \phi_d^*(p) \phi_c(p) .$$

For $\int \Phi_1^* \tau \Phi_2 (dx''_{pq}) \neq 0$, Φ_1 and Φ_2 can differ in at most two states. If Φ_2 differs by two states from Φ_1 and contains one more of c_1 and c_2 states and one less of d_1 and d_2 states than Φ_1, let $\Phi_2 \triangleq \Phi_1(c_1 c_2; d_1 d_2)$. Let $\Phi'_2 = \tau_1 \Phi_2$ match Φ_1 except in two positions and let d_1 and c_1 correspond

$$\int \Phi_1^* O_{ij} \Phi_1(c_1 c_2; d_1 d_2)(dx''_{pq}) = \frac{2^{M'}}{\Theta} \Big[U_{ij\tau_1} \big(\delta_{d_1 i_p} \delta_{d_2 i_q} \phi^*_{d_1}(p) \phi^*_{d_2}(q) \phi_{c_1}(p) \phi_{c_2}(q) + \\ + \delta_{d_1 i_q} \delta_{d_2 i_p} \phi^*_{d_2}(p) \phi^*_{d_1}(q) \phi_{c_2}(p) \phi_{c_1}(q) \big) + U_{ij(p,q)\tau_1} \big(\delta_{d_1 i_p} \delta_{d_2 i_q} \phi_{d_1}(p) \phi_{d_2}(q) \phi_{c_2}(p) \phi_{c_1}(q) + \\ + \delta_{d_1 i_q} \delta_{d_2 i_p} \phi^*_{d_2}(p) \phi^*_{d_1}(q) \phi_{c_1}(p) \phi_{c_2}(q) \big) \Big]$$

If Φ_2 and Φ_1 differ only by one state, set $d_2 = c_2$ equal to whatever the δ-function implies. Hence,

$$g^{\alpha\beta}_{ij}(1; \Phi_1 \chi_1, \Phi_1(c;d) \chi_2) = \delta^{\alpha\beta} \rho_{\sigma_{ij}} \sqrt{2^{M_2 - M_1}}\, U^{\alpha}_{ij\tau_1} \phi^*_d(1) \phi_c(1) \qquad \text{(Bb-15)}$$

$$g^{\alpha\beta}_{ij}(1,2; \Phi_1 \chi_1, \Phi_1(c;d) \chi_2) = \delta^{\alpha\beta} \rho_{\sigma_{ij}} \sqrt{2^{M_2 - M_1}} \sum_{i_p \neq d_1} \big[U_{ij\tau_1} \phi^*_{d_1}(1) \phi^*_{i_p}(2) \phi_{c_1}(1) \phi_{i_p}(2) + \\ + U_{ij(d_1,p)\tau_1} \phi^*_{d_1}(1) \phi^*_{i_p}(2) \phi_{i_p}(1) \phi_{c_1}(2) \big] \qquad \text{(Bb-16)}$$

$$g^{\alpha\beta}_{ij}(1,2; \Phi_1 \chi_1, \Phi_1(c_1 c_2; d_1 d_2) \chi_2) = \delta^{\alpha\beta} \rho_{\sigma_{ij}} \sqrt{2^{M_2 - M_1}} (\epsilon + (1,2)) \big[U_{ij\tau_1} \phi^*_{d_1}(1) \phi^*_{d_2}(2) \phi_{c_1}(1) \phi_{c_2}(2) + \\ + U_{ij(d_1,d_2)\tau_1} \phi^*_{d_1}(1) \phi^*_{d_2}(2) \phi_{c_1}(2) \phi_{c_2}(1) \big] \qquad \text{(Bb-17)}$$

Using Bb-13 through 17 any operator involving one- and two-electron operators can be evaluated for the expansion in terms of G_i^{α} of the exact wave function of Ba-13.

Bc: The Set of $G_i^\alpha \Phi \chi$ With the Same Φ ; Hund's Rule and the Heisenberg Hamiltonian

Consider the matrix elements of the Hamiltonian between the various G_i^α for the same space state Φ. $H = \sum_i h_i + \sum_{i<j} g_{ij}$

Using Bb-2, 4, 13, 14

$$E_{ij}^{\alpha\beta} = \delta^{\alpha\beta} \Big\{ \delta_{ij} \Big[\sum_{p=1}^{N} \langle \phi_{ip} | h | \phi_{ip} \rangle + \frac{1}{2} \sum_{p,q}{}' \langle \phi_{ip} \phi_{iq} | g | \phi_{ip} \phi_{iq} \rangle \Big] + \\ + \frac{1}{2} \sum_{p,q}{}' U'^{\alpha}_{ij(p,q)} \langle \phi_{ip} \phi_{iq} | g | \phi_{iq} \phi_{ip} \rangle \Big\} \tag{Bc-1}$$

It is common to define

$$\begin{aligned} J_{pq} &= \langle \phi_{ip} \phi_{iq} | g | \phi_{ip} \phi_{iq} \rangle = \text{"Coulomb integral"} \\ K_{pq} &= \langle \phi_{ip} \phi_{iq} | g | \phi_{iq} \phi_{ip} \rangle = \text{"Exchange integral"} \end{aligned} \tag{Bc-2}$$

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

1	2
3	4
5	6
⋮	⋮
2n-1	2n
⋮	⋮

I call a tableau of this form the lowest-in-2n. So the possible α and i for the above case are in a one-to-one correspondence with the possible α' 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

$U_{ij(p,q)} = - \frac{1}{2} \delta_{ij}$ if $p \leq 2n$ and q is any letter not in the same row as p. Thus,

$$E_{ij}^{\alpha} = A\,\delta_{ij} + \frac{1}{2} \sum_{p,q>2n}{}' U^{\alpha}_{ij(p,q)} K_{pq} \tag{Bc-3}$$

where A (which is independent of α, i, and j) is

$$A = \Big[\sum_{p=1}^{N} \langle 1p|h|1p\rangle + \frac{1}{2}\sum_{p,q\le 2n}\left(J_{pq} - \frac{1}{2}K_{pq}\right) + \frac{1}{2}\sum_{\substack{p\le 2n\\ q>2n}}\left(J_{pq} + J_{qp} - \frac{1}{2}K_{pq} - \frac{1}{2}K_{qp}\right) + \\ + \frac{1}{2}\sum_{p,q>2n}{}' J_{pq}\Big] \qquad \text{(Bc-4)}$$

where we used $J_{pp} = K_{pp}$. We now delete the A and obtain

$$E_{ij}^{\alpha} = \frac{1}{2}\sum_{p,q>2n}{}' U_{ij(p,q)} K_{pq} \qquad \text{(Bc-5)}$$

For α with more than one non-zero $G_i^{\alpha}\Phi\chi$ we must diagonalize E_{ij}^{α}. In so doing we take a unitary transformation on the $G_i^{\alpha}\Phi\chi$ of the same α which subjects each matrix $U_{ij(p,q)}^{\alpha}$ 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 ($|U_{ii}|^2 \le \sum_j U_{ij}U_{ij}^{*} = 1$); thus, if V is the transformed U, $E_i^{\alpha} = \sum_{p,q>2n} V_{ii(p,q)}^{\alpha} K_{pq}$ and $|V_{ii}| \le 1$.

The K_{pq} 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 acheive 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 $G_i^{\alpha}\Phi\chi$. If $p > 2n$, $U_{ij(p,p+1)}^{\alpha} = -\delta_{ij}$; thus, $U_{ij(p,q)}^{\alpha} = -\delta_{ij}$ for $p > q > 2n$. So the $E_{ij}^{\alpha} = -\delta_{ij}\sum_{p,q>2n}{}' K_{pq}$ 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} and would thus be quite unlikely.

with the same orbital occupations is the state of highest spin. We have not stated how the ϕ_{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 ϕ_i which minimize the energy, then the best ϕ_i are actually different for different j and same α of G_j^α. The point is that regardless of which ϕ_i are used the α of the maximum spin for which $G_j^\alpha \Phi \chi \neq 0$ 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_i 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 Bc-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 system of electrons and despite the fact that for atoms it is more resricted than Hund's rule.

Consider $\sum_i E^{\alpha}_{ii} = \sum_{p>q>2n} K_{pq} \sum_i U_{ii(p,q)}$. Pretend that the closed shells are not there; then we have r = N-2n electrons and r different one-electron states; so, all f^{α} of the $G^{\alpha}_i \Phi\chi$ are allowed. Then $\sum_i U^{\alpha}_{ii(p,q)} = \chi^{\alpha}_t$ = character of transposition (same for all transpositions). $E_{ave} \overset{\Delta}{=} \frac{1}{f^{\alpha}} \sum_i E^{\alpha}_{ii} = \frac{\chi^{\alpha}_t}{f^{\alpha}} \sum_{p>q>2n} K_{pq}$. From appendix B

$\frac{\chi^{\alpha}_t}{f^{\alpha}} = -\frac{2}{r(r-1)}\left[\frac{r(r-4)}{4} + S(S+1)\right]$; thus,

$$E^{\alpha}_{AVE} = -\frac{2}{r(r-1)}\left[\frac{r(r-4)}{4} + S(S+1)\right] \sum_{p>q>2m} K_{pq} \qquad \text{(Bc-6)}$$

$$S(S+1) G^{\alpha}_1 \Phi\chi = \hat{S}^2 G^{\alpha}_1 \Phi\chi = \left(\sum_{k=1}^{r} \hat{S}_k\right)\left(\sum_{l=1}^{r} \hat{S}_l\right) G^{\alpha}_1 \Phi\chi = \left(\frac{3}{4}r + \sum_{k,l}{}' \hat{S}_k \cdot \hat{S}_l\right) G^{\alpha}_1 \Phi\chi \qquad \text{(Bc-7)}$$

thus,

$$E_{ave} = -\left[+ \frac{1}{2} + \frac{2}{r(r-1)} \sum_{k,l}{}' \hat{S}_k \cdot \hat{S}_l \right]\left(\sum_{p>q>2m} K_{pq}\right) \qquad \text{(Bc-8)}$$

If we now set all the K_{pq} equal

$E^{\alpha}_{ave} = -\left[\frac{r(r-1)}{4} + \sum_{k,l}{}' \hat{S}_k \cdot \hat{S}_l\right] K = -\sum_{k,l}{}' \left(\frac{1}{4} + \hat{S}_k \cdot \hat{S}_l\right) K$ and if we now let $K = K_{kl}$ for the kl term, etc.,

$$E^{\alpha}_{ave} = -\sum_{k,l}{}' \left(\frac{1}{4} + \hat{S}_k \cdot \hat{S}_l\right) K_{kl} \qquad \text{(Bc-9)}$$

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^{α} different states of the same α (same total spin) are at f^{α} different energies; so, the differences between E_{ave} for different α 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, $q > 2n$ are the same as for the case where there are only $r = N-2n$ electrons (e.g., see the lemma to 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 ϕ_i for different G_i^{α} and then we require that the $G_i^{\alpha}\Phi\chi$ or linear combinations (with the same $\Phi\chi$) form eigenstates of the N-electron system. It is all right to form $G_i^{\alpha}\Phi\chi$ with some set of ϕ_i and then to vary the ϕ_i to try to obtain a good approximation to an eigenstate, but we must do this for each α separately. For the same basic set of ϕ_i we would vary, among other parameters, the scale of the different ϕ_i, (we do not change the type of symmetry of the ϕ_i) to minimize the energy and find the ground state of each, but the scale changes of the same ϕ_i for different α would in general be different. Hence, the integrals in (Bc-1) are different for different α. 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.

Bd: The Hartree-Fock Equations for the $G_i^\alpha \Phi \chi$

Let us try to find the lowest energy $G_i^\alpha \Phi \chi$ for a given type of Φ. Each ϕ_i 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 ϕ_i in Φ 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 Φ'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 $\delta E_{ii}^\alpha = 0$ under the constraint that the N-n occupied one-electron functions remain orthogonal (this also preserves the norm of the many-electron function). Thus,

$$\delta \left(E_{ii}^\alpha - \sum_{p,q} \epsilon_{pq} \langle \phi_p | \phi_q \rangle \right) = 0 \tag{Bd-1}$$

Taking the complex conjugate of this equation, we see that ϵ_{pq} is Hermitian . Using Bc-1

$$\delta\Big[\sum_{p=1}^{N}\langle i_p|h|i_p\rangle+\frac{1}{2}\sum_{p,q}{}'\big(\langle i_p i_q|g|i_p i_q\rangle+U'_{ii(p,q)}\langle i_p i_q|g|i_q i_p\rangle\big)-\sum_{p,q}\epsilon_{pq}\langle i_p|i_q\rangle\Big]=0$$

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

$$\sum_{p=1}^{N}\Big[\langle\delta\phi_{i_p}|h|i_p\rangle+\sum_{q\neq p}\big(\langle\delta\phi_{i_p}\phi_{i_q}|g|i_p i_q\rangle+U'_{ii(p,q)}\langle\delta\phi_{i_p}\phi_{i_q}|g|i_q i_p\rangle\big)+ -\sum_{q}\epsilon_{pq}\langle\delta\phi_{i_p}|i_q\rangle\Big]=0 \tag{Bd-2}$$

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

$$\Big\{h|\phi_{i_p}\rangle+\sum_{q\neq p}\big(\langle\phi_{i_q}|g|\phi_{i_q}\rangle|\phi_{i_p}\rangle+U'_{ii(p,q)}\langle\phi_{i_q}|g|\phi_{i_p}\rangle|\phi_{i_q}\rangle\big)-\sum_{q}\epsilon_{pq}|\phi_{i_q}\rangle\Big\}=0 \tag{Bd-3}$$

If the coefficient of $\delta\phi_{i_p}$ 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 ϕ_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 ϕ_{i_p} after solving for the integrals from the previous approximation (this is impractical except for atoms). Consider $\langle\phi_{i_p}|$ and $\langle\phi_{i_r}|$ on equation Bd-3 (and use $\langle\phi_{i_p}|\phi_{i_q}\rangle=\delta_{i_p i_q}$)

$$\langle i_p|h|i_p\rangle + \sum_{q\neq p}\left(\langle i_p i_q|g|i_p i_q\rangle + U'_{ii}(p,q)\langle i_p i_q|g|i_q i_p\rangle\right) = \epsilon_{pp} \qquad \text{(Bd-4)}$$

$$\langle i_r|h|i_p\rangle + \sum_{q\neq p}\left(\langle i_r i_q|g|i_p i_q\rangle + U'_{ii}(p,q)\langle i_r i_q|g|i_q i_p\rangle\right) = \epsilon_{pr} \qquad \text{(Bd-5)}$$

Since h and g are invariant under the spatial symmetry group, $\epsilon_{pr} = 0$ unless ϕ_{i_p} and ϕ_{i_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 $\epsilon_{pr} = 0$ for different symmetry functions since $\langle\phi_{i_p}|\phi_{i_q}\rangle = 0$ by symmetry. A transformation among the occupied states which leaves $G_i^\alpha\Phi\chi$ invariant (except for a phase factor) leaves E_{ii}^α unchanged. If $\phi_i = \sum_j S_{ij}\phi_j$ (where the sum is over all different states), then

$$\sum_{p,q}\epsilon_{pq}\langle\phi_p|\phi_q\rangle = \sum_{i,j}\langle\phi_i|\phi_j\rangle\sum_{p,q}S^*_{pi}\epsilon_{pq}S_{qj}\,.$$

$\underline{\underline{S}}$ is unitary (since we require that orthonormality of the ϕ_i be preserved) and $\underline{\underline{\epsilon}}$ is Hermitian; thus, an $\underline{\underline{S}}$ can be chosen which will diagonalize $\underline{\underline{\epsilon}}$. We wish to preserve the symmetry properties of the ϕ_i (which is possible since the only non-diagonal ϵ_{pq} are between states of the same symmetry). It is now necessary to see in which cases the transformation $\underline{\underline{S}}$ leaves $G_i^\alpha\Phi\chi$ invariant.

Consider a set of functions which are transformed by $\underline{\underline{S}}$ among themselves; then

$$\begin{aligned}\Phi = \phi_{i_1}(1)\phi_{i_2}(2)\cdots\phi_{i_N}(N) &= \sum_{j_1\cdots j_N} S_{j_1 i_1}S_{j_2 i_2}\cdots S_{j_N i_N}\,\phi_{j_1}(1)\phi_{j_2}(2)\cdots\phi_{j_N}(N) \qquad \text{(Bd-6)}\\ &= \sum_{\tau} S_{\tau(i_1),i_1}\cdots S_{\tau(i_N),i_N}\left(\tau\phi_{i_1}(1)\cdots\phi_{i_N}(N)\right)\end{aligned}$$

and consider $G_i^\alpha(\tau\Phi)\chi$. Assume that <u>all</u> of the functions in the transformed set are doubly-occupied orbitals. If the functions ϕ_{i_p} and ϕ_{i_q} are transposed, then we must use for τ in Bd-6 say

say $\tau = (2p-1,2q-1)(2p,2q)$ or $(2p-1,2q)(2p,2q-1)$ (Bd-7)

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

$G_i^\alpha(\tau\Phi)\chi = \sum_r b_{\sigma_{r\alpha}}(O_{r\alpha}\tau\Phi)(\omega_{\bar{r}\bar{i}}\chi)$

$O_{r\alpha}\tau\Phi = \sum_j U_{ij\tau}O_{rj}\Phi$ (from Aa-4), but by THM Ab-2

$O_{rj}\Phi = \delta_{ij}\,O_{ri}\Phi$ (since Φ is ordered and S_i is lowest-in-2n). Thus,

$O_{ri}\tau\Phi = U_{ii\tau}O_{ri}\Phi$. From appendix C $U_{ij(2p-1,2q-1)(2p,2q)} = \delta_{ij}$, and thus for any permutation τ composed of transpositions of type Bd-7 $U_{ij\tau} = \delta_{ij}$ thus, $G_i^\alpha(\tau\Phi)\chi = G_i^\alpha\Phi\chi$, and the $G_i^\alpha\Phi\chi$ 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 ϵ_{pq} between doubly-occupied orbitals are zero**. The ϵ_{pq} between doubly-occupied orbitals and singly-occupied

---

* In case anyone is worrying about numbering the electrons when they are indistinguishable, the particles in a function such as Φ 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 ϵ_{pq} between functions transforming as different basis functions of 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 irreducible representation and if this irreducible representation occurs only once* then the $\underline{\underline{\epsilon}}$ can be completely diagonalized** (e.g., for the (argon)d^n and (argon)$d^{n-2}s^2$ configurations of the <u>first</u> row transition elements). A case such as lithium would involve off-diagonal ϵ_{pq}. These same problems occur for the same reasons for Slater determinants despite the common derivation (e.g., MESSIAH (1962) p. 777) showing that the ϵ_{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 restriction which is justified because of the vast simplification in calculations (making them feasible) and because the restriction usually fits with whatever experimental information is available. If <u>no</u> restrictions are made on the spin-orbitals other than their orthonormality, then indeed a transformation can always be performed to diagonalize the ϵ_{pq};

---

* This is true even if some of the states of this irreducible representation are doubly-occupied, the reason is that ϵ_{pq} is non-zero only between states which transform as the same basis function of the same irreducible representation, thus an irreducible representation <u>must</u> occur twice in order to have off-diagonal ϵ_{pq}.

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

*** Actually we required that the ϕ_i be basis functions of the irreducible representations of the spatial symmetry group; however, even without this additional assumption the unitary transformations of Φ in order to leave E^{α}_{11} 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 $\alpha = [2^n, 1^{N-2n}]$, see p. Cb-1,2) and thus, the ϵ_{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 ϵ_{pq} to be diagonalized. From Bd-3

$$h|\phi_{i_p}\rangle + \sum_{q\neq p}\left(\langle\phi_{i_q}|g|\phi_{i_q}\rangle|\phi_{i_p}\rangle + U''_{ii(p,q)}\langle\phi_{i_q}|g|\phi_{i_p}\rangle|\phi_{i_q}\rangle\right) = \epsilon_{pp}|\phi_{i_p}\rangle \qquad \text{(Bd-8)}$$

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

$\triangleq -\frac{1}{2}$ if $i_p = i_q$ and $p,q \leq 2n$ (to correct for the extra Coulomb term)

$\triangleq -1$ if $i_p = i_q$ and $p,q > 2n$

Using the results of appendix A for $U_{ij(p,q)}$ as in section Bc, if $p \leq 2n$

$$h|\phi_{i_p}\rangle + \sum_{q}\left(\langle\phi_{i_q}|g|\phi_{i_q}\rangle|\phi_{i_p}\rangle - \tfrac{1}{2}\langle\phi_{i_q}|g|\phi_{i_p}\rangle|\phi_{i_q}\rangle\right) = \epsilon_{pp}|\phi_{i_p}\rangle \qquad \text{(Bd-9)}$$

if $p > 2n$

$$h|\phi_{i_p}\rangle + \sum_{q\leq 2n}\left(\langle\phi_{i_q}|g|\phi_{i_q}\rangle|\phi_{i_p}\rangle - \tfrac{1}{2}\langle\phi_{i_q}|g|\phi_{i_p}\rangle|\phi_{i_q}\rangle\right) + \sum_{q>n}\left(\langle\phi_{i_q}|g|\phi_{i_q}\rangle|\phi_{i_p}\rangle + {}\right. \qquad \text{(Bd-10)}$$
$$\left. + U''_{ii(p,q)}\langle\phi_{i_q}|g|\phi_{i_p}\rangle|\phi_{i_q}\rangle\right) = \epsilon_{pp}|\phi_{i_q}\rangle$$

Now, in order for equations Bd-9 and 10 to be consistent with the restrictions on the ϕ_{i_p}, the equations above must transform as ϕ_{i_p}; so, the term operating on ϕ_{i_p} must be invariant under transformations of the spatial symmetry group. The sum $\sum_i \phi_i^*(x_1)\phi_i(x_2)$ 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)

---

* $R\sum_i \phi_i^*(x_1)\phi_i(x_2) = \sum_{j,k}\phi_j^*(x_1)\phi_k(x_2)\sum_i R_{ji}^* R_{ki}$; R is unitary since the orthonormality 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, $\sum_i R_{ji}^* R_{ki} = \delta_{jk}$. 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 $U_{ii(p,q)}$ must be the same for all q of the same shell (e.g., if there is only one unfilled shell then $U_{ii(p,q)}$ must be independent of p and q for $p,q > 2n$; the only case is $U_{ii(p,q)} = -1$ (maximum spin)). In cases other than these we <u>cannot</u> 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 U''_{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 ϕ_{i_p} 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 (α or β) is associated with each orbital. One gets an equation similar to Bd-8 except that $U''_{ii(p,q)}$ 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 <u>not</u> 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 α spin and once with a β 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 ϕ_{i_p} 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 ϵ_{pq} is needed when the irreducible representation used in the singly-occupied set is repeated.

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

$$\Big[H_{\mu\nu} + \sum_{q \neq p} c^*_{q\sigma} c_{q\lambda} \big(g_{\mu\sigma;\nu\lambda} + U'_{\lambda\lambda(p,q)} g_{\mu\sigma;\lambda\nu}\big)\Big] c_{p\nu} = \epsilon_{pp} S_{\mu\nu} c_{p\nu} \,, \qquad \text{(Bd-11)}$$

where $\phi_{i_p} = \sum_{\mu=1}^{M} c_{p\mu} \chi_\mu$, $M \geq N-m$, χ_μ is a fixed basis function

$H_{\mu\nu} = \langle \chi_\mu | h | \chi_\nu \rangle$, $g_{\mu\sigma;\nu\lambda} = \langle \chi_\mu \chi_\sigma | g | \chi_\nu \chi_\lambda \rangle$

Substituting for the $U'_{ii(p,q)}$, if $p \leq 2n$

$$\Big[H_{\mu\nu} + \sum_{q} c^*_{q\sigma} c_{q\lambda} \big(g_{\mu\sigma;\nu\lambda} - \tfrac{1}{2} g_{\mu\sigma;\lambda\nu}\big)\Big] c_{p\nu} = \epsilon_{pp} S_{\mu\nu} c_{p\nu} \qquad \text{(Bd-12)}$$

---

* 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-8 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-8 has one exchange term between singly-occupied orbitals and one-half exchange term between a singly-occupied orbital and each occurence 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 $p > 2n$

$$\Big[H_{\mu\nu} + \sum_{q\leq 2m} C^{*}_{q\sigma} C_{q\lambda}\left(g_{\mu\sigma;\nu\lambda} - \tfrac{1}{2}\, g_{\mu\sigma;\lambda\nu}\right) + \sum_{q>2m} C^{*}_{q\sigma} C_{q\lambda}\big(g_{\mu\sigma;\nu\lambda} + {} + U''_{\lambda\lambda(p,q)}\, g_{\mu\sigma;\lambda\nu}\big)\Big] C_{p\nu} = \epsilon_{pp} S_{\mu\nu} C_{p\nu} \qquad \text{(Bd-13)}$$

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 x 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 ϕ_i being a basis function of an irreducible representation only if $\alpha = [2^{\overset{P}{n}}, 1^{N-2m}]$ in which case $U''_{ii(p,q)} = -1$, 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 $G^{\alpha}_{i}\Phi\chi$ is left invariant (see discussion after Bd-7). Under these transformations the Hartree-Fock Hamiltonian (Be-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 ϵ_{pr} (Bd-5) are non-zero. So, instead of Bd-8 the equation for the new orbitals is

$$h|\phi_{1p}\rangle + \sum_q \left(\langle\phi_{1q}|g|\phi_{1q}\rangle|\phi_{1p}\rangle + U''_{11(p,q)}\langle\phi_{1q}|g|\phi_{1p}\rangle|\phi_{1q}\rangle \right) =$$
$$= \sum_q \Big[\langle\phi_{1q}|h|\phi_{1p}\rangle + \sum_r \big(\langle\phi_{1q}\phi_{1r}|g|\phi_{1p}\phi_{1r}\rangle + U''_{11(p,r)}\langle\phi_{1q}\phi_{1r}|g|\phi_{1r}\phi_{1p}\rangle \Big] |\phi_{1q}\rangle \quad \text{(Bd-14)}$$

which is extremely inconvenient to use.

Be: Perturbation of the Hartree-Fock Equations for $G_i^{\alpha}\Phi\chi$

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 ϕ_{i_p} and ϕ_{i_q} with a two-electron function $\Theta_{i_p i_q}$. If $i_p = i_q$ then the $\Theta_{i_p i_p}$ can be found by solving an inhomogeneous differential equation. If $i_p \neq i_q$ then $\Theta_{i_p i_q}$ is expressed as a sum over the excited states of the Hartree-Fock Hamiltonian. This allows us to accurately determine the $\Theta_{i_p i_p}$ (which should be large since the states were artificially forced to be the same) and to determine $\Theta_{i_p i_q}$ by a simpler less accurate method more in keeping with the lesser importance of $\Theta_{i_p i_q}$. It should be stressed that we <u>do</u> include systems with half-filled shells and maximum spin.

From Bd-8, the equation for the Hartree-Fock molecular orbitals is

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

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

$\triangleq -\frac{1}{2}$ if $i_p = i_q$ and $p,q \leq 2n$

$\triangleq -1$ if $i_p = i_q$ and $p,q > 2n$

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

$(h + V_p) | i_p \rangle = \epsilon_{pp} | i_p \rangle$. Take $\langle i_p |$ and obtain

$$\epsilon_{pp} = \left[\langle i_p | h | i_p \rangle + \sum_q \left(\langle i_p i_q | g | i_p i_q \rangle + U''_{ii(p,q)} \langle i_p i_q | g | i_q i_p \rangle\right)\right]$$

Let $H_o \triangleq \sum_{p=1}^{N} (h_p + V_p)$; so, $H_o \phi_{i_1} \cdots \phi_{i_N} = \left(\sum_{p=1}^{N} \epsilon_{pp}\right) \phi_{i_1} \cdots \phi_{i_N}$.

Then multiply by χ and operate with $\bar{G}_i^{\alpha}$; thus,

$$H_o \bar{G}_i^{\alpha} \Phi \chi = \left(\sum_{p=1}^{N} \epsilon_{pp}\right) \bar{G}_i^{\alpha} \Phi \chi \quad \text{AND} \quad E_o = \sum_{p=1}^{N} \epsilon_{pp} \tag{Be-1}$$

The perturbation is $H_1 = \sum_{p<q=1}^{N} g_{pq} - \sum_{p=1}^{N} V_p$. The first-order energy correction is

$$E_1 = -\frac{1}{2} \sum_{p,q} \left(\langle pq | g | pq \rangle + U''_{ii(p,q)} \langle pq | g | qp \rangle\right) \tag{Be-2}$$

This first-order correction merely corrects for the electron-electron interaction energies being counted twice in E_o . Let $\langle \psi_o | \psi_o \rangle = 1$, then (e.g., SINANOǦLU (1961a)) $(H_o - E_o) X_1 = (E_1 - H_1) \psi_o$, where X_1 is the first-order correction to ψ_o .

From above

$$E_1 - H_1 = \sum_{p>q=1}^{N} \left(J_{pq} + U'_{ii(p,q)} K_{pq} - g_{pq}\right) + \sum_{p=1}^{N} \left[V_p - \sum_{q=1}^{N} \left(J_{pq} + U''_{ii(pq)} K_{pq}\right)\right]$$

let $X_1 \triangleq Y - W$; $m_{pq} \triangleq J_{pq} + U'_{ii(p,q)} K_{pq} - g_{pq}$;

$$k_{pq} \underset{p\neq q}{\hat{=}} J_{pq} + U'_{\lambda\lambda(p,q)} K_{pq} - \langle \lambda_q | g | \lambda_q \rangle - U'_{\lambda\lambda(p,q)} \langle \lambda_q | g | \lambda_p \rangle P_{pq} .$$

we have

$$(H_0 - E_0) Y = \Big(\sum_{p<q} m_{pq}\Big)\psi_0 \; ; \; (H_0 - E_0) W = \Big(\sum_{p,q}{}' k_{pq}\Big)\psi_0 \qquad \text{(Be-3)}$$

ψ_0 is non-degenerate (except for the 2S + 1 different spin states); so, $\langle \psi' | \sum_{p<q} m_{pq} | \psi_0 \rangle = 0$ for all solutions (i.e., ψ_0) of the homogeneous equation, and thus solutions to the inhomogeneous equation for Y exist. We have $Y = \frac{1}{H_0 - E_0}\Big(\sum_{p<q} m_{pq}\Big)\psi_0 = \Big(\sum_{p,q} M_{pq}\Big)\psi_0$, where $M_{pq} \hat{=} \frac{m_{pq}}{H_0 - E_0}$. 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 $G_1^\alpha \Phi \chi \neq 0$ if Φ is written in standard order. Thus, $\Psi_0 = \bar{G}_1^\alpha \Phi_0 \chi$.

Since the one-electron orbitals and spin functions form a complete set of functions on the one-electron space, then the product functions $\psi = \phi_1 \cdots \phi_N s_1 \cdots s_N$ form a complete set of functions on the N-electron space. Thus, since the $G_i^\alpha \Phi \chi$ is obtained by a non-singular transformation from the ψ , then the set of G_i for all α , λ , and ordered Φ and χ is complete. Thus, $1 = \sum_{\beta,\lambda,\Phi,M_S} |\bar{G}_\lambda^\beta \Phi \chi_M \rangle \langle \bar{G}_\lambda^\beta \Phi \chi_M |$ and

$$Y = \sum_{\beta,\lambda,\Phi,\chi_{M'}} |\bar{G}_\lambda^\beta \Phi \chi_{M'} \rangle \langle \bar{G}_\lambda^\beta \Phi \chi_{M'} | \sum_{p<q} M_{pq} | \bar{G}_1^\alpha \Phi_0 \chi_M \rangle$$

$$= \sum_{\lambda,\Phi} |\bar{G}_\lambda^\alpha \Phi \chi_M \rangle \langle \bar{G}_\lambda^\alpha \Phi \chi_M | \sum_{p<q} M_{pq} | \bar{G}_1^\alpha \Phi_0 \chi_M \rangle \qquad \text{(Be-4)}$$

From before if $\Phi = \Phi_0$, then $\langle \psi | \sum_{p<q} M_{pq} | \psi_0 \rangle = 0$. From Bb-12 (or by just expanding) $\langle \bar{G}_\lambda^\alpha \Phi \chi_S | \sum_{p<q} M_{pq} | \bar{G}_1^\alpha \Phi_0 \chi_S \rangle = \rho_{\sigma_{\lambda 1}} \frac{\Theta^\alpha}{\sqrt{2^{m_0} 2^m}} \langle \Phi | \Big(\sum_{p<q} M_{pq}\Big) O_{\lambda 1}^\alpha | \Phi_0 \rangle$

Use Ba-12 for $\bar{G}_i^\alpha \Phi \chi$ and Ba-8 for $C_{m_0}^\alpha$, then

$$Y = C_{m_0}^\alpha \sum_{\lambda,\Phi,\tau} \frac{\rho_{\sigma_{\lambda 1}}}{2^m} |G_\lambda^\alpha \Phi \chi \rangle U_{\lambda 1 \tau} \langle \tau^{-1} \Phi | \sum_{p<q} M_{pq} | \Phi_0 \rangle$$

. Let $\Phi = \tau \Phi'$, then

$$G_{\underline{i}}^{\alpha}\Phi\chi = \sum_{r} \zeta_{\sigma_{r\lambda}} (O_{r\lambda}\,\tau\Phi')(\omega_{\bar{r}\bar{q}}\chi_{q}) = \sum_{t} U_{\lambda t\tau} \sum_{r} \zeta_{\sigma_{r\lambda}} (O_{rt}\Phi')(\omega_{\bar{r}\bar{q}}\chi_{\bar{q}})$$
$$= \sum_{t} U_{\lambda t\tau}\, \zeta_{\sigma_{\lambda t}}\, G_{\underline{t}}^{\alpha}\Phi'\chi$$

Thus, $\sum_{\lambda} \zeta_{\sigma_{\lambda 1}} | G_{\underline{\lambda}}\Phi\chi\rangle U_{\lambda 1\tau} = \sum_{\lambda}\sum_{t} \zeta_{\sigma_{\lambda 1}}\zeta_{\sigma_{\lambda t}} U_{\lambda t\tau} U_{\lambda 1\tau} | G_{\underline{t}}^{\alpha}\Phi'\chi\rangle$

$$= \sum_{t} \zeta_{\sigma_{t1}} | G_{\underline{t}}\Phi'\chi\rangle \sum_{\lambda} U_{\lambda t\tau} U_{\lambda 1\tau} = G_{\underline{1}}\Phi'\chi = G_{\underline{1}}(\tau^{-1}\Phi\chi)$$

So, $Y = C_{M_0}^{\alpha} \sum_{\Phi,\tau} \frac{1}{2^m} | G_{\underline{1}}(\tau\Phi)\chi\rangle\langle\tau\Phi | \sum_{p<q} M_{pq} | \Phi_0\rangle$; here τ goes over all permutations of S_N. If we restrict the τ to ones yielding different Φ, then we get a factor of 2^n and

$$Y = C_{M_0}^{\alpha} \sum_{\text{DIFF}\,\Phi'} | G_{\underline{1}}\Phi'\chi\rangle\langle\Phi' | \sum_{p<q} M_{pq} | \Phi_0\rangle \qquad \text{(Be-5)}$$

But, for $\langle\Phi' | M_{pq} | \Phi_0\rangle \neq 0$, Φ must match Φ_0 in all positions excepting p and q. Let $\Phi' = \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}} \phi_k(p)\phi_m(q)$ differ from Φ_0 only by having ϕ_k rather than ϕ_{i_p} in position p and ϕ_m rather than ϕ_{i_q} in position q (of course ϕ_k may equal ϕ_{i_p}, etc.); then

$$Y = C_{M_0}^{\alpha} \sum_{p<q} \sum_{k,m} | G_{\underline{1}} \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}} \phi_k\phi_m\chi\rangle\langle km | M_{pq} | i_p i_q\rangle \qquad \text{(Be-6)}$$

Define $\Theta'_{i_p i_q}(p,q) \triangleq \sum_{\substack{k,m\\ \text{ALLOWED}}} \phi_k(p)\phi_m(q)\langle km | M_{pq} | i_p i_q\rangle$, where $i_p \leq i_q$ and the sum over k and m only goes over the ϕ_k and ϕ_m such that $G_{\underline{1}} \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}} \phi_k\phi_m\chi \neq 0$ and also for which the set ϕ_k and ϕ_m is not equal to the set ϕ_{i_p} and ϕ_{i_q} (see the remark after Be-4). Then $Y = C_{M_0}^{\alpha} \sum_{p<q} | G_{\underline{1}} \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}} \Theta'_{i_p i_q}\chi\rangle$. There are three possibilities for $\Theta'_{i_p i_q}$; both i_p and i_q are doubly-occupied, one is doubly-occupied, or none is doubly-occupied; I use FF, FH 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 Φ_0

$k > i_{2M_0}$ denote that k can be any orbital not doubly-occupied in Φ_0

(i.e., k can be a singly-occupied orbital in Φ_0 or else an orbital not contained in Φ_0)

$$\Theta'^{HH}_{i_p i_q}(p,q) = \sum_{\substack{k,m \in \{x : x = i_p,\ x = i_q,\ x > i_N\} \\ k \neq m}} \phi_k(p)\phi_m(q) \langle \phi_k \phi_m | M_{12} | \phi_{i_p} \phi_{i_q} \rangle \tag{Be-7}$$

Neither k nor m can be a doubly-occupied orbital of Φ_o since Φ 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 Φ_o other than i_p or i_q since we would have $n = n_o + 1$ (Φ has $n_o + 1$ 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 $n = n_o + 1$

$$\Theta'^{FH}_{i_p i_q}(p,q) = \sum_{k,m} \phi_k(p)\phi_m(q) \langle \phi_k \phi_m | M_{12} | \phi_{i_p} \phi_{i_q} \rangle \tag{Be-8}$$

$$\begin{array}{ll}
k > i_N \text{ AND } m > i_N & \\
k > i_N \text{ AND } i_{2M_0} < m \leq i_N & \uparrow \\
i_{2M_0} < k \leq i_N \text{ AND } m > i_N & \text{DOUBLY-EXCITED} \\
\hline
k = i_q \text{ AND } i_{2M_0} < m \leq i_N & \text{SINGLY EXCITED} \\
i_{2M_0} < k \leq i_N \text{ AND } m = i_q & \downarrow \\
k = i_p \text{ AND } m > i_N & \\
k > i_N \text{ AND } m = i_p &
\end{array}$$

The other possibilities yield either some orbital triply-occupied or else result in $n > n_o$ (note that $n < n_o$ allows $G_1 \Phi \chi = 0$ and is thus acceptable)

$$\Theta'^{FF}_{i_p i_q}(p,q) = \sum_{k,m} \phi_k(p)\phi_m(q) \langle k m | M | i_p i_q \rangle \tag{Be-9}$$

$$i_p \neq i_q \left\{ \begin{array}{ll}
k > i_N ;\ m > i_N & \\
k > i_N ;\ i_{2M_0} < m \leq i_N & \\
i_{2M_0} < k \leq i_N ;\ m > i_N & \uparrow \\
i_{2M_0} < k \leq i_N ;\ m > i_N & \text{DOUBLY EXCITED} \\
\hline
k > i_{2M_0} ;\ m = i_p \text{ OR } i_q & \text{SINGLY EXCITED} \\
k = i_p \text{ OR } i_q ;\ m > i_{2M_0} & \downarrow
\end{array} \right.$$

$$\Theta'^{FF}_{i_p i_q}(p,q) = \sum_{k,m} \phi_k(p)\phi_m(q)\langle km|M|i_p i_q\rangle$$

$$i_p = i_q \begin{cases} k > i_N\,;\ m > i_N \\ k > i_N\,;\ i_{2M_0} < m \le i_N \\ i_{2M_0} < k \le i_N\,;\ m > i_N \qquad \text{DOUBLY EXCITED} \\ \hline k > i_{2M_0}\,;\ m = i_p \qquad \text{SINGLY EXCITED} \\ k = i_p\,;\ m > i_{2M_0} \end{cases}$$

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

$$W = \frac{1}{H_0 - E_0}\sum_{p<q}(k_{qp} + k_{pq})|\bar{G}_1 \Phi_0 \chi\rangle = \sum_{p<q}(\bar{K}_{pq} + \bar{K}_{qp})|\bar{G}_1 \Phi_0 \chi\rangle\,,$$

where $\bar{K}_{pq} \hat{=} \frac{k_{pq}}{H_0 - E_0}$. The development proceeds as for Y except that since $\bar{K}_{pq}$ is a one-electron operator, the excited Φ' in Be-5 can differ from Φ_0 in at most one place. Φ' must differ from Φ_0 in at least one place since $\langle \bar{G}_1^{\alpha}\Phi_0\chi|\sum_{p<q}(\bar{K}_{pq}+\bar{K}_{qp})|G_1^{\alpha}\Phi_0\chi\rangle = 0$ (remember $\bar{G}_i^{\alpha}\Phi_0\chi = \delta_{i1}\bar{G}_1^{\alpha}\Phi_0\chi$). Instead of Be-6 we get

$W = C^{\alpha}_{M_0}\sum'_{p,q}\sum_{k}|G_1 \frac{\Phi_0}{\phi_{i_p}}\phi_k \chi\rangle\langle k q|\bar{K}_{pq}|i_p i_q\rangle$. We obtain

$W = C^{\alpha}_{M_0}\sum_{p<q}|G_1 \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}}\Theta''_{i_p i_q}\rangle\,,$ where

$$\Theta''_{i_p i_q}(p,q) \hat{=} \sum_{\substack{k \\ \text{ALLOWED}}}[\phi_k(p)\phi_{i_q}(q)\langle k i_q|\bar{K}_{pq}|i_p i_q\rangle + \phi_{i_p}(p)\phi_k(q)\langle i_p k|\bar{K}_{pq}|i_p i_q\rangle]$$

Analyzing as before

$$\Theta''^{HH}_{i_p i_q}(p,q) = \sum_{k,m}\phi_k(p)\phi_m(q)\langle km|\bar{K}|i_p i_q\rangle \qquad \text{(Be-11)}$$

$$k > i_N\,;\ m = i_q$$
$$k = i_p\,;\ m > i_N$$

$$\Theta''^{FH}_{i_p i_q}(p,q) = \sum_{k,m}\phi_k(p)\phi_m(q)\langle km|\bar{K}|i_p i_q\rangle \qquad \text{(Be-12)}$$

$$k = i_p\,;\ m > i_N$$
$$k > i_{2M_0}\,;\ m = i_q$$

$$\Theta''^{FF}_{i_p i_q}(p,q) = \sum_{k,m}\phi_k(p)\phi_m(q)\langle km|\bar{K}|i_p i_q\rangle \qquad \text{(Be-13)}$$

$$k = i_p\,;\ m > i_{2M_0}$$
$$k > i_{2M_0}\,;\ m = i_p$$

Next we need the relation between $\langle km|M|i_p i_q\rangle$ and $\langle km|\bar{K}|i_p i_q\rangle$.

If $k \neq i_p$,

$$\begin{aligned} \langle k\, i_q | M | i_p i_q \rangle &= - \langle k\, i_q | g | i_p i_q \rangle \\ \langle i_q k | M | i_p i_q \rangle &= - \langle i_q k | g | i_p i_q \rangle \\ \langle k\, i_q | \bar{K} | i_p i_q \rangle &= - \langle k\, i_q | g | i_p i_q \rangle - U'_{11(p,q)} \langle i_q k | g | i_p i_q \rangle \end{aligned} \qquad \text{(Be-14)}$$

Let $\Theta'''_{i_p i_q} \triangleq \Theta'_{i_p i_q} - \Theta''_{i_p i_q}$ then $X_1 = C^{\alpha}_{n_0} \sum_{p<q} | G_1 \frac{\Phi_0}{\phi_{i_p} \phi_{i_q}} \Theta'''_{i_p i_q}(p,q) \chi \rangle$

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

$$| G_1 \frac{\Phi_0}{\phi_{i_p} \phi_{i_q}} \phi_k \phi_{i_p} \chi \rangle = U_{11(p,q)} | G_1 \frac{\Phi_0}{\phi_{i_q}} \phi_k \chi \rangle \qquad \text{(Be-15)}$$

then the singly-excited q states of Θ'' would cancel those of Θ' leaving Θ''' with no singly-excited q states. $| G_1 \frac{\Phi_0}{\phi_{i_p} \phi_{i_q}} \phi_k \phi_{i_p} \chi \rangle = | G_1 (p,q) \frac{\Phi_0}{\phi_{i_q}} \phi_k \chi \rangle = = \sum_r \zeta_{\sigma_{r1}} (\sigma_{r1}(p,q) \frac{\Phi_0}{\phi_{i_q}} \phi_k)(\omega_{\bar{r}\bar{q}} \chi) = \sum_j \zeta_{\sigma_{1j}} U_{1j(p,q)} | G_j \frac{\Phi_0}{\phi_{i_q}} \phi_k \chi \rangle$. Let $\frac{\Phi_0}{\phi_{i_q}} \phi_k \triangleq \Phi_q$. From appendix D for $p,q \leq 2n_0$ and for $p,q > 2n_0$ $\sigma_{r1}(p,q) \Phi_q = U_{11(p,q)} \sigma_{r1} \Phi_q$. Thus, Be-15 is true for these cases and Θ'''^{HH} and Θ'''^{FF} have no singly-excited states. However, Θ'''^{FH} does have singly-excited states. Let Θ^{FH} be the part of Θ'''^{FH} containing only doubly-excited states and let $\Theta^{FF} \triangleq \Theta'''^{FF}$ and $\Theta^{HH} \triangleq \Theta'''^{HH}$. Then,

$$\begin{aligned} X_1 = C^{\alpha}_{n_0} \Big\{ & \sum_{p<q} | G_1 \frac{\Phi_0}{\phi_{i_p} \phi_{i_q}} \Theta_{i_p i_q} \rangle + \\ & + \sum_{\substack{p \leq 2n_0 \\ q > 2n_0}} \sum_{j \neq 1} \zeta_{\sigma_{1j}} U_{1j(p,q)} \Big[\sum_{k > i_N} | G_j \frac{\Phi_0}{\phi_{i_q}} \phi_k \chi \rangle \langle k \phi_{i_p} | M | i_p i_q \rangle + \\ & \qquad + \sum_{k > i_{2n_0}} | G_j \frac{\Phi_0}{\phi_{i_p}} \phi_k \chi \rangle \langle i_q k | M | i_p i_q \rangle \Big] \Big\} \end{aligned}$$

We are considering the case of $\alpha = [2^M, 1^{N-2M}]$ (maximum spin for the configuration); for this case $G^{\alpha}_i \Phi \chi = \delta_{i1} G^{\alpha}_1 \Phi \chi$. Consider the singly-excited states in the above equation; if ϕ_{i_q} is excited then

$G_j \frac{\Phi_0}{\phi_{i_q}} \phi_k \chi = 0$ if $j \neq 1$ for the same reason as above.

Thus*,

$$X_1 = C^{\alpha}_{M_0}\Big[\sum_{p<q}\Big|G_{\underline{1}}\frac{\Phi_0}{\phi_{1p}\phi_{1q}}\Theta_{1p1q}\Big\rangle + \sum_{\substack{p\le 2M_0\\ q>2M_0}}\sum_{j\neq 1} \zeta_{G_{1j}}\, U_{1j(p,q)}^{\times} \tag{Be-16}$$

$$\times \sum_{k>1_{2M_0}}\Big|G_{\underline{j}}\frac{\Phi_0}{\phi_{1p}}\phi_k\Big\rangle\langle 1_q k|M|1_p 1_q\rangle\Big]$$

if only closed shells are present then

$$\boxed{X_1 = C^{\alpha}_{M_0}\sum_{p<q}\Big|G_{\underline{1}}\frac{\Phi_0}{\phi_{1p}\phi_{1q}}\Theta_{1p1q}\Big\rangle} \tag{Be-17}$$

Using Be-7, 8, 9 and remembering that Θ 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 ϕ_1 is a singly-occupied orbital and associated with the spin α in the Slater determinant and if ϕ_2 is associated with spin β, then the excited state obtained by replacing $\phi_1\alpha$ with $\phi_1\beta$ and $\phi_2\beta$ with $\phi_2\alpha$ 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 $\alpha = [2^m, 1^{N-2m}]$ (maximum spin) and M = S the excited state is a mixture of S and S + 1). Thus, the Brillouin theorem is somewhat emptier 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. Bd-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.

$\langle \phi_{1_r}(p) | \Theta^{HH}_{1_p 1_q}(p,q) \rangle = \langle \phi_{1_r}(q) | \Theta^{HH}_{1_p 1_q}(p,q) \rangle = 0$, I.E., Θ^{HH} IS ORTHOGONAL TO ALL OCCUPIED STATES

$\langle \phi_{1_r}(p) | \Theta^{FH}_{1_p 1_q}(p,q) \rangle = \sum_{m > 1_N} \phi_m(q) \langle 1_r m | M | 1_p 1_q \rangle$ IF $2M_0 < r \leq N$

= 0 otherwise

$\langle \phi_{1_r}(q) | \Theta^{FH}_{1_p 1_q}(p,q) \rangle = \sum_{m > 1_N} \phi_m(p) \langle m 1_r | M | 1_p 1_q \rangle$ IF $2M_0 < r \leq N$

= 0 otherwise

$\langle \phi_{1_r}(q) | \Theta^{FF}_{1_p 1_q}(p,q) \rangle = \sum_{\substack{m > 1_{2M_0} \\ m \neq 1_r}} \phi_m(p) \langle 1_r m | M | 1_p 1_q \rangle$ IF $2M_0 < r \leq N$ AND $1_p \neq 1_q$

$= \sum_{m > 1_N} \phi_m(p) \langle 1_r m | M | 1_p 1_q \rangle$ IF $2M_0 < r \leq N$ AND $1_p = 1_q$

= 0 otherwise

$\langle \phi_{1_r}(p) \phi_{1_t}(q) | \Theta^{HH}_{1_p 1_q}(p,q) \rangle = 0$

$\langle \phi_{1_r}(p) \phi_{1_t}(q) | \Theta^{FH}_{1_p 1_q}(p,q) \rangle = 0$

$\langle \phi_{1_r}(p) \phi_{1_t}(q) | \Theta^{FF}_{1_p 1_q}(p,q) \rangle = \langle 1_r 1_t | M | 1_p 1_q \rangle$ IF $2M_0 < r, t \leq N$, $r \neq t$ AND $1_p \neq 1_q$

= 0 otherwise

If only closed shells are present, then

$\langle \phi_{1_r}(q) | \Theta^{FF}_{1_p 1_q}(p,q) \rangle = 0$ i.e., Θ^{FF} is orthogonal to all occupied states (Be-18)

$\langle \phi_{1_r} \phi_{1_t} | \Theta^{FF}_{1_p 1_q} \rangle = 0$

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 SINANOǦLU (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

$$\Theta_{1_p 1_q} = \sum_{k, m > 1_N} \phi_k \phi_m \langle k m | \frac{m_{pq}}{h'_p + h'_q - e_p - e_q} | 1_p 1_q \rangle, \qquad \text{(Be-19)}$$

where h'_p is the one-electron Hartree-Fock operator and e_i is one-electron Hartree-Fock energy. This is equivalent to (if the equation has a solution)

$$(h'_p + h'_q - e_p - e_q)\Theta_{i_p i_q}(p,q) = m_{pq}\,\phi_{i_p}(p)\,\phi_{i_q}(q)$$

$$\text{AND}\ \langle \phi_{i_r}(q)|\Theta_{i_p i_q}(p,q)\rangle = \langle \phi_{i_r}(p)|\Theta_{i_p i_q}(p,q)\rangle = 0 \qquad \text{(Be-20)}$$

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 inhomogenity is orthogonal to the solutions of the homogeneous equation. Since

$\langle i_p i_q|m_{ij}|i_p i_q\rangle = J_{pq} + U'_{ii(p,q)}K_{pq} - J_{pq} = U'_{ii(p,q)}K_{pq} \neq 0$ and

$\langle i_q i_p|m_{ij}|i_p i_q\rangle = \delta_{i_p i_q} J_{pq} - K_{pq} \neq 0$, then the inhomogeneous equation for $\Theta_{i_p i_p}$ has a solution, but for $\Theta_{i_p i_q}$ where $i_p \neq i_q$ it does not. The reason for wanting a differential equation for $\Theta_{i_p i_q}$ is that the expansions of the form Be-19 are slowly convergent with a significant contribution from the continuum (e.g., SHULL and LÖWDIN (1955)); whereas, for two-particle differential equations quite accurate results can be obtained. However, the $\Theta_{i_p i_p}$ terms are likely to be very much more important than the terms for $i_p \neq i_q$ 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 $\Theta_{i_p i_q}$, it probably would be worth solving only for $\Theta_{i_p i_p}$, 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 M > N-n basis functions, after diagonalizing the matrix we select the N-n ones with the lowest ϵ_{pp} as the Hartree-Fock orbitals the others are the left-over or virtual orbitals) can be used for the expansion of $\Theta_{i_p i_q}$ for $i_p \neq i_q$ (see LÖWDIN (1955b) p. 1504).

For systems with half-closed shells then similarly to the above case

$$(h_p' + h_q' - e_p - e_q)\theta_{i_p i_p}(p,q) = m_{pq}\,\phi_{i_p}(p)\phi_{i_q}(q) \quad , \text{ NOTE } i_p = i_q \qquad \text{(Be-21)}$$

where we require that $\langle \phi_{i_r}(q)|\theta_{i_p i_p}(p,q)\rangle = 0$ for all doubly-occupied i_r and $\langle \phi_{i_r}(p)\phi_{i_t}(q)|\theta_{i_p i_p}(p,q)\rangle = 0$ for all occupied states

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

$$E_2 = \langle \psi_0 | H_1 - E_1 | X_1 \rangle$$

$$E_3 = \langle X_1 | H_1 - E_1 | X_1 \rangle$$

Let X_1^S involve only singly-excited states and X_1^D involve only doubly-excited states (see Be-16). X_1 is orthogonal to ψ_0, so, $\langle \psi_0 | E_1 | X_1 \rangle = 0$. Also, $\langle \psi_0 | \sum_p V_p | X_1^D \rangle = 0$ since V_p is a one-electron operator.

$$E_2^D = \langle \psi_0 | \sum_{r<t} g_{rt} X_1^D \rangle = \frac{\theta^\alpha}{\beta^\alpha 2^{n_0} \langle \chi_{\bar q} | \omega_{\bar q \bar q} | \chi_{\bar q} \rangle} \sum_{p<q} \langle G_{\underline{1}} \Phi_0 \chi | \sum_{r<t} g_{rt} | G_{\underline{1}} \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}} \theta_{i_p i_q} \chi \rangle$$

$$= \frac{\theta^\alpha}{2^{n_0}} \sum_{p<q} \langle O_{11}^\alpha \Phi_0 | \sum_{r<t} g_{rt} | \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}} \theta_{i_p i_q} \rangle = \sum_{p<q} \sum_{r<t} \sum_{\substack{\tau \\ \mathrm{DIFF}}} U_{11\tau} \langle \tau \Phi_0 | g_{rt} | \frac{\Phi_0}{\phi_{i_p}\phi_{i_q}} \theta_{i_p i_q} \rangle$$

where τ goes over all τ giving different Φ_0.

$$E_2^{SV} = \langle \psi_0 | \sum_r V_r | X_1^S \rangle = \sum_{\substack{p \le 2n_0 \\ q > 2n_0}} \sum_r \sum_{\substack{\mathrm{DIFF} \\ \tau}} \sum_{j \ne 1} \zeta_{\sigma_{1j}} U_{j1\tau} \sum_{k > i_{2n_0}} \langle \tau \Phi_0 | V_r | \frac{\Phi_0}{\phi_{i_p}} \phi_k \rangle \langle i_q k | M | i_p i_q \rangle$$

if $k > i_N$, then get zero since τ must be ϵ and $U_{ji\epsilon} = 0$

if $i_{2n_0} < k \le i_N$, let $k = i_t$, then $r = p$ or $r = t$;

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

Thus,

$$E_2^{SV} = \sum_{\substack{p \le 2n_0 \\ q > 2n_0}} \sum_{j \ne 1} \zeta_{\sigma_{1j}} U_{1j(p,q)} U_{j1(p,t)} \sum_{t > 2n_0} \langle i_p | V_r | i_t \rangle \langle i_q i_t | M | i_p i_q \rangle$$

Also,

$$E_2^{Sg} = \langle \psi_0 | \sum_{r<t} g_{rt} | \chi_1^S \rangle =$$

$$= \sum_{\substack{p \le 2n_0 \\ q > 2n_0}} \frac{1}{2} \sum_{r,t}{}' \sum_{\substack{\text{DIFF} \\ \tau}} \sum_{j \ne 1} \zeta_{\sigma_{1j}} U_{1j(p,q)} U_{j1\tau} \sum_{k > i_{2n_0}} \langle \tau \Phi_0 | g_{rt} | \frac{\Phi_0}{\phi_{i_p}} \phi_k \rangle \langle i_q k | M | i_p i_q \rangle$$

if $k > i_N$ let $p = r$ then $\tau = \epsilon$ or (p,t);

if $k = i_m$ where $2n_0 < m \le N$ and $p = r$ then $\tau = \epsilon$, (p,t), (m,p), or $(m,p)(p,t)$

Thus,

$$E_2^{Sg} = \frac{1}{2} \sum_{\substack{p \le 2n_0 \\ q > 2n_0}} \sum_{j \ne 1} \zeta_{\sigma_{1j}} U_{1j(p,q)} \sum_t \Big[\sum_{k > i_N} \underset{p \ne t}{U_{j1(p,t)}} \langle i_p i_t | g | i_t k \rangle \langle i_q k | M | i_p i_q \rangle +$$

$$+ \sum_{2n_0 < m \le N} \Big(U_{j1(m,p)} \langle i_p i_t | g | i_m i_t \rangle \langle i_q i_m | M | i_p i_q \rangle +$$

$$+ \big(\underset{p \ne t}{U_{j1(p,t)}} + \underset{p \ne t}{U_{j1(m,p)(p,t)}} \big) \langle i_t i_p | g | i_m i_t \rangle \langle i_q i_m | M | i_p i_q \rangle \Big) \Big]$$

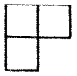
The analysis for E_3 is similar with more terms.

Bf: Construction from the $G_i^\alpha \Phi\chi$ of Basis Functions of Irreducible Representations of the Spatial Symmetry Group

The $G_i^\alpha \Phi\chi$ 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 $G_i^\alpha \Phi\chi$ 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 $\sum\limits_{i<j} \frac{e^2}{r_{ij}}$ 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_i^{α} 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 $O_{11}^{[1^3]}$ is just the determinant of space states.

Consider $S = \frac{1}{2}$ then $\alpha = [2,1] =$ . We wish the maximum L_z;

if $\Phi = \phi_2\phi_2\phi_2$, then $G_i^{\alpha}\Phi\chi = 0$ (THM Ab-1), but for $\Phi = \phi_2\phi_2\phi_1$ $G_1^{\alpha}\Phi\chi \neq 0$. Thus we will consider the L = 5 eigenstates. Recall that $G_1^{\alpha}\Phi\chi = \sum_r \zeta_{\sigma_{r1}}(O_{r1}^{\alpha}\Phi)(\omega_{\bar{r}\bar{1}}^{\bar{\alpha}}\chi)$. Since the $L^{\pm}$ operates only on the $O_{ri}\Phi$, we need only consider the $O_{ri}\Phi$ (remember that states of different M_s are obtained by simply using a different number of α and β in the ordered χ).

L = 5,

$L_z = 5$: $(55)_{r1} = O_{r1}^{[2,1]}\phi_2\phi_2\phi_1$

$L_z = 4$: $(54)_{r1} = \sqrt{\frac{1}{10}}\, O_{r1}^{[2,1]} L^{-}\phi_2\phi_2\phi_1 = \sqrt{\frac{1}{10}}\, O_{r1}^{[2,1]}(2\phi_2\phi_1\phi_1 + 2\phi_1\phi_2\phi_1 + \sqrt{6}\phi_2\phi_2\phi_0)$

but we want the standard order $\phi_1\phi_1\phi_2$, thus

$O_{r1}\,\phi_2\phi_1\phi_1 = O_{r1}(13)\,\phi_1\phi_1\phi_2 = U_{11(1,3)}O_{r1}\,\phi_1\phi_1\phi_2$

$(54)_{r1} = \sqrt{\frac{1}{10}}\, O_{r1}^{[2,1]}(2(U_{11(1,3)} + U_{11(2,3)})\phi_1\phi_1\phi_2 + \sqrt{6}\,\phi_2\phi_2\phi_0$

$U_{11(1,3)} = U_{11(2,3)} = -\frac{1}{2}$

$(54)_{r1} = \frac{-2}{\sqrt{10}}\, O_{r1}^{[2,1]}\phi_1\phi_1\phi_2 + \sqrt{\frac{6}{10}}\, O_{r1}^{[2,1]}\phi_2\phi_2\phi_0$

$L_z = 3$: $(53)_{r1} = \frac{1}{\sqrt{18}\sqrt{10}}\, O_{r1}^{[2,1]}[-2\sqrt{6}\phi_0\phi_1\phi_2 - 2\sqrt{6}\phi_1\phi_0\phi_2 + 2\sqrt{6}\phi_1\phi_2\phi_0 + 2\sqrt{6}\phi_2\phi_1\phi_0 + 6\phi_2\phi_2\phi_{\bar{1}}]$

$= \frac{1}{\sqrt{180}}[2\sqrt{6}(U_{11\epsilon} + U_{11(1,2)} - U_{11(1,3)} - U_{11(132)})\, O_{r1}\phi_2\phi_1\phi_0 +$

$+ 2\sqrt{6}(U_{12\epsilon} + U_{12(12)} - U_{12(13)} - U_{12(132)})\, O_{r2}\phi_2\phi_1\phi_0 + 6\, O_{r1}\phi_2\phi_2\phi_{\bar{1}}]$

$= \sqrt{\frac{6}{5}}\, O_{r1}\phi_2\phi_1\phi_0 + \sqrt{\frac{2}{5}}\, O_{r2}\phi_2\phi_1\phi_0 + \frac{1}{\sqrt{5}}\, O_{r1}\phi_2\phi_2\phi_{\bar{1}}$

(it should be noted that the norm of $O_{r1}\,\phi_k\phi_l\phi_m$ where $k = l \neq m$ is twice as much as it is if $k > l > m$; thus, the norm of (53) is the same as that of (54) and (55)) (see Ba-1)

$$L_z = 2: \quad (52)_{r1} = \frac{1}{\sqrt{24+\sqrt{5}}}\Big[\sqrt{6}\,O_{r1}(2\phi_1\phi_1\phi_0 + \sqrt{6}\phi_2\phi_0\phi_0 + \sqrt{6}\phi_2\phi_1\phi_{\bar{1}}) + \sqrt{2}\,O_{r2}(2\phi_1\phi_1\phi_0 + $$
$$+\sqrt{6}\phi_2\phi_0\phi_0 + \sqrt{6}\phi_2\phi_1\phi_{\bar{1}}) + O_{r1}(2\phi_1\phi_2\phi_{\bar{1}} + 2\phi_2\phi_1\phi_{\bar{1}} + 2\phi_2\phi_2\phi_{\bar{2}})\Big]$$
$$= \frac{1}{\sqrt{120}}\Big[O_{r1}(2\sqrt{6}\,\phi_1\phi_1\phi_0 + 6\,U_{11(1,3)}\phi_0\phi_0\phi_2 + \sqrt{2}\sqrt{6}\,U_{21(13)}\phi_0\phi_0\phi_2 + 2\phi_2\phi_2\phi_{\bar{2}}) +$$
$$+ O_{r1}(6\phi_2\phi_1\phi_{\bar{1}} + 2\phi_2\phi_1\phi_{\bar{1}} + 2\phi_2\phi_1\phi_{\bar{1}}) + O_{r2}(\sqrt{2}\sqrt{6}\,\phi_2\phi_1\phi_{\bar{1}})\Big]$$
$$= \frac{1}{\sqrt{30}}\Big[O_{r1}\phi_2\phi_2\phi_{\bar{2}} + \sqrt{6}\,O_{r1}\phi_1\phi_1\phi_0 - 3\,O_{r1}\phi_0\phi_0\phi_2 + 5\,O_{r1}\phi_2\phi_1\phi_{\bar{1}} + \sqrt{3}\,O_{r2}\phi_2\phi_1\phi_{\bar{1}}\Big]$$

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 $(L\ L_z)$ and the $(L\ L_z)_{ri}$ terms occurring in $G_i^{\alpha}\Phi\chi$. 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 SHORTLY (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^{α} 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 $G_f^{\alpha}\Phi\chi$ wave function is found for each spin and number of electrons by using the variational principle ($\Phi = \phi_{1a}\phi_{2a}\cdots\phi_{Ma}\phi_{1b}\cdots\phi_{mb}$, where the ϕ_{ia} are mutually orthonormal orbitals and the ϕ_{ib} 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 ϕ_{ia} (the ϕ_{ia} are eigenfunctions of the one-electron Hamiltonian H^a and the ϕ_{ib} are eigenfunctions of H^b). The ϕ_{ia} and ϕ_{ib} which lead to the minimum total energy are called the GF orbitals and the one-electron equations for the GF orbitals (e.g., $H^a\phi_{ia} = \epsilon_{ia}\phi_{ia}$) 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 H_2 (comparing with the Hartree-Fock and exact energies given in FRAGA 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 Cf.

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 α spin are varied independently of the orbitals of β 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 Bd-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 results.

Another approach which does not use the SODS restriction is the extended valence bond method* (SLATER (1963) p. 195 ff): A simple 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 $(A(1)B(2)+B(1)A(2))(\alpha(1)\beta(2)-\alpha(2)\beta(1))$ where A(1) is an atomic orbital on atom a and B(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 orthogonal 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^{α} such functions). With these assumptions a fairly simple energy expression 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 H_2O (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 $G_i^{\alpha}\Phi\chi$, where $\alpha = \left[\frac{N}{2} + S, \frac{N}{2} - S\right]$. Define $n \triangleq \frac{N}{2} + S$. We let Φ contain $n = \frac{N}{2} + S$ orthogonal orbitals ϕ_{ia} and $N-n \triangleq m$ orthogonal orbitals ϕ_{ib} . Let $S(ia;jb) \triangleq \int \phi_{ia}^{*}\phi_{jb}\,dx$ (for SODS $S(ia;jb) = \delta_{ij}$, for alternate orbitals $S(ia;jb) = \lambda\,\delta_{ij}$). We wish to determine the ϕ_{ia} and ϕ_{jb} such that the energy of the system is a minimum.

Let $\Phi = \phi_{1a}(1)\phi_{2a}(2)\cdots\phi_{ma}(m)\phi_{1b}(m+1)\phi_{2b}(m+2)\cdots\phi_{mb}(N)$ (Ca-1)

and $\chi = \alpha(1)\alpha(2)\cdots\alpha(N)\,\beta(m+1)\beta(m+2)\cdots\beta(N)$ (Ca-2)

(we are here considering $M_s = S$). Then by THM Ab-2 $G_i^{\alpha}\Phi\chi = \delta_{if}G_f^{\alpha}\Phi\chi$. We will consider $G_f^{\alpha}\Phi\chi$ in the following.*

$$S_f^{\alpha} = \begin{array}{|c|c|}
\hline
1 & m+1 \\
\hline
2 & m+2 \\
\hline
\cdot & \vdots \\
\cdot & \vdots \\
\cline{2-2}
\cdot & N \\
\hline
m \\
\cline{1-1}
\end{array}$$

The energy, $E = \dfrac{\langle G_f^{\alpha}\Phi\chi | H | G_f^{\alpha}\Phi\chi\rangle}{\langle G_f^{\alpha}\Phi\chi | G_f^{\alpha}\Phi\chi\rangle}$ (Ca-3)

where $H = \sum_{\lambda=1}^{N} h_{\lambda} + \frac{1}{2}\sum_{\lambda,\gamma=1}^{N}{}' g_{\lambda\gamma}$

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

The normalization integral:

From (Ac-3) $\langle G_f^{\alpha}\Phi\chi | G_f^{\alpha}\Phi\chi\rangle = f^{\alpha}\langle\Phi|O_{ff}^{\alpha}|\Phi\rangle\langle\chi|\omega_{\bar{f}\bar{f}}^{\bar{\alpha}}|\chi\rangle$

From p. Ba-4 $\langle\chi|\omega_{\bar{f}\bar{f}}|\chi\rangle = (\frac{N}{2}+S)!\,(\frac{N}{2}-S)!\,/\,\Theta^{\alpha}$

$\Theta^{\alpha}\langle\Phi|O_{ff}|\Phi\rangle = \sum_{t} U_{fft}\,\langle\Phi|t|\Phi\rangle$

Due to the orthogonality among the a states and among the b states, then either ϕ_{ia} is in the i position of $\Phi' = t\Phi$ or else it is in a position > n; similarly either ϕ_{ib} is in the n+i position of Φ' or

---

* But see section Ce.

else it is in a position $\leq n$. If τ takes no element $\leq n$ to a position $> n$, then $\langle\Phi|\tau|\Phi\rangle \neq 0$ only for $\tau = \epsilon$. If τ takes one element, say i, from a position $\leq m$ to a position $> n$, then it must take one element say j from a position $> n$ to a position $\leq n$; $\langle\Phi|\tau|\Phi\rangle \neq 0$ only for $\tau = (i,j)$. If τ 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 $\leq n$; $\langle\Phi|\tau|\Phi\rangle \neq 0$ only for $\tau = (i,p)(j,q)$; $(i,j)(i,p)(j,q)$; $(p,q)(i,p)(j,q)$; and $(i,j)(p,q)(i,p)(j,q) = (i,q)(j,p)$. If τ 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 $\leq n$; $\langle\Phi|\tau\Phi\rangle \neq 0$ only if $\tau = \tau_a\tau_b\tau_r$ where $\tau_r = (p_1,q_1)(p_2,q_2)\ldots(p_r,q_r)$ and τ_a is any permutation of $\{p_1, p_2, \ldots, p_r\}$ and τ_b is any permutation of $\{q_1, q_2, \ldots, q_r\}$ (the choice of the specific set of r disjoint transpositions for τ_r is arbitrary, the permutations τ_a and τ_b take τ_r through the $(r!)^2$ different permutations for which $\langle\Phi|\tau\Phi\rangle \neq 0$). Now we evaluate $U_{ff\tau_a\tau_b\tau_r}$. From (Aa-1), $U_{fs\tau_a} = \delta_{fs}\zeta_{\tau_a}$ and $U_{fs\tau_b} = \delta_{fs}\zeta_{\tau_b}$; thus, $U_{ff\tau_a\tau_b\tau_r} = \zeta_{\tau_a}\zeta_{\tau_b}U_{ff\tau_r}$, but from appendix H* $U_{ff\tau_r} = \frac{1}{\binom{m}{r}}$. Thus,

$$\Theta^\alpha\langle\Phi|O^\alpha_{ff}|\Phi\rangle = \sum_{r=1}^{m}\frac{1}{\binom{m}{r}}\sum_{p_1<p_2<\cdots<p_r}\sum_{q_1<q_2<\cdots<q_r}\sum_{\tau_a\tau_b}\zeta_{\tau_a}\zeta_{\tau_b}\langle\phi_{p_1a}\cdots\phi_{p_ra}\phi_{q_1b}\cdots\phi_{q_rb}|\tau_a\tau_b|\phi_{q_1b}\cdots\phi_{q_rb}\phi_{p_1a}\cdots\phi_{p_ra}\rangle$$

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

$$\langle\phi_{p_1a}\phi_{p_2a}\cdots\phi_{p_ra}|\tau_a|\phi_{q_1b}\phi_{q_2b}\cdots\phi_{q_rb}\rangle = S(p_1a;q_{\tau_a^{-1}(1)}b)\cdots S(p_ra;q_{\tau_a^{-1}(r)}b)$$

The sum $\sum_{\tau_a}\zeta_{\tau_a}$ times the above is recognized as the determinant of the r-dimensional matrix $S(p_ia;q_jb)$ which from (Aa-8) is written as

---

* You must note that the proof in appendix H is incomplete.

$r!\, O_{11} S(p_1 a; q_1 b) \cdots S(p_r a; q_r b)$. Thus,

$$\Theta^{\alpha}\langle \Phi | O_{\ell\ell} | \Phi \rangle = \sum_{r=0}^{m} \frac{(r!)^2}{\binom{M}{r}} \sum_{P_r, Q_r} \left[O_{11}^{[1^r]} S(p_1 a; q_1 b) \cdots S(p_r a; q_r b) \right] \times \\ \times \left[O_{11}^{[1^r]} S(q_1 b; p_1 a) \cdots S(q_r b; p_r a) \right] \tag{Ca-4}$$

A standard order of the p_i and of the q_i is presumed to have been defined. Since S(ia;jb) = S* (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

$$\rho(1,1') \hat{=} N \int (G_\ell \Phi \chi)^* (G_\ell \Phi \chi)(dx_1')(ds) = \ell^{\alpha} \langle \chi | \omega_{\ell\bar{\ell}} | \chi \rangle \sum_i (1,i) \int \Phi^* O_{\ell\ell} \Phi (dx_i')$$

Let $$\int \Phi^* O_{\ell\ell} \Phi (dx_i') \hat{=} \sum_j \phi_{k_i}^*(i') \phi_j(i) D_j^{k_i} \tag{Ca-5}$$

where k_i is the ith state in Φ ; I shall often omit the k. Using the same reasoning as in deriving the normalization integral

$$D_{ia}^{ia} = \sum_{r=0}^{m} \frac{(r!)^2}{\binom{M}{r}} \sum_{\substack{P_r, Q_r \\ i \notin P_r}} \left[O_{11}^{[1^r]} S(p_1 a; q_1 b) \cdots S(p_r a; q_r b) \right] \times \\ \times \left[O_{11}^{[1^r]} S(q_1 b; p_1 a) \cdots S(q_r b; p_r a) \right] \tag{Ca-6}$$

I will always write m 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}^{ia} this number is the smaller of m and n-1 since $i \notin P_r$. For D_{ja}^{ia}, $j \neq i$, LET $\tau = \tau_2 \tau_a \tau_b \tau_{r-1} \tau_1$ WHERE $\tau_1 = (ia, sb)$ AND $\tau_2 = (ia, ja)$

$$\begin{matrix} ia \\ ja \\ sb \end{matrix} \xrightarrow{\tau_1} \begin{matrix} sb \\ ja \\ ia \end{matrix} \xrightarrow{\tau_{r-1}} \begin{matrix} sb \\ ja \\ ia \end{matrix} \xrightarrow{\tau_a \tau_b} \begin{matrix} tb \\ ja \\ \mu a \end{matrix} \xrightarrow{\tau_2} \begin{matrix} ja \\ tb \\ \mu a \end{matrix}$$

$$D_{ja}^{ia} = - \sum_{r=1}^{m} \frac{(r!)^2}{\binom{M}{r}} \sum_{\substack{P_r, Q_r \\ i \in P_r \\ j \notin P_r}} \left[(ia, ja) O_{11}^{[1^r]} S(p_1 a; q_1 b) \cdots \right] \left[O_{11}^{[1^r]} S(q_1 b; p_1 a) \cdots \right] \tag{Ca-7}$$

where the -1 is due to $\zeta_{\tau_2} = -1$ and (ia, ja) transposes ia and ja. For D_{jb}^{ia} , $\tau = \tau_a \tau_b \tau_r \tau_1$ where $\tau_1 =$ (ia, jb)

$$\begin{matrix} ia \\ jb \end{matrix} \xrightarrow{\tau_1} \begin{matrix} jb \\ ia \end{matrix} \xrightarrow{\tau_r} \begin{matrix} jb \\ ia \end{matrix} \xrightarrow{\tau_a \tau_b} \begin{matrix} jb \\ sa \end{matrix}$$

Thus,

$$D^{ia}_{jb} = \sum_{r=0}^{m} \frac{r!(r+1)!}{\binom{M}{r+1}} \sum_{\substack{P_r, Q_r \\ i \notin P_r \\ j \notin Q_r}} \left[O_{11}^{[1^r]} S(pa;qb)\cdots\right]\left[O_{11}^{[1^{r+1}]} S(jb;ia)S(q b;p a)\cdots\right] \qquad \text{(Ca-8)}$$

where the $\binom{M}{r+1}$ 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^{i}_{j} is Hermitian.

The second-order density matrix: From section Bb

$$\pi(1,2;1',2') \triangleq N(N-1)\int(G_g \Phi X)^*(G_g \Phi X)(dx''_{12})(ds)$$

$$= g^{\alpha}\langle X|\omega_{gg}|X\rangle \sum_{i,j}{}' (1,i)(2,j)\int \Phi^* O_{gg} \Phi\, (dx''_{ij})$$

$$\int \Phi^* O_{gg} \Phi\,(dx''_{ij}) \triangleq \sum_{s,t} \phi^*_{k_i}(i)\,\phi^*_{k_j}(j)\,\phi_s(i')\,\phi_t(j')\,D^{k_i,k_j}_{s,t}$$

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

$$D^{ia,ja}_{ia,ja} = \sum_{r=0}^{m} \frac{(r!)^2}{\binom{M}{r}} \sum_{\substack{P_r, Q_r \\ i,j \notin P_r}} \left| O_{11}^{[1^r]} S(pa;qb)\cdots \right|^2$$

$$D^{ia,ja}_{ia,ta} = -\sum_{r=1}^{m} \frac{(r!)^2}{\binom{M}{r}} \sum_{\substack{P_r, Q_r \\ i,t \notin P_r \\ j \in P_r}} \left[(ja,ta)O_{11}^{[1^r]} S(pa;qb)\cdots\right]\left[O_{11}^{[1^r]} S(qb;pa)\cdots\right]$$

$$D^{ia,ja}_{ka,ta} = \sum_{r=2}^{m} \frac{(r!)^2}{\binom{M}{r}} \sum_{\substack{P_r, Q_r \\ k,t \notin P_r \\ i,j \in P_r}} \left[(ia,ka)(ja,ta)O_{11}^{[1^r]} S(pa;qb)\cdots\right]\left[O_{11}^{[1^r]} S(qb;pa)\cdots\right]$$

$$D^{ia,ja}_{ia,kb} = \sum_{r=0}^{m} \frac{r!(r+1)!}{\binom{M}{r+1}} \sum_{\substack{P_r, Q_r \\ i,j \notin P_r \\ k \in Q_r}} \left[O_{11}^{[1^r]} S(pa;qb)\cdots\right]\left[O_{11}^{[1^{r+1}]} S(kb;ja)S(qb;pa)\cdots\right]$$

$$D^{ia,ja}_{ka,tb} = -\sum_{r=0}^{m} \frac{r!(r+1)!}{\binom{M}{r+1}} \sum_{\substack{P_r, Q_r \\ j,k \notin P_r \\ i \in P_r \\ t \notin Q_r}} \left[(ka,ia)O_{11}^{[1^r]} S(pa;qb)\cdots\right]\left[O_{11}^{[1^{r+1}]} S(tb;ja)S(qb;pa)\cdots\right]$$

$$D^{ia,ja}_{kb,tb} = \sum_{r=0}^{m} \frac{r!(r+2)!}{\binom{M}{r+2}} \sum_{\substack{P_r, Q_r \\ i,j \notin P \\ k,t \notin Q}} \left[O_{11}^{[1^r]} S(pa;qb)\cdots\right]\left[O_{11}^{[1^{r+2}]} S(kb;ia)S(tb;ja)S(qb;pa)\cdots\right]$$

$$D^{ia,kb}_{ia,kb} = \sum_{r=0}^{m} \frac{(r!)^2}{\binom{m}{r}} \sum_{\substack{P_r,Q_r\\ i\notin P\\ k\notin Q}} \left| O_{11}^{[1^r]} S(p,a;q,b)\cdots \right|^2$$

$$D^{ia,kb}_{kb,ia} = \sum_{r=0}^{m} \frac{(r!)^2}{\binom{m}{r+1}} \sum_{\substack{P_r,Q_r\\ i\notin P\\ k\in Q}} \left| O_{11}^{[1^r]} S(p,a;q,b)\cdots \right|^2$$

$$D^{ia,jb}_{ia,kb} = -\sum_{r=1}^{m} \frac{(r!)^2}{\binom{m}{r}} \sum_{\substack{P_r,Q_r\\ i\notin P\\ k\in Q\\ j\notin Q}} \left[(kb,jb) O_{11}^{[1^r]} S(p,a;q,b)\cdots\right]\left[O_{11}^{[1^r]} S(q,b;p,a)\cdots\right]$$

$$D^{ia,jb}_{tb,ia} = -\sum_{r=1}^{m} \frac{(r!)^2}{\binom{m}{r+1}} \sum_{\substack{P_r,Q_r\\ i\notin P_r\\ j\in Q_r\\ t\notin Q_r}} \left[O_{11}^{[1^r]} S(p,a;q,b)\cdots\right]\left[(jb,tb) O_{11}^{[1^r]} S(q,b;p,a)\cdots\right]$$

where we used $\tau = \tau_2 \tau_a \tau_b \tau_r \tau_1$ and $\tau_2 =$ (jb,kb)

$$\begin{tikzcd}
\begin{matrix} ia \\ \\ jb \\ tb \end{matrix} \arrow[r, "\tau_1"] & \begin{matrix} tb \\ \\ jb \\ ia \end{matrix} \arrow[r, "\tau_r"] & \begin{matrix} ib \\ \\ sa \\ ia \end{matrix} \arrow[r, "\tau_a\tau_b"] & \begin{matrix} tb \\ \\ ua \\ ia \end{matrix} \arrow[r, "\tau_2"] & \begin{matrix} tb \\ \\ ia \\ ua \end{matrix}
\end{tikzcd}$$

$$D^{ia,kb}_{ja,tb} = \sum_{r=0}^{m} \frac{[(r+1)!]^2}{\binom{m}{r+1}} \sum_{\substack{P_r,Q_r\\ i,j\notin P\\ k,t\notin Q}} \left[O_{11}^{[1^{r+1}]} S(ja;kb) S(p,a;q,b)\cdots\right] \times$$
$$\times \left[O_{11}^{[1^{r+1}]} S(tb;ia) S(q,b;p,a)\cdots\right]$$

where we used $\tau = \tau_2 \tau_a \tau_b \tau_r \tau_1$

$$\begin{tikzcd}
\begin{matrix} ia \\ ja \\ kb \\ tb \end{matrix} \arrow[r, "\tau_1"] & \begin{matrix} kb \\ ja \\ ia \\ tb \end{matrix} \arrow[r, "\tau_r"] & \begin{matrix} kb \\ ja \\ ia \\ tb \end{matrix} \arrow[r, "\tau_a\tau_b"] & \begin{matrix} ub \\ ja \\ va \\ tb \end{matrix} \arrow[r, "\tau_2"] & \begin{matrix} ja \\ ub \\ tb \\ va \end{matrix}
\end{tikzcd}$$

$$D^{ia,sb}_{tb,ja} = \sum_{r=0}^{m} \frac{[(r+1)!]^2}{\binom{m}{r+2}} \sum_{\substack{P_r,Q_r\\ i,j\notin P\\ s,t\notin Q}} \left[O_{11}^{[1^{r+1}]} S(ja;sb) S(p,a;q,b)\cdots\right] \times$$
$$\times \left[O_{11}^{[1^{r+1}]} S(tb;ia) S(q,b;p,a)\cdots\right]$$

Note that $D^{is_1,js_2}_{ks_3,ts_4} = (D^{ks_3,ts_4}_{is_1,js_2})^* = D^{js_2,is_1}_{ts_4,ks_3}$ and

$D^{is_1,js_1}_{ks_3,ts_4} = -D^{js_1,is_1}_{ks_3,ts_4}$; using these relations all other D are obtained from the above ones.

Using the preceding work

$$\frac{\langle G_\gamma \Phi X|H|G_\gamma \Phi X\rangle}{f^\alpha \langle X|\omega_{\overline{\gamma\gamma}}|X\rangle} = \sum_{i,j=1}^{N} \langle \phi_i|h|\phi_j\rangle D^i_j + \frac{1}{2}\sum_{i,j}{}' \sum_{k,t}{}' \langle \phi_i \phi_j|g|\phi_k \phi_t\rangle D^{i,j}_{k,t} \qquad \text{(Ca-9)}$$

where the i, j, k, t denote the position of the state in Φ . But, $D^{ia,ja}_{k,t} = -D^{ia,ja}_{t,k}$; so, regardless of what we define the, so far undefined, $D^{ia,ia}_{k,t}$ to be, then $D^{ia,ia}_{k,t} = -D^{ia,ia}_{t,k}$. Also, $\langle \phi_i \phi_i|g|\phi_k \phi_t\rangle = \langle \phi_i \phi_i|g|\phi_t \phi_k\rangle$; thus, we can write

$$\frac{\langle G_\gamma \Phi X|H|G_\gamma \Phi X\rangle}{f^\alpha \langle X|\omega_{\overline{\gamma\gamma}}|X\rangle} = \sum_{i,j=1}^{N} \langle i|h|j\rangle D^i_j + \frac{1}{2}\sum_{i,j,k,t=1}^{N} \langle ij|g|kt\rangle D^{i,j}_{k,t} \qquad \text{(Ca-10)}$$

In the next section the density matrices derived here will be utilized in solving for the best DODS $G^\alpha_\Gamma \Phi X$ approximation to the ground state of an N-electron molecule.

Cb: The GF Equations

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

$$I = \Big[\theta^{\alpha} \langle \Phi | H O_{ff} | \Phi \rangle - E \theta^{\alpha} \langle \Phi | O_{ff} | \Phi \rangle - \sum_{i,j=1}^{n} \epsilon_{ij}^{a} \langle \phi_{ia} | \phi_{ja} \rangle + \\ - \sum_{i,j=1}^{m} \epsilon_{ij}^{b} \langle \phi_{ib} | \phi_{jb} \rangle \Big] \qquad \text{(Cb-1)}$$

Consider $\delta \langle \Phi | (H-E) O_{ff} | \Phi \rangle = \langle \delta \Phi | (H-E) O_{ff} | \Phi \rangle + \langle \Phi | (H-E) O_{ff} | \delta \Phi \rangle$

$$= \langle \delta \Phi | (H-E) O_{ff} | \Phi \rangle + \langle \delta \Phi | (H-E) O_{ff} | \Phi \rangle^{*}$$

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

Consider the variation of ϕ_{ia}

$$\delta(\text{OTHER TERM}) = \sum_{\substack{j=1 \\ j \neq i}}^{M} \left[\epsilon_{ij}^{a} \langle \delta \phi_{ia} | \phi_{ja} \rangle + \epsilon_{ji}^{a} \langle \phi_{ja} | \delta \phi_{ia} \rangle \right] + \epsilon_{ii} \, \delta \langle \phi_{ia} | \phi_{ia} \rangle$$

Since this is rea, by taking the complex conjugate we obtain $\epsilon_{ij}^{a} = \epsilon_{ji}^{a*}$, i.e., ϵ^{a} AND ϵ^{b} are Hermitian. Take some unitary transformation of the ϕ_{ia}, e.g., $\phi_{ia} = S_{ij}^{a} \phi'_{ja}$ then $\Phi = \sum_{\tau_a} c_{\tau_a} (\tau_a \Phi')$, where $\Phi = \phi'_1 \phi'_2 \cdots$

Also, $\tau_a \Phi = \sum_{\tau'_a} c_{\tau'_a} (\tau_a \tau'_a \Phi') = \sum_{\tau''_a} c_{\tau_a^{-1} \tau''_a} (\tau''_a \Phi')$. But, $\langle \Phi | \tau_a \Phi \rangle = \delta_{\tau_a, \epsilon}$ and $\langle \Phi' | \tau_a \Phi' \rangle = \delta_{\epsilon, \tau_a}$ (since the unitary transformation preserves the orthonormality of the ϕ_i). Thus,

$$\sum_{\tau'_a} c_{\tau'_a} c_{\tau_a^{-1} \tau'_a} = \delta_{\tau_a, \epsilon} \qquad \text{(Cb-2)}$$

Moreover, $O_{rf} \tau \Phi = U_{ff \tau} O_{rf} \Phi = \zeta_{\tau} O_{rf} \Phi$ if τ involves only elements $\leqslant$ n (or > n); so, $O_{rf} \Phi = \left(\sum_{\tau} c_{\tau} \zeta_{\tau} \right) O_{rf} \Phi'$. Hence,

$$\langle \Phi | (H-E) O_{ff} | \Phi \rangle = \langle O_{rf} \Phi | (H-E) | O_{rf} \Phi \rangle =$$

$$= \langle O_{rf} \Phi' | (H-E) | O_{rf} \Phi' \rangle \sum_{\tau_a, \tau'_a} c_{\tau_a} \zeta_{\tau_a} c_{\tau'_a} \zeta_{\tau'_a}$$

Recall that $\zeta_{\tau_a} \zeta_{\tau'_a} = \zeta_{\tau_a} \zeta_{\tau_a'^{-1}} = \zeta_{\tau_a \tau_a'^{-1}} = \zeta_{\tau''_a}$, where $\tau''_a \triangleq \tau_a \tau_a'^{-1}$;

thus, $\sum_{\tau_a, \tau'_a} c_{\tau_a} c_{\tau'_a} \zeta_{\tau_a} \zeta_{\tau'_a} = \sum_{\tau''_a} \zeta_{\tau''_a} \sum_{\tau_a} c_{\tau_a} c_{\tau_a''^{-1} \tau_a} = 1$ using Cb-2 .

---

* The θ^{α}, a constant, is used to simplify later formulae

Therefore, $\langle \Phi | (H-E)\, O_{ff} | \Phi \rangle = \langle \Phi' | (H-E)\, O_{ff} | \Phi' \rangle$. Thus,

$$I = \langle \Phi | (H-E)\, O_{ff} | \Phi \rangle - \sum_{p,q=1}^{m} \sum_{i,j=1}^{m} S^{a}_{ip} S^{a}_{jq} \epsilon^{a}_{ij} \langle \phi'_{pa} | \phi'_{qa} \rangle - \sum_{i,j=1}^{m} \epsilon^{b}_{ij} \langle \phi_{ib} | \phi_{jb} \rangle$$

Hence, the transformed equation is the same as the original except that ϵ_{ij} is replaced by $\bar{\epsilon}^{a}_{ij} \triangleq \sum_{p,q=1}^{M} S_{pi} S_{qj} \epsilon^{a}_{pq}$. Since ϵ^{a} is Hermitian and S^{a} is unitary, an S^{a} exists which diagonalizes ϵ^{a}. Similarly for ϵ^{b}. In the following we use the ϕ'_{i} defined by these transformations, thus

$$\boxed{I = \theta^{\alpha} \langle \Phi | (H-E) O_{ff} | \Phi \rangle - \sum_{i=1}^{N} \epsilon_{ii} \langle \phi_{i} | \phi_{i} \rangle} \qquad \text{(Cb-3)}$$

Using the same analysis as on p. Bd-2, we see that ϕ^{*}_{k} and ϕ_{k} can be varied independently; we vary ϕ^{*}_{k}. We are now interested in the coefficient of ϕ^{*}_{ka} in I.

$\theta^{\alpha} \langle \Phi | O_{ff} | \Phi \rangle \underline{(ka)} \triangleq$ [sum of terms in $\theta^{\alpha} \langle \Phi | O_{ff} | \Phi \rangle$ involving ka] =

$= \sum_{r=0}^{m} \frac{(r!)^2}{\binom{m}{r}} \sum_{\substack{P_r, Q_r \\ k \in P_r}} \left| O_{11}^{[1^r]} S(p_1 a; q_1 b) \cdots \right|^2$. Now relax the restriction that the p_i be ordered (i.e., $p_1 < p_2 < \cdots < p_r$); as a consequence we must divide by r!. Let $p_1 = k$ and multiply by r. Expand the determinant

$$r!\left[O_{11}^{[1^r]} S(ka; q_1 b) S(p_2 a; q_2 b) \cdots\right] = \sum_{\tau \in S_r} \delta_{\tau}\, S(ka; q_{\tau(1)} b) S(p_2 a; q_{\tau(2)} b) \cdots =$$

$$= \sum_{\tau(1)=1}^{r} \sum_{\tau' \in S_{r-1}} \delta_{(1,\tau(1))\tau'}\, S(ka; q_{\tau(1)} b) S(p_2 a; q_{(1,\tau(1))\tau'(2)} b) \cdots =$$

$$= (r-1)! \sum_{\mu = q_1}^{q_r} \delta_{(q_1,\mu)}\, S(ka; \mu b) \left[(q_1 b, \mu b)\, O_{11}^{[1^{r-1}]} S(p_2 a; q_2 b) \cdots \right]$$

In the second equation we set $\tau = (1, \tau(1))\tau'$, where τ' involves only elements 2 to r. Let

$$\theta^{\alpha} \langle \Phi | O_{ff} | \Phi \rangle \underline{(ka)} = \sum_{\mu=1}^{m} S(ka; \mu b)\, F(ka, \mu) \quad ,$$

where

$$F(ka;\mu) \triangleq \sum_{r=1}^{m} \frac{r!}{\binom{m}{r}} \sum_{\substack{P(r), Q(r) \\ k \notin P' \\ \mu \in Q}} \delta_{(q_1,\mu)} \left[(q_1 b, \mu b)\, O_{11}^{[1^{r-1}]} S(p_2 a; q_2 b) \cdots \right] \times$$
$$\times \left[O_{11}^{[1^r]} S(q_1 b; ka) S(q_2 b; p_2 a) \cdots \right]$$

P(r) denotes the unordered set which is ordered in P_r, and P'(r) has p_1 deleted. Similarly, $D^{ia}_{ia}\,\underline{(ka)} = \sum_{\mu=1}^{M} S(ka;\mu b)\,F^{ia}_{ia}(ka;\mu)$ if $i \neq k$,

where (for $i \neq k$)

$$F^{ia}_{ia}(ka;\mu) \triangleq \sum_{r=1}^{M} \frac{r!}{\binom{M}{r}} \sum_{\substack{P(r),Q(r)\\ i,k\notin P'\\ \mu\in Q}} \zeta_{(q_1,\mu)} \left[(q_1 b,\mu b)\, O_{11}^{[1^{r-1}]} S(p_2 a;q_2 b)\cdots\right] \times \left[O_{11}^{[1^{r}]} S(q_1 b;ka) S(q_2 b;p_2 a)\cdots\right]$$

Continuing in the same manner let $D^{i}_{j}\,\underline{(ka)} = \sum_{\mu=1}^{M} S(ka;\mu b)\,F^{i}_{j}(ka;\mu)$ if $i \neq ka$.

$$F^{ia}_{ja}(ka;\mu) \triangleq -\sum_{r=2}^{M} \frac{r!}{\binom{M}{r}} \sum_{\substack{P(r),Q(r)\\ i\in P'\\ j,k\notin P'\\ \mu\in Q}} \zeta_{(q_1,\mu)} \left[(q_1 b,\mu b)(ia,ja)\, O_{11}^{[1^{r-1}]} S(p_2 a;q_2 b)\cdots\right] \times \left[O_{11}^{[1^{r}]} S(q_1 b;ka) S(q_2 b;p_2 a)\cdots\right]$$

$$F^{ia}_{jb}(ka;\mu) = \sum_{r=1}^{M} \frac{(r+1)!}{\binom{M}{r+1}} \sum_{\substack{P'(r),Q(r)\\ i,k\notin P'\\ j\notin Q\\ \mu\in Q}} \zeta_{(q_1,\mu)} \left[(q_1 b;\mu b)\, O_{11}^{[1^{r-1}]} S(p_2 a;q_2 b)\cdots\right] \times \left[O_{11}^{[1^{r+1}]} S(jb;ia) S(q_1 b;ka) S(q_2 b;p_2 a)\cdots\right]$$

Defining $F^{i,j}_{s,t}$ in a similar way we obtain $D^{i,j}_{s,t}\,\underline{(ka)} = \sum_{\mu=1}^{M} S(ka;\mu b)\,F^{i,j}_{s,t}(ka;\mu)$ if $i,j \neq ka$. Thus, the coefficient of $\delta\phi^{*}_{ka}$ in $\delta\Phi$ is (after multiplying by $\delta\phi^{*}_{ka}$)

$$\begin{aligned} \delta I\,\underline{(ka)} = \Big\{ &\sum_{j=1}^{N} \langle \delta\phi_{ka}|h|\phi_j\rangle D^{ka}_{j} + \sum_{j,s,t=1}^{N} \langle \delta\phi_{ka}\,\phi_j|g|\phi_s\phi_t\rangle D^{ka,j}_{s,t} + \\ &+ \sum_{\mu=1}^{M} \langle \delta\phi_{ka}|\phi_{\mu b}\rangle \Big[\sum_{\substack{i,j=1\\ i\neq ka}}^{N} \langle \phi_i|h|\phi_j\rangle F^{i}_{j}(ka;\mu) + \frac{1}{2} \sum_{\substack{i,j,s,t=1\\ i,j\neq ka}}^{N} \langle \phi_i\phi_j|g|\phi_s\phi_t\rangle F^{i,j}_{s,t}(ka;\mu) + \\ &- E F(ka;\mu)\Big] - \epsilon^{a}_{kk} \langle \delta\phi_{ka}|\phi_{ka}\rangle \Big\} = 0 \end{aligned}$$

(Cb-4)

The differential-integral equation which is the coefficient of $\delta\phi^{*}_{ka}$ must be equal to zero since the $\delta\phi^{*}_{ka}$ 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 $\delta\phi^*_{ka}$ in Cb-4 the GF equation for ϕ_{ka} since it results from using $G_f^{\alpha}\Phi$ as the wave function. We will write the GF equation as

$$H(ka)|\phi_{ka}\rangle = \epsilon^{a}_{kk}|\phi_{ka}\rangle \qquad \text{(Cb-5)}$$

where* $H(ka) = \Big\{ \sum_{j=1}^{N} D_j^{ka}\, h\, P_{j,ka} + \sum_{j,s,t=1}^{N} \langle\phi_j|g|\phi_t\rangle D_{s,t}^{ka,j} P_{s,ka} +$

$$+ \sum_{\mu=1}^{m} \Big[\sum_{\substack{i,j=1\\ i\neq ka}}^{N} \langle\phi_i|h|\phi_j\rangle F_j^{i}(ka;\mu) + \sum_{\substack{i,j,s,t=1\\ i,j\neq ka}}^{N} \langle\phi_i\phi_j|g|\phi_s\phi_t\rangle F_{s,t}^{i,j}(ka;\mu) +$$

$$- EF(ka;\mu)\Big] P_{\mu b,ka} \Big\}$$

and $P_{ka,j}$ changes ϕ_{ka} into ϕ_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 ϕ_i in terms of a fixed set of $M > n$ basis functions, χ_i. Then in the variation of ϕ_i the coefficients of the expansion are varied; ROOTHAAN (1951) described this approach for the SODS closed-shell approximation of molecules. If the basis set χ_i 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 $\phi_{ia} = \sum_{\mu} C^{a}_{\mu i}\chi_{\mu}$ and $\phi_{ib} = \sum_{\mu} C^{b}_{\mu i}\chi_{\mu}$. Let

$$\langle\chi_\mu|h|\chi_\nu\rangle \hat{=} h_{\mu\nu}\, ; \langle\chi_\mu\chi_\nu|g|\chi_\sigma\chi_\lambda\rangle \hat{=} g_{\mu\nu;\sigma\lambda}\, ; \langle\chi_\mu|\chi_\nu\rangle \hat{=} S_{\mu\nu} \qquad \text{(Cb-6)}$$

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

$$\sum h_{\mu\nu} C_{\nu j} D_j^{ka} + \sum_{j,s,t} g_{\mu\nu;\sigma\lambda} C^{*}_{\nu j} C_{\sigma s} C_{\lambda t} D_{s,t}^{ka,j} + \qquad \text{(Cb-7)}$$

$$+ \sum_{\mu=1}^{m} S_{\mu\nu} C^{b}_{\nu\mu} \Big[\sum_{i,j} \langle i|h|j\rangle F_j^{i}(ka;\mu) + \frac{1}{2}\sum_{ijst} \langle ij|g|st\rangle F_{st}^{ij}(ka;\mu) - EF(ka;\mu)\Big] =$$

$$= \epsilon^{a}_{kk} S_{\mu\nu} C^{a}_{\nu k}$$

Every term on the left side contains a factor of $C^{a}_{\nu k}$, thus we can

---

* The E appearing here is just $\langle\Phi|H O_{ff}|\Phi\rangle / \langle\Phi|O_{ff}|\Phi\rangle$ 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 Cb-7 as

$$\boxed{H_{\mu\nu}(a)\,C^{a}_{\nu k} = \epsilon^{a}_{kk}\,S_{\mu\nu}\,C^{a}_{\nu k}} \qquad \text{(Cb-8)}$$

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

Let

$$f(ka;\mu) \triangleq \sum_{r=1}^{m} \frac{r!}{\binom{M}{r}} \sum_{\substack{P'(r),Q(r)\\ k\in P'\\ \mu\in Q}} \zeta^{p}_{(q_1,\mu)} \left[(q_1b,\mu b)\,O_{11}^{[1^{r-1}]}\,S(p_2a;q_2b)\cdots\right]\times \times\left[O_{11}^{[1^{r}]}\,S(q_1b;ka)\,S(q_2b;p_2a)\cdots\right]$$

= 0 since the second determinant has the same column twice

Then $F(a;\mu) \triangleq F(ka;\mu) + f(ka;\mu)$ is independent of k. Thus, $\underline{\langle\Phi|O_{ff}|\Phi\rangle(ka)} = \sum_{\mu=1}^{m} S(ka;\mu b)\,F(a;\mu)$. Also, $f^{\lambda}_{j}(ka;\, u)$ and $f^{i,j}_{s,t}(ka;\mu)$ are defined similarly and are zero for the same reason. Thus, $F^{i}_{j}(ka;\mu)$, $F^{ij}_{st}(ka;\mu)$, and $F(ka;\mu)$ in Cb-4 and Cb-5 can be replaced with $F^{i}_{j}(a;\mu)$, $F^{ij}_{st}(a;\mu)$, and $F(a;\mu)$. Let

$$D^{ka}_{ja} \triangleq \sum_{\nu=1}^{m} S(\nu b;ka)\,F^{ka}_{ja}(ka;\nu) \qquad \text{, where}$$

$$F^{ka}_{ja}(ka;\nu) \triangleq -\sum_{r=1}^{m} \frac{r!}{\binom{M}{r}} \sum_{\substack{P'(r),Q(r)\\ j,k\notin P'\\ \nu\in Q}} \zeta^{p}_{(q_1,\nu)} \left[O_{11}^{[1^{r}]}\,S(ja;q_1b)\,S(p_2a;q_2b)\cdots\right]\times \times\left[(q_1b,\mu b)\,O_{11}^{[1^{r-1}]}\,S(q_2b,p_2a)\cdots\right]$$

$$f^{ka}_{ja}(ka;\nu) \triangleq -\sum_{r=1}^{m} \frac{r!}{\binom{M}{r}} \sum_{\substack{P'(r),Q(r)\\ j\notin P'\\ k\in P'\\ \nu\in Q}} \zeta^{p}_{(q_1,\mu)} \left[O_{11}^{[1^{r}]}\,S(ja;q_1b)\,S(p_2a;q_2b)\cdots\right]\times \times\left[(q_1b,\mu b)\,O_{11}^{[1^{r-1}]}\,S(q_2b;p_2a)\cdots\right]$$

Then $\sum_{\nu} S(\nu b;ka)\,f^{ka}_{ja}(ka;\nu) = 0$ since this corresponds to the det S(ib;ja) in D^{ka}_{ja} having the ka column twice; thus, $D^{ka}_{ja} = \sum_{v=1}^{m} S(vb;ka)\, F^{a}_{ja}(a;v)$ where $F^{a}_{ja}(a;v) \triangleq F^{ka}_{ja}(ka;v) + f^{ka}_{ja}(ka;v)$. The other D^{ka}_{j} and $D^{ka,i}_{j,s}$

with j,s≠ka are treated similarly. First consider,

$$\sum_{j} \langle \delta\phi_{ka} | h | \phi_{ja} \rangle D^{ka}_{ja} = \langle \delta\phi_{ka} | h | \phi_{ka} \rangle D^{ka}_{ka} + \sum_{j \neq ka} \langle \delta\phi_{ka} | h | \phi_{j} \rangle \sum_{v} S(vb;ka)\, F^{a}_{ja}(a;v)$$

$$= \langle \delta\phi_{ka} | h | \phi_{ka} \rangle \Big[D^{ka}_{ka} - \sum_{v} S(vb;ka)\, F^{a}_{ka}(a;v) \Big] + \sum_{j} \langle \delta\phi_{ka} | h | \phi_{ja} \rangle \sum_{v} S(vb;ka)\, F^{a}_{ja}(a;v)$$

But

$$D^{ka}_{ka} - \sum_{v} S(vb;ka)\, F^{a}_{ka}(a;v) = \Big\{ \sum_{r=0}^{m} \frac{r!}{\binom{m}{r}} \sum_{\substack{P(r),Q(r) \\ k \notin P}} \Big| O_{11}^{[1^r]} S(p_1a;q_1b) \cdots \Big|^2 + $$

$$+ \sum_{r=1}^{m} \frac{r!}{\binom{m}{r}} \sum_{P(r),Q(r)} \Big| O_{11}^{[1^r]} S(ka;q_1b)\, S(p_2a;q_2b) \cdots \Big|^2 \Big\}$$

$$= \sum_{r=0}^{m} \frac{r!}{\binom{m}{r}} \sum_{P(r),Q(r)} \Big| O_{11}^{[1^r]} S(p_1a;q_1b) \cdots \Big|^2 \overset{k \notin P'}{=} \Theta^{\alpha} \langle \Phi | O_{ff} | \Phi \rangle$$

Similarly,

$$\sum_{j,s,t} \langle \delta\phi_{ka}\, \phi_j | g | \phi_s \phi_t \rangle D^{ka,j}_{s,t} = \Big\{ \sum_{j}\sum_{t \neq ka} \Big[\langle \delta\phi_{ka}\, \phi_j | g | \phi_{ka} \phi_t \rangle D^{ka,j}_{ka,t} + \langle \delta\phi_{ka}\, \phi_j | g | \phi_t \phi_{ka} \rangle D^{ka,j}_{t,ka} \Big] +$$

$$+ \sum_{j} \sum_{\substack{t,s \\ \neq ka}} \langle \delta\phi_{ka}\, \phi_j | g | \phi_s \phi_t \rangle D^{ka,j}_{s,t} \Big\}$$

$$= \sum_{j,t,s} \langle \delta\phi_{ka}\, \phi_j | g | \phi_s \phi_t \rangle \sum_{v} S(vb;ka)\, F^{a,j}_{s,t}(a;v) +$$

$$+ \sum_{j,t} \Big\{ \langle \delta\phi_{ka}\, \phi_j | g | \phi_{ka} \phi_t \rangle \Big[D^{ka,j}_{ka,t} - \sum_{v} S(vb;ka)\, F^{a,j}_{ka,t}(a;v) \Big] +$$

$$+ \langle \delta\phi_{ka}\, \phi_j | g | \phi_t \phi_{ka} \rangle \Big[D^{ka,j}_{t,ka} - \sum_{v} S(vb;ka)\, F^{a,j}_{t,ka}(a;v) \Big] \Big\}$$

But,

$$D^{ka,ja}_{ka,ta} - \sum_{v} S(vb;ka)\, F^{a,ja}_{ka,ta}(a;v) = \Big\{ -\sum_{r=1}^{m} \frac{r!}{\binom{m}{r}} \sum_{\substack{P(r),Q(r) \\ k,t \notin P \\ j \in P}} \Big[(ja,ta)\, O_{11}^{[1^r]} S(p_1a;q_1b) \cdots \Big] \times$$

$$\times \Big[O_{11}^{[1^r]} S(q_1b;p_1a) \cdots \Big] +$$

$$- \sum_{r=2}^{m} \frac{r!}{\binom{m}{r}} \sum_{\substack{P(r),Q(r) \\ k,t \notin P \\ j \in P}} \Big[(ja,ta)\, O_{11}^{[1^r]} S(ka;q_1b)\, S(p_2a;q_2b) \cdots \Big] \Big[O_{11}^{[1^r]} S(q_1b;ka)\, S(q_2b;p_2a) \cdots \Big] \Big\}$$

$$= D^{ja}_{ta}$$

Similarly, for the others we obtain

$$D^{ka,j}_{ka,t} - \sum_{v} S(vb;ka)\, F^{a,j}_{ka,t}(a;v) = D^{j}_{t} \qquad \text{(Cb-9)}$$

Also,

$$D^{ka,jb}_{tb,ka} - \sum_{v} S(vb;ka)F^{a,jb}_{tb,ka}(a;v) = \Big\{ + \sum_{r=1}^{m} \frac{r!}{\binom{M}{r+1}} \sum_{\substack{P(r),Q(r)\\ k\notin P\\ t\in Q\\ j\notin Q}} \left[(tb,jb)O_{11}^{[1^r]}S(p,a;q,b)\cdots\right]\times \left[O_{11}^{[1^r]}S(q,b;p,a)\cdots\right] +$$

$$+\sum_{r=0}^{m}\frac{(r+1)!(r+1)}{\binom{M}{r+2}}\sum_{\substack{P(r),Q(r)\\ k\notin P\\ j,t\notin Q}}\left[O_1^{[1^{r+1}]}S(ka;jb)S(p,a;q,b)\cdots\right]\left[O_1^{[1^{r+1}]}S(tb;ka)S(q\,b;p,a)\cdots\right]\Big\}$$

The second term $= + \sum_{r=1}^{m}\frac{r!\,r}{\binom{M}{r+1}}\sum_{\substack{P(r-1),Q(r-1)\\ k\notin P\\ j,t\notin Q}}\left[O_1^{[1^r]}S(ka;jb)S(p,a;q,b)\cdots\right]\times\left[O_{11}^{[1^r]}S(tb;ka)S(q\,b;p,a)\cdots\right]$

(LET $r \to r-1$)

$$= + \sum_{r=1}^{m}\frac{r!}{\binom{M}{r+1}}\sum_{\substack{P(r),Q(r)\\ k\in P\\ t\in Q\\ j\notin Q}}\left[(tb,jb)O_{11}^{[1^r]}S(p,a;q,b)\cdots\right]\left[O_{11}^{[1^r]}S(q\,b;p,a)\cdots\right]$$

Thus, $D^{ka,jb}_{tb,ka} - \sum_{v} S(vb;ka)F^{a,jb}_{tb,ka}(a;v) = -D^{jb}_{tb}$ Moreover, for the other cases where jb or tb is replaced by a term with a, this result follows immediately from Cb-9 since $D^{ka,j}_{ta,ka} = -D^{ka,j}_{ka,t}$. Hence,

$D^{ka,j}_{t,ka} - \sum_{v} S(vb;ka)F^{a,j}_{t,ka}(a;v) = -D^{j}_{t}$. Now consider the coefficient of $\langle\delta\phi_{ka}|\phi_{ub}\rangle$ in Cb-4. Since $F^{ka}_{j}(ka;u) = 0$ and $F^{ka,j}_{s,t}(ka;u) = 0$, then as before, we define f^{i}_{j} (ka;u) (which is zero) and F^{i}_{j} (a;u) $\hat{=}$ $= F^{i}_{j}$ (ka;u) + (ka;u). Thus, Cb-4 becomes

$$\Big\{\langle\delta\phi_{ka}|h|\phi_{ka}\rangle\theta^{\alpha}\langle\Phi|O_{ff}|\Phi\rangle + \sum_{j=1}^{N}\langle\delta\phi_{ka}|h|\phi_j\rangle\sum_{v=1}^{m}S(vb;ka)F^{a}_{j}(a;v) +$$

$$+\sum_{j,t}\left[\langle\delta\phi_{ka}\phi_j|g|\phi_{ka}\phi_t\rangle D^{j}_{t} - \langle\delta\phi_{ka}\phi_j|g|\phi_t\phi_{ka}\rangle D^{j}_{t}\right] +$$

$$+\sum_{j,s,t}\langle\delta\phi_{ka}\phi_j|g|\phi_s\phi_t\rangle\sum_{v=1}^{m}S(vb;ka)F^{a,j}_{s,t}(a;v) +$$

$$+\sum_{u=1}^{m}\langle\delta\phi_{ka}|\phi_{ub}\rangle\left[\sum_{i,j=1}^{N}\langle i|h|j\rangle F^{i}_{j}(a;u) + \frac{1}{2}\sum_{i,j,s,t=1}^{N}\langle ij|g|st\rangle F^{i,j}_{s,t}(a;u) - EF(a;u)\right]\Big\} =$$

$$= \epsilon^{a}_{kk}\langle\delta\phi_{ka}|\phi_{ka}\rangle$$

The coeff of $\langle \delta\phi_{ka} | \phi_{ub} \rangle = \sum_{i} \Big[\langle i|h|ka\rangle F^{i}_{ka}(ka;u) + \sum_{j \neq ka} \langle i|h|j\rangle F^{i}_{j}(ka;u) +$

$$+ \sum_{j,t} \langle ij|g|ka,t\rangle F^{i,j}_{ka,t}(ka;u) + \sum_{\substack{j,s,t \\ s,t \neq ka}} \langle ij|g|st\rangle F^{i,j}_{s,t}(ka;u) - EF(ka;u) \Big]$$

Here $F^{i}_{ka}(ka;u) = 0$ and $F^{i,j}_{ka,t} = 0$ as before. We let

$F^{ia}_{ja}(ka;u) = \sum_{v} S(vb;ka) F^{ia}_{ja}(ka;u,v)$, where

$$F^{ia}_{ja}(ka;u,v) \triangleq - \sum_{r=2}^{m} \frac{r!}{\binom{m}{r}} \sum_{\substack{P'(r),Q(r) \\ i \in P' \\ j,k \notin P' \\ u,v \in Q}} f_{(g_1,u)} f_{(g_1,v)} \Big[(g_1 b; ub) O_{11}^{[1^{r-1}]} S(p_2 a; g_2 b) \cdots \Big] \times$$
$$\times \Big[(g_1 b, vb) O_{11}^{[1^{r-1}]} S(g_2 b; p_2 a) \cdots \Big]$$

Next define $f^{ia}_{ka}(ka;u,v)$ which is zero, and $F^{ia}_{ja}(a;u,v)$. Do the same for the other F^{i}_{j} and $F^{i,j}_{s,t}$. Thus, Cb-4 becomes

$$\Big\{ \langle \delta\phi_{ka}|h|\phi_{ka}\rangle \theta^{\alpha} \langle \Phi|O_{ff}|\Phi\rangle + \sum_{j,t} \langle \delta\phi_{ka}\phi_j|g|\phi_{ka}\phi_t - \phi_t\phi_{ka}\rangle D^{j}_{t} + \tag{Cb-10}$$
$$+ \sum_{j=1}^{N} \sum_{v=1}^{m} \Big[\langle \delta\phi_{ka}|h|\phi_j\rangle S(vb;ka) F^{a}_{j}(a;v) + \sum_{s,t} \langle \delta\phi_{ka}\phi_s|g|\phi_j\phi_t\rangle S(vb;ka) F^{a,s}_{j,t}(a;v) \Big] +$$
$$+ \sum_{u,v=1}^{m} \langle \delta\phi_{ka}|\phi_{ub}\rangle S(vb;ka) \Big[\sum_{i,j=1}^{N} \langle i|h|j\rangle F^{i}_{j}(a;u,v) + \frac{1}{2} \sum_{i,j,s,t=1}^{N} \langle ij|g|st\rangle F^{i,j}_{s,t}(a;u,v) +$$
$$- EF(a;u,v) \Big] \Big\} = \epsilon^{a}_{kk} \langle \delta\phi_{ka}|\phi_{ka}\rangle$$

Hence, (see Cb-8)

$$\boxed{\begin{aligned} H^{a}_{\mu\nu} = \Big\{ & h_{\mu\nu} \theta^{\alpha} \langle \Phi|O_{ff}|\Phi\rangle + \sum_{\sigma,\lambda} (g_{\mu\sigma;\nu\lambda} - g_{\mu\sigma;\lambda\nu}) \sum_{j,t=1}^{N} C^{*}_{\sigma j} C_{\lambda t} D^{j}_{t} + \\ & + \sum_{\sigma,\lambda} \Big[h_{\mu\lambda} S_{\sigma\nu} \sum_{j=1}^{N} \sum_{v=1}^{m} C_{\lambda j} C^{*}_{\sigma v} F^{a}_{j}(a;v) + \\ & \quad + \sum_{\eta\xi} g_{\mu\eta;\lambda\xi} S_{\sigma\nu} \sum_{j,s,t}^{N} \sum_{v}^{m} C^{*}_{\eta s} C_{\lambda j} C_{\xi t} C^{*}_{\sigma v} F^{a,s}_{j,t}(a;v) \Big] + \\ & + \sum_{\sigma,\lambda} S_{\mu\lambda} S_{\sigma\nu} \sum_{u,v}^{m} C_{\lambda u} C^{*}_{\sigma v} \Big[\sum_{i,j} \langle i|h|j\rangle F^{i}_{j}(a;u,v) + \frac{1}{2} \sum_{ijst} \langle ij|g|s,t\rangle F^{ij}_{st}(a;u,v) + \\ & - EF(a;u,v) \Big] \Big\} \end{aligned}} \tag{Cb-11}$$

And we see that $H^{a}_{\mu\nu}$ is, indeed, independent of k. Thus, the n ϕ_{ia} are eigenfunctions of $H^{a}_{\mu\nu}$ and the m ϕ_{ib} are eigenfunctions of $H^{b}_{\mu\nu}$ ($H^{b}_{\mu\nu}$ is obtained from Cb-11 by letting a ↔ b and the limits on v and u become 1 to n). $H^a_{\mu\nu}$ is the GF Hamiltonian and the ϕ_{ia} are the GF orbitals.

Notice that by using the Roothaan procedure we avoid the coupling of equations present in Cb-5 because of exchange and nonorthogonality. Since H is Hermitian and independent of k, then the ϕ_{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 ϕ_{ia} and ϕ_{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 ϕ_{ia} and ϕ_{ib}. It probably will result that the ϕ_{ia} selected will usually be those with the lowest ϵ_{ii} as apparently* occurs in the SODS HF approach.

---

* I should mention that although everyone takes it for granted in the usual SODS HFR approach (which is only applicable for closed shell systems, let N = 2n) that the n states of lowest ϵ_{kk} are the occupied ones (e.g., ROOTHAAN (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)

$$E' = 2\Big[\sum_{i=1}^{n} \epsilon_{ii} - \frac{1}{2}\sum_{i,j=1}^{n}(J_{ij} - \frac{1}{2}K_{ij})\Big] = 2\Big[\sum_{i=1}^{n}\langle i|h|i\rangle + \frac{1}{2}\sum_{i,j=1}^{n}(J_{ij} - \frac{1}{2}K_{ij})\Big]$$

Thus, the criterion for choosing the occupied states among the eigenstates of the HFR Hamiltonian should be: The n states of lowest $e_i = \Big[\langle i|h|i\rangle + \frac{1}{2}\sum_{j=1}^{n}(J_{ij} - \frac{1}{2}K_{ij})\Big]$ are occupied. We see then that if i = n is one of the n states of lowest ϵ_{ii} and if p is some state with higher ϵ_{pp} then $e_n > e_p$ and state p is occupied and state n is empty if

$$\frac{1}{2}\sum_{j=1}^{n}(J_{pj} - J_{nj} - \frac{1}{2}K_{pj} + \frac{1}{2}K_{nj}) > \epsilon_{pp} - \epsilon_{nn}$$

. This is true only when p is more favorable then 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 GF 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 ϕ_{ia} and ϕ_{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 <u>one</u> set of GFR equations. For the various compounds with the same total number of electrons and the same total spin we merely substitute different values for $\langle\chi_\mu|h|\chi_\nu\rangle$ and $\langle\chi_\mu\chi_\nu|g|\chi_\sigma\chi_\lambda\rangle$.

My reasoning for expecting that the GF orbitals are localized:

Consider first the situation where there are no interactions between the electrons. The Hamiltonian $H = \sum_i h_i$ 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, $H\Phi = E\Phi$, $\Phi = \phi_1(1)\phi_2(2)\cdots\phi_N(N)$, $h_\lambda\phi_\lambda = \epsilon_\lambda\phi_\lambda$, $E = \sum_\lambda \epsilon_\lambda$. Since H is invariant under permutations, we can write $HG_i^{\alpha}\Phi\chi = EG_i^{\alpha}\Phi\chi$ just as before. The point is that since the ϕ_i are all eigenfunctions of the Hermitian operator 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 j is essentially the same as the potential due to an electron at any of the positions for which ϕ_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 $(m_{pq}-k_{pq}-k_{qp})$ 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., Y_{ℓ}^{m} for the finite nuclei, $e^{ik\cdot r}$ 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 Φ and χ as in Ca-1 and Ca-2, we obtain $G_i^\alpha \Phi \chi = \delta_{if} G_f^\alpha \Phi \chi$, this is not true if any other permutation of the states in χ is used. Since these new $G_i^\alpha \Phi \chi'$ are linearly independent of $G_f^\alpha \Phi \chi$, then we should have considered the Hamiltonian matrix over this complete space and diagonalized it. Then we might have tried all possible sets of ϕ_{1a} and ϕ_{1b} 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 $G_i^\alpha \Phi \chi'$ separately and vary the ϕ_{1a} and ϕ_{1b} in order to minimize the energy. Those $G_i^\alpha \Phi \chi'$ giving the lowest energies would be examined and a compromise set of ϕ_{1a} and ϕ_{1b} 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 ϕ_{1a} and among the ϕ_{1b} (I will call this the anti-anti case). This requirement is sufficient to restrict the possible $G_i^\alpha \Phi \chi'$ to just the $G_f^\alpha \Phi \chi$ 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 etal (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 (1962) 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^{α} state for $\bar{\alpha} = [n,m]$ is antisymmetrical for interchanges of the n ϕ_{ia} orbitals and also for interchanges of the m ϕ_{ib} orbitals; thus, the order of the subgroup of S_N which transforms only antisymmetrically related (and thus correlated) orbitals into each other is n!m! This is the maximum possible <u>antisymmetricity</u> for any state G_i of the same spin. The next best state is

$$S_{nb} = \begin{array}{|c|c|}\hline 1 & n \\ \hline 2 & n+2 \\ \hline 3 & \\ \cline{1-1} & \\ \cline{2-2} & m+n \\ \hline n-1 & \\ \cline{1-1} n+1 & \\ \cline{1-1}\end{array}$$

for which the antisymmetricity is (n-1)!(m-1)!. It is clear that the states of higher antisymmetricity 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 <u>factor</u> of at least nm e.g., 9 for Benzene) the fewest possible close encounters by electrons 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 $G_i^{\alpha}\Phi\chi'$ different from $G_f\Phi\chi$, we can vary the ϕ_i to get the best possible $G_f\Phi\chi$, 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 $G_f\Phi\chi$ 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 α or β ; 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 $G_f\phi_a\phi_b\chi$ and $RG_f\phi_a\phi_b\chi = C_R G_f\phi_a\phi_b\chi$, where $|C_R| = 1$. Take ϕ_a and ϕ_b to be real and note that $G_f\phi_a\phi_b\chi = G_f\phi_b\phi_a\chi$. Define

$$\psi_1 \overset{\Delta}{=} \frac{1}{\sqrt{2(1+S)}}(\phi_a+\phi_b) \qquad \text{and} \qquad \psi_2 \overset{\Delta}{=} \frac{1}{\sqrt{2(1-S)}}(\phi_a-\phi_b)$$

where $S = \langle\phi_a|\phi_b\rangle$; then ψ_1 and ψ_2 are orthonormal. Thus,

$$\phi_a = \sqrt{\frac{1+S}{2}}\,\psi_1 + \sqrt{\frac{1-S}{2}}\,\psi_2 \qquad \text{and} \qquad \phi_b = \sqrt{\frac{1+S}{2}}\,\psi_1 - \sqrt{\frac{1-S}{2}}\,\psi_2$$

Hence, $G_f\,\phi_a\phi_b\chi = (\frac{1+S}{2})G_f\,\psi_1\psi_1\chi - (\frac{1-S}{2})G_f\,\psi_2\psi_2\chi$. Define the functions $\psi_3, \psi_4, \ldots$ such that the set of all ψ_i is complete and orthonormal. Expand $R\psi_1$ and $R\psi_2$ as $R\,\psi_1 = \sum_i A_{1i}\psi_i$ and $R\,\psi_2 = \sum_i A_{2i}\psi_i$ Therefore,

$$RG_f\phi_a\phi_b\chi = \Big\{\sum_{i<j}\big[(1+S)A_{1i}A_{1j} - (1-S)A_{2i}A_{2j}\big]G_f\psi_i\psi_j\chi +$$
$$+ \sum_i \frac{1}{2}\big[(1+S)A_{1i}^2 - (1-S)A_{2i}^2\big]G_f\psi_i\psi_i\chi\Big\} = C_R\Big[\frac{(1+S)}{2}G_f\psi_1\psi_1\chi - \frac{(1-S)}{2}G_f\psi_2\psi_2\chi\Big]$$

The various $G_f\,\psi_1\psi_2\chi$ are orthogonal and, hence, linearly independent; thus, we may equate coefficients.

From $i = j > 2$ $\quad\sqrt{1+S}\,A_{1i} = \pm\sqrt{1-S}\,A_{2i}$

From $j > i > 2$ $\quad$ either all of the signs in the above equation are + or all are -

From $j > 2$, $i = 1$ $\quad(1+S)A_{1j}A_{11} = (1-S)A_{2j}A_{21} = \pm\sqrt{1-S}\,\sqrt{1+S}\,A_{ij}A_{21}$;

thus, either all $A_{ij} = 0$ for $j > 2$ or $\sqrt{1+S}\,A_{11} = \pm\sqrt{1-S}\,A_{21}$

From $i = j = 1$ $\quad(1+S)A_{11}^2 - (1-S)A_{21}^2 = C_R(1+S)$ thus, we cannot have the second of the above choices and $A_{ij} = 0$ 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: $(1+S)A_{12} - (1-S)A_{22}^2 = -C_R(1-S)$

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 $\langle R\psi_i | R\psi_j \rangle = \delta_{ij}$. We let $A_{ij} = \alpha_{ij} e^{i\theta_{ij}}$, where α_{ij} is real and positive. We <u>let $C_R = 1$</u> 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

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

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

Hence, either $R\phi_a = \pm\phi_a$ and $R\phi_b = \pm\phi_b$ (Ce-6)

or $R\phi_a = \pm\phi_b$ and $R\phi_b = \pm\phi_a$ (Ce-7)

We note that the ϕ_a function may transform into the ϕ_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).

Cf: The Alternate Orbital Method

Löwdin suggested a simple way to remove much of the correlation error in the π electron M. O. of alternate systems (e.g., benzene). In these systems there is a natural set of n excited states which can be mixed with the n filled (π 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 Itoh (1955), Itoh and Yoshizumi (1955)) and for cyclic systems of 2n atoms (where n is odd) (Paunez et al(1962)). The general theory is presented in Paunez etal (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^{i}_{j} and D^{ij}_{kt} of equation Ca-10 (or to allow different mixing parameters we simply let $S(ia;jb) = \lambda_i \delta_{ij}$). In all cases even if different parameters are used $D^{ia}_{ja} = 0$ if $i \neq j$, $D^{ia,ja}_{ka,ta} = 0$ unless $\{i,j\} = \{k,t\}$, $D^{ia,ja}_{ka,tb} = 0$ if $k = i$ or j, $D^{ia,kb}_{ia,tb} = 0$ and $D^{ia,kb}_{tb,ia} = 0$ if $k \neq t$;

If all λ are equal we obtain (N must be equal to 2n for the alternate orbital method)

$$D^{ia}_{ia} = \sum_{r=0}^{m-1} \frac{(m-r)}{m} \lambda^{r} = \frac{1}{1-\lambda^2}\left[1 - \frac{\lambda^2(1-\lambda^{2m})}{m(1-\lambda^2)}\right] \xrightarrow[\lambda\to 1]{} \frac{m+1}{2}$$

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

$$D^{ia,ja}_{ia,ja} = \sum_{r=0}^{m-2} \frac{\lambda^{2r}}{\binom{m}{r}} \binom{m-2}{r} \triangleq A = \frac{1}{1-\lambda^2} - \frac{2\lambda^2}{(m-1)(1-\lambda^2)^2} + \frac{2\lambda^2(1-\lambda^{2m})}{m(m-1)(1-\lambda^2)^3} \xrightarrow[\lambda\to 1]{} \frac{m+1}{3}$$

$$D^{ia,jb}_{ia,jb} = \sum_{r=0}^{m-2} \frac{\lambda^{2r+1}}{\binom{m}{r+1}} \binom{m-2}{r} \triangleq \lambda B = \lambda\left[\frac{(1+\lambda^{2m})}{(m-1)(1-\lambda^2)^2} - \frac{(1+\lambda^2)(1-\lambda^{2m})}{m(m-1)(1-\lambda^2)^3}\right] \xrightarrow[\lambda\to 1]{} \frac{m+1}{6}$$

$$D^{ia,ja}_{ib,jb} = \sum_{r=0}^{m-2} \frac{\lambda^{2r+2}}{\binom{m}{r+2}} \binom{m-2}{r} \triangleq \lambda^2 C = \lambda^2\left[-\frac{\lambda^{2m-2}}{(1-\lambda^2)} - \frac{2\lambda^{2m-2}}{(m-1)(1-\lambda^2)^2} + \frac{2(1-\lambda^{2m})}{m(m-1)(1-\lambda^2)^3}\right] \xrightarrow[\lambda\to 1]{} \frac{m+1}{3}$$

$$D^{ia,jb}_{ja,ib} = \sum_{r=0}^{m-2} \frac{\lambda^{2r+2}}{\binom{m}{r+1}} \binom{m-2}{r} = \lambda^2 B \quad \text{IF } i \neq j \qquad \xrightarrow[\lambda\to 1]{} \frac{m+1}{6}$$

$$D^{ia,ib}_{ia,ib} = \sum_{r=0}^{m-1} \frac{\lambda^{2r}}{\binom{m}{r}} \binom{m-1}{r} \triangleq D = \frac{1}{1-\lambda^2} - \frac{\lambda^2(1-\lambda^{2m})}{m(1-\lambda^2)^2} \longrightarrow \frac{m+1}{2}$$

$$D^{ia,jb}_{ia,jb} = \sum_{r=0}^{m-2} \frac{\lambda^{2r}}{\binom{m}{r}} \binom{m-2}{r} = A \quad \text{IF } i \neq j \qquad \longrightarrow \frac{m+1}{3}$$

$$D^{ia,ib}_{ib,ia} = \sum_{r=0}^{m-1} \frac{\lambda^{2r}}{\binom{m}{r+1}} \binom{m-1}{r} \triangleq F = \frac{-\lambda^{2m}}{1-\lambda^2} + \frac{(1-\lambda^{2m})}{m(1-\lambda^2)^2} \longrightarrow \frac{m+1}{2}$$

$$D^{ia,jb}_{jb,ia} = \sum_{r=0}^{m-2} \frac{\lambda^{2r}}{\binom{m}{r+1}} \binom{m-2}{r} = B \quad \text{IF } i \neq j \qquad \longrightarrow \frac{m+1}{6}$$

$$D^{ia,jb}_{ib,ja} = \sum_{r=0}^{m-2} \frac{\lambda^{2r+2}}{\binom{m}{r+2}} \binom{m-2}{r} = \lambda^2 C \quad \text{IF } i \neq j \qquad \longrightarrow \frac{m+1}{3}$$

all other cases are zero.

From here one next evaluates the $\langle |h| \rangle$ and $\langle |g| \rangle$ integrals in equation Ca-10 then minimizes the energy with respect to λ. 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 is the HF energy (the norm of the wave function is n + 1).

LÖWDIN (1962b, p.270) states that if DODS are used, then for an S = 0 state a transformation on the ϕ_{ia} and another on the ϕ_{ib} can be found for which S(ia;jb) $= \lambda_i \delta_{ij}$, 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) $= \lambda_i \delta_{ij}$ 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 ϕ_{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 Da 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, γ and Γ^2 (which have unsummed spins), for $C_f \Phi \chi$

$$\bar{G}_f \Phi \chi = \frac{1}{\sqrt{f^\alpha \langle \Phi | O_{ff} | \Phi \rangle \langle \chi | \omega_{\bar{f}\bar{f}} | \chi \rangle}} G_f \Phi \chi$$

$$\int (G_f \Phi \chi)^* (G_f \Phi \chi)(dx_1')(ds_1') = \sum_r \sum_t \delta_{\sigma_{rf}} \delta_{\sigma_{tf}} \left(\int O_{rf} \Phi O_{rf} \Phi (dx_1') \right) \left(\int \omega_{\bar{r}\bar{f}} \chi \omega_{\bar{t}\bar{f}} \chi (ds_1') \right)$$

$$\int \omega_{\bar{r}\bar{f}} \chi \omega_{\bar{t}\bar{f}} \chi (ds_1') = \frac{1}{\theta^\alpha} \sum_\tau \sum_{\tau'} U_{\bar{r}\bar{f}\tau} U_{\bar{t}\bar{f}\tau'} \int \tau \chi \, \tau' \chi (ds_1')$$

$$= \frac{1}{\theta^\alpha} \sum_\tau \sum_{\tau''} U_{\bar{r}\bar{f}\tau^{-1}} U_{\bar{t}\bar{f}\tau^{-1}\tau''}(1,\tau(1)) \int \chi \tau'' \chi (ds'_{\tau(1)})$$

But, $\int \chi \tau'' \chi (ds'_{\tau(1)})$ = 0 unless $\tau'' = \sigma_a \sigma_b$, where σ_a involves only elements ≤ n

= 1 IF $\tau'' = \sigma_a \sigma_b$ and σ_b only elements > n

For convenience let M_s = S in the following

$$\int \omega_{\bar{r}\bar{f}} \chi \omega_{\bar{t}\bar{f}} \chi (ds_1') = \left(\left(\tfrac{N}{2}+S\right)! \left(\tfrac{N}{2}-S\right)! / \theta^\alpha \right) \sum_\tau U_{\bar{r}\bar{f}\tau^{-1}} U_{\bar{t}\bar{f}\tau^{-1}} \; S^{(1)}_{\tau(1)} S^{(1)}_{\tau(1)}$$

$$\int O_{rf} \Phi O_{tf} \Phi (dx_1') = \frac{1}{\theta^\alpha} \sum_\tau \sum_{\tau''} U_{rf\tau^{-1}} U_{tf\tau^{-1}\tau''}(1,\tau(1)) \int \Phi \tau'' \Phi (dx'_{\tau(1)})$$

$$\underline{\gamma} \triangleq N \int (G_f \Phi \chi)(G_f \Phi \chi)(dx_1')(ds_1')$$

$$= \frac{\left(\frac{N}{2}+S\right)! \left(\frac{N}{2}-S\right)!}{(\theta^\alpha)^2} \sum_r \sum_t \delta_{\sigma_{rf}} \delta_{\sigma_{tf}} \sum_\tau \sum_{\tau_1} \sum_{\tau_2} U_{\bar{r}\bar{f}\tau^{-1}} U_{\bar{t}\bar{f}\tau^{-1}} U_{rf\tau_1^{-1}} U_{rf\tau_1^{-1}\tau_2} \times$$

$$\times S^{(1)}_{\tau(1)} S^{(1)}_{\tau(1)} (1,\tau_1(1)) \int \Phi \tau_2 \Phi (dx'_{\tau_1(1)})$$

But, $U_{\bar{r}\bar{f}\tau} = \delta_{\sigma_{rf}} \delta_\tau U_{rf\tau}$ from appendix I.

Thus, $\sum_r \delta_{\sigma_{rf}} U_{\bar{r}\bar{f}\tau^{-1}} U_{rf\tau_1^{-1}} = \delta_\tau U_{ff\tau\tau_1^{-1}}$

$$\gamma = \frac{\left(\frac{N}{2}+S\right)! \left(\frac{N}{2}-S\right)!}{\theta^\alpha} \sum_\tau \sum_{\tau_1} \sum_{\tau_2} U_{ff\tau\tau_1^{-1}} U_{ff\tau\tau_1^{-1}\tau_2} S^{(1)}_{\tau(1)} S^{(1)}_{\tau(1)} (1,\tau_1(1)) \int \Phi \tau_2 \Phi (dx'_{\tau_1(1)})$$

where γ is the normalized $\underline{\gamma}$. $\gamma(1,1') = [\alpha(1)\alpha(1') \rho_\alpha(1,1') + \beta(1)\beta(1') \rho_\beta(1,1')]$,

where $\rho_\alpha(1) \triangleq \dfrac{N}{\theta^\alpha f^\alpha \langle \Phi | O_{ff} | \Phi \rangle} \sum_{\substack{\tau \\ \tau(1) \le M}} \sum_{\tau_1} \sum_{\tau_2} U_{ff\tau\tau_1^{-1}} U_{ff\tau\tau_1^{-1}\tau_2}(1,\tau_1(1)) \int \Phi \tau_2 \Phi (dx'_{\tau(1)})$

For $\rho_\beta(1)$ the sum over τ is for $\tau(1) > M$. $\rho_\alpha(1) + \rho_\beta(1) = \rho(1)$ since

$$\sum_\tau U_{fr\tau} U_{ft\tau} = \delta_{rt}$$

$$\underline{\Gamma}^2 = N(N-1)\int (G_\gamma \Phi \chi)(G_\gamma \Phi \chi)(dx''_{12})(d\sigma''_{12}) =$$

$$= \sum_r \sum_t \zeta_{G_{r\gamma}} \zeta_{G_{t\gamma}} \left(\int o_{r\gamma}\Phi\, o_{t\gamma}\Phi(dx''_{12})\right)\left(\int \omega_{F\bar{\gamma}}\chi\, \omega_{t\bar{\gamma}}\chi(d\sigma''_{12}\right)$$

$$\int \omega_{F\gamma}\chi\, \omega_{t\bar{\gamma}}\chi(d\sigma''_{12}) = \frac{1}{\theta^\alpha}\sum_\tau \sum_{\tau'} U_{F\bar{\gamma}\tau^{-1}} U_{t\bar{\gamma}\tau^{-1}\tau'}(1,\tau(1))(2,\tau(2))\int \chi\tau'\chi(d\sigma''_{\tau(1),\tau(2)})$$

$$= \frac{(\frac{N}{2}+S)!\,(\frac{N}{2}-S)!}{\theta^\alpha}\sum_\tau \Big[U_{F\bar{\gamma}\tau^{-1}} U_{t\bar{\gamma}\tau^{-1}} \underset{\tau(1)}{S(1)}\,\underset{\tau(2)}{S(2)}\,\underset{\tau(1)}{S(1)}\,\underset{\tau(2)}{S(2)} + \underset{S(\tau(1))\neq S(\tau(2))}{U_{F\bar{\gamma}\tau^{-1}} U_{t\bar{\gamma}\tau^{-1}(\tau(1),\tau(2))}\underset{\tau(1)}{S(1)}\,\underset{\tau(2)}{S(2)}\,\underset{\tau(2)}{S(1)}\,\underset{\tau(1)}{S(2)}}\Big]$$

$$\Gamma^2 = \Big[\alpha(1)\alpha(2)\alpha(1)\alpha(2)\Pi_{\alpha\alpha}(1,2) + \beta(1)\beta(2)\beta(1)\beta(2)\Pi_{\beta\beta}(1,2) + \alpha(1)\beta(2)\alpha(1)\beta(2)\Pi_{\alpha\beta}(1,2) +$$

$$+ \beta(1)\alpha(2)\beta(1)\alpha(2)\Pi_{\beta\alpha}(1,2) + \alpha(1)\beta(2)\beta(1)\alpha(2)\Pi_{\substack{\alpha\beta\\ \beta\alpha}}(1,2) + \beta(1)\alpha(2)\alpha(1)\beta(2)\Pi_{\substack{\beta\alpha\\ \alpha\beta}}(1,2)\Big]$$

$$\Pi_{\alpha\alpha}(1,2) \overset{\Delta}{=} \frac{N(N-1)}{\theta^\alpha f^\alpha \langle\Phi|O_{\gamma\gamma}|\Phi\rangle}\sum_{\substack{\tau\\ \tau(1),\tau(2)\le M}} \sum_{\tau_1}\sum_{\tau_2} U_{\gamma\gamma\tau\tau_1^{-1}} U_{\gamma\gamma\tau\tau_1^{-1}\tau_2}(1,\tau_1(1))(2,\tau_1(2))\int \Phi\tau_2\Phi(dx''_{\tau_1(1),\tau_1(2)})$$

$$\Pi_{\alpha\beta}(1,2) \overset{\Delta}{=} \frac{N(N-1)}{\theta^\alpha f^\alpha \langle\Phi|O_{\gamma\gamma}|\Phi\rangle}\sum_{\substack{\tau\\ \tau(1)\le M\\ \tau(2)>M}} \sum_{\tau_1}\sum_{\tau_2} U_{\gamma\gamma\tau\tau_1^{-1}} U_{\gamma\gamma\tau\tau_1^{-1}\tau_2}(1,\tau_1(1))(2,\tau_1(2))\int \Phi\tau_2\Phi(dx''_{\tau_1(1),\tau_1(2)})$$

$$\Pi_{\substack{\alpha\beta\\ \beta\alpha}}(1,2) \overset{\Delta}{=} \frac{N(N-1)}{\theta^\alpha f^\alpha \langle\Phi|O_{\gamma\gamma}|\Phi\rangle}\sum_{\substack{\tau\\ \tau(1)\le M\\ \tau(2)>M}} \sum_{\tau_1}\sum_{\tau_2} U_{\gamma\gamma\tau\tau_1^{-1}} U_{\gamma\gamma(\tau(1),\tau(2))\tau\tau_1^{-1}\tau_2}(1,\tau_1(1))(2,\tau_1(2)) \times$$

$$\times \int \Phi\tau_2\Phi(dx''_{\tau_1(1),\tau_1(2)})$$

Recall that $\theta^\alpha f^\alpha = N!$

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 Φ is known γ and Γ^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 γ and Γ^2 that we must hope to try to extract the reasons for the stability of the various magnetic structures.

Db: 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 $E(S) \leqslant E(S')$ if $S < S'$. 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_i 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 x_1, x_2, x_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 $0 < x_1 < x_2 < x_3 < L$ 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, $\alpha = [2,1^2]$, $S_3 = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline \multicolumn{1}{|c|}{2} & \\ \cline{1-1} \multicolumn{1}{|c|}{3} & \\ \cline{1-1} \end{array}$, $S_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline \multicolumn{1}{|c|}{2} & \\ \cline{1-1} \multicolumn{1}{|c|}{4} & \\ \cline{1-1} \end{array}$ then consider $0 < x_1 < x_2 < L$. $G_2^\alpha \Phi\chi$ has no required nodes in this region but $G_3^\alpha \Phi\chi$ has additional required nodes for $x_1 = x_3$ and $x_2 = x_3$ thus $E_{22}^\alpha \leq E_{33}^\alpha$.

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_f^α where $\alpha = [\frac{N}{2}+S, \frac{N}{2}-S]$ (i.e., $M_s = S$) is that assumption true, and $G_f^\alpha \Phi\chi$ is never (vide infra) the lowest energy state of spin for a one-dimensional problem.

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

Proof: $S_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \vdots & \vdots \\ \hline 2n-1 & 2n \\ \hline \end{array}$ or $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \vdots & \vdots \\ \hline 2n-1 & 2n \\ \hline \end{array}$ and, thus, has no necessarily antisymmetric variables; and, hence, can be nodeless through all of configuration space. Every other S_i for all i and α has at least one occasion where adjacent letters are in the same column and thus has at least one required node. QED

---

* 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 $F_1^\alpha \Phi\chi = \sum_r (O_{r1}^\alpha \Phi)(\omega_{r1}^\alpha \chi)$ WITH $\alpha = [N]$ (thus, the total spin = $N/2$). This is because $\langle F_1^\alpha \Phi\chi | H | F_1^\alpha \Phi\chi \rangle = \text{CONST} \langle O_{r1}^\alpha \Phi | H | O_{r1}^\alpha \Phi \rangle$ and any other α 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 $S_f^{[2^n]} = \begin{array}{|c|c|}\hline 1 & n+1 \\ \hline 2 & n+2 \\ \hline 3 & n+3 \\ \hline \vdots & \vdots \\ \hline n & N \\ \hline \end{array}$ (for an even number of electrons)

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

$$S_1^{[n+1,n-1]} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \vdots & \vdots \\ \hline 2n-3 & 2n-2 \\ \hline 2n-1 & \multicolumn{1}{c}{} \\ \cline{1-1} 2m & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}$$

In spite of THM Db-1 it is <u>not</u> 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 α (i.e., $\alpha = [\frac{N}{2}+S, \frac{N}{2}-S]$) can be constructed from any linear combination of the f^α different $G_i^\alpha \Phi \chi$; since there is no pair of coordinates which is antisymmetric in every tableau of shape α , then the wave function for any spin can be made nodeless. It is apparent that the use of the $G_i^\alpha \Phi \chi$ provides a clear insight into many-electron problem, at least for the Hamiltonian A-1.

Section E

Explanation of the 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 Ea 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 <u>the big assumption</u>: 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)

$$H = 2 \sum_{i,j} J_{ij} S_i \cdot S_j \qquad \text{(E-1)}$$

where i and j refer to different <u>atoms</u> and the sum is over the <u>different</u> pairs. In analogy to expressions such as Bc-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 $\hat{S}_i$ rather than S_i, 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 Ec 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: MgF_2(1), MnF_2(2)(3), FeF_2(2)(3), CoF_2(2)(3), NiF_2(2)(3), ZnF_2(2)(3), PdF_2(6), TiO_2(1)(7), CrO_2(4)(8), MnO_2(17), TaO_2(5), GeO_2(1), and SnO_2(1). The following six substances have a structure which is essentially the same as rutile but with a slightly lower symmetry and slight (~.05 to .3 A) displacements of the atoms: VO_2(12), NbO_2(13)(14), MoO_2(9)(10), TcO_2(11), WO_2(9)(10) and ReO_2(11). The following two compounds have structures directly related to the rutile structure but with a significant distortion; CrF_2(16) and CuF_2(15). The following compounds are reported to have the rutile structure (but the work is rather old, Goldschmidt 1926, ref. 18 PbO_2, RuO_2, OsO_2, IrO_2, and TeO_2(from the Jahn-Teller theorem we would expect RuO_2 and OsO_2 to be strongly distorted, TeO_2 seems to me to be an inappropriate compound

---

(1) W. H. Baur, Acta Cryst. 9, 515-520 (1956)
(2) J. W. Stout and S. A. Reed, J. Amer. Chem. Soc. 76, 5279-5281(1954)
(3) W. H. Baur, Acta Cryst. 11, 488-490 (1958)
(4) O. Glemser, U. Hauschild and F. Trüpel, Zeit. anorg. Chem. 227, 113-126 (1954)
(5) N. Schoenberg, Acta Chem. Scand. 8, 240-245 (1954)
(6) N. Bartlett and R. Maitland, Acta Cryst. 11, 747-748 (1958)
(7) D. T. Cromer and K. Herrington, J. Amer. Chem. Soc. 77, 4708-4709 (1955)
(8) K. A. Wilhelmi and O. Jonsson, Acta Chem. Scand. 12, 1532-3 (1958)
(9) A. Magnéli, Arkiv Kemi, Min. Geol. 24A, No. 2 (1947)
(10) A. Magnéli and G. Andersson, Acta Chem. Scand. 9, 1378-1381 (1955)
(11) W. H. Zachariasen, A.C.A. Program & Abstracts of Winter Meeting (1951)F-4; this is footnote 4 of reference 10.
(12) G. Andersson, Acta Chem. Scand. 10, 623-628 (1956)
(13) B. O. Marinder, Arkiv Kemi 19, 435-446 (1962)
(14) N. Terao, Jap. J. Appl. Phys. 2, 156-174 (1963)
(15) C. Billy and H. M. Haendler, J. Amer. Chem. Soc. 79, 1049-1051 (1957)
(16) K. H. Jack and R. Maitland, Proc. Chem. Soc. (1957) 232
(17) J. Brenet, Compt. Rend. 230, 1360-1362, 1950-1952 (1950)
(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 ($\rho = 2\,\Omega$-cm at the transition), the high temperature form has $\rho = .02\,\Omega$-cm at the transition). On the other hand MnO_2, MnF_2,FeF_2,

---

(19) K. Brandt, Arkiv Kemi, Min. Geol. <u>17A</u>, #15 (1943).

(20) R. A. Erickson, Phys. Rev. <u>90</u>, 779-785 (1953).

(21) R. A. Erickson, Phys. Rev. <u>85</u>, 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.

(22) A. Yoshimori, J. Phys. Soc. Japan <u>14</u>, 807-821 (1959).

(23) E. Hoschek and W. Klemm, Z anorg. Chem. <u>242</u>, 63-69 (1939).

(24) J. Jaffray and A. Dumas, J. Recherches Centre Natl. Recherche Sci. Labs. Bellevue (Paris) <u>5</u>, 360 (1953-54); as quoted by ref. 26.

(25) N. Perakis and J. Wucher, Compt. Rend. <u>235</u>, 354-356 (1952); reference 27 gives the wrong volume and page number for this reference.

(26) F. J. Morin, Phys. Rev. Lett. <u>3</u>, 34-36 (1959).

CoF_2, and NiF_2 have Neel points of about 90°K, 70°K, 80°K, 40°K, and 70°K respectively ((27) p. 14 ff, I have averaged and rounded off the experimental values). Ferromagnetic CrO_2 has a Curie point at 116°C (28). Erickson found that the temperature dependence of the magnetic structure factor of the (100) diffraction for MnF_2, FeF_2, CoF_2, and NiF_2 are consistent with localized spins on the transition element of magnitude 5/2, 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 CrO_2 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 MnF_2, CoF_2, and NiF_2 but not quite quenched in FeF_2.

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 VO_2 which shows a transition from a semiconductor to a poor conductor (26) and CrO_2 which has a resistivity

---

(27) T. Nagamiya, K. Yosida, and R. Kubo, Adv. in Phys. 4, 1-112(1955)
(28) A. Michel and J. Bénard, Compt. Rend. 200, 1316-1318 (1935)
(29) C. Guillaud, A. Michel, J. Bénard, & M. Fallot, Compt. Rend. 219, 58-60 (1944)

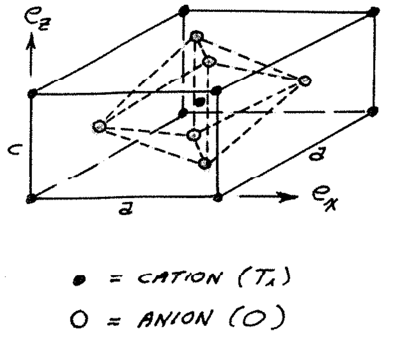
of 10^{-2} to 10^{-3} Ω-cm at room temperature (30)) suggests that the set of all extravalence electrons of all the transition element ions may be considered as somewhat independent of the other electrons (31) in the solid (the ones which are responsible for the stability of the structure) and, in addition, 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).

---

(30) T. J. Swoboda, etal, J. Appl. Phys. 32, 374S-375S (1961)
(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 $\pm ce_z$, $J = J_2$), next-nearest neighbors (eight at $\pm\frac{a}{2}e_x \pm \frac{a}{2}e_y \pm \frac{c}{2}e_z$, $J = J_1$), and the favorable (for superexchange because of the intervening oxygen atoms) pair of next-next-next-nearest neighbors (two at $+ae_x - ae_y$ and $-ae_x + ae_y$ for one sublattice and at $+ae_x + ae_y$ and $-ae_x - ae_y$ 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,

$$H = 2J_1 \sum_{\langle i,i'\rangle} S_i \cdot S_i' + 2J_2\Big[\sum_{\langle i,j\rangle} S_i \cdot S_j + \sum_{\langle i',j'\rangle} S_{i'}' \cdot S_{j'}'\Big] + 2J_4\Big[\sum_{\langle i,j\rangle} S_i \cdot S_j + \sum_{\langle i',j'\rangle} S_{i'}' \cdot S_{j'}'\Big]$$

where, <i,j> indicates all pairs of a certain type, e.g., for J_2 we have the sum over all nearest neighbor pairs.

For J_1: $i = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{c}$;

$i' = (n_1 \pm \frac{1}{2})\,\vec{a}_1 + (n_2 \pm \frac{1}{2})\,\vec{a}_2 + (n_3 \pm \frac{1}{2})\vec{c}$

For J_2: $i' = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{c}$; $j = n_1\vec{a}_1 + n_2\vec{a}_2 + (n_3 \pm 1)\vec{c}$

For J_4: $i = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{c}$; $j = (n_1 \pm 1)\vec{a}_1 + (n_2 \mp 1)\vec{a}_2 + n_3\vec{c}$ for one sublattice

$i' = n_1\vec{a}_1 + n_2\vec{a}_2 + (n_3 + \frac{1}{2})\vec{c}$; $j_1' = (n_1 \pm 1)\vec{a}_1 + (n_2 \pm 1)\vec{a}_2 + (n_3 \pm \frac{1}{2})\vec{c}$ for other sublattice

2. $S_i = \sum_{k} S(k)e^{ik\cdot r_i}$; $S_i' = \sum_{k} S'(k)e^{ik\cdot r_i'}$

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

(3i) $$2J_1\sum_{\langle l,l'\rangle} S_l\cdot S'_{l'} = 2J_1\sum_{l} S_l\cdot\sum_{\substack{\text{NEAR}\\ \text{NEIGH}}} S_l = 2J_1\sum_{k,k'} S(k)S'(k')\sum_{l} e^{i(k\cdot r_l+k'\cdot r_l)}\,8\gamma_{K1}$$

where

$$8\gamma_{K1} = \Big[e^{\frac{i}{2}(k_xa+k_ya+k_zc)}+e^{\frac{i}{2}(k_xa-k_ya+k_zc)}+e^{\frac{i}{2}(-k_xa+k_ya+k_zc)}+e^{\frac{i}{2}(-k_xa-k_ya+k_zc)}+$$
$$+e^{\frac{i}{2}(k_xa+k_ya-k_zc)}+e^{\frac{i}{2}(k_xa-k_ya-k_zc)}+e^{\frac{i}{2}(-k_xa+k_ya-k_zc)}+e^{\frac{i}{2}(-k_xa-k_ya-k_zc)}\Big]=$$

$$= 8\cos\frac{k_xa}{2}\cos\frac{k_ya}{2}\cos\frac{k_zc}{2}$$

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

$$\sum_{l} e^{i(k+k')\cdot r_l} = e^{ig\cdot r_{l0}}N\delta_{k+k',g}$$

The unit cell is simple tetragonal; thus, so is the first Brillouin zone; so, the only way for k + k' = g = 0 is for k to be on the surface of the Brillouin zone. Note that both k and g-k must be in first Brillouin zone. The term $e^{ig\cdot r_{l0}}$ is due to not requiring that the coordinate system origin be placed on a lattice site of the sublattice under consideration. r'_{l0} 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 $e^{ig\cdot r'_{l0}}$ or $e^{ig\cdot r_{l0}}$ depending on which sublattice we sum over. However, it turns out (vide infra) that the only occasions when g ≠ 0 have $\gamma_{K1} = 0$, and if g = 0 then $e^{ig\cdot r'_{l0}} = e^{ig\cdot r_{l0}}$. Similarly (remember to count each pair only once)

3(iii) $$2J_2\sum_{\langle l,j\rangle} S_l\cdot S_j = J_2\sum_{k,k'} S(k)\cdot S(k')\sum_{l} e^{i(k+k')\cdot r_l}\,2\gamma_{K2}$$

where, $2\gamma_{K2} = e^{ik_zc}+e^{-ik_zc} = 2\cos k_zc$

(3iv) $2J_4 \sum_{\langle i,j\rangle} S_i \cdot S_j = 2J_4 \sum_{k,k'} S(k)\cdot S(k') \frac{1}{2} \sum_i e^{i(k+k')\cdot r_i}\, 2\gamma_{K4}$

$$2\gamma_{K4} = e^{i(k_x a - k_y a)} + e^{-i(k_x a - k_y a)} = 2\cos(k_x - k_y)a$$

$$2\gamma'_{K4} = 2\cos(k_x + k_y)a$$

(3v) thus,

$$\begin{aligned} H = N \Big\{ & 16 J_1 \sum_k \gamma_{K1} \Big[S(k)\cdot S'(-k) + \sum_{g\neq 0} S(k)\cdot S'(g-k)\, e^{ig\cdot r_{i0}} \Big] + \\ & + 2J_2 \sum_k \gamma_{K2} \Big[S(k)\cdot S(-k) + \sum_{g\neq 0} S(k)\cdot S(g-k) e^{ig\cdot r_{i0}} + \\ & \qquad + S'(k)\cdot S'(-k) + \sum_{g\neq 0} S'(k)\cdot S'(g-k)\, e^{ig\cdot r'_{i0}} \Big] + \\ & + 2J_4 \sum_k \Big[\gamma_{K4} \big(S(k)\cdot S(-k) + \sum_{g\neq 0} S(k)\cdot S(g-k)\, e^{ig\cdot r_{i0}} \big) + \\ & \qquad + \gamma'_{K4} \big(S'(k)\cdot S'(-k) + \sum_{g\neq 0} S'(k)\cdot S'(g-k) e^{ig\cdot r'_{i0}} \big) \Big] \Big\} \end{aligned}$$

$\gamma_{K1} = \cos\frac{k_x a}{2} \cos\frac{k_y a}{2} \cos\frac{k_z c}{2}$; $\gamma_{K2} = \cos k_z c$

$\gamma_{K4} = \cos(k_x - k_y)a$; $\gamma'_{K4} = \cos(k_x + k_y)a$

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

(4i) S_i is real, thus $S(-k) = S^*(k)$

(4ii) $S_i \cdot S_i = S_i^2 = S^2$ is a scalar and thus is left invariant by lattice translation; $S_i^2 = \sum_{k,k'} S(k)\cdot S(k')\, e^{i(k+k')\cdot r_i}$

$$= \sum_k \Big[S(k)\cdot S(-k) + \sum_{k'\neq k} S(k)\cdot S(k') e^{i(k+k')\cdot r_i} \Big]$$

Hence, $\sum_k \sum_{k'\neq -k} S(k)\cdot S(k') e^{i(k+k')\cdot r_i}$ is invariant under lattice translations.

(4iii) Consider first the terms with k' = k (i.e., assume $S(k)\cdot S(k') - 0$ if k' = k or -k) then $\sum_k S(k)\text{-}S(k)\, e^{i2k\cdot r_i}$ is invariant under lattice translations; thus,

$$\sum_k \Big\{ \big[S(k)\cdot S(k) + S^*(k)\cdot S^*(k) \big] \cos(2k\cdot r_i) + \big[S(k)\cdot S(k) - S^*(k)\cdot S^*(k) \big] \sin 2k\cdot r_i \Big\}$$

is invariant under lattice translations.

Thus, either

(4iiiα) $S(k)\cdot S(k) = 0$ (and thus $S(-k)\cdot S(-k) = 0$); no restriction on k

(4iiiβ) $2k\cdot r_i = n2\pi$, thus, $k_i = 0, \pi$; no restriction on $S(k)\cdot S(k)$

(4iiiγ) $S(k)\cdot S(k)$ = pure imaginary and $r_i = 0$ for some i; then

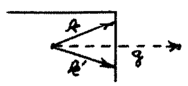
$\sin 2k\cdot r_j = 0$, hence, $2k\cdot r_j = n\pi$ or $k_i = 0, \pm\frac{\pi}{2}, \pi$ IF $r_i = 0$; we must include at least one $k_i = \pm\frac{\pi}{2}$ to get cases different from β

(4iv) Condition (iii) is too restrictive, we should only require invariance for $\sum_{\substack{k\\ \text{UNIT}\\ \text{CELL}}}\sum_{K} S(k)\cdot S(K-k)e^{iK\cdot r_i}$ where $K = k+k'$, for all K or

$\sum_{k}(S(k)\cdot S(K-k)e^{iK\cdot r_i} + S^*(k)\cdot S^*(K-k)e^{-iK\cdot r_i}$, for each K; K goes all possible values of k+k'. Thus, either,

(4iv α) $\sum_{k} S(k)\cdot S(K-k) = 0$; no restriction on K

(4iv β) $K\cdot r_j = 2\pi n$, thus $K_i = 0,\ 2\pi = k_i + k'_1$



(4iv γ) $\sum_{k} S(k)\cdot S(K-k)$ = imag and $r_i = 0$ for some i; then

$\sin(K\cdot r_j) = 0$; thus, $K\cdot r_j = m\pi$; $K_i = 0, \pm\pi,\ 2\pi = k_i + k'_i$

(4v) in all cases we get $S_i^2 = \sum_{k}\left[S(k)\cdot S(-k) + \sum_{g\in G} S(k)\cdot S(g-k)e^{ig\cdot r_i}\right]$

$G = \left\{\left(\frac{2\pi}{a},0,0\right), \left(0,\frac{2\pi}{a},0\right), \left(0,0,\frac{2\pi}{a}\right), \left(\frac{2\pi}{a},\frac{2\pi}{a},0\right), \left(\frac{2\pi}{a},0,\frac{2\pi}{a}\right), \left(0,\frac{2\pi}{a},\frac{2\pi}{a}\right), \left(\frac{2\pi}{a},\frac{2\pi}{a},\frac{2\pi}{a}\right)\right\}$

For $k \neq \frac{g}{z}$ at most one term of $\sum_{g}$ is allowed and even then k must lie on the surface of the Brillouin zone.

(5i) Now minimize H with respect to S(k) and S'(k) holding all $S_i^2 = S^2$; thus, consider $H - \sum_{i}\lambda_i S_i^2 - \sum_{i}\lambda'_i S'^2_i$ (see 3v and 4v).

$$\frac{\partial}{\partial S(k)}: N\Big\{16J_1\gamma_{K1}\Big[S'(-k) + \sum_{g\neq 0} S'(g-k)e^{ig\cdot r_{10}}\Big] + 2J_2\gamma_{K2}\Big[2S(-k) + 2\sum_{g\neq 0} S(g-k)e^{ig\cdot r_{10}}\Big] + 2J_4\gamma_{K4}\Big[2S(-k) + 2\sum_{g\neq 0} S(g-k)e^{ig\cdot r_{10}}\Big]\Big\} = \sum_{i}\lambda_i\Big[2S(-k) + 2\sum_{g\in G} S(g-k)e^{ig\cdot r_i}\Big]$$

$$\frac{\partial}{\partial S'(k)}: N\Big\{16J_1\gamma_{K1}\Big[S(-k) + \sum_{g\neq 0} S(g-k)e^{ig\cdot r_{10}}\Big] + 2J_2\gamma_{K2}\Big[2S'(-k) + 2\sum_{g\neq 0} S'(g-k)e^{ig\cdot r'_{10}}\Big] + 2J_4\gamma'_{K4}\Big[2S'(-k) + \sum_{g\neq 0} S'(g-k)e^{ig\cdot r'_{10}}\Big] = \sum_{i}\lambda'_i\Big[2S'(-k) + 2\sum_{g\in G} S'(g-k)e^{ig\cdot r'_i}\Big]$$

(5ii): Note that $\sum_i \lambda_i e^{ig\cdot r_i} = e^{ig\cdot r_{io}} \sum_i \lambda_i$, where r_{io} is the vector to some atom position of the unprimed sublattice, and $\sum_i \lambda_i' e^{ig\cdot r_i'} = e^{ig\cdot r_{io}'} \sum_i \lambda_i'$, where r_{io}' is some atom position of the primed sublattice. But, $e^{ig\cdot r_{io}'} = e^{ig\cdot r_{io}} e^{\frac{i}{2}(g_x a + g_y a + g_z c)} = (-1)^{n_g} e^{ig\cdot r_{io}}$, where n_g is the number of the $g_i = 2\pi$ (0 to 3). Cases where $n_g \neq 1$ can occur only for $k = \frac{g}{2}$. Let $N\lambda \triangleq \sum_i \lambda_i$, $N\lambda' \triangleq \sum_i \lambda_i'$

(5iii): Consider k not on surface of the Brillouin zone

$$16 J_1 \gamma_{k1} S'(-k) + (4J_2\gamma_{k2} + 4J_4\gamma_{k4}) S(-k) = 2\lambda S(-k)$$

$$16 J_1 \gamma_{k1} S(-k) + (4J_2\gamma_{k2} + 4J_4\gamma_{k4}') S'(-k) = 2\lambda' S'(-k)$$

If we multiply the first equation by S(k) and the second equation by S'(k) and sum over k, we get $\frac{2H}{N} = 2(\lambda+\lambda')S^2$. Thus, $H = N(\lambda + \lambda')S^2$; hence, H is proportional to $\lambda + \lambda'$. But, if S(k) and S'(k) are non-zero, then the determinant equals zero or S(k) = 0 and S'(k) = 0

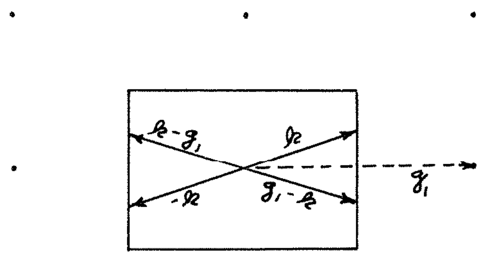
$$\begin{vmatrix} A+B-\lambda & C \\ C & A-D-\lambda \end{vmatrix} = 0$$

$$A = 2J_2\gamma_{k2}; \quad B = 2J_4\gamma_{k4}; \quad C = 8J_1\gamma_{k1}; \quad D = 2J_4\gamma_{k4}'$$

Thus, $\lambda\lambda' - \lambda'(A+B) - \lambda(A+D) + (A+B)(A+D) - C^2 = 0$. Thus, we have λ' as function of λ , but H is proportional to $\lambda + \lambda'$ and we want H to be a minimum; thus, we take $\frac{dH}{d\lambda} = NS^2(1 + \frac{d\lambda'}{d\lambda}) = 0$; thus, $\frac{d\lambda'}{d\lambda} = -1$. Substituting above we get $\lambda' + \lambda\frac{d\lambda'}{d\lambda} - \frac{d\lambda'}{d\lambda}(A+B) - (A+D) = 0$. Hence, $\lambda' - \lambda = (A+D) - (A+B) = D-B$. Substituting in the secular equation and solving, we obtain $\lambda = (A+B) \pm C$ and $\lambda' = (A+D) \pm C$. Let, $\mu \triangleq \lambda + \lambda' = 2A + (B+D) \pm 2C = 4J_2\gamma_{k2} + 2J_4(\gamma_{k4} + \gamma_{k4}') \pm 16J_1\gamma_{k1}$; then $H = N\mu S^2$; also, $\begin{pmatrix} \mp C & C \\ C & \mp C \end{pmatrix}\begin{pmatrix} S(-k) \\ S'(-k) \end{pmatrix} = 0$; thus, $S(-k) = \pm S'(-k)$ or $S^*(k) = \pm S^{*\prime}(k)$. Hence, $S(k) = \pm S'(k)$. Now note that the only non-zero S(k) and S'(k) are those for which μ 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 $\gamma_{K'1} = \gamma_{K1}$, thus $k'_x = \pm k_x$ or $\pm k_y$ or $\pm k_z$; $\gamma_{K'2} = \gamma_{K2}$, thus, $k'_z = \pm k_z$; $\gamma_{K'4} + \gamma'_{K'4} = \gamma_{K4} + \gamma'_{K4}$, thus $k'_x = \pm k_x$ and $k'_y = \pm k_y$ or $k'_x = \pm k_y$ and $k'_y = \pm k_x$. 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 homomorphic 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 defenerate 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-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

$$16 J_1 \gamma_{K1} S'(g_1-k)e^{ig_1\cdot r_{10}} + (4J_2\gamma_{K2}+4J_4\gamma_{K4})S(g_1-k)e^{ig_1\cdot r_{10}} = 2\lambda S(g_1-k)e^{ig_1\cdot r_{10}}$$

$$16 J_1 \gamma_{K1} S(g_1-k)e^{ig_1\cdot r_{10}} + (4J_2\gamma_{K2}+4J_4\gamma'_{K4})S'(g_1-k)e^{ig_1\cdot r'_{10}} = 2\lambda' S'(g_1-k)e^{ig_1\cdot r'_{10}}$$

$e^{ig_1\cdot r'_{10}} = -e^{ig\cdot r_{10}}$; thus, $\begin{vmatrix} A-B-\lambda & C \\ C & -A-D+\lambda' \end{vmatrix} = 0$

A = $2J_2\gamma_{K2}$; B = $2J_4\gamma_{K4}$; C = $8J_1\gamma_{K1}$; D = $2J_4\gamma'_{K4}$

Solving as in (iii) $\lambda' = \lambda + (D-B)$, $\lambda = (A+B) \pm iC$;

thus, there is no real solution unless C = 0. But if k is on B. Z. surface, then at least one $k_i = \pi$, and since $\gamma_{K1} = \cos\frac{k_x a}{2} \cdot \cos\frac{k_y a}{2} \cdot \cos\frac{k_z c}{2}$ then $\gamma_{K1} = 0$ and, thus, C = 0; thus, $\mu = \lambda+\lambda' = 2A+B+D = 4J_2\gamma_{K2} + 2J_4(\gamma_{K4}+\gamma'_{K4})$

<u>No relation between S(k) and S'(k) is required</u> since the minimization conditions separate into a term involving only S(k) and a term involving only S'(k)

(5v) Consider k on surface of Brillouin zone and such that k = g/2 for some g ≠ 0. Thus, we consider

$$k \in \left\{(\tfrac{\pi}{a},0,0), (0,\tfrac{\pi}{a},0), (0,0,\tfrac{\pi}{a}), (\tfrac{\pi}{a},\tfrac{\pi}{a},0), (\tfrac{\pi}{a},0,\tfrac{\pi}{a}), (0,\tfrac{\pi}{a},\tfrac{\pi}{a}), (\tfrac{\pi}{a},\tfrac{\pi}{a},\tfrac{\pi}{a})\right\}$$

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 $\gamma_{K1} = 0$. Thus, from 5i we obtain

$$(4J_2\gamma_{K2}+4J_4\gamma_{K4})S(k)e^{ig_1\cdot r_{10}} = 2\lambda S(k)e^{ig_1\cdot r_{10}}$$

$$(4J_2\gamma_{K2}+4J_4\gamma'_{K4})S'(k)e^{ig_1\cdot r'_{10}} = 2\lambda' S'(k)e^{ig_1\cdot r'_{10}}$$

Thus, $\mu = \lambda + \lambda' = 4J_2\gamma_{K2} + 2J_2(\gamma_{K4} + \gamma'_{K4})$ and there is no required relation between S(k) and S'(K).

(5vi): 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:

$$\boxed{\mu = 4J_2\gamma_{K2} + 2J_4(\gamma_{K4} + \gamma'_{K4}) \pm 16J_1\gamma_{K1}}$$, where $H = N\mu S^2$,

If $\gamma_{K1} \neq 0$ THEN $S'(k) = \pm S(k)$

If $\gamma_{K1} = 0$ then no required relation exists between S'(k) and S(k), The above ± cases are not independent, the pluses correspond and the minuses correspond.

(5vii): In any case S(k) and S'(k) must satisfy one of the conditions 4ivα 4ivβ, 4ivγ. The cases β and γ involve only a few discrete k values; however, for case α there is no restriction on the possible values of k. If the state has S(k) ≠ 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 4iiiα, 4iiiβ, and 4iiiγ are used.

(5viiiα): Since for case α there is no restriction on the possible k values, we shall find the minimum of μ with respect to variations of k.

$$\mu = 4J_2\cos k_z c + 4J_4(\cos k_x a)(\cos k_y a) \pm 16J_1\left(\cos\frac{k_x a}{2}\right)\left(\cos\frac{k_y a}{2}\right)\left(\cos\frac{k_z c}{2}\right)$$

$$\frac{\partial}{\partial k_z} : -4c\left[J_2 \sin k_z c \pm \frac{4}{2} J_1 \cos\frac{k_x a}{2} \cos\frac{k_y a}{2} \sin\frac{k_z c}{2}\right] = 0$$

$$A: \quad \sin\frac{k_z c}{2}\left[J_2 \cos\frac{k_z c}{2} \pm J_1 \cos\frac{k_x a}{2} \cos\frac{k_y a}{2}\right] = 0$$

$$\frac{\partial}{\partial k_x} : -4a\left[J_4 \sin k_x a \cos k_y a \pm \frac{4}{2} J_1 \sin\frac{k_x a}{2} \cdot \cos\frac{k_y a}{2} \cdot \cos\frac{k_z c}{2}\right] = 0$$

$$\frac{\partial}{\partial k_y} : -4a\left[J_4 \cos k_x a \sin k_y a \pm \frac{4}{2} J_1 \cos\frac{k_x a}{2} \cdot \sin\frac{k_y a}{2} \cdot \cos\frac{k_z c}{2}\right] = 0$$

adding or subtracting these we get

$\left[J_4 \sin (k_x \pm k_y)a \pm 2J_1 \sin\frac{(k_x \pm k_y)a}{2} \cos\frac{k_z c}{2}\right] = 0$ Thus,

B:
C: $\sin (k_x \pm k_y)\frac{a}{2}\left[J_4 \cos(k_x \pm k_y)\frac{a}{2} \pm J_1 \cos\frac{k_z c}{2}\right] = 0$

A, B, and C are all necessary conditions for μ to be a minimum with respect to variations in k.

(5viii β) To satisfy A we must have either

I: $k_z c = 0$ or II: $J_2 \cos\frac{k_z c}{2} = \mp J_1 \cos\frac{k_x a}{2} \cdot \cos\frac{k_y a}{2}$. To satisfy B we must have either ρ: $k_x a = -k_y a$, σ: $k_x a = k_y a = \pi$, OR

τ: $J_4 \cos(k_x + k_y)\frac{a}{2} = \mp J_1 \cos\frac{k_z c}{2}$.

To satisfy C we must have either

r: $k_x a = k_y a$ or s: $J_4 \cos(k_x - k_y)\frac{a}{2} = \mp J_1 \cos\frac{k_z c}{2}$. We consider all possible combinations of these conditions

$\alpha 1 = \text{I}\rho\text{r}$: $\underline{k_z c = 0}$; $k_x a = +k_y a = -k_y a$, thus, $\underline{k_x a = k_y a = 0}$

$\alpha 5 = \text{I}\rho\text{s}$: $\underline{k_z c = 0}$; $\underline{k_x a = -k_y a}$, $\underline{J_4 \cos k_x a = \mp J_1}$

$\alpha 3 = \text{I}\sigma\text{r}$: $\underline{k_z c = 0}$; $\underline{k_x a = k_y a = \pi}$

$\text{I}\sigma\text{s}$: No solution

$\alpha 6 = \text{I}\tau\text{r}$: $\underline{k_z c = 0}$; $\underline{k_x a = k_y a}$, $\underline{J_4 \cos k_x a = \mp J_1}$

$\alpha 7 = \text{I}\tau\text{s}$: $\underline{k_z c = 0}$; $\mp J_1 = J_4 \cos(k_x + k_y)\frac{a}{2} = J_4 \cos (k_x - k_y)\frac{a}{2}$, thus, $\underline{k_y = 0}$ and $\underline{J_4 \cos k_x\frac{a}{2} = \mp J_1}$ or the same equations with k_x and k_y interchanged

$\alpha 8 = \text{II}\rho\text{r}$: $\underline{k_x a = k_y a = 0}$; $\underline{J_2 \cos\frac{k_z c}{2} = \mp J_1}$

$\alpha 10 = \text{II}\rho\text{s}$: $\underline{k_x a = -k_y a}$; $J_2 \cos\frac{k_z c}{2} = \mp J_1\left(\cos\frac{k_x a}{2}\right)^2$ AND

$J_4 \cos k_x a = \mp J_1 \cos\frac{k_z c}{2}$; THUS,

$J_1\left(\cos\frac{k_x a}{2}\right)^2 = J_2 \frac{J_4}{J_1} \cos k_x a = \frac{J_2 J_4}{J_1}\left[2\left(\cos\frac{k_x a}{2}\right)^2 - 1\right]$ OR

$\boxed{(\cos \frac{k_x a}{2})^2} = -\left(\frac{J_2 J_4}{J_1}\right)/\left(1-2\left(\frac{J_2 J_4}{J_1}\right)^2\right) = \boxed{1/\left(2-\frac{J_1^2}{J_2 J_4}\right)}$;

thus, $\boxed{J_1 \cos \frac{k_z c}{2}} = \mp J_4 \left[2(\cos \frac{k_x a}{2})^2 - 1\right] = \boxed{\mp J_4 \left[\left(\frac{J_1^2}{J_2 J_4}\right)/\left(2-\frac{J_1^2}{J_2 J_4}\right)\right]}$

$\alpha 4 = \mathrm{II}\sigma r$: $\boxed{k_x a = k_y a = \pi}$; $\boxed{k_z c = \pi}$

$\mathrm{II}\sigma s$: No solution

$\alpha 9 = \mathrm{II}\gamma r$: $\boxed{k_x a = k_y a}$, $J_2 \cos \frac{k_z c}{2} = \mp J_1 (\cos \frac{k_x a}{2})^2$, and

$\boxed{J_4 = \mp J_1 \cos \frac{k_z c}{2}}$; thus, $J_1 (\cos \frac{k_x a}{2})^2 = \frac{J_2 J_4}{J_1}$;

$\boxed{(\cos \frac{k_x a}{2})^2 = \frac{J_2 J_4}{J_1^2}}$

$\alpha 2 = \mathrm{II}\gamma s$: $\mp J_1 \cos \frac{k_z c}{2} = J_4 \cos (k_x + k_y)\frac{a}{2} = J_4 \cos (k_x - k_y)\frac{a}{2}$; thus

$\boxed{k_y = 0}$ and $k_x \neq 0$ (for $k_x = k_y > 0$, No solution);

$\mp J_1 \cos \frac{k_z c}{2} = J_4 \cos \frac{k_x a}{2}$ and $J_2 \cos \frac{k_z c}{2} = \mp J_1 \cos \frac{k_x a}{2}$

$J_1 \cos \frac{k_z c}{2} = J_4 \frac{J_2}{J_1} \cos \frac{k_z c}{2}$; thus, $\cos \frac{k_z c}{2} = 0$

$\boxed{k_z c = \pi}$ and thus, $\boxed{k_x a = \pi}$

Now for the energies:

$$\mu = 4J_2 \cos k_z c + 4J_4 \cos k_x a \cos k_y a \pm 16 J_1 \cos \frac{k_x a}{2} \cdot \cos \frac{k_y a}{2} \cdot \cos \frac{k_z c}{2}$$

$\alpha 1 = \mathrm{I}\rho r$: $\mu = 4J_2 + 4J_4 \pm 16 J_1$

$\alpha 5 = \mathrm{I}\rho s$: $\mu = 4J_2 + 4J_4 \left(\mp \frac{J_1}{J_2}\right)^2 \mp \frac{16}{2} J_1 \left[1 + \cos k_x a\right] = 4J_2 - 4\frac{J_1^2}{J_4} \pm 8 J_1$

$\alpha 3 = \mathrm{I}\sigma r$: $\mu = 4J_2 + 4J_4$

$\alpha 6 = \mathrm{I}\gamma r$: $\mu = 4J_2 + 4J_4 \left(\mp \frac{J_1}{J_4}\right)^2 \mp 8 J_1 \left[1 \mp \frac{J_1}{J_4}\right] = 4J_2 - 4\frac{J_1^2}{J_4} \pm 8 J_1$

$\alpha 7 = \mathrm{I}\gamma s$: $\mu = 4J_2 + 4J_4 \left(2(\cos \frac{k_x a}{2})^2 - 1\right) \pm 16 J_1 \left(\mp \frac{J_1}{J_4}\right) = 4J_2 - 4J_4 - 8\frac{J_1^2}{J_4}$

$\alpha 8 = \mathrm{II}\rho r$: $\mu = 4J_2 \left(2\frac{J_1^2}{J_2^2} - 1\right) + 4J_4 \pm 16 J_1 \left(\mp \frac{J_1}{J_2}\right) = -4J_2 + 4J_4 - 8\frac{J_1^2}{J_2}$

$\alpha 10 = \mathrm{II}\rho s$: $\mu = 4J_2 \left(2\left(\frac{J_1}{J_2}/D\right)^2 - 1\right) + 4J_4 \left(\frac{2}{D} - 1\right)^2 \pm 16 J_1 \frac{1}{D}\left(\mp \frac{J_1}{J_2}/D\right)$

$= -4J_2 + 4J_4 - 8J_4/D$ where $D = 2 - \frac{J_1^2}{J_2 J_4}$

$\alpha 4 = \mathrm{II}\sigma r$: $\mu = -4J_2 + 4J_4$

$\alpha 9 = \mathrm{II}\gamma r$: $\mu = -4J_2 + 4J_4 - 24\frac{J_2 J_4^2}{J_1^2} + 16\frac{J_2^2 J_4^3}{J_1^4}$

$\alpha 2 = \mathrm{II}\gamma s$: $= -4J_2 - 4J_4$

(6i): The conditions 4iiiα, 4iiiβ, 4iiiγ, 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 $\gamma_{k1} \neq 0$ 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 $S_i^2 = S$ be invariant under lattice translations; for cases α and β $S'(k) = \pm S(k)$ implies that $S_i = S^2$; however, case γ is not sufficient to guarantee that $S_i' = S^2$ if $S_i = S^2$. In order for case γ 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, $2(k_x \frac{a}{2} + k_y \frac{a}{2} + k_z \frac{c}{2}) = 2\pi n$ and we have the additional requirement that $k_i = \pm \frac{\pi}{2}$ must occur in pairs (e.g., $k = (\frac{\pi}{2}, \frac{\pi}{2}, \pi)$, $(\frac{\pi}{2}, -\frac{\pi}{2}, 0)$, etc.). This means that states $\gamma 1$, $\gamma 2$, $\gamma 7$, and $\gamma 8$ 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 μ 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*.

---

* $\gamma 10$ 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

	$k_x a$	$k_y a$	$k_z c$	γ_{K1}	γ_{K2}	γ_{K4}	γ'_{K4}	μ**
$\beta 1$	0	0	0	1	1	1	1	$4J_2 + 4J_4 \pm 16J_1$
$\beta 2$	0	0	π	0	-1	1	1	$-4J_2 + 4J_4$
$\beta 3$	π	0	0	0	1	-1	-1	$4J_2 - 4J_4$
$\beta 4$	π	0	π	0	-1	-1	-1	$-4J_2 - 4J_4$
$\beta 5$	π	π	0	0	1	1	1	$4J_2 + 4J_4$
$\beta 6$	π	π	π	0	-1	1	1	$-4J_2 + 4J_4$
$\gamma 1$	0	0	$\pi/2$	$\sqrt{2}/2$	0	1	1	$4J_4 \pm 8\sqrt{2}J_1$ NOT ALLOWED*
$\gamma 2$	$\pi/2$	0	0	$\sqrt{2}/2$	1	0	0	$4J_2 \pm 8\sqrt{2}J_1$ NOT ALLOWED*
$\gamma 3$	$\pi/2$	0	$\pi/2$	$1/2$	0	0	0	$\pm 8J_1$
$\gamma 4$	$\pi/2$	0	π	0	-1	0	0	$-4J_2$
$\gamma 5$	$\pi/2$	$\pi/2$	0	$1/2$	1	1	-1	$4J_2 \pm 8J_1$
$\gamma 6$	$\pi/2$	$-\pi/2$	0	$1/2$	1	-1	1	$4J_2 \pm 8J_1$
$\gamma 7$	$\pi/2$	$\pi/2$	$\pi/2$	$\sqrt{2}/4$	0	1	-1	$\pm 4\sqrt{2}J_1$ NOT ALLOWED*
$\gamma 8$	$\pi/2$	$-\pi/2$	$\pi/2$	$\sqrt{2}/4$	0	-1	1	$\pm 4\sqrt{2}J_1$ NOT ALLOWED*
$\gamma 9$	$\pi/2$	$\pi/2$	π	0	-1	1	-1	$-4J_2 + 4J_4$ ***
$\gamma 10$	$\pi/2$	$-\pi/2$	π	0	-1	-1	1	$-4J_2 - 4J_4$ ***
$\gamma 11$	$\pi/2$	π	0	0	1	0	0	$4J_2$
$\gamma 12$	$\pi/2$	π	$\pi/2$	0	0	0	0	0
$\gamma 13$	$\pi/2$	π	π	0	-1	0	0	$-4J_2$
$\gamma 14$	π	0	$\pi/2$	0	0	-1	-1	$-4J_4$
$\gamma 15$	π	π	$\pi/2$	0	0	1	1	$4J_4$
$\alpha 1$	0	0	0	SAME AS $\beta 1$				
$\alpha 2$	π	0	π	SAME AS $\beta 4$				
$\alpha 3$	π	π	0	SAME AS $\beta 5$				
$\alpha 4$	π	π	π	SAME AS $\beta 6$				
$\alpha 5$	–	$-k_x a$	0	$J_4 \cos k_x a = \mp J_1$				$4J_2 - 4(J_1^2/J_4) \pm 8J_1$
$\alpha 6$	–	$+k_x a$	0	$J_4 \cos k_x a = \mp J_1$				SAME AS $\alpha 5$
$\alpha 7$	–	0	0	$J_4 \cos(k_x a/2) = \mp J_1$				$4J_2 - 4J_4 - 8(J_1^2/J_4)$
$\alpha 8$	0	0	–	$J_2 \cos(k_z c/2) = \mp J_1$				$-4J_2 + 4J_4 - 8(J_1^2/J_2)$
$\alpha 9$	–	$k_x a$	–	$J_4 = \mp J_1 \cos(k_z c/2)$; $J_1 \cos(k_x a/2) = \sqrt{J_2 J_4}$				$-4J_2 + 4J_4 - 24(J_2 J_4^2/J_1^2) + 16(J_2^2 J_4^3/J_1^4)$
$\alpha 10$	–	$-k_x a$	–	$J_1 \cos(k_z c/2) = \mp J_1^2/J_2 D$; $(\cos(k_x a/2))^2 = 1/D$ $D = 2 - J_1^2/J_2 J_4$				$-4J_2 + 4J_4 + \frac{8}{D}\frac{J_1^2}{J_2} - 16\frac{J_4}{D^2} =$ $= -4J_2 + 4J_4 - \frac{8}{D}J_4$

---

* See 6ii

** S'(k) = ± S(k) if $\gamma_{K1} \neq 0$

*** $\gamma 9$ and $\gamma 10$ have $\gamma_{K1} = 0$, 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 γ 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 β and γ states we obtain the lowest states for the various j_2 and j_4 values as shown in Figure Eb-1. Comparing the various α 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 5viii 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 α, β, and γ 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, etal (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-

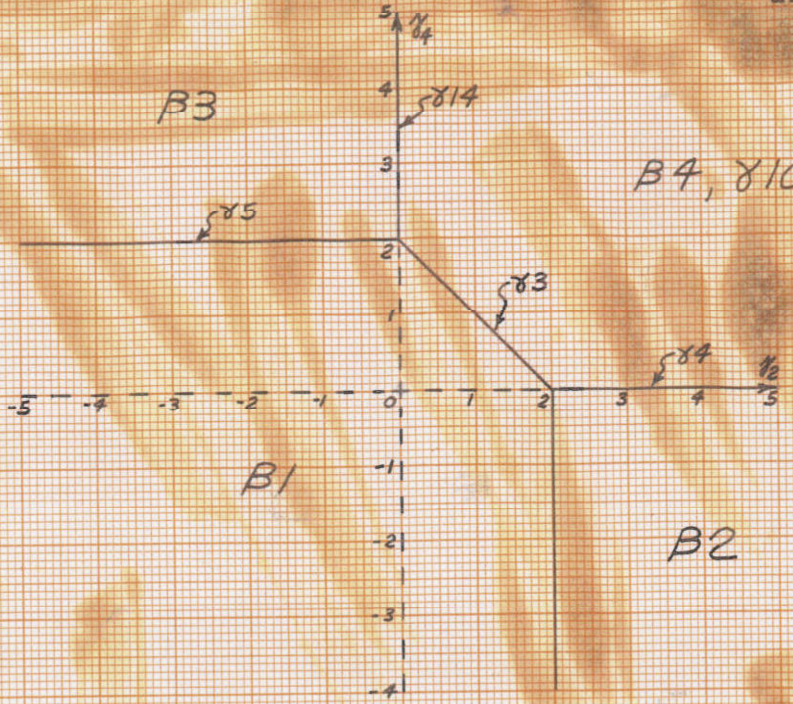


Figure 1: The Lowest Allowed States Among the β and γ Cases

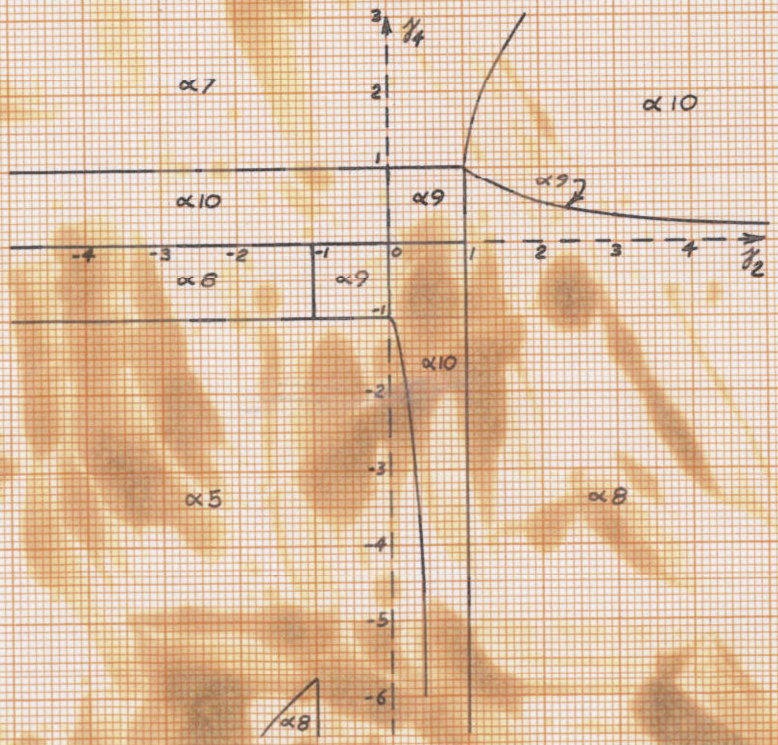


Figure 2: The Lowest States Among α5 Through α10

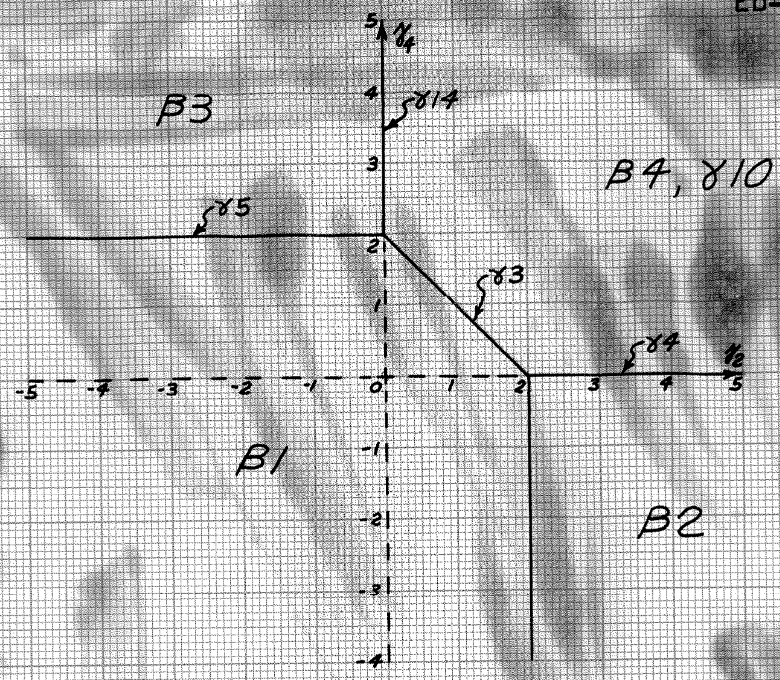


Figure 1: The Lowest Allowed States Among the β and γ Cases

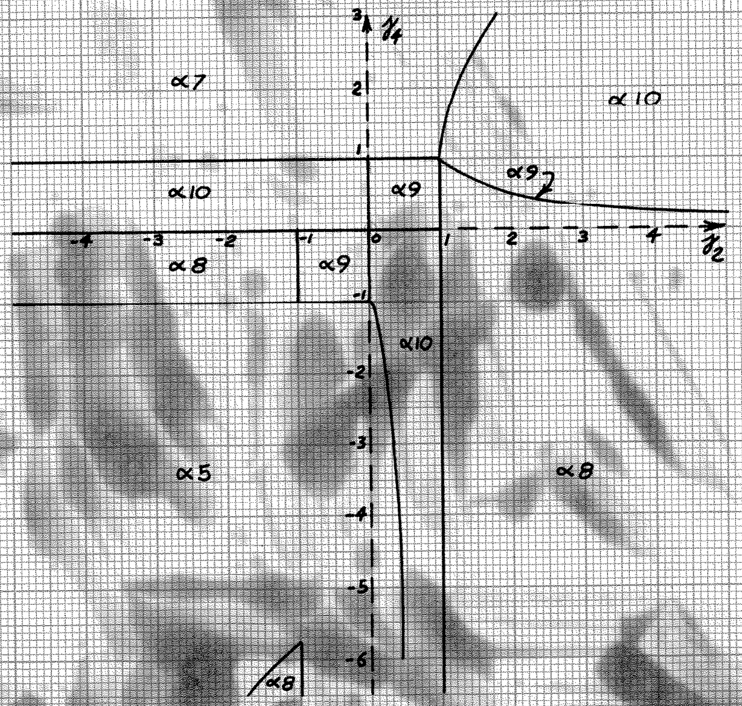


Figure 2: The Lowest States Among α5 Through α10

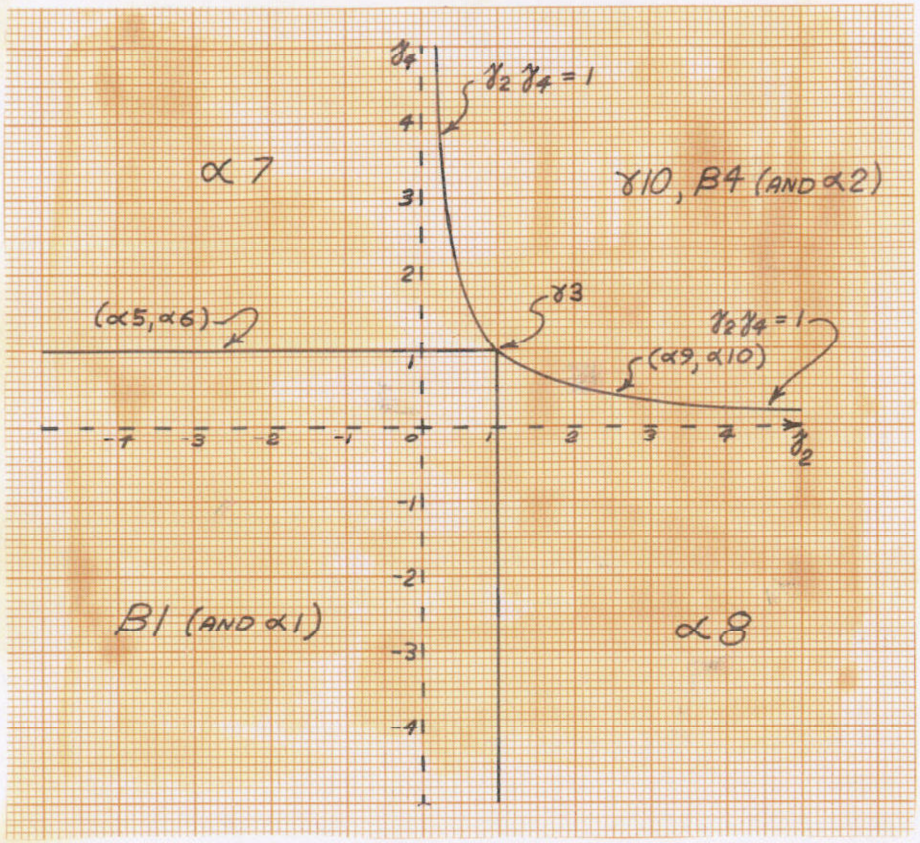


Figure 3: The Stable Magnetic Arrangements For the Rutile Structure

γ3 is stable at the point $j_2 = j_4 = 1$ and is degenerate there with β1, β4, γ10, α7, and α8

α5 and α6 are stable on the line $j_4 = +1$, $j_2 \leq +1$ and are degenerate there with α7 and β1; on this line α5 and α6 are identical with β1

α9 and α10 are stable on the line $j_2 j_4 = 1$, $0 < j_4 \leq 1$ and are degenerate there with α8 and β4; on this line α9 and α10 are identical with α8

α1 is identical with β1

α2 is identical with β4

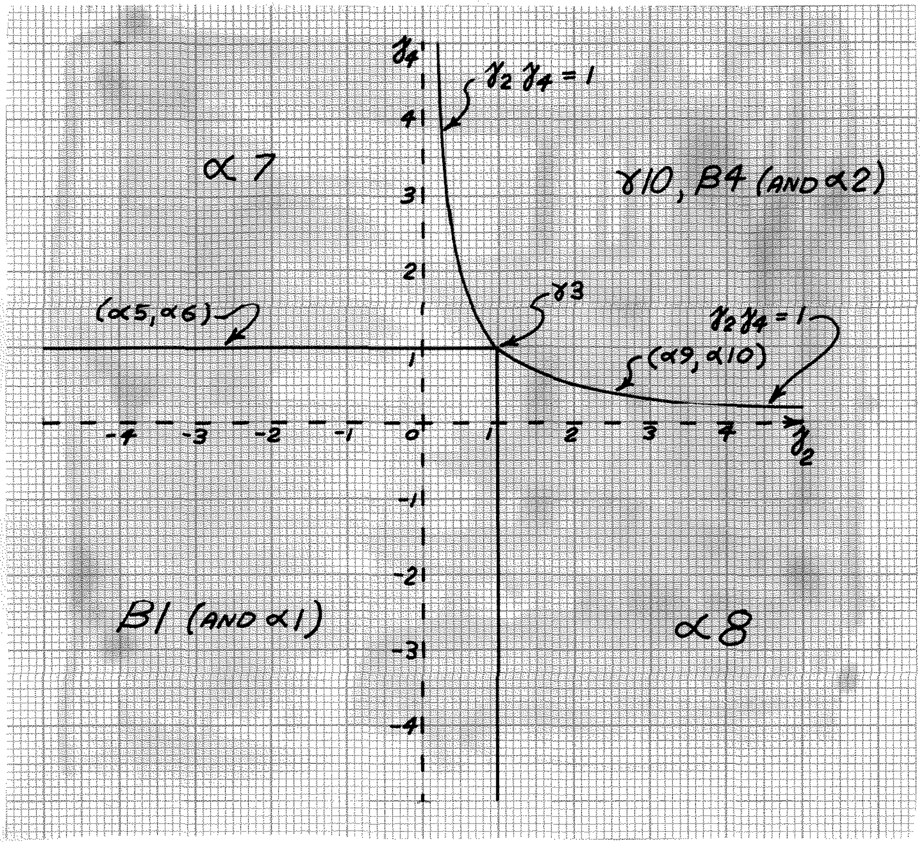


Figure 3: The Stable Magnetic Arrangements For the Rutile Structure

γ3 is stable at the point $j_2 = j_4 = 1$ and is degenerate there with β1, β4, γ10, α7, and α8

α5 and α6 are stable on the line $j_4 = +1$, $j_2 \leq +1$ and are degenerate there with α7 and β1; on this line α5 and α6 are identical with β1

α9 and α10 are stable on the line $j_2 j_4 = 1$, $0 < j_4 \leq 1$ and are degenerate there with α8 and β4; on this line α9 and α10 are identical with α8

α1 is identical with β1

α2 is identical with β4

action. From the start he used the same Lagrange multipliers for each sublattice. He also failed to discuss the g ≠ 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 γ 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 γ-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 ±c; one with the same spin and one with the opposite spin. (This results since $S_i = \sum_k S(k)e^{ikr} = 2\mu_4 \cos(k\cdot r + \frac{\pi}{4})$ 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 μ of $(-4\, J_4^{\#} + 4J_4^{//})$. Thus, for $\gamma 3$

$$\mu = \pm 8J_1 - 4(J_4^{//} - J_4^{\#}) = \pm 8J_1 - 4xJ_1 \qquad \text{(Ec-1)}$$

$$\text{where } x = \frac{J_4^{\#} - J_4^{//}}{J_1} \geq 0$$

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

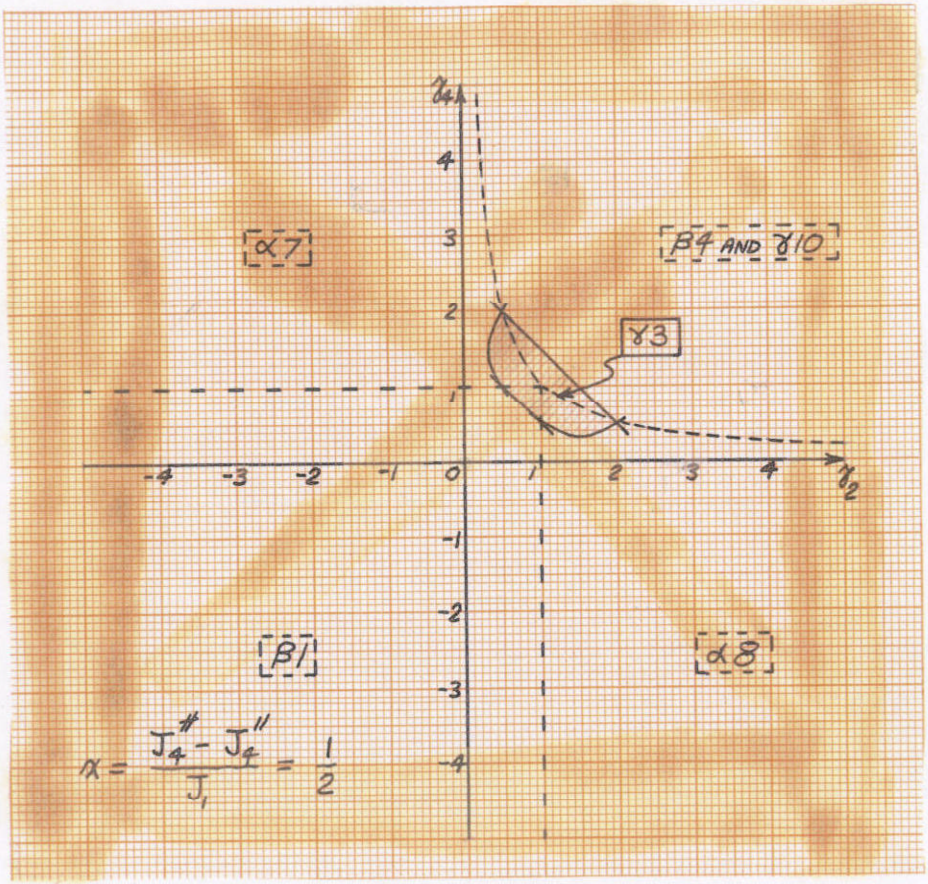


Figure Ec-1: The Stability Region For the γ3 Magnetic Arrangement For x = 1/2

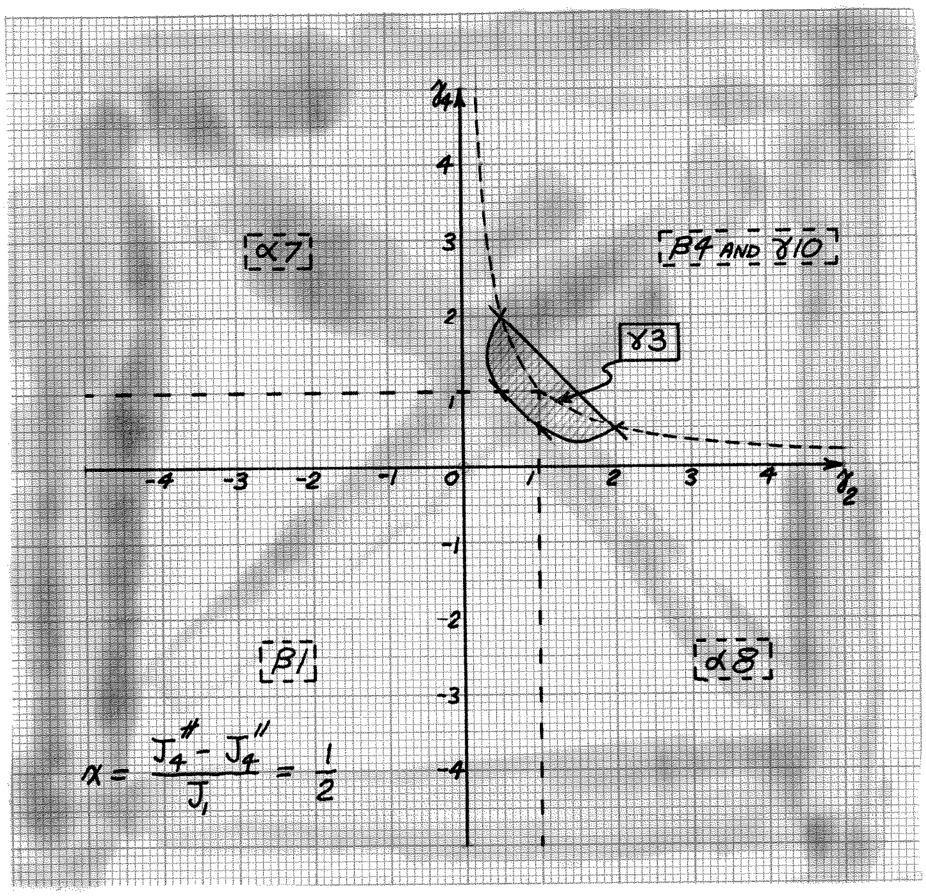
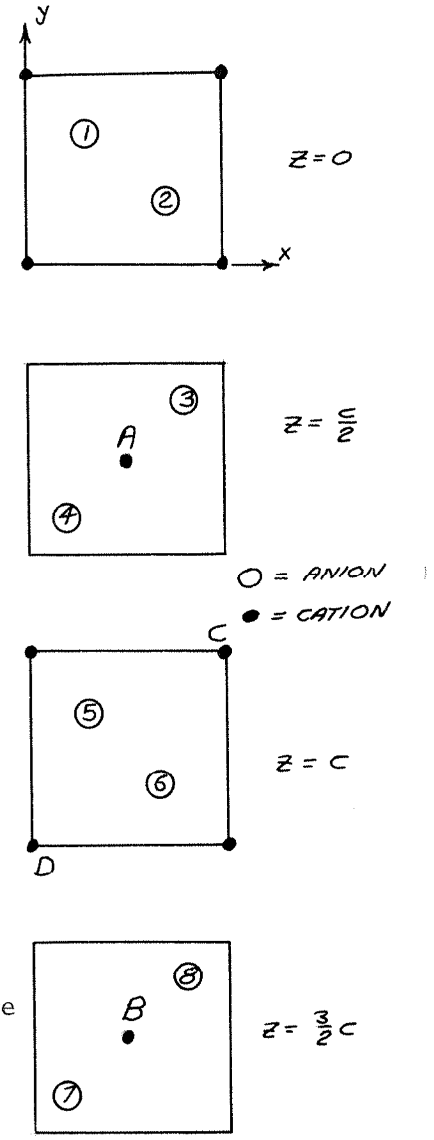
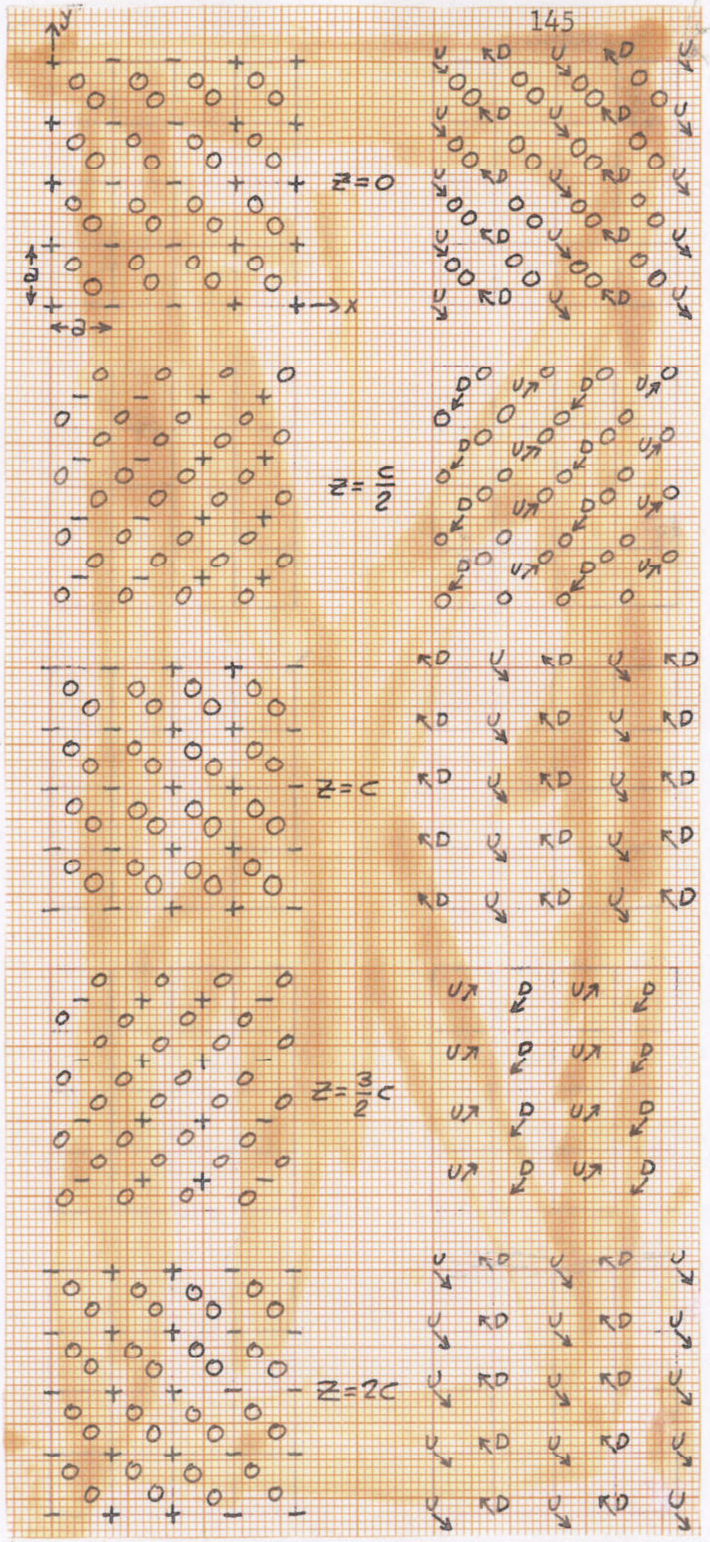


Figure Ec-1: The Stability Region For the γ3 Magnetic Arrangement For x = 1/2 (e.g., anions 1, 2, 3, 4, 5, and 6 "surround" cation A). If A and B displace toward each other along the line between them, they conflict with anions 5 and 6 which are already very close together. The extravalence electrons 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, <u>if any</u>, distorted structures can result from the γ3 magnetic arrangement.



From before, $S_i = 2\mu_4 \cos(k \cdot r_i + \frac{\pi}{4})$, $S'_i = -2\mu_4 \cos(k \cdot r_i + \frac{\pi}{4})$ and $k_x a = \frac{\pi}{2}$, $k_y a = 0$, $k_z c = \frac{\pi}{2}$. 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 $\vec{a}_x' = \frac{1}{2} a_x - \frac{1}{2} a_z$. 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.

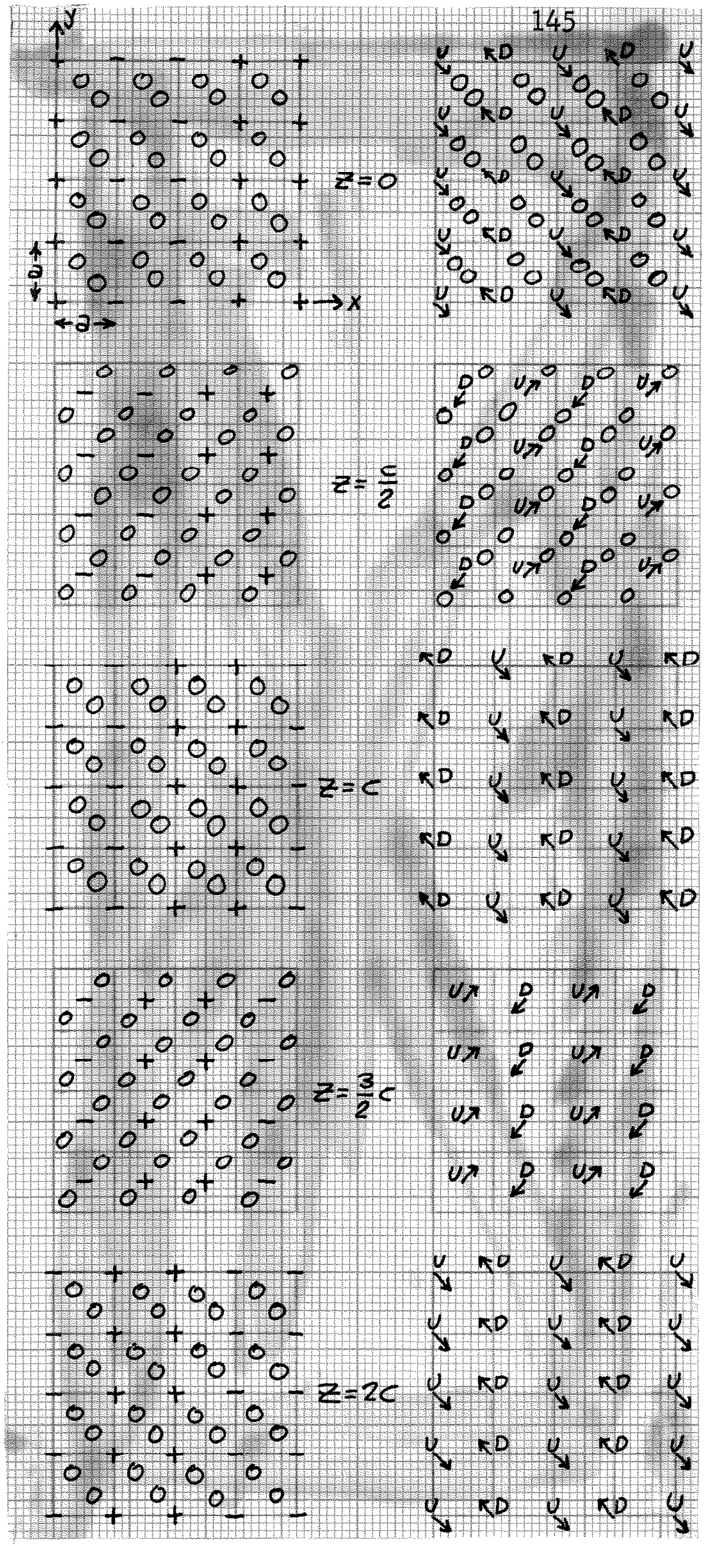
0 indicates an anion in the same plane.

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 γ 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.

O 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 $\gamma 3$ Magnetic Arrangement and the Resulting (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 $\alpha 8$ 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 explanation 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 $\beta 4$ (quite possible noting the region in which VO_2 is stable), then the $\gamma 10$ structure would result ($\beta 4$ and $\gamma 10$ have same energy for the undistorted 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 literature; 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 exchange integrals for the various compounds; it should be noted that there is no known relation between these exchange integrals and the actual 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°K, but VO_2 has a Néel point of 340°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 superexchange type J_1 interaction) we may calculate for VO_2 the $J_{\#} - J_{//}$, i.e., the change in exchange parameter with distance times the distance. We use $kT_N = \frac{S(S+1)}{3}\mu$. YOSHIMORI (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 $\alpha 8$ differ by my j_4 being replaced by his $2j_3$). Thus, for MnO_2 $\frac{kT_N}{J_1} = \frac{5}{4}\left[4j_2 - 4j_4 + \frac{8}{j_2}\right] = 11.75$; $T_N = 90°K$, thus, $2 + x = 46$ and $x = 44 = (J_{\#} - J_{//})/J_1$; thus, $J_{\#} - J_{//} = \frac{44}{1600} = \frac{1}{36}$ ev. This value seems unreasonably high. If a value of 10 times the previous J_1 is used, the new $x = 2.6$ which is much more reasonable and $J_{\#} - J_{//} = \frac{2.6}{160} = \frac{1}{60}$ ev. The displacement of the V from the center of symmetry is 0.24 Å(ANDERSSON (1956)); thus, (averaging) J_2 changes as $\frac{1}{30}$ ev per Å; note that the effect of changes in J_4 has been neglected; inclusion of J_4 would decrease the calculated change in J_2.

The point to be made about the results of section E is that although 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 neutron 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 VO_2 does not occur (and similarly for the one for TaO_2) then we may simply discard the theory.

APPENDIX A:

Summary: We prove that for tableau S_i^{α} and S_j^{α} like this if $p \leq 2n$ and if q is not in the same row as p then $U^{\alpha}_{ij(p,q)} = -\frac{1}{2}\delta_{ij}$

1	2
3	4
⋮	⋮
2n-1	2n
2n+1	⋮
⋮	

Lemma: Let $S_i = \sigma_{ij} S_j$, in order for $U_{ij\tau}$ to be non-zero; then for any expansion of τ in elementary transpositions it must be possible to get some expansion of σ_{ij} in elementary transpositions by deleting elementary transpositions from the expansion of

Proof: Remember from Aa-1 that all diagonal elements of $U_{ij(k,k+1)}$ are non-zero but the off-diagonal elements are non-zero iff $S_i = (k,k+1)S_j$. $U_{ij\tau} = \sum_{sm..pq} U_{is\tau_1} U_{sm\tau_2} \cdot\cdot U_{pq\tau_{r-1}} U_{qj\tau_r}$, where: $\tau = \tau_1\tau_2 \cdot\cdot \tau_{r-1}\tau_r$ is an expansion of τ in elementary transpositions. But $U_{qj\tau_r} \neq 0$ iff $q = j$ or $q = \tau_r j \hat{=} rj$. If $q = j$ then $U_{pq\tau_{r-1}} \neq 0$ iff $p = j$ or $p = \tau_{r-1} j \hat{=} (r-1)j$. If $q = rj$ then $U_{pq\tau_{r-1}} \neq 0$ iff $p = rj$ or $p = (r-1)rj$. We carry this on until we arrive at τ_1, by which time we have 2^{n-1}(not all different) possibilities for s; then we find that either $s = i$ or $s = \tau_1 i$ for non-zero contributions to $U_{ij\tau}$. Thus, all non-zero contributions to $U_{ij\tau}$ result from ordered selections of elementary transpositions from τ such that $S_i = \tau_a\tau_b \ldots \tau_c S_j$. We note that $\sigma_{ij} = \tau_a\tau_b \ldots \tau_c$ so $\tau_a\tau_b \ldots \tau_c$ is an expansion of σ_{ij} in elementary transpositions. If no such set of $\tau_a\tau_b \cdot\cdot \tau_c = \sigma_{ij}$ exists then $U_{ij\tau} = 0$. Note that this result is independent of the particular choice of the expansion of τ in elementary transpositions. QED

THEOREM a: If $d_{p,p+1} = \pm 1$, then $U_{ii(p,q)} = U_{ii(p+1,q)}$

Proof: $(p,q) = (p,p+1)(p+1,q)(p+1,p)$

$$U_{ii(p,q)} = \sum_{jk} U_{ij(p,p+1)} U_{jk(p+1,q)} U_{ki(p,p+1)} =$$
$$= \left(U_{ii(p,p+1)}\right)^2 U_{ii(p+1,q)} = U_{ii(p+1,q)}$$

Define a tableau lowest-in-2n as one like

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

Theorem bi: If S_i^α and S_j^α are lowest-in-2n and $p \leq 2n < q$, then

$$U^\alpha_{ij(p,q)} = U^\alpha_{ij(2n,q)}$$

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 for $S_k = (p,p+1)S_i$; 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 $S_j = \tau S_i$, then $S_m = \tau S_k$; denote j by τi and m by τk. S_k and S_m are the same in p and p+1; so,

$$U_{mj(p+1,p)} = U_{\tau k,\tau j(p+1,p)} = U_{ki(p+1,p)};$$

$$U_{ij(p,q)} = \left[\left(U_{ii(p,p+1)}\right)^2 U_{ij(p+1,q)} + \left(U_{i,pi(p,p+1)}\right)^2 U_{pi,pj(p+1,q)}\right]$$

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

$$U_{pi,pj(p+1,q)} = U_{pi,pj(p+2,q)} = U_{ij(p+2,q)};$$

$\left(U_{i,pi(p,p+1)}\right)^2 = 1 - \left(U_{ii(p,p+1)}\right)^2$; so, $U_{ij(p,q)} = U_{ij(p+2,q)}$. We continue this until $U_{ij(p,q)} = U_{ij(2n,q)}$ QED

Theorem bii: If $p < q \leq 2n$ 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 $U_{ij(q-1,q)} = 0$.

Theorem c: For a tableau like S_i, $U_{ii(p,q)} = -\frac{1}{2}$

$$S_i = \begin{array}{|c|c|} \hline & p \\ \hline p+1 & q+ \\ \hline p+2 & \\ \cline{1-1} \vdots & \\ \cline{1-1} q & \\ \cline{1-1} \end{array}$$

Proof: This follows immediately from repeated application of part a.

Theorem di: For a tableau like S_i,

$$S_i = \begin{array}{|c|c|} \hline & p \\ \hline p+1 & p+2 \\ \hline \end{array}$$

$U_{ii(p,p+2)} = -\frac{1}{2}$

Proof: Directly from Theorem a

Theorem dii: For a tableau like S_i,

$$S_i = \begin{array}{|c|c|} \hline & p \\ \hline p+1 & p+3 \\ \hline p+2 & \\ \cline{1-1} \end{array}$$

$U_{ii(p,p+3)} = -\frac{1}{2}$

Proof: $U_{ii(p,p+3)} = \left[\left(U_{ii(p+3,p+2)}\right)^2 U_{ii(p,p+2)} + \left(U_{i,(p+2)i;(p+2,p+3)}\right)^2 U_{(p+2)i,(p+2)i;(p,p+2)}\right]$

since $U_{i(p+2)i;\ (p,p+2)} = 0$. But from Theorem di $U_{(p+2)i,(p+2)i;\ (p,p+2)} = -\frac{1}{2}$, and from Theorem c $U_{ii(p,p+2)} = -\frac{1}{2}$. From equation Aa-1 $(U_{i,(p+2)i;\ (p+2,p+3)})^2 = 1 - (U_{ii(p+3,p+2)})^2$; thus, $U_{ii(p,p+3)} = -\frac{1}{2}$

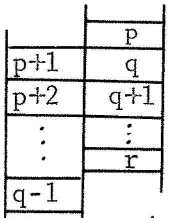
Theorem diii: For a tableau like S_i,

$$S_i = \begin{array}{|c|c|} \hline & p \\ \hline p+1 & q \\ \hline \vdots & \\ \cline{1-1} q-1 & \\ \cline{1-1} \end{array}$$

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

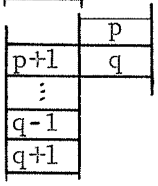
Proof: All the steps in Theorem dii are the same in this case (presuming the cases for small q to be already done).

Theorem div: For a tableau like S_i, $U_{ii(p,r)} = -\frac{1}{2}$. S_i =



Proof: This follows from Theorem diii and repeated application of Theorem A.

Theorem ei: For a tableau like S_i, S_i =

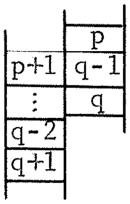


$$U_{ii(p,q+1)} = -\frac{1}{2}$$

Proof: $U_{ii(p,q+1)} = \left[(U_{ii(q,q+1)})^2 U_{ii(p,q)} + (U_{i,qi(q,q+1)})^2 U_{qi,qi(p,q)}\right]$;

$U_{ii(p,q)} = -\frac{1}{2}$ from Theorem d; $U_{qi,qi(p,q)} = -\frac{1}{2}$ from Theorem c; thus, $U_{ii(p,q+1)} = -\frac{1}{2}$

Theorem eii: For a tableau like S_i, S_i =

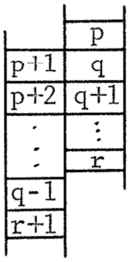


$$U_{ii(p,q+1)} = -\frac{1}{2}$$

Proof: $U_{ii(p,q)} = (U_{ii(q,q+1)})^2 U_{ii(p,q)} + (U_{i,qi(q,q+1)})^2 U_{qi,qi(p,q)}$;

$U_{ii(p,q)} = -\frac{1}{2}$ from div; $U_{qi,qi(p,q)} = -\frac{1}{2}$ from Theorem ei; thus, $U_{ii(p,q)} = -\frac{1}{2}$

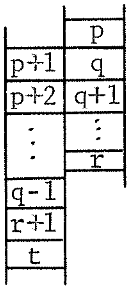
Theorem eiii: For a tableau like S_i, S_i =



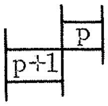
$$U_{ii(p,r+1)} = -\frac{1}{2}$$

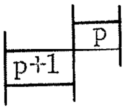
Proof: All the steps are the same as in Theorem eii (presuming the cases for lower r to be already due)

Theorem eiv: For a tableau like S_i, $U_{ii(p,t)} = -\frac{1}{2}$ S_i =



Proof: This follows from Theorem eiii and repeated application of Theorem a.

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 

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: 

if $q > p$ then $U_{ii(p,q)} = -\frac{1}{2}$.

Nowconsider $U_{ij(p,q)}$ and require $i \neq j$.

Theorem gi: For the tableau of Theorem c, $U_{ij(p,q)} = 0$ since there is only one tableau lowest-in-2n for that shape.

Theorem gii: For the tableau of Theorem di $U_{ij(p,q)} = 0$; same reason as for Theorem gi.

Theorem giii: For the tableau of Theorem dii, $S_j = (p+2,p+3)S_i$

$$U_{ij(p,p+3)} = \left[U_{ii(p+3,p+2)} U_{ii(p+2,p)} U_{ij(p+3,p+2)} + U_{ij(p+3,p+2)} U_{jj(p+2,p)} U_{jj(p+3,p+2)}\right]$$

But, from parts c and d $U_{jj(p+2,p)} = U_{ii(p+2,p)} = -\frac{1}{2}$ and from equation Aa-1 $U_{ii(p+3,p+2)} = -U_{jj(p+3,p+2)}$. Thus, $U_{ij(p,p+3)} = 0$.

Theorem giv: Continuing as in Theorem diii, $U_{ij(p,q)} = 0$ since all steps in above case are still valid.

Theorem gv: The equivalent of div, e and f for $U_{ij(p,q)}$ follow just as in the above cases where we always take

$$U_{ij(p,q)} = \left[U_{ii(q-1,q)} U_{ii(q-1,p)} U_{ij(q-1,q)} + U_{ij(q,q-1)} U_{jj(q-1,p)} U_{jj(q-1,q)}\right]$$

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 = 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, $U_H = aE$ where U_H represents H, a is a constant, and E represents ϵ. If $G = S_N$ and C is the class of transpositions, then we obtain $\sum_{\tau \in C} U^{\alpha}_{ij\tau} = a^{\alpha}\delta_{ij}$. Taking the trace $t\chi^{\alpha}_{t} = f^{\alpha}a^{\alpha}$, where χ^{α}_{t} 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^{α} is the degree of the representation α , $t = N(N-1)/2$. From LOMONT (1959) p. 263:

$f^{\alpha} = \binom{N}{k} - \binom{N}{k-1}$; p. 264: $\chi_t = -\binom{N-2}{k} - \binom{N-2}{k-2} + \binom{N-2}{k-1} + \binom{N-2}{k-3}$

$N = (N-k) + k = [\frac{N}{2}+S] + [\frac{N}{2}-S]$

$$\chi_t = -\frac{(N-2)!(\frac{N}{2}-S-1)!(\frac{N}{2}-S-2)!(\frac{N}{2}-S-3)!(\frac{N}{2}+S)!(\frac{N}{2}+S-1)!(\frac{N}{2}+S-2)!\,A}{(\frac{N}{2}-S)!(\frac{N}{2}-S-1)!(\frac{N}{2}-S-2)!(\frac{N}{2}-S-3)!(\frac{N}{2}+S)!(\frac{N}{2}+S-1)!(\frac{N}{2}+S-2)!(\frac{N}{2}+S+1)!}$$

$$\begin{aligned}A &\triangleq (\tfrac{N}{2}+S-1)(\tfrac{N}{2}+S)(\tfrac{N}{2}+S+1)+(\tfrac{N}{2}-S-1)(\tfrac{N}{2}-S)(\tfrac{N}{2}+S-1)-(\tfrac{N}{2}-S)(\tfrac{N}{2}+S)(\tfrac{N}{2}+S+1)-(\tfrac{N}{2}-S-2)(\tfrac{N}{2}-S-1)(\tfrac{N}{2}-S)\\ &= [(\tfrac{N}{2}+S)(\tfrac{N}{2}+S+1)+(\tfrac{N}{2}-S)(\tfrac{N}{2}-S-1)](2S+1) + 2[(\tfrac{N}{2}-S)(\tfrac{N}{2}-S-1)-(\tfrac{N}{2}+S)(\tfrac{N}{2}+S+1)]\\ &= (2S+1)[\tfrac{N^2}{2}-2N+2S(S+1)]\end{aligned}$$

$$\chi_t = \frac{-(N-2)!(2S+1)[\frac{N^2}{2}-2N+2S(S+1)]}{(\frac{N}{2}-S)!(\frac{N}{2}+S+1)!}$$

$$f^{\alpha} = \frac{N![\frac{N}{2}+S+1-(\frac{N}{2}-S)]}{(\frac{N}{2}+S+1)!(\frac{N}{2}-S)!} = \frac{N!(2S+1)}{(\frac{N}{2}+S+1)!(\frac{N}{2}-S)!}$$

$t = N(N-1)/2$

Thus, $\frac{t\chi_t}{f^{\alpha}} = -\frac{[\frac{N^2}{2}-2N+2S(S+1)]}{2} = -\frac{N(N-4)}{4} - S(S+1)$. Thus,

$\sum_{\tau \in C} U^{\alpha}_{ij\tau} = -\delta_{ij}[\frac{N(N-4)}{4} + S(S+1)]$; (this above formula is given by LOWDIN (1955a) p. 1476). Also, $X_t = -\frac{f^{\alpha}}{t}[\frac{N(N-4)}{4} + S(S+1)]$.

APPENDIX C:

Theorem: If S_i or S_j is lowest-in-2n and if $p < q \leq n$, then

$U_{ij(2p-1,2q-1)(2p,2q)} = \delta_{ij}$. (Note that the proof of this theorem is not complete.)

Proof: Since U_{ij} is unitary, it is sufficient to prove that

$U_{ii(2p-1,2q-1)(2p,2q)} = 1$. $(2p-1,2p)(2p-1,2q-1)(2p,2q) = (2p,2q-1)(2p-1,2q)(2p-1,2p)$ and $U_{ik(2p-1,2p)} = \delta_{ik}$ 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 $S_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}$

$(23)(14) = \underbrace{(23)(123)}_{(13)}(34)(23)(12)$; 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 $S_{\underline{k}}$; denote $(k,k+1)(m,m+1)S_1$ as $S_{\underline{km}}$, etc.

$U_{11(23)(14)} = (-\frac{1}{2})U_{11(1,3)} + (-\frac{\sqrt{3}}{2})\, U_{1\underline{2}(1,3)} = \frac{1}{4} + \frac{3}{4} = 1$

(ii) Consider p-q = 2, then $S_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline \end{array}$

$(25)(16) = (2,3,4)\overbrace{(5,4,3,2)(1,2,3,4,5)}^{(1,5)}(6,5,4)(4,3,2)(1,2)$

$$U_{11(25)(16)} = \Big\{(-\tfrac{1}{2})^3 U_{11(1,5)} + (-\tfrac{1}{2})^2(-\tfrac{\sqrt{3}}{2})\big[U_{\underline{2}1(1,5)} + U_{1\underline{2}(1,5)} + U_{1\underline{4}(1,5)}\big] +$$
$$+(-\tfrac{1}{2})(-\tfrac{\sqrt{3}}{2})^2\big[U_{\underline{2}\underline{2}(1,5)} + U_{\underline{2}\underline{4}(1,5)} + U_{1,\underline{24}(1,5)}\big] + (-\tfrac{\sqrt{3}}{2})^3 U_{\underline{2}\,\underline{24}(1,5)}\Big\}$$

From appendix A, $U_{11(1,5)} = -\frac{1}{2}$, I call terms like this no-jump terms. By direct evaluation, $U_{\underline{2},1(1,5)} = U_{1,\underline{2}(1,5)} = U_{1,\underline{4}(1,5)} = 0$; I call these one-jump terms.

$U_{\underline{2},\underline{2}(1.5)} = U_{\underline{2},\underline{4}(1,5)} = U_{1,\underline{24}(1,5)} = -\frac{1}{2}$; I call these two-jump terms. $U_{\underline{2},\underline{24}(1,5)} = -\frac{\sqrt{3}}{3}$; I call this a three-jump term. Thus,

$U_{11(25)(16)} = 1$

(iii) Consider p-q = 3, then $S_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline 7 & 8 \\ \hline \end{array}$

(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,

$U_{\underline{2},\underline{2}(1,7)} = U_{\underline{2},\underline{4}(1,7)} = U_{\underline{2},\underline{6}(1,7)} = U_{\underline{4},\underline{2}(1,7)} = U_{\underline{4},\underline{4}(1,7)} =$
$= U_{\underline{4},\underline{6}(1,7)} = -\frac{1}{2}$

$U_{\underline{24},1(1,7)} = U_{1,\underline{24}(1,7)} = U_{1,\underline{26}(1,7)} = U_{1,\underline{46}(1,7)} = 0$

$U_{\underline{24},\underline{24}(1,7)} = -\frac{1}{2}$

$U_{\underline{24},\underline{26}(1,7)} = U_{\underline{24},\underline{46}(1,7)} = U_{\underline{2},\underline{246}(1,7)} = U_{\underline{4},\underline{246}(1,7)} = -\frac{1}{3}$

$U_{\underline{24},\underline{246}(1,7)} = -\frac{2}{9}\sqrt{3}$. Thus, $U_{11(2,7)(1,8)} =$

$$= \left(-\tfrac{1}{2}\right)^5\left(-\tfrac{1}{2}\right) + \left(-\tfrac{1}{2}\right)^3\left(-\tfrac{\sqrt{3}}{2}\right)^2\left(-\tfrac{1}{2}\cdot 6\right) + \left(-\tfrac{1}{2}\right)\left(-\tfrac{\sqrt{3}}{2}\right)^4\left[-\tfrac{1}{2} - \tfrac{1}{3}\cdot 4\right] + \left(-\tfrac{\sqrt{3}}{2}\right)^5\left(-\tfrac{2\sqrt{3}}{9}\right) = 1$$

(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 SODS Hartree-Fock State.

Theorem: Consider $\alpha = [2^n, 1^{N-2n}]$. Let Φ have the standard order and obtain Φ_q from Φ by replacing the state ϕ_{iq} with some state ϕ_k for which $G_1^{\alpha}\Phi_q\chi \neq 0$. In the following cases

$$O_{r1}(p,q)\,\Phi_q = U_{11(p,q)}O_{r1}\Phi_q$$

(i) p and $q > 2n$,

PROOF(i): k is not a doubly-occupied state of Φ_o since then Φ_q would have a triply occupied state; k is not a singly occupied state of Φ_o since then Φ_f would have $n' = n + 1$ (THM Ab-1). Thus, k is not occupied in Φ_o; so, $O_{rj}\Phi_q = \delta_{j1}O_{r1}\Phi_q$;

$$O_{r1}(p,q)\Phi_q = \sum_j U_{1j(p,q)}O_{rj}\Phi_q = U_{11(p,q)}O_{r1}\Phi_q$$

(ii) p and $q \leq 2n_o$, $i_{2n_o} < k \leq i_N$ and p and q not in the same row of S_1^{α}.

PROOF (ii): Let $k = i_m$ and let $i_{q-1} = i_q$, then $(m,q-1)\Phi_q = \Phi'$ has the standard order and $n = n_o$. $O_{r1}\Phi_q = U_{11(m,q-1)}O_{r1}\Phi'$;

$$O_{r1}(p,q)\Phi_q = U_{11(p,q)(m,q-1)}O_{r1}\Phi';$$

$(p,q)(m,q-1) = (q,q-1)(p,q-1)(m,q)(q,q-1)$. Thus,

$U_{11(p,q-1)(p,q-1)(m,q)(q,q-1)} = U_{11(p,q-1)(m,q)}$. If $p < q-1$, then $U_{11(p,q-1)(m,q)} = U_{11(p,q-1)}U_{11(m,q)}$ from the Lemma of appendix A.

If $p > q-1$, use $(p,q-1)(m,q-1)(p-1,q)(p,q-1) = (m,p)(p-1,q)$; then,

$U_{11(p,q-1)(m,q)} = U_{11(p,q-1)(m,q-1)(p-1,q)(p,q-1)} = U_{11(m,p)(p-1,q)} = U_{11(m,p)}U_{11(p-1,q)}$, where appendix C and the Lemma of appendix A have been used. From appendix A, $U_{11(m,p)} = U_{11(m,q)}$ and $U_{11(p-1,q)} = U_{11(p,q)}$. Thus,

$$O_{r1}(p,q)\Phi_q = \frac{U_{11(p,q)}\,U_{11(m,q-1)}}{U_{11(m,q-1)}}\,O_{r1}\,\Phi_q = U_{11(p,q)}\,O_{r1}\,\Phi_q$$

(iii) $p,q \leq 2n_o$ and $k > i_N$.

PROOF (iii): Let $i_{q-1} = i_q$ then $(q-1,\ 2n-1)(q,2n)\Phi_q = \Phi'$ has the standard order. Using appendix C, $O_{ri}\Phi_q = \sum_j U_{1j(q,2n)(q-1,2n-1)}O_{rj}\Phi' =$

$= U_{11(q,2n)(q-1,2n-1)}O_{r1}\Phi'$;

$O_{r1}(p,q)\Phi_q = \sum_j U_{1j(p,q)(q,2n)(q-1,2n-1)}O_{rj}\Phi'$;

$(q,q-1)(p,q)(q,2n)(q-1,2n-1) = (p,q-1)(q-1,2n)(q,2n-1)(q,q-1)$.

Using the Lemma of appendix A and the appendix C

$U_{1j(p,q)(q,2n)(q-1,2n-1)} = U_{1j(p,q-1)(q-1,2n)(q,2n-1)} =$

$= U_{11(p,q-1)}U_{11(q-1,2n)(q,2n-1)}\ \delta_{j1}$. Thus,

$O_{r1}(p,q)\Phi_q = U_{11(p,q)}O_{r1}\Phi_q$.

APPENDIX E:

I have not developed general simple formulae relating the proportional $\omega_{rj}\chi$ (for different j) to each other. We know from THM Ab-2 that the proportional $\omega_{rj}\chi$ all have the same array if the convention for χ is used. As an example of how the proportionality constants can be found for any particular case we consider the partition [4,2] and $\chi = \alpha\alpha\beta\beta\beta\beta$. Recall that $\omega_{rj}\chi = \omega_{rj}\tau\chi = \sum_{i} U_{ji\tau}\,\omega_{ri}\chi$

$$S_1 = \begin{array}{|c|c|c|c|}\hline 1&2&3&4\\\hline 5&6\\\cline{1-2}\end{array}\ ;\ S_2 = \begin{array}{|c|c|c|c|}\hline 1&2&3&5\\\hline 4&6\\\cline{1-2}\end{array}\ ;\ S_3 = \begin{array}{|c|c|c|c|}\hline 1&2&4&5\\\hline 3&6\\\cline{1-2}\end{array}$$

$$S_4 = \begin{array}{|c|c|c|c|}\hline 1&2&3&6\\\hline 4&5\\\cline{1-2}\end{array}\ ;\ S_5 = \begin{array}{|c|c|c|c|}\hline 1&2&4&6\\\hline 3&5\\\cline{1-2}\end{array}\ ;\ S_6 = \begin{array}{|c|c|c|c|}\hline 1&2&5&6\\\hline 3&4\\\cline{1-2}\end{array}$$

$$(U_{jj\tau}-1)\,\omega_{rj}\chi + \sum_{i\neq j} U_{ji\tau}\,\omega_{ri}\chi = 0$$

$$(1-U_{11(45)})\,\omega_{r1}\chi = U_{12(45)}\,\omega_{r2}\chi$$

$$(1-U_{22(34)})\,\omega_{r2}\chi = U_{23(34)}\,\omega_{r3}\chi\ ;\ (1-U_{22(56)})\,\omega_{r2}\chi = U_{25(56)}\,\omega_{r5}\chi$$

$$(1-U_{33(56)})\,\omega_{r3}\chi = U_{34(56)}\,\omega_{r4}\chi$$

$$(1-U_{44(34)})\,\omega_{r4}\chi = U_{45(34)}\,\omega_{r5}\chi\ ;\ (1-U_{44(45)})\,\omega_{r4}\chi = U_{46(45)}\,\omega_{r6}\chi$$

BUT IF $S_j = (k,k+1)S_i$, THEN $U_{ij(k,k+1)} = \sqrt{1-U^2_{ii(k,k+1)}}$

$$= \sqrt{1-U^2_{jj(k,k+1)}}\ ;$$

THUS, $\sqrt{1-U_{11(45)}}\,\omega_{r1}\chi = \sqrt{1+U_{11(45)}}\,\omega_{r2}\chi$. NOTE THAT $U_{11(45)} = -U_{22(45)}$; THUS, $\sqrt{1+U_{22(45)}}\,\omega_{r1}\chi = \sqrt{1-U_{22(45)}}\,\omega_{r2}\chi$.

$$\omega_{r2}\chi = \frac{\sqrt{1-U_{11(45)}}}{\sqrt{1+U_{11(45)}}}\,\omega_{r1}\chi$$

$$\omega_{r3}\chi = \frac{\sqrt{1-U_{22(34)}}}{\sqrt{1+U_{22(34)}}}\,\omega_{r2}\chi = \frac{\sqrt{(1-U_{22(34)})(1-U_{11(45)})}}{\sqrt{(1+U_{22(34)})(1+U_{11(45)})}}\,\omega_{r1}\chi$$

$$\omega_{r4}\chi = \frac{\sqrt{1-U_{33(56)}}}{\sqrt{1+U_{33(56)}}}\,\omega_{r3}\chi = \frac{\sqrt{(1-U_{33(56)})(1-U_{22(34)})(1-U_{11(45)})}}{\sqrt{(1+U_{33(56)})(1+U_{22(34)})(1+U_{11(45)})}}\,\omega_{r1}\chi$$

$$\omega_{r5}\chi = \frac{\sqrt{1-U_{22(56)}}}{\sqrt{1+U_{22(56)}}}\,\omega_{r2}\chi = \frac{\sqrt{(1-U_{33(56)})(1-U_{11(45)})}}{\sqrt{(1+U_{33(56)})(1+U_{11(45)})}}\,\omega_{r1}\chi$$

$$\omega_{r5}\chi = \frac{\sqrt{1-U_{44(34)}}}{\sqrt{1+U_{44(34)}}}\,\omega_{r4}\chi$$

$$\omega_{r6}\chi = \frac{\sqrt{1-U_{44(45)}}}{\sqrt{1+U_{44(45)}}}\,\omega_{r4}\chi = \frac{\sqrt{(1-U_{44(45)})(1-U_{33(56)})(1-U_{22(34)})(1-U_{11(45)})}}{\sqrt{(1+U_{44(45)})(1+U_{33(56)})(1+U_{22(34)})(1+U_{11(45)})}}\,\omega_{r1}\chi$$

Thus,

$$\omega_{r2}\chi = \sqrt{\frac{5}{3}}\,\omega_{r1}\chi$$

$$\omega_{r3}\chi = \sqrt{\frac{4\cdot 5}{2\cdot 3}}\,\omega_{r1}\chi$$

$$\omega_{r4}\chi = \sqrt{\frac{4\cdot 4\cdot 5}{2\cdot 2\cdot 3}}\,\omega_{r1}\chi$$

$$\omega_{r5}\chi = \sqrt{\frac{4\cdot 5}{2\cdot 3}}\,\omega_{r1}\chi$$

$$\omega_{r6}\chi = \sqrt{\frac{3\cdot 4\cdot 4\cdot 5}{1\cdot 2\cdot 2\cdot 3}}\,\omega_{r1}\chi$$

APPENDIX F: The Matrices for the Orthogonal Representation for N = 3 and 4

N = 3, $\alpha = [2,1]$

$S_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$

$S_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}$

$$U_{\epsilon} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ; \quad U_{(1,2)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$U_{(2,3)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & +\frac{1}{2} \end{pmatrix} ; \quad U_{(13)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & +\frac{1}{2} \end{pmatrix}$$

$$U_{(123)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} ; \quad U_{(1,3,2)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

N = 4, $\alpha = [2,2]$

$S_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline\end{array}$

$S_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline\end{array}$

$$U_{\epsilon} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ; U_{(1,2)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; U_{(2,3)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

$$U_{(3,4)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; U_{(1,3)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} ; U_{(2,4)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

$$U_{(1,4)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} ; U_{(1,2)(3,4)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ; U_{(1,3)(2,4)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$U_{(14)(23)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ; U_{(1,2,3)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} ; U_{(2,3,4)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

$$U_{(1,2,4)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} ; U_{(1,3,4)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} ; U_{(1,3,2,4)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$U_{(1,4,3,2)} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} ; U_{(1,3,4,2)} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

$U_{1\gamma\tau} = U_{\gamma 1 \tau^{-1}}$: $(1,2,3)(1,3,2) = \epsilon$, $(2,3,4)(2,4,3) = \epsilon$

$(1,2,4)(1,4,2) = \epsilon$, $(1,3,4)(1,4,3) = \epsilon$, $(1,3,2,4)(1,4,2,3) = \epsilon$

$(1,4,3,2)(1,2,3,4) = \epsilon$, $(1,3,4,2)(1,2,4,3) = \epsilon$

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.

$N = 4$, $\alpha = [3,1]$: $S_1 = \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4\\\cline{1-1}\end{array}$ $S_2 = \begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3\\\cline{1-1}\end{array}$ $S_3 = \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2\\\cline{1-1}\end{array}$

$\alpha = [2,1^2]$: $S_1 = \begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1} 3\\\cline{1-1}\end{array}$ $S_2 = \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1} 4\\\cline{1-1}\end{array}$ $S_3 = \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array}$

The upper sign applies to $\alpha = [3, 1]$ the lower sign to $\alpha = [2, 1^2]$

$$U_\epsilon = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ; U_{(1,2)} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \mp 1 \end{pmatrix} ; U_{(2,3)} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \mp\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & +\frac{\sqrt{3}}{2} & \pm\frac{1}{2} \end{pmatrix}$$

$$U_{(3,4)} = \begin{pmatrix} \mp\frac{1}{3} & \frac{2}{3}\sqrt{2} & 0 \\ \frac{2}{3}\sqrt{2} & \pm\frac{1}{3} & 0 \\ 0 & 0 & \pm 1 \end{pmatrix} ; U_{(1,3)} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \mp\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \pm\frac{1}{2} \end{pmatrix} ; U_{(2,4)} = \begin{pmatrix} \mp\frac{1}{3} & -\frac{\sqrt{2}}{3} & \pm\frac{\sqrt{6}}{3} \\ -\frac{\sqrt{2}}{3} & \pm\frac{5}{6} & +\frac{\sqrt{3}}{6} \\ \pm\frac{\sqrt{6}}{3} & +\frac{\sqrt{3}}{6} & \pm\frac{1}{2} \end{pmatrix}$$

$$U_{(1,4)} = \begin{pmatrix} \mp\frac{1}{3} & -\frac{\sqrt{2}}{3} & \mp\frac{\sqrt{6}}{3} \\ -\frac{\sqrt{2}}{3} & \pm\frac{5}{6} & -\frac{\sqrt{3}}{6} \\ \mp\frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & \pm\frac{1}{2} \end{pmatrix} ; U_{(1,2)(3,4)} = \begin{pmatrix} -\frac{1}{3} & \pm\frac{2}{3}\sqrt{2} & 0 \\ \pm\frac{2}{3}\sqrt{2} & +\frac{1}{3} & 0 \\ 0 & 0 & -1 \end{pmatrix} ; U_{(1,3)(2,4)} = \begin{pmatrix} -\frac{1}{3} & \mp\frac{\sqrt{2}}{3} & +\frac{\sqrt{6}}{3} \\ \mp\frac{\sqrt{2}}{3} & -\frac{2}{3} & \mp\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & \mp\frac{\sqrt{3}}{3} & 0 \end{pmatrix}$$

$$U_{(1,4)(2,3)} = \begin{pmatrix} -\frac{1}{3} & \mp\frac{\sqrt{2}}{3} & -\frac{\sqrt{6}}{3} \\ \mp\frac{\sqrt{2}}{3} & -\frac{2}{3} & \pm\frac{\sqrt{3}}{3} \\ -\frac{\sqrt{6}}{3} & \pm\frac{\sqrt{3}}{3} & 0 \end{pmatrix} ; U_{(1,2,3)} = \begin{pmatrix} +1 & 0 & 0 \\ 0 & -\frac{1}{2} & \pm\frac{\sqrt{3}}{2} \\ 0 & \mp\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} ; U_{(2,3,4)} = \begin{pmatrix} -\frac{1}{3} & \pm\frac{2}{3}\sqrt{2} & 0 \\ \mp\frac{\sqrt{2}}{3} & -\frac{1}{6} & \pm\frac{\sqrt{3}}{2} \\ +\frac{\sqrt{6}}{3} & \pm\frac{\sqrt{3}}{6} & +\frac{1}{2} \end{pmatrix}$$

$$U_{(1,2,4)} = \begin{pmatrix} -\frac{1}{3} & \mp\frac{\sqrt{2}}{3} & \frac{\sqrt{6}}{3} \\ \mp\frac{\sqrt{2}}{3} & +\frac{5}{6} & \pm\frac{\sqrt{3}}{6} \\ -\frac{\sqrt{6}}{3} & \mp\frac{\sqrt{3}}{6} & -\frac{1}{2} \end{pmatrix} ; U_{(1,3,4)} = \begin{pmatrix} -\frac{1}{3} & \pm\frac{2}{3}\sqrt{2} & 0 \\ \mp\frac{\sqrt{2}}{3} & -\frac{1}{6} & \mp\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{6}}{3} & \mp\frac{\sqrt{3}}{6} & \frac{1}{2} \end{pmatrix} ; U_{(1,3,2,4)} = \begin{pmatrix} \mp\frac{1}{3} & -\frac{\sqrt{2}}{3} & \pm\frac{\sqrt{6}}{3} \\ -\frac{\sqrt{2}}{3} & \mp\frac{2}{3} & -\frac{\sqrt{3}}{3} \\ \mp\frac{\sqrt{6}}{3} & +\frac{\sqrt{3}}{3} & 0 \end{pmatrix}$$

$$U_{(1,4,3,2)} = \begin{pmatrix} \mp\frac{1}{3} & -\frac{\sqrt{2}}{3} & \mp\frac{\sqrt{6}}{3} \\ +\frac{2}{3}\sqrt{2} & \mp\frac{1}{6} & -\frac{\sqrt{3}}{6} \\ 0 & +\frac{\sqrt{3}}{2} & \mp\frac{1}{2} \end{pmatrix} ; U_{(1,3,4,2)} = \begin{pmatrix} \mp\frac{1}{3} & +\frac{2}{3}\sqrt{2} & 0 \\ -\frac{\sqrt{2}}{3} & \mp\frac{1}{6} & -\frac{\sqrt{3}}{2} \\ \pm\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{6} & \mp\frac{1}{2} \end{pmatrix}$$

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 C_1a, $G_2a, \ldots, G_na$ where n is the order of the group. Then $G_jG_ka = G_ia$ if G_jG_k, and we see that the set of quantities G_1a, $G_2a, \ldots, G_na$ are transformed into each other by the operations of the group. We obtain a set of orthonormal basis vectors $b_1, b_2, \ldots, b_m$ by, say, using the Gram-Schmidt procedure on the previous set ($m \leq n$). G_ib_j equals a linear combination of b_k, say, $G_ib_j = D^i_{kj}b_k$; if $G_iG_p = G_q$; then $G_iG_pb_j = D^p_{kj}G_ib_k = D^p_{kj}D_{rk}b_r = G_qb_j = D^q_{rj}b_r$; thus, $D^q_{rj} = D^i_{rk}D^p_{kj}$ (since the b_r 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 $D^1 = 1$, $D^2 = 1, \ldots, D^n = 1$, 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_1c, G_2c, ..., G_nc is orthonormalized to obtain $f_1, f_2, \ldots, f_\mu$, where $\mu < m$, then the previous representation is said to have been reduced. If the set f_1, f_2, ..., f_μ is completed to give the space of b_1, $\ldots b_m$ and if the new matrices representing the group elements are determined, then it will be found that: for every matrix, $D^i_{pq} = 0$ if $p \leq \mu$, $q > \mu$ or if $p > \mu$ and $q \leq \mu$. 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 ($\in$) and the inversion (i). If f ($\vec{x}$) is some function of $\vec{x}$, then $\epsilon f(x) = f(x)$, $i\ f(x) = f(-x)$. Thus, $f(\vec{x})$ and $f(-\vec{x})$ correspond to the b_1, $b_2, \ldots$. $\epsilon\ f(\vec{x}) = f(\vec{x}), \epsilon\ f(-\vec{x}) = f(-\vec{x})$, $i\ f(\vec{x}) = f(-\vec{x})$, $i\ f(-\vec{x}) = f(\vec{x})$; thus, if $G_1 = \in$ 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 $\in\in\ = \in$, $ii = \in$, $\in i = i\in\ = i$, and we see that $D^1D^1 = D^1$, $D^2D^2 = D^1$, $D^1D^2 = D^2D^1 = D^2$. But $[f(\vec{x}) - f(-\vec{x})]$ is contained in the space spanned by $f(\vec{x})$ and $f(-\vec{x})$; $\epsilon[f(\vec{x}) - f(-\vec{x})] = [f(\vec{x}) - f(-\vec{x})]$, $i\ [f(\vec{x}) - f(-\vec{x})] = -\ [f(\vec{x}) - f(-\vec{x})]$. Thus, the one dimensional space $[f(\vec{x}) - f(-\vec{x})]$ is invariant under the group and the previous representation has been reduced. The complementary space is $[f(\vec{x}) + f(-\vec{x})]$ for which $\epsilon [f(\vec{x}) + f(-\vec{x})] = i\ [f(\vec{x}) + f(-\vec{x})] = [f(\vec{x}) + f(-\vec{x})]$.

Consider the homogeneous differential equation $H\,\psi_i = E_i\,\psi_i$; consider transformations of the space of eigenfunctions, ψ_i. The set of transformations (say h), leaving H invariant forms a group. (If $G_i \in h$ and $G_j \in h$, then $G_iH = HG_i$ and $G_jH = HG_j$, thus, $G_iG_jH = G_iHG_j = HG_iG_j$; hence, if $G_iG_j \overset{\Delta}{=} G_k$ then $G_k \in h$; the unit element is in the set since it does not change H and $\in H = H\in$; if $G_i \in h$ then consider HG_i^{-1}, we have $G_iHG_i^{-1} = HG_iG_i^{-1} = H$, thus $G_i^{-1}(G_iHG_i^{-1}) = G_i^{-1}H$ or $G_i^{-1}H = HG_i^{-1}$, hence $G_i^{-1} \in h$; thus, the set is a group). Let $G_j \in h$ operate on $H\,\psi_i = E_i\psi_i$; then $E_iG_j\psi_i = G_jH\psi_i = HG_j\psi_k$; thus, $G_j\psi_i$ is an eigenfunction of H degenerate with the eigenfunction ψ_i. Continuing in this way the functions of the set $G_1\psi_i$, $G_2\psi_i, \ldots, G_n\psi_i$ 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 ϕ_i among themselves and the transformation matrices form a representation of the group h. All of the ϕ_i are necessarily degenerate. Approaching the other direction if $\chi_1, \chi_2, \ldots, \chi_\mu$ are basis functions of an irreducible representation of h and if χ_1 is an eigenfunction of H then $\chi_2, \ldots, \chi_\mu$ must all be eigenfunctions of H and degenerate with χ_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 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_i be the group containing the identity and the inversion. This group has two irreducible representations both of dimension one. For one, the trivial representation, $D^{\epsilon} = D^{i} = 1$ and the basis function is invariant under i; such a function is said to have even parity.

For the other irreducible representation $D^{\epsilon} = 1$, $D^{i} = -1$ 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 ϵ, t, t^2, ..., 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 $t^m \phi_k = e^{imk} \phi_k$ (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 $t(ma)\,\phi_k(x) = \phi_k(x+ma) = e^{imk}\phi_k(x)$. So far, the m and k are dimensionless, it is convenient to let $ma = M$ and $\frac{k}{a} = K$ 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

$$t(m_1a_1,\ m_2a_2,\ m_3a_3)\,\phi_{k_1k_2k_3}(x_1x_2x_3) = e^{i(k_1m_1 + k_2m_2 + k_3m_3)}\phi_{k_1k_2k_3}(x_1,x_2,x_3).$$

We let $M = m_1\vec{a}_1 + m_2\vec{a}_2 + m_3\vec{a}_3$ ($\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ not required to be orthogonal) and $K = k_1\vec{a}^1 + k_2\vec{a}^2 + k_3\vec{a}^3$ where $\vec{a}_i \cdot \vec{a}^j = \delta_{ij}$ then,

$$M \cdot K = m_1k_1 + m_2k_2 + m_3k_3 \text{ and } t(M)\,\phi_{\vec{K}}(\vec{x}) = \phi_{\vec{k}}(\vec{x} + \vec{M}) = e^{i\vec{K}\cdot\vec{M}}\phi_{\vec{K}}(\vec{x}).$$

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 $\phi_{\vec{K}}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_N)$ be the ground state eigenfunction 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 $\vec{k}'$. Let H' be the interaction between the solid and the X-ray or neutron. The transition matrix element is $I = \langle \phi_{\vec{K}}(\vec{x}_1, .., \vec{x}_N)\, \chi_{\vec{k}'}(\vec{x}_m) | H' | \phi_{\vec{K}}(\vec{x}_1, .., \vec{x}_N)\, \chi_{\vec{k}}(\vec{x}_m) \rangle$. This quantity is invariant under lattice translation (H' has the symmetry of the lattice); so, we transform with T(M), $I = T(\vec{M}) I = e^{-i\vec{M}\cdot\vec{K}} e^{-i\vec{k}'\cdot\vec{M}} e^{i\vec{M}\cdot\vec{K}} e^{i\vec{k}\cdot\vec{M}} I = e^{i(\vec{k}-\vec{k}')\cdot\vec{M}} I$; thus, $e^{i(\vec{k} - \vec{k}')\cdot M} = 1$ or $(\vec{k} - \vec{k}')\cdot\vec{M} = 2\pi n$ (n is an integer). Thus, if $(\vec{k} - \vec{k}') \overset{\wedge}{=} 2\pi\vec{g} = 2\pi(g_1\vec{a}^1 + g_2\vec{a}^2 + g_3\vec{a}^3)$, then the g_i are integers; such vectors as $\vec{g}$ are called reciprocal lattice vectors, and we see that the $2\pi\vec{q}$ 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 permutations of the particles; thus, ψ_i is a basis function of an irreducible representation of the symmetric group, S_N (see section Aa for some discussion of S_N). If we consider some perturbation of the problem, H', then the transition matrix element is $\langle \psi_j | H' | \psi_i \rangle$; but if the particles are still identical, then H' is also invariant under the operations of S_N. Thus, $\langle \psi_j | H' | \psi_i \rangle \neq 0$ only when the ψ_i and ψ_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 HAMERMESH (1962).

APPENDIX H:

Summary:

(i)* Let $S_1^{\alpha} = \begin{array}{|c|c|c|c|}\hline 1 & 2 & \cdots & n \\ \hline n+1 & n+2 & \cdots & N \\ \hline \end{array}$ where $\alpha = [n, N-n]$

Then if τ_r 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^{α}, $U^{\alpha}_{11\tau_r} = \frac{(-1)^r}{\binom{m}{r}}$

(ii)* if $S_f^{\alpha} = \begin{array}{|c|c|}\hline 1 & n+1 \\ \hline 2 & n+2 \\ \hline \vdots & \vdots \\ \hline \vdots & N \\ \hline n & \\ \hline \end{array}$ where $\alpha = [2^n, 1^{N-2n}]$

then $U_{ff\tau_r} = \frac{1}{\binom{m}{r}}$

(iii) for the α of (ii) if τ involves only elements $\leq n$ or else only elements $> n$, $U_{ff\tau\tau_r} = \delta_{\tau} \frac{1}{\binom{m}{r}}$. 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 $i \leq n < j$ and $\alpha = [n, N-n]$.

We note that $U_{kl(n,n-1,..,i)} = \delta_{kl}$, $U_{kl(n+1,n+2,..,j)} = \delta_{kl}$

$$U_{tl(i,j)} = U_{tl(i,i+i,..,n)(n,j)(n,n-1,..,i)} = U_{rl(i,i+1,..,n)(n,j)} =$$
$$= U_{tl(i,i+1,..,n)(j,j-1,..,n+1)(n,n+1)(n+1,n+2,..,j)} =$$
$$= U_{tl(i,i+i,..n)(j,j-1,..,n+1)(n,n+1)} .$$

Consider r = 2, let $\tau = (k,m)(i,j)$, where $k < i \leq n < j < m$

$U_{tl(k,m)(i,j)} = U_{tl(k,m)(i,i+1,..,n)(j,j-1,..,n+1)(n,n+1)}$; consider $U_{il(n,n-1,..,k)(i,i+i,..,n)}$, $k < i$; thus, by the lemma of appendix A the elementary transpositions involving letters $\leq i-1$ 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 $\leqslant i-1$ give unity contributions to $U_{rf\ell}$, for this reason I replace these transpositions by $\in$ or neglect; thus, $U_{ij(n,n-1,\ldots k)(i,i+1,\ldots,n)} =$

$= U_{ij(n,n-1,\ldots,i)(i-1,i)(i,i+1,\ldots,n)} = U_{ij(n,i-1)}$

$U_{t1(k,m)(i,j)} = U_{t1(k,k+1,\ldots,n)(n,m)(n,i-1)(j,j-1,\ldots,n+1)(n,n+1)}$

$= U_{t1(k,k+1,\ldots,n)(m,m-1,\ldots,n+1)(n,n+1)(n,i-1)(n+1,j+1)(n,n+1)} =$

$= U_{t1(k,k+1,\ldots,n)(m,m-1,\ldots,n+1)(i-1,\ldots,n-1)(j+1,\ldots,n+2)(n,n+1)(n,n-1)(n+1,n+2)(n,n+1)}$

$= U_{t1(k,k+1,\ldots,n)(m,m-1,\ldots,n+1)(i-1,i,\ldots,n-1)(j+1,j,\ldots,n+2)(n-1,n+2)(n,n+1)}$

Thus, $U_{11(k,1)(m,p)} = U_{ii(n-1,n+2)(n,n+1)}$

Theorem App. H-1: If the 2r inequalities $p < k < \ldots < i \leqslant n < j < \ldots < m < q$ are satisfied and $\alpha = [n, N-n]$, then

(i) $U_{t1(p,q)(k,m)\ldots(i,j)} =$

$= U_{t1(p,\ldots,n)(q,\ldots,n+1)(k-1,\ldots,n-1)(m+1,\ldots,n+2)\ldots(i-r+1,\ldots,n-r+1)(j+r-1,\ldots,n+r)}A_r$

where $A_r \triangleq (n,n+1)(n,n-1,\ldots,n-r+1)(n+1,n+2,\ldots,n+r)(n,n+1)(n,\ldots,n-r+2)(n+1,\ldots,n+r-1)\cdots(n,n+1)(n,n-1)\times$
$\times(n+1,n+2)(n,n+1)$

(ii) $U_{11(p,q)(k,m)\ldots(i,j)} = (-1)^r/\binom{n}{r}$

Proof (i):

(i) It is true for r = 1 and 2. Assume it is true for $r-1 \triangleq s$.

$U_{t1(i_r,j_r)(i_s,j_s)\ldots(i,j)} =$

$= U_{t1(i_r,j_r)(i_s,i_s+1,\ldots,n)(j_s,j_s-1,\ldots,n+1)\ldots(i-r+2,\ldots,n-r+2)(j+r-2,\ldots,n+r-1)}A_{r-1} =$

$= U_{t1(i_r,\ldots,n)(j_r,\ldots,n+1)(n,n+1)(n+1,\ldots,j_r)(j_s,j_s-1,\ldots,n+1)(n,\ldots,i_r)(i_s,\ldots,n)\ldots}A_{r-1}$

$= U_{t1(i_r,\ldots,n)(j_r,\ldots,n+1)(i_s-1,\ldots,n-1)(j_s+1,\ldots,n+2)(n,n+1)(n-1,n)(n+2,n+1)(n-1,\ldots,i_s-1)(n+2,\ldots,j_s+1)\ldots}A_{r-1}$

$= U_{t1(i_r,\ldots,n)(j_r,\ldots,n+1)(i_s-1,\ldots,n-1)(j_s+1,\ldots,n+2)\ldots(i-r+1,\ldots,n-r+1)(j+r-1,\ldots,n+r)}\times$
$\times(n,n+1)(n,n-1)(n+1,n+2)(n-1,n-2)(n+2,n+3)\cdots(n-r+2,n-r+1)(n+r-1,n+r)A_{r-1}$ QED

Proof (ii): $U_{11A_1} = U_{11(n,n+1)} = -\frac{1}{n}$ (from A_a-1)

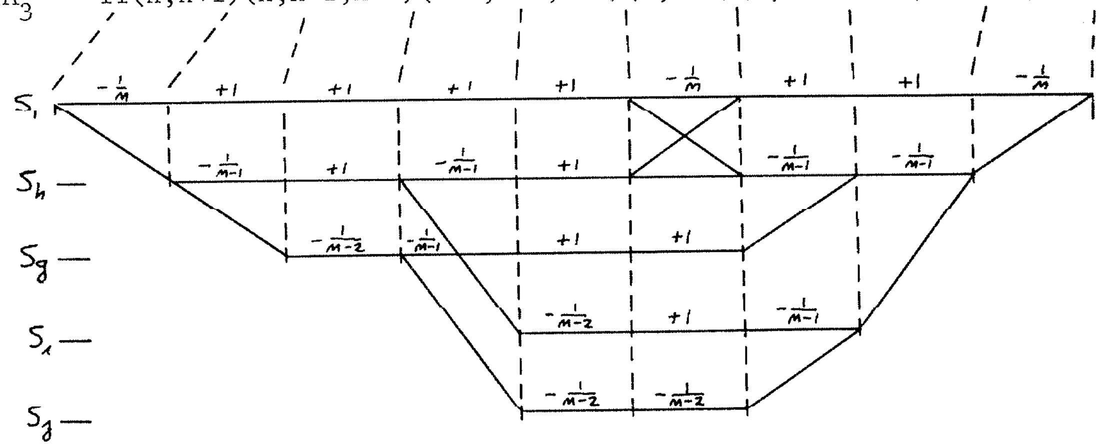
$U_{11A_2} = U_{11(n,n+1)(n,n-1)(n+1,n+2)(n,n+1)} =$

$$= \left(U_{11(n,n+1)}\right)^2 U_{11(n,n-1)(n+1,n+2)} + \left(U_{1h(n,n+1)}\right)^2 U_{hh(n,n-1)(n+1,n+2)}$$

$$= \left(\frac{1}{n}\right)^2 + \left(1-\frac{1}{n^2}\right)\left(\frac{1}{n-1}\right)^2 = \frac{1}{n^2} + \frac{(n+1)}{n^2(n-1)} = \frac{2n}{n^2(n-1)} = \frac{1}{\binom{n}{2}}$$

where $S_h = (n,n+1)S_1$; $S_h = \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & \cdots & n-1 & n+1 \\ \hline n & n+2 \\ \cline{1-2}\end{array}$

$U_{11A_3} = U_{11(n,n+1)(n,n-1,n-2)(n+1,n+2,n+3)(n,n+1)(n,n-1)(n+1,n+2)(n,n+1)}$



$S_g = \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & & n & n+1 \\ \hline n-1 & n+2 \\ \cline{1-2}\end{array}$; $S_x = \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & & n-1 & n+2 \\ \hline n & n+1 \\ \cline{1-2}\end{array}$; $S_y = \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & & n & n+2 \\ \hline n-1 & n+1 \\ \cline{1-2}\end{array}$

$$U_{11A_3} = -\frac{2}{n^2(n-1)} - \frac{(n-2)(n+1)(n-1)}{(n-1)^3(n-1)n^2} - \frac{2(n+1)(n-1)n(n-2)}{(n-2)(n-1)^2n^2(n-1)^2} + \frac{(n+1)(n-1)n(n-2)n(n-2)}{(n-2)^3n^2(n-1)^2(n-1)^2} =$$

$$= -\frac{1}{n(n-1)}\left[\frac{2}{n} + \frac{(n-2)(n+1)}{n(n-1)^2} + \frac{2(n+1)}{(n-1)^2} + \frac{n(n+1)}{(n-2)(n-1)^2}\right] = \frac{-6}{n(n-1)(n-2)} = -1\Big/\binom{n}{3}$$

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 $\frac{1}{\binom{n}{r}}$. Thus, the proof is incomplete.

Theorem App. H-2: If $p < k < \ldots < i \leq n < j < \ldots < m < g$ and $\alpha = [2^n, 1^{N-2n}]$,

then

(i) $U_{tf(p,q)(k,m)..(i,j)} = (-1)^{g+1-p}(-1)^{m+1-k}\cdots(-1)^{j+1-i}\, U^{\bar{\alpha}}_{\bar{f}1(p,g)(k,m)\cdots(i,j)}$

(ii) $U_{ff(p,q)(k,m)..(i,j)} = \frac{1}{\binom{n}{r}}$

Proof (i): the proof is the same as for THM 1i where the elementary transpositions which were +1 before are -1 which causes the factor of the parity of the (p,..,n) etc. to appear.

Proof (ii): the parity factor is squared for t = f so $U_{ff()()\cdot\cdot()} = U_{ffA_r}$ just as in THM 1ii. We follow the same steps as before except that all terms are positive.

Theorem App. H-3: If p, k, .., i are all different and $\leq n$, and j,..m,q are all different and $> n$, then

(i) if $\alpha = [n, N-n]$, $U_{11(p,q)(k,m)..(i,j)} = (-1)^r / \binom{m}{r}$

(ii) if $\alpha = [2^n, 1^{N-2n}]$, $U_{ff(p,q)(k,m)..(i,j)} = 1/\binom{m}{r}$

Proof: $U_{11\,(p,q)(k,m)\ldots(i,j)\tau} = U_{11(p,q)(k,m)\ldots(i,j)}$ if τ either involves letters $\leq n$ or if τ involves letters $> n$. But the similarity transformation τ rearranges the p, q,..,i among the transpositions. Thus, the p, k,.., i can be put in the order of THM-1. Similarly, if τ involves only elements $> n$ the q, m,..,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 τ involves only elements $\leq n$ or else elements $> n$ then

$$U^{\alpha}_{ff\tau\tau_r} = \zeta_\tau \frac{1}{\binom{m}{r}}$$

Proof:

$$U_{ff\tau\tau_r} = U_{ff\tau} U_{ff\tau_r} = \zeta_\tau U_{ff\tau_r} = \zeta_\tau \frac{1}{\binom{m}{r}}$$ QED

APPENDIX I:

Theorem: $U^{\bar{\alpha}}_{\bar{r}\bar{s}\tau} = \zeta_{\tau}\,\zeta_{\sigma_{rs}} U_{rs\tau}$

Proof: Consider $\tau = (k,k+1)$, then $U_{\bar{r}\bar{s}(k,k+1)} = [-\delta_{rs} + (1-\delta_{rs})]\,U_{rs(k,k+1)}$ from (Aa-9). We write this as $U_{\bar{r}\bar{s}(k,k+1)} = \zeta_{(k,k+1)}\,\zeta_{\sigma_{rs}}\,U_{rs(k,k+1)}$ since $\zeta_{\sigma_{rs}} = -1$ if $r \neq s$ and $U_{rs(k,k+1)} \neq 0$. $U_{\bar{r}\bar{s}\tau} = \sum_{i,j\cdots m} U_{\bar{r}\bar{i}\tau_1} U_{\bar{i}\bar{j}\tau_2} \cdots U_{\bar{m}\bar{s}\tau_m}$ where $\tau_1, \tau_2, \cdots, \tau_m$ are elementary transpositions. But $\zeta_{\sigma_{ri}}\,\zeta_{\sigma_{ij}} \cdots \zeta_{\sigma_{ms}} = \zeta_{\sigma_{rs}}$ and $\zeta_{\tau_1}\,\zeta_{\tau_2} \cdots \zeta_{\tau_m} = \zeta_{\tau}$. QED

REFERENCES:

Andersson, G., 1956, Acta Chem. Scand. 10, 623-628.

Brillouin, L., 1933, La Methode du Champ Self-Consistent, Actualites 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.

Condon, E. U. and G. H. Shortley, 1963, The Theory of Atomic Spectra, Cambridge University Press. This is a paperback version of the 1951 edition.

Daudel, R., H. Brion, and S. Odiot, 1955, J. Chem. Phys. 23, 2080-2083.

Edmiston, C. and K. Ruedenberg, 1963, Rev. Mod. Phys. 35, 457-465.

Erickson, R. A., 1953, Phys. Rev. 90, 779-785.

Fraga, S., and B. J. Ransil, 1961, J. Chem. Phys. 35, 1967-1977.

Griffith, J. S., 1961, The Theory of Transition-Metal Ions, Cambridge University Press.

Guillaud, C., A. Michel, J. Benard, and M. Fallot, 1944, Compt. Rend. 219, 58-60.

Hamermesh, M., 1962, Group Theory, Addison-Wesley.

Harris, F. E., and H. S. Taylor, 1963, J. Chem. Phys. 38, 2591-2596.

de Heer, J., 1962, J. Chem. Phys. 37, 2080-2083.

Itoh, T. and H. Hoshizumi, 1955, J. Phys. Soc. Japan 10, 201-207.

Koster, G. F., 1953, Quarterly Progress Reports, MIT Solid State and Molecular Theory Group, July 15, 1953, p. 37-38

Kotani, J., A. Amemiya, E. Ishiguro, T. Kimura, Table of Molecular Integrals, Maruzen Co., Ltd., Tokyo, 1955.

Lennard-Jones, J. E., 1949, Proc. Roy. Soc. A198, 1-13, 14-26.

Lieb, E., and D. Mattis, 1962, Phys. Rev. 125, 164-172.

Lomont, J. S., 1959, Applications of Finite Groups, Academic Press.

Löwdin, P. O., 1955 abc, Phys. Rev. 97, 1474-1489, 1490-1508, 1509-1520.

Löwdin, P. O., 1959, Adv. Chem. Phys. 2, 207-322.

Löwdin, P. O., 1962a, Rev. Mod. Phys. 34, 80-86.

Löwdin, P. O., 1962b, J. Appl. Phys. Supp 33, 251-280.

Merrifield, D. P., 1961, MIT Thesis. As quoted on p. 224 of Slater (1963).

Messiah, A., 1962, Quantum Mechanics Vol. II, Wiley, New York. This is in a paperback edition now.

Møller, C., and M. S. Plesset, 1934, Phys. Rev. 46, 618-622.

Pauncz, R., J. de Heer, and P. O. Lowdin, 1962ab, J. Chem. Phys. 36, 2247-2256, 2257-2265

Pratt, Jr., G. W., 1956, Phys. Rev. 102, 1303-1307.

Roothaan, C. C. J., 1951, Rev. Mod. Phys. 23, 69-89.

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.

Shull, H. and P. O. Lowdin, 1955, J. Chem. Phys. 23, 1362.

Sinanoğlu, O., 1961a, Proc. Roy. Soc. A260, 379-392.

Sinanoğlu, O., 1961b, Proc. Nat. Acad. Science 47, 1217-1226.

Sinanoğlu, O., 1961cd, Phys. Rev. 122, 491-492, 493-499.

Sinanoğlu, O., 1961e, J. Chem. Phys. 34, 1237-1240.

Sinanoğlu, O., 1962ab, J. Chem. Phys. 36, 706-717, 3198-3208.

Sinanoğlu, O., and D. F. Tuan, 1963, J. Chem. Phys. 38, 1740-1748.

Slater, J. C., 1960, Quantum Theory of Atomic Structure Vol. II, McGraw-Hill, New York.

Slater, J. C. 1963, *Quantum Theory of Molecules and Solids* Vol. I, McGraw-Hill, New York.

Wannier, G. H., 1937, Phys. Rev. 52, 191-197.

Watson, R. E., 1959, *Iron Series Hartree-Fock Calculations*, Technical Report No. 12, MIT Solid State and Molecular Theory Group.

Watson, R. E., 1960, Phys. Rev. 119, 1934-1939.

Watson, R. E., and A. J. Freeman, 1960, Phys. Rev. 120, 1125-1134.

Wigner, E., 1934, Phys. Rev. 46, 1002-1011.

Yoshimori, A., 1959, J. Phys. Soc. Japan, 14, 807-821.

Yoshizumi, H. and T. Itoh, 1955, J. Chem. Phys. 23, 412-413.

Ziman, J. M., 1960, Electrons and Phonons, Oxford, Clarendon Press.