# A. BFEECT OF INTERATOMIG <br> IWTERACTONS ON THE ZARO. BaNDUDTH HUEBMD HAMLTONLN 

# B. THEONY OF SIPMERECHANCE WHERMCTORS IN MBCNETIC IMSULATORS 

Dissertation for the Dagece of Ph. D. PACHORNS STATE UNVERBTY REM Silla ? 1975

## 

This is to certify that the
thesis entitled
A. Effect of Interatomic Interactions on the Zero-Bandwidth Hubbard Hamiltonian.
B. Theory of Superexchange Interactions In Magnetic Insulators.
presented by

Rem Sing Tu
has been accepted towards fulfillment
of the requirements for
Ph.D. degree in Physics

## $\underset{\text { Major professor }}{\text { How as A. Copland }}$

Date August 28, 1975

# ABSTRACT <br> A. EFFECT OF INTERATOMIC INTERACTIONS ON THE ZERO-BANDWIDTH HUBBARD HAMILTONIAN <br> B. THEORY OF SUPEREXCHANGE INTERACTIONS IN MAGNETIC INSULATORS 

By

Rem Sing Tu

The works of theoretical solid state physics can be divided roughly into two types of problems. The first type is to find the thermodynamic properties from a given model Hamiltonian. The second type is to find out an appropriate model Hamiltonian for a given problem or system. Part $A$ is of the first type, and part $B$ is of the second type. Therefore, the two main subjects of this thesis, unrelated as they may seem, can be regarded from a general theoretical point of view as being two different aspects of the same branch of physics.

In part $A$, we consider the linear-chain zerobandwidth Hubbard Hamiltonian with added nearest-neighbor interaction, with a magnetic field present. By the transfer matrix method, exact expressions for thermodynamic
quantities are obtained in simple closed form for the half-filled band. Recently, the half-filled band Hubbard model for the linear chain was proposed to explain the properties of the organic salt NMP-TCNQ. It was shown that the susceptibility $X$ versus temperature $T$ obtained from the Hubbard model disagreed in an essential way with the experiment. The experimental susceptibility rapidly becomes too small with increasing $T$, showing in particular what appears as a Curie-Weiss law with a moment appreciably reduced from the theoretical value. Since the nearest-neighbor Coulomb interaction causes a transition to a ground state of zero magnetic moment if large enough, it seemed possible that adding the Coulomb interaction might reduce the discrepancy between the experiment and theory. The answer we find is unfortunately negative. In part $B$, we study Anderson's theory of superexchange. It is thought that the exchange interaction between magnetic ions in magnetic insulator is described essentially by the Heisenberg Hamiltonian; also the exchange parameter $J$ is of 4 th order in the overlap between para- and dia-magnetic ions. However, the Wannier functions are not uniquely defined in the superexchange problem. Therefore, if one uses an arbitrary set to calculate $J$, one has to go to the 4 th order perturbation theory in order to exhaust all the terms of the 4 th order
in overlap. Anderson suggested that there exists "the exact" Wannier function which makes the perturbation theory converge rapidly. By using this set, the exchange parameter was presumed to come mainly from 1st and 2nd order perturbation terms, the 3 rd and 4 th order perturbation term being negligible. He proposed to use the Hartree-Fock method which put all the electrons in the magnetic ions spin parallel and doubly occupy the diamag-netic-ion orbitals to find the Wannier functions. However, his Hartree-Fock leads to a magnetic solution, that is, the spatial function of the spin-orbital depends on the spin. This is inconsistent with Anderson's requirement that they be nonmagnetic. In this work, we use a different variational approach, namely the thermal single determinantal approximation (TSDA) to substitute for his HartreeFock method. We first investigate a 3-site 4-electron linear cluster, and then generalize to a crystal. We find that there exists nonmagnetic solutions which make 3rd and 4 th order perturbation term vanish in both cases. The exchange parameter therefore comes only from 1st and 2nd order perturbation terms. Hence Anderson's idea is fulfilled. His "exact Wannier function" turn out to be the TSDA solution. In the 3 -site case, we also show the size of the contribution to $J$ from each order in perturbation
theory is very sensitive to the choice of Wannier functions. The generalization of this type of consideration to more realistic model containing more than one electron on a magnetic site is important and interesting.

# A. EFFECT OF INTERATOMIC INTERACTIONS ON THE ZERO-BANDWIDTH HUBBARD HAMILTONIAN <br> B. THEORY OF SUPEREXCHANGE INTERACTIONS IN MAGNETIC INSULATORS 

## By

Rem Sing Tu

## A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
Department of Physics
1975

## ACKNOWLEDGMENTS

It is a great pleasure to express my gratitude to Professor T. A. Kaplan for having given me the opportunity to collaborate closely with him in physics research. This collaboration has been a continuous stimulation for me to learn, especially because of Professor Kaplan's deep insight into the problems of concern and his way of enjoying teaching and doing research. This thesis would not have been possible without his continuous friendship, his time and effort.

I also learned a lot of solid state from other people in Michigan State. I would especially like to acknowledge many interesting and instructive hours spent in the company of Professor S. D. Mahanti and Dr. M. A. Barma.

I am greatly indebted to Professor T. O. Woodruff and the Physics Department of Michigan State University for providing me with a research assistantship. I also appreciate the encouragement from the committee and the fellow students at Michigan State.

## TABLE OF CONTENTS

Page
ACKNOWLEDGMENTS ..... ii
LIST OF FIGURES ..... iv
APPENDICES ..... v
PART A
I . INTRODUCTION ..... 2
II. SOLUTION OF THE PROBLEM ..... 8
REFERENCES ..... 17
PART B
I. INTRODUCTION ..... 20
II. THE 3-SITE MODEL ..... 26
III. THE ONE-ELECTRON SITES IN THE THERMAL SINGLE DETERMINANT APPROXIMATION ..... 35
IV. RESULTS OF THE 3-SITE MODEL ..... 43
V. GENERALIZATION OF THE 3-SITE
RESULTS TO A MANY-ATOM LATTICE ..... 45
VI. SUMMARY AND DISCUSSION ..... 53
REFERENCES ..... 54
Page
Figure
A1. Inverse susceptibility $X^{-1}$ versus temperature $T$ ..... 14
A2. Specific heat $C_{v}$ versus temperature $T$ ..... 16
B1. Three-site model ..... 27
B2. Linear-chain mode1 ..... 47
B3. Peroskite Structure $A B F_{3}$, showing only the B -ions (.) and the F-ion (0) ..... 47

## APPENDICES

APPENDI X Page
A. DIAGONALIZATION OF THE TRANSFER MATRIX ..... 56
B. WAVEFUNCTIONS OF 3-SITE 4-ELECTRON LINEAR CLUSTER ..... 59
C. MATRIX ELEMENT OF THE HAMILTONIAN
OPERATOR WITH RESPECT TO DETERMINANTAL WAVEFUNCTIONS ..... 61
D. VARIOUS ORDERS OF ENERGY CORRECTION ON THE PERTURBATION THEORY ..... 65
E. THE MATRIX ELEMENT FOR NEAREST-NEIGHBOR
HOPPING OF ONE ELECTRON ..... 67
F. DEGENERATE PERTURBATION APPROACH VIA
EFFECTIVE HAMILTONIAN ..... 68
G. THE EFFECTIVE HAMILTONIAN IN SECOND
ORDER PERTURBATION THEORY ..... 71

## PART A

EFFECT OF INTERATOMIC INTERACTIONS ON THE ZERO-BANDWIDTH HUBBARD HAMILTONIAN

## I. INTRODUCTION

It is well known that in a crystal the energy levels of the electrons are grouped in bands. We consider the case of a crystal of $N$ atoms and an average of $N$ electrons filling exactly half of one nondegenerate band; and we disregard the presence of all the other bands. To do so, we define an orthonormal complete set of $N$ Wannier functions for this band. The Wannier functions are localized at the lattice sites, i.e., each of them is appreciably different from zero only in the neighborhood of a lattice site. We then define operators $c_{i \sigma}^{+}$and $c_{i \sigma}$ which respectively create and destroy an electron in the Wannier function at site $i$ with $\operatorname{spin} \sigma$. The $c_{i \sigma}^{+}$'s and $c_{i \sigma}{ }^{\prime} s$ satisfy the usual fermion anticommutation relations and $n_{i \sigma}=c_{i \sigma}^{+} c_{i \sigma}$ is the number operator of site $i$ and spin $\sigma$. The Hubbard Hamiltonian ${ }^{1}$ is written in terms of these operators as

$$
\begin{equation*}
H=\sum_{i j \sigma} b_{i j} c_{i \sigma}^{+} c_{j \sigma}+U \sum_{i}^{\sum} n_{i \uparrow} n_{i \downarrow} \tag{1}
\end{equation*}
$$

The $b_{i j}\left(=b_{j i}^{*}\right)$ and $U$ are constant parameters and have precise physical meaning. For simplicity we take $b_{i j}=b$
for $i$ and $j$ nearest neighbors, and zero otherwise. b is called the transfer or hopping integral. $U$ is the intrasite Coulomb repulsion energy. The Hubbard model has found wide use in the theoretical description of electronic states in magnetic insulators. It was studied earlier than Hubbard did; e.g., des Cloizeaux ${ }^{2}$ discussed it in the late $50^{\prime}$ s. Hubbard and also Gutzwiller ${ }^{3}$ reintroduced it in 1963. Presumably Hubbard's name is attached because he was the only one who tried to give a serious derivation. His derivation leads to completely unsatisfactory behavior of the $b_{i j}$ as a function of distance between sites $i$ and $j$ as shown by N. Silva and T. A. Kaplan ${ }^{4}$. They present an essentially different theory which yields a satisfactory result. Although the derivation aspect is an important one, one can take this Hamiltonian phenomenologically as a model and study its physical predictions. That is, $H$ is given, and is to be studied as a function of the parameters $b_{i j}$ and $U$, as well as temperature. Some exact solutions can be found ${ }^{5-9}$. This is an interesting problem essentially because it is probably the simplest model such that special cases yield pure band-1ike behavior and atomic-like behavior; and of course the question of how electrons go from one type to the other has been of interest in solid state physics for many decades.

When $U=0$, we get a very familiar simple example of noninteracting fermions. $H$ can be written in terms of Bloch function occupation numbers as

$$
\begin{equation*}
\sum_{\underline{\mathbf{k}} \sigma}^{\varepsilon} \underline{\mathbf{k}}^{{ }^{n} \underline{\underline{k}}} \tag{2}
\end{equation*}
$$

$n_{\underline{k} \sigma}$ are defined as $n_{\underline{k} \sigma}=a_{\underline{k} \sigma}^{+}{ }^{a_{\underline{k} \sigma}}$, where $a_{\underline{k} \sigma}^{+}$and $a_{\underline{k} \sigma}$ are respectively the creation and annihilation operators for an electron in the Bloch function with crystal momentum $\underline{k}$ and spin $\sigma . a_{\underline{k} \sigma}^{+}$is related to the Wannier function creation operators by

$$
\begin{equation*}
\mathrm{a}_{\underline{k} \sigma}^{+}=\mathrm{N}^{-1 / 2} \sum_{\mathrm{i}} \mathrm{e}^{\mathrm{ik} \cdot \underline{R}_{i}} \mathrm{c}_{\mathrm{i} \sigma}^{+} \tag{3}
\end{equation*}
$$

$\varepsilon_{\underline{k}}$ are the one-electron energies of the band in question whose width is proportional to $\mathrm{b}_{\mathrm{i}}$ in general

$$
\begin{equation*}
\varepsilon_{\underline{k}}=\sum_{j} b_{i j} e^{i \underline{k} \cdot \underline{R}_{i j}} . \tag{4}
\end{equation*}
$$

Thus the energy eigenfunctions are single Slater-determinants with Bloch functions occupied in all possible ways. The calculation of all physical properties is tractable in the manner discussed in any elementary solid state textbook.

When $b_{i j}=0, H$ becomes

$$
\begin{equation*}
U \Sigma N_{i \uparrow} N_{i \downarrow} \tag{5}
\end{equation*}
$$

Two electrons with opposite spins occupying the same site interact with an energy $U$. They do not interact if they are on different sites. The complete set of eigenstates in this limit is given by the set of Slater determinants obtained by occupying Wannier functions in all possible ways. This was pointed out by Kaplan ${ }^{10}$ in 1968 and Kaplan and Argyres ${ }^{11}$ in 1970.

The half-filled band Hubbard model for the linear chain was recently proposed to explain the properties of the organic salt NMP-TCNQ. These organic solids are composed of two types of molecules, a donor and acceptor giving rise to the presence of unpaired electrons in the crystal. In this material the TCNQ molecules are presumed to be simple minus ions with the extra electron per TCNQ being the source of the observed electronic properties of the system. These molecules are large and flat, and stacked in linear arrays. These salts are highly anisotropic displaying a very pronounced one-dimensional behavior, the unpaired electron moving along the chains made up of the acceptor molecules. The one-dimensionality is clearly displayed by the conductivity measurements by Shchegolev ${ }^{13}$. It was shown ${ }^{14}$ that the susceptibility $x$ versus temperature $T$ obtained from the Hubbard model with $T$ independent parameters, $b$ and $U$, disagreed in an essential way with the experiment ${ }^{12}$. On the other hand, the
introduction of a T-dependence into $b$, phenomenologically, such that b increases appreciably with $T$ (with $U \simeq$ constant) can correct $X$ versus $T$. Since the physical interpretation becomes drastically modified ${ }^{15}$, it was concluded that understanding the physics beyond the Hubbard model is the essence of understanding NMP-TCNQ. The difficulty with the constant $b$ and $U$ is that when one uses values which fit $\chi$ to experiment at $\mathrm{T}=0$ the experimental susceptibility rapidly becomes too small with increasing $T$, showing in particular what appears as a Curie-Weiss law with a moment appreciably reduced from the theoretical value ${ }^{14}$. Since large enough nearest-neighbor Coulomb interaction $V\left(V>\frac{U}{2}\right)$ actually causes ${ }^{16}$ a transition to a ground state of zero magnetic moment, it seemed possible that values of $V$ smaller than this critical value but still appreciable might importantly reduce the discrepancy.

In this part, we consider a zero bandwidth modified Hubbard Hamiltonian in the half-filled chain. A summary of the results appeared earlier ${ }^{17,18}$. Our model Hamiltonian is in the form

$$
\begin{equation*}
\mathrm{H}=\mathrm{U} \underset{\mathrm{i}}{\sum} \mathrm{n}_{\mathrm{i} \uparrow} \mathrm{n}_{\mathrm{i} \downarrow}+\mathrm{V} \sum_{\mathrm{i}} \mathrm{n}_{\mathrm{i}} \mathrm{n}_{\mathrm{i}+1} \tag{6}
\end{equation*}
$$

The second term of the above Hamiltonian represents the intermolecular electron repulsion. It may be considered as a first step towards taking into account the long range
character of the electron-electron interaction. This Hamiltonian is related to Hubbard's by putting $\mathrm{b}_{\mathrm{ij}}=0$ in Hubbard case and adding the interatomic interaction. In the half-filled case, as we show below, the ground state configuration consists of one electron per site if $\mathrm{V}<\mathrm{U} / 2$, but it consists of alternating pairs if $V=\mathrm{U} / 2$. This effect was first pointed out by Bari ${ }^{16}$ who investigated the role of electron-lattice interactions in a very narrow half-filled band. His Hamiltonian which incorporates elec-tron-electron and electron-lattice interactions can be decoupled via a canonical transformation, and in one dimension reduces to the above equation.

## II. SOLUTION OF THE PROBLEM

In the grand canonical case, one must consider

$$
\begin{align*}
& H-\mu N=U \sum_{i}^{\sum} n_{i \uparrow} n_{i \downarrow}+V \sum_{i}\left(n_{i \uparrow}+n_{i \downarrow}\right)\left(n_{i+1 \uparrow}+n_{i+1 \downarrow}\right) \\
& -\mu \sum_{i}\left(n_{i \uparrow}+n_{i \downarrow}\right) \tag{1}
\end{align*}
$$

Here $\mu$ is the chemical potential and we assume periodic boundary conditions such that $n_{N+1 \sigma}=n_{1 \sigma}$ ( $\sigma$ equals either spin direction). Since we are only interested in the case $\sum_{i}\left(n_{i \uparrow}+n_{i \downarrow}\right)=N, \mu$ can be set by finding the average number of particles. This Hamiltonian is rewritten by Kaplan and Argyres ${ }^{11}$ in terms of "spins"

$$
\begin{equation*}
\mathrm{s}_{\mathrm{i}} \equiv \mathrm{n}_{\mathrm{i} \uparrow}+\mathrm{n}_{\mathrm{i} \downarrow}-1 \tag{2}
\end{equation*}
$$

We find that ${ }^{19}$

$$
\begin{align*}
H-\mu N & =\frac{U}{2} \sum_{i} S_{i}^{2}+V \sum_{i} S_{i} S_{i+1} \\
& -\left(\mu-\frac{U}{2}-2 V\right) \sum_{i} S_{i}+N(U-\mu) \tag{3}
\end{align*}
$$

We drop the last constant ( $\mathrm{S}_{\mathrm{i}}$-independent) term and add the magnetic field energy. We obtain, for a linear chain
in the magnetic field,

$$
\begin{align*}
H-\mu N & =\frac{U}{2} \sum_{i} S_{i}^{2}+V \sum_{i} S_{i} S_{i+1} \\
& +\left(\frac{U}{2}+2 V-\mu\right) \sum_{i} S_{i}+\frac{h}{2} \sum_{i}\left(n_{i \uparrow}-n_{i \downarrow}\right) \tag{4}
\end{align*}
$$

$h$ is $q \mu_{B} x$ the magnetic field.
This Hamiltonian is quite analogous to a spin-1 Ising Hamiltonian in a magnetic field. However, one must be careful since the "spin value" $S_{i}=0$ can occur in two ways. This arises from the fact that a singly occupied electron site is two-fold spin degenerate. This means that, if we treat the thermodynamics of $H$ from equation (5) of this section, we must take special care in counting the states. From this, the grand canonical partition function is the following:

$$
\begin{align*}
Z & =\sum_{n_{i \uparrow}}^{1} \sum_{\sum}^{1} \sum_{i \downarrow}=0 \\
& \left.\left.+y S_{i}+\text { z }_{i}^{2}-\mu M_{i}\right)\right] \tag{5}
\end{align*}
$$

where

$$
\begin{array}{lll}
x=\beta V & y=\beta\left(\frac{U}{2}+2 V-\mu\right) & z=\frac{U}{2} \\
\mu=\beta \frac{h}{2} & M_{i}=n_{i \uparrow}-n_{i \downarrow}
\end{array}
$$

Let us introduce the notation $\xi_{i}=1,2,3,4$ corresponding respectively to $\left(n_{i \uparrow} n_{i \downarrow}\right)=(1,1),(1,0),(0,1)$, $(0,0)$ and note that we can write the partition function
as following with the periodic boundary condition

where

$$
\begin{align*}
\overline{\mathrm{T}}_{\xi_{1} \xi_{2}} & =\exp \left\{-\left[x S_{1} S_{2}+\frac{y}{2}\left(S_{1}+S_{2}\right)\right.\right. \\
& \left.\left.+\frac{z}{2}\left(S_{1}^{2}+S_{2}^{2}\right)-\frac{u}{2}\left(M_{1}+M_{2}\right)\right]\right\}=T_{\xi_{2} \xi_{1}} \tag{8}
\end{align*}
$$

Let a $4 \times 4$ matrix $\bar{T}$ be so defined that its matrix alements are given by

$$
\begin{align*}
\bar{T}_{\xi \xi^{\prime}} & \equiv\langle\xi| T\left|\xi^{\prime}\right\rangle=\exp \left\{-\left[x S S^{\prime}+\frac{y}{2}\left(S+S^{\prime}\right)\right.\right. \\
& \left.\left.+\frac{z}{2}\left(S^{2}+S^{\prime 2}\right)-\frac{u}{2}\left(M+M^{\prime}\right)\right]\right\} . \tag{9}
\end{align*}
$$

Thus an explicit representation for $\bar{T}$ is
$\bar{T}=\left(\begin{array}{llll}e^{-(x+y+z)} & e^{-\frac{1}{2}(y+z-u)} & e^{-\frac{1}{2}(y+z+u)} & e^{x-z} \\ e^{-\frac{1}{2}(y+z-u)} & e^{u} & 1 & e^{-\frac{1}{2}(-y+z-u)} \\ e^{-\frac{1}{2}(y+z+u)} & 1 & e^{-u} & e^{-\frac{1}{2}(-y+z+u)} \\ e^{x-z} & e^{-\frac{1}{2}(-y+z-u)} & e^{-\frac{1}{2}(-y+z+u)} & e^{-(x-y+z)}\end{array}\right)$
From (T),

$$
\begin{equation*}
Z=\operatorname{Tr} \bar{T}^{N}=\lambda_{0}^{N}+\lambda_{1}^{N}+\lambda_{2}^{N}+\lambda_{3}^{N} \tag{11}
\end{equation*}
$$

where $\lambda_{i}$ are the eigenvalues of $\bar{T} . \bar{T}$ is called a "transfer matrix" ${ }^{20}$.

In the following, we limit ourselves to the halffilled band case. This condition on the number gives

$$
\left\langle_{i}^{\Sigma} S_{i}\right\rangle \equiv \frac{\partial}{\partial \beta \mu} \log Z=0
$$

Referring to eqs. (3) and (6), one sees that $y$ is formally an effective magnetic field acting on the $S_{i}$. Therefore, $y=0$, or $^{19}$

$$
\begin{equation*}
\mu=\frac{U}{2}+V \tag{12}
\end{equation*}
$$

The transfer matrix becomes
$T=\left[\begin{array}{llll}e^{-(x+z)} & e^{-\frac{1}{2}(z-u)} & e^{-\frac{1}{2}(z+u)} & e^{x-z} \\ e^{-\frac{1}{2}(z-u)} & e^{u} & 1 & e^{-\frac{1}{2}(z-u)} \\ e^{-\frac{1}{2}(z+u)} & 1 & e^{-u} & e^{-\frac{1}{2}(z+u)} \\ e^{x-z} & e^{-\frac{1}{2}(z-u)} & e^{-\frac{1}{2}(z+u)} & e^{-(x+z)}\end{array}\right]$
Immediately, one sees that an eigenvector of (13) is the transpose of ( $1,0,0,-1$ ) with eigenvalue $-2 e^{-z} \sin h x(\neq 0)$; furthermore, it can be seen that det $T=0$ so that the eigenvalues are obtained in simple closed form. For $N \rightarrow \infty$, the free energy per particle is $f=-k T \log \lambda_{m}$, where $\lambda_{m}$ is the eigenvalue of maximum magnitude, which is

$$
\begin{align*}
\lambda_{m} & =e^{-z} \cosh x+\cosh u+\left[\left(e^{-z} \cosh x\right.\right. \\
& \left.-\cosh u)^{2}+4 e^{-z} \cosh u\right] \tag{14}
\end{align*}
$$

All these results are shown in Appendix A. From this, the zero-field magnetic susceptibility is

$$
\begin{aligned}
x & =-\left(\frac{\partial^{2} f}{\partial h^{2}}\right)_{h=0}=k T\left[-\frac{1}{\lambda_{m}^{2}}\left(\frac{\partial \lambda_{m}}{\partial h}\right)^{2}+\frac{1}{\lambda_{m}} \frac{\partial^{2} \lambda_{m}}{\partial h^{2}}\right]_{h=0} \\
& =\frac{B}{e^{-2} \cosh x+1+\sqrt{\Delta}}\left[1+\frac{2 e^{-z}-e^{-z} \cosh x+1}{\sqrt{\Delta}}\right]
\end{aligned}
$$

where

$$
\Delta \equiv\left(e^{-z} \cosh x-1\right)^{2}+4 e^{-z}
$$

Similarly, the specific heat is given by

$$
\begin{equation*}
\frac{C v}{k}=-\left(\frac{\beta}{\lambda_{m}} \frac{\partial \lambda_{m}}{\partial \beta}\right)^{2}+\frac{\beta^{2}}{\lambda_{m}} \frac{\partial^{2} \lambda_{m}}{\partial \beta^{2}}=\beta^{2} \frac{d^{2}}{d \beta^{2}} \log \lambda_{m} \tag{16}
\end{equation*}
$$

To understand these results we investigate the ground state. Without the magnetic field, the electronic Hamiltonian in the case of the half-filled band can be rewritten in the form

$$
\begin{align*}
H_{e} & =\frac{U}{2} \sum_{i} S_{i}^{2}+V \sum_{i} S_{i} S_{i+1}+N V \\
& =\left(\frac{U}{2}-V\right) \sum_{i} S_{i}^{2}+\frac{V}{2} \sum_{i}\left(S_{i}+S_{i+1}\right)^{2}+N V \tag{17}
\end{align*}
$$

The variable $S_{i}$ can take on the values $-1,0,1$ and each summation in the above equation is a positive quantity. We first consider $U>2 V$; the minimum energy is obtained by simultaneously minimizing each summation in this case. This is obtained by requiring $S_{i}{ }^{2}=0$ and $S_{i}+S_{j}=0$ for
all i. Thus for $U>2 V$, a minimum energy eigenstate $\Psi_{1}$ has each site singly occupied. For the case $U<2 V$ one must maximize $\sum_{i} S_{i}^{2}$ while minimizing $\sum_{i}\left(S_{i}+S_{j}\right)^{2}$. This leads to the condition $S_{i}^{2}=1$ and $\left(S_{i}+S_{j}\right)^{2}=0$ for all $i$ and $j$. These conditions imply $S_{i}=1$ for all $i$ on sublattice $A$ and $S_{i}=-1$ for all $i$ on sublattice $B$. Thus, $n_{i \uparrow}+n_{i \downarrow}=2$ for $i$ on sublattice $A$ and $n_{i \uparrow}+n_{i \downarrow}=0$ for $i$ on sublattice B. A ground state $\Psi_{2}$ consists in this case of alternating empty and doubly occupied sites. For $U$ > 2V the ground-state energy is $N V$ and for $U<2 V$, it is $\frac{1}{2} N U$. This difference shows up strikingly in the zero-field susceptibility $x$ as shown in Fig. Al. As is seen, for $V / U \leq 0.5$, $X$ is very similar to the atomic limit of the Hubbard model $(V / U=0)$. But for $V / U>0.5$, a marked change occurs at low $T$, since in this case $X \rightarrow 0$ (rather than $\infty$ ) as $T=0$, due to the fact that $\left\langle\Psi_{2}\right| S^{2}\left|\Psi_{2}\right\rangle=0$ (whereas $\left.<\Psi_{1}\left|S^{2}\right| \Psi_{1}\right\rangle$ is not zero for many of the $2^{N}$ degenerate state $\Psi_{1}$ ). In fact, it is easy to see from (14) that asymptotically at low T

$$
\begin{equation*}
x \simeq 4 \beta \exp \left[-\beta\left(2 \mathrm{~V}-\frac{\mathrm{U}}{2}\right)\right], 2 \mathrm{~V}>\mathrm{U} \tag{18}
\end{equation*}
$$

so that $X$ is exponentially small at low $T$ when $2 V>U$. In the other case, equation (14) gives the Curie law

$$
\begin{equation*}
X \simeq \beta \quad 2 V \leq U \tag{19}
\end{equation*}
$$



Figure Al. Inverse susceptibility $\chi^{-1}$ versus temperature $T$.

In Fig. A2, we see that when $V$ turns on, the peak of the specific heat becomes narrower and it moves to lower $T$; the lowest value occurs when $2 \mathrm{~V}=\mathrm{U}$. As V continues to increase, the peak moves back to higher $T$, with a considerable additional sharpening. Clearly, the area under the large $-V / U$ peak is appreciably greater than that for the small -V/U peaks; this is consistent with the easily proven facts,

$$
\int_{0}^{\infty} \frac{C_{v}}{k T} \quad d T=\left\{\begin{array}{cc}
\log 2 & V \leq U / 2  \tag{20}\\
2 \log 2 & V \geq U / 2
\end{array}\right.
$$

It is interesting to note that the correct low-T behavior, equation (18) is rather different from what one might have guessed from continuity given equation (19), namely susceptibility $\sim \beta \exp [-\beta(V-U / 2)]$ for $2 V \geq U$. Finally, it is the insensitivity of $X$ vs $T$ for $V<U / 2$ that shows that the addition of $V$ to the zero-bandwidth Hubbard model will not significantly improve the theory in connection with the experimental $X$ found for NMP-TCNQ. That is, the experimental $X^{-1}$ vs. $T$ is nearly straight over $T \sim 30^{\circ} \mathrm{K}$ to $T \sim 200^{\circ} \mathrm{K}$ with a slope corresponding to $\sim 3 k T$ rather than either the kT we find at $\mathrm{kT} \rightarrow 0$ or the maximum of $\check{2} 2 \mathrm{kT}$ at higher T . Furthermore, the insensitivity of the theoretical $x$ to the addition of hopping terms to the zero-bandwidth Hubbard model in the Curie Weiss region ${ }^{14}$ suggests that perhaps adding $b$ and $V$ also will not help.


Figure A2. Specific heat $C_{v}$ versus temperature $T$

## REFERENCES

1. J. Hubbard, Proc. Roy. Soc. (London) A276, 238 (1963); A277, 237 (1964); A281, 401 (1964).
2. J. des Cloizeaux, J. Phys. Radium 20, 606 (1959).
3. M. C. Gutzwiller, Phys. Rev. Letters 10, 159 (1963).
4. N. P. Silva and T. A. Kaplan, Bull. Amer. Phys. Soc. 18, 450 (1973); N. P. Silva and T. A. Kaplan, AIP Conf. Proc. No. 19, Magnetism and Magnetic Materials (1973) (p. 656)
5. E. H. Lieb and F. Y. Wu, Phys. Rev. Letters 20, 1445 (1968).
6. C. N. Yang, Phys. Rev. Letters 19, 1312 (1967)
7. M. Takahashi, Progr. Theoret. Phys. 43, 1619 (1970).
8. A. A. Orchinnikov. Sov. Phys. JETP 30, 1160 (1970).
9. H. Shiba and P. A. Pincus, Phys. Rev. B5 1966 (1972).
10. T. A. Kaplan, Bull. Amer. Phys. Soc. 13, 386 (1968).
11. T. A. Kaplan and P. N. Argyres, Phys. Rev. B1, 2457 (1970).
12. A. J. Epstein, S. Etmemad, A. F. Garito and A. J. Heeger, Phys. Rev. B5, 952 (1972).
13. I. F. Shchegoler, Phys. Stat. Sol. (a) 12, 9 (1972).
14. D. Cabib and T. A. Kaplan, AIP Conf Proc. No. 10, Magnetism and Magnetic Materials (1972).
15. T. A. Kaplan, Bull. Amer. Phys. Soc. 18, 399 (1973).
16. R. A. Bari, Phys. Rev. B3, 2662 (1971)
17. R. S. Tu and T. A. Kaplan, Phys. Stat. Sol. (b) 63, 659 (1974).
18. Essentially similar results were found independently by G. Beni and P. Pincus, Phys. Rev. B9, 2963 (1974).
19. This differs from the first equation which appears in ref. 17 because that equation is derived from ref. 11 while it is derived from ref. 16 here. The difference only comes from a different choice of $\mathrm{b}_{\mathrm{ii}}$, which merely causes a constant shift in the chemical potential.
20. See, e.g. K. Huang, Statistical Mechanics, New York, Wiley, 1963.

## PART B

## THEORY OF SUPEREXCHANGE INTERACTIONS <br> IN MAGNETIC INSULATORS

## I. INTRODUCTION

There is a large class of materials called magnetic insulators; some examples: $\mathrm{MnO}, \mathrm{EuO}, \mathrm{MnF}_{2}, \mathrm{KMnF}_{3}$, $\mathrm{KNiF}_{3}, \mathrm{Fe}_{2} \mathrm{O}_{3}, \mathrm{Cr}_{2} \mathrm{O}_{3}$. They have an extremely low electrical conductivity; they are presumed to have localized electronic magnetic moments on the metal ions, the nonmetal ions being diamagnetic. The magnetism in these crystals arises either from incomplete 3d- or $4 f$ - electron shells. The outer s-electrons are always importantly involved in the binding energy of the system. The s-electrons from the metal atom are pictured as being transferred to fluorine or oxygen atoms; e.g. in $\mathrm{MnF}_{2}$, the manganese (neutral Mn is $3 \mathrm{~d}^{5} 4 \mathrm{~s}^{2}$ ) are considered $\mathrm{Mn}^{2+}$ while each fluorine is $\mathrm{F}^{-}$; in Mn0 we presumably have $\mathrm{Mn}^{2+}$ and $0^{2-}$ ions. Furthermore, the closed shell ion $\mathrm{F}^{-}$or $0^{--}$(both $1 \mathrm{~s}^{2} 2 \mathrm{~s}^{2} 2 \mathrm{p}^{6}$ ) are pictured as having much larger ionic radii than the positively charged cation. Hence, the 3 d electrons on the $\mathrm{Mn}^{2+}$ are prevented from overlapping very strongly with their neighboring Mn 3d-electrons, and therefore one might treat the overlap as small. At high temperature, the atomic moments behave paramagnetically. But at lower temperature, they undergo a phase transition to a magnetically ordered phase. The
critical temperatures $T_{c}$ range from $\sim 1^{\circ} \mathrm{K}$ to $\simeq 1000^{\circ} \mathrm{K} . \quad \mathrm{T}_{\mathrm{c}}$ has a different name for different magnetic ordering. For ferromagnetic crystals, the magnetic moments of the constituent magnetic ion align parallel to one another, and $T_{c}$ is called the Curie temperature. Substances of these kind are $\mathrm{Fe}, \mathrm{Co}, \mathrm{CrO}_{2}, \mathrm{Eu} 0, \mathrm{GdBr}_{3}$ ( Fe and Co being metals, however). In an antiferromagnet the spins are ordered in an antiparallel arrangement with zero net moment at temperature below the ordering or Neel temperature. For example, Mn0, Fe 0 , and Cr are antiferromagnets ( Cr being a metal). If however, one of the magnetizations is stronger than the other, it is to be expected that the difference between the two magnetizations will give rise to a strong magnetism. These substances are called ferrimagnetic, such as $\mathrm{Fe}_{3} 0_{4}$. Other types of ordering are also observed, e.g. spiral or helical ordering. ${ }^{1,2}$

Clearly the existence of such a $T_{c}$ implies interactions between the atomic moments and it is the purpose of this thesis to contribute to the theory of these interactions. If $T_{c}>1^{\circ} \mathrm{K}$ then the electron-electron and electron-core Coulomb interaction plus the electronic kinetic energy are generally accepted as giving the important source in the Hamiltonian of these interactions. Also essential is the fermion nature of electrons. The effective interactions that arise in this way are called exchange interactions.

They are generally thought to be essentially of the form

$$
\begin{equation*}
-\sum_{i j} J_{i j} \vec{S}_{i} \cdot \vec{S}_{j} \tag{1}
\end{equation*}
$$

where $\vec{S}_{i}$ is the spin of the $i^{\text {th }}$ ion and $J_{i j}$ is the exchange parameter for ion $i$ and $j$. If the interaction $J_{i j}$ involves only the overlap of free-ion 3 d -states associated respectively with ions $i$ and $j$, it is called direct exchange. If the exchange couplings exist between ions separated by one or several diamagnetic groups, Kramers ${ }^{3}$ pointed out that the magnetic ions could cause spin-dependent perturbations in the wave functions of intervening ions thereby transmitting the exchange effect over large distances. The latter effect is called superexchange. One of the stumbling blocks in the theory of superexchange is the derivation of the Heisenberg Hamiltonian(1). The first formulation is in terms of the nonorthogonal atomic orbitals. The second formulation uses orthogonal "atomic" orbitals, namely Wannier functions. Because of the considerable mathematical advantage when dealing with a macroscopic system, we follow Anderson's approach which uses Wannier functions. However, the Wannier functions are not uniquely defined in the superexchange problem. Therefore, the convergence of the perturbation series will depend on the choice of the Wannier functions. A major presumption of Anderson's theory of superexchange ${ }^{4}$ is that the perturbation theory defined in terms of
"the exact" Wannier functions would be rapidly convergent. (The small perturbation parameter is the nearest neighbor overlap $\Delta$ ). His Hartree-Fock (HF) definition of thesc exact Wannier states was shown by Silva and Kaplan ${ }^{5}$ to be unsatisfactory; in particular these states do not satisfy his requirement that they be nonmagnetic (nonmagnetic wavefunctions are by definition products $W(r) \sigma$ of spatial and spin functions in which the spatial functions are independent of spin $\sigma$ ). In the present paper, we nevertheless investigate the presumption of rapidity of convergence using a different variational definition, namely the nonmagnetic localized solutions in the thermal single determinant approximation (TSDA). 6,7,8

Our investigation is made first within a previously studied 4-electron 3 -orbital 3 -site ${ }^{9}$ model of superexchange. Then we generalize to a 3-dimensional crystal. The function space is defined to have atomic functions $a_{i}$ centered on magnetic atoms and $b_{i}$ centered on diamagnetic atoms (with one orbital at each atom); $a_{i}$ and $b_{i}$ are real; since they are presumed to be free-atom (or ion) states, they are not orthogonal (interatomic overlap integrals are nonzero). The states $a_{i}$ are related, for the crystal, by a lattice translation operation $L$, and similarly, for the $b_{i}$. We assume that each $a_{i}$ and $b_{i}$ is inversion invariant about its respective atomic site. We also assume nearest neighbor overlap, the next nearest
neighbor and more distant overlap being taken to be zero. In case of 3 sites, we have the inversion operation $I$ through the central site instead of the lattice translation operation L. This situation is shown in Fig. B1. Wannier functions $A_{i}$ and $B_{i}$ are constructed in this space. We require the constructed Wannier function to be real, orthogonal and to satisfy $A_{i} \rightarrow a_{i} ; B_{i} \rightarrow b_{i}$ as the overlap $\Delta \rightarrow 0$. The Wannier functions so constructed turn out to be not unique. Adding the spin to these Wannier functions, we obtain our complete set of orthonormal one-electron states. Occupying these one-electron states with electrons, we get the many-electron states, namely the Slater determinants. The model Hamiltonian is defined as the projection onto these many-electron states of the usual Hamiltonian containing electronic kinetic energy, electron-electron, and electronnucleon Coulomb interactions.

The Hamiltonian is divided into two parts. $H_{o}$, the unperturbed part is diagonal in the basis defined in the previous paragraph, and is of zero order in the overlap $\Delta$. The perturbation part is $V \equiv H-H_{o}$ which is of first order in the overlap. In the unperturbed ground state, there is by definition one electron on each magnetic site and two electrons on each diagmatic site. Because the Wannier functions as defined above are not unique, this division of $H_{0}$ and $V$ is correspondingly not unique. As is well known, for the three-site model, the exchange, i.e., the splitting between
the lowest singlet and triplet is $0\left(\Delta^{4}\right)$. Therefore, it has contributions through $4^{\text {th }}$ order perturbation theory. Each order is very sensitive to the choice of Wannier functions, as will be shown. Although the sum of all perturbation terms through fourth order is independent of this choice. We show however, that with the TSDA choice, the total energy to $0\left(\Delta^{4}\right)$ is given exactly by the $1^{\text {st }}$ and $2^{\text {nd }}$ order perturbation theory ${ }^{10}$ both in the 3 -site case and in a 3-dimensional crystal.

Previous attempts at calculating the exchange parameter $J$ within the general low-order perturbation approach have failed to give agreement with the experiment. In the most recent and elaborate attempt (for $\mathrm{KMnF}_{3}$ ), Fuchikami's ${ }^{11}$ straightforward perturbation calculation led to a J which is an order of magnitude too smal1 ${ }^{12}$. But her ${ }^{11}$ Wannier functions were apparently chosen arbitrarily and she considered perturbation theory only through second order. Hence, our finding of extraordinary sensitivity to this choice the relative size of $1^{s t}$ through $4^{\text {th }}$ order perturbation terms (we give simple example where the exchange constant comes only from fourth order perturbation theory!) suggest that she might not have obtained all the leading contributions to J.

## II. THE 3-SITE MODEL

The system we are considering in this section is a single linear cluster with a diamagnetic atom in the center and magnetic atoms on each end. It is a four-electron system. The Hamiltonian of this system is the usual kinetic energy of electrons, the Coulomb interaction between electron and electron, and between electron and nuclei:

$$
\begin{equation*}
H=\sum_{i=1}^{4} h(i)+\frac{1}{2} \sum_{i j=1}^{4} v(i, j) \tag{1}
\end{equation*}
$$

Here $h_{i}$ is the kinetic energy of the "i" electron plus its interaction with the atomic cores, and $v(i, j)$ is the Coulomb interaction between electrons $i$ and $j$.

Imagine three atomic orbitals $a_{0}, b_{0}, a_{1}$ localized at three atomic sites as indicated in Fig. Bl. We assume that the central one, $b_{0}$ is invariant under $I$, inversion through the central site $I b_{0}=b_{0}$, and $I a_{0}=a_{1}$. Among these orbitals, we assume they are real, and have only nearest neighbor overlap $\Delta$, i.e.

$$
\begin{equation*}
\left\langle a_{0} \mid b_{0}\right\rangle \equiv \int a_{0} b_{0} d v=\Delta, \quad a_{0} a_{1}=0 \tag{2}
\end{equation*}
$$

From these one can construct the following set of orthogonal functions:


$$
\begin{align*}
& A_{0}=C\left[a_{0}+\mu b_{0}+\nu a_{1}\right] \\
& \underline{B_{0}}=D\left[b_{0}+\gamma\left(a_{0}+a_{1}\right)\right]  \tag{3}\\
& \underline{A_{1}}=C\left[a_{1}+\mu b_{0}+v a_{0}\right]
\end{align*}
$$

Here $C$ and $D$ are normalization factors, while $\mu, \nu, \gamma$ are parameters assumed to be real and of $0(\Delta) . A_{0}, B_{0}, A_{1}$ satisfy the same symmetry properties as $a_{0}, b_{0}, a_{1}, i . e .$, $I B_{0}=B_{0}, I A_{0}=\underline{A_{1}}$. These "Wanner functions" will have the properties $\underline{A}_{0} \rightarrow a_{0}, \underline{B_{0}} \rightarrow b_{0}, \underline{A}_{1} \rightarrow a_{1}$ as $\Delta \rightarrow 0$. The expressions for $\mu, \nu, \gamma$ dictated by orthogonality are somewhat complicated; to leading order in $\Delta$ they give

$$
\begin{align*}
& \nu=-\left(\frac{\mu^{2}}{2}+\mu \Delta\right) \\
& \gamma=-(\mu+\Delta) . \tag{4}
\end{align*}
$$

$A_{0}, B_{0}, A_{1}$ are not uniquely defined because the set

$$
\begin{align*}
& A_{0}=\frac{1}{1+\lambda^{2} / 2}\left(A_{0}+\lambda B_{0}-\frac{\lambda^{2}}{2} A_{1}\right) \\
& B_{0}=\frac{1-\lambda^{2} / 2}{1+\lambda^{2} / 2}\left(B_{0}-\frac{\lambda}{1-\lambda^{2} / 2}\left(A_{0}+A_{1}\right)\right]  \tag{5}\\
& A_{1}=\frac{1}{1+\lambda^{2} / 2}\left(A_{1}+\lambda B_{0}-\frac{\lambda^{2}}{2} A_{0}\right)
\end{align*}
$$

are orthonormal to $0(\Delta)$ provided $\lambda \rightarrow 0$ as $\Delta \rightarrow 0$, and still satisfy the other symmetry conditions. After adding the spin to these "Wanner function", we have a six dimensional space.

Occupying the six spin orbitals $A_{0} \sigma, B_{0} \sigma, A_{1} \sigma$ $(\sigma=\uparrow \downarrow)$, one obtains 15 four-electron determinants. From these we constructed those linear combinations $\Phi_{1} \ldots \Phi_{15}$ which are eigenfunctions of the total spin operators $\vec{S}^{2}$, $S_{z}$ and the inversion $I$. The Hamiltonian matrix $\left\langle\Phi_{i}\right| H \mid \Phi_{j}>$ is then reduced into block form. This is shown in Appendix B. For small enough $\Delta$, the lowest singlet can be shown to come from a $4 \times 4$ submatrix connecting the states

$$
\begin{align*}
\Phi_{1}= & \frac{1}{\sqrt{2}}\left[\left|A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right\rangle-\left|A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow A_{1} \uparrow\right\rangle\right] \\
\Phi_{2}= & \frac{1}{2}\left[\left|A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right\rangle-\left|A_{0} \uparrow A_{0} \downarrow B_{0} \downarrow A_{1} \uparrow\right\rangle\right. \\
& \left.+\left|A_{0} \uparrow B_{0} \downarrow A_{1} \uparrow A_{1} \downarrow\right\rangle-\left|A_{0} \downarrow B_{0} \uparrow A_{1} \uparrow A_{1} \downarrow\right\rangle\right]  \tag{6}\\
\Phi_{3}= & \frac{1}{\sqrt{2}}\left[\left|A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow\right\rangle+\left|B_{0} \uparrow B_{0} \downarrow A_{1} \uparrow A_{1} \downarrow\right\rangle\right] \\
\Phi_{4}= & \left|A_{0} \uparrow A_{0} \downarrow A_{1} \uparrow A_{1} \downarrow\right\rangle .
\end{align*}
$$

Similarly, the lowest triplet comes from a $2 \times 2$ submatrix connecting the states

$$
\begin{align*}
\Phi_{5}= & \frac{1}{\sqrt{2}}\left[\left|A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right\rangle+\left|A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow A_{1} \uparrow\right\rangle\right] \\
\Phi_{6}= & \frac{1}{2}\left[\left|A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right\rangle+\left|A_{0} \uparrow A_{0} \downarrow B_{0} \downarrow A_{1} \uparrow\right\rangle\right. \\
& \left.+\left|A_{0} \uparrow B_{0} \downarrow A_{1} \uparrow A_{1} \downarrow\right\rangle+\left|A_{0} \downarrow B_{0} \uparrow A_{1} \uparrow A_{1} \downarrow\right\rangle\right] . \tag{7}
\end{align*}
$$

Here | ... > is the normalized antisymmetrized product of
spin orbitals indicated. The matrix elements $H_{i j}=\left\langle\Phi_{i}\right| H\left|\Phi_{j}\right\rangle$ can be classified according to their order of magnitude as follows:
(1) $0\left(\Delta^{\circ}\right)$

$$
\begin{aligned}
H_{11}= & \left\langle A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right| H \mid A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow>-\left\langle A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right. \\
& \left.|H| A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow A_{1} \uparrow\right\rangle \\
H_{22}= & \left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right| H \mid A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow>-\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right. \\
& |H| A_{0} \downarrow B_{0} \uparrow A_{1} \uparrow A_{1} \downarrow>+\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right| H\left|A_{0} \uparrow B_{0} \downarrow A_{1} \uparrow A_{1} \downarrow\right\rangle \\
& -\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right| H\left|A_{0} \uparrow A_{0} \downarrow B_{0} \downarrow A_{1} \uparrow\right\rangle \\
& \\
H_{33}= & \left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow\right| H \mid A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow>+\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow\right. \\
& |H| B_{0} \uparrow B_{0} \downarrow A_{1} \uparrow A_{1} \downarrow>
\end{aligned}
$$

$$
\begin{equation*}
H_{44}=\left\langle A_{0} \uparrow A_{0} \downarrow A_{1} \uparrow A_{1} \downarrow\right| H\left|A_{0} \uparrow A_{0} \downarrow A_{1} \uparrow A_{1} \downarrow\right\rangle \tag{8}
\end{equation*}
$$

$$
H_{55}=\left\langle A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right| H \mid A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow>+\left\langle A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right.
$$

$$
\left.|H| A_{0}+B_{0} \uparrow B_{0} \downarrow A_{1} \uparrow\right\rangle
$$

$$
H_{66}=\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right| H\left|A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right\rangle+\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right.
$$

$$
|H| A_{0} \downarrow B_{0} \uparrow A_{1} \uparrow A_{1} \downarrow>+\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right| H \mid A_{0} \uparrow B_{0} \downarrow A_{1} \uparrow
$$

$$
A_{1} \downarrow>+\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right| H\left|A_{0} \uparrow A_{0} \downarrow B_{0} \downarrow A_{1} \uparrow\right\rangle
$$

(2) $0(\Delta)$

$$
\begin{align*}
H_{12}= & \sqrt{2}\left[<A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow|H| A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow>-<A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right. \\
& \left.|H| A_{0} \downarrow B_{0} \uparrow A_{1} \uparrow A_{1} \downarrow>\right] \\
H_{23}= & \sqrt{2}\left[<A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow|H| A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow>+<A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow\right. \\
& \left.|H| B_{0} \uparrow B_{0} \downarrow A_{1} \uparrow A_{1} \downarrow>\right] \\
H_{24}= & 2<A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow|H| A_{0} \uparrow A_{0} \downarrow A_{1} \uparrow A_{1} \downarrow>  \tag{9}\\
H_{56}= & \sqrt{2}\left[<A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow|H| A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow>+<A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow\right. \\
& \left.A_{1} \downarrow|H| A_{0} \downarrow B_{0} \uparrow A_{1} \uparrow A_{1} \downarrow>\right]
\end{align*}
$$

(3) $0\left(\Delta^{2}\right)$

$$
\begin{align*}
& H_{13}=2\left\langle A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right| H \mid A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow A_{1} \downarrow> \\
& H_{14}=\sqrt{2}\left\langle A_{0} \uparrow B_{0} \uparrow B_{0} \downarrow A_{1} \downarrow\right| H \mid A_{0} \uparrow A_{0} \downarrow A_{1} \uparrow A_{1} \downarrow>  \tag{10}\\
& H_{34}=\sqrt{2}\left\langle A_{0} \uparrow A_{0} \downarrow B_{0} \uparrow B_{0} \downarrow\right| H \mid A_{0} \uparrow A_{0} \downarrow A_{1} \uparrow A_{1} \downarrow>
\end{align*}
$$

The detailed calculation of each matrix element uses the technique given in Appendix $C$. Each diagonal submatrix of $H_{i j}$ is found to have nondegeneracy among the diagonal elements. Because of our requirement that the ground state must be singly occupied on the magnetic sites
in the zero order overlap limit, we obtain the inequalities

$$
\begin{align*}
E_{3}^{\circ}-E_{1}{ }^{\circ} & =\left\langle A_{0} A_{0}\right| v\left|A_{0} A_{0}\right\rangle-\left\langle A_{0} A_{1}\right| v\left|A_{0} A_{1}\right\rangle>0 \\
E_{2}^{\circ}-E_{1}{ }^{\circ} & \left.=\left\langle A_{0}\right| h\left|A_{0}\right\rangle-<B_{0}|h| B_{0}\right\rangle+\left\langle A_{0} A_{0}\right| v\left|A_{0} A_{0}\right\rangle \\
& -\left\langle B_{0} B_{0}\right| v\left|B_{0} B_{0}\right\rangle+\left\langle A_{0} A_{1}\right| v\left|A_{0} A_{1}\right\rangle \\
& -<A_{0} B_{0}|v| A_{0} B_{0} \gg 0 .  \tag{11}\\
E_{4}^{\circ}-E_{1}{ }^{\circ} & \left.=2\left[\left\langle A_{0}\right| h\left|A_{0}\right\rangle-<B_{0}|h| B_{0}\right\rangle\right]+2\left\langle A_{0} A_{0}\right| v\left|A_{0} A_{0}\right\rangle \\
& \left.\left.-<B_{0} B_{0}|v| B_{0} B_{0}\right\rangle+3<A_{0} A_{1}|v| A_{0} A_{1}\right\rangle \\
& -4<A_{0} B_{0}|v| A_{0} B_{0} \gg 0
\end{align*}
$$

where

$$
\begin{align*}
\langle a| h|b\rangle= & \int a^{*}(1) h(1) b(1) d^{3} v_{1} \\
\langle a b| v|f g\rangle= & \int a(1) * b(2) * v(1,2) f(1) \\
& g(2) d v_{1} d v_{2} . \tag{12}
\end{align*}
$$

Hence, we apply nondegenerate perturbation theory up to $4^{\text {th }}$ order to calculate the singlet-triplet energy difference keeping all terms in the energy through $0\left(\Delta^{4}\right)$. We get

$$
\begin{equation*}
\Delta E_{s t}=\sum_{i=1}^{4} \Delta E_{s t}^{(i)}=\sum_{i=1}^{4}\left(\underset{s}{(i)}-E_{t}^{(i)}\right) \tag{13}
\end{equation*}
$$

where $i$ is the order or perturbation theory. The formulas for various order corrections to the energy from perturbation theory are displayed in Appendix $D$. Up to $0\left(\Delta^{4}\right)$, each $\Delta E_{s t}^{(i)}$ can be written as
$E(1)=2\left\langle A_{0} A_{1}\right| v\left|A_{1} A_{0}\right\rangle$

$$
E_{s t}^{(4)}=-\left(\frac{H_{12}}{H_{22^{-H}}{ }^{-H}}\right)^{2}\left(\frac{H_{23}^{2}}{H_{33^{-H}} 11}+\frac{H^{2} 24}{H_{44^{-H}}{ }^{-H_{11}}}\right)
$$

Each $E_{s t}^{(i)}$ is a function of $\lambda$ because of (5) but the total energy difference to $0\left(\Delta^{4}\right)$ is independent of $\lambda$.

$$
\begin{align*}
\Delta E_{s t} & =2\left\langle\underline{A}_{0} \underline{A}_{1}\right| v\left|\underline{A}_{1} \underline{A}_{0}\right\rangle+\frac{4\left\langle\underline{A}_{0} \underline{A}_{0}\right| v\left|\underline{B}_{0} \underline{A}_{1}\right\rangle}{w_{21}} p-\frac{z}{w_{21}^{2}} p^{2} \\
& -\frac{4}{w_{31}}\left(x+\frac{p q}{w_{21}}\right)^{2}-\frac{2}{w_{41}}\left(y+\frac{2 r p}{w_{21}}\right)^{2} \tag{15}
\end{align*}
$$

where $W_{m n}=E_{m}^{(0)}-E_{n}^{(0)}$ is obtained from $E_{m}^{(0)}-E_{n}^{(0)}$ given in (11) with the replacements $A_{0} \rightarrow \underline{A}_{0}, B \rightarrow \underline{B}_{0}, A_{1} \rightarrow \underline{A}_{1}$, and $p=\left\langle\underline{A}_{0}\right| h\left|\underline{B}_{0}\right\rangle+\left\langle\underline{A}_{0} \underline{A}_{0}\right| v\left|\underline{B}_{0} \underline{A}_{0}\right\rangle+\left\langle\underline{A}_{0} \underline{B}_{0}\right| v\left|\underline{B}_{0} \underline{B}_{0}\right\rangle$ $+\left\langle\underline{A}_{0} \underline{A}_{1}\right| v\left|\underline{B}_{0} \underline{A}_{1}\right\rangle$
$q=\left\langle\underline{A}_{0}\right| h\left|\underline{B}_{0}\right\rangle+\left\langle\underline{A}_{0} \underline{B}_{0}\right| v\left|\underline{B}_{0} \underline{B}_{0}\right\rangle+2\left\langle\underline{A}_{0} \underline{A}_{1}\right| v\left|\underline{B}_{0} \underline{A}_{1}\right\rangle$
$r=\left\langle\underline{A}_{0}\right| h\left|\underline{B}_{0}\right\rangle+\left\langle\underline{A}_{0} \underline{A}_{0}\right| v\left|\underline{B}_{0} \underline{A}_{0}\right\rangle+2\left\langle\underline{A}_{0} \underline{A}_{1}\right| v\left|\underline{B}_{0} \underline{A}_{1}\right\rangle$

$$
\begin{aligned}
& x=\left\langle\underline{A}_{0}\right| h\left|\underline{A}_{1}\right\rangle+\left\langle\underline{A}_{0} \underline{A}_{0}\right| v\left|\underline{A}_{1} \underline{A}_{0}\right\rangle+2\left\langle{\underset{\sim}{A}}_{0}^{B} \underline{\sim}_{0}\right| v\left|\underline{A}_{1}{\underset{\sim}{B}}_{0}\right\rangle
\end{aligned}
$$

$$
\begin{aligned}
& -2<{\underset{\sim}{A}}_{0}{\underset{\sim}{B}}_{0}|v|{\underset{-}{B}}_{0}^{A_{0}}>\quad-2<A_{0} B_{0}|v| B_{0} A_{1}>
\end{aligned}
$$

The $\lambda$ independence is expected because changing $\lambda$ simply amounts to changing the basis set.

In the following two sections we examine aspects of the behavior of the four individual terms $E_{S t}^{(i)}$, $i=1,2,3,4$ under different choices of the Wannier functions $A_{0}, B_{0}, A_{1}, i . e .$, of $\lambda$. We find, in particular, an a priori way of determining $\lambda$ which makes $\Delta \mathrm{E}_{\mathrm{st}}^{(3)}=\Delta \mathrm{E}_{\mathrm{st}}^{(4)}=0$; we also show that these individual terms $\Delta \mathrm{E}_{\mathrm{st}}^{(\mathrm{i})}$ depend very sensitively on $\lambda$. The utility of these findings, while not apparent from the present simple example (where the complete answer to $0\left(\Delta^{4}\right)$, (15) has been obtained), will be seen in connection with the generalization to a crystal.

## III. THE ONE-ELECTRON STATES IN THE THERMAL SINGLE DETERMINANT APPROXIMATION

The well-known method of determining one-electron states is the Hartree-Fock approximation. As we noted earlier, this was Anderson's approach and it fails to give one-electron states that satisfy the requirement that they be nonmagnetic. This requirement is clearly important, since the magnetism is to be predicted by the result (effective Hamiltonian $=-\Sigma J_{i j} \vec{S}_{i} \cdot \vec{S}_{j}$ ) of our perturbation theory. We therefore turn to a recently introduced development ${ }^{6}$, the thermal single-determinant approximation (TSDA), which is closely related to the Hartree-Fock approximation in that it also determines "best" one-electron states. But in the variational context, it is better (more precisely, it is never worse) than the Hartree-Fock approximation. ${ }^{4}$

The idea of TSDA is as follows. We consider the minimum (or variational) principle of quantum statistical mechanics for a system with a Hamiltonian operator $H$ and number of particle operator $N$,

$$
\begin{equation*}
F(\rho) \equiv \operatorname{Tr}\left[\rho\left(H-\mu N+\beta^{-1} \log \rho\right)\right] \geq F_{e} \tag{1}
\end{equation*}
$$

Here $\rho$ is an arbitrary Hermitian and nonnegative operator with unit trace, (i.e. a density operator) $\beta=1 / \mathrm{kT}, \mu$ is the chemical potential, and $F_{e}$ is the exact grand-canonical free energy for the system, namely

$$
\begin{align*}
& F_{e}=F\left[\rho_{e}\right]=\beta^{-1} \log \operatorname{Tr} e^{-\beta(H-\mu N)} ;  \tag{2}\\
& \rho_{e}=e^{-\beta(H-\mu N)} / \operatorname{Tr} e^{-\beta(H-\mu N)} \tag{3}
\end{align*}
$$

is the exact grand-canonical density matrix. For a system of interacting fermions, we have

$$
\begin{equation*}
\left.H=\Sigma\langle i| h \left\lvert\, j>c_{i}^{\dagger} c_{j}+\frac{1}{2} \underset{i j}{\sum} \sum_{k \ell}^{\Sigma}\langle i j| v\right. \right\rvert\, k 1>c_{i}{ }^{\dagger} c_{j}{ }^{\dagger} c_{\ell} c_{k} \tag{4}
\end{equation*}
$$

where $|i\rangle,|j\rangle \ldots$ is any complete and orthonormal set of oneparticle state, and $c_{i}$ are the corresponding Fermion destruction operators. The free energy in the TSDA is the minimum of $F(\rho)$ for a trial density matrix of the form

$$
\begin{equation*}
\rho=e^{-\beta(\tilde{H}-\mu N)} / \operatorname{Tr} e^{-\beta(\tilde{H}-\mu N)} \tag{5}
\end{equation*}
$$

where

$$
\begin{equation*}
\tilde{H}=\sum h_{i i} n_{i}+\frac{1}{2} \sum_{i j} \hat{v}_{i j, j i} n_{i} n_{j} \tag{6}
\end{equation*}
$$

is a function of the occupation-number operators $n_{i}$ corresponding to a complete orthonomal set of one-particle states |i>. Here

$$
\begin{align*}
h_{i j} & =\langle i| h|j\rangle \\
v_{i j, k l} & =\langle i j| v|k l\rangle  \tag{7}\\
\hat{v}_{i j, k l} & =\langle i j| v|k l\rangle-\langle i j| v|1 k\rangle
\end{align*}
$$

Requiring stationarity of $F(\rho)$ under arbitrary variations of the states $\mid i>$, we find that the one-particle states |i> are determined by the system of TSDA equations,

$$
\begin{equation*}
\left\langle<n_{i}-n_{j} \gg h_{i j}+\sum_{k} \hat{v}_{i k, j k} \ll\left(n_{i}-n_{j}\right) n_{k} \gg=0\right. \tag{8}
\end{equation*}
$$

plus the condition that they form a complete orthonormal set. The double brackets in the TSDA equations denote the average over the trial density matrix,

$$
\begin{equation*}
\ll 0 \gg=\operatorname{Tr}[\rho 0] \tag{9}
\end{equation*}
$$

where $\rho$ is given by equations (5) and (6). Thus the equations obtained by first putting $0=n_{i}$ into (9) and then $0=n_{i} n_{j}$ form, together with (8), a set that must be solved self-consistently: Given the average in (8), solution of these equations yields a set of one-electron state |i> which then determines $\tilde{H}$ and $\rho$ and therefore the average occurring in (8). And the latter must match those "given" values that started the process.

In our 3-site problem i, j, $k$ run over 6 values corresponding to our 6-dimensional single-electron function
-10 $2+10$
space. Although this was derived from the grand canonical ensemble, these equations are applicable to our present considerations. Because we are interested only in $T \ll U$, the smallest unperturbed excitation energy from the ground state, the fluctuations in the number of particles is negligible. The reason for this restriction to low $T$ is that we are after the best one-electron states which will define our perturbation expansion for the low-lying manyelectron states. Let $n_{i \sigma}$ be the occupation number for site i with spin $\sigma$. Then equation (6) becomes

$$
\begin{equation*}
\tilde{H}=\Sigma h_{i i} n_{i \sigma}+\frac{1}{2} \Sigma \hat{v}_{i \sigma j \sigma^{\prime}, i \sigma j \sigma^{\prime}} n_{i \sigma^{\prime}} n_{j \sigma^{\prime}} \tag{10}
\end{equation*}
$$

and (8) becomes

$$
\begin{equation*}
\left\langle<n_{i \sigma^{\prime}}-n_{j \sigma^{\prime}} \gg h_{i j}+\sum_{\ell \sigma^{\prime}}^{\Sigma} \hat{\mathbf{v}}_{i \sigma \ell \sigma^{\prime}, j \sigma \ell \sigma^{\prime}} \ll\left(n_{i \sigma^{-}}-n_{j \sigma}\right) n_{\ell \sigma^{\prime}} \gg=0\right. \tag{1}
\end{equation*}
$$

Now for the 3 -site case, $i=1,2,3$ correspond to $A_{0}, B_{0}, A_{1}$ respectively. For $i=A_{0}, j=A_{1}$ we get

$$
\begin{align*}
& \ll n_{A_{0} \uparrow}{ }^{-n_{A_{1} \uparrow}} \gg h_{A_{0} A_{1}}+\ll\left(n_{A_{0} \uparrow}{ }^{-n_{A_{1}}}\right)^{n_{A_{0}} \downarrow}{ }^{\gg} v_{A_{0}} A_{0}, A_{1} A_{0} \\
& +\ll\left(\mathrm{n}_{\mathrm{A}_{0} \uparrow}-\mathrm{n}_{\mathrm{A}_{1} \uparrow}\right) \mathrm{n}_{\mathrm{A}_{0} \uparrow \gg \hat{\mathrm{v}}_{\mathrm{A}_{0} \mathrm{~A}_{0}, \mathrm{~A}_{1} \mathrm{~A}_{0}}+\ll\left(\mathrm{n}_{\mathrm{A}_{0} \uparrow}-\mathrm{n}_{\mathrm{A}_{1} \uparrow}\right)}  \tag{12}\\
& \mathrm{n}_{\mathrm{A}_{1} \downarrow \gg} \mathrm{v}_{\mathrm{A}_{0} \mathrm{~A}_{1}, \mathrm{~A}_{\mathrm{o}} \mathrm{~A}_{\mathrm{o}}}+\ll\left(\mathrm{n}_{\mathrm{A}_{0} \uparrow}-\mathrm{n}_{\mathrm{A}_{1} \uparrow}\right) \mathrm{n}_{\mathrm{A}_{1} \uparrow} \gg \hat{\mathrm{v}}_{\mathrm{A}_{0} \mathrm{~A}_{1}}, \mathrm{~A}_{1} \mathrm{~A}_{1}
\end{align*}
$$

$$
\begin{aligned}
& \hat{v}_{A_{0} B_{0}, A_{1} B_{o}}=0
\end{aligned}
$$

we see that from the symmetry $\mathrm{v}_{\mathrm{A}_{0} A_{0}, A_{1} A_{0}}=\mathrm{v}_{\mathrm{A}_{0} A_{1}, A_{1} A_{1}}$ and the left-hand side of the above equation vanishes identically. Therefore, for $i=A_{0}, j=A_{1}$ the TSDA equation is satisfied by symmetry. For $i=A, j=B$, we get the TSDA equation

$$
\begin{align*}
& =\ll\left(\mathrm{n}_{\mathrm{A}_{0} \uparrow}-\mathrm{n}_{\mathrm{B}_{0} \uparrow}\right) \mathrm{n}_{\mathrm{A}_{1} \uparrow}{ }^{\gg \mathrm{v}_{\mathrm{A}_{0}} \mathrm{~A}_{1}, \mathrm{~A}_{1} \mathrm{~B}_{\mathrm{o}} .} \tag{13}
\end{align*}
$$

If we substitute equation (5) of section $I$ into this equation and use (5) and (6),it becomes a transcendental equation for $\lambda$ thus obtaining $A_{0}, B_{0}, A_{1}$. For $k T \ll U$ we can

$\left.\left.<n_{A_{0} \uparrow}\right\rangle \simeq \frac{1}{2}<n_{A_{0}} \uparrow^{n} A_{A_{0}}\right\rangle^{>} \simeq 0$ at $T \rightarrow 0$. The above TSDA equation then becomes

$$
\begin{equation*}
\left\langle A_{0}\right| h\left|B_{0}>+<A_{0} A_{0}\right| v\left|B_{0} A_{0}>+<A_{0} B_{0}\right| v \mid B_{0} B_{0}>+\left\langle A_{0} A_{1}\right| v \mid B_{0} A_{1}>=0 \tag{14}
\end{equation*}
$$

to the leading order in $\Delta$. To the leading order the following matrix elements are
$h_{A_{0} B_{0}} \simeq \underline{h}_{\underline{A}_{0} \underline{B}_{0}}+\lambda\left(\underline{h}_{\underline{B}_{0} \underline{B}_{0}}{ }^{-h_{A_{0}} \underline{A}_{0}}\right)$



Substituting the above equation into (14) and solving for $\lambda$ we get the TSDA value of $\lambda$ to leading order in the overlap $\Delta$

$$
\begin{equation*}
\lambda_{T}=-\frac{p}{w_{21}} \tag{16}
\end{equation*}
$$

where $w_{21}$ is defined below (15) in section $I I$, and $p$ is given in (16) in Section II. To the leading order of overlap $\Delta$

$$
\begin{align*}
& \underline{A}_{0} \simeq a_{0}+\mu b_{0} \\
& \underline{B}_{0} \simeq b_{0}-(\Delta+\mu)\left(a_{0}+a_{1}\right)  \tag{17}\\
& \underline{A}_{1} \simeq a_{1}+\mu b_{0}
\end{align*}
$$

Substituting this into (16), we get a useful relation

$$
\begin{equation*}
\lambda_{\mathrm{T}}+\mu=\theta \tag{18}
\end{equation*}
$$

where

$$
\begin{equation*}
\theta=\frac{\left.\left\langle a_{0}\right| \tilde{h}\left|b_{0}\right\rangle-\Delta<a_{0}|\tilde{h}| a_{0}\right\rangle}{w_{21}^{o}} \tag{19}
\end{equation*}
$$

and

$$
\begin{equation*}
\tilde{h} \equiv h+\int d^{3} \vec{r}^{\prime} \frac{e^{2}}{\left|\vec{r}-\vec{r}^{\prime}\right|}\left[a_{o}^{2}\left(\vec{r}^{\prime}\right)+b_{o}^{2}\left(\vec{r}^{\prime}\right)+a_{1}^{2}\left(\vec{r}^{\prime}\right)\right] \tag{20}
\end{equation*}
$$

$w_{21}^{\circ}$ is obtained from $w_{21}$ simply by letting ${\underset{\sim}{A}}^{A_{0}} a_{0}, \underline{B}_{0} \rightarrow b_{0}$, $\mathrm{A}_{1} \rightarrow \mathrm{a}_{1}$.

Although $\lambda_{\mathrm{T}}$ is linearly dependent on $\mu, \theta$ is independent of $\mu$. The solution of TSDA, $A_{0}, B_{o}, A_{1}$, should be independent of the choice of $\underline{A}_{0}, \underline{B}_{0}, \underline{A}_{1}$ for given nonorthogonal atomic orbitals $a_{0}, b_{0}, a_{1}$ or independent of $\mu$ from equation (3) of Section II (since ${\underset{\sim}{A}}_{0},{\underset{O}{B}}^{0}, \underline{A}_{1}$ is simply a basis set). We can verify that $A_{0}, B_{0}, A_{1}$ do satisfy this by the useful relation (18). Take the TSDA of $A$ and express it in terms of the nonorthogonal atomic orbitals. That is, substituting (3) into (5) in Section II we get

$$
\begin{align*}
A & =\frac{1}{1+\lambda_{T}^{2} / 2}\left\{\left[C-\lambda_{T} D(\Delta+\mu)\right] a_{0}+\left(C \mu+\lambda_{T} D\right) b_{0}\right. \\
& \left.-\left[C\left(\frac{\mu^{2}}{2}+\mu \Delta\right)+D \lambda_{T}(\Delta+\mu)+C \lambda_{T}^{2} / 2\right] a_{1}\right\} \tag{21}
\end{align*}
$$

with $\quad C^{2}=1-2 \mu \Delta-\mu^{2}$ and $D^{2}=1$.
The coefficient of $a_{0}$ to $O\left(\Delta^{2}\right)$ is
$\frac{\mathrm{C}-\lambda_{\mathrm{T}} \mathrm{D}(\Delta+\mu)}{1+\lambda_{\mathrm{T}}^{2} / 2}=1-\mu \Delta-\frac{\mu^{2}}{2}-(\theta-\mu)(\Delta+\mu)-\frac{\lambda_{\mathrm{T}}^{2}}{2}=1-\theta \Delta-\frac{\theta^{2}}{2} ;$
The coefficient of $a_{1}$ to $0\left(\Delta^{2}\right)$ is

$$
\begin{equation*}
-\left[\frac{\mu^{2}}{2}+\mu \Delta-+(\theta-\mu)(\Delta+\mu)+\frac{(\theta-\mu)^{2}}{2}\right]=-\left(\theta \Delta+\frac{\theta^{2}}{2}\right) . \tag{23}
\end{equation*}
$$

To the $0\left(\Delta^{2}\right)$, we thus obtain
$A=\left(1-\theta \Delta-\frac{\theta^{2}}{2}\right) a_{0}+\theta b_{0}-\left(\theta \Delta+\frac{\theta^{2}}{2}\right) a_{1}$
Similarly, we can prove that $B_{o}$ is independent of $\mu$. Therefore this yields the Wanner functions in TSDA. To $0\left(\Delta^{2}\right)$ these are ${ }^{13}$
$A=\left(1-\theta \Delta-\frac{\theta^{2}}{2}\right)\left[a_{0}+\theta b_{0}-\left(\theta \Delta+\frac{\theta^{2}}{2}\right) a_{1}\right]$
$B=\left(1-\theta^{2}+\Delta^{2}\right)\left[b_{0}-(\Delta+\theta)\left(a_{0}+a_{1}\right)\right]$.
IV. RESULTS OF THE 3-SITE MODEL

To the leading order, eq. (9) of Section II yield

$$
\begin{aligned}
\mathrm{H}_{12}=\mathrm{H}_{56} & =\left\langle A_{0}\right| h\left|B_{0}\right\rangle+\left\langle A_{0} A_{0}\right| v\left|B_{0} A_{0}\right\rangle+\left\langle A_{0} B_{0}\right| v\left|B_{0} B_{0}\right\rangle \\
& +\left\langle A A_{1}\right| v\left|B_{0} A_{1}\right\rangle
\end{aligned}
$$

This is the left hand side of the TSDA equation (14) in the last Section. Therefore, from equation (8), (9), (10) and (14) of Section II, we immediately have to $0\left(\Delta^{4}\right)$ and zeroth order in $\exp (-\beta U)$

$$
\begin{align*}
& \Delta \mathrm{E}_{\mathrm{st}}^{(3)}\left(\lambda_{\mathrm{T}}\right)=0 \quad, \quad \Delta \mathrm{E}_{\mathrm{St}}^{(4)}\left(\lambda_{\mathrm{T}}\right)=0 \\
& \Delta \mathrm{E}_{\mathrm{St}}^{(2)}\left(\lambda_{\mathrm{T}}\right)=-\frac{\mathrm{H}_{13}{ }^{2}}{\mathrm{H}_{33}-\mathrm{H}_{11}}-\frac{\mathrm{H}_{14}^{2}}{\mathrm{H}_{44}-\mathrm{H}_{11}} . \tag{1}
\end{align*}
$$

Therefore using the TSDA states, the $2^{\text {nd }}$ order perturbation theory exhausts to $0\left(\Delta^{4}\right)$, the total energy splitting:

$$
\begin{equation*}
\Delta E_{s t}=\Delta E_{s t}^{(1)}\left(\lambda_{T}\right)+\Delta E_{s t}^{(2)}\left(\lambda_{T}\right) \tag{2}
\end{equation*}
$$

To show how sensitive is the dependence of the individual perturbation terms on the choice of $\lambda$, we consider the following simple example. Choose

$$
\begin{align*}
& A_{0}=a_{0} \\
& B_{0}=b_{0}-\Delta\left(a_{0}+a_{1}\right)  \tag{3}\\
& A_{1}=a_{1}
\end{align*}
$$

instead of the TSDA states. Then dropping all intersite Coulomb matrix elements, one can see very easily that

$$
\begin{align*}
& \Delta \mathrm{E}_{\mathrm{st}}^{(1)}=0, \quad \Delta \mathrm{E}_{\mathrm{st}}^{(2)}=0, \quad \Delta \mathrm{E}_{\mathrm{st}}^{(3)}=0 \\
& \Delta \mathrm{E}_{\mathrm{st}}=\Delta \mathrm{E}_{\mathrm{st}}^{(4)}=-\frac{4\left|<\mathrm{a}_{0}\right| \mathrm{h}\left|\mathrm{~b}_{0}>\right|^{4}}{\mathrm{w}_{21}}\left(\frac{1}{\mathrm{w}_{31}}+\frac{2}{w_{41}}\right) \tag{4}
\end{align*}
$$

That is, the total splitting to $0\left(\Delta^{4}\right)$ comes entirely from the fourth order perturbation rather than from first and second order as is the case where TSDA orbitals are used.

In summary, equation (1) or (2) shows that,
within the model considered here, Anderson's hope that there exist "exact Wannier functions" which would lead to rapid convergence of the perturbation series has been fulfilled, the exact Wannier functions being the TSDA localized states given by (16) of Section III or alternately by (25) of Section III. In fact, the rapidity of convergence is probably better than expected, in that the $3^{r d}$ and $4^{\text {th }}$ order perturbation terms vanish identically to $O\left(\Delta^{4}\right)$ using the TSDA states.

## V. GENERALIZATION OF THE 3-SITE RESULTS TO A MANY-ATOM LATTICE

Similarly to the 3 -site case, the function space is defined to have atomic functions $a_{i}$ centered on magnetic atoms and $b_{i}$ centered on diamagnetic atoms. These functions are real; the $a_{i}$ are connected by lattice translations, and similarly for the $b_{i}$. We also assume nearest neighbor cation-anion overlap of these functions; the next nearest neighbor and more distant overlaps are assumed to be zero. The Wannier functions $\underline{A}_{i}, \underline{B}_{i}$ are constructed in this space. We require the constructed Wannier functions to be orthogonal, and to approach the atomic functions $a_{i}$ and $b_{i}$ as the overlap approach zero. We also take them to satisfy "maximum similarity to the $a_{i}\left(b_{i}\right)$ " i.e., we require $\underline{A}_{i}\left(\underline{B}_{i}\right)$ to transform 1ike $a_{i}\left(b_{i}\right)$ under all lattice symmetry operations that leave the point $R_{i}$ unchanged. For simplicity we take the $a_{i}$ to be s-functions; then we require the $\underline{A}_{i}$ to be real and invariant under the symmetry group of rotations that leave the point $R_{i}$ fixed. Despite these restrictions, the Wannier functions so constructed are not uniquely defined (as is also true for the 3-site mode1).

For illustrative purposes, suppose we have a periodic linear chain with $2 N+1$ unit cells, with one a-atom and one $b$-atom in each cell, as shown in Fig. B2. The Wannier functions in cell "o" on atom a and b respectively will be

$$
\begin{align*}
{\underset{\sim}{0}} & =a_{0}+\lambda_{1}\left(b_{-1}+b_{0}\right)+\lambda_{2}\left(b_{-2}+b_{1}\right)+\ldots+\lambda_{i}\left(b_{-i}+b_{i-1}\right)+\ldots \\
& +\lambda_{N}\left(b_{-N}+b_{N-1}\right)+\lambda_{N+1} b_{N}+v_{1}\left(a_{-1}+a_{1}\right)+v_{2}\left(a_{2}+a_{-2}\right)+\ldots \\
& +v_{i}\left(a_{-i}+a_{i}\right)+\ldots+v_{N}\left(a_{-N}+a_{N}\right)  \tag{1}\\
\underline{B}_{0} & =b_{0}+\lambda_{1}^{\prime}\left(a_{0}+a_{1}\right)+\lambda_{2}^{\prime}\left(a_{-1}+a_{2}\right)+\ldots+\lambda_{i}^{\prime}\left(a_{-i+1}+a_{i}\right)+\ldots \\
& +\lambda_{N}^{\prime}\left(a_{-N+1}+a_{N}\right)+v_{1}^{\prime}\left(b_{-1}+b_{1}\right)+v_{2}^{\prime}\left(b_{-2}+b_{2}\right)+\ldots \\
& +v_{i}^{\prime}\left(b_{-i}+b_{i}\right)+\ldots+v_{N}^{\prime}\left(b_{-N}+b_{N}\right) \tag{2}
\end{align*}
$$

There are $2 \mathrm{~N}+1$ parameters $\left(\lambda_{1}, \lambda_{2}, \ldots \lambda_{N}, \lambda_{N+1} \nu_{1} \ldots \nu_{N}\right)$ in $\underline{A}_{0}$; and similarly $2 N+1$ parameters in ${\underset{\sim}{B}}\left(\lambda_{1}^{\prime}, \lambda_{2}^{\prime}, \ldots\right.$ $\left.\lambda_{N}^{\prime}, \lambda_{N+1}^{\prime} ; v_{1}^{\prime} \ldots v_{N}^{\prime}\right)$.

A 3-dimensional example for which our assumptions would reasonably apply is indicated in Fig. B3. There the dots represent cation (magnetic-ion) sites with wave-functions $a_{i}$, the open circles stand for axion (diamagnetic-ion) sites with wave-functions $b_{i}$. This is


Figure B2. Linear-chain model


Figure $B 3$. Peroskite structure $\mathrm{ABF}_{3}$, showing only the $B$-ions (.) and the F-ion (0).
pertinent to magnetic materials of the perovskite structure, e.g. $\mathrm{KMnF}_{3}, \mathrm{KNiF}_{3}{ }^{14}$. In the latter example, $\cdot=\mathrm{N}_{\mathrm{i}}{ }^{2+}$, $0=\mathrm{F}^{-}$; the potassium ions, which are not thought to contribute appreciably to the superexchange, are not shown.

The model Hamiltonian is written in terms of these Wannier functions as follows:

$c_{i \sigma}$ are the destruction operators corresponding respectively to the various Wannier functions. The hamiltonian is divided into two parts, $H=H_{0}+V$ where

$$
\begin{equation*}
H_{o}=\sum_{i} h_{i j} n_{i}+\frac{1}{2} \sum_{i j} v_{i j, i j} n_{i} n_{j} \tag{4}
\end{equation*}
$$

which is clearly of zero order in the overlap $\Delta$; here $n_{i}=\sum_{\sigma} n_{i \sigma}$. The perturbation part $V=H-H_{o}$ is $1^{\text {st }}$ order in the overlap. In the unperturbed ground state there is by definition one electron on each magnetic site $\sum_{\sigma} n_{a \sigma}=1$, and the electrons on each diamagnetic site $\sum_{\sigma} n_{b \sigma}=2$. Our purpose is to derive for small overlap the appropriate spin Hamiltonian which will describe the lowlying magnetic states (including the magnetic ordering and thermal magnetic properties). We therefore look for our unperturbed states in a temperature region in which there is no magnetic ordering, i.e. $T \gg$ the magnetic ordering temperature. Similarly, we want our states to relate to
the physical situation where the number of electrons on a magnetic ion is approximately 1 , hence we consider $\mathrm{kT} \ll \Delta \mathrm{E}_{\text {inter }}$ the unperturbed energy it costs to hop an electron from one atom to another. Under this condition, we get

$$
\begin{align*}
&<n_{a_{i} \sigma^{\prime}} \simeq \frac{1}{2} \quad<n_{b_{i} \sigma^{\prime}} \simeq 1 \\
&<n_{a_{i} \uparrow} n_{a_{i} \nmid}> \simeq 0 \quad<n_{a_{i} \sigma^{n} a_{j} \sigma^{\prime}}>\simeq \frac{1}{4}  \tag{5}\\
&<n_{a_{i} \sigma^{n_{b}} \sigma^{\prime}}>\simeq \frac{1}{2}
\end{align*}
$$

The TSDA equations between different magnetic sites are satisfied exactly by symmetry, a similar result being valid for different diamagnetic sites. The only nonzero TSDA equation are between the magnetic and diamagnetic sites. In the case of a linear chain, we have $N+1$ equations, namely between $a_{0}$ and $b_{i}, i=0,1, \ldots N$. From equation (11) of Section III let $i=a_{o}, j=b_{j}$ and substitute the conditions (5). We get

$$
\begin{align*}
<a_{o}|h| b_{j}> & +\sum_{a i}<a_{o} a_{i}|v| b_{j} a_{i}>+2 b_{i} \neq b_{j}<a_{o} b_{i}|v| b_{j} b_{i}> \\
& +<a_{o} b_{j}|v| b_{j} b_{j}>=\frac{1}{2} \sum_{a_{i} \neq a_{o}}^{\sum_{o}}<a_{i}|v| a_{i} b_{j}> \\
& +b_{i}^{\sum \neq b_{j}}<a_{o} b_{i}|v| b_{i} b_{j}> \tag{6}
\end{align*}
$$

Besides $N+1$ of these equations, there are orthogonality
conditions of the following types:

$$
\begin{align*}
& \left\langle A_{0} \mid A_{i}\right\rangle=2\left(\lambda_{i}+\lambda_{i+1}\right) \Delta+2 v_{i} \quad i=1,2, \ldots N \\
& <B_{0}\left|A_{i}\right\rangle=\left\langle b_{0} \mid a_{i}\right\rangle+\left(v_{i}+v_{i}^{\prime}+v_{i-1}+v_{i-1}^{\prime}\right) \Delta+\left(\lambda_{i}+\lambda_{i}^{\prime}\right) \\
& <B_{0}\left|A_{-N}\right\rangle=2\left(v_{N}^{\prime}+\nu_{N}\right) \Delta+\lambda_{N+1}^{\prime}+\lambda_{N+1}  \tag{7}\\
& <B_{0}\left|B_{i}\right\rangle=2\left(\lambda_{i}^{\prime}+\lambda i_{i+1}^{\prime}\right) \Delta+2 v_{i}^{\prime} .
\end{align*}
$$

Therefore, we have totally $4 \mathrm{~N}+2$ equations to determine $4 \mathrm{~N}+2$ parameters. To the leading order of overlap, we find

$$
\begin{equation*}
\lambda_{1}=\frac{f-\Delta \varepsilon^{\prime}}{\varepsilon^{\prime}-\varepsilon} \quad \lambda_{1}^{\prime}=\frac{f-\Delta \varepsilon}{\varepsilon-\varepsilon^{\prime}} \tag{8}
\end{equation*}
$$

where

$$
\begin{aligned}
f & =\left\langle a_{0}\right| \tilde{h}\left|b_{0}\right\rangle \\
\varepsilon^{\prime} & =\left\langle a_{0}\right| \tilde{h}\left|a_{0}\right\rangle \quad \varepsilon=\left\langle b_{0}\right| \tilde{h}\left|b_{0}\right\rangle
\end{aligned}
$$

and

$$
\tilde{h}=h+\int d^{3} r \frac{e^{2}}{|\vec{r}-\vec{r} \cdot|}\left[\sum_{\ell} a_{\ell}^{2}+b_{o}^{2}+2 \sum_{\ell \neq 0} b_{l}^{2}\right]
$$

The effective Hamiltonian $H_{S}$ defined in Appendix $F$ can be expanded as follows:

$$
\begin{align*}
H_{S} & =P\left[H_{0}+V-V Q \frac{1}{H-E} Q V\right] P  \tag{9}\\
& =P\left[H_{0}+V-V \frac{1}{H_{0}-E_{0}} V+V_{\frac{1}{H_{0}-E_{0}}} V \frac{1}{H_{0}-E_{0}} V \ldots\right] P
\end{align*}
$$

where $P$ is the projection operator, defined in Appendix $F$,

$$
\begin{aligned}
& 1 \\
& 1 \\
& 1 \\
& 1
\end{aligned}
$$

which projects onto the ground state manifold. Looking at the $3^{\text {rd }}$ order perturbation term, we should begin with the ground state and come back to the ground state. Each term in the expansion will be proportional to the product of three matrix elements of $V$, and has the form $\left\langle g^{\prime}\right| v|x\rangle\langle x| v|y\rangle\langle y| v|g\rangle$ where $\left|g^{\prime}\right\rangle,|g\rangle$ are in the groundstate manifold and $|x\rangle,|y\rangle$ are excited states. For the factor $\langle y| v|g\rangle$, either one electron or two electrons can be hopped from the ground state. If two electrons are hopped, tinis factor is of $0\left(\Delta^{2}\right)$. If only one electron is hopped, and the electron goes from one a-site to another a-site, this factor is also of $0\left(\Delta^{2}\right)$. If the electron goes from a b-site to an a-site, this factor is the following (see Appendix E):

$$
\begin{equation*}
\left\langle\mathrm{c}_{\mathrm{a}_{\mathrm{o}}^{+}}^{\mathrm{c}_{\mathrm{b}}} \Phi_{\mathrm{o}}^{\circ}\right| \mathrm{H}\left|\Phi_{\mathrm{g}}^{\circ}\right\rangle=\ldots+0\left(\Delta^{3}\right) \tag{10}
\end{equation*}
$$

where ... is the expression on the left hand side of TSDA equation (6). The rest of the terms are higher order in the overlap. Therefore, if we use the TSDA basis, this factor is of order $\Delta^{3}$ instead of $\Delta$. The total product is therefore higher than $0\left(\Delta^{4}\right)$. For the $4^{\text {th }}$ order perturbation term, we can apply the same analysis. Therefore, both the $3^{\text {rd }}$ and $4^{\text {th }}$ order perturbation terms are of higher order as long as we use the TSDA states as basis-functions. Using this basis, we therefore need apply only through $2^{\text {nd }}$ order
perturbation theory to give all the leading-order contributions $\left(0\left(\Delta^{4}\right)\right)$ to the effective Hamiltonian $H_{s}$. As shown in Appendix G, this gives the Heisenberg spin Hamiltonian
where $\quad H_{S}=-\sum_{i j} J_{i j} \vec{S}_{i} \cdot \vec{S}_{j}$
where

$$
\begin{align*}
J_{i j} & =J_{i j}^{P}+J_{i j}^{k}+J_{i j}^{N} \\
J_{i j}^{k} & =-\frac{2\left|t_{i j}\right|^{2}}{U_{i j}}, J_{i j}^{P}=v_{i j, j i} \\
J_{i j}^{N} & =-\sum_{\ell}^{\sum} \frac{\left|v_{i j, \ell \ell}\right|^{2}}{U_{i j, \ell}} \tag{12}
\end{align*}
$$

here

$$
\begin{align*}
t_{i j} & =h_{i j}+\sum_{\ell} 2 v_{i \ell, j \ell}+v_{i j, j j}+\sum_{k}^{i, j} v_{i k, j k} \\
v_{i j, \ell \ell} & =\left\langle c_{j \bar{\sigma}}^{+} c_{\ell \bar{\sigma}} c_{i \sigma} c_{\ell \sigma} v\right| H_{o}|v\rangle \\
U_{i j} & =\left\langle c_{i}^{+} c_{j} v\right| H_{o}|v\rangle-\varepsilon . \tag{13}
\end{align*}
$$

$\varepsilon=E_{0}, i, j$ are on a-sites, $\ell$ is on $a b-s i t e$, and $|v\rangle$
is an unperturbed ground state.
VI. SUMMARY AND DISCUSSION

In principle, one has to go through fourth order perturbation theory in order to exhaust all terms of $0\left(\Delta^{4}\right)$, a very complex task for realistic models of magnetic insulators. We showed that the nonmagnetic Wannier function solutions of the TSDA equations make the $3^{\text {rd }}$ and $4^{\text {th }}$ order perturbation terms vanish, for the simple models considered. This result proves Anderson's idea that there exist "exact Wannier functions" which make $3^{\text {rd }}$ and $4^{\text {th }}$ order perturbation terms negligible, the main contribution to the exchange parameter coming from $1^{s t}$ and $2^{\text {nd }}$ order perturbation terms.

The generalization of this type of consideration to more realistic models (containing more than one electron on a magnetic site and describing a crystal) is felt to be of considerable importance. The reason is that, in our opinion, a conclusive evaluation of Anderson's general perturbation theoretical approach as a practical means of calculating spin-Hamiltonian parameters is impossible without such considerations. That is to say, one must correctly evaluate all terms to $0\left(\Delta^{4}\right)$; hopefully an appropriate choice of Wannier functions will greatly simplify that task.

## REFERENCES

1. S. Chikazumi and S. H. Charap, Physics of Magnetism (John Wiley \& Sons, Inc., New York) p. 440.
2. A. H. Morrish, The Physical Principles of Magnetism, (John Wiley \& Sons, Inc., New York) p. 478.
3. H. A. Kramers, Physica 1, 182 (1934).
4. P. A. Anderson, Solid State Physics 14, 99 (1963).
5. N. P. Silva and T. A. Kaplan, AIP Conf. Proc. No. 18, Magnetism and Magnetic Mat'1s. (1973), p.656.
6. T. A. Kaplan and Petros N. Argyres, Phys. Rev. B1, 2457 (1970).
7. In Ref. (5) another variational approach was studied. However, the present one is simpler and adequate to our more limited purpose.
8. T. A. Kaplan, Bull. Am. Phys. Soc. 13, 386 (1968).
9. Ken-Ichiro Gondaira, Tukito Tanabe, J. Phys. Soc. Japan 21, 1527 (1966), referred to as GT.
10. GT also carried out perturbation calculations using these basis functions, but they did not discuss the corresponding $3^{\text {rd }}$ and $4^{\text {th }}$ order perturbationtheoretic terms.
11. N. Fuchikami, J. Phys. Soc., Japan 28871 (1970).
12. In Ref. 11, agreement with experiment is claimed. But to obtain that, the carefully calculated energy denominator $U$ was reduced by a factor of crudely 2; the physical origin of this reduction, if it properly exists, is outside the simple low-order ( ~fourth-order) perturbation theory.
13. These TSDA Wannier functions are the same as those derived by GT by a different method. Their method also differs from Anderson's HF theory although they use the term "Hartree-Fock" to describe their approach. However, if there is more than one electron on each magnetic site, their "Hartree-Fock" approach does not take care of Hand's rule while TSDA does. Therefore, it is expected that the TSDA will have different solutions in such cases.
14. D. E. Rimmer, J. Phys. C (Solid St. Phys.) 2 329, (1969).
15. R. D. Mattuck, A Guide to Feynman Diagrams in the ManyBody Problem, (McGraw-Hill Book Company, New York).
16. L. I. Schiff, Quantum Mechanics, (McGraw-Hill Book Company, New York).
17. T. A. Kaplan, unpublished lecture notes for a course in Magnetism given at Michigan State University in 1971 and 1972.

APPENDICES

## APPENDIX A <br> DIAGONALIZATION OF THE TRANSFER MATRIX

The transfer matrix $\bar{T}$ is in the following form:
$\bar{T}=\left[\begin{array}{llll}A & B & C & D \\ B & e^{-u} & 1 & B \\ C & 1 & e^{u} & C \\ D & B & C & A\end{array}\right] \quad \begin{array}{ll}A=e^{-(x+z)} \\ & B=e^{-\frac{1}{2}(z+u)} \\ C h e r e & C e^{-\frac{1}{2}(z-u)} \\ D=e^{x-z}\end{array}$
If we make a similarity transformation on $\bar{T}$ with the unitary matrix

$$
S=\left[\begin{array}{rrrr}
\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}}  \tag{A2}\\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
\frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}}
\end{array}\right]
$$

the T is transformed into the following form:

$$
S^{-1} \mathrm{TS}=\left(\begin{array}{cccc}
A+D & \sqrt{2} B & \sqrt{2} C & 0  \tag{A3}\\
\sqrt{2} B & e^{-u} & 1 & 0 \\
\sqrt{2} C & 1 & e^{u} & 0 \\
0 & 0 & 0 & A-D
\end{array}\right)
$$

Immediately, one of the eigenvalues is

$$
\begin{equation*}
\lambda_{0}=A-D=-2 e^{-z} \sinh x \tag{A4}
\end{equation*}
$$

The other three eigenvalues are found from the following characteristic determinant

$$
\left|\begin{array}{rcc}
(A+D)-\lambda & \sqrt{2} B & \sqrt{2} C  \tag{A5}\\
\sqrt{2} B & e^{-u}-\lambda & 1 \\
\sqrt{2} C & 1 & e^{u}-\lambda
\end{array}\right|=0
$$

By using

$$
\begin{aligned}
& A+D=2 e^{-z} \cosh x \\
& B^{2}+C^{2}=2 e^{-z} \cosh u
\end{aligned}
$$

one can easily prove

$$
\left|\begin{array}{ccc}
A+D & \sqrt{2} B & 2 C  \tag{A6}\\
\sqrt{2} B & e^{-u} & 1 \\
\sqrt{2} C & 1 & e^{u}
\end{array}\right|=0
$$

Therefore, another eigenvalue is $\lambda_{1}=0$. The rest of the eigenvalues are found from the quadratic equation

$$
\begin{equation*}
\lambda^{2}-2 \lambda\left(e^{-2} \cosh x+\cosh u\right)+4 e^{-z} \cosh u(\cosh x-1)=0 \tag{A7}
\end{equation*}
$$

The solutions are:

$$
\begin{align*}
\lambda_{2,3} & =\left(e^{-z} \cosh x+\cosh u\right) \pm\left[\left(e^{-z} \cosh x-\cosh u\right)^{2}\right. \\
& \left.+4 e^{-z} \cosh u\right]^{1 / 2} \tag{A8}
\end{align*}
$$

$\lambda_{2}>\lambda_{3}$ and $\lambda_{2}>\lambda_{1}$ are obvious. Also

$$
\lambda_{2}-\left|\lambda_{0}\right|>\left[\left(e^{-z} \cosh x+\cosh u\right)+\left(e^{-z} \cosh x-\cosh u\right)\right]
$$

$$
-2 e^{-z} \sinh x=2 e^{-z} \cosh x-2 e^{-z} \sinh x=e^{-(z+x)}>0
$$

Hence $\lambda_{2}>\left|\lambda_{0}\right|$

Therefore, $\lambda_{2}$ is the eigenvalue of maximum magnitude.

## APPENDIX B <br> WAVEFUNCTIONS OF 3－SITE <br> 4－ELECTRON LINEAR CLUSTER

Let＂＂丹＂denote the space orbital which is occu－ pied by a spin up electron and＂$\downarrow$＂by a spin down electron， ＂ 1 ＂denote the space orbital which is occupied by two alec－ trons with the order of spin up followed by spin down．If the space orbit isn＇t occupied，we denote it by＂ 0 ＂and the order of space orbit is ${\underset{\sim}{A}}_{0}, \underline{B}_{0}, \underline{A}_{1}$ ．With these conven－ tins we get the following 154 －electron wavefunctions on three sites：

$$
\begin{aligned}
& \left.\Phi_{1}=\frac{1}{\sqrt{2}}[|\uparrow \downarrow \downarrow\rangle-|\uparrow| \downarrow\rangle\right] \\
& \Phi_{2}=\frac{1}{2}[|\Phi \uparrow \downarrow\rangle-|\Phi \downarrow \uparrow\rangle+|\uparrow \downarrow \ddagger\rangle-|\downarrow \uparrow \Psi\rangle] \\
& \left.\Phi_{3}=\frac{1}{\sqrt{2}}[|I\{0\rangle+| 0 £ I\rangle\right] \\
& \Phi_{4}=| \pm 0 t\rangle \\
& \Phi_{5}=\frac{1}{\sqrt{2}}[|\uparrow \uparrow \downarrow\rangle+|\downarrow \Phi \uparrow\rangle] \\
& \Phi_{6}=\frac{1}{2}[| \pm \uparrow \downarrow\rangle+|\Psi \downarrow \uparrow\rangle+|\uparrow \downarrow \uparrow\rangle+|\downarrow \uparrow \uparrow\rangle] \\
& \Phi_{7}=\frac{1}{2}[|\uparrow \uparrow \downarrow\rangle+|\uparrow \downarrow \uparrow\rangle-|\uparrow \downarrow \pm\rangle-|\downarrow \uparrow \downarrow\rangle] \\
& \Phi_{8}=|\uparrow 亡 \uparrow\rangle
\end{aligned}
$$

$$
\begin{aligned}
& \psi_{9}=\frac{1}{\sqrt{2}}[|\leq \uparrow \uparrow\rangle+\mid \uparrow \uparrow\{ \rangle] \\
& \left.\psi_{10}=\frac{1}{\sqrt{2}}[|s \uparrow \uparrow\rangle-|\uparrow \uparrow| s\rangle\right] \\
& \left.\psi_{11}=\mid \downarrow \text { さ } \downarrow\right\rangle \\
& \Psi_{12}=\frac{1}{\sqrt{2}}[|\leq \downarrow \downarrow\rangle+|\downarrow \downarrow \tau\rangle] \\
& \left.\psi_{13}=\frac{1}{\sqrt{2}}[|z \downarrow \downarrow>-| \downarrow \downarrow \pm\rangle\right] \\
& \left.\psi_{14}=\frac{1}{\sqrt{2}}[|\neq \pm 0>-| 0 £ \&\rangle\right] \\
& \psi_{15}=\frac{1}{2}[| \pm \uparrow \downarrow\rangle-|\leq \downarrow \uparrow\rangle-|\uparrow \downarrow f\rangle+|+\uparrow \tau\rangle]
\end{aligned}
$$

These are eigenstates of $\vec{S}^{2}, S_{z}$ and I with eigenvalues given as follows:

| $\Psi_{i}$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| $\mathrm{~S}^{2}$ | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 0 |
| $\mathrm{~S}_{\mathrm{z}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | -1 | -1 | -1 | 0 | 0 |
| I | 1 | 1 | 1 | 1 | -1 | -1 | 1 | -1 | -1 | 1 | -1 | -1 | 1 | -1 | -1 |

From the above table, we see that the $15 \times 15$ matrix factors into one $4 \times 4$ submatrix from the singlets, four $2 \times 2$ submatrices, corresponding respectively to $\left(S^{2}, S_{z}, I\right)=(2,0,-1),(2,1,-1),(2,-1,-1)$, ( $0,0,-1$ ), and three 1 x 1 matrices.

## APPENDIX C <br> MATRIX ELEMENTS OF THE HAMILTONIAN OPERATOR WITH RESPECT TO DETERMINANTAL WAVEFUNCTIONS

The Hamiltonian operator for an $n$-electron system is taken to have the form

$$
H=\sum_{i} h(i)+\sum_{i<j} v(i, j)
$$

where $h(i)$ is the one-electron interaction term, $v(i, j)$ is the two-electron term. A Slater determinant of one-electron states $\Phi_{1} \ldots \Phi_{n}$ is defined as

$$
\Phi=\frac{1}{\sqrt{n!}}\left|\begin{array}{cccc}
\Phi_{1}(1) & \Phi_{1}(2) & \ldots & \Phi_{1}(n)  \tag{C2}\\
\Phi_{2}(1) & \Phi_{2}(2) & \ldots & \Phi_{2}(n) \\
\Phi_{n}(1) & \Phi_{n}(2) & & \Phi_{n}(n)
\end{array}\right| .
$$

We will assume that the $\phi_{i}$ are members of an orthonormal set. Equivalently, we may write

$$
\begin{align*}
\Phi= & \mathrm{R}\left[\phi_{1}(1) \quad \phi_{2}(2) \cdots \phi_{\mathrm{n}}(\mathrm{n})\right]=\frac{1}{\sqrt{\mathrm{n}!}} \\
& \sum_{\mathrm{p}}^{\Sigma}(-1)^{\mathrm{p}} \mathrm{P}\left[\phi_{1}(1) \phi_{2}(2) \cdots \phi_{\mathrm{n}}(\mathrm{n})\right] \tag{C3}
\end{align*}
$$

where the antisymmetrization operator $R$ is defined by this equation. The summation is over all n ! distinct permutations $P$ of the electrons among themselves, and ( -1$)^{\mathrm{p}}$ is +1 if the permutation entails an even number of pair interchanges, -1 if it involves an odd number. Let $\Phi$ and $\Phi^{\prime}$ be two different $n-e l e c t r o n$ wavefunctions, with $\Phi^{\prime}$ built up from $\phi_{1}{ }^{\prime} \ldots \phi_{\mathrm{n}}{ }^{\prime}$. Before calculating the matrix element, we assume that, by interchange of rows, $\Phi$ and $\Phi$ ' have already been put into maximum coincidence with each other. We then may distinguish four cases:

Case 1: $\quad \Phi$ and $\Phi^{\prime}$ are identical
Case 2: $\quad \Phi$ and $\Phi^{\prime}$ differ in one spin orbital with $\phi_{\mathrm{m}}$ entering $\Phi$ and $\phi_{\mathrm{p}}$ entering $\Phi^{\prime}$

Case 3: $\Phi$ and $\Phi^{\prime}$ differ in two spin orbitals with $\phi_{m}$ and $\phi_{\mathrm{n}}$ entering $\Phi$ where $\phi_{\mathrm{p}}$ and $\phi_{q}$ enter $\Phi^{\prime}$

$$
\begin{aligned}
\text { Case 4: } & \Phi \text { and } \Phi^{\prime} \text { differ in three or more spin } \\
& \text { orbitals }
\end{aligned}
$$

the overlap integral $S$ is straightforward to derive

$$
S=\left\langle\Phi, \Phi^{\prime}\right\rangle=\left\{\begin{array}{l}
1 \text { if } \Phi^{\prime}=\Phi^{\prime} \quad \text { (case 1) }  \tag{C4}\\
0 \text { otherwise (case } 2,3,4)
\end{array}\right.
$$

The matrix element of $H$ can be broken up into two pieces, one from the one-electron part, the other from the twoelectron part. The matrix element of the one-electron part is the following:

$$
\begin{aligned}
& \mathrm{I}=\langle\Phi| \sum_{\mathrm{i}} \mathrm{~h}(\mathrm{i})\left|\Phi^{\prime}\right\rangle \\
& =\frac{1}{n!} \sum_{p}(-1)^{p} P[\phi(1) \phi(2) \quad \ldots]^{*} \sum_{i} h(i)
\end{aligned}
$$

$$
\begin{aligned}
& =\int\left[\phi_{1}(1) \phi_{2}(2) \quad \ldots\right]^{*} \sum_{i} h(i) \sum_{Q}(-1)^{Q} \\
& Q\left[\phi_{1}^{\prime}(1) \phi_{2}^{\prime}(2) \ldots\right] d \tau(1) d \tau(2) \ldots \\
& =\int\left[\phi_{1}(1) \phi_{2}(2) \ldots\right]^{*} \sum_{i} h(i)\left[\phi_{1}^{\prime}(1) \phi_{2}^{\prime}(2) \ldots\right] \\
& d \tau(1) \quad d \tau(2) \ldots
\end{aligned}
$$

The $4^{\text {th }}$ line follows from the fact that each of the permutations $P$ merely affects the labeling of the variables of integration. The $6^{\text {th }}$ line follows from the fact that $\mathrm{Ch}(\mathrm{i})$ is a sum of one-electron operators; any nontrivial permutation $P$ produces two noncoincidences of one-electron states one of which integrates to zero because of the orthogonality. We then obtain immediately

$$
I=\left\{\begin{array}{cl}
\sum\left\langle\Phi_{i}\right| h(1)\left|\phi_{i}\right\rangle & \text { for case } 1  \tag{C6}\\
i & \text { for case } 2 \\
\left.<\phi_{n}|h(1)| \phi_{\mathrm{p}}\right\rangle & \text { for case } 3 \text { and } 4
\end{array}\right.
$$

For the matrix element of the two-electron part, the reduction proceeds similarly although now of all the permutations $Q$ we must retain for each term $v(i, j)$ the permutation that interchanges electrons $i$ and $j$

$$
\begin{align*}
& I I \equiv\langle\Phi| \underset{i<j}{ } v(i, j)\left|\Phi^{\prime}\right\rangle \\
& \left.=\int\left[\begin{array}{lll} 
\\
\phi_{1}(1) \phi_{2}(2) & \ldots
\end{array}\right]^{*} \underset{i<j}{\sum} v_{i j}\right) \underset{P}{\sum(-1)^{P}} \\
& P\left[\phi_{1}^{\prime}(1) \phi_{2}^{\prime}(2) \quad . . .\right] d \tau(1) d \tau(2) \ldots  \tag{CT}\\
& \left.=\left[\begin{array}{lll}
\phi_{1} & (1) & \phi_{2}(2) \\
\ldots
\end{array}\right]^{*} \underset{i<j}{\sum} v_{i j}\right) \\
& {\left[\phi_{1}^{\prime}(1) \quad \phi_{2}^{\prime}(2) \quad \ldots \phi_{i}^{\prime}(i) \quad \phi_{j}^{\prime}(j) \quad \ldots\right.} \\
& \left.-\phi_{1}^{\prime}(1) \quad \phi_{2}^{\prime}(2) \quad \ldots \phi_{i}^{\prime}(j) \quad \ldots \phi_{j}^{\prime}(i) \ldots\right] \\
& d \tau(1) d \tau(2) \ldots . . . . \quad d \tau(i) \ldots d \tau(j) \ldots
\end{align*}
$$

This gives different answers for each of cases 1, 2, 3 and 4.

## APPENDIX D <br> VARIOUS ORDERS OF ENERGY CORRECTION ON THE PERTURBATION THEORY

Suppose the total Hamiltonian $H$ is divided into unperturbed part $H_{o}$ and perturbed part $V$ :

$$
H=H_{0}+\lambda V
$$

and

$$
\begin{aligned}
H_{n} & =E_{n} \Phi_{n} \\
H_{0} \Phi_{n}^{\circ} & =E_{n}^{\circ} \Phi_{n}^{\circ}
\end{aligned}
$$

therefore

$$
E_{n}=E_{n}^{\circ}+\lambda E_{n}^{(1)}+\lambda^{2} E_{n}^{(2)}+\ldots
$$

If $E_{n}^{\circ}$ is nondegenerate we have the formula up to fourth ${ }^{15,16}$ order, as follows:

Let

$$
\begin{aligned}
E_{n}^{(1)} & =\left\langle\Phi_{n}^{0}\right| v\left|\Phi_{n}^{0}\right\rangle \\
v_{i j} & \left.=\left.\left\langle\Phi_{i}^{0}\right| v\right|_{j} ^{\circ}\right\rangle
\end{aligned}
$$

then

$$
E_{n}^{(2)}=-\sum_{i \neq n} \frac{\left|v_{i n}\right|^{2}}{E_{i}^{(0)}-E_{n}^{(0)}}
$$

$E_{n}^{(3)}=\sum_{i \neq n} \sum_{j \neq n} \frac{v_{n i} v_{i j} v_{j n}}{\left(E_{i}^{(0)}-E_{n}^{(0)}\right)\left(E_{j}^{(0)}-E_{n}^{(0)}\right)}-v_{n n} \sum_{i \neq n} \frac{\left|v_{i n}\right|^{2}}{\left(E_{i}^{(0)}-E_{n}^{(0)}\right)^{2}}$,

$$
\begin{aligned}
& E_{n}^{(4)}=\sum_{i \neq n} \frac{\left|v_{i n}\right|^{2}}{E_{i}^{(0)}-E_{n}^{(0)}} \sum_{j \neq n} \frac{\left|v_{j n}\right|^{2}}{\left(E_{j}^{\circ}-E_{n}^{\circ}\right)^{2}} \\
& +v_{n n} \sum_{i \neq n} \sum_{j \neq n} \frac{v_{n i} v_{i j} v_{j n}}{\left(E_{i}^{(0)}-E_{n}^{(0)}\right)^{2}\left(E_{j}^{(0)}-E_{n}^{(0)}\right)} \\
& +v_{n n} \sum_{i \neq n} \sum_{j \neq n} \frac{v_{n i} v_{i j} v_{j n}}{\left(E_{i}^{(0)}-E_{n}^{(0)}\right)\left(E_{j}^{(0)}-E_{n}^{(0)}\right)} \\
& -\sum_{i \neq n} \sum_{j \neq n} \sum_{k \neq n} \frac{v_{n i} v_{i j} v_{j k} v_{k n}}{\left(E_{i}^{(0)}-E_{n}^{(0)}\right)\left(E_{j}^{(0)}-E_{n}^{(0)}\right)\left(E_{k}^{(0)}-E_{n}^{(0)}\right)} \\
& -v_{n n} \sum_{i \neq n} \frac{\left|v_{i n}\right|^{2}}{\left(E_{i}^{(0)}-E_{n}^{(0)}\right)^{3}}
\end{aligned}
$$

## APPENDIX E

## THE MATRIX ELEMENT FOR NEAREST -NEIGHBOR HOPPING OF ONE ELECTRON

$$
\begin{aligned}
& =\left\langle A_{0} \downarrow\right| h \mid B_{o} \downarrow>+\sum_{\ell \sigma}^{\ell \sigma \neq B_{0}}\left\langle A_{o} \downarrow \Psi_{\ell \sigma}\right| \hat{v}\left|B_{o} \downarrow \Psi_{\ell \sigma}\right\rangle \\
& =\left\langle A_{0}\right| h\left|B_{0}\right\rangle+\sum_{i \sigma}\left\langle A_{o} \downarrow A_{i \sigma}\right| \hat{v}\left|B_{o} \downarrow A_{i \sigma}\right\rangle+\sum_{j \sigma}^{j \sigma \neq B_{o} \downarrow} \\
& <A_{o} \downarrow B_{j \sigma}|\hat{v}| B_{o} \downarrow B_{j 0}> \\
& =\left\langle A_{0}\right| h\left|B_{0}\right\rangle+\sum_{i}\left\langle A_{0} A_{i}\right| v\left|B_{0} A_{i}\right\rangle+2 \underset{j \neq 0}{\sum}\left\langle A_{0} B_{j}\right| v\left|B_{o} B_{j}\right\rangle \\
& +\left\langle\mathrm{A}_{0} \mathrm{~B}_{0}\right| v\left|\mathrm{~B}_{0} \mathrm{~B}_{0}\right\rangle-\sum_{\mathrm{i} \sigma}\left\langle\mathrm{~A}_{0} \downarrow \mathrm{~A}_{\mathrm{i} \sigma}\right| v \mid \mathrm{A}_{i \sigma} \mathrm{~B}_{0} \downarrow> \\
& -\sum_{j \sigma}^{\Sigma\left\langle A_{o}\right.} \downarrow B_{j \sigma}|v| B_{j \sigma} B_{o}^{\downarrow>} \\
& =\left\langle A_{0}\right| h\left|B_{0}\right\rangle+\sum_{i}\left\langle A_{0} A_{i}\right| v\left|B_{0} A_{i}\right\rangle+2 \sum_{j \neq 0}\left\langle A_{o} B_{j}\right| v\left|B_{o} B_{j}\right\rangle \\
& +\left\langle A_{0} B_{0}\right| v\left|B_{0} B_{0}\right\rangle-\sum_{i \sigma}^{i \neq 0}<A_{o} \downarrow A_{i \sigma}|v| A_{i \sigma} B_{o} \downarrow> \\
& -2 \underset{j \neq 0}{\sum}<A_{0} B_{j}|v| B_{j} B_{o}>
\end{aligned}
$$

# APPENDIX F <br> DEGENERATE PERTURBATION APPROACH VIA EFFECTIVE HAMILTONIAN 

A Hamiltonian can be partitioned into an unperturbed part $H_{o}$ and a perturbation part $V$

$$
\begin{equation*}
\mathrm{H}=\mathrm{H}_{\mathrm{o}}+\mathrm{V} \tag{F1}
\end{equation*}
$$

We want to solve

$$
\begin{equation*}
H \Phi_{n}=E_{n} \Phi_{n} \tag{F2}
\end{equation*}
$$

and we know

$$
\begin{equation*}
H_{0} \Phi_{n}{ }^{\circ}=E_{n}^{0} \Phi_{n}^{\circ} \tag{F3}
\end{equation*}
$$

we can write

$$
\begin{equation*}
\Phi_{\mathrm{n}}=\Phi_{\mathrm{n}}^{\circ}+\eta_{\mathrm{n}}, \quad E_{\mathrm{n}}=\mathrm{E}_{\mathrm{n}}^{\circ}+\Delta_{\mathrm{n}} . \tag{F4}
\end{equation*}
$$

Assume (F3) has a degenerate ground level $\mathrm{E}_{1}^{0}=\mathrm{E}_{2}^{0}=\mathrm{E}_{3}^{0}=\ldots$ $\mathrm{E}_{\mathrm{g}}^{0}=\varepsilon$; $\quad \mathrm{E}_{\mathrm{n}}^{0}>\varepsilon$ for $\mathrm{n}>\mathrm{g}$. We also assume $\mathrm{H}_{\mathrm{o}}$ is hermitean so that we can take $\Phi_{n}^{\circ}$ such that

$$
\begin{equation*}
\left\langle\Phi_{\mathrm{n}}^{\circ} \mid \Phi_{\mathrm{m}}^{\circ}\right\rangle=\delta_{\mathrm{nm}} \tag{F5}
\end{equation*}
$$

It is natural to call the degenerate ground state manifold $\left(\Phi_{1}^{\circ}, \Phi_{1}^{\circ}, \ldots \Phi_{g}^{\circ}\right)=$ subspace $G$, and all the excited states $\left(\Phi_{\mathrm{g}+1}^{\circ} \cdots\right)=$ subspace X . From now on, $\Phi_{\mathrm{n}}^{\circ}$ stands for the projection of $\Phi_{n}$ onto $G$, this projection being normalized to unity. Thus $\eta_{n}$ is in $X$ and

$$
\begin{equation*}
\left\langle\Phi_{\mathrm{n}} \mid \Phi_{\mathrm{n}}\right\rangle=1+\left\langle n_{\mathrm{n}} \mid n_{\mathrm{n}}\right\rangle \tag{F6}
\end{equation*}
$$

That is, $\Phi_{\mathrm{n}}$ is not normalized to unity; however, we can always normalize $\Phi_{n}$ by dividing by the square root of (F6) of course.

Substituting (F4) into (F1) gives

$$
\left(H_{0}+V\right)\left(\Phi_{n}^{\circ}+\eta_{n}\right)=\left(E_{n}^{o}+\Delta_{n}\right)\left(\Phi_{n}^{o}+\eta_{n}\right)
$$

or $\quad\left(H-E_{n}\right) \eta_{n}=\left(E_{n}-H\right) \Phi_{n}^{\circ}=\left(\Delta_{n}-V\right) \Phi_{n}^{\circ}$
or $\quad\left(H_{o}-E_{n}^{\circ}\right) \eta_{n}+\left(V-\Delta_{n}\right) \eta_{n}=-\left(V-\Delta_{n}\right) \Phi_{n}^{o}$.

If one defines idempotent operators $P$ and $Q$ such that $P \Psi$ belongs to $G$ and $Q \Psi$ belongs to $X$, then we get $P+Q=1$, $P^{2}=P, Q^{2}=Q$, and $P Q=Q P=0$. Clearly we also have $P \eta_{n}=0$, $\theta \eta_{n}=\eta_{n}, P \Phi_{n}^{\circ}=\Phi_{n}^{\circ}$, and $Q \Phi_{n}^{\circ}=0$, for $n \leq g$. Since $H_{o}$ is diagonal in $\left\{\Phi_{n}^{\circ}\right\}$, equation (F7) can be split into two equations with these relations by applying the operators $P$ and $Q$ on both sides.

$$
\begin{align*}
& P V Q \eta_{n}=P\left(V-\Delta_{n}\right) P \Phi_{n}^{\circ}  \tag{F9a}\\
& Q\left(H-E_{n}\right) Q n_{n}=-Q V P \Phi_{n}^{\circ} \tag{F9b}
\end{align*}
$$

(F9b) gives

$$
\begin{equation*}
Q \eta_{n}=-\frac{1}{Q\left(H-E_{n}\right) Q} \quad Q V P \Phi_{n}^{\circ} \tag{F10}
\end{equation*}
$$

and (F9a) becomes

$$
\begin{equation*}
\left[P V P-P V Q \frac{1}{Q\left(H-E_{n}\right) Q} \quad Q V P\right] \Phi_{n}^{\circ}=\Delta_{n} P \Phi_{n}^{\circ} \tag{F11}
\end{equation*}
$$

By adding $\mathrm{PH}_{0} \mathrm{P}$ to the left hand side and $E_{n}$ to the right we have

$$
\begin{equation*}
H_{s} \Phi_{n}^{\circ}=E_{n} \Phi_{n}^{\circ} \tag{F12}
\end{equation*}
$$

where the "effective Hamiltonian" $H_{s}$ is

$$
\begin{equation*}
H_{s} \equiv P\left[H-V \frac{1}{Q\left(H-E_{n}\right) Q} V\right] P \tag{F13}
\end{equation*}
$$

This is of course a gxg matrix. If $V$ is "small", we expand
$\frac{1}{Q\left(H-E_{n}\right) Q}=\frac{1}{Q\left(H_{0}-E_{n}^{\circ}\right) Q} \sum_{i=0}^{\infty}(-1)^{i}\left[\frac{1}{Q\left(H-E_{n}^{\circ}\right) Q} Q\left(V-\Delta_{n}\right) Q\right]^{i}$
Up to the $4^{\text {th }}$ order, the effective Hamiltonian become $H_{s}=P\left\{H_{o}+V-V \frac{1}{H_{0}-E_{0}} V+V \frac{1}{H_{0}-E_{0}} V \frac{1}{H_{0}-E_{0}} V\right.$
$\left.-V \frac{1}{H_{0}-E_{0}} V \frac{1}{H_{0}-E_{0}} V \frac{1}{H_{0}-E_{0}} V\right\}_{P}$
One might note ${ }^{17}$ that this perturbation theory differs in an essential way from the "standard" textbook degenerate perturbation theory ${ }^{16}$, in that here the "proper zero-order states" $\Phi_{n}^{\circ}$ are allowed to change as higher-order corrections are added, whereas in the standard method these $\Phi_{n}^{\circ}$ are fixed by the lowest order correction that removes the degeneracy. This difference can drastically affect the rate of convergence of the expansion.

## APPENDIX G

THE EFFECTIVE HAMILTONIAN IN SECOND ORDER PERTURBATION THEORY

From equation (3) of Section $V$

$$
\begin{align*}
H= & \sum_{i j} \sum_{\sigma}^{\sum} h_{i j} c_{i \sigma}^{+} c_{j \sigma}+\frac{1}{2} \sum_{i j}^{\sum} \sum_{k \ell} \sum_{\sigma \sigma} v_{i j, k \ell} \\
& c_{i \sigma}^{+} c_{j \sigma}^{+}, c_{\ell \sigma^{\prime}} c_{k \sigma} \tag{Gi}
\end{align*}
$$

$H_{0}$ is defined as all terms of $0\left(\Delta^{\circ}\right)$ :

$$
\begin{align*}
H_{0}= & \underset{i}{\sum \sum{ }_{\sigma} h_{i i} c_{i \sigma}^{+} c_{i \sigma}+\frac{1}{2} \sum_{i j}^{\sum} \underset{\sigma \sigma^{\prime}}{\Sigma} v_{i j, i j}} \\
& c_{i \sigma^{+}}^{+} c^{+}{ }_{\sigma^{\prime}} c_{j \sigma^{\prime}} c_{i \sigma} \\
= & \sum_{i} h_{i i} n_{i}+\sum_{i}^{\Sigma} v_{i i, i i} n_{i \uparrow} n_{i \downarrow} \\
+ & \frac{1}{2} \underset{i j}{\sum_{i j}^{\prime}} v_{i j, i j} n_{i} n_{j} \tag{G2}
\end{align*}
$$

The perturbation term is defined as $\mathrm{V}=\mathrm{H}-\mathrm{H}_{\mathrm{O}}$. The projection operator $P$ projects on those states with

$$
n_{i \sigma}=\left\{\begin{array}{lc}
1 & \text { if i on a s-site }  \tag{GB}\\
2 & \text { i on } b-\text { site }
\end{array}\right.
$$

If $D_{\nu}$ are single Slater determinants with Wanner functions singly occupied on magnetic sites, and doubly occupied on diamagnetic sites, then we have

$$
\begin{gather*}
\left\langle D_{\nu}\right| H_{o}\left|D_{\mu}\right\rangle=E_{\nu}^{o} \delta_{\nu \mu} \equiv \varepsilon \delta_{\nu \mu}  \tag{G4}\\
E_{\nu}^{\circ} \equiv \varepsilon=\left\langle D_{\nu}\right| H\left|D_{\nu}\right\rangle=N\left(h_{a a}+h_{b b}+v_{b b}, b b\right)+\frac{1}{2} \sum_{a a^{\prime}} \\
v_{a a^{\prime}, a a^{\prime}}+2 \sum_{b b^{\prime}}^{v_{b b^{\prime}}, b b^{\prime}}+\sum_{a i} v_{a b}, a b \tag{GS}
\end{gather*}
$$

If there are $N$ a-sites (magnetic sites), the degeneracy of $E_{V}^{\circ}$ will be $2^{N}$. In order to apply the perturbation theory, the one-electron and two-electron interaction terms are analyzed as follows:

$$
\begin{equation*}
H_{1}=\sum_{i j} \sum_{\sigma} h_{i j} c_{i \sigma}^{+} c_{j \sigma}=Q_{i}+i \nrightarrow j \tag{G6}
\end{equation*}
$$

$Q_{i}$ refers to the terms for which $i=j$ and $i \neq j$ refers to terms with $i \neq j$.

$$
\begin{align*}
& H_{2}=\frac{1}{2} \sum_{i j}^{\sum \sum \sum_{k \ell \sigma \sigma^{\prime}}} v_{i k, j \ell} c_{i \sigma}^{+} c_{k \sigma^{\prime}}^{+} c_{\ell \sigma^{\prime}} c_{j \sigma} \\
& =\frac{1}{2} \sum_{i j} v_{i i, j j} c_{i \sigma}^{+} c_{i \sigma}^{+} c_{j \sigma}, c_{j \sigma}+\frac{1}{2} \sum_{j \ell}^{\prime} \sum_{i} \sum_{\sigma \sigma^{\prime}} \\
& \mathrm{v}_{\mathrm{ii}, \mathrm{j} \ell}^{+} \mathrm{c}_{\mathrm{i} \sigma}^{+} \mathrm{c}_{\mathrm{i} \sigma}^{+}, \mathrm{c}_{\ell \sigma}, \mathrm{c}_{\mathrm{j} \sigma}+\frac{1}{2} \sum_{\mathrm{ik}}^{\mathrm{\prime}} \sum_{j} \sum_{\sigma \sigma^{\prime}} \mathrm{v}_{\mathrm{ik}, \mathrm{jj}} \\
& c_{i \sigma}^{+} c_{k \sigma^{\prime}}^{+} c_{j \sigma^{\prime}} c_{j \sigma}+\frac{1}{2} \sum_{i k}^{\prime} \sum_{j \ell}^{\prime} \sum_{\sigma \sigma^{\prime}} v_{i k, j \ell} \\
& c_{i \sigma}^{+} c_{k \sigma^{\prime}}^{+}{ }_{\ell \sigma^{\prime}} c_{j \sigma}=A+B+C+D \tag{G7}
\end{align*}
$$

where $\bar{\sigma}=-\sigma$, and $A, B, C, D$ stand respectively for the four sums of the previous expression, taken in the same order as they appear. They may be represented by differen types of graphs, in which each line corresponds to a $C^{+} C$ pair with the same spin index.
A:


B :

C:


D:

$+D_{2}\left[\underset{\sigma^{\circ}}{ } \int_{j=k}^{\ell}\right.$

$+Y[i=\ell \underset{\sigma}{\underset{\sigma}{\leftrightarrows}} \underset{\sim}{\sigma}=k]$

Let us first consider the leading term PHP in the effective Hamiltonian, eq. (F15). The projection operators demand that only matrix elements of $H$ among the G-states occur. Clearly $C_{i \sigma}^{+} C_{j \sigma} G$ is in $G$ if and only if $i=j$; thus only the diagram $\circlearrowleft$ contributes from $H$. As to the two-electron interaction, we see similarly that only $A_{0}, D_{0}$ and $Y$ contribute. Hence,
$H_{s}(1)=P H P=P \sum_{i \sigma} c_{i \sigma}^{+} c_{i \sigma}{ }^{P+} \frac{1}{2} P \sum_{i j}^{\sum} \sum_{\sigma \sigma^{\prime}} v_{i j}, i j n_{i \sigma} n_{j \sigma^{\prime}} P$

$$
\begin{equation*}
+\frac{1}{2} P \sum_{i j \prime} \sum_{\sigma \sigma^{\prime}} v_{i j, j i} c_{i \sigma^{\prime}}^{+} c_{j \sigma^{\prime}}^{+} c_{i \sigma^{\prime}} c_{j \sigma^{\prime}} \tag{G9}
\end{equation*}
$$

The first two terms of the above equation only give a constant, while the last term will give a Heisenberg Hamiltonian, partially from the flip of a part of the electron spins on different a-sites. We can see this from the following: Here $i, j$ indicate different a-sites, and we use

$$
\begin{array}{ll}
\mathrm{Pn}_{i} P=\frac{1}{2}+S_{i z}, & \mathrm{Pn}_{i} P=\frac{1}{2}-S_{i z} \\
\mathrm{Pc}_{i \uparrow}^{+} \mathrm{c}_{i \downarrow} P=S_{i+}, & \mathrm{Pc}_{i \downarrow}^{+} \mathrm{c}_{i \uparrow} \mathrm{P}=\mathrm{S}_{\mathrm{i}-} \tag{G10}
\end{array}
$$

Clearly

$$
\begin{gathered}
\frac{1}{2} \sum_{i j} \sum_{\sigma \sigma^{\prime}} v_{i j, j i} c_{i \sigma}^{+} c_{j \sigma}^{+}, c_{i \sigma}, c_{j \sigma} \\
=-\frac{1}{2} \sum_{i j} \sum_{\sigma} v_{i j, j i}\left[n_{i \sigma} n_{j \sigma}+c_{i \sigma}^{+} c_{i \sigma} c_{j \sigma}^{+} c_{j \sigma}\right]=-\frac{1}{2} \sum_{i j} v_{i j, j i}
\end{gathered}
$$

$$
\begin{align*}
& {\left[n_{i \uparrow} n_{j \uparrow}+n_{i \nless} n_{j \downarrow}+S_{i+} S_{j-}+S_{i-} S_{j+}\right]=\frac{1}{2} \sum_{i j} v_{i j, j i}} \\
& {\left[\left(\frac{1}{2}+S_{i z}\right)\left(\frac{1}{2}+S_{j z}\right)+\left(\frac{1}{2}-S_{i z}\right)\left(\frac{1}{2}-S_{j z}\right)+S_{i+} S_{j-}\right.} \\
& \left.+S_{i-} S_{j+}\right]=-\frac{1}{2} \sum_{i j} v_{i j, j i}\left(\frac{1}{2}+2 \vec{S}_{i} \cdot \vec{S}_{j}\right) \tag{G11}
\end{align*}
$$

The contribution from $2^{\text {nd }}$ order perturbation theory has the form

$$
\begin{equation*}
H_{s}^{(2)}=-P H Q \frac{1}{Q\left(H_{0}-\varepsilon\right) Q} Q H P \tag{G12}
\end{equation*}
$$

But

$$
H=H+A+B_{1}+B_{2}+C_{1}+C_{2}+D_{0}+D_{1}+D_{2}+d_{2}+Y
$$

and we see from the diagram that $Q A P=0 . \quad Q Y P=0, \quad Q B_{2} P=0$, and $Q D_{0} P=0$. Therefore

$$
\begin{align*}
& \mathrm{QHP}=\mathrm{Q}\left(\mathrm{H}_{1}+\mathrm{B}_{1}+\mathrm{C}_{1}+\mathrm{D}_{1}+\mathrm{C}_{2}+\mathrm{D}_{2}+\mathrm{d}_{2}\right) \mathrm{P}  \tag{G13}\\
& \mathrm{PHQ}=[\mathrm{QHP}]^{\dagger}=\mathrm{P}\left(\mathrm{H}_{1}+\mathrm{C}_{1}+\mathrm{B}_{1}+\mathrm{D}_{1}+\mathrm{B}_{2}+\mathrm{D}_{2}+\mathrm{d}_{2}\right) \mathrm{Q} \tag{G14}
\end{align*}
$$

For the $d_{2}$ graph, the following operation will flip the spin on different a-sites.


Although this operation gives a contribution to the Heisenberg Hamiltonian, we see immediately that the overlap is higher order than $0\left(\Delta^{4}\right)$ no matter whether we consider a linear chain or perovskite lattice. Therefore, we can drop $\mathrm{d}_{2}$. Let

$$
\mathrm{R} \equiv \mathrm{H}_{1}+\mathrm{B}_{1}+\mathrm{C}_{1}+\mathrm{D}_{1}+\mathrm{D}_{2}
$$

Aside from numerical factors, the effect of operator $R$ is equal to $c_{i \sigma}^{+} c_{k \sigma}$ where $i$ has to be on an a-site, and $k$ can be on an a-or b-site. Therefore

$$
\begin{align*}
& \mathrm{QHP}=\mathrm{Q}\left(\mathrm{R}+\mathrm{C}_{2}\right) \mathrm{P} \\
& \mathrm{PHQ}=\mathrm{P}\left(\mathrm{R}+\mathrm{B}_{2}\right) \mathrm{Q} \tag{G15}
\end{align*}
$$

Because we start from a ground state, and must come back to a ground state, it is very easy to see from the graph that
$P R Q \frac{1}{Q(H-\varepsilon) Q} Q C_{2} P=P B_{2} Q \frac{1}{Q(H-\varepsilon) Q} Q R P=0$.
Therefore $H_{s}{ }^{(2)}$ arises from two kinds of term. The first one is called the Nesbet term; it comes from the "double hopping" i.e.

$$
-\mathrm{PB}_{2} \mathrm{Q} \frac{1}{Q\left(\mathrm{H}_{0}-\varepsilon\right) \mathrm{Q}} \mathrm{QC}_{2} \mathrm{P}
$$

$=-\sum_{i j \ell \sigma \sigma^{\prime}}^{\sum \sum_{i j, \ell}-\varepsilon} \frac{\left|v_{i j, \ell \ell}\right|^{2}}{U_{\ell \sigma}}+c_{\ell \bar{\sigma},}^{+} c_{j \sigma}, c_{i \sigma}, c_{i \sigma}^{+} c_{j \bar{\sigma}}^{+} c_{\ell \bar{\sigma}} c_{\ell \sigma}$

Here $i, j$ have to be on different a-sites but $\ell$ is on one $b-$ site, and $U_{i j, \ell}=\left\langle c_{j}^{+}{ }_{j} c_{\ell \bar{\sigma}} C_{i \sigma}^{+}{ }^{c}{ }_{\ell \sigma} v\right| H_{o}|v\rangle$. $|v\rangle$ is a ground state. The above expression can further become

$$
=\sum_{i j} \sum_{\ell} \frac{\left|v_{i j, \ell \ell}\right|^{2}}{U_{i j, \ell}-\varepsilon}\left[\frac{1}{2}+2 S_{i z} S_{j z}+S_{i+} S_{j-}+S_{i-} S_{j+}\right]
$$

$$
\begin{equation*}
=\sum_{i j} \sum_{\ell} \frac{\left|v_{i j, \ell \ell}\right|^{2}}{U_{i j, \ell}-\varepsilon}\left[\frac{1}{2}+2 \vec{S}_{i} \cdot \vec{S}_{j}\right] \tag{G18}
\end{equation*}
$$

Another term in the $2^{\text {nd }}$ order perturbation theory is called the kinetic exchange term. It comes from the "single hopping", and has the form

$$
\begin{aligned}
& =-\sum_{i j} \sum_{\ell} \frac{\left|v_{i j, \ell \ell}\right|^{2}}{U_{i j, \ell}-\varepsilon}\left[-1+n_{i \uparrow} n_{j \downarrow}+n_{i \neq} n_{j \uparrow}-S_{i+} S_{j-}-S_{i-} S_{j+}\right] \\
& =-\sum_{i j} \sum_{\ell} \frac{\left|v_{i j, \ell \ell}\right|^{2}}{U_{i j, \ell}-\varepsilon}\left[-1+\left(\frac{1}{2}+S_{i z}\right)\left(\frac{1}{2}-S_{j z}\right)+\left(\frac{1}{2}-S_{i z}\right)\right. \\
& \left.\left(\frac{1}{2}+S_{j z}\right)-S_{i+} S_{j-}-S_{i-} S_{j+}\right]
\end{aligned}
$$

$$
\begin{equation*}
\sum_{\gamma} \frac{\langle\nu| H|\gamma\rangle\langle\gamma| H|\mu\rangle}{E_{\gamma \nu}} \tag{G19}
\end{equation*}
$$

where $|\nu\rangle,|\mu\rangle$ are ground states, and $|\gamma\rangle$ is an excited state. But

$$
\begin{gathered}
\langle\gamma| H|\mu\rangle=\langle\gamma| H_{1}+B_{1}+C_{1}+D_{1}+D_{2}|\mu\rangle=\langle\gamma| \mathrm{R}|\mu\rangle ; \\
|\gamma\rangle=c_{i \sigma}^{+} c_{k \sigma}|\mu\rangle \\
\langle\gamma| H|\mu\rangle=\left\langle c_{i \sigma}^{+} c_{k \sigma} \mu\right| H|\mu\rangle .
\end{gathered}
$$

If $k$ is on $a$-site, this matrix element is $0\left(\Delta^{3}\right)$ for the TSDA basis from eq. (10) of Section V. If $k$ is on an a-site, $D_{2}$ is $0\left(\Delta^{2}\right)$, therefore, we can drop $D_{2}$. Hence the operator corresponding to (G19) is

$$
\begin{equation*}
-P\left(H_{1}+C_{1}+B_{1}+D_{1}\right) Q \frac{1}{Q\left(H_{0}-\varepsilon\right) Q} Q\left(H_{1}+B_{1}+C_{1}+D_{1}\right) P . \tag{G20}
\end{equation*}
$$

Clearly

$$
\begin{aligned}
& \mathrm{Q}\left(\mathrm{H}_{1}+\mathrm{B}_{1}+\mathrm{C}_{1}+\mathrm{D}_{1}\right) \mathrm{P}
\end{aligned}
$$

where

$$
\begin{aligned}
\tau_{i j} & =h_{i j}+v_{i i, j i} n_{i \bar{\sigma}}+v_{i j, j j} n_{j \bar{\sigma}} \\
& +\sum_{\ell}^{i j} v_{i \ell, j \ell} n_{\ell}
\end{aligned}
$$

and $\Sigma^{i, j}$ means sum over $\ell$ with $\ell \neq i, j$. Taking the matrix element of (G20) between ground states, we were able to do the intermediate state sum over the $\gamma$ which contribute in order $\Delta^{4}$. Because the energy denominator is a constant for all such $\gamma$, we obtain

Here $i, j$ run over a -sites only, and

$$
\begin{aligned}
& \mathrm{U}_{\mathrm{ij}}=\left\langle\mathrm{c}_{\mathrm{i}}^{+} \mathrm{c}_{\mathrm{j}}^{\left.\mu\left|\mathrm{H}_{\mathrm{o}}\right| \mu\right\rangle-\varepsilon}\right. \\
& \overline{\mathrm{h}}_{\mathrm{ij}}=\mathrm{h}_{\mathrm{ij}}+\sum_{\ell}^{\mathrm{i} j} \mathrm{v}_{\mathrm{i} \ell, \mathrm{j} \ell}
\end{aligned}
$$

Using the expression of $\tau_{i j \sigma}$ and making a little bit of algebraic manipulation, the part of the effective Hamiltonia coming from this term is found to be

$$
\begin{aligned}
& -\sum_{i j}{ }^{\prime} \frac{1}{U_{i j}} \sum_{\sigma}\left[\left(\bar{h}_{i j}+v_{j i}^{*} n_{j \bar{\sigma}}\right) n_{i \sigma}\left(\bar{h}_{j i}+v_{j i} n_{j \sigma}\right)-\left|\bar{h}_{i j}\right|^{2}\right. \\
& \left.n_{i \sigma} n_{j \sigma}-\left(\bar{h}_{i j}+v_{j i}^{*}\right) \quad c_{i \sigma}^{+} c_{i \sigma} \bar{c}_{j \sigma}^{+} c_{j \sigma}\left(\bar{h}_{j i}+v_{j i}\right)\right] \\
& =-\Sigma^{\prime} \frac{\left|\bar{h}_{i j}\right|^{2}}{U_{i j}}-\Sigma^{\prime} \frac{1}{U_{i j}} \sum_{\sigma}\left[\bar{h}_{i j} v_{j i}{ }^{n}{ }_{i \sigma} n_{j} \bar{\sigma}+\bar{h}_{j i} v_{j i}^{*}{ }^{n}{ }_{i} \sigma^{n}{ }_{j} \bar{\sigma}\right.
\end{aligned}
$$

$$
80
$$

$$
\begin{aligned}
& +\left|v_{j i}\right|^{2} n_{i \sigma} n_{j \sigma}-\left|\bar{h}_{i j}+v_{j i}^{*}\right|^{2}\left(n_{i \sigma} n_{j \sigma}+c_{i \sigma}^{+} c_{i \sigma} c_{j \sigma}^{+} c_{j \sigma}\right) \\
& \left.+\left(\bar{h}_{i j} v_{j i}+\bar{h}_{j i} v_{j i}^{*}+\left|v_{j i}\right|^{2}\right) n_{i \sigma}^{n}{ }_{j \sigma}\right] \\
& =-\frac{1}{2} \sum \frac{\left|t_{i j}\right|^{2}}{U_{i j}}+\sum_{i j}^{\prime} \frac{2\left|t_{i j}\right|^{2}}{U_{i j}} \vec{S}_{i} \cdot \vec{S}_{j}
\end{aligned}
$$

(G23)
where

$$
t_{i j}=h_{i j}+\sum_{\ell} 2 v_{i \ell, j \ell}+v_{i j, j j}+\sum_{k}^{i j} v_{i k, j k}
$$



