

## Exchange Interaction and Heisenberg's Spin Hamiltonian\*

TADASHI ARAI

*Argonne National Laboratory, Argonne, Illinois*

(Received June 26, 1961)

By the use of representation matrices  $U(P)$  of the permutation group, the energy matrix of many-electron systems has been expanded such that  $E = \sum_P J_P \bar{U}(P)$ , and it has been proved that all the coefficients  $J_P$  are bounded and determined uniquely. This means that the expansion is mathematically valid even though nonorthogonal orbitals are used and no matter how large the overlaps between the orbitals are. Furthermore, it has been shown that the *nonorthogonality catastrophe* which was pointed out by Inglis and Slater does not appear and values of the coefficients can be evaluated correctly even if higher permutations are omitted. Then we find the Heisenberg spin operator as the first-order approximation of the expansion.

### I. INTRODUCTION

DESPITE the success of qualitative features of the Heitler-London and Heisenberg methods in understanding important concepts in the theory of ferro- and antiferromagnetism, the mathematical basis of the methods has not been established rigorously. Instead, they are rather regarded as a model which is not obtainable from the first principles of quantum mechanics. According to Slater,<sup>1</sup> the methods are, by their nature, incapable of being rigorously applied to many-electron systems for the following two reasons: First, use of nonorthogonal orbitals gives divergence in both the energy matrix  $\int \Phi^* \mathcal{H} \Phi d\tau$  and the overlap  $\int \Phi^* \Phi d\tau$  when the number  $N$  of electrons increases infinitely, so that the energy value  $E$  obtained correctly by

$$E = \frac{\int \Phi^* \mathcal{H} \Phi d\tau}{\int \Phi^* \Phi d\tau} \quad (1)$$

can be quite different from one calculated by neglecting the higher permutations. This implies that the usual derivation of the effective spin Hamiltonian,

$$\mathcal{H} = -2 \sum_{i>j} J_{ij} S_i S_j, \quad (2)$$

is by no means justified.<sup>2</sup> Here  $S_i$  is the spin angular momentum vector of electrons in atom  $i$ , and  $J_{ij}$  is the exchange integral connecting atoms  $i$  and  $j$ . Secondly, use of orthogonal orbitals gives always positive exchange integrals  $J_{ij}$  and therefore cannot explain the existence of antiferromagnetic substances, although, because of the orthogonality of orbitals, it is possible to obtain the spin Hamiltonian (2) correctly.

Recently, Mizuno and Izuyama<sup>3</sup> have given a rigorous mathematical proof that the nonorthogonality catastrophe

does not appear in the manner Inglis<sup>4</sup> pointed out; actually, the numerator and denominator of the energy expression (1) can be factored exactly but in somewhat complicated fashion. By the use of Schwarz's inequality relation, they have shown that these factors can be left out, but instead each term in the energy expression (1) should be bounded. After evaluating the boundary conditions, they found that the exact values cannot be so different from values obtained by neglecting overlap integrals as long as overlaps between nearest neighbor atoms are not so large.

This result may give some justification to recent calculations<sup>5</sup> on the superexchange interaction, where a small number of electrons, say 4 electrons, is taken out of a crystal and the interaction between them is calculated correctly, while a part of the energy, which is proportional to  $S_i S_j$ , is defined as the exchange interaction  $J_{ij}$ . There it was not assumed *a priori* that the exchange interaction could be described by the Heisenberg spin operator, but the energy expression (1) was reduced to the form (2).

It is still not possible to see, however, that the energy expression can be described in the form (2). In the 4-electron system, two of them are coupled to each other and therefore the system is essentially a two-spin system and the energy could be written as the quadratic form  $S_i S_j$  with respect to spin angular momentum vectors. In many-electron systems, we should expect higher terms as long as nonorthogonal orbitals are used as basis. Then the question arises whether such an expansion really exists in many-electron systems and also whether the series converges quickly so that the form (2) can be a good approximation.

By the use of representation matrices  $U(P)$  of the permutation group,<sup>6</sup> in Sec. II, we shall expand the

\* This work was supported in part by the U. S. Air Force through its European Office under contract with Uppsala University, and in part by the U. S. Atomic Energy Commission.

<sup>1</sup> J. C. Slater, *Revs. Modern Phys.* **25**, 199 (1953).

<sup>2</sup> For the standard derivation of the operator, see P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1947), 3rd ed., Chap. IX; for the application to ferromagnetism, see J. H. Van Vleck, *Theory of Electric and Magnetic Susceptibilities* (Oxford University Press, New York, 1932), Chap. XII, and also J. H. Van Vleck, *Revs. Modern Phys.* **17**, 27 (1945).

<sup>3</sup> Y. Mizuno and T. Izuyama, *Progr. Theoret. Phys. (Kyoto)* **22**, 344 (1959).

<sup>4</sup> D. R. Inglis, *Phys. Rev.* **46**, 135 (1934). The same difficulty had been pointed out earlier by J. C. Slater, *Phys. Rev.* **35**, 509 (1930).

<sup>5</sup> For instance, J. Yamashita and J. Kondo, *Phys. Rev.* **109**, 730 (1959); F. Keffer and T. Oguchi, *ibid.* **115**, 1428 (1959); J. Kondo, *Progr. Theoret. Phys. (Kyoto)* **22**, 41, 819 (1959).

<sup>6</sup> Use of representation matrices of the permutation group in the energy calculation has been introduced by R. Serber, *Phys. Rev.* **45**, 461 (1934); T. Yamanouchi, *Proc. Phys. Math. Soc. Japan* **18**, 623 (1936); **20**, 547 (1938). The method is also reviewed by M. Kotani *et al.*, *Table of Molecular Integrals* (Maruzen Company, Ltd., Tokyo, 1955), Chap. I.

energy expression (1) of  $N$  electrons such that

$$E = \sum_P J_P \tilde{U}(P), \quad (3)$$

where  $P$  is a permutation operator of electronic coordinates and the summation  $\sum_P$  includes all  $N!$  different permutations  $P$ . Then we shall show that the effective spin Hamiltonian  $\mathcal{H}_{\text{eff}}$  is given by

$$\mathcal{H}_{\text{eff}} = \sum_P (-1)^P J_{P^{-1}P^\sigma}, \quad (4)$$

where  $P^\sigma$  permute the spin coordinates. The expression (4) is rigorous under the following assumption: Since we are interested in spin-degenerate states only, all the states have a common electronic configuration and the spatial parts of the wave functions are all common and there will not be any difference from state to state. Only differences in the wave functions occur in the spin parts. The main task in Sec. II is to show that all the coefficients  $J_P$  exist and are determined uniquely even if the number  $N$  of electrons becomes infinite. This means that the expansion (4) is mathematically valid.

Furthermore, in Sec. III, we shall show that the *nonorthogonality catastrophe* does not appear and that values of the coefficients  $J_P$  are bounded and can be evaluated even if we omit the higher permutations. Then we will find that the Heitler-London and Heisenberg methods are reliable even though truncation is inevitable and that the Heisenberg spin Hamiltonian will be suitable as the first-order approximation of the effective Hamiltonian.

## II. EFFECTIVE HAMILTONIAN FOR MANY-SPIN SYSTEM

In the first paragraph of this section, we shall explain the representation of wave functions as used here. From the orthogonality relation of representation matrices, we further derive and prove Lemma 1, which is essential for the expansion of the energy matrix. In the second paragraph, the energy expression will be expanded in terms of representation matrices as it will appear in Eq. (28). Then it will be proved that the expansion is really valid, since the coefficients  $J_P$  exist and are determined uniquely. This eliminates the non-orthogonality catastrophe in its most literal interpretation. It is claimed that, in the Heitler-London and Heisenberg methods, there is no other expansion than the obvious one which is expressed schematically as  $E_0 - \infty + \infty - \infty$ . In the new expansion (28), however, all terms involved are bounded.

In the last paragraph, it will be shown that the energy expansion can be written in the form of an effective Hamiltonian which will appear in Eq. (33). This means that the problem of many-electron systems is converted to that of many spins. Since the first two terms of the effective Hamiltonian are equivalent to the Heisenberg spin Hamiltonian, the Heisenberg model will not be in conflict with the first principles. Furthermore, we shall show that each term in the effective Hamiltonian is

Hermitian and therefore the expansion is physically sound.

### 1. Representation of Wave Functions

We shall consider a crystal which consists of  $N'$  atoms and  $N$  electrons. For the moment, we assume only Coulomb interaction between electrons and nuclei. The Hamiltonian is therefore given by

$$\mathcal{H} = \sum_{i=1}^N (P_i^2/2m) - e^2 \sum_{A=1}^{N'} \sum_{i=1}^N (Z_A/r_{Ai}) + e^2 \sum_{i=2}^N \sum_{j=1}^{i-1} 1/r_{ij}, \quad (5)$$

where  $P_i^2 = -\hbar^2 \nabla_i^2$ ,  $m$  is the mass of an electron,  $r_{Ai}$  is the distance between electron  $i$  and nucleus  $A$  with nuclear charge  $Z_A$ , and  $r_{ij}$  is the distance between the electrons  $i$  and  $j$ .

Since  $\mathcal{H}$  does not involve spin operators, the spin angular momentum operator  $S^2$  and its component  $S_Z$  of  $N$  electrons commute with  $\mathcal{H}$ , so that eigenfunctions of  $\mathcal{H}$  are simultaneously eigenfunctions of  $S^2$  and  $S_Z$  and are classified by the spin angular quantum numbers  $S$  and  $M$ . This implies that eigenfunctions of  $\mathcal{H}$  which pertain to particular values of  $S$  and  $M$  are written as<sup>7</sup>

$$\Phi = \Theta \Psi, \quad \Psi = \frac{1}{\sqrt{(N!)}} \sum_P U(P)^* P \psi, \quad (6)$$

where  $\Theta$  is the row matrix  $(\Theta_1, \Theta_2, \dots, \Theta_f)$ , which consists of  $f_S$  linearly independent and orthonormal eigenfunctions of  $S^2$  and  $S_Z$  with the specified values of  $S$  and  $M$ , and

$$f_S = \binom{N}{\frac{1}{2}N - S} - \binom{N}{\frac{1}{2}N - S - 1}.$$

$U(P)$  is the  $f \times f$  irreducible representation matrix for the permutation  $P$  of electronic coordinates, while  $\psi$  is a function of space coordinates of  $N$  electrons. The summation  $\sum_P$  covers all  $N!$  different permutations  $P$ . Therefore, it is evident that  $\Psi$  is a  $f \times f$  matrix, and  $\Phi$  a row matrix with  $f$  elements. A set of the  $N!$  different permutations forms a group. We denote the group by  $\sigma_N$ .

The representation matrix  $U(P)$  is defined by

$$P^\sigma \Theta = (-1)^P \Theta U(P)^*, \quad (7)$$

where  $P^\sigma$  is the operator, which transforms spin coordinates  $1, 2, \dots, N$  of  $\Theta$  into  $1', 2', \dots, N'$ , while  $P$  changes space coordinates  $1, 2, \dots, N$  of  $\psi$  into  $1', 2', \dots, N'$ . As long as  $f$  functions  $\Theta_i$  are normalized,  $U(P)$  is unitary and given by

$$U(P)^* = (-1)^P \int \Theta^\dagger P^\sigma \Theta d\sigma. \quad (8)$$

<sup>7</sup>The detailed description of the method will be found in Kotani's book (see reference 6).

If  $P\epsilon\sigma_N$  and  $Q\epsilon\sigma_N$ , then we find that

$$R=PQ\epsilon\sigma_N. \quad (9)$$

By Eq. (7), therefore, we obtain that

$$U(P)U(Q)=U(PQ)=U(R). \quad (10)$$

If, for example, the one-particle approximation is applied, the spatial function  $\psi$  may be written as a product of  $N$  atomic orbitals which are not orthogonal to each other:

$$\psi(1,2,\dots,N)=\psi_{11}(1)\psi_{12}(2)\cdots\psi_{A1}(x_A)\psi_{A2}(x_A+1)\cdots\psi_{Ah}(x_A+h_A-1)\cdots\psi_{Mh}(N), \quad (11)$$

where  $\psi_{A1}, \psi_{A2}, \dots, \psi_{Ah}$  are atomic orbitals localized at the atom  $A$ . We shall denote by  $h_A$  the number of electrons localized at the atom  $A$ . Then

$$N=\sum_{A=1}^{N'} h_A.$$

However, we want to make clear that the following arguments are more general and entirely free from the choice of the function  $\psi$ . Even an exact eigenfunction of  $\mathcal{H}$  can be written in the form (6), and the proof concerning the existence of the expansion (3) is correct. But in order to make the expansion physically sensible, we have to make certain assumptions.

(a) We assume that in  $\psi$ , each electron is localized at a particular atom, since the exchange interaction is a quantity characteristic to the Heitler-London scheme. If  $\psi$  is expanded in terms of atomic orbitals, therefore,  $\psi$  will be written such that

$$\psi=\sum_{n_{11}n_{12}\cdots} A(n_{11}n_{12}\cdots n_{A1}n_{A2}\cdots n_{Mh}) \times \psi_{n_{11}}(1)\psi_{n_{12}}(2)\cdots\psi_{n_{A1}}(x_A) \times \psi_{n_{A2}}(x_A+1)\cdots\psi_{n_{Mh}}(N), \quad (12)$$

where  $\psi_{n_{A1}}, \psi_{n_{A2}}, \dots, \psi_{n_{Ah}}$  are atomic orbitals centered at the atom  $A$ . In  $\psi$ , the probability of identifying the electrons  $x_A, x_A+1, \dots, x_A+h-1$  as the atomic orbitals  $\psi_{A1}, \psi_{A2}, \dots, \psi_{Ah}$  will be dominant.<sup>8</sup>

(b) We assume that the essential approximation we have to make is that the space parts  $\psi$  of the wave functions are the same for all the spin degenerate states we are interested in.

(c) Furthermore, we assume, for the moment, that the  $N!$  functions  $P\psi$  obtained by operating  $N!$  different permutations  $P$  are linearly independent. Later in this section, however, we shall remove this restriction and show that the following arguments are applicable even if  $\psi$  has symmetry.

It is easily shown that spatial functions  $\psi$  are common to  $2S+1$  magnetic substates ( $M=S, S-1, \dots, -S$ ) for a state with given  $S$  and that representations for

<sup>8</sup> Existence of such function in an exact eigenfunction of  $\mathcal{H}$  has been discussed by the author in *Revs. Modern Phys.* **32**, 370 (1960).

those  $2S+1$  states are equivalent. Therefore, assumption (b) is rigorous for those cases. If spin-degenerate states belong to different spin quantum numbers  $S$ , however, the representations are no longer equivalent, and the spatial functions  $\psi$  are not usually common. We shall identify the representations by  $\mathfrak{D}_S$  and their matrices by the superior such that  $U^S(P)$ . Assumption (b) will be a good approximation if differences in the correlation energies of spin-degenerate states are small.

As long as  $N!$  functions  $P\psi$  are linearly independent,  $f_S$  functions obtained by Eq. (6) will be linearly independent. Since there are  $(2S+1)$  substates for a state with given  $S$ , the total number of linearly independent functions obtained by Eq. (6) is  $\sum_S(2S+1)f_S$ . It is easily proved that

$$\sum_S(2S+1)f_S=2^N. \quad (13)$$

Since, in an  $N$ -electron system, we find  $2^N$  linearly independent spin functions, Eq. (13) shows that all linearly independent functions for spin-degenerate states are obtained by Eq. (6).

In general, the representations  $\mathfrak{D}_S$  ( $S=0$  or  $\frac{1}{2}, \dots, \frac{1}{2}N$ ) do not exhaust all irreducible representations of  $\sigma_N$ . We shall denote  $\mathfrak{D}_\rho$  when we want to include all the irreducible representations of  $\sigma_N$ .<sup>9</sup> The total number of  $\mathfrak{D}_\rho$  is equal to the number of classes in  $\sigma_N$  and we have the relation

$$\sum_\rho f_\rho^2=N!. \quad (14)$$

Since the representations are irreducible, we have the following orthogonality relations:

$$(f_\rho/N!) \sum_P U_{km}^\rho(P)U_{n\ell}^\tau(P^{-1})=\delta_{\rho\tau}\delta_{k\ell}\delta_{m\ell}. \quad (15)$$

These relations show that the  $N!\times N!$  matrix  $\Gamma$ ,

$$\Gamma_{\rho km, P}=(f_\rho/N!)^{\frac{1}{2}}U_{km}^\rho(P), \quad (16)$$

whose rows are numbered by  $(\rho km)$  and whose columns are numbered by  $P$ , is unitary.

Now we find the following lemma.

*Lemma 1.* If we have the relation

$$\sum_P a_P U^\rho(P)=0, \quad (17)$$

which is valid for all the irreducible representations  $\mathfrak{D}_\rho$  of  $\sigma_N$ , then all the coefficients  $a_P$  must be identically zero.

By the use of the matrix  $\Gamma$ , defined by Eq. (16), Eq. (17) is written as

$$\Gamma A=0, \quad (18)$$

<sup>9</sup> The rest of the representations  $\mathfrak{D}_\rho$  of  $\sigma_N$ , which are not included in  $\mathfrak{D}_S$ , cannot find basic functions in spin space. This is because we have only a limited number  $2^N$  of independent spin functions at our disposal and in general this number of independent functions is insufficient to form bases for all the irreducible representations of  $\sigma_N$ . If we use a representation  $\mathfrak{D}_\rho$  which is different from  $\mathfrak{D}_S$ , the wave functions defined by Eq. (6) vanish, because the spin functions which satisfy Eq. (7) vanish. Formally, however, we can apply the following way of writing to all the representations  $\mathfrak{D}_\rho$  of  $\sigma_N$  and we can omit the part which is obtained from the representations  $\mathfrak{D}_\rho$  not included in  $\mathfrak{D}_S$ .

where  $A$  is a column matrix:

$$A = \begin{bmatrix} a_{P_1} \\ a_{P_2} \\ \vdots \end{bmatrix}.$$

By multiplying Eq. (18) by  $\Gamma^\dagger$ , we find  $A=0$  because  $\Gamma$  is unitary.

## 2. Expansion of Energy Matrix

We can assume that the wave functions  $\Phi$  given by Eq. (6) are orthogonal and the  $f \times f$  overlap matrix  $C = \int \Phi^\dagger \Phi d\tau$  is diagonal. If this is not true, then there exists a unitary matrix  $V$ , with which  $C$  can be brought into diagonal form by the similarity transformation such that

$$D = V^\dagger C V = \int (\Phi V)^\dagger (\Phi V) d\tau,$$

since  $C^\dagger = C$ . This means that, instead of  $\Phi$ , we can use the orthogonal functions  $\Phi V$ , which are given by

$$\Phi V = (\Theta V)(V^\dagger \Psi V),$$

$$V^\dagger \Psi V = \frac{1}{\sqrt{(N!)}} \sum_P V^\dagger U(P)^* V P \Psi.$$

This implies that, instead of  $\Theta$ , we should use  $\Theta' = \Theta V$  as the basis. Therefore, we can write the overlap matrix as  $D = \int \Phi'^\dagger \Phi' d\tau$ , where  $D$  is diagonal.

Now we prove that all the diagonal elements are nonvanishing; namely, there exists an  $\epsilon$  such that

$$D_{ii} > \epsilon > 0, \quad \text{for } i=1, 2, \dots, f. \quad (19)$$

If, for any small number  $\epsilon$ , we find that

$$D_{ii} = \int \Phi_i^* \Phi_i d\tau = \int |\Phi_i|^2 d\tau < \epsilon,$$

then  $\Phi_i$  must vanish. This means that  $\Psi_{ki} = 0$ , for  $k=1, 2, \dots, f$  since all the spin functions  $\Theta_i$  in Eq. (6) are linearly independent. Since  $N!$   $P\Psi$ 's are linearly independent and Eq. (6) is written as

$$0 = \Psi_{ki} = \frac{1}{\sqrt{(N!)}} \sum_P U_{ki}(P)^* P \Psi, \quad (20)$$

we find that  $U_{ki}(P)^* = 0$  for all  $k$  and  $P$ . This would imply that the representation is neither unitary nor irreducible, which is contradictory to the original method described in Eq. (6).

Since Eq. (19) is valid,  $D^{-1}$  exists and its elements are bounded such that

$$0 < D_{ii}^{-1} < \epsilon^{-1} < \infty, \quad \text{for } i=1, 2, \dots, f.$$

Here we can define  $D^{-\frac{1}{2}}$  as the diagonal matrix whose elements  $(D^{-\frac{1}{2}})_{ii}$  are given by  $D_{ii}^{-\frac{1}{2}}$ . Then the functions defined by

$$\Phi^N = \Phi D^{-\frac{1}{2}},$$

are orthonormal, and the energy matrix is given by

$$\begin{aligned} E &= \int (\Phi D^{-\frac{1}{2}})^\dagger \mathcal{H} (\Phi D^{-\frac{1}{2}}) d\tau \\ &= E_0 + D^{-\frac{1}{2}} H D^{-\frac{1}{2}}, \end{aligned} \quad (21)$$

where

$$H = \int \Phi^\dagger (\mathcal{H} - E_0) \Phi d\tau,$$

and

$$E_0 = \int \Psi^* \mathcal{H} \Psi d\tau.$$

The nonorthogonality catastrophe tells us that, since overlap integrals  $\int P \Psi^* \Psi d\tau$  are generally not zero, both  $H$  and  $D$  diverge and the evaluation of  $E$  in Eq. (21) can be illustrated by the calculation of  $0 \times \infty \times 0$ , when the number  $N$  becomes infinite. On the other hand, if we neglect higher permutations,  $H$  and  $D$  behave as if they are finite numbers  $H_0$  and  $D_0$ . It is entirely unjustified to evaluate  $E$  by  $D_0^{-\frac{1}{2}} H_0 D_0^{-\frac{1}{2}}$ , since the main contribution  $H - H_0 = \infty$  is neglected there.

From Eqs. (6), (9), and (10), it is easily shown that

$$H = \sum_P H_P \tilde{U}(P), \quad H_P = \int P \Psi^* (\mathcal{H} - E_0) \Psi d\tau, \quad (22)$$

and

$$D = \sum_P S_P \tilde{U}(P), \quad S_P = \int P \Psi^* \Psi d\tau, \quad (23)$$

where the summation  $\sum_P$  is taken over all  $N!$  permutations  $P$ . Because of the properties of a group and its representation shown by Eqs. (9) and (10), it is possible to expand  $D^{-1}$  formally in the following fashion.

$$D^{-1} = [\sum_P S_P \tilde{U}(P)]^{-1} = \sum_P T_P \tilde{U}(P). \quad (24)$$

We shall show that the expansion exists and the coefficients  $T_P$  are uniquely determined.

From Eqs. (23) and (24), it is found that

$$1 = \tilde{D}^{-1} \tilde{D} = \sum_P \sum_Q T_Q S_Q^{-1} U(P).$$

By applying Lemma 1,<sup>10</sup> we find

$$\sum_Q T_Q S_Q^{-1} P = \delta_{EP}, \quad \text{for } P=1, 2, \dots, N!,$$

since the representation matrix  $U(E)$  of the identity element  $E$  is unit. The expression is written as

$$T S = \delta, \quad (25)$$

where  $S$  is a  $N! \times N!$  matrix

$$S = \int \Phi^\dagger \Phi d\tau,$$

<sup>10</sup> Since the equation written above is valid for all the irreducible representations  $\mathfrak{D}_\rho$  of  $\sigma_N$ , we are able to apply Lemma 1. See the note in reference 9. The matrices  $T, S$ , etc., whose elements are given by  $T_P, S_Q^{-1}$ , etc., pertain to an  $N!$ -dimensional space, because the number of independent permutations  $P$  or  $Q$  is  $N!$ .

and  $T$ ,  $\delta$  and  $\phi$  are the following row matrices:

$$\begin{aligned} T &= (T_{P_1}, T_{P_2}, \dots, T_{P_N}), \\ \delta &= (1, 0, \dots, 0), \\ \phi &= (P_1\psi, P_2\psi, \dots, P_N\psi). \end{aligned}$$

Since  $S^\dagger = S$ , it is possible to diagonalize  $S$  by a similarity transformation such that

$$\Delta = V^\dagger S V = \int (\phi V)^\dagger (\phi V) d\tau.$$

Since  $(\phi V)_i$  is a linear combination of  $N!$   $P\psi$  and is similar to  $\Psi_{ki}$  given by Eq. (20), there exists an  $\epsilon$  such that

$$\Delta_{ii} > \epsilon > 0, \quad \text{for } i=1, 2, \dots, N!$$

Equation (25) is therefore written as

$$T' = T V = \delta V \Delta^{-1},$$

and  $\sum |T_i|^2$  is calculated as follows

$$\begin{aligned} \sum_i |T_i|^2 &= T T^\dagger = T' T'^\dagger = \delta V \Delta^{-2} V^\dagger \delta^\dagger \\ &= \sum_i |V_{1i}|^2 / \Delta_{ii}^2 \leq 1/\epsilon^2 < \infty. \end{aligned}$$

That is, the coefficients  $T_P$  exist and are always uniquely determined by Eq. (25) because  $|S| \neq 0$ .

Since the expansion (24) of  $D^{-1}$  exists,  $D^{-\frac{1}{2}}$  is also written as

$$D^{-\frac{1}{2}} = \sum_P t_P \tilde{U}(P), \tag{26}$$

where the coefficients  $t_P$  are determined uniquely.

This is proved as follows. By using a technique similar to Eq. (25), it is found that  $t_P$  should satisfy the relation

$$t = T, \tag{27}$$

where  $t$  and  $T$  are the  $N! \times N!$  matrices whose  $QP$  elements are given by  $t_{Q^{-1}P}$  and  $T_{Q^{-1}P}$ , respectively. Since  $D^{-\frac{1}{2}}$  is real and diagonal, we find that  $t^\dagger = t$ . Then  $t$  can be brought into a diagonal form  $t'$  by a similarity transformation such that

$$t' = V^\dagger t V.$$

Under the same transformation,  $T$  is also diagonalized and Eq. (27) is written as

$$t' t' = T',$$

where  $T' = V^\dagger T V$  and  $t_{ii}' = \sqrt{T_{ii}'}$  (or  $-\sqrt{T_{ii}'}$ ). The matrix  $t$  obtained by  $t = V t' V^\dagger$  satisfies Eq. (27).<sup>11</sup>

<sup>11</sup> The uniqueness of  $t$  can be proved as follows. We consider a different similarity transformation. Then we obtain a different diagonal matrix  $T'' = W^\dagger T W$ . Since  $T = V T' V^\dagger = W T'' W^\dagger$ , we find that

$$U T' = T'' U, \tag{a}$$

where  $U = W^\dagger V$  is unitary. Let us consider the  $ij$  element of the above Eq. (a). Then  $U_{ij} T_{ij}' = T_{ij}'' U_{ij}$ , since  $T'$  and  $T''$  are diagonal. This indicates that  $T_{ij}' = T_{ij}''$  or  $U_{ij} = 0$ . In either case, we have the relation

$$U_{ij} t_{ij}' = t_{ij}'' U_{ij}, \tag{b}$$

where  $t_{ij}' = \sqrt{T_{ij}'}$  and  $t_{ij}'' = \sqrt{T_{ij}''}$ . Therefore  $V t' V^\dagger = W t'' W^\dagger$ , and regardless of the unitary matrix used, we obtain the same  $t$ .

If we put  $t_{ii}' = -\sqrt{T_{ii}'}$  for all  $i$ , then the sign of  $D^{-\frac{1}{2}}$  will change. Although  $D^{-\frac{1}{2}}$  is positive definite, we do not know the sign of

Furthermore, we find that

$$\sum_i |t_i|^2 = T_E < 1/\epsilon < \infty;$$

consequently, all the coefficients  $t_P$  exist and are obtainable.

By the use of the expansions (22) and (26), the energy expression (21) can be expanded formally as follows:

$$E - E_0 = \sum_P J_P \tilde{U}(P), \tag{28}$$

where

$$J_P = \sum_Q \sum_R t_Q H_{Q^{-1}R} t_{R^{-1}P}. \tag{29}$$

Now we shall prove that, as long as the values of all the integrals  $H_P$  are bounded, all the coefficients  $J_P$  also exist and are determined uniquely. In Appendix A, we shall show that all  $H_P$  are indeed bounded.

By adopting the matrix notation similar to the one used in Eq. (27), Eq. (29) is written as

$$J = t H t,$$

where the  $QP$  elements of the  $N! \times N!$  matrices  $J$  and  $H$  are given by  $J_{Q^{-1}P}$  and  $H_{Q^{-1}P}$ , respectively. By the similarity transformation, which brings  $t$  into the diagonal form  $t' = V^\dagger t V$ , we find that

$$J' = V^\dagger J V = t' H' t',$$

where all the elements  $H_{ij}'$  of  $H' = V^\dagger H V$  are bounded after the unitary transformation by  $V$ . As the  $ij$  element  $J_{ij}'$  is written as

$$J_{ij}' = t_{ii}' H_{ij} t_{jj}',$$

all the coefficients  $J_{ij}'$  and also  $J_{ij}$  must exist. Since, apart from the factor 1 or  $-1$ , the  $N!$  coefficients  $t_P$  are uniquely determined,  $N! J_P$  are also determined uniquely by Eq. (29). As  $J_P$  is quadratic with respect to  $t_P$ , the undetermined signs of  $t_P$  does not affect the signs of  $J_P$ .

So far, we have assumed that the  $N!$  functions  $P\psi$  are linearly independent. In Appendix B, we shall show that even if  $N! P\psi$  are not linearly independent, the arguments remain correct. Therefore, the expansion (28) is always valid regardless of the symmetry of  $\psi$ . This proves the existence of an energy expansion in which all terms are bounded and determined uniquely. Consequently, the nonorthogonality catastrophe does not appear in its most strict sense, and the Heitler-London and Heisenberg approaches can have possibilities of being rigorously applied to many-electron systems, as we will shown in detail in Sec. III.

### 3. Effective Spin Hamiltonian

Now we shall convert the energy expansion into an effective spin Hamiltonian. First let us consider the

of  $\sum_P t_P \tilde{U}(P)$  without calculating its value, since it is not possible to see simply from Eq. (27) which sign should be assigned. Since, in physical quantities,  $t$  appears always in the quadratic form, it is not necessary to determine the signs. From Eq. (b), it is easily seen that the signs of  $t$  must be all plus or all minus.

energy expansion (28). Here the representation matrices  $\tilde{U}(P)$  are given by Eq. (8): namely,

$$\tilde{U}(P^{-1}) = (-1)^P \int \Theta^\dagger P^\sigma \Theta d\sigma. \quad (8')$$

Since we assumed that the spin functions  $\Theta_i$  used as the basis are simultaneous eigenfunctions of  $S^2$  and  $S_z$ , the representation is irreducible. Furthermore, the coefficients  $J_P$  are independent of  $\Theta$  and are common for all possible spin states classified by spin quantum numbers  $S$  and  $M$ , as long as a common  $\psi$  is used as the space part of the wave functions. Therefore, the energy expressions (28) for all the spin states are written in one expression as follows:

$$E = \sum_P J_P \tilde{W}(P), \quad (30)$$

where  $\tilde{W}(P)$  is a supermatrix in which all the irreducible representations  $\tilde{U}^{SM}(P)$  for all possible values of  $S$  and  $M$  appear in the form

$$\begin{bmatrix} \tilde{U}^{\frac{1}{2}N, \frac{1}{2}N}(P) & & & \\ & \tilde{U}^{\frac{1}{2}N, \frac{1}{2}N-1}(P) & & \\ & & \ddots & \\ & & & \tilde{U}^{\frac{1}{2}N-1, \frac{1}{2}N-1}(P) & \\ & & & & \ddots \end{bmatrix}.$$

Because of the relation (13),  $\tilde{W}(P)$  is a  $2^N \times 2^N$  matrix and given by

$$\tilde{W}(P^{-1}) = (-1)^P \int \Theta^\dagger P^\sigma \Theta d\sigma, \quad (31)$$

where  $\Theta$  is the row matrix  $(\Theta_1, \Theta_2, \dots, \Theta_F)$  consisting of  $F = 2^N$  linearly independent spin functions. Since the  $2^N$  functions form a complete set in  $N$ -spin systems,<sup>12</sup> the evaluation of the energy of the Hamiltonian is equivalent to the eigenvalue problem of the operator

$$\mathcal{H}_{\text{eff}} = \sum_P (-1)^P J_P P^{-1} P^\sigma. \quad (4)$$

By a unitary transformation, Eq. (30) is written as

$$E' = V^\dagger E V = \sum_P J_P \tilde{W}'(P),$$

where

$$\tilde{W}'(P) = V^\dagger \tilde{W}(P) V = (-1)^P \int (\Theta V)^\dagger P^\sigma (\Theta V) d\sigma.$$

This shows that the effective Hamiltonian (4) is valid not only to the eigenstates of  $S^2$  and  $S_z$ , but also to any spin function, since  $\Theta_1, \Theta_2, \dots, \Theta_F$  are complete and  $V$  is arbitrary. Even if we consider a spin-dependent interaction  $\mathcal{H}'$ , therefore, we can use the effective Hamiltonian, and the eigenvalue problem of the total Hamiltonian,

$$\mathcal{H}_{\text{total}} = \mathcal{H} + \mathcal{H}',$$

<sup>12</sup> As we shall show in Appendix B, a part of the energy matrix  $E$  given by Eq. (28) will vanish if the  $N!$   $P$ 's are not linearly independent, and therefore the corresponding part of spin functions in  $\Theta$  will disappear in Eq. (30). Consequently, we may not obtain  $2^N$  independent solutions from the Hamiltonian (4).

is equivalent to  $\mathcal{H}_{\text{eff}} + \mathcal{H}'$ , although eigenfunctions of  $\mathcal{H}_{\text{total}}$  are usually not simultaneous eigenfunctions of  $S^2$  and  $S_z$ .

The  $N!$  permutations  $P^\sigma$  can be expressed by products of permutations of pairs such that

$$P_M^\sigma = (i_1 j_1)(i_2 j_2) \cdots (i_M j_M), \quad (32)$$

where  $M$  indicates the number of electrons involved in the permutation. According to Dirac,<sup>13</sup> each of the pair permutations  $(ij)$  is equivalent to the spin operator

$$(ij) = \frac{1}{2}(1 + 4s_i s_j).$$

Therefore,  $P_M^\sigma$  is described by a product of  $\frac{1}{2}(1 + 4s_i s_j)$ . The expression (32) is not uniquely given, because of the relations

$$(ij)^2 = 1, \quad (ij)(jk) = (ik)(ij).$$

By the use of the commutation relations of the spin operator, however, it is easily shown that all different representations (32) of a permutation  $P_M^\sigma$  are equivalent in terms of spin operators. Thus each of  $N!$  permutation operators  $P_M^\sigma$  is written as a function  $\Theta_M$  of operators  $s_{i_1} s_{i_2} \cdots s_{i_M}$ .

$$P_M^\sigma = \Theta_M(s_{i_1}, s_{i_2}, \dots, s_{i_M}; P_M).$$

The expression is unique up to the commutation relations of spin operators.

The effective Hamiltonian (4) is therefore written as

$$\begin{aligned} \mathcal{H}_{\text{eff}} = & J_E - \frac{1}{2} \sum_{i>j} J_{ij} (1 + 4s_i s_j) \\ & + \sum_{M \geq 3} \sum_{P_M} (-1)^P J_P \Theta_M(s_{i_1}, s_{i_2}, \dots, s_{i_M}; P_M). \end{aligned} \quad (33)$$

If we can prove that, in ordinary crystals, coefficients  $J_{ij}$  for permutations of single pairs are dominant as compared with  $J_P$  for  $M \geq 3$ , the energy levels are described by

$$\mathcal{H}_{\text{eff}}^{(1)} = J_E - \frac{1}{2} \sum_{i>j} J_{ij} (1 + 4s_i s_j),$$

and the higher terms may be neglected. The approximate Hamiltonian is equivalent to the Heisenberg spin Hamiltonian (2).

Here,  $s_i$  is the spin operator of electron  $i$ . If each atom carries more than one electron and the coefficients  $J_{A_i B_j}$  for the permutations between electrons  $A_1, A_2, \dots, A_h$  in atom  $A$  and  $B_1, B_2, \dots, B_h$  in atom  $B$  are equal to each other, then the approximate Hamiltonian is

$$\mathcal{H}_{\text{eff}}^{(1)} = -2 \sum_{A>B} J_{AB} S_A S_B.$$

provided the intra-atomic exchanges are negligible. Here  $J_{AB}$  is the exchange interaction connecting atoms

<sup>13</sup> P. A. M. Dirac, Proc. Roy. Soc. (London) **A123**, 714 (1929). Also, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1947), 3rd ed., Chap. IX, p. 222.

$A$  and  $B$  and  $S_A$  is the spin angular momentum vector of atom  $A$ .

Finally, we shall show that the exchange interaction energy  $J_{ij}$  thus defined is real and a physically sensible quantity to discuss. By taking the Hermitian conjugate of Eq. (28), we obtain that

$$(E - E_0)^\dagger = \sum_P J_{P^{-1}} \tilde{U}(P).$$

Since  $(E - E_0)^\dagger = (E - E_0)$ , use of Lemma 1 will give us that  $J_{P^{-1}} = J_P$ . For permutations  $P = (ij)$  of single pairs,  $P^{-1} = P$  and therefore  $J_{ij}$  is always real. Furthermore,  $J_P \tilde{U}(P) + J_{P^{-1}} \tilde{U}(P^{-1})$  is Hermitian and the diagonal elements are real, since the  $ii$  element is written as

$$J_P \tilde{U}_{ii}(P) + J_{P^{-1}} \tilde{U}_{ii}(P)^*,$$

for all  $N!$  permutations  $P$ . This means that, even if we take the summation  $\sum_M$  in Eq. (33) up to a limited number, say  $M_0$ , the truncated Hamiltonian is still Hermitian and the energy levels obtained are always real, and, therefore, the expansion is physically sensible.

### III. CALCULATION OF EXCHANGE INTERACTION ENERGY

In Sec. II, we have shown that the effective Hamiltonian for spin-degenerate states is written as Eq. (4) or Eq. (33) where the coefficients  $J_P$  represent values of the exchange interaction and can be obtained in principle by solving the  $N!$  linear equations (25). In practice, however, this is not possible, since the number  $N$  of electrons in a crystal is effectively infinite. Here we shall demonstrate a possibility of evaluation  $J_P$  without solving Eq. (25) and show that all the coefficients  $J_P$  are bounded and exchange interactions between nearest-neighbor atoms are indeed dominant.

This again eliminates the nonorthogonality catastrophe in its orthodox interpretation. The argument used is the following. Even if an energy expansion is used which leads to convergency such as the one given by Eq. (28) and even if overlap integrals between atoms are relatively small, the convergency will be extremely slow. Third, fourth, and even higher terms will be dominant as compared with the second term (the Heisenberg exchange term). Therefore, the Heisenberg model can never be a useful one. Although many calculations indicate that the Heisenberg term is dominant, such calculations have no meaning according to exponents of the orthodox view, since effects of overlaps from distant electrons are neglected. An exact treatment supposedly might change the whole situation.

In contrast to this criticism, the results given in this section and written clearly in Eqs. (77) and (86) and the subsequent text, show that each term  $J_P$  on the energy expansion cannot be far from the values obtained by omitting the higher permutations. As long as overlap integrals between atoms are small, therefore, the Heisenberg model will be a good approximation.

In order to prove the results, we first have to separate a few electrons in the energy matrix and treat that part explicitly. Then the energy matrix will be expressed as a sum of products of Coulomb or exchange integrals and overlap matrices. This result is expressed in Eqs. (37) and (38) and our problem of expanding the energy matrix is reduced to the expansion of overlap matrices. The aim of the latter half of the first paragraph is to prove Lemma 4. We have derived an inequality relation between the length of expansion coefficients  $d_p$  of various overlap matrices having the same number of electrons. This will be presented in Lemma 4 or 4'. The Lemma is actually a generalized form of Schwarz's inequality and the essential means for comparing and bounding expansion coefficients of the overlap and energy matrices. To prove the lemma, however, we have to pass through Lemmas 2 and 3.

After Lemma 4 is verified, it is no longer difficult to evaluate upper and lower bounds of  $\sum |d_p|^2$ , as is carried out in the second paragraph. When the overlap integrals are not exceedingly large ( $\Delta < 1$ ), the results will be explicitly given in Eqs. (54)–(57) and (60)–(63). Even if  $\Delta \geq 1$ , the bounds can, of course, be given by considering overlap between nearest neighbor atoms explicitly. Since such extension is trivial, we shall omit the mathematics in this paper. In any case, the results together with the energy expansions (64) and (65) show that the effects of overlap from distant atoms and higher permutations are not essential. On the other hand, if one is not satisfied with the accuracy of numerical results for some particular problem, one can easily improve the accuracy by taking more permutations and it is always possible to evaluate the upper limit of the error due to the truncation of the energy expansion.

Finally, in the last paragraph, we derive a complete expansion of the exchange interactions  $J_p$  in terms of  $d_p$ ; this will appear in Eq. (76). By using the results in the second paragraph, we can at last obtain the promised relations (77) and (86), which will show that Heisenberg term is indeed dominant. Since the complete energy expansion (33) is available and each of the coefficients  $J_p$  can be evaluated exactly the same way as the Heisenberg exchange interaction  $J_p$ , we are able to increase the accuracy whenever necessary. In this point, there is a sharp difference from and great advantage over the existing Heisenberg theory.

#### 1. Separation of Permutations

First in this paragraph, we shall separate interacting electrons from others. For simplicity, we shall assume that each atom carries a valence electron and the spatial function  $\psi$  is written as

$$\psi(1, 2 \cdots N) = \psi_1(1) \psi_2(2) \cdots \psi_N(N),$$

where the inferior  $i$  of  $\psi_i$  indicates that the orbital  $\psi_i$  is centered at atom  $I$ . Here the one-particle approxima-

tion is applied because it simplifies notations, but the more general case (12) can be treated in exactly the same manner.

The Hamiltonian (5) is written as

$$\mathcal{H} = \sum_i \mathcal{H}_i + \sum_{i>j} \mathcal{H}_{ij},$$

where

$$\mathcal{H}_i = \mathcal{H}_i^0 - \langle i | \mathcal{H}_i^0 | i \rangle, \quad \mathcal{H}_i^0 = (P_i^2/2m) - e^2/r_{Ii},$$

and

$$\mathcal{H}_{ij} = \mathcal{H}_{ij}^0 - \langle ij | \mathcal{H}_{ij}^0 | ij \rangle, \quad \mathcal{H}_{ij}^0 = -e^2/r_{Ji} - e^2/r_{Ij} + e^2/r_{ij}.$$

Then the energy matrix  $H$ , given by Eq. (22), is divided as follows:

$$H = \sum_i H_i + \sum_{i>j} H_{ij},$$

where

$$H_i = \sum_P (H_i)_P \tilde{U}(P), \\ H_{ij} = \sum_P (H_{ij})_P \tilde{U}(P),$$

and

$$(H_{ij})_P = \int P \psi^* \mathcal{H}_{ij} \psi d\tau.$$

Consequently, the energy expression is written as

$$E = \sum_i E_i + \sum_{i>j} E_{ij}, \quad (34)$$

where  $E_i = D^{-\frac{1}{2}} H_i D^{-\frac{1}{2}}$  and  $E_{ij} = D^{-\frac{1}{2}} H_{ij} D^{-\frac{1}{2}}$ .

We shall denote  $(N-1)!$  permutations of electrons  $1, 2, \dots, h-1; h+1, \dots, N$  by  $Q[h]$ . All the  $N!$  permutations are given by

$$P = Q[h] \cdot P(k \leftarrow h), \quad k = 1, 2, \dots, N,$$

where  $P(k \leftarrow h) = (hk)$ . The wave function  $\Phi$ , the overlap matrix  $D$ , and  $E_h$  are written as

$$\Phi = \frac{1}{\sqrt{N}} \sum_k \psi_k(h) \cdot \Phi[k(h)] \quad (35)$$

$$D = \sum_k \sum_{Q[h]} S_{Q[h]P} \tilde{U}(Q[h] \cdot P(k \leftarrow h)) \\ = \sum_k \langle k | h \rangle S[k(h) | h(h)], \quad (36)$$

and

$$H_h = \sum_k \sum_{Q[h]} (H_h)_{Q[h]P} \tilde{U}(Q[h] \cdot P(k \leftarrow h)) \\ = \sum_k \langle k | \mathcal{H}_h | h \rangle S[k(h) | h(h)],$$

where

$$\psi_k(h) \cdot \Phi[k(h)] = [(N-1)!]^{-\frac{1}{2}} \sum_{Q[h]} \Theta \\ \times U(Q[h] \cdot P(k \leftarrow h))^* \cdot Q[h] \cdot P(k \leftarrow h) \psi,$$

and

$$S[k(h) | h(h)] = \int \Phi[k(h)]^\dagger \Phi[h(h)] d\tau.$$

In Eq. (35), all the possible permutations  $(kh)$  of electron  $h$  are considered separately and therefore  $\Phi[k(h)]$  is a function of the space coordinates of electrons  $1, 2, \dots, h-1; h+1, \dots, N$  made up of orbitals  $\psi_1, \psi_2, \dots, \psi_{k-1}, \psi_{k+1}, \dots, \psi_N$ .<sup>14</sup> The energy matrix  $E_h$  is then given by

$$E_h = \sum_k \langle k | \mathcal{H}_h | h \rangle D^{-\frac{1}{2}} S[k(h) | h(h)] D^{-\frac{1}{2}}. \quad (37)$$

Similarly,  $E_{hl}$  is written as

$$E_{hl} = \sum_k \sum_{m'} \langle km | \mathcal{H}_{hl} | hl \rangle D^{-\frac{1}{2}} \\ \times S[k(h)m(l) | h(h)l(l)] D^{-\frac{1}{2}}, \quad (38)$$

where

$$S[k(h)m(l) | h(h)l(l)] = \int \Phi[k(h)m(l)]^\dagger \Phi[h(h)l(l)] d\tau. \\ \psi_k(h) \psi_m(l) \Phi[k(h)m(l)] \\ = [(N-2)!]^{-\frac{1}{2}} \sum_{Q[hl]} \Theta \cdot U(Q[hl] \cdot P(km \leftarrow hl))^* \\ \times Q[hl] P(km \leftarrow hl) \psi,$$

and  $Q[hl]$  are  $(N-2)!$  permutations of electrons  $1, 2, \dots, h-1; h+1, \dots, l-1; l+1, \dots, N$ , while  $P(km \leftarrow hl)$  is a permutation taking  $h$  to  $k$  and  $l$  to  $m$ . The functions  $\Phi[k(h)m(l)]$  are related to  $\Phi[k(h)]$  as follows:

$$\Phi[k(h)] = (N-1)^{-\frac{1}{2}} \sum_m \psi_m(l) \Phi[k(h)m(l)].$$

More generally, we define  $\Phi[k_1(h_1)k_2(h_2)\dots k_n(h_n)]$  by

$$\psi_{k_1}(h_1) \psi_{k_2}(h_2) \dots \psi_{k_n}(h_n) \Phi[k_1(h_1)k_2(h_2)\dots k_n(h_n)] \\ = [(N-n)!]^{-\frac{1}{2}} \sum_Q \Theta \cdot U(Q[h_1 h_2 \dots h_n] \\ \cdot P(k_1 k_2 \dots k_n \leftarrow h_1 h_2 \dots h_n))^* Q[h_1 h_2 \dots h_n] \\ \cdot P(k_1 k_2 \dots k_n \leftarrow h_1 h_2 \dots h_n) \psi, \quad (39)$$

where, by  $Q[h_1 h_2 \dots h_n]$  we denote all possible permutations in which  $n$  electrons  $h_1 h_2 \dots h_n$  are not involved, while  $P(k_1 k_2 \dots k_n \leftarrow h_1 h_2 \dots h_n)$  interchanges space coordinates  $h_1 h_2 \dots h_n$  with  $k_1 k_2 \dots k_n$ . Similarly, we define  $S$  by

$$S[k_1(h_1)k_2(h_2)\dots k_n(h_n) | k_1'(h_1')k_2'(h_2')\dots k_n'(h_n')] \\ = \int \Phi[k_1(h_1)k_2(h_2)\dots k_n(h_n)]^\dagger \\ \times \Phi[k_1'(h_1')k_2'(h_2')\dots k_n'(h_n')] d\tau. \quad (40)$$

Now we like to prove the most important relation: Schwarz's inequality applied to the lengths  $\sum |d_p|^2$  of expansion coefficients  $d_p$  of overlap matrices. We shall derive and prove the relation in Lemma 4, but before proving it, we need to verify the following two lemmas.

<sup>14</sup> The spin functions  $\Theta$  in  $\Phi[k(h)]$  include the spin coordinate of electron  $h$ .



*Lemma 2.* The overlap matrix

$$S[k_1(h_1)k_2(h_2)\cdots k_n(h_n) | k_1(h_1)k_2(h_2)\cdots k_n(h_n)]$$

is independent of  $h_1h_2\cdots h_n$  and therefore

$$\begin{aligned} S[k_1(h_1)k_2(h_2)\cdots k_n(h_n) | k_1(h_1)k_2(h_2)\cdots k_n(h_n)] \\ = S[k_1(k_1)k_2(k_2)\cdots k_n(k_n) | \\ \times k_1(k_1)k_2(k_2)\cdots k_n(k_n)]. \quad (41) \end{aligned}$$

From the definition (39) of  $\Phi$ , it is found that

$$\begin{aligned} \int \Phi[k_1(h_1)k_2(h_2)\cdots k_n(h_n)]^\dagger \\ \times \Phi[k_1(h_1)k_2(h_2)\cdots k_n(h_n)] d\tau \\ = \sum_Q S_{P^{-1}QP} \tilde{U}(P^{-1}QP), \quad (42) \end{aligned}$$

where the summation  $\sum_Q$  is taken over all the  $(N-n)!$  permutations  $Q[h_1h_2\cdots h_n]$ . To simplify the notations, we will abbreviate  $Q[h_1h_2\cdots h_n]$ ,  $Q[k_1k_2\cdots k_n]$ , and  $P(k_1k_2\cdots k_n \leftarrow h_1h_2\cdots h_n)$  by  $Q$ ,  $Q'$ , and  $P$ , respectively. Then it is easily recognized that  $P^{-1}QP$  are permutations of electrons not including  $k_1k_2\cdots k_n$ , and equivalent to  $Q'$ . There are one-to-one correspondences between  $Q$  and  $Q'$ , and if all the  $(N-n)!$  different permutations  $Q$  are taken into account, the  $(N-n)!$  different permutations  $Q'$  are generated by  $Q' = P^{-1}QP$ . Therefore, the summation (42) is equivalent to

$$\begin{aligned} \sum_{Q'} S_{Q'} \tilde{U}(Q') = \int \Phi[k_1(k_1)k_2(k_2)\cdots k_n(k_n)]^\dagger \\ \times \Phi[k_1(k_1)k_2(k_2)\cdots k_n(k_n)] d\tau. \end{aligned}$$

This proves the lemma.

$$\begin{aligned} D^{-\frac{1}{2}} S[k_1(h_1)k_2(h_2)\cdots k_n(h_n) | k_1'(h_1')k_2'(h_2')\cdots k_n'(h_n')] D^{-\frac{1}{2}} \\ = \sum_P d[k_1(h_1)k_2(h_2)\cdots k_n(h_n) | k_1'(h_1')k_2'(h_2')\cdots k_n'(h_n')]_P \tilde{U}(P). \quad (47) \end{aligned}$$

Then it is found that

*Lemma 4.* The following inequality relation holds between the coefficients  $d_P$ :

$$\begin{aligned} \sum_P |d[k_1(h_1)k_2(h_2)\cdots k_n(h_n) | k_1'(h_1')k_2'(h_2')\cdots k_n'(h_n')]_P|^2 \\ \leq \{ \sum_P |d[k_1(k_1)k_2(k_2)\cdots k_n(k_n) | k_1(k_1)k_2(k_2)\cdots k_n(k_n)]_P|^2 \\ \times \sum_P |d[k_1'(k_1')k_2'(k_2')\cdots k_n'(k_n')]_P|^2 \}^{\frac{1}{2}}. \quad (48) \end{aligned}$$

Since the expression  $D^{-\frac{1}{2}} S[k_1(h_1)k_2(h_2)\cdots k_n(h_n) | k_1(h_1)k_2(h_2)\cdots k_n(h_n)] D^{-\frac{1}{2}}$  is Hermitian, it can be brought into a diagonal form  $L^\rho$  by a similarity transformation. From the definition (40) of  $S$ , it is found that

$$\begin{aligned} L^\rho = V^\dagger D^{-\frac{1}{2}} S[k_1(h_1)k_2(h_2)\cdots k_n(h_n) | k_1(h_1)k_2(h_2)\cdots k_n(h_n)] D^{-\frac{1}{2}} V \\ = \int \{ \Phi[k_1(h_1)k_2(h_2)\cdots k_n(h_n)] D^{-\frac{1}{2}} V \}^\dagger \{ \Phi[k_1(h_1)k_2(h_2)\cdots k_n(h_n)] D^{-\frac{1}{2}} V \} d\tau. \end{aligned}$$

Since  $L^\rho$  is diagonal, the  $f_s$  functions,

$$\phi_i \equiv \{ \Phi[k_1(h_1)k_2(h_2)\cdots k_n(h_n)] D^{-\frac{1}{2}} V \}_i, \quad i = 1, 2, \dots, f_s,$$

are orthogonal to each other and therefore we obtain that

$$\sum_k \sum_m |L_{km}^\rho|^2 = \int \{ \sum_i \phi_i(1, 2, \dots, N) \phi_i^*(1'2' \dots N') \}^* \{ \sum_i \phi_i(1, 2, \dots, N) \phi_i^*(1'2' \dots N') \} d\tau d\tau'.$$

*Lemma 3.* If the following expansion of a matrix  $L$ ,

$$L^\rho = \sum_P d_P \tilde{U}^\rho(P), \quad (44)$$

is valid for all representations of  $\sigma_N$ , then  $\sum_P |d_P|^2$  is invariant under a unitary transformation of  $L$  and equal to  $\sum_\rho \sum_k \sum_m (f_\rho/N!) |L_{km}^\rho|^2$ .

By the use of the unitary matrix  $\Gamma$  defined by Eq. (16), the expansion (44) is written as

$$\Gamma d = l,$$

where  $d$  and  $l$  are the column matrices, in which the  $P$ th and the  $(\rho hm)$ th elements are given by  $d_P$  and  $(f_\rho/N!)^{\frac{1}{2}} L_{km}^\rho$ , respectively. Then

$$\sum_P |d_P|^2 = d^\dagger d = l^\dagger l = \sum_\rho \sum_k \sum_m (f_\rho/N!) |L_{km}^\rho|^2.$$

By a unitary transformation  $L'^\rho = L^\rho V$ ,  $\sum_k \sum_m |L_{km}^\rho|^2$  is invariant, since

$$\begin{aligned} \sum_k \sum_m |L_{km}'^\rho|^2 = \sum_k \sum_m \sum_l \sum_n L_{kl}^\rho L_{kn}^{\rho*} V_{lm} V_{nm}^* \\ = \sum_k \sum_l |L_{kl}^\rho|^2. \end{aligned}$$

This means that  $l^\dagger l$  and  $\sum_P |d_P|^2$  are also invariant under the unitary transformation.

By the use of the expansion (26),  $D^{-\frac{1}{2}} S[k(h) | h(h)] D^{-\frac{1}{2}}$  and  $D^{-\frac{1}{2}} S[k(h)m(l) | h(h)l(l)] D^{-\frac{1}{2}}$  can be written as

$$D^{-\frac{1}{2}} S[k(h) | h(h)] D^{-\frac{1}{2}} = \sum_P d[k(h) | h(h)]_P \tilde{U}(P), \quad (45)$$

$$\begin{aligned} D^{-\frac{1}{2}} S[k(h)m(l) | h(h)l(l)] D^{-\frac{1}{2}} \\ = \sum_P d[k(h)m(l) | h(h)l(l)]_P \tilde{U}(P), \quad (46) \end{aligned}$$

or, more generally,

Similarly,

$$\sum_k \sum_m |L_{km}'{}^\rho|^2 = \int \left\{ \sum_i \phi_i'(1, 2, \dots, N) \phi_i'^*(1'2' \dots N') \right\}^* \left\{ \sum_i \phi_i'(1, 2, \dots, N) \phi_i'^*(1'2' \dots N') \right\} d\tau d\tau',$$

where

$$L'{}^\rho = \int \phi'^{\dagger} \phi' d\tau,$$

and

$$\phi_i' = \{ \Phi[k_1'(h_1') k_2'(h_2') \dots k_n'(h_n')] D^{-\frac{1}{2}} V' \}_i, \quad i=1, 2, \dots, f_S$$

are orthogonal to each other.

Now we define  $L''{}^\rho$  by

$$L''{}^\rho = \int \phi^{\dagger} \phi' d\tau = \int \{ \phi[k_1(h_1) k_2(h_2) \dots k_n(h_n)] D^{-\frac{1}{2}} V \}^{\dagger} \{ \phi'[k_1'(h_1') k_2'(h_2') \dots k_n'(h_n')] D^{-\frac{1}{2}} V' \} d\tau.$$

Application of the Schwarz inequality relation to the above expressions for  $L{}^\rho$ ,  $L'{}^\rho$ , and  $L''{}^\rho$  leads to

$$\sum_k \sum_m |L_{km}''{}^\rho|^2 \leq \left\{ \sum_k \sum_m |L_{km}{}^\rho|^2 \cdot \sum_{k'} \sum_{m'} |L_{k'm'}'{}^\rho|^2 \right\}^{\frac{1}{2}}. \quad (49)$$

The relations is valid for all the representations of  $\sigma$ . Here we expand  $L{}^\rho$  such that

$$L{}^\rho = \sum_P d_P \tilde{U}(P).$$

By Lemma 3, we obtain the relation

$$\sum_P |d_P|^2 = \sum_\rho \sum_k \sum_m (f_\rho/N!) |L_{km}{}^\rho|^2,$$

and furthermore we find that the sum  $\sum_P |d_P|^2$  is equal to

$$\sum_P |d[k_1(h_1) k_2(h_2) \dots k_n(h_n) | k_1(h_1) k_2(h_2) \dots k_n(h_n)]_P|^2,$$

since  $D^{-\frac{1}{2}} S[k_1(h_1) k_2(h_2) \dots k_n(h_n) | k_1(h_1) k_2(h_2) \dots k_n(h_n)] D^{-\frac{1}{2}}$  are obtained from  $L{}^\rho$  by unitary transformations. By Lemma 2, it is also evident that

$$d[k_1(h_1) k_2(h_2) \dots k_n(h_n) | k_1(h_1) k_2(h_2) \dots k_n(h_n)]_P = d[k_1(k_1) k_2(k_2) \dots k_n(k_n) | k_1(k_1) k_2(k_2) \dots k_n(k_n)]_P.$$

This means that we have the relation

$$\sum_P |d[k_1(k_1) k_2(k_2) \dots k_n(k_n) | k_1(k_1) k_2(k_2) \dots k_n(k_n)]_P|^2 = \sum_\rho \sum_k \sum_m (f_\rho/N!) |L_{km}{}^\rho|^2.$$

The similar relations will be found for  $L'{}^\rho$  and  $L''{}^\rho$ . Therefore, Eq. (49) can be written as Eq. (48).

From the discussion given in Sec. II, it is clear that

$$\left\{ \sum_P |d[k_1(k_1) k_2(k_2) \dots k_n(k_n) | k_1(k_1) k_2(k_2) \dots k_n(k_n)]_P|^2 \right\}^{\frac{1}{2}}$$

is bounded. We shall denote the maximum and minimum values for arbitrary  $k_1 k_2 \dots k_n$  by  ${}^n l_{\max}$  and  ${}^n l_{\min}$ . Then Lemma 4 is written as

*Lemma 4'.*

$$\sum_P |d[k_1(h_1) k_2(h_2) \dots k_n(h_n) | k_1'(h_1') k_2'(h_2') \dots k_n'(h_n')]_P|^2 \leq ({}^n l_{\max})^2. \quad (50)$$

Henceforth, we will deal mostly with the case where  $h_1 = h_1' = k_1'$ ,  $h_2 = h_2' = k_2'$ ,  $\dots$  and  $h_n = h_n' = k_n'$ . In such a case, we shall not repeat the indices  $h_1, h_2, \dots, h_n$  in  $S$  and  $d$  and will use the following abbreviations:

$$\begin{aligned} S[k_1 k_2 \dots k_n | h_1 h_2 \dots h_n] &\equiv S[k_1(h_1) k_2(h_2) \dots k_n(h_n) | h_1(h_1) h_2(h_2) \dots h_n(h_n)], \\ d[k_1 k_2 \dots k_n | h_1 h_2 \dots h_n] &\equiv d[k_1(h_1) k_2(h_2) \dots k_n(h_n) | h_1(h_1) h_2(h_2) \dots h_n(h_n)]. \end{aligned}$$

## 2. Expansion of Overlap Matrices

Now we shall evaluate the upper and lower bounds  ${}^n l_{\max}$  and  ${}^n l_{\min}$  of  $\sum_P |d_P|^2$  defined above. This will be carried out easily if we expand the overlap matrices  $S[h_1 \dots h_{n-1} | h_1 \dots h_{n-1}]$  in terms of

$$S[h_1 \dots h_{n-1} k_n | h_1 \dots h_{n-1} h_n].$$

First, we consider  ${}^1 l_{\max}$  and  ${}^1 l_{\min}$ . By the use of the

expansion (45), Eq. (36) is written as

$$1 = \sum_k \langle k | h \rangle D^{-\frac{1}{2}} S[k | h] D^{-\frac{1}{2}} = \sum_P \sum_k \langle k | h \rangle d[k | h]_P \tilde{U}(P).$$

Applying Lemma 1 to the equation, we find that

$$1 = \sum_k \langle k | h \rangle d[k | h]_E, \quad (51)$$

$$0 = \sum_k \langle k | h \rangle d[k | h]_P. \quad (52)$$

Therefore, we obtain that

$$\begin{aligned} \sum_P |d[h|h]_P|^2 &= 1 - \sum_{k'} \{ \langle k|h \rangle d[k|h]_E + \langle k|h \rangle^* d[k|h]_E^* \} \\ &+ \sum_{k'} \sum_{k''} \langle k|h \rangle \langle k''|h \rangle^* \{ \sum_P d[k|h]_P d[k''|h]_P^* \}. \end{aligned} \quad (53)$$

where the summation  $\sum_{k'}$  includes  $k=1, 2, \dots, h-1; h+1, \dots, N$ .

From the definition of  ${}^n l_{\max}$  and Lemma 4', it is evident that

$$d[k|h]_E \leq {}^n l_{\max},$$

and

$$\begin{aligned} \sum_P d[k|h]_P \cdot d[k'|h]_P^* \\ \leq \{ \sum_P |d[k|h]_P|^2 \cdot \sum_P |d[k'|h]_P|^2 \}^{\frac{1}{2}} \leq ({}^n l_{\max})^2. \end{aligned}$$

By the use of the above two relations in Eq. (53), it is easily seen that the relation

$$\sum_P |d[h|h]_P|^2 \leq 1 + 2\Delta {}^n l_{\max} + \Delta^2 ({}^n l_{\max})^2$$

is valid for any  $h$ , where  $\Delta$  is the maximum value of  $\Delta(h) = \sum_{k'} \langle k|h \rangle$ ,  $h=1, 2, \dots, N$ . This means that

$$({}^n l_{\max})^2 \leq 1 + 2\Delta {}^n l_{\max} + \Delta^2 ({}^n l_{\max})^2.$$

If  $\Delta < 1$ , the relation is written as

$${}^n l_{\max} \leq (1 - \Delta)^{-1}.$$

Similarly, we find that

$$({}^n l_{\min})^2 \geq 1 - 2\Delta {}^n l_{\max} - \Delta^2 ({}^n l_{\max})^2.$$

Therefore,

$$\begin{aligned} (1 - 4\Delta + 2\Delta^2)(1 - \Delta)^{-2} \\ \leq \sum_P |d[h|h]_P|^2 \leq (1 - \Delta)^{-2}. \end{aligned} \quad (54)$$

From Eqs. (51) and (52), it is also found that

$$(1 - 2\Delta)(1 - \Delta)^{-1} \leq d[h|h]_E \leq (1 - \Delta)^{-1}, \quad (55)$$

$$d[h|h]_P \leq \Delta(1 - \Delta)^{-1}, \quad (56)$$

and also

$$\sum_P |d[k|h]_P|^2 \leq (1 - \Delta)^{-2}, \quad (57)$$

provided  $\Delta < 1$ .

Now we consider  ${}^2 l_{\max}$  and  ${}^2 l_{\min}$ . By the definition (39) of  $\Phi$ , it is immediately found that

$$D^{-\frac{1}{2}} S[k|k] D^{-\frac{1}{2}} = \sum_{m'} \langle m|l \rangle D^{-\frac{1}{2}} S[km|kl] D^{-\frac{1}{2}},$$

or, in terms of the representation matrices,

$$\sum_P d[k|k]_P \tilde{U}(P) = \sum_P \sum_{m'} \langle m|l \rangle d[km|kl]_P \tilde{U}(P),$$

where  $\sum_{m'}$  excludes  $m=k$ . By Lemma 1, therefore, we obtain that

$$d[kl|kl]_P = d[k|k]_P - \sum_{m''} \langle m|l \rangle d[km|kl]_P, \quad (58)$$

and also

$$\begin{aligned} \sum_P |d[kl|kl]_P|^2 &= \sum_P |d[k|k]_P|^2 - \sum_{m''} \{ \langle m|l \rangle \sum_P d[k|k]_P^* \\ &\cdot d[km|kl]_P + \langle m|l \rangle^* \sum_P d[k|k]_P d[km|kl]_P^* \} \\ &+ \sum_{m''} \sum_{m'''} \{ \langle m|l \rangle \langle m'|l \rangle^* \sum_P d[km|kl]_P \\ &\cdot d[km'|kl]_P^* \}, \end{aligned} \quad (59)$$

where the summation  $\sum_{m''}$  excludes  $m=k$  and  $l$ .

By the use of Lemma 4' and the Schwarz inequality relation in Eqs. (58) and (59), it is found that

$$\begin{aligned} \{ (1 - 4\Delta + 2\Delta^2) - 2\Delta_k (1 - \Delta_k)^{-1} - \Delta_k^2 (1 - \Delta_k)^{-2} \} (1 - \Delta)^{-2} \\ \leq \sum_P |d[kl|kl]_P|^2 \leq (1 - \Delta_k)^{-2} (1 - \Delta)^{-2}, \end{aligned} \quad (60)$$

$$\begin{aligned} (1 - 2\Delta - 2\Delta_k + 2\Delta\Delta_k)(1 - \Delta_k)^{-1} (1 - \Delta)^{-1} \\ \leq d[kl|kl]_E \leq (1 - \Delta_k)^{-1} (1 - \Delta)^{-1}, \end{aligned} \quad (61)$$

$$d[kl|kl]_P \leq \{ \Delta + \Delta_k (1 - \Delta_k)^{-1} \} (1 - \Delta)^{-1}, \quad (62)$$

and

$$\sum_P |d[km|hl]_P|^2 \leq (1 - \Delta_k)^{-2} (1 - \Delta)^{-2}, \quad (63)$$

where  $\Delta_k$  is the maximum value of

$$\Delta_k(l) = \sum_{m''} \langle m|l \rangle, \quad l=1, 2, \dots, k-1; k+1, \dots, N.$$

These relations are valid as long as  $\Delta < 1$  since  $\Delta_k < \Delta$ . Furthermore, similar relations will be found for the coefficients  $d[k_1 k_2 \dots k_n | h_1 h_2 \dots h_n]_P$ .

From the results given by Eqs. (54)–(57) and (60)–(63), it is evident that the expansions

$$E_k = \sum_P \sum_k \langle k | \mathfrak{C}_k | h \rangle d[k|h]_P \tilde{U}(P), \quad (64)$$

and

$$E_{hl} = \sum_P \sum_k \sum_{m'} \langle km | \mathfrak{C}_{hl} | hl \rangle d[km|hl]_P \tilde{U}(P), \quad (65)$$

obtained from Eqs. (37), (38), (45), and (46), converge rapidly and provide a practical way to evaluate the energy since the higher permutations are indeed negligible. On the other hand, if  $\Delta$  is close to or larger than 1, we have to take the interaction between the nearest neighbor atoms explicitly. Then, the summation in  $\Delta$  starts from the next nearest neighbor atoms. Since the overlap integrals between atoms decrease exponentially, a satisfactory solution will be obtained by this approach.

### 3. Heisenberg's Spin Hamiltonian

Finally, we shall introduce a more explicit expression for the exchange interactions  $J_P$  and show that all  $J_P$  are bounded in a similar fashion as  $d_P$ . Here the energy matrix  $E_{hl}$  given by Eq. (38) will be considered since  $E_h$  can be treated exactly the same as  $E_{hl}$ .

Following the definition of  $S[k_1 k_2 \dots k_n | h_1 h_2 \dots h_n]$  and  $\Phi[k_1(h_1) k_2(h_2) \dots k_n(h_n)]$  as given by Eqs. (39)

and (40), the expression (38) is written as

$$\begin{aligned}
E_{h_1 h_2} &= E_{h_1 h_2}^{\text{I}} + E_{h_1 h_2}^{\text{II}}, \\
E_{h_1 h_2}^{\text{I}} &= \langle h_2 h_1 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle D^{-\frac{1}{2}} S[h_2 h_1 | h_1 h_2] D^{-\frac{1}{2}} + \sum_{h_3}'' \langle h_1 h_3 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle \langle h_2 | h_3 \rangle D^{-\frac{1}{2}} S[h_1 h_3 h_2 | h_1 h_2 h_3] D^{-\frac{1}{2}} \\
&\quad + \sum_{h_3}'' \langle h_3 h_2 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle \langle h_1 | h_3 \rangle D^{-\frac{1}{2}} S[h_3 h_2 h_1 | h_1 h_2 h_3] D^{-\frac{1}{2}} \\
E_{h_1 h_2}^{\text{II}} &= \sum_{h_3}'' \sum_{h_4}''' \langle h_1 h_3 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle \langle h_4 | h_3 \rangle D^{-\frac{1}{2}} S[h_1 h_3 h_4 | h_1 h_2 h_3] D^{-\frac{1}{2}} \\
&\quad + \sum_{h_3}'' \sum_{h_4}''' \langle h_3 h_1 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle \langle h_4 | h_3 \rangle D^{-\frac{1}{2}} S[h_3 h_1 h_4 | h_1 h_2 h_3] D^{-\frac{1}{2}} \\
&\quad + \sum_{h_3}'' \sum_{h_4}''' \langle h_3 h_2 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle \langle h_4 | h_3 \rangle D^{-\frac{1}{2}} S[h_3 h_2 h_4 | h_1 h_2 h_3] D^{-\frac{1}{2}} \\
&\quad + \sum_{h_3}'' \sum_{h_4}''' \langle h_2 h_3 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle \langle h_4 | h_3 \rangle D^{-\frac{1}{2}} S[h_2 h_3 h_4 | h_1 h_2 h_3] D^{-\frac{1}{2}} \\
&\quad + \sum_{h_3}'' \sum_{h_4}''' \sum_{h_5}'' \sum_{h_6}''' \langle h_3 h_4 | \mathfrak{J}C_{h_1 h_2} | h_1 h_2 \rangle \langle h_5 h_6 | h_3 h_4 \rangle D^{-\frac{1}{2}} S[h_3 h_4 h_5 h_6 | h_1 h_2 h_3 h_4] D^{-\frac{1}{2}},
\end{aligned} \tag{66}$$

where the summations  $\sum_{h_3}''$  and  $\sum_{h_4}'''$  exclude  $h_3 = h_1, h_2$  and  $h_4 = h_1, h_2, h_3$ , respectively,  $\sum_{h_5}''$  and  $\sum_{h_6}'''$  exclude  $h_5 = h_3, h_4$  and  $h_6 = h_3, h_4, h_5$ , while  $\sum_{h_4}'''$  excludes  $h_1$  and  $h_3$  (or  $h_2$  and  $h_5$ ).

If we expand the energy expression  $E_{h_1 h_2}$  in terms of  $\tilde{U}(P)$  and take a summation over  $h_1$  and  $h_2$ , we will obtain the desired expression for  $J_P$  as a sum of products of exchange integrals  $\langle km | \mathfrak{J}C_{kl} | hl \rangle$  and  $d_P$ . Before that, it is convenient to consider the expansion term by term.

Let us first consider the matrix  $D^{-\frac{1}{2}} S[lk | kl] D^{-\frac{1}{2}}$ . According to the definition, we find that

$$S[lk | kl] = \tilde{U}(kl) S[kl | kl]. \tag{67}$$

Also we shall expand the expression  $D^{-\frac{1}{2}} \tilde{U}(R) D^{\frac{1}{2}}$ , as follows:

$$D^{-\frac{1}{2}} \tilde{U}(R) D^{\frac{1}{2}} = \{1 + \sum_P a(R)_P \tilde{U}(P)\} \tilde{U}(R). \tag{68}$$

By using the above relations together with the expansions (45)~(47), therefore, our matrix will be written as

$$\begin{aligned}
&D^{-\frac{1}{2}} S[lk | kl] D^{-\frac{1}{2}} \\
&= D^{-\frac{1}{2}} \tilde{U}(kl) D^{\frac{1}{2}} \cdot D^{-\frac{1}{2}} S[kl | kl] D^{-\frac{1}{2}} \\
&= \{1 + \sum_Q a(kl)_Q \tilde{U}(Q)\} \tilde{U}(kl) \cdot \{\sum_{Q'} d[kl | kl]_{Q'} \tilde{U}(Q')\} \\
&= \sum_P \{d[kl | kl]_{P(kl)} + \sum_Q a(kl)_Q d[kl | kl]_{PQ^{-1}(kl)}\} \tilde{U}(P).
\end{aligned}$$

The matrix  $D^{-\frac{1}{2}} S[lk | kl] D^{-\frac{1}{2}}$  can also be expanded directly as  $\sum_P d[lk | kl]_P \tilde{U}(P)$ . Comparison of the two expansions shows that

$$\begin{aligned}
d[lk | kl]_P &= d[kl | kl]_{P(kl)} \\
&\quad + \sum_Q a(kl)_Q d[kl | kl]_{PQ^{-1}(kl)}. \tag{69}
\end{aligned}$$

In case of  $P = (kl)$ , this relation becomes

$$\begin{aligned}
d[lk | kl]_{(kl)} &= d[kl | kl]_E \\
&\quad + \sum_Q a(kl)_Q d[kl | kl]_{(kl)Q^{-1}(kl)}. \tag{70}
\end{aligned}$$

In Appendix C, we shall examine the expansion (68) and prove that  $\sum_P |a(R)_P|^2$  is small as compared with

unity. The results (A11), (A20), (A21), and Eqs. (60)~(63) together with Schwarz's inequality can be applied to Eq. (70), and it will be found that

$$A_1 - (3M_{\max})^{\frac{1}{2}} B_1 \leq d[lk | kl]_{(kl)} \leq B_1 + (3M_{\max})^{\frac{1}{2}} B_1. \tag{71}$$

Similarly, the following relation will be obtained from Eq. (69) when  $P \neq (kl)$ :

$$|d[lk | kl]_P| \leq \frac{1}{2}(B_1 - A_1) + (3M_{\max})^{\frac{1}{2}} B_1, \tag{72}$$

where

$$\begin{aligned}
A_1 &= 1 - \Delta(1 - \Delta)^{-1} - \Delta_k(1 - \Delta_k)^{-1}(1 - \Delta)^{-1}, \\
B_1 &= (1 - \Delta_k)^{-1}(1 - \Delta)^{-1}.
\end{aligned}$$

The other matrices can be handled in the same manner, and it will be found that

$$\begin{aligned}
A_2 - (3M_{\max})^{\frac{1}{2}} B_2 &\leq d[hkl | hkl]_{(kl)} \\
&\leq B_2 + (3M_{\max})^{\frac{1}{2}} B_2, \tag{73}
\end{aligned}$$

$$|d[hkl | hkl]_P| \leq \frac{1}{2}(B_2 - A_2) + (3M_{\max})^{\frac{1}{2}} B_2, \tag{74}$$

for  $P \neq (kl)$ .

where

$$\begin{aligned}
A_2 &= 1 - \Delta(1 - \Delta)^{-1} - \Delta_k(1 - \Delta_k)^{-1}(1 - \Delta)^{-1} \\
&\quad - \Delta_{kl}(1 - \Delta_{kl})^{-1}(1 - \Delta_k)^{-1}(1 - \Delta)^{-1}, \\
B_2 &= (1 - \Delta_{kl})^{-1}(1 - \Delta_k)^{-1}(1 - \Delta)^{-1}.
\end{aligned}$$

Let us now come back to the total energy,

$$\sum_{h_1 > h_2} E_{h_1 h_2},$$

and expand it as follows:

$$\sum_{h_1 > h_2} E_{h_1 h_2} = \sum_P J_P \tilde{U}(P). \tag{75}$$

By using expansions (46)~(47), the energy matrix  $E_{h_1 h_2}$  can also be expanded in terms of  $\tilde{U}(P)$ . Comparison of the resulting expansion with Eq. (75) shows that

the exchange interaction energy  $J_{ij}$  between electrons  $i$  and  $j$  is given by

$$\begin{aligned}
J_{ij} &= J_{ij}^I + J_{ij}^{II} + J_{ij}^{III}, \\
J_{ij}^I &= \langle ji | \mathfrak{C}_{ij} | ij \rangle d[ji | ij]_{(ij)} + \sum_h'' \{ \langle jh | \mathfrak{C}_{ih} | ih \rangle \langle i | j \rangle + \langle ih | \mathfrak{C}_{jh} | jh \rangle \langle j | i \rangle \} d[jih | ijh]_{(ij)} \\
J_{ij}^{II} &= \frac{1}{2} \sum_{h_1}' \sum_{h_2}'' \langle h_2 h_1 | \mathfrak{C}_{h_1 h_2} | h_1 h_2 \rangle d[h_2 h_1 | h_1 h_2]_{(ij)} + \sum_{h_1}' \sum_{h_2}'' \sum_{h_3}''' \langle h_2 h_3 | \mathfrak{C}_{h_1 h_3} | h_1 h_3 \rangle \langle h_1 | h_2 \rangle d[h_2 h_1 h_3 | h_1 h_2 h_3]_{(ij)} \\
J_{ij}^{III} &= \sum_{h_1}' \sum_{h_2}'' \sum_{h_3}''' \sum_{h_4}'''' \langle h_1 h_3 | \mathfrak{C}_{h_1 h_2} | h_1 h_2 \rangle \langle h_4 | h_3 \rangle d[h_1 h_3 h_4 | h_1 h_2 h_3]_{(ij)} \\
&\quad + \sum_{h_1}' \sum_{h_2}'' \sum_{h_3}''' \sum_{h_4}'''' \langle h_3 h_1 | \mathfrak{C}_{h_1 h_2} | h_1 h_2 \rangle \langle h_4 | h_3 \rangle d[h_3 h_1 h_4 | h_1 h_2 h_3]_{(ij)} \\
&\quad + \frac{1}{2} \sum_{h_1}' \sum_{h_2}'' \sum_{h_3}''' \sum_{h_4}'''' \sum_{h_5}'''''' \sum_{h_6}'''''' \langle h_3 h_4 | \mathfrak{C}_{h_1 h_3} | h_1 h_2 \rangle \langle h_5 h_6 | h_3 h_4 \rangle d[h_3 h_4 h_5 h_6 | h_1 h_2 h_3 h_4]_{(ij)},
\end{aligned} \tag{76}$$

where  $\sum_{h_1}' \sum_{h_2}''$  in  $J_{ij}^{II}$  excludes  $h_1, h_2 = i, j$  and  $j, i$  as well as  $h_1 = h_2$ .

By applying relations (70)~(73), we find that

$$\begin{aligned}
&\langle ji | \mathfrak{C}_{ij} | ij \rangle [A_1 - (3M_{\max})^{\frac{1}{2}} B_1] \\
&\quad + \sum_{h''}'' [\langle jh | \mathfrak{C}_{ih} | ih \rangle \langle i | j \rangle + \langle ih | \mathfrak{C}_{jh} | jh \rangle \langle j | i \rangle] \\
&\quad \quad \times [A_2 - (3M_{\max})^{\frac{1}{2}} B_2] \\
&\leq J_{ij}^I \leq \langle ji | \mathfrak{C}_{ij} | ij \rangle B_1 [1 + (3M_{\max})^{\frac{1}{2}}] \\
&\quad + \sum_{h''}'' [\langle jh | \mathfrak{C}_{ih} | ih \rangle \langle i | j \rangle + \langle ih | \mathfrak{C}_{jh} | jh \rangle \langle j | i \rangle] \\
&\quad \quad \times B_2 [1 + (3M_{\max})^{\frac{1}{2}}]. \tag{77}
\end{aligned}$$

$$J_{ij}^{II} = \frac{1}{2} \sum_{h_2}' \langle h_2 i | \mathfrak{C}_{ih} | ih_2 \rangle \sum_T' d[h_2 i | ih_2]_{T(ij)T^{-1}} + \sum_{h_2}' \sum_{h_3}' \langle h_2 h_3 | \mathfrak{C}_{ih} | ih_3 \rangle \langle i | h_2 \rangle \sum_T' d[h_2 ih_3 | ih_2 h_3]_{T(ij)T^{-1}}, \tag{78}$$

where  $\sum_{h_2}'$  and  $\sum_{h_3}'$  exclude  $h_2 = i$  and  $h_3 = i, h_2$ , while  $\sum_T'$  excludes  $T = E$ , when  $h_2 = j$ , and  $T$ , which shifts  $i$  and  $j$  to  $k$  and  $i$ , when  $h_2 = k$ . By  $T(ij)T^{-1}$ ,  $(ij)$  is shifted to  $(i_T j_T)$ .

Here the summations  $\sum_{T''} d[h_2 i | ih_2]_{T(ij)T^{-1}}$  and

$$\sum_{T''} d[h_2 ih_3 | ih_2 h_3]_{T(ij)T^{-1}}$$

converge rapidly, since  $d[hi | ih]_{T(ij)T^{-1}}$  decreases proportionally to  $\exp(-R)$ , when the distance  $R$  between  $ih$  and  $T(ij)T^{-1}$  increases. This will be explained as follows. For simplicity, we shall show that  $d[k | k]_{(lm)}$  decreases proportionally to  $\exp(-R)$ , where  $R$  is the distance between  $k$  and  $m$ . The extension to  $d[ij | ij]_{(lm)}$  or  $d[ji | ij]_{(lm)}$  is trivial.

Let us consider the expansion

$$\begin{aligned}
1 &= D^{-\frac{1}{2}} D D^{-\frac{1}{2}} = D^{-\frac{1}{2}} S [k | k] D^{-\frac{1}{2}} \\
&\quad + \sum_{h'} \langle h | k \rangle D^{-\frac{1}{2}} S [h | k] D^{-\frac{1}{2}} \\
&= \sum_P d[k | k]_P \tilde{U}(P) + \sum_{h'} \langle h | k \rangle \sum_P d[h | k]_P \tilde{U}(P). \tag{79}
\end{aligned}$$

By taking the term  $P = (lm)$ , we find that

$$\begin{aligned}
d[k | k]_{(lm)} &= - \sum_{h'} \langle h | k \rangle d[h | k]_{(lm)} \\
&= - \langle l | k \rangle d[l | k]_{(lm)} - \langle m | k \rangle d[m | k]_{(lm)} \\
&\quad - \sum_{h''}'' \langle h | k \rangle d[h | k]_{(lm)}, \tag{80}
\end{aligned}$$

where  $\sum_{h''}''$  excludes  $h = k, l$  and  $m$ . From Eq. (67), we

The second term  $J_{ij}^{II}$  is a small correction term to  $J_{ij}^I$ . This will be seen if we consider a perfect crystal. Here we assume that there are  $N$  permutations  $T$  of electrons, which are also elements of the translation group of the crystal. As we shall discuss later in detail and will give in Eqs. (91) and (92), the summation over lattice points  $h_1$  in  $J_{ij}^{II}$  can be replaced by the summation over translation  $T$  such that

notice that

$$d[h | k]_{(lm)} \approx d[k | k]_{(kh)(lm)}.$$

Furthermore, we are able to show that

$$d[k | k]_{(kh)(lm)} \approx d[k | k]_{(kh)} \cdot d[k | k]_{(lm)}, \tag{81}$$

as long as neither  $k$  or  $h$  is equal to  $l$  or  $m$ . Since both  $d[k | k]_{(kh)}$  and  $d[k | k]_{(lm)}$  are bounded as given by Eq. (56) and smaller than 1 when  $\Delta < 1$ , the above two relations lead to

$$|d[k | k]_{(lm)}| > |d[h | k]_{(lm)}|.$$

Therefore, we obtain that

$$\begin{aligned}
|d[k | k]_{(lm)} + \sum_{h''}'' \langle h | k \rangle d[h | k]_{(lm)}| \\
> (1 - \Delta) |d[k | k]_{(lm)}|.
\end{aligned}$$

Use of this result in Eq. (80) gives us the relation

$$\begin{aligned}
|\langle l | k \rangle| |d[l | k]_{(lm)}| + |\langle m | k \rangle| |d[m | k]_{(lm)}| \\
> (1 - \Delta) |d[k | k]_{(lm)}|,
\end{aligned}$$

or

$$|d[k | k]_{(lm)}| < \Delta (1 - \Delta)^{-2} \{ |\langle l | k \rangle| + |\langle m | k \rangle| \}. \tag{82}$$

This shows that  $d[k | k]_{(lm)}$  should decrease proportionally to  $\exp(-R)$ , when the distance  $R$  between  $k$  and  $l, m$  increases, since the overlap integrals  $\langle l | k \rangle$  and  $\langle m | k \rangle$  decrease as  $\exp(-R)$ . This will be true if Eq. (81) is valid.

The relation (81) will be justified as follows. By using the expansion (79),  $D^{-\frac{1}{2}}S[k|k]D^{-\frac{1}{2}}$  is written as

$$\begin{aligned} D^{-\frac{1}{2}}S[k|k]D^{-\frac{1}{2}} &= 1 - \sum_h \langle h|k \rangle D^{-\frac{1}{2}}S[h|k]D^{-\frac{1}{2}} \\ &= 1 - \sum_h \langle hk|kh \rangle D^{-\frac{1}{2}}S[hk|kh]D^{-\frac{1}{2}} \\ &\quad - \sum_h \sum_l \langle hl|kh \rangle D^{-\frac{1}{2}}S[hl|kh]D^{-\frac{1}{2}}. \end{aligned}$$

Because of the relations (68) and (A20), it is evident that the main contribution to  $d[k|k]_{(kh)}$  comes from the term

$$\begin{aligned} -\langle hk|kh \rangle D^{-\frac{1}{2}}S[hk|kh]D^{-\frac{1}{2}} \\ = -\langle hk|kh \rangle D^{-\frac{1}{2}}\tilde{U}(kh)D^{\frac{1}{2}}D^{-\frac{1}{2}}S[kh|kh]D^{-\frac{1}{2}}. \end{aligned}$$

Since  $D^{-\frac{1}{2}}\tilde{U}(kh)D^{-\frac{1}{2}} \approx \tilde{U}(kh)$  and  $d[kh|kh]_E \approx 1$ , we find that

$$d[k|k]_{(kh)} \approx -\langle hk|kh \rangle, \quad (83)$$

and also

$$d[k|k]_{(kh)(lm)} \approx -\langle hk|kh \rangle d[kh|kh]_{(lm)}. \quad (84)$$

On the other hand,  $D^{-\frac{1}{2}}S[k|k]D^{-\frac{1}{2}}$  will be expanded as

$$\begin{aligned} D^{-\frac{1}{2}}S[k|k]D^{-\frac{1}{2}} &= D^{-\frac{1}{2}}S[kh|kh]D^{-\frac{1}{2}} \\ &\quad + \sum_l \langle l|h \rangle D^{-\frac{1}{2}}S[kl|kh]D^{-\frac{1}{2}}. \end{aligned}$$

Here, the main part of the second term is  $\sum_l \langle l|h \rangle \tilde{U}(hl)$  and this will not give appreciable contribution to  $d[k|k]_{(lm)}$ . Therefore,

$$d[k|k]_{(lm)} \approx d[kh|kh]_{(lm)}. \quad (85)$$

The relations (83)~(85) justify Eq. (81).

This proves that the second term  $J_{ij}^{\text{II}}$  of the exchange interaction should also be bounded. The first term of  $J_{ij}^{\text{II}}$  is less than

$$\begin{aligned} \frac{1}{4} \sum'_{h_2} \langle h_2 i | \mathfrak{C}_{ih_2} | ih_2 \rangle (B_1 - A_1) \\ + \frac{1}{2} \sum'_{h_2} \sum'_{h_3} \langle h_2 h_3 | \mathfrak{C}_{ih_3} | ih_3 \rangle \langle i | h_2 \rangle (B_2 - A_2), \quad (86) \end{aligned}$$

and the higher terms will decrease proportionally to  $\exp(-R)$ . From this proof, it is clear that the arguments remain correct even if the crystal is not perfect.

The third term  $J_{ij}^{\text{III}}$  is a contribution from  $E_{h_1 h_2}^{\text{II}}$ . It is possible to bound  $J_{ij}^{\text{III}}$  like  $J_{ij}^{\text{II}}$ , although the expression will become very complicated. Without carrying out the direct method, however, it is easily seen that  $J_{ij}^{\text{III}}$  is small for the following reasons. First,  $E_{h_1 h_2}^{\text{II}}$  is smaller than  $E_{h_1 h_2}^{\text{I}}$  by a factor  $\Delta$ . Secondly, the main part of  $E_{h_1 h_2}^{\text{II}}$  belongs to the three electron exchange interaction terms  $J_{(ij)(jk)} \tilde{U}((ij)(jk))$ . The difference between  $E_{h_1 h_2}^{\text{II}}$  and the three electron exchange interaction terms will be found to be small, while the main part  $J_{(ij)(jk)}^{\text{III}}$  of the three electron exchange terms can be calculated in a similar way as  $J_{ij}^{\text{I}}$ .

Up to this moment, we have shown that the energy matrix  $E$  can be expanded as  $\sum_P J_P \tilde{U}(P)$  and that the terms  $J_{ij}$  for permutations of single pairs  $(ij)$  are equivalent to the Heisenberg exchange terms. Furthermore,  $J_{ij}$  can be given by Eq. (76) and the values

should be bounded as shown by Eqs. (77) and (86) and the subsequent text, when  $\Delta < 1$ . Those arguments show that exact values of  $J_{ij}$  should not be so different from values obtained by omitting higher permutations as long as  $\Delta$  is small as compared with unity. In addition, it has been shown that  $J_P$  for higher permutations  $P$  is smaller than  $J_{ij}$  by a factor  $\Delta$ . The foregoing arguments disprove the nonorthogonality catastrophe. Although we have not given explicit calculations in cases where  $\Delta \geq 1$ , it is evident from the discussion in Sec. III.2, that  $d_P$  can be bounded similar to Eqs. (54)~(57) and (60)~(63). The exchange interaction energy  $J_P$  will be calculated similarly, and higher permutations and exchange interaction energy between distant electrons will vanish. However, effects from second- and third-nearest neighbor atoms will not be negligible when  $\Delta$  is not small as compared with unity. The existence of antiferromagnetic substances such as MnO indicates this possibility. We shall discuss such a case in a future communication.

Finally, we shall show that the total energy calculated is proportional to the number  $N$  of electrons in crystals regardless of whether we use the energy expansion (3), the effective Hamiltonian (4), or the Heisenberg spin Hamiltonian (2). This assures that the use of the Heisenberg approximation as well as the existence of the expansion proposed here is really justified.

We define the energy  $E(i)$  attributed to the  $i$ th electron by

$$E(i) = \sum_R \frac{1}{M_R} J_R \tilde{U}(R) = \sum_M \sum_{R_M} \frac{1}{M_R} J_{R_M} \tilde{U}(R_M), \quad (87)$$

where the summation  $\sum_R$  includes all permutations which shift electron  $i$  from orbital  $\psi_i$  to other orbitals. Furthermore, we classify the permutations as follows. By  $R_M$ , we consider all possible permutations of  $M_R$  electrons by which orbitals  $\psi_i, \psi_{i_2}, \psi_{i_3}, \dots, \psi_{i_M}$  are replaced by  $\psi_j, \psi_{j_2}, \dots, \psi_{j_M}$  such that  $\psi_i \neq \psi_j$  and  $\psi_{i_k} \neq \psi_{j_k}$  for  $k=2, 3, \dots, M$ , while the rest of the electrons are not permuted. The fact  $1/M_R$  must be inserted in Eq. (87) since the same permutation will be taken into account  $M_R$  times when we consider  $E(i_k)$  for  $k=2, 3, \dots, M$ .

For simplicity, we shall consider  $E(i)$  in a perfect crystal and show that  $E(i)$  is finite and the total energy of the crystal is given by  $NE(i)$ . Furthermore, it will be shown that the main part of  $E(i)$  is described by  $\frac{1}{2} \sum_j J_{ij} \tilde{U}(ij)$ . This indicates that the exchange interactions between two electrons are indeed dominant as compared with the effect of simultaneous permutations of more than two electrons as long as overlap integrals are small; and the expansion (4) or (33) converges rapidly in many-electron systems. The perfect crystal here means that there exist  $N$  permutations  $T$  of electrons, which are at the same time elements of the translation group of the crystal. By applying  $T$  to  $\psi$ ,

electrons 1, 2,  $\dots$ ,  $N$  are shifted to orbitals  $\psi_{T_1}, \psi_{T_2}, \dots, \psi_{T_N}$ , but  $\psi_{T_1}, \psi_{T_2}, \dots, \psi_{T_N}$  are equivalent to  $\psi_1, \psi_2, \dots, \psi_N$  in the following sense:

$$\int \psi_{T_i}^* \psi_{T_j} d\tau = \int \psi_i^* \psi_j d\tau, \quad (88)$$

for any pair of  $i$  and  $j$ .

Let us consider the summation

$$F(i) = \sum_{h_2}' E_{ih_2} = \sum_{h_2}' \sum_{k_1}' \sum_{k_2}' \langle k_1 k_2 | \mathcal{H}_{ih_2} | ih_2 \rangle \times \sum_P d[k_1 k_2 | ih_2]_P \tilde{U}(P). \quad (89)$$

From the results given by Eqs. (60)–(63), it is evident that  $F(i)$  converges and is not very different from  $\frac{1}{2} \sum_j' J_{ij} \tilde{U}(ij)$ , if  $\Delta$  is small as compared with unity. Since  $N!$  permutations  $P$  can be divided as  $(1/M_R) \times TRT^{-1}$ , Eq. (89) is written as

$$F(i) = \sum_R \sum_T \frac{1}{M_R} J_{TRT^{-1}}(i) \tilde{U}(TRT^{-1}), \quad (90)$$

where

$$J_{TRT^{-1}}(i) = \sum_{h_2}' \sum_{k_1}' \sum_{k_2}' \langle k_1 k_2 | \mathcal{H}_{ih_2} | ih_2 \rangle \times d[k_1 k_2 | ih_2]_{TRT^{-1}}, \quad (91)$$

and, in the summation  $\sum_T$ , only independent elements  $TRT^{-1}$  are taken. Since  $T_i T_j = T_j T_i$ , the representation matrices  $\tilde{U}(T)$  of  $T$  are constant,  $\tilde{U}(T)$  of  $\sigma_N$  can be written as  $\tilde{U}(Q)$  given by Eq. (A5).<sup>15</sup> Because of the symmetry relation (88), it is also evident that

$$J_{TRT^{-1}}(i) = J_R(T_i^{-1}),$$

where electron  $i$  is shifted to  $T_i^{-1}$  by translation  $T$ . Therefore, we find that

$$\sum_T J_{TRT^{-1}}(i) = \sum_T J_R(T_i^{-1}) = J_R, \quad (92)$$

and the invariant part of  $F(i)$  is written as

$$F(i) = \sum_R \frac{1}{M_R} J_R \tilde{U}(R),$$

which is equivalent to the expression (87). This proves that  $E(i)$  is finite and not far different from  $\frac{1}{2} \sum_j' J_{ij} \times \tilde{U}(ij)$ . From Eq. (88), it is also clear that  $E(i)$  is independent of  $i$  and the total energy is given by  $NE(i)$ . If we apply a similar technique to the second part  $E_{h_1 h_2}^{\text{II}}$  of the energy matrix (66), it will be easily proved that the contribution from  $E_{h_1 h_2}^{\text{II}}$  to  $E(i)$  is really a small fraction,  $\sim \Delta \times E(i)$ .

In conclusion, use of the effective spin Hamiltonian (4) is practical since the series converges rapidly. If  $\Delta$

is small, the Hamiltonian becomes equivalent to the Heisenberg spin operator. If not, we may need more terms and the Heisenberg model will become somewhat less accurate, but it will still be possible to obtain a satisfactory solution by the use of the effective Hamiltonian (4).

#### ACKNOWLEDGMENTS

Part of this work has been carried out during the author's stay at Uppsala University, Sweden. He would like to thank Professor P. O. Löwdin for his hospitality and many useful discussions. He is also grateful to Dr. C. Herring for his valuable comments.

#### APPENDIX A

Here we shall show that all the energy integrals  $H_P$  defined by Eq. (22) are bounded. For simplicity, we assume the one particle approximation where  $\psi$  is given by Eq. (11) and we write  $\psi$  as

$$\psi(1, 2, \dots, N) = \psi_1(1) \psi_2(2) \cdots \psi_N(N).$$

The following proof will be easily extended to the more general case where  $\psi$  is given by Eq. (12) as long as  $\psi$  is normalized.

By the definition of  $E_0$ , it is found that

$$H_E = 0, \\ H_{(12)} = \langle 2 | \mathcal{H}(1) | 1 \rangle \langle 1 | 2 \rangle + \langle 1 | \mathcal{H}(2) | 2 \rangle \langle 2 | 1 \rangle + \langle 21 | V_{12} | 12 \rangle, \quad (A1)$$

where (12) is the permutation of orbitals 1 and 2. The definition of the other notations will be found in Eq. (A3). In general, we denote by  $(1, 2, \dots, M)$  the permutation of  $M$  electrons, by which orbitals 1, 2,  $\dots$ ,  $M$  are replaced by  $P_1, P_2, \dots, P_M$ , respectively, such that  $P_i \neq i$  for  $i = 1, 2, \dots, M$ , while electrons  $M+1, \dots, N$  are not permuted. Then we find that

$$H_{(1, 2, \dots, M)} = \sum_{i=1}^M \langle P_i | \mathcal{H}(i) | i \rangle \prod_{\substack{k=1 \\ (k \neq i)}}^M \langle P_k | k \rangle + \frac{1}{2} \sum_{i=1}^M \sum_{\substack{j=1 \\ (j \neq i)}}^M \langle P_i P_j | V_{ij} | ij \rangle \prod_{\substack{k=1 \\ (k \neq i \text{ or } j)}}^M \langle P_k | k \rangle, \quad (A2)$$

where

$$\langle P_i | \mathcal{H}(i) | i \rangle = \langle P_i | \mathcal{H}'(i) | i \rangle + \sum_{l=M+1}^N \langle P_{il} | V_{il} | il \rangle, \\ \mathcal{H}'(i) = \mathcal{H}_0(i) - \langle i | \mathcal{H}_0(i) | i \rangle, \\ \mathcal{H}_0(i) = \frac{P_i^2}{2m} - e^2 \sum_{A=1}^{N'} \frac{Z_A}{r_{Ai}}, \quad (A3)$$

$$V_{ij} = \frac{e^2}{r_{ij}} - e^2 \left\langle ij \left| \frac{1}{r_{ij}} \right| ij \right\rangle,$$

and

$$\left\langle P_i P_j \left| \frac{1}{r_{ij}} \right| ij \right\rangle = \int \psi_{P_i}^*(i) \psi_{P_j}^*(j) \frac{1}{r_{ij}} \psi_i(i) \psi_j(j) d\tau.$$

<sup>15</sup> Here we have applied the same argument as the one used in Appendix B. Since the first  $g \times g$  parts and the diagonal elements of  $F(i)$  and  $E(i)$  can have nonvanishing elements and the rest of them vanish identically.

Here  $\langle P_i P_j | V_{ij} | ij \rangle$  decreases in proportion to  $e^{-R}$  when the distance  $R$  between the centers of the electrons  $i$  and  $j$  increases, while the number of electrons  $j$  localized between the distances  $R$  and  $R + \Delta R$  from the center electron  $i$  will be proportional to  $R^2$ . This means that both  $\langle P_i | \mathcal{H}(i) | i \rangle$  and

$$\sum_{j=1}^M \sum_{(j \neq i)} \langle P_i P_j | V_{ij} | ij \rangle$$

should converge to finite values even if  $N \rightarrow \infty$ . We shall denote the maximum value by  $I$ . The overlap integrals  $\langle k | l \rangle$  are smaller than 1, since the atomic orbitals are normalized. We denote the maximum value by  $S$ . From Eq. (A2), therefore, we find that

$$H_{(1,2 \dots M)} \leq \{H_{(1,2 \dots M)}\}_{\max} = MI(S + \frac{1}{2})S^{M-2},$$

and also

$$\frac{\{H_{(1,2 \dots M+1)}\}_{\max}}{\{H_{(1,2 \dots M)}\}_{\max}} = \frac{M+1}{M} S, \quad S < 1.$$

For  $M=2$ ,  $H_{(12)}$  is clearly finite as shown in Eq. (A1). Since  $S \approx 0.1 \sim 0.2$ , the maximum value of  $H_P$  for  $M > 2$  decreases rapidly and vanishes in practice. This proves that all the energy integrals  $H_P$  are bounded.

#### APPENDIX B

In this Appendix, we shall show that, even if  $N!$  functions  $P\psi$  are not linearly independent,  $D^{-1}$  and the expansion (3) or (28) exist and therefore assumption (c) used in Sec. II is not necessary. For simplicity, we consider the case where  $\psi(1,2 \dots N)$  is symmetric with respect to exchange of two electrons in  $r$  pairs (1,2), (3,4),  $\dots$ ,  $(2r-1, 2r)$ .

We denote by  $h$  the subgroup of  $\sigma_N$ , consisting of  $2^r$  elements

$$Q = (1,2)^{\lambda_1} (3,4)^{\lambda_2} \dots (2r-1, 2r)^{\lambda_r},$$

where  $\lambda_1, \lambda_2, \dots, \lambda_r$  can take the value 0 and 1 independently of other  $\lambda$ 's. Then

$$Q\psi = \psi.$$

If we select  $s = N!/2^r$  elements  $R_1, R_2, \dots, R_s$  of the  $N!$  elements of  $\sigma_N$  in such a way that  $R_u^{-1}R_v \notin h$  for any pair of  $R_u$  and  $R_v$ , all  $N!$  elements of  $\sigma_N$  can be obtained in the form  $RQ$  where  $Q \in h$ . We assume that all the  $S$  functions  $R\psi$  are linearly independent. The overlap matrix  $D$  will then be given by

$$D = \sum_R \sum_Q S_{RQ} \tilde{U}(RQ) = \sum_R \sum_Q S_R \tilde{U}(RQ), \quad (\text{A4})$$

because

$$S_{RQ} = \int RQ\psi^* \psi d\tau = \int R\psi^* \psi d\tau = S_R.$$

Since  $Q_i Q_j = Q_j Q_i$  and  $Q^2 = E$ , all the irreducible representations of  $h$  are one-dimensional and the matrices

are either (1) or  $(-1)$ . Then it is possible to make the matrices of the representation  $\mathfrak{D}_S$  of  $\sigma_N$  in such a way that the matrices for elements of the subgroup  $h$  appear in diagonal form:

$$U_{kh}(Q) = \pm \delta_{kh} \quad \text{for } Q \in h.$$

Here we shall number the rows and columns of the matrices such that the first  $g$  diagonal elements are  $+1$  for all the  $2^r$  matrices  $U(Q)$ .

$$U_{kk}(Q) = 1, \quad 1 \leq k \leq g, \quad Q \in h \quad (\text{A5})$$

while the other diagonal elements  $U_{kk}(Q)$ ,  $g < k \leq f_s$  take the value  $-1$  for at least one of the elements  $Q$ .

*Lemma 5.* The first  $g$  members of  $\Phi$  with given  $S$  are linearly independent and the rest of them vanish identically.

To prove this lemma, we shall consider a possible linear combination among  $\Phi_m$  ( $m=1, 2, \dots, g$ ):

$$0 = \sum_{m=1}^g C_m \Phi_m.$$

Since the spin functions  $\Theta$  are linearly independent, we find that

$$0 = \sum_m \sum_P C_m U_{km}(P)^* P\psi = \sum_R \{ \sum_m \sum_Q C_m U_{km}(RQ)^* \} R\psi.$$

As the  $s$  functions  $R\psi$  are assumed to be linearly independent, the coefficients must vanish;

$$0 = \sum_m \sum_Q C_m U_{km}(RQ)^*,$$

where  $U_{km}(RQ)^* = U_{km}(R)^*$  because of the structure of  $U(Q)$  given by Eq. (A5). Therefore the relation

$$0 = \sum_m C_m U_{km}(R)^*$$

holds for all  $N!$  permutations of  $\sigma_N$ . Multiplying this by  $\tilde{U}_{kn}(P)$  and summing over  $P$ , we obtain

$$0 = \sum_m C_m \sum_P \tilde{U}_{kn}(P) U_{km}(P)^* = \sum_m C_m \{ (N! / f_S^N) \delta_{mn} \} \quad \text{for } n=1, 2, \dots, g.$$

This implies that  $C_m = 0$ , for  $m=1, 2, \dots, g$ , and the  $g$  functions  $\Phi_m$  ( $m=1, 2, \dots, g$ ) are linearly independent.

If  $m > g$ , there exists at least one  $Q_0$  of  $h$  for which  $U_{kk}(Q_0) = -1$ . Then we find that

$$\begin{aligned} \Phi_m &= \frac{1}{\sqrt{(N!)}} \sum_k \Theta_k \sum_P U_{km}(P)^* P\psi \\ &= \frac{1}{\sqrt{(N!)}} \sum_k \Theta_k \sum_P \frac{1}{2} \{ U_{km}(P)^* P\psi + U_{km}(PQ_0)^* PQ_0\psi \}, \\ &= 0, \end{aligned}$$

since  $U_{km}(PQ_0)^* = \sum_j U_{kj}(P)^* U_{jm}(Q_0)^* = -U_{km}(P)^*$ .

From Lemma 5, it is evident that with the exception of the first  $g \times g$  part, all the elements of  $D$  vanish. It is



still possible to assume that, since  $D^\dagger = D$ ,  $D$  is diagonal. As  $g$  functions  $\Phi_i$  ( $i=1, 2, \dots, g$ ) are linearly independent, there exists an  $\epsilon$  such that

$$D_{ii} \geq \epsilon > 0, \quad \text{for } i=1, 2, \dots, g,$$

while

$$D_{ii} = 0, \quad \text{for } i > g.$$

We define  $D^{-1}$  as the diagonal matrix, in which the first  $g$  elements are given by  $D_{ii}^{-1}$  ( $i=1, 2, \dots, g$ ), respectively, and the rest of the elements vanish.

If we assume that  $D^{-1}$  can be expanded as

$$D^{-1} = \sum_P T_P \tilde{U}(P) = \sum_R \sum_Q T_{RQ} \tilde{U}(RQ),$$

then we find that  $T_{RQ} = T_R$  for all  $2^r$  permutations  $Q$ , since the first  $g \times g$  parts of  $\tilde{U}(RQ)$  are invariant with respect to  $Q$  while the rest of the matrices vanish. The coefficients  $T_R$  should satisfy the relation

$$\sum_R T_R S_{R^{-1}R'} = \delta_{RR'} \quad \text{for } R' = 1, 2, \dots, s, \quad (\text{A6})$$

instead of Eq. (25). Here  $s$  functions  $R\psi$  are assumed to be linearly independent and the  $g \times g$  determinant  $|S|$  does not vanish. Therefore, we find that  $T_R$  are all bounded and determined uniquely, by Eq. (A6).

Likewise,  $D^{-\frac{1}{2}}$  can be expanded in terms of  $\tilde{U}(P)$  and therefore  $J$  can be calculated uniquely, as we have proved in Sec. II.

In case  $\psi(1, 2, \dots, N)$  is antisymmetric with respect to the permutations  $Q$ , the diagonal elements of the invariant parts of the matrices  $U(Q)$ , from which the nonvanishing wave functions are made up, have the value  $-1$ . The more general case can also be treated in a similar way.

### APPENDIX C

The expansion of  $D^{-\frac{1}{2}} \tilde{U}(R) D^{\frac{1}{2}}$  given by Eq. (68) will be the subject of this Appendix. First, we consider a commutation relation between  $D$  and  $\tilde{U}(R)$ . Since  $D$  is expanded in terms of  $\tilde{U}(P)$  as shown in Eq. (23), we find that

$$\tilde{U}(R)D = \sum_P S_P \tilde{U}(PR) = \sum_P S_P \tilde{U}(R^{-1}PR) \cdot \tilde{U}(R).$$

By introducing the notation

$$X(R) = \sum_P S_P \tilde{U}(R^{-1}PR) - D,$$

the commutation relation is written as

$$\tilde{U}(R)D = (D + X(R))\tilde{U}(R). \quad (\text{A7})$$

The matrix  $X(R)$  for any permutation  $R$  can be calculated similar to  $E_h$ . As an example, we shall consider the case where  $R = (hl)$ . The  $N!$  permutations  $P$  are divided as follows.

$$P = Q[hl]P[km].$$

The matrix  $X(R)$  is then written as

$$\begin{aligned} X(R) &= \sum_P S_{RP} \tilde{U}(P) - \sum_P S_P \tilde{U}(P) \\ &= \sum_{P'} \sum_Q (S_{RQP'R^{-1}} - S_{QP'}) \tilde{U}(QP') \\ &= \sum_m'' \langle m|h \rangle S[hm|hl]_{hl} - \sum_m'' \langle m|l \rangle S[hm|hl] \\ &\quad + \sum_k'' \langle k|l \rangle S[kl|hl]_{hl} - \sum_k'' \langle k|h \rangle S[kl|hl] \\ &\quad + \sum_k'' \sum_m''' \langle km|lh \rangle S[km|hl]_{hl} \\ &\quad - \sum_k'' \sum_m''' \langle km|hl \rangle S[km|hl], \end{aligned}$$

where  $P'$  is an abbreviation of  $P(km \leftarrow hl)$  and both  $\sum_k''$  and  $\sum_m''$  exclude  $h$  and  $l$ , while  $\sum_m'''$  excludes  $h$ ,  $l$ , and  $k$ . In  $S[km|hl]_{hl}$ , orbitals  $\psi_h$  and  $\psi_l$  are interchanged. Since the overlap matrices  $S$  can be expanded such that

$$\sum_m'' \langle m|h \rangle S[hm|hl]_{hl} = \sum_m'' \sum_r'' \langle mr|hm \rangle S[hmr|hlm]_{hl},$$

the matrix  $M = D^{-\frac{1}{2}} X(R) D^{-\frac{1}{2}}$  is given by

$$\begin{aligned} M &= D^{-\frac{1}{2}} X(R) D^{-\frac{1}{2}} \\ &= \sum_m'' \sum_r'' \langle mr|hm \rangle D^{-\frac{1}{2}} S[hmr|hlm]_{hl} D^{-\frac{1}{2}} \\ &\quad + \sum_k'' \sum_t'' \langle kt|lk \rangle D^{-\frac{1}{2}} S[klt|hlk]_{hl} D^{-\frac{1}{2}} \\ &\quad + \sum_k'' \sum_m'' \sum_t'' \sum_r''' \langle kmtr|lhkm \rangle \\ &\quad \quad \quad \times D^{-\frac{1}{2}} S[kmtr|hllkm]_{hl} D^{-\frac{1}{2}} \\ &\quad - \sum_m'' \sum_r'' \langle mr|lm \rangle D^{-\frac{1}{2}} S[hmr|hlm] D^{-\frac{1}{2}} \\ &\quad - \sum_k'' \sum_t'' \langle kt|hk \rangle D^{-\frac{1}{2}} S[klt|hlk] D^{-\frac{1}{2}} \\ &\quad - \sum_k'' \sum_m'' \sum_t'' \sum_r''' \langle kmtr|hlkm \rangle \\ &\quad \quad \quad \times D^{-\frac{1}{2}} S[kmtr|hllkm] D^{-\frac{1}{2}}. \quad (\text{A8}) \end{aligned}$$

We shall evaluate the maximum values of the diagonal elements of matrices such as

$$D^{-\frac{1}{2}} S[hmtr|hlm] D^{-\frac{1}{2}} = \int \{ \Phi[hmtr] D^{-\frac{1}{2}} \}^\dagger \{ \Phi[hlm] D^{-\frac{1}{2}} \} d\tau.$$

By applying Lemma 2 and the Schwarz inequality relation, we obtain the following equation:

$$\begin{aligned} &| \{ D^{-\frac{1}{2}} S[k_1 k_2 \dots k_n | k_1' k_2' \dots k_n'] D^{-\frac{1}{2}} \}_{ii} | \\ &\quad \leq \{ D^{-\frac{1}{2}} S[k_1 k_2 \dots k_n | k_1 k_2 \dots k_n] D^{-\frac{1}{2}} \}_{ii}^{\frac{1}{2}} \\ &\quad \times \{ D^{-\frac{1}{2}} S[k_1' k_2' \dots k_n' | k_1' k_2' \dots k_n'] D^{-\frac{1}{2}} \}_{ii}^{\frac{1}{2}} \\ &\quad \leq n!_{\max}, \quad (\text{A9}) \end{aligned}$$

where by  $\{X\}_{ii}$  we denote the  $ii$  element of matrix  $X$

and by  ${}^n t_{\max}$  the maximum value of the diagonal elements of  $D^{-\frac{1}{2}}S[k_1 k_2 \cdots k_n | k_1 k_2 \cdots k_n]D^{-\frac{1}{2}}$  for arbitrary  $k_1 k_2 \cdots k_n$ . Since Eq. (A9) is similar to Eqs. (48) and (50), the discussion in Sec. III.2 can be applied to the present case and we find that  ${}^n t_{\max} = {}^n l_{\max}$ . This means that

$$\begin{aligned} {}^1 t_{\max} &< (1-\Delta)^{-1} \\ {}^n t_{\max} &\leq ({}^{n-1} t_{\max})(1-\Delta_{h_1 h_2 \cdots h_{n-1}})^{-1} \end{aligned} \quad (\text{A10})$$

where  $\Delta_{h_1 h_2 \cdots h_{n-1}} = \sum_t t^{(n-1)} |\langle t | h_n \rangle|$  and the summation  $\sum_t t^{(n-1)}$  excludes  $t = h_1 h_2 \cdots h_{n-1}$ .

From Eqs. (A8)~(A10), we obtain that

$$M_{\max} \leq 2\Delta_h^2 (1-\Delta_h)^{-2} (1-\Delta)^{-1} + 2\Delta_l^2 (1-\Delta_l)^{-2} (1-\Delta)^{-1} + 2\Delta^2 \Delta_l \Delta_h (1-\Delta)^{-4}, \quad (\text{A11})$$

where  $M_{\max}$  is the maximum value of the diagonal elements  $|M_{ii}|$ , that is,

$$|M_{ii}| \leq M_{\max}. \quad (\text{A12})$$

Now we shall consider a commutation relation between  $D^{\frac{1}{2}}$  and  $U(R)$ . Let us write the relation as

$$\tilde{U}(R)D^{\frac{1}{2}} = [D^{\frac{1}{2}} + Y(R)]\tilde{U}(R). \quad (\text{A13})$$

By using the matrix  $Y(R)$ ,  $D^{-\frac{1}{2}}\tilde{U}(R)D^{\frac{1}{2}}$  is written as

$$D^{-\frac{1}{2}}\tilde{U}(R)D^{\frac{1}{2}} = [1 + D^{-\frac{1}{2}}Y(R)]\tilde{U}(R).$$

Comparing the relation with Eq. (68), we find that

$$N = D^{-\frac{1}{2}}Y(R) = \sum_P a(R)_P \tilde{U}(P).$$

We shall prove in the following that  $\sum_P |a(R)_P|^2$  is smaller than  $3M_{\max}$  and therefore the dominant term in  $D^{-\frac{1}{2}}\tilde{U}(R)D^{\frac{1}{2}}$  is  $\tilde{U}(R)$ .

From Eqs. (A7) and (13), we obtain that

$$D + X(R) = [D^{\frac{1}{2}} + Y(R)][D^{\frac{1}{2}} + Y(R)^\dagger],$$

or, by the use of  $M$  and  $N$ , it can be written as

$$1 + M = (1 + N)(1 + N^\dagger), \quad (\text{A14})$$

since  $Y(R)^\dagger = Y(R)$ . By taking the  $ii$  element of the matrix equation (A14), we find that

$$M_{ii} = 2N_{ii} + \sum_k |N_{ik}|^2, \quad (\text{A15})$$

where  $N_{ii}^* = N_{ii}$ . This means that

$$M_{\max} \geq 2N_{ii}, \quad \text{for any } i. \quad (\text{A16})$$

The lower bound of  $N_{ii}$  is obtained as follows. From the definition of  $M$  and Eq. (A7),

$$M = D^{-\frac{1}{2}}\tilde{U}(R)D\tilde{U}(R)^\dagger D^{-\frac{1}{2}} - 1.$$

The  $ii$  element is given by

$$\begin{aligned} M_{ii} &= \sum_k |\tilde{U}_{ik}|^2 (D_{kk} D_{ii}^{-1} - 1) \\ &= \sum_k |\tilde{U}_{ik}|^2 (D_{kk}^{\frac{1}{2}} D_{ii}^{-\frac{1}{2}} - 1)(D_{kk}^{\frac{1}{2}} D_{ii}^{-\frac{1}{2}} + 1). \end{aligned}$$

Similarly,

$$N_{ii} = \sum_k |\tilde{U}_{ik}|^2 (D_{kk}^{\frac{1}{2}} D_{ii}^{-\frac{1}{2}} - 1).$$

Let us assume that  $N_{jj}$  is the minimum of the diagonal elements  $N_{ii}$  and  $D_{ll}$  the maximum of  $D_{ii}$ . We denote  $N_{jj}$  and  $D_{ll}$  by  $N_{\min}$  and  $D_{\max}$ . Then

$$N_{\min} = N_{jj} \geq \sum_k |\tilde{U}_{jk}|^2 (D_{kk}^{\frac{1}{2}} D_{\max}^{-\frac{1}{2}} - 1), \quad (\text{A17})$$

since  $(D_{kk}^{\frac{1}{2}} D_{\max}^{-\frac{1}{2}} - 1) \leq (D_{kk}^{\frac{1}{2}} D_{jj}^{-\frac{1}{2}} - 1)$ . We also find that

$$-M_{\max} \leq \sum_k |\tilde{U}_{lk}|^2 (D_{kk}^{\frac{1}{2}} D_{\max}^{-\frac{1}{2}} - 1)(D_{kk}^{\frac{1}{2}} D_{\max}^{-\frac{1}{2}} + 1).$$

From the definition of  $M_{\max}$  given by Eq. (A11), it is evident that the above relation is valid regardless of the structure of the representation matrix  $U$ . This means that this relation is valid, even if we interchange the  $j$ th and  $l$ th rows of  $U$ . Then it follows that

$$\begin{aligned} -M_{\max} &\leq \sum_k |\tilde{U}_{jk}|^2 (D_{kk}^{\frac{1}{2}} D_{\max}^{-\frac{1}{2}} - 1) \\ &\quad \times (D_{kk}^{\frac{1}{2}} D_{\max}^{-\frac{1}{2}} + 1). \end{aligned} \quad (\text{A18})$$

Comparing Eq. (A18) with Eq. (A17), we find that

$$N_{\min} > -M_{\max}, \quad (\text{A19})$$

since  $1 < D_{kk}^{\frac{1}{2}} D_{\max}^{-\frac{1}{2}} + 1 < 2$ .

From Eqs. (A15), (A16), and (A19), the following relation is found:

$$\sum_k |N_{ik}|^2 = M_{ii} - 2N_{ii} \leq 3M_{\max}.$$

By applying Lemma 3, therefore, we obtain that

$$\begin{aligned} \sum_P |a(R)_P|^2 &= \sum_\rho \sum_m \sum_k (f_\rho/N!) |N_{mk}^\rho|^2 \\ &\leq \sum_\rho \sum_m (f_\rho/N!) \cdot 3M_{\max}. \end{aligned}$$

Since  $\sum_\rho \sum_m (f_\rho/N!) = 1$ , we find the relation

$$\sum_P |a(R)_P|^2 \leq 3M_{\max}, \quad (\text{A20})$$

and also,

$$|a(R)_E| \leq (3M_{\max})^{\frac{1}{2}}, \quad (\text{A21})$$

where  $M_{\max}$  is given by Eq. (A11).