DIAGRAMMATICAL METHODS FOR MANY-FERMION SYSTEMS

J. Paldus

1. Introduction

2. Basic theory (time-independent formalism)

   2.1. Second quantization representation
   2.2. Wick's theorem
   2.3. Representation of operators
   2.4. Hole-particle formalism
   2.5. Diagrammatical representation
      2.5.1. Basic concepts
      2.5.2. General rules
      2.5.3. Spin-independent operators

3. Applications

   3.1. Matrix elements of 1- and 2-particle operators
   3.2. Time-independent RS perturbation theory
   *3.3. Stability of HF Solutions
   3.4. Coupled cluster approaches
   *3.5. Green-function approaches
   *3.6. Open-shells

*4. Spin-adapted graphical techniques

   4.1. Angular momentum graphs
   4.2. Relationship between orbital and spin graphs
   4.3. Evaluation of spin-adapted quantities

---

*These topics are optional

Preface

These lecture notes have been prepared during my lecture series delivered at the Institute of Theoretical Chemistry of the Catholic University in Nijmegen during the Spring of 1981. In these lectures we succeeded to cover all the basic material as indicated in the attached outline. Unfortunately, we were unable to follow this material with the exploitation of graphical techniques of spin algebras leading to a general spin-adapted formalism (sections indicated by an asterisk in the original outline). The principal reasons for this omission was the lack of time and the fact that only part of the audience was familiar with angular momentum graphical methods from the lecture series given earlier by Dr. Wormer. Since I hope that the spin-adapted version of the diagrammatic formalism can be filled in at some future date, the appropriate Chapter was included in the Contents, but is absent in the body of the text. Of course, the corresponding sections in the Applications are also missing.

We have also decided to attach to these notes a number of solved examples which date back to my lecture series given at the University of Louis Pasteur in Strasbourg. The notational conventions used in these examples do not always conform with those used in the present notes, but the reader should have no difficulties in following them. These examples give also an ample illustration of the use of the diagrammatic method for the evaluation of matrix elements of one- and two-particle operators between spin-orbital N-electron states (Examples 16 and 18) so that we did not repeat them in Chapter VII. A brief summary of the developments in Chapter IX on the many-body perturbation theory follows these examples.

Finally, we decided to attach to these notes a recent preprint concerning the relationship between variational and coupled cluster approaches, which nicely complements Chapter X on the coupled cluster approach.

It is hoped that these notes will be of some help to those who faithfully followed my lectures as well as to those who for various reasons were unable to join us. The author would very much welcome any comments, remarks or criticism (even a vitriolic one) which would help him to prepare a better set of notes when the opportunity arises again.

It has been a great pleasure for me to spend the past academic term visiting the Institute of Theoretical Chemistry in Nijmegen and to have the opportunity to interact with its members. I would like to use this occasion to express my sincere gratitude to the Catholic University in Nijmegen for giving me this opportunity and to the members of the Institute for their very kind reception and stimulating milieu. I am particularly indebted to the head of the Institute, Professor Ad van der Avoird, for his kind hospitality which made our stay so enjoyable and to my former collaborator and friend, Paul Wormer, as well as to his wife Jenny, for their tireless help making our move to Nijmegen so smooth and pleasant as well as for all their kindness and friendship making our stay in Holland the most wonderful experience.

Nijmegen, June 30, 1981.

J.P.

Contents

A. Theory

1. Second quantization formalism 1
- 1.1. N-particle spaces, basic concepts and notation 1'
- 1.2. Occupation number representation 4'
- 1.3. Creation and annihilation operators 6
- 1.4. Number operator 8
- 1.5. Fock space 9
- 1.6. Anti-commutation relations 11
- 1.7. Examples 13
- 1.8. Canonical basis transformation 16'
- 1.9. Field operators 18

2. Wick's theorem 21
- 2.1. Normal product of operators 21
- 2.2. Contraction of operators 22
- 2.3. Basic rules for n-products and contractions 23'
- 2.4. Wick's theorem 25
- 2.5. Generalized Wick's theorem 29
- 2.6. Vacuum mean values of operator products 30

3. Representation of operators 32
- 3.1. Coordinate representation of operators 32
- 3.2. Occupation number representation of operators 34
  - 3.2.1. One-particle operators 36
  - 3.2.2. Two-particle operators 37'
- 3.3. Operator representation and Wicks theorem 40'
- 3.4. Spin-independent operators 43
- 3.5. Canonical transformation of operators 44

4. Hole-particle formalism 45'
- 4.1. Hole-particle operators 47
- 4.2. Anti-commutation relations 48
- 4.3. Normal products and contractions 48'
- 4.4. Wicks theorem and examples 51
- 4.5. Normal product form of operators 55

5. Diagrammatic representation 58
5.1. Basic concepts 58
5.2. Examples 61'
5.2.1. Hamiltonian matrix elements between monoexcited configurations 61'
5.2.2. Action of a one-particle operator on mono-excited configurations 67
5.3. Skeletons and weights
5.4. Graphical representation of operators 71'
5.5. General rules for resulting diagrams 73'
Appendix A: Proof of Thm. 3. 75a
5.6. Standard and normal product operator forms 77
5.7. Hugenholtz and Brandow diagrams 79'
5.8. Spin-independent operators 82
5.9. Summary of rules 83

6. Spin-adapted diagrammatic formalism*

B. Applications

7. Evaluation of Hamiltonian matrix elements

8. Independent particle model approximation HF 1
8.1. Exponential form of the IPM wavefunction HF2'
8.2. Hartree-Fock Equations HF4
8.3. Stability conditions HF6
8.4 Brueckner orbitals HF8

9. Perturbation Theory
9.1. Rayleigh-Schrödinger perturbation formula P1
9.2. Renormalization terms and bracketing technique P3
9.2.1. Eigenvalue expansion P3'
9.2.2. Eigenfunction expansion P7
9.3. Ground State Perturbation Theory P8
9.4. Diagrammatic representation P11
9.5. Perturbation theory diagrams P13
9.5.1. Second and third order energy contributions P15'
9.5.2. First and second order wavefunction contributions P19'
9.6. Linked and unlinked diagrams P20
9.7. Exclusion Principle Violating Diagrams P23
9.8. Factorization lemma P24'
9.9. Linked Cluster Theorem P28
9.10. Connected cluster structure of the wavefunction P30
9.11. Hartree-Fock diagrams P32
9.12. Spin-independent formalism P35
9.13 Summary of rules for perturbation theory diagrams P37

10. Coupled Cluster Approach
10.1. Coupled Cluster Ansatz CC 2
10.2. Diagrammatic representation CC 3'
10.3. Connected and disconnected clusters CC 5'
10.4. Connected Cluster Theorem CC 6'
10.5. Spin-orbital form of CCA Equations CC 11
10.5.1. CPMET Equations (Spinorbital form) CC 13
10.5.2. ECPMET Equations (Spinorbital form) CC 16'
10.6. Relationship with other approaches CC 19'
10.6.1. Variational approach (CI) CC 19'
10.6.2. Perturbative approach CC 25'

C. Appendices.

1. Solved examples.
2. Relationship between configuration interaction and coupled cluster approaches.
3. Summary of definitions and formulas for Perturbation Theory.

1. Second quantization formalism

The second quantization formalism enables an alternative representation of quantum mechanical states and observables, which is quite distinct from the well known Schrödinger wavefunction or Heisenberg matrix formalisms. The name of this formalism, which seems to imply some physical procedure analogous to a transition from the classical to quantum mechanics (sometimes refered to as the "first quantization") is sometimes confusing to the uninitiated. We thus haste to say that the term "second quantization" originated in early attempts of a quantum mechanical description of electromagnetic fields and is nowadays used for historical reasons to designate a quantum mechanical representation of states and observables, which is much more appropriately refered to as the "occupation number representation".

The primary advantages of this representation lie in the fact that it automatically yields the states of correct symmetry for the systems of identical particles and thus enables one to bypass various combinatorial considerations which are unavoidable in conventional formalisms. Moreover, it enables a simple graphical representation of its basic quantities and a formulation of simple rules simplifying many quantum mechanical calculations.

In this chapter we shall introduce the concepts of the occupation number representation and of the basic operators used in the second quantization formalism, namely the creation and annihilation operators. We shall start from the well known configuration wavefunction representation of N-particle states and

will introduce the necessary operators purely heuristically without paying much attention to the mathematical rigor. Only after this initial outline, which should clarify our basic objectives, we shall introduce the concept of the Fock space and will make our definitions more satisfactory from the mathematical viewpoint. We shall derive the basic properties of creation and annihilation operators in a simplest way possible and will also briefly introduce their field counterparts for the sake of completeness.

1.1. N-particle spaces; basic concepts and notation

In this text we are solely dealing with the systems of identical fermions and we shall assume that the reader is familiar with basic concepts of quantum mechanical description of these systems in the usual wavefunction or abstract Dirac formalism.

Consider, thus, a general complete system of discrete orthonormal one-particle states $|A\rangle$, where $\langle A|B\rangle = \delta_{AB}$, and where δ_{AB} is the Kronecker symbol ($\delta_{AA}=1$, $\delta_{AB}=0$ for $A \neq B$). In the coordinate representation the abstract one-particle states $|A\rangle$ have the form

$$\varphi_A(\vec{x}) \equiv \langle \vec{x}|A\rangle , \tag{1.1}$$

where $\vec{x} \equiv (\vec{r}, s)$ designates the spatial ($\vec{r}$) and internal (s) coordinates. When the considered particle is an electron, $\vec{r}$ designates its three-dimensional position vector and s its spin variable. The one-electron wavefunction $\varphi_A(\vec{x})$ or the corresponding abstract ket $|A\rangle$ are refered to as a <u>spin-orbital</u>. We shall always label the spin-orbitals by the capital letters of the Latin alphabet reserving the lower case letters for corresponding spatial components, referred to as <u>orbitals</u>.

The wavefunction and the abstract ket representations are simply related as Eq. (1) indicates. The orthonormality conditions for the spinorbitals (1) is easily obtained using the resolution of identity for the one-particle coordinate space $E_3 \otimes S$, $\vec{r} \in E_3$, $s \in S$, $\vec{x} \equiv (\vec{r}, s)$

$$\hat{1} = \int d\vec{x}\ |\vec{x}\rangle\langle\vec{x}| , \tag{1.2}$$

where $\hat{1}$ designates the identity operator in $E_3 \otimes S$ and the integral symbol implies the integration over the three dimensional Euclidean space and the summation over the discrete (usually spin) variable s.

Thus

$$\langle A|B\rangle = \int d\vec{x}\,\langle A|\vec{x}\rangle\langle\vec{x}|B\rangle = \int d\vec{x}\,\varphi_A^*(\vec{x})\varphi_B(\vec{x}) = \delta_{AB}, \tag{1.3}$$

where $M^* = \langle A|\vec{x}\rangle$ designates the complex conjugate of $M = \langle\vec{x}|A\rangle$.

Any single particle state $\psi(\vec{x}) \equiv \langle\vec{x}|\psi\rangle$ can be expanded as

$$\psi(\vec{x}) = \sum_A c_A \varphi_A(\vec{x}) \tag{1.4a}$$

or

$$|\psi\rangle = \sum_A c_A |A\rangle , \tag{1.4b}$$

where the summation generally extends over all states of a given complete orthonormal system of discrete states. The generalized Fourier coefficients c_A are given as

$$c_A = \langle A|\psi\rangle = \int d\vec{x}\,\varphi_A^*(\vec{x})\,\psi(\vec{x}) \tag{1.5}$$

where we have used again the resolution of identity, Eq. (2).

A complete orthonormal system $\{|A\rangle\}$ or

$\{\varphi_A(\vec{x})\}$ may thus be regarded as a basis for a one-particle space $\mathcal{V}_1$ assuming some arbitrary but fixed ordering of these states.*)

Consider, next, a system of N identical fermions, say, a system of N electrons. Its state space $\mathcal{F}_N$ can be described as an antisymmetric component of the N-th tensor power of one-particle Hilbert spaces $\mathcal{V}_1$,

$$\mathcal{F}_N = \mathcal{A}\,\mathcal{V}_N , \tag{1.6}$$

where

$$\mathcal{V}_N = \underbrace{\mathcal{V}_1 \otimes \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_1}_{N\text{-times}} = \mathcal{V}_1^{\otimes N} \tag{1.7}$$

and $\mathcal{A}$ designates the projector onto the totally antisymmetric subspace of $\mathcal{V}_N$. Choosing a canonical tensor product basis for $\mathcal{V}_N$ consisting of monomials

$$\varphi_{A_1}(\vec{x}_1) \otimes \varphi_{A_2}(\vec{x}_2) \otimes \cdots \otimes \varphi_{A_N}(\vec{x}_N), \tag{1.8a}$$

or

$$|A_1\rangle \otimes |A_2\rangle \otimes \cdots \otimes |A_N\rangle , \tag{1.8b}$$

which we write for simplicity as

$$\varphi_{A_1}(\vec{x}_1)\,\varphi_{A_2}(\vec{x}_2) \cdots \varphi_{A_N}(\vec{x}_N) \tag{1.9a}$$

or

$$|A_1\rangle\,|A_2\rangle \cdots |A_N\rangle \equiv |A_1 A_2 \ldots A_N\rangle , \tag{1.9b}$$

we can write a general normalized basis vector for $\mathcal{F}_N$ as

$$\Phi_K(1,2,\ldots,N) \equiv \sqrt{N!}\;\mathcal{A}\,\varphi_{A_1}(\vec{x}_1) \cdots \varphi_{A_N}(\vec{x}_N) . \tag{1.10}$$

Here K designates the index set

$$K \equiv \{A_1, A_2, \ldots, A_N\} , \tag{1.11}$$

*) Clearly, the states $|A\rangle$ and $\varphi_A(\vec{x})$ belong to very different vector spaces. However, these spaces are simply related by a natural isomorphism (see Eq. (1)) and we thus use the same symbol $\mathcal{V}_1$ for both. Same convention is used for corresponding N-particle spaces.

and the antisymmetrizer $\mathcal{A}$ has the form

$$\mathcal{A} = (N!)^{-1} \sum_{P \in \mathfrak{S}_N} (-1)^p\, P , \tag{1.12}$$

where the sum extends over N! permutations P of the symmetric group $\mathfrak{S}_N$. The permutations P,

$$P \equiv \begin{pmatrix} 1 & 2 & \cdots & N \\ p_1 & p_2 & \cdots & p_N \end{pmatrix}, \tag{1.13}$$

whose parity is designated as $(-1)^p$, act either on the spin-orbital labels A_i or on the coordinates $\vec{x}_i$ $(i=1,\ldots,N)$ but not on both simultaneously.

In the abstract bra-ket notation we shall write the state (10) as

$$|\Phi_K\rangle \equiv |\Phi_K(1,2,\ldots,N)\rangle \equiv |\{A_1 A_2 \ldots A_N\}\rangle$$
$$= \sqrt{N!}\;\mathcal{A}\,|A_1 A_2 \ldots A_N\rangle . \tag{1.14}$$

We shall refer to the states (10) or (14) as <u>configurations</u>. The single particle states appearing in K are said to be <u>occupied</u> in this configuration while those not present in K as <u>unoccupied</u> spin-orbitals relative to $|\Phi_K\rangle$. In order that $|\Phi_K\rangle$ or Φ_K form a basis for $\mathcal{F}_N$, only linearly independent <u>ordered configurations K</u>, for which

$$A_1 < A_2 < \cdots < A_N \tag{1.15}$$

holds, are considered and properly ordered. The abstract and x-representation configurations are again simply related

$$\Phi_K \equiv \Phi_K(1,2,\ldots,N) \equiv \Phi_K(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_N)$$
$$= (N!)^{-1/2} \det \| \varphi_{A_1}(\vec{x}_1)\,\varphi_{A_2}(\vec{x}_2) \cdots \varphi_{A_N}(\vec{x}_N) \|$$
$$= (N!)^{-1/2} \sum_{P \in \mathfrak{S}_N} (-1)^p\, \varphi_{A_1}(\vec{x}_{p1})\,\varphi_{A_2}(\vec{x}_{p2}) \cdots \varphi_{A_N}(\vec{x}_{pN})$$

$$= (N!)^{-1/2} \sum_{P \in \mathfrak{S}_N} (-1)^P \varphi_{A_{P_1}}(\vec{x}_1)\, \varphi_{A_{P_2}}(\vec{x}_2) \cdots \varphi_{A_{P_N}}(\vec{x}_N)$$

$$= \langle \vec{x}_1\ \vec{x}_2\ \cdots\ \vec{x}_N \mid \Phi_K(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N) \rangle$$

$$= \langle 1\ 2\ \cdots\ N \mid \Phi_K \rangle$$

$$= \langle 1\ 2\ \cdots\ N \mid \{A_1\ A_2\ \cdots\ A_N\} \rangle \qquad (1.16)$$

Our notation stresses that the configurations are antisymmetric in the occupied single particle states by enclosing them in curly brackets. In this way no confusion between configurations or basis states of $\mathcal{F}_N$ and simple products of one-particle states, Eq. (9b), can arise. Thus, for example, we have

$$|\{AB\}\rangle = |\{A(1)\ B(2)\}\rangle = 2^{-1/2} \begin{vmatrix} |A(1)\rangle & |A(2)\rangle \\ |B(1)\rangle & |B(2)\rangle \end{vmatrix}$$

$$= 2^{-1/2} \left(|A(1)\rangle |B(2)\rangle - |B(1)\rangle |A(2)\rangle \right)$$

$$= 2^{-1/2} \left(|A(1)\ B(2)\rangle - |B(1)\ A(2)\rangle \right)$$

$$= 2^{-1/2} \left(|AB\rangle - |BA\rangle \right). \qquad (1.17)$$

Clearly,

$$|\{A\}\rangle \equiv |A\rangle. \qquad (1.18)$$

We also recall that expanding the Slater determinant in (16) about, say, the first column, we obtain

$$\langle \vec{x}_1\ \vec{x}_2 \cdots \vec{x}_N \mid \{A_1\ A_2\ \cdots\ A_N\} \rangle$$

$$= N^{-1/2} \sum_{j=1}^{N} \langle \vec{x}_1 \mid A_j \rangle\, (-1)^{j-1} \langle \vec{x}_2 \cdots \vec{x}_N \mid \{A_1 \cdots A_{j-1} A_{j+1} \cdots A_N\} \rangle. \qquad (1.19)$$

Similarly, expanding about the i-th row we get

$$\langle 1\ 2\ \cdots\ N \mid \{A_1\ A_2\ \cdots\ A_N\} \rangle$$

$$= N^{-1/2} \sum_{\ell=1}^{N} (-1)^{\ell+i} \langle \ell \mid A_i \rangle \langle 1\ 2 \cdots (\ell-1)\,(\ell+1) \cdots N \mid \{A_1 \cdots A_{i-1} A_{i+1} \cdots A_N\} \rangle. \qquad (1.20)$$

Now, the ordered configurations $|\Phi_K\rangle$ form a basis of the N-particle space $\mathcal{F}_N$ so that any N-electron state $|\Psi\rangle = |\Psi(1,2,\ldots,N)\rangle$ can be expanded as follows

$$|\Psi\rangle = \sum_K C_K |\Phi_K\rangle, \qquad (1.21)$$

where

$$C_K = \langle \Phi_K | \Psi \rangle. \qquad (1.22)$$

Formally the above expansion immediately results by applying the resolution of identity, for the N-particle space $\mathcal{F}_N$, to $|\Psi\rangle$,

$$\hat{1} = \sum_K |\Phi_K\rangle\langle\Phi_K|$$

$$= \sum_{A_1 < A_2 < \cdots < A_N} |\{A_1 A_2 \cdots A_N\}\rangle\langle\{A_1 A_2 \cdots A_N\}|$$

$$= (N!)^{-1} \sum_{A_1, A_2, \ldots, A_N} |\{A_1 A_2 \cdots A_N\}\rangle\langle\{A_1 A_2 \cdots A_N\}|, \qquad (1.23)$$

where the last equation follows immediately from the antisymmetry of configurations $|\Phi_K\rangle$, Eq. (14) or (16),

$$P\, |\{A_1\ A_2 \cdots A_N\}\rangle = (-1)^P |\{A_1\ A_2 \cdots A_N\}\rangle. \qquad (1.24)$$

Let us, finally, also mention that a very similar formalism also applies to the systems of identical bosons. In this case the antisymmetric product states (configurations) $|\Phi_K\rangle$ of the fermion theory must be replaced by the corresponding totally symmetrized products $|\Phi_K\rangle^S$ of spin-orbitals $|A_i\rangle$,

$$|\Phi_K\rangle^S = \mathcal{S}\, |A_1\ A_2 \cdots A_N\rangle, \qquad (1.25)$$

where

$$\mathcal{S} = (N!)^{-1} \sum_{P \in \mathfrak{S}_N} P. \qquad (1.26)$$

However, in contrast with the fermionic states, each spin-orbital $|A_i\rangle$ may now be contained in K or $|\Phi_K\rangle^S$ any number of

times, so that an ordered configuration states, which we can designate as

$$|\Phi_K\rangle^S \equiv |[A_1\, A_2 \cdots A_N]\rangle\,, \tag{1.27}$$

we have that

$$A_1 \leq A_2 \leq \cdots \leq A_N\,. \tag{1.28}$$

Clearly, the family of all index sets K, Eq.(28), is much larger than the fermionic one where the condition (15) holds and which is thus a proper subfamily of the bosonic family of index sets K.

1.2. Occupation number representation

In order to specify uniquely any given ordered configuration $|\Phi_K\rangle$, we have only to know which spinorbitals are occupied and which are unoccupied in $|\Phi_K\rangle$. Thus, the index set K, Eq. (11), uniquely defines $|\Phi_K\rangle$.

In the occupation number representation one generally defines the occupation number n_A of the spin-orbital $|A\rangle$ in the configuration $|\Phi_K\rangle$ as the number of how many times the spin-orbital $|A\rangle$ is occupied in $|\Phi_K\rangle$. For fermionic systems the occupation numbers n_A for any configuration can only take on two values, namely 0 and 1. Clearly, $n_A = 1$ if $|A\rangle$ is occupied in $|\Phi_K\rangle$, i.e. $A \in K$, and $n_A = 0$ if $|A\rangle$ is unoccupied in $|\Phi_K\rangle$, $A \notin K$. Using this terminology we can thus state that any configuration is fully determined by the occupation numbers of all spinorbitals pertaining to this configuration.

We note that bosonic configurations $|\Phi_K\rangle^S$, Eq. (25), are also uniquely defined by the occupations of the spin-orbitals $|A\rangle$. In fact, the ordering of spin-orbitals is not essential in this case and the occupation numbers are only restricted by the total number of particles,

$$0 \leq n_A \leq N\,. \tag{1.29}$$

Since the configurations $|\Phi_K\rangle$ (or, $|\Phi_K\rangle^S$) may be fully determined by the spinorbital occupation numbers n_A, we can also designate these configurations by simply listing the occupation numbers of individual spinorbitals, arranged in some definite order, as follows

$$|\, n_1\, n_2 \cdots \quad n_A \cdots \,\rangle\,. \tag{1.30}$$

This notation is often used in the second quantization formalism and is especially useful for systems of bosons. However, in the case of fermion systems, where the spinorbitals can be at most singly occupied, it is often simpler to use a formerly introduced notation which simply lists the occupied spin-orbitals. In other words, we write

$$|\Phi_K\rangle = |\{A_1\, A_2 \cdots A_N\}\rangle = |n_1\, n_2 \cdots \, n_A \cdots \,\rangle \tag{1.31}$$

where

$$\begin{aligned} n_A &= 1 \quad \text{if} \quad A = A_1, A_2, \cdots, A_N \\ \text{and} \quad n_A &= 0 \quad \text{if} \quad A \neq A_i,\ (i=1,\cdots,N)\,. \end{aligned} \tag{1.32}$$

Thus, for example

$$|\{1\,2\,7\,10\,12\}\rangle = |1\,1\,0\,0\,0\,0\,1\,0\,0\,1\,0\,1\,0\,0 \cdots 0 \cdots\rangle, \tag{1.33}$$

or, in the bosonic case

$$|[1\,1\,1\,3\,3\,5]\rangle = |3\,0\,2\,0\,1\,0\,0 \cdots 0 \cdots\rangle^{(S)}\,. \tag{1.34}$$

The former notation is clearly often much simpler for the fermionic systems than the standard occupation number notation. Sometimes, however,

it may be convenient to combine both notations. For example, the expansions (19) and (20) can be rewritten as follows using the occupation number notation

$$\langle 1\,2\,\dots N \,|\, \{A_1 A_2 \dots A_N\}\rangle =$$

$$= \langle 1\,2\,\dots N \,|\, n_1 n_2 \dots n_A \dots\rangle$$

$$= (N!)^{-1/2} \sum_A \langle 1\,|A\rangle\, n_A (-1)^{s(A)} \langle 2 \dots N \,|\, n_1 \dots (n_A - 1) \dots\rangle \tag{1.35}$$

and

$$= (N!)^{-1/2} \sum_{l=1}^{N} (-1)^{l-1+s(A)} \langle l\,|A\rangle \langle 1 \dots (l-1)(l+1) \dots N \,|\, n_1 \dots (n_A - 1)\rangle \tag{1.36}$$

where

$$s(A) = \sum_{I<A} n_I\,, \tag{1.37}$$

respectively. The last expression, s(A), gives clearly the number of ~~orbitals~~ occupied spinorbitals which are preceding the spinorbital $|A\rangle$ in the chosen ordering, i.e., the number of rows in a Slater determinant (16) which are above a given row associated with the spin-orbital $|A\rangle$. The unoccupied spinorbitals are deleted from the infinite sum in Eq. (35) by including the factor n_A.

To summarize, we can state that given a complete orthonormal system of spinorbitals $|A\rangle$, arranged in some definite order, we can form an orthonormal basis for an N-electron space as normalized antisymmetrized products of spinorbitals $|A\rangle$, each basis vector (ordered configuration) being determined by the spin-orbital occupation numbers n_A. An arbitrary N-electron state may then be represented using Eq. (21). This "occupation number representation" clearly depends on the chosen system of single particle states or spinorbitals $|A\rangle$.

1.3. Creation and annihilation operators

In nonrelativistic quantum mechanics any operator associated with some observable leaves invariant the total particle number of the system considered. In other words, these operators commute with the total particle number operator. It may thus seem very strange that we shall start our exposition by introducing the operators which do change the particle number of the system. We shall see later, however, that even in a nonrelativistic theory it may be formally advantageous to express the operators through certain basic operators which do not commute with the total particle number operator. We will now define these basic operators of the occupation number representation or second quantization formalism, namely the so called annihilation and creation operators. Let us also add that at this stage, however inapropriate this may seem from a mathematical standpoint, we shall define these operators purely heuristically and will ignore any questions concerning their domain of definition and range.

We define the annihilation operator $\hat{X}_{A_i}$ for an electron in the state $|A_i\rangle$ as follows

$$\hat{X}_{A_i} |\{A_i\, A_{j_2}\, A_{j_3} \dots A_{j_N}\}\rangle = |\{A_{j_2}\, A_{j_3} \dots A_{j_N}\}\rangle$$

$$\hat{X}_{A_i} |\{A_{j_1}\, A_{j_2} \dots A_{j_N}\}\rangle = 0 \quad \text{if} \quad i \neq j_1, j_2, \dots, j_N\,. \tag{1.38}$$

Thus, the annihilation operator $\hat{X}_{A_i}$ destroys (annihilates) a particle in the state A_i, assuming there is a particle present in this state (i.e., A_i is occupied) and, if there is no particle present in the state A_i, the annihilation operator X_{A_i} makes the whole N-particle state to

vanish.

In the first equation (38) we have assumed for simplicity that the state A_i appears as the first one in the configuration considered. This can always be achieved using the antisymmetry property of N-fermion states, since

$$|\{A_{i_1} A_{i_2} \cdots A_{i_N}\}\rangle = (-1)^P |\{A_{j_1} A_{j_2} \cdots A_{j_N}\}\rangle, \tag{1.39}$$

where $(-1)^P$ is the parity of the permutation P,

$$P \equiv \begin{pmatrix} j_1 & j_2 & \cdots & j_N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix}. \tag{1.40}$$

Thus, for example

$$\begin{aligned} \hat{X}_A |\{ABCD\cdots\}\rangle &= |\{BCD\cdots\}\rangle \\ \hat{X}_A |\{BACD\cdots\}\rangle &= -|\{BCD\cdots\}\rangle \\ \hat{X}_A |\{BCAD\cdots\}\rangle &= |\{BCD\cdots\}\rangle, \text{ etc.} \end{aligned} \tag{1.41}$$

Using this antisymmetry property and expressing our basis vectors in the form (30), in which the occupation numbers of all spinorbitals are listed in some definite order, we can rewrite definition (38) of the annihilation operator $\hat{X}_A$ in the following compact form

$$X_A |n_1 n_2 \cdots n_A \cdots\rangle = (-1)^{s(A)} n_A |n_1 n_2 \cdots (n_A - 1) \cdots\rangle, \tag{1.42}$$

where $s(A)$ is defined in Eq. (37). The exponent $s(A)$, which equals the number of occupied states preceding A, is identical with the number of elementary transpositions which are needed to shift the state A from its position in the antisymmetrized product into the first position assumed in definition (38). We could, of course, include the same phase factor in definition (38) should the annihilated state be in an arbitrary rather than the first position.

Definition (38) is unambiguous whenever $\hat{X}_A$ acts on an N-electron state where $N \geq 2$. It also applies to a one-electron case

$$\hat{X}_A |B\rangle = 0 \quad \text{when} \quad A \neq B. \tag{1.43}$$

In order to get an unambiguous result for an arbitrary single electronic state we define a so-called <u>vacuum state</u> $|0\rangle$. This state, also referred to as the <u>true</u> or <u>physical vacuum</u> may be regarded as the basis vector spanning the one-dimensional zero-particle space $\mathcal{F}_0$. Clearly, $\mathcal{F}_0$ is isomorphic with $\mathcal{C}$, the complex number field. Since no particles are present in this state, all occupation numbers must vanish, so that in the occupation number notation, Eq. (30), we can write

$$|0\rangle \equiv |0\ 0\ 0\ \cdots\ 0\ \cdots\rangle. \tag{1.44}$$

We also assume this state to be normalized so that

$$\langle 0|0\rangle = 1. \tag{1.45}$$

The introduction of this state, however formal or artificial it may seem, is very essential for the full generality and simplicity of the second quantization formalism. We can now complete the relationship (43) writing

$$\hat{X}_A |A\rangle = |0\rangle. \tag{1.46}$$

Finally, in line with the general definition (38) or (42) we have that

$$\hat{X}_A |0\rangle = 0. \tag{1.47}$$

We next introduce <u>creation operators.</u> This is most simply achieved by defining the creation operator $\hat{X}_A^\dagger$, associated with the state A, as a hermitian conjugate of the corresponding

annihilation operator $\hat{X}_A$, as already the notation employed suggests.

Let us first see what will result when $\hat{X}_A^\dagger$ acts on the vacuum state $|0\rangle$. Taking a hermitean conjugate of (46),

$$\langle A | \hat{X}_A^\dagger = \langle 0 | \tag{1.48}$$

and forming the scalar product with $|0\rangle$, we get

$$\langle A | \hat{X}_A^\dagger | 0 \rangle = \langle 0 | 0 \rangle = 1 , \tag{1.49}$$

where we have used the normalization property, Eq. (45), in the last equation. Invoking the completeness and orthonormality of the spin-orbitals $|A\rangle$, Eq. (49) implies that

$$\hat{X}_A^\dagger | 0 \rangle = | A \rangle . \tag{1.50}$$

Thus, as might have been expected, creating a particle in state A into the vacuum yields a single particle state $|A\rangle$.

In general the action of $\hat{X}_A$ on an N-electron state yields an (N+1)-electron state

$$X_A^\dagger | \{ A_1 A_2 \dots A_N \} \rangle = | \{ A\, A_1 A_2 \dots A_N \} \rangle \tag{1.51}$$

assuming that $A \neq A_i$, $i = 1, \dots, N$. Again, if the created state should appear in other than the first place in the resulting state, an appropriate sign factor must be introduced. However, if the state A is already present among the states A_i, $i = 1, \dots, N$, the resulting (N+1)-particle state will vanish in view of its antisymmetry property. Thus

$$\hat{X}_A^\dagger | \{ A_1 A_2 \dots A_N \} \rangle = 0 \quad \text{if} \quad A = A_i \text{ for some } i, i = 1, \dots, N. \tag{1.52}$$

Using the occupation number notation for our basis states we can ~~write~~ ~~define~~ write

for the creation operators $X_A^\dagger$ a relationship analogous to Eq. (42)

$$\hat{X}_A^\dagger | n_A n_2 \dots n_A \dots \rangle = (-1)^{s(A)} (1 - n_A) | n_1 n_2 \dots (n_A + 1) \dots \rangle , \tag{1.53}$$

where $s(A)$ is again given by Eq. (37).

Clearly, we could have also used the above relationship (or Eqs. (51, 52) for that matter) as a definition of the creation operators and then show subsequently that they are in fact hermitean conjugates of the corresponding annihilation operators.

1.4 Number operator

Assume now that we wish to construct an operator that preserves the particle number. It will have to consist, generally speaking, of products of the same numbers of creation and annihilation operators. Indeed, creating and subsequently annihilating (or vice versa) a particle in the same spinorbital state should have no effect on a given N-particle state as long as legitimate (N±1)-particle intermediate states result after the first operator is applied. For The operator product of the type $(\hat{X}_A^\dagger \hat{X}_A)$, this will be the case whenever spinorbital A is occupied in the initial state. Thus, the sum of all such operators must determine the total number of particles present. Let us now examine these qualitative considerations more precisely.

Consider the effect of the operator product $(\hat{X}_A^\dagger X_A)$ on an arbitrary N-particle basis state. We have to distinguish two possibilities, namely the case that (i) A is occupied and (ii) A is not occupied in the N-particle basis state considered:

(i) Assuming that $A \in K$ we can write the N-particle state $|\Phi_K\rangle$ as $|\{B_1 B_2 \cdots B_{i-1} A B_{i+1} \cdots B_N\}\rangle$. Operating with the product $(\hat{X}_A^\dagger \hat{X}_A)$ we get

$$(\hat{X}_A^\dagger \hat{X}_A)|\Phi_K\rangle = \hat{X}_A^\dagger \hat{X}_A |\{B_1 \cdots B_{i-1} A B_{i+1} \cdots B_N\}\rangle$$

$$= \hat{X}_A^\dagger \hat{X}_A (-1)^{i-1} |\{A B_1 \cdots B_{i-1} B_{i+1} \cdots B_N\}\rangle$$

$$= (-1)^{i-1} X_A^\dagger |\{B_1 \cdots B_{i-1} B_{i+1} \cdots B_N\}\rangle$$

$$= (-1)^{i-1} |\{A B_1 \cdots B_{i-1} B_{i+1} \cdots B_N\}\rangle$$

$$= |\{B_1 \cdots B_{i-1} A B_{i+1} \cdots B_N\}\rangle = |\Phi_K\rangle \quad (1.54)$$

(ii) Supposing, next, that $A \notin K$, we get clearly that

$$(\hat{X}_A^\dagger \hat{X}_A)|\Phi_K\rangle = 0 \quad (1.55)$$

which we can write formally as

$$(\hat{X}_A^\dagger \hat{X}_A)|\Phi_K\rangle = 0 \cdot |\Phi_K\rangle . \quad (1.56)$$

Thus, ignoring the fact that the intermediate vectors in our calculation do not belong to $\mathcal{F}_N$, we find that an arbitrary configuration $|\Phi_K\rangle$ or, equivalently, any occupation number representation state (30), is an eigenstate of the operator $(\hat{X}_A^\dagger \hat{X}_A)$ with an eigenvalue equal to the occupation number n_A. We can thus write both relationships (54) and (56) as follows

$$(\hat{X}_A^\dagger \hat{X}_A)|\Phi_K\rangle = n_A |\Phi_K\rangle \quad (1.57a)$$

or

$$(\hat{X}_A^\dagger \hat{X}_A)|n_1 n_2 \cdots n_A \cdots\rangle = n_A |n_1 n_2 \cdots n_A \cdots\rangle . \quad (1.57b)$$

The operator $(\hat{X}_A^\dagger \hat{X}_A)$ is thus appropriately referred to as the number operator of the state A, and is designated by $\hat{N}_A$,

$$\hat{N}_A = \hat{X}_A^\dagger \hat{X}_A . \quad (1.58)$$

We see immediately that $\hat{N}_A$ is hermitean (while X_A and $X_A^\dagger$ are clearly not hermitean) which is consistent with the fact that it represents an observable. Its eigenvalues must be real and tell us whether or not a given spinorbital state is occupied or unoccupied in $|\Phi_K\rangle$.

Designate, now, the sum of all possible number operators by $\hat{N}$,

$$\hat{N} \equiv \sum_A \hat{N}_A = \sum_A \hat{X}_A^\dagger \hat{X}_A . \quad (1.59)$$

This operator is called the <u>total number operator</u> since its eigenvalue for an arbitrary basis state is given by the sum of the occupation numbers of all single particle states and, obviously, is equal to the total number of electrons in a given state. Indeed,

$$\hat{N}|\Phi_K\rangle = N|\Phi_K\rangle , \qquad |\Phi_K\rangle \in \mathcal{F}_N \quad (1.60)$$

and

$$N = \sum_A n_A . \quad (1.61)$$

1.5. Fock space

We will now make our concepts and definitions more satisfactory from the mathematical point of view.

The annihilation and creation operators, as defined in Section 1.3, transform states from $\mathcal{F}_N$ to $\mathcal{F}_{N-1}$ and vice versa. Consequently, if we wish to use these operators we cannot restrict our considerations to the Hilbert space $\mathcal{F}_N$ of N-electron, quadratically integrable, antisymmetric states even when considering a nonrelativistic N-particle system.

A possible way out of these difficulties is to consider a much more general space $\mathcal{F}$, given as a direct sum of zero-,

one-, two-, three-, etc. particle spaces,

$$\mathcal{F} \equiv \mathcal{F}_0 \oplus \mathcal{F}_1 \oplus \mathcal{F}_2 \oplus \cdots \oplus \mathcal{F}_N \oplus \cdots . \tag{1.62}$$

This space is refered to as a <u>Fock Space.</u> The vectors of this space represent states of systems of identical particles (in our case fermions) whose number is not necessarily fixed. They may be represented as infinite sequences

$$(|\Phi^{(0)}\rangle, |\Phi^{(1)}\rangle, |\Phi^{(2)}\rangle, \cdots, |\Phi^{(N)}\rangle, \cdots) \tag{1.63}$$

where $|\Phi^{(N)}\rangle \in \mathcal{F}_N$, $|\Phi^{(0)}\rangle \equiv c|0\rangle$ and $(c \in \mathbb{C})$

$$\sum_N \langle \Phi^{(N)} | \Phi^{(N)} \rangle = \sum_N \| \Phi^{(N)} \|^2 < \infty . \tag{1.64}$$

Equivalently, the vectors from $\mathcal{F}$ may be regarded as column vectors,

$$|\Psi\rangle = \begin{pmatrix} |\Psi^{(0)}\rangle \\ |\Psi^{(1)}\rangle \\ \vdots \\ |\Psi^{(N)}\rangle \\ \vdots \end{pmatrix} \tag{1.65}$$

with the scalar product defined as follows

$$\langle \Psi | \Phi \rangle = \sum_{N=0}^{\infty} \langle \Psi^{(N)} | \Phi^{(N)} \rangle . \tag{1.66}$$

Any particle number preserving operator $\hat{H}$ acting on $\mathcal{F}$ can be writen as

$$\hat{H} = \hat{H}^{(0)} \oplus \hat{H}^{(1)} \oplus \cdots \oplus \hat{H}^{(N)} \oplus \cdots \tag{1.67}$$

and its action on $|\Psi\rangle \in \mathcal{F}$ is given by

$$\hat{H}|\Psi\rangle = \begin{pmatrix} \hat{H}^{(0)} |\Psi^{(0)}\rangle \\ \hat{H}^{(1)} |\Psi^{(1)}\rangle \\ \vdots \\ \hat{H}^{(N)} |\Psi^{(N)}\rangle \\ \vdots \end{pmatrix} . \tag{1.68}$$

The action of annihilation and creation operators on the vectors from $\mathcal{F}$ may then be defined as follows

$$\hat{X}_A |\Psi\rangle = \begin{pmatrix} X_A^{(1)} |\Psi^{(1)}\rangle \\ X_A^{(2)} |\Psi^{(2)}\rangle \\ \vdots \\ X_A^{(N+1)} |\Psi^{(N+1)}\rangle \\ \vdots \end{pmatrix}, \quad \hat{X}_A^{\dagger} |\Psi\rangle = \begin{pmatrix} 0 \\ X_A^{(1)\dagger} |\Psi^{(0)}\rangle \\ X_A^{(2)\dagger} |\Psi^{(1)}\rangle \\ \vdots \\ X_A^{(N)\dagger} |\Psi^{(N-1)}\rangle \\ \vdots \end{pmatrix} \tag{1.69}$$

where

$$\hat{X}_A^{(N)} : \mathcal{F}_N \to \mathcal{F}_{N-1} \qquad (\hat{X}_A^{(0)} : \mathcal{F}_0 \to 0) \tag{1.70a}$$

and

$$\hat{X}_A^{(N)\dagger} : \mathcal{F}_{N-1} \to \mathcal{F}_N . \tag{1.70b}$$

Since all $\mathcal{F}_N$ are naturally embedded in $\mathcal{F}$ (i.e., represented by sequences of the type $(0, 0, \cdots, 0, |\Phi^{(N)}\rangle, 0, \cdots)$) and all finite sequences span the pre-Hilbert space $\tilde{\mathcal{F}}$, we can extend the operators defined by ( ) on $\tilde{\mathcal{F}}$ to $\mathcal{F}$, since $\tilde{\mathcal{F}}$ is everywhere dense in $\mathcal{F}$. Indeed, in the fermion case considered, the annihilation and creation operators thus defined can be easily shown*) to be bounded on $\tilde{\mathcal{F}}$ so that the extension from $\tilde{\mathcal{F}}$ to $\mathcal{F}$ follows immediately.

---

*) We will show in the next section that

$$\hat{X}_A \hat{X}_A^{\dagger} + \hat{X}_A^{\dagger} \hat{X}_A = \langle A | A \rangle .$$

Taking a mean value in $|\Phi\rangle \in \mathcal{F}$ we get

$$\langle \Phi | \hat{X}_A \hat{X}_A^{\dagger} | \Phi \rangle + \langle \Phi | \hat{X}_A^{\dagger} \hat{X}_A | \Phi \rangle = \langle A | A \rangle \langle \Phi | \Phi \rangle \quad (*)$$

or, using the fact that $X_A^{\dagger}$ is a hermitean conjugate of X_A,

$$\langle \hat{X}_A^{\dagger} \Phi | \hat{X}_A^{\dagger} \Phi \rangle + \langle \hat{X}_A \Phi | \hat{X}_A \Phi \rangle = \|A\|^2 \, \|\Phi\|^2$$

since

$$\langle \hat{X}_A \Phi | \Psi \rangle = \langle \Phi | \hat{X}_A^{\dagger} \Psi \rangle .$$

However, $(*)$ immediately implies that

$$\| X_A \Phi \| \leq K \|\Phi\| , \quad \| \hat{X}_A^{\dagger} \Phi \| \leq K \|\Phi\| , \quad K \text{ bounded.}$$

Note also that the resolution of identity on $\mathcal{F}$ can be written as

$$\bigoplus_{N=0}^{\infty} \sum_{K(N)} |\Phi_K^{(N)}\rangle\langle\Phi_K^{(N)}| = \hat{1}\,, \qquad (1.71)$$

where the sum extends over all ordered configurations of $\mathcal{F}_N$. Equivalently, we can write using (23)

$$\hat{1} = \bigoplus_{N=0}^{\infty} (N!)^{-1} \sum_{A_1,\dots,A_N} |\{A_1 \cdots A_N\}\rangle\langle\{A_1 \cdots A_N\}|\,. \qquad (1.72)$$

As a result of these considerations, the domain and range of $\hat{X}_A$ or $\hat{X}_A^\dagger$ are both subspaces of $\mathcal{F}$ so that we can easily define their products. Moreover, since $\tilde{\mathcal{F}}$ is everywhere dense in $\mathcal{F}$ we can restrict all our proofs of various properties of these operators to a general N-particle component of $\tilde{\mathcal{F}}$. We also note that once we derive the basic algebraic properties of these operators, we will only use the operator algebraic manipulations and the simplest vector of $\mathcal{F}$, namely the vacuum state $|0\rangle$. Since the appropriate particle symmetry will be automatically contained in the algebra of these operators, we will be able to avoid the usual combinatorial considerations associated with the wavefunction formalism.

1.6 Anticommutation relations

We shall now derive basic algebraic properties of the creation and annihilation operators, which fully characterize the algebra of these operators, namely their commutation (or, rather, anti-commutation) relations. We have to distinguish three cases: (a) both operators are annihilators,
(b) both are creators, and
(c) one is a creation and one annihilation operator.

Let us start with case (a). In order to find the desired properties of two annihilators, we compare the result obtained with products*) $X_A X_B$ and $X_B X_A$ acting on an arbitrary N-particle component of the vector of the Fock space. In order to avoid many indices we shall label the spinorbitals of a defining complete orthonormal set with ~~latin~~ capitals from the beginning of the Latin alphabet. We can also assume, without restricting the generality of our considerations, that the one-particle state labels A and B are the first state labels in the configuration K considered (if, of course, $A, B \in K$), as implied by Eqs. (24) or (39).

We have to distinguish two cases depending on the occupancy of states A and B:

(i) If both A and B are occupied, $(A, B \in K)$, we have

$$\begin{aligned} X_A X_B |\{ABCD\dots\}\rangle &= -X_A X_B |\{BACD\dots\}\rangle \\ &= -X_A |\{ACD\dots\}\rangle \\ &= -|\{CD\dots\}\rangle \end{aligned} \qquad (1.73a)$$

and

$$\begin{aligned} X_B X_A |\{ABCD\dots\}\rangle &= X_B |\{BCD\dots\}\rangle \\ &= |\{CD\dots\}\rangle\,. \end{aligned} \qquad (1.73b)$$

(i) If at least one of the states A and B

*) We drop the carret signs on the operator symbols if no confusion can arrise.

is unoccupied in $|\Phi_K\rangle$, the result vanishes *)

$$X_A X_B |\Phi_K\rangle = X_B X_A |\Phi_K\rangle = 0 , \quad A,B \notin K . \tag{1.74}$$

Adding both relations in cases (i) or (ii) we find that for an arbitrary ket $|\Phi_K\rangle$ the result vanishes

$$(X_A X_B + X_B X_A) |\Phi_K\rangle = 0 . \tag{1.75}$$

We can thus write this relationship as an operator relation

$$X_A X_B + X_B X_A = 0 , \tag{1.76}$$

which indicates that the annihilation operators anticommute. We define a special symbol for this quantity, a so called <u>anticommutator</u> of two operators $\hat{P}$ and $\hat{Q}$,

$$[\hat{P}, \hat{Q}]_+ = \hat{P}\hat{Q} + \hat{Q}\hat{P} \tag{1.77}$$

in analogy with the well known commutator notation

$$[\hat{P}, \hat{Q}]_- = \hat{P}\hat{Q} - \hat{Q}\hat{P} . \tag{1.78}$$

Obviously,

$$[\hat{P}, \hat{Q}]_+ = [\hat{Q}, \hat{P}]_+ \tag{1.79}$$

while

$$[\hat{P}, \hat{Q}]_- = -[\hat{Q}, \hat{P}]_- . \tag{1.80}$$

Using this symbol we express the relationship (76) as

$$[X_A, X_B]_+ = 0 . \tag{1.81}$$

In a similar way we can find that the anticommutator of two creation operators, corresponding to the case (b) above, also vanishes

$$[X_A^\dagger, X_B^\dagger]_+ = 0 . \tag{1.82}$$

*) The reader can examine in detail the three possible cases : (α) $A \in K, B \notin K$, (β) $A \notin K, B \in K$ and (γ) $A,B \notin K$.

Indeed, if at least one of the spinorbitals A and B is occupied, we have

$$X_A^\dagger X_B^\dagger |\Phi_K\rangle - X_B^\dagger X_A^\dagger |\Phi_K\rangle = 0 , \quad A \text{ or } B \in K , \tag{1.83}$$

while in the case that both A and B are unoccupied, $A,B \notin K$, we get

$$X_A^\dagger X_B^\dagger |\{CD\dots\}\rangle = |\{ABCD\dots\}\rangle \tag{1.84a}$$
$$X_B^\dagger X_A^\dagger |\{CD\dots\}\rangle = -|\{ABCD\dots\}\rangle . \tag{1.84b}$$

In deriving relations (81) and (82) we have tacitly assumed that $A \neq B$. It is, however, easy to see that these relations remain valid even when $A \equiv B$. In fact, in this case the result must always vanish since no single particle state can be doubly occupied (Pauli principle), so that in fact

$$(X_A)^2 = (X_A^\dagger)^2 = 0 . \tag{1.85}$$

Consider, finally, case (c). In this case we have to consider separately the diagonal and off-diagonal cases since no analogous relationship to Eq. (85) exists. Considering, first, the diagonal case, we recall that generally (cf., Section 1.4)

$$X_A^\dagger X_A |\Phi_K\rangle = n_A |\Phi_K\rangle , \tag{1.86}$$

so that we need only to examine the action of these operators in reversed order. Assuming that $A \in K$, we get clearly a vanishing result

$$X_A X_A^\dagger |\Phi_K\rangle = 0 , \quad A \in K \tag{1.87}$$

while for A unoccupied we find

$$X_A X_A^\dagger |\{BC\dots\}\rangle = X_A |\{ABC\dots\}\rangle = |\{BC\dots\}\rangle . \tag{1.88}$$

Thus, generally

$$X_A X_A^\dagger |\Phi_K\rangle = (1 - n_A) |\Phi_K\rangle . \tag{1.89}$$

Adding Eqs. (88) and (89) we thus find that the anticommutator of X_A and $X_A^\dagger$ reproduces an arbitrary state vector so that

$$[X_A, X_A^\dagger]_+ = \hat{1} \,. \tag{1.90}$$

It thus remains to consider the off-diagonal ($A \neq B$) case (c). We can again distinguish four different cases according to the occupancy of both states:

(i) $A, B \in K$

$$X_A X_B^\dagger \,|\{A\,B\,C\cdots\}\rangle = 0 \tag{1.91a}$$
$$X_B^\dagger X_A \,|\{A\,B\,C\cdots\}\rangle = X_B^\dagger \,|\{B\,C\cdots\}\rangle = 0 \tag{1.91b}$$

(ii) $A \notin K$, $B \in K$

$$X_A X_B^\dagger \,|\{B\,C\cdots\}\rangle = 0 \tag{1.92a}$$
$$X_B^\dagger X_A \,|\{B\,C\cdots\}\rangle = 0 \tag{1.92b}$$

(iii) $A \in K$, $B \notin K$

$$\begin{aligned} X_A X_B^\dagger \,|\{A\,C\,D\cdots\}\rangle &= X_A \,|\{B\,A\,C\,D\cdots\}\rangle \\ &= -X_A \,|\{A\,B\,C\,D\cdots\}\rangle \\ &= -\,|\{B\,C\,D\cdots\}\rangle \end{aligned} \tag{1.93a}$$
$$\begin{aligned} X_B^\dagger X_A \,|\{A\,C\,D\cdots\}\rangle &= X_B^\dagger \,|\{C\,D\cdots\}\rangle \\ &= \,|\{B\,C\,D\cdots\}\rangle \end{aligned} \tag{1.93b}$$

(iv) $A, B \notin K$

$$X_A X_B^\dagger \,|\{C\,D\cdots\}\rangle = 0 \tag{1.94a}$$
$$X_B^\dagger X_A \,|\{C\,D\cdots\}\rangle = 0 \tag{1.94b}$$

Thus, adding both results in each case we find that X_A and $X_B^\dagger$ exactly anticommute if $A \neq B$,

$$[X_A, X_B^\dagger]_+ = 0 \qquad \text{if } A \neq B. \tag{1.95}$$

We can thus summarize the desired anticommutation relations for fermionic creation and annihilation operators as follows

$$[X_A, X_B]_+ = [X_A^\dagger, X_B^\dagger]_+ = 0$$
$$[X_A, X_B^\dagger]_+ = \langle A|B\rangle \,. \tag{1.96}$$

These relationships express the statistics governing fermionic systems and automatically take care of the antisymmetry properties of pertinent states. We note that corresponding relationships for boson creation and annihilation operators are obtained when the anticommutators are replaced by commutators.

The above relationships remain in force even for non-orthogonal complete systems of single particle states if we interpret the symbol $\langle A|B\rangle$ as the overlap integral of spinorbitals A and B. However, in this text we shall always assume the spin-orbitals to be orthonormal. Nevertheless we find it convenient to employ the symbol $\langle A|B\rangle$ in place of the usual Kronecker delta δ_{AB}.

1.7. Examples

We now present several examples which illustrate certain properties and usefulness of the above derived relations.

1. Show that $\hat{N}_A$ is idempotent,

$$N_A^2 = N_A \,. \tag{1.97}$$

Indeed, using relations (96) we get

$$\hat{N}_A^2 = (X_A^\dagger X_A)^2 = X_A^\dagger X_A X_A^\dagger X_A = X_A^\dagger (1 - X_A^\dagger X_A) X_A$$
$$= X_A^\dagger X_A - (X_A^\dagger)^2 (X_A)^2 = \hat{N}_A \,, \tag{1.98}$$

where in the last equation we have used the fact that the square of X_A or $X_A^\dagger$ vanishes, Eq. (85).

2. Show that the number operators commute,

$$[\hat{N}_A, \hat{N}_B]_- = 0 . \tag{1.99}$$

Applying anticommutation relations to shift the B dependent operators to the left we get

$$\begin{aligned}
\hat{N}_A\hat{N}_B &= X_A^\dagger X_A X_B^\dagger X_B = X_A^\dagger\left(\langle A|B\rangle - X_B^\dagger X_A\right) X_B \\
&= \langle A|B\rangle\, X_A^\dagger X_B - X_A^\dagger X_B^\dagger X_A X_B \\
&= \langle A|B\rangle\, X_A^\dagger X_B - (-X_B^\dagger X_A^\dagger)(-X_B X_A) \\
&= \langle A|B\rangle\, X_A^\dagger X_B - X_B^\dagger(\langle A|B\rangle - X_B X_A^\dagger) X_A \\
&= \langle A|B\rangle\,(X_A^\dagger X_B - X_B^\dagger X_A) + X_B^\dagger X_B X_A^\dagger X_A
\end{aligned} \tag{1.100}$$

Now, the first term on the right hand side of the final equation always vanishes since either $A \neq B$, in which case $\langle A|B\rangle = 0$, or $A = B$, in which case the terms in parentheses cancel. Since the last term equals $\hat{N}_B\hat{N}_A$, the commutation property (99) follows.

3. Show that

$$[\hat{X}_A, \hat{N}_B]_- = \langle A|B\rangle\, \hat{X}_A \tag{1.101a}$$

and

$$[\hat{X}_A^\dagger, \hat{N}_B]_- = -\langle A|B\rangle\, X_A^\dagger . \tag{1.101b}$$

To proove the first relationship we simply have to "anticommute" the operator X_A "through" the operator $\hat{N}_B$,

$$\begin{aligned}
[\hat{X}_A, \hat{N}_B]_- &= X_A X_B^\dagger X_B - X_B^\dagger X_B X_A \\
&= (\langle A|B\rangle - X_B^\dagger X_A) X_B - X_B^\dagger X_B X_A \\
&= \langle A|B\rangle X_B + X_B^\dagger X_B X_A - X_B^\dagger X_B X_A \\
&= \langle A|B\rangle X_B = \langle A|B\rangle X_A .
\end{aligned} \tag{1.102}$$

The second relationship may be either derived in the same way as the first one or we can simply take a hermitean conjugate of (101a) obtaining

$$[X_A, N_B]_-^\dagger = \langle B|A\rangle\, X_A^\dagger , \tag{1.103}$$

and since

$$[\hat{P}, \hat{Q}]_-^\dagger = [\hat{Q}^\dagger, \hat{P}^\dagger]_- = -[\hat{P}^\dagger, \hat{Q}^\dagger]_- , \tag{1.104}$$

where $\hat{P}, \hat{Q}$ are arbitrary two operators, we get (101b).

4. An arbitrary basis ket (configuration) $|\{A_1 A_2 \cdots A_N\}\rangle$ may be expressed as

$$|\{A_1 A_2 \cdots A_N\}\rangle = X_{A_1}^\dagger X_{A_2}^\dagger \cdots X_{A_N}^\dagger |0\rangle = \left(\prod_{i=1}^{N} X_{A_i}^\dagger\right)|0\rangle, \tag{1.105}$$

where the product symbol designates an ordered product.

This relationship follows immediately when we recall Eqs. (50) and (51). Indeed

$$|\{A_N\}\rangle = |A_N\rangle = X_{A_N}^\dagger |0\rangle, \tag{1.106}$$

$$|\{A_{N-1} A_N\}\rangle = X_{A_{N-1}}^\dagger |\{A_N\}\rangle = X_{A_{N-1}}^\dagger X_{A_N}^\dagger |0\rangle, \tag{1.107}$$

and, generally

$$|\{A_i A_{i+1} \cdots A_N\}\rangle = X_{A_i}^\dagger |\{A_{i+1} \cdots A_N\}\rangle = \left(\prod_{j=i}^{N} X_{A_j}^\dagger\right)|0\rangle, \tag{1.108}$$

which is identical with (105) when $i = 1$.

5. Simplify following expressions

(i) $X_A X_B^\dagger X_C^\dagger |0\rangle$ (1.109a)

(ii) $\langle 0| X_A X_B^\dagger X_C X_D^\dagger |0\rangle$ (1.109b)

(iii) $\langle 0| X_A X_B X_C^\dagger X_D X_E^\dagger |0\rangle$ (1.109c)

In these problems we exploit the fact that the result of any annihilation operator acting on the true vacuum state vanishes, Eq. (47), or

the corresponding conjugate relationship, i.e.

$$X_A|0\rangle = 0 = \langle 0|X_A^\dagger . \tag{1.110}$$

We thus use the anticommutation relations (96) to move the annihilators to the rhs and/or the creators to the lhs. Thus,
(i) we anticommute X_A to the right hand side which yields

$$\begin{aligned} X_A X_B^\dagger X_C^\dagger|0\rangle &= (\langle A|B\rangle - X_B^\dagger X_A) X_C^\dagger|0\rangle \\ &= \langle A|B\rangle X_C^\dagger|0\rangle - X_B^\dagger(\langle A|C\rangle - X_C^\dagger X_A)|0\rangle \\ &= \langle A|B\rangle X_C^\dagger|0\rangle - \langle A|C\rangle X_B^\dagger|0\rangle \\ &= \langle A|B\rangle|C\rangle - \langle A|C\rangle|B\rangle . \end{aligned} \tag{1.111}$$

(ii) Similarly, anticommuting annihilators to the rhs we find

$$\begin{aligned} \langle 0|X_A X_B^\dagger X_C X_D^\dagger|0\rangle &= \langle 0|X_A X_B^\dagger(\langle C|D\rangle - X_D^\dagger X_C)|0\rangle \\ &= \langle C|D\rangle\langle 0|X_A X_B^\dagger|0\rangle \\ &\quad - \langle 0|X_A X_B^\dagger X_D^\dagger X_C|0\rangle \end{aligned} \tag{1.112}$$

and, since the last term vanishes in view of Eq. (110), we get finally

$$\begin{aligned} \langle 0|X_A X_B^\dagger X_C X_D^\dagger|0\rangle &= \langle C|D\rangle\langle 0|X_A X_B^\dagger|0\rangle \\ &= \langle C|D\rangle\langle 0|(\langle A|B\rangle - X_B^\dagger X_A)|0\rangle \\ &= \langle C|D\rangle\langle A|B\rangle\langle 0|0\rangle \\ &\quad - \langle 0|X_B^\dagger X_A|0\rangle\langle C|D\rangle \\ &= \langle C|D\rangle\langle A|B\rangle \\ &= \langle A|B\rangle\langle C|D\rangle . \end{aligned} \tag{1.113}$$

(iii) In this case we see immediately that the result must vanish since the numbers of creation and annihilation operators differ. This will always be the case when one calculates a vacuum mean value of the product of an odd number of operators. In this case we shall always be left with the vacuum mean values of the product of either only creation or only annihilation operators,

$$\langle 0|\prod_i X_{A_i}|0\rangle = 0 = \langle 0|\prod_i X_{A_i}^\dagger|0\rangle , \tag{1.114}$$

which clearly vanish. In the given case we can write

$$\begin{aligned} \langle 0|X_A X_B X_C^\dagger X_D X_E^\dagger|0\rangle &= \langle 0|X_A X_B X_C^\dagger(\langle D|E\rangle - X_E^\dagger X_D)|0\rangle \\ &= \langle D|E\rangle\langle 0|X_A X_B X_C^\dagger|0\rangle \\ &= \langle D|E\rangle\langle B|C\rangle\langle 0|X_A|0\rangle \\ &= 0 \end{aligned} \tag{1.115}$$

6. Calculate the overlap between the following two-particle configurations

$$|\Phi\rangle = |\{A_1 A_2\}\rangle \tag{1.116a}$$

$$|\Psi\rangle = |\{B_1 B_2\}\rangle , \tag{1.116b}$$

assuming orthonormality* of $|A_i\rangle, |B_i\rangle$, $i=1,2$.

Using the result of Ex. 4, Eq. (1.105), we can write $|\Phi\rangle$ as

$$|\Phi\rangle = X_{A_1}^\dagger X_{A_2}^\dagger|0\rangle \tag{1.117a}$$

so that

$$\langle\Phi| = \langle 0|X_{A_2} X_{A_1} , \tag{1.118}$$

and, similarly

$$|\Psi\rangle = X_{B_1}^\dagger X_{B_2}^\dagger|0\rangle . \tag{1.117b}$$

The overlap $\langle\Phi|\Psi\rangle$ of the two configurations (116) may thus be expressed as a vacuum mean value of the product of creation and annihilation operators,

$$\langle\Phi|\Psi\rangle = \langle 0|X_{A_2} X_{A_1} X_{B_1}^\dagger X_{B_2}^\dagger|0\rangle . \tag{1.119}$$

*) Interpreting $\langle A|B\rangle$ as a spinorbital overlap integral rather than the Kronecker delta, we obtain in fact a general result valid even for non-orthogonal spinorbitals (see comments at the end of Sec. 1.6).

We shall see later that in fact an evaluation of any integral of the type $\langle \Phi | \hat{O} | \Psi \rangle$, where $\hat{O}$ is some operator, reduces to the same type of problem. Applying the same technique as in a foregoing example, we get

$$\langle \Phi | \Psi \rangle = \langle 0 | X_{A_2} (\langle A_1 | B_1 \rangle - X_{B_1}^+ X_{A_1}) X_{B_2}^+ | 0 \rangle$$

$$= \langle A_1 | B_1 \rangle \langle 0 | X_{A_2} X_{B_2}^+ | 0 \rangle - \langle 0 | X_{A_2} X_{B_1}^+ X_{A_1} X_{B_2}^+ | 0 \rangle, \quad (1.120)$$

and, using the result of Ex. 5(ii) in the last term we get finally

$$\langle \Phi | \Psi \rangle = \langle A_1 | B_1 \rangle \langle 0 | (\langle A_2 | B_2 \rangle - X_{B_2}^+ X_{A_2} | 0 \rangle - \langle A_2 | B_1 \rangle \langle A_1 | B_2 \rangle$$

$$= \langle A_1 | B_1 \rangle \langle A_2 | B_2 \rangle - \langle A_2 | B_1 \rangle \langle A_1 | B_2 \rangle$$

$$= \begin{vmatrix} \langle A_1 | B_1 \rangle & \langle A_1 | B_2 \rangle \\ \langle A_2 | B_1 \rangle & \langle A_2 | B_2 \rangle \end{vmatrix} . \quad (1.121)$$

This is a well known result valid for N-particle configurations.

7. Show that any two configurations which differ in only one spin-orbital, say

$$|\Phi_1\rangle = |\{A\, A_2\, A_3 \cdots A_N\}\rangle \quad (1.122a)$$

$$|\Phi_2\rangle = |\{B\, A_2\, A_3 \cdots A_N\}\rangle , \quad (1.122b)$$

may be related as follows

$$|\Phi_2\rangle = X_B^+ X_A |\Phi_1\rangle \quad \text{or} \quad |\Phi_1\rangle = X_A^+ X_B |\Phi_2\rangle \quad (1.123)$$

This property follows immediately from basic relations (38), (51). Equivalently, we can write both states as follows using the result of Ex. 4

$$|\Phi_1\rangle = X_A^+ X_{A_2}^+ \cdots X_{A_N}^+ |0\rangle \quad (1.124a)$$

$$|\Phi_2\rangle = X_B^+ X_{A_2}^+ \cdots X_{A_N}^+ |0\rangle , \quad (1.124b)$$

and substituting (124a) into the first equation (123) we find

$$|\Phi_2\rangle = (X_B^+ X_A) X_A^+ X_{A_2}^+ \cdots X_{A_N}^+ |0\rangle$$

$$= X_B^+ (X_A X_A^+) |\{A_2 \cdots A_N\}\rangle .$$

$$= X_B^+ (\langle A | A \rangle - X_A^+ X_A) |\{A_2 \cdots A_N\}\rangle$$

$$= X_B^+ |\{A_2 \cdots A_N\}\rangle - X_A^+ X_A |\{A_2 \cdots A_N\}\rangle . \quad (1.125)$$

However, since A must be different from all A_i, $i = 2, \cdots, N$ (otherwise $|\Phi_1\rangle$ vanishes), the last term gives zero using (38) and the first relationship (123) follows. The second relationship (123) may be proved analogously.

It should be noted that the operators of the type $(X_A^+ X_B)$ may be generally interpreted as "excitation operators", i.e. the operators annihilating a particle in some state A and creating a particle in the state B, which is equivalent to "exciting" (or "de-exciting") a particle from the state A to the state B.

1.8. Canonical basis transformation

A linear transformation of the spin-orbital basis, which we used to define creation and annihilation operators of the second quantization formalism, implies a corresponding transformations of these annihilation and creation operators. A linear transformation which preserves the anticommutation relations (96) is called a canonical transformation.

Consider a general linear transformation of the annihilation operators X_A defined on some given orthonormal spin-orbital basis $\{|A\rangle\}$,

$$\tilde{X}_A = \sum_B C_{AB} X_B \,, \tag{1.126}$$

where we designated the new operators by a tilda. Since the creation operators $X_A^\dagger$ are hermitean conjugates of X_A, the transformation (125) implies the following transformation rules for creators

$$\tilde{X}_A^\dagger = \sum_B C_{AB}^* X_B^\dagger \,. \tag{1.127}$$

Examine now the conditions which must be satisfied by a linear transformation (126), defined by the matrix $\mathbb{C}$,

$$\mathbb{C} = \| C_{AB} \| \,, \tag{1.128}$$

in order to be canonical. We thus have to find the transformation laws for the anticommutators defining the basic relations (96). Considering first the anticommutators of the like operators (i.e., both creators or both annihilators) we find easily that they vanish for any linear transformation, since

$$[\tilde{X}_A, \tilde{X}_B]_+ = \tilde{X}_A \tilde{X}_B + \tilde{X}_B \tilde{X}_A = \sum_{C,D} \left(C_{AC} C_{BD} X_C X_D + C_{BD} C_{AC} X_D X_C \right)$$
$$= \sum_{C,D} C_{AC} C_{BD} [X_C, X_D]_+ = 0 \,, \tag{1.129}$$

and similarly for the creation operators. Second, for the anticommutator of unlike operators, we find

$$[\tilde{X}_A, \tilde{X}_B^\dagger]_+ = \sum_{C,D} C_{AC} C_{BD}^* [X_C, X_D^\dagger]_+$$
$$= \sum_{C,D} C_{AC} C_{BD}^* \langle C|D\rangle = \sum_C C_{AC} C_{BC}^* \tag{1.130}$$

using (96). Thus, we must require that

$$\sum_C C_{AC} C_{BC}^* = \langle A|B\rangle \,, \tag{1.131}$$

or, in matrix notation

$$\mathbb{C}\mathbb{C}^\dagger = \mathbb{I} = \mathbb{C}^\dagger \mathbb{C} \,, \quad \mathbb{I} = \| \langle A|B\rangle \| \equiv \| \delta_{AB} \| \tag{1.132}$$

in order that the transformed operators $\tilde{X}_A$ satisfy anticommutation relations (96). We also see immediately that for any linear transformation satisfying the conditions (131) or (132) the transformed operators satisfy relations (96).

We can, therefore, conclude that a linear transformation (126) is canonical if and only if it is unitary. This result could have indeed been expected since we assumed the spinorbitals, on which we define the second quantization operators, to be orthonormal. Since the spinorbitals $|\tilde{A}\rangle$, corresponding to the transformed operators $\tilde{X}_A$, are given by a conjugate transformation

$$|\tilde{A}\rangle = \tilde{X}_A^\dagger |0\rangle = \sum_B C_{AB}^* X_B^\dagger |0\rangle = \sum_B C_{AB}^* |B\rangle \,, \tag{1.133}$$

the new set $\{|\tilde{A}\rangle\}$ will be orthonormal if and only if $\mathbb{C}$ is unitary. We can also write that

$$\tilde{X}_A \equiv X_{\tilde{A}} \quad \text{and} \quad \tilde{X}_A^\dagger \equiv X_{\tilde{A}}^\dagger \,. \tag{1.134}$$

These considerations also indicate that as long as we consider a complete set of spin-orbitals it is immaterial which set is

actually used since all such sets are unitarily equivalent. Correspondingly, all the observables must be invariant with respect to an arbitrary canonical transformation. Let us verify this statement for the total particle number operator $\hat{N}$ defined in Sec. 1.4. (this is, so far, the only operator of an observable which we know how to express in second quantized form). We easily verify that

$$\tilde{\hat{N}} = \sum_A \tilde{X}_A^\dagger \tilde{X}_A = \sum_A \sum_C \sum_D C_{AC}^* C_{AD} X_C^\dagger X_D$$

$$= \sum_{C,D} \left(\sum_A C_{AC}^* C_{AD} \right) X_C^\dagger X_D$$

$$= \sum_{C,D} \langle C|D\rangle X_C^\dagger X_D = \sum_C X_C^\dagger X_C = \hat{N}, \qquad (1.135)$$

using the unitarity of $\mathbb{C}$, Eq. (132). We shall see later that the same property applies to any observable.

1.9 Field operators

The operators $X_A^\dagger$ and X_A are associated with some given discrete basis of a one-particle Hilbert space. Any such operator creates ($X_A^\dagger$) or annihilates (X_A) a particle occupying the spin-orbital $|A\rangle$. It is sometimes convenient to define analogous operators which are associated with a continuum of single particle states, for example, with the x-representation states $|\vec{x}\rangle$, which we may regard as the eigenstates of the position operator $\hat{x}$. Such operators are referred to as <u>field operators</u>.

We define such operators through the discrete operators X_A as follows

$$\hat{\psi}(\vec{x}) = \sum_A \langle \vec{x}|A\rangle X_A, \qquad (1.136a)$$

so that for the creation operators we have

$$\hat{\psi}^\dagger(\vec{x}) = \sum_A \langle A|\vec{x}\rangle X_A^\dagger. \qquad (1.136b)$$

The anticommutation relations satisfied by these operators are then easily found to be

$$[\hat{\psi}(\vec{x}), \hat{\psi}(\vec{x}')]_+ = \sum_{A,B} \langle \vec{x}|A\rangle\langle \vec{x}'|B\rangle [X_A, X_B]_+ = 0 \qquad (1.137a)$$

$$[\hat{\psi}^\dagger(\vec{x}), \hat{\psi}^\dagger(\vec{x}')]_+ = \sum_{A,B} \langle A|\vec{x}\rangle\langle B|\vec{x}'\rangle [X_A^\dagger, X_B^\dagger]_+ = 0 \qquad (1.137b)$$

$$[\hat{\psi}(\vec{x}), \hat{\psi}^\dagger(\vec{x}')]_+ = \sum_{A,B} \langle \vec{x}|A\rangle\langle B|\vec{x}'\rangle [X_A, X_B^\dagger]_+$$

$$= \sum_{A,B} \langle \vec{x}|A\rangle\langle B|\vec{x}'\rangle\langle A|B\rangle$$

$$= \sum_A \langle \vec{x}|A\rangle\langle A|\vec{x}'\rangle = \langle \vec{x}|\vec{x}'\rangle$$

$$= \delta(\vec{x} - \vec{x}'), \qquad (1.137c)$$

where $\delta(\vec{x})$ is the Dirac delta function.

Let us next find how these operators act on a general vector (65) from the Fock space. We shall write the x-representation of $|\Psi\rangle \in \mathcal{F}$ as Ψ,

$$\Psi \equiv \begin{pmatrix} \langle 0|\Psi\rangle \\ \langle \vec{x}_1|\Psi\rangle \\ \langle \vec{x}_1 \vec{x}_2|\Psi\rangle \\ \vdots \\ \langle \vec{x}_1 \cdots \vec{x}_N|\Psi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \Psi^{(0)} \\ \Psi^{(1)}(\vec{x}_1) \\ \Psi^{(2)}(\vec{x}_1, \vec{x}_2) \\ \vdots \\ \Psi^{(N)}(\vec{x}_1, \cdots, \vec{x}_N) \\ \vdots \end{pmatrix} \qquad (1.138a)$$

with the N-particle component expressed in any one of the following forms

$$\Psi^{(N)}(\vec{x}_1, \cdots, \vec{x}_N) \equiv \Psi^{(N)}(1, 2, \cdots, N) \equiv \langle \vec{x}_1 \cdots \vec{x}_N | \Psi^{(N)} \rangle$$

$$\equiv \langle \vec{x}_1 \cdots \vec{x}_N | \Psi \rangle \equiv \langle 1\ 2 \cdots N | \Psi \rangle . \qquad (1.138b)$$

Using the definition (136a) we can thus write

$$\hat{\psi}(\vec{x})\,|\Psi\rangle = \sum_A \langle \vec{x}|A\rangle\, X_A |\Psi\rangle$$

$$= \sum_A \bigoplus_{N=0}^{\infty} \sum_{\{n'\}} \langle \vec{x}|A\rangle\, X_A |n_1' n_2' \cdots n_A' \cdots\rangle$$

$$\times \langle n_1' n_2' \cdots n_A' \cdots |\Psi\rangle$$

$$= \sum_A \bigoplus_{N=0}^{\infty} \sum_{\{n\}} \langle \vec{x}|A\rangle\, n_A (-1)^{s(A)} |n_1 n_2 \cdots (n_A - 1) \cdots\rangle$$

$$\times \langle n_1 n_2 \cdots n_A \cdots |\Psi\rangle \qquad (1.139)$$

where we have used the resolution of identity for the Fock space, Eq. (71), expressed in the occupation number notation (30), (31) and the definition (42). For the N-particle component of (139) in the x-representation we get

$$\langle 1\ 2 \cdots N | \hat{\psi}(\vec{x}) | \Psi \rangle =$$

$$= \sum_A \bigoplus_{N'=0}^{\infty} \sum_{\{n'\}} \langle \vec{x}|A\rangle\, n_A (-1)^{s(A)} \langle 1\ 2 \cdots N | n_1' n_2' \cdots (n_A' - 1) \cdots\rangle$$

$$\times \langle n_1' n_2' \cdots n_A' \cdots |\Psi\rangle . \qquad (1.140)$$

In fact, only the (N+1) particle component of the direct sum in (140) gives a nonvanishing contribution, since the matrix element $\langle 1\ 2 \cdots N | n_1' n_2' \cdots (n_A' - 1) \cdots\rangle$ must vanish unless $\sum n_i' = N + 1$. However, we shall keep formally the whole direct sum over all N.

We now recall the expansion formula (35), which we can rewrite for the (N+1) particle case as follows

$$\langle \vec{y}\ \vec{x}_1 \cdots \vec{x}_N | n_1 n_2 \cdots n_A \cdots\rangle = \qquad (1.141)$$

$$= (N+1)^{-1/2} \sum_B \langle \vec{y}|B\rangle\, n_B (-1)^{s(B)} \langle \vec{x}_1 \cdots \vec{x}_N | n_1 \cdots (n_B - 1) \cdots\rangle .$$

Multiplying this expression by $\langle A|\vec{y}\rangle$ and integrating over $\vec{y}$ we get

$$\int d\vec{y}\, \langle A|\vec{y}\rangle \langle \vec{y}\ \vec{x}_1 \cdots \vec{x}_N | n_1 n_2 \cdots n_A \cdots\rangle$$

$$= (N+1)^{-1/2} \sum_B \int d\vec{y}\, \langle A|\vec{y}\rangle \langle \vec{y}|B\rangle\, n_B (-1)^{s(B)} \langle \vec{x}_1 \cdots \vec{x}_N | n_1 \cdots (n_B - 1) \cdots\rangle$$

$$= (N+1)^{-1/2} \sum_B \langle A|B\rangle\, n_B (-1)^{s(B)} \langle 1\ 2 \cdots N | n_1 \cdots (n_B - 1) \cdots\rangle$$

$$= (N+1)^{-1/2}\, n_A (-1)^{s(A)} \langle 1\ 2 \cdots N | n_1 \cdots (n_A - 1) \cdots\rangle \qquad (1.142)$$

where we have used the resolution of identity (2). Using this relationship in (140) we finally get

$$\langle 1\ 2 \cdots N | \hat{\psi}(\vec{x}) | \Psi \rangle \equiv \langle \vec{x}_1 \cdots \vec{x}_N | \hat{\psi}(\vec{x}) | \Psi \rangle =$$

$$= \sum_A \bigoplus_{N=0}^{\infty} \sum_{\{n'\}} \langle \vec{x}|A\rangle\, (N+1)^{+1/2} \int d\vec{y}\, \langle A|\vec{y}\rangle \langle y\ \vec{x}_1 \cdots \vec{x}_N | n_1' n_2' \cdots\rangle$$

$$\times \langle n_1' n_2' \cdots |\Psi\rangle$$

$$= (N+1)^{1/2} \sum_A \int d\vec{y}\, \langle \vec{x}|A\rangle \langle A|\vec{y}\rangle \langle \vec{y}\ \vec{x}_1 \cdots \vec{x}_N | \Psi\rangle$$

$$= (N+1)^{1/2} \int d\vec{y}\, \langle \vec{x}|\vec{y}\rangle \langle \vec{y}\ \vec{x}_1 \cdots \vec{x}_N | \Psi\rangle \qquad (1.143)$$

$$= (N+1)^{1/2} \langle \vec{x}\ \vec{x}_1 \cdots \vec{x}_N | \Psi\rangle = (N+1)^{1/2}\, \Psi^{(N+1)}(\vec{x}, \vec{x}_1, \cdots, \vec{x}_N)$$

We shall examine the physical interpretation of

of these formal results shortly. Using the definition (136) we can also rewrite the expansion (35) into a very compact form

$$\langle 1\,2\,\dots\,N \mid n_1 n_2 \dots \rangle =$$

$$= N^{-\frac{1}{2}} \sum_A \langle 1|A\rangle\, n_A (-1)^{s(A)} \langle 2\,3\dots N \mid n_1 n_2 \dots (n_A - 1) \dots \rangle$$

$$= N^{-\frac{1}{2}} \sum_A \langle 1|A\rangle \langle 2\,3\dots N \mid X_A \mid n_1 n_2 \dots n_A \dots \rangle$$

$$= N^{-\frac{1}{2}} \langle 2\,3\dots N \mid \hat{\psi}(1) \mid n_1 n_2 \dots \rangle \qquad (1.144)$$

and, by induction, we get

$$\langle 1\,2\dots N \mid n_1 n_2 \dots \rangle = (N!)^{-\frac{1}{2}} \langle 0 \mid \hat{\psi}(N) \dots \hat{\psi}(1) \mid n_1 n_2 \dots \rangle \qquad (1.145)$$

Since the ket state in this equation is arbitrary, we can also write

$$\mathcal{A} \mid 1\,2\dots N\rangle = (N!)^{-\frac{1}{2}} \psi^\dagger(1) \dots \psi^\dagger(N) \mid 0\rangle \qquad (1.146)$$

which represents the x-representation version of the relationship (105).

In a similar way we find the result of the action of the creation operator $\hat{\psi}^\dagger(\vec{x})$ on $|\Psi\rangle \in \mathcal{F}$. Considering directly the N-particle component of the resulting state, we get using (136b) and (53)

$$\langle 1\,2\dots N \mid \hat{\psi}^\dagger(\vec{x}) \mid \Psi\rangle =$$

$$= \sum_A \bigoplus_{N'=0}^{\infty} \sum_{\{n'\}} \langle A|\vec{x}\rangle \langle 1\,2\dots N \mid X_A^\dagger \mid n_1' n_2' \dots \rangle$$

$$\times \langle n_1' n_2' \dots \mid \Psi\rangle$$

$$= \sum_A \bigoplus_{N'=0}^{\infty} \sum_{\{n'\}} \langle A|\vec{x}\rangle\, (1-n_A) (-1)^{s(A)} \langle 1\,2\dots N \mid n_1' n_2' \dots (n_A'+1) \dots \rangle$$

$$\times \langle n_1' n_2' \dots n_A' \dots \mid \Psi\rangle, \qquad (1.147)$$

which is analogous to (140). We now use expansion (36) which we rewrite first in the following form

$$\langle 1\,2\dots N \mid n_1 n_2 \dots (n_A+1) \dots \rangle (1-n_A) = \qquad (1.148)$$

$$= N^{\frac{1}{2}} \sum_{\ell=1}^{N} (-1)^{\ell-1+s(A)} \langle \ell|A\rangle \langle 1\,2\dots(\ell-1)(\ell+1)\dots N \mid n_1 \dots n_A \dots \rangle,$$

where we included the factor $(1-n_A)$ indicating that A is assumed to be occupied (i.e., $n_A = 0$ and $n_A + 1 = 1$) on the left hand side. Substituting into (147) we find, $(N \geq 1)$

$$\langle 1\,2\dots N \mid \hat{\psi}^\dagger(\vec{x}) \mid \Psi\rangle =$$

$$= \sum_A \bigoplus_{N'=0}^{\infty} \sum_{\{n'\}} \langle A|\vec{x}\rangle (-1)^{2s(A)} N^{-\frac{1}{2}} \sum_{\ell=1}^{N} (-1)^{\ell-1} \langle \vec{x}_\ell | A\rangle$$

$$\times \langle \vec{x}_1 \dots \vec{x}_{\ell-1}\, \vec{x}_{\ell+1} \dots \vec{x}_N \mid n_1' n_2' \dots n_A' \dots \rangle$$

$$\times \langle n_1' n_2' \dots n_A' \dots \mid \Psi\rangle$$

$$= \sum_A \sum_{\ell=1}^{N} N^{-\frac{1}{2}} (-1)^{\ell-1} \langle \vec{x}_\ell | A\rangle \langle A|\vec{x}\rangle \langle \vec{x}_1 \dots \vec{x}_{\ell-1}\, \vec{x}_{\ell+1} \dots \vec{x}_N \mid \Psi\rangle$$

$$= N^{-\frac{1}{2}} \sum_{\ell=1}^{N} (-1)^{\ell-1} \delta(\vec{x} - \vec{x}_\ell) \langle \vec{x}_1 \dots \vec{x}_{\ell-1}\, \vec{x}_{\ell+1} \dots \vec{x}_N \mid \Psi\rangle. \qquad (1.149)$$

For the special case when $N = 0$ we have that

$$\langle 0 \mid \psi^\dagger(\vec{x}) \mid \Psi\rangle = 0 \qquad (1.150)$$

which follows immediately from (136b).

Let us now examine the physical significance of the field operators defined by (136a). We shall first convince ourselves that $\hat{\psi}^\dagger(\vec{x})$ is a creation operator by showing that $\hat{\psi}^\dagger(\vec{x})|0\rangle$ represents a one particle state. To see this, we first rewrite the total number operator (59) in terms of the field operators

$$\hat{N} = \sum_A X_A^\dagger X_A = \sum_{A,B} X_A^\dagger X_B \langle A|B\rangle$$

$$= \sum_{A,B} X_A^\dagger X_B \int d\vec{x}\, \langle A|\vec{x}\rangle \langle \vec{x}|B\rangle$$

$$= \int d\vec{x} \left(\sum_A \langle A|\vec{x}\rangle X_A^\dagger \right) \left(\sum_B \langle \vec{x}|B\rangle X_B \right)$$

$$= \int d\vec{x}\, \hat{\psi}^\dagger(\vec{x})\, \hat{\psi}(\vec{x}) .$$

Second, we evaluate the result of $\hat{N}$ acting on the state $\hat{\psi}^\dagger(\vec{x})|0\rangle$,

$$\hat{N}\hat{\psi}^\dagger(\vec{x})|0\rangle = \int d\vec{x}'\, \hat{\psi}^\dagger(\vec{x}')\hat{\psi}(\vec{x}')\hat{\psi}^\dagger(\vec{x})|0\rangle$$

$$= \int d\vec{x}'\, \hat{\psi}^\dagger(\vec{x}') \left[\delta(\vec{x}-\vec{x}') - \hat{\psi}^\dagger(\vec{x})\hat{\psi}(\vec{x}') \right] |0\rangle$$

$$= \hat{\psi}^\dagger(\vec{x})|0\rangle$$

since

$$\hat{\psi}(\vec{x})|0\rangle = \sum_A \langle \vec{x}|A\rangle X_A|0\rangle = 0 .$$

Thus, $\hat{\psi}^\dagger(\vec{x})|0\rangle$ is an eigenstate of $\hat{N}$ with the eigenvalue $N=1$. In a completely analogous way we can show that $\hat{\psi}^\dagger(\vec{x}_1)\hat{\psi}^\dagger(\vec{x}_2)|0\rangle$ is a two-particle state, etc.

Let us now evaluate the probability amplitude for finding a particle in the state $\hat{\psi}^\dagger(\vec{x})|0\rangle$ at $\vec{x}_1$,

$$\langle \vec{x}_1 | \hat{\psi}^\dagger(\vec{x})|0\rangle = (1!)^{-\frac{1}{2}} \langle 0|\hat{\psi}(\vec{x}_1)\hat{\psi}^\dagger(\vec{x})|0\rangle$$

$$= \langle 0|\, \delta(\vec{x}-\vec{x}_1) - \hat{\psi}^\dagger(\vec{x})\hat{\psi}(\vec{x}_1)|0\rangle$$

$$= \delta(\vec{x}-\vec{x}_1),$$

where we have used Eq. (146) in the first equation and the anticommutation relation (137c) in the second equation. We thus find that $\hat{\psi}^\dagger(\vec{x})$ creates a particle localized at $\vec{x}$. We also see that the one-particle state $\hat{\psi}^\dagger(\vec{x})|0\rangle$ cannot be normalized, since

$$\langle 0|\psi(\vec{x})\,\psi^\dagger(\vec{x})|0\rangle = \delta(0),$$

as could have been expected for an eigenstate belonging to a continuum.

2. Wick's theorem

We shall now formulate and prove the time-independent version of Wick's theorem, which effectively expresses the combinatorial properties of an arbitrary product of creation and annihilation operators in a very compact form. This theorem enables a very efficient calculation of various quantities (particularly the average values of operators or transition moments) involving products of creation and annihilation operators and it also is easily amenable to a diagrammatic interpretation. Before formulating this theorem we have to introduce the necessary concepts of a normal product and a contraction of operators.

2.1 Normal product of operators

Consider an arbitrary product of creation and annihilation operators

$$\hat{M}_{B_1}\hat{M}_{B_2}\cdots\hat{M}_{B_N}, \qquad (2.1)$$

where $\hat{M}_{B_i} = X_{B_i}$ or $M_{B_i} = X_{B_i}^\dagger$. The <u>normal product</u> (or, simply, <u>n-product</u>) of these operators is defined as the product of the same operators whose ordering is rearranged in such a way that all the creators are to the left of all the annihilators. When rearranging a given product (1) to the normal product form we permute the operators as if they all exactly anticommute. In other words, we attach the phase (+1) or (-1) to the operators in the normal product form according to whether an even or an odd number of transpositions are necessary to achieve the desired ordering.

We shall denote the n-product of the operators (1) as follows

$$n[\hat{M}_{B_1}\hat{M}_{B_2}\cdots\hat{M}_{B_N}], \tag{2.2}$$

so that we can formally express the above definition as

$$n[\hat{M}_{B_1}\hat{M}_{B_2}\cdots\hat{M}_{B_N}] = (-1)^P X^\dagger_{C_1} X^\dagger_{C_2}\cdots X^\dagger_{C_k} X_{C_{k+1}}\cdots X_{C_N}, \tag{2.3}$$

where $(-1)^P$ is the parity of the permutation P,

$$P \equiv \begin{pmatrix} B_1 & B_2 & \cdots & B_N \\ C_1 & C_2 & \cdots & C_N \end{pmatrix}. \tag{2.4}$$

For example, we have

$$n[X^\dagger_A X_B X^\dagger_C] = -X^\dagger_A X^\dagger_C X_B. \tag{2.5}$$

It might seem that this definition is ambiguous since it does not specify explicitly the ordering within each group of like operators. For example, in the above example (5) we can also write

$$n[X^\dagger_A X_B X^\dagger_C] = X^\dagger_C X^\dagger_A X_B. \tag{2.6}$$

However, both results are equivalent since the like operators (in this case creators) exactly anticommute.

The following simple properties of the n-products are a direct consequence of our definition:

(i) For a single operator

$$n[M] = M \tag{2.7}$$

(ii) The operation of the normal product is idempotent

$$n[n[\hat{M}_1\cdots\hat{M}_N]] = n[\hat{M}_1\cdots\hat{M}_N]. \tag{2.8}$$

In particular, when the operators are already in the n-product form, the n-product symbol may be droped

$$n[X^\dagger_{A_1}\cdots X^\dagger_{A_k} X_{B_1}\cdots X_{B_\ell}] = X^\dagger_{A_1}\cdots X^\dagger_{A_k} X_{B_1}\cdots X_{B_\ell}. \tag{2.9}$$

This is always the case if the product (1) consists of the operators of only one kind, specifically

$$n[X^\dagger_{A_1}\cdots X^\dagger_{A_N}] = X^\dagger_{A_1}\cdots X^\dagger_{A_N} \tag{2.10a}$$

$$n[X_{B_1}\cdots X_{B_M}] = X_{B_1}\cdots X_{B_M}. \tag{2.10b}$$

(iii) The operators inside the normal product can be arbitrarily permuted as long as the sign is changed accordingly. Thus

$$n[\hat{M}_1\hat{M}_2\cdots\hat{M}_k] = (-1)^P n[\hat{M}_{p_1}\hat{M}_{p_2}\cdots\hat{M}_{p_k}] \tag{2.11}$$

where $(-1)^P$ designates the parity of the permutation P,

$$P \equiv \begin{pmatrix} 1 & 2 & \cdots & k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}. \tag{2.12}$$

In the special case of two operators we have

$$n[\hat{M}_1\hat{M}_2] = -n[M_2 M_1]. \tag{2.13}$$

Finally, we shall find it convenient to define the n-product of an empty set of operators as unity,

$$n[\varnothing] = 1. \tag{2.14}$$

2.2 Contraction of operators

A contraction (or a pairing) $\hat{M}_1\hat{M}_2$ of two operators M_1 and M_2 is defined as a difference between an ordinary product and a normal product of these operators

$$\hat{M}_1\hat{M}_2 = \hat{M}_1\hat{M}_2 - n[\hat{M}_1\hat{M}_2]. \tag{2.15}$$

There are clearly four basic possibilities for a contraction of $\hat{M}_1$ and $\hat{M}_2$:

(i) $\hat{M}_1 = X_A$, $\hat{M}_2 = X_B$

$$X_A X_B = X_A X_B - n[X_A X_B] = X_A X_B - X_A X_B = 0. \tag{2.16}$$

(ii) $M_1 = X_A^\dagger$, $M_2 = X_B^\dagger$

$$\underbrace{X_A^\dagger X_B^\dagger} = X_A^\dagger X_B^\dagger - n[X_A^\dagger X_B^\dagger] = 0 \qquad (2.17)$$

(iii) $M_1 = X_A^\dagger$, $M_2 = X_B$

$$\underbrace{X_A^\dagger X_B} = X_A^\dagger X_B - n[X_A^\dagger X_B] = 0 \qquad (2.18)$$

(iv) $M_1 = X_A$, $M_2 = X_B^\dagger$

$$\underbrace{X_A X_B^\dagger} = X_A X_B^\dagger - n[X_A X_B^\dagger] = X_A X_B^\dagger + X_B^\dagger X_A =$$

$$= [X_A, X_B^\dagger]_+ = \langle A|B\rangle . \qquad (2.19)$$

Thus, a contraction of two operators always vanishes, except in the case

$$\underbrace{X_A X_A^\dagger} = 1 . \qquad (2.20)$$

However, a contraction of any two operators, just like their anticommutator, is a c-number (i.e, a scalar operator $c\cdot\hat{1}$).

To state the Wick theorem we need finally the concept of a <u>normal product with contractions</u> (pairings), i.e. of a normal product of operators some of which are already contracted. We will define such a product to equal the product of all pairings, times the n-product of uncontracted operators, and the sign factor given by the parity of a permutation which effect the transition from the original order to the final order of the operators. Schematically this definition can be illustrated as follows

$$n[\hat{M}_1 \cdots \underbrace{\hat{M}_{i_1} \cdots \hat{M}_{i_2} \cdots \hat{M}_{j_2}} \cdots \underbrace{\hat{M}_{j_1} \cdots \hat{M}_{i_3} \cdots \hat{M}_{i_4} \cdots \hat{M}_{j_3}} \hat{M}_{j_4} \cdots \hat{M}_N]$$

$$= (-1)^p \underbrace{\hat{M}_{i_1}\hat{M}_{j_1}}\, \underbrace{\hat{M}_{i_2}\hat{M}_{j_2}} \cdots \underbrace{M_{i_k}M_{j_4}} \cdot n[\hat{M}_{m_1} \cdots \hat{M}_{i_{j-1}} \hat{M}_{i_{j+1}} \cdots \hat{M}_{m_\ell}], \qquad (2.21)$$

where we have assumed that the n-product contains k contractions of 2k operators (only four of these are

shown explicitly) and where $(-1)^p$ is the parity of permutation P,

$$P \equiv \begin{pmatrix} 1 & 2 & \cdots & 2k & (2k+1) & \cdots & N \\ i_1 & j_1 & \cdots & j_k & m_1 & \cdots & m_\ell \end{pmatrix}, \quad N = 2k+\ell . \qquad (2.22)$$

We must stress that the permutation P must not affect the order of each pair of contracted operators since

$$n[\underbrace{M_1 M_2}] \neq -n[\underbrace{M_2 M_1}] \qquad (2.23)$$

in contrast to an analogous relationship (13) for the uncontracted operators.

Let us now illustrate these concepts on a few examples which will also indicate their usefulness:

1. Evaluate (i) $n[\underbrace{X_A X_B} X_C^\dagger]$ (2.24)

(ii) $n[\underbrace{X_A X_B X_C} X_D^\dagger X_E^\dagger X_F^\dagger X_G X_H]$ (2.25)

(iii) $n[\underbrace{X_A X_B} \underbrace{X_C^\dagger X_D} \underbrace{X_E^\dagger X_F^\dagger}]$ (2.26)

Using the above definition of the normal product with contractions, we find easily that

(i)

$$n[\underbrace{X_A X_B} X_C^\dagger] = -\underbrace{X_A X_C^\dagger}\, n[X_B]$$

$$= -\langle A|C\rangle X_B \qquad (2.27)$$

where we have used (7) and (19) in the last equation. Similarly for the two remaining cases we find

(ii)

$$n[\underbrace{X_A X_B X_C} X_D^\dagger \underbrace{X_E^\dagger X_F^\dagger} X_G X_H^\dagger]$$

$$= -\underbrace{X_A X_D^\dagger}\, \underbrace{X_B X_E^\dagger}\, \underbrace{X_C X_F^\dagger}\, n[X_G X_H^\dagger]$$

$$= \langle A|D\rangle\langle B|E\rangle\langle C|F\rangle X_H^\dagger X_G \qquad (2.28)$$

(iii) $n[\underbrace{X_A X_B} \underbrace{X_C^\dagger X_D} \underbrace{X_E^\dagger X_F^\dagger}] =$

$$= -\underbrace{X_A X_C^\dagger}\, \underbrace{X_B X_F^\dagger}\, \underbrace{X_D X_E^\dagger}\, n[\varnothing] = -\langle A|C\rangle\langle B|F\rangle\langle D|E\rangle$$

(2.29)

This last example shows the usefulness of a formal definition (14).

2. A contraction of two neighboring operators can be taken out of the n-product without any sign change,

$$n[\cdots \hat{M}_i \hat{M}_{i+1} \cdots] = M_i M_{i+1}\, n[\cdots], \qquad (2.30)$$

since this operation always involves an even number of transpositions.

3. The fully contracted n-product (i.e., the n-product of an even number of operators in which all operators are contracted) is a c-number the absolute of which is given by the product of all pairings (arranged in an arbitrary order) and the sign factor equals $(-1)^q$, where q is the number of "intersections" of the pairing symbols inside the n-product.

For example in (29) we have $q=1$ (or $q=3$ if we draw the pairing $\underbrace{X_D X_E^\dagger}$ in such a way that it intersects the pairing $\underbrace{X_B X_F^\dagger}$) so that the sign factor is (-1) in agreement with the obtained result. The same rule can also be used in example 1(ii) since all contracted operators are grouped together.

We now have all the necessary concepts which are needed to formulate Wick's theorem. However, it will prove useful to study first in more detail the introduced concepts and their usefulness.

2.3. Basic rules for n-products and contractions

In order to express simply some important properties of the n-products we designate a general product (1) by a simple symbol $\hat{P}(A)$. We then postulate the linear property

$$n[\lambda \hat{P}(A) + \mu \hat{P}(B)] = \lambda\, n[\hat{P}(A)] + \mu\, n[\hat{P}(B)], \qquad (2.31)$$

where λ and μ are arbitrary (complex) numbers. delete

This implies immediately a general property

$$n\Big[\sum_i \lambda_i \hat{P}(A_i)\Big] = \sum_i \lambda_i\, n[\hat{P}(A_i)], \qquad (2.32)$$

as well as the distributive law

$$n[\hat{P}(A)\{\hat{P}(B) + \hat{P}(C)\}] = n[\hat{P}(A)\hat{P}(B)] + n[\hat{P}(A)\hat{P}(C)], \qquad (2.33)$$

so that in general

$$n\Big[\Big(\sum_i \lambda_i \hat{P}(A_i)\Big)\Big(\sum_j \mu_j \hat{P}(B_j)\Big)\Big] = \sum_{ij} \lambda_i \mu_j\, n[\hat{P}(A_i)\hat{P}(B_j)]. \qquad (2.34)$$

delete

The n-product concept is particularly useful when we calculate mean values of the operator products in the true vacuum state $|0\rangle$ or, more generally, the action of $\hat{P}(A)$ on $|0\rangle$. When we recall how the basic operators act on the vacuum state, namely that

$$X_A|0\rangle = 0 = \langle 0|X_A^\dagger, \qquad (2.35)$$

we can immediately state the following rules:

R1: A general n-product $n[\hat{M}_1 \cdots \hat{M}_N]$ yields a vanishing result when operating on the true vacuum state $|0\rangle$,

$$n[\hat{M}_1 \cdots \hat{M}_N]|0\rangle = 0, \qquad (2.36)$$

if there is <u>at least one annihilator</u> among the operators $\hat{M}_i$, $i=1,2,\ldots,N$.

R2: A true vacuum mean value of an arbitrary n-product vanishes,

$$\langle 0|\, n[\hat{M}_1 \cdots \hat{M}_N]\,|0\rangle = 0, \qquad (2.37)$$

assuming that $N \geq 1$.

Indeed, if there is at least one annihilator in the n-product, it will appear in the

rightmost position when the n-product operation is carried out and we obtain a vanishing result in view of (35). Similarly the relations (35) yield immediately the second rule (37). We can also extend these rules to the n-products with contractions:

R3: A normal product with contractions yields zero when operating on the vacuum state $|0\rangle$,

$$n[\hat{M}_1 \hat{M}_2 \cdots \underbrace{\hat{M}_{i_1} \cdots \hat{M}_{j_1}} \cdots \hat{M}_N]\,|0\rangle = 0, \qquad (2.38)$$

assuming there is <u>at least one uncontracted annihilator</u> in the product.

R4: A vacuum mean value of the n-product with pairings vanishes,

$$\langle 0|\, n[M_1 M_2 \cdots \underbrace{M_{i_1} \cdots M_{j_1}} \cdots \hat{M}_N]\,|0\rangle = 0, \qquad (2.39)$$

<u>unless all operators are contracted,</u>

since all the pairings are c-numbers and may be taken outside the n-product and, in case of R4, outside the mean value expression, reducing thus the rules 3 and 4 to the rules 1 and 2, respectively. The last rule also immediately leads to the following corollary:

R4': A vacuum mean value of the n-product with contractions of an <u>odd</u> number of operators always vanishes,

since in such a case there must exist at least one uncontracted operator.

Let us consider as an example a vacuum mean value of the product of two operators, $\langle 0|\hat{M}_1 \hat{M}_2|0\rangle$. The basic properties of creation and annihilation operators immediately imply that this mean value can give a non-vanishing

result only if $\hat{M}_1 = X_A$ and $\hat{M}_2 = X_B^+$ when

$$\langle 0|\hat{M}_1 \hat{M}_2|0\rangle = \langle 0|X_A X_B^+|0\rangle = \langle A|B\rangle. \qquad (2.40)$$

Using, on the other hand, the definition of a contraction, Eq. (15), we can rewrite the product as follows

$$\hat{M}_1 \hat{M}_2 = n[\hat{M}_1 \hat{M}_2] + \underbrace{\hat{M}_1 \hat{M}_2}, \qquad (2.41)$$

and applying R2 we find

$$\langle 0|\hat{M}_1 \hat{M}_2|0\rangle = \langle 0|\underbrace{\hat{M}_1 \hat{M}_2}|0\rangle$$

$$= \underbrace{\hat{M}_1 \hat{M}_2}\,\langle 0|0\rangle = \underbrace{\hat{M}_1 \hat{M}_2}. \qquad (2.42)$$

When we recall, finally, that the only non-vanishing contractions are of the type $\underbrace{X_A X_B^+}$, we recover immediately the result (40). This latter derivation using Eq. (41) is in fact the simplest application of Wick's theorem which we are now going to formulate and prove. In the trivial example given above this alternative derivation of Eq. (40) may seem rather cumbersome. However, we shall see later, that Wick's theorem enables the most convenient and systematic evaluation of more complicated expressions of this type.

2.4. Wick's theorem

The time-independent version of Wick's theorem may be stated as follows:

T1: An arbitrary product of creation and annihilation operators $\hat{M}_1 \hat{M}_2 \cdots \hat{M}_N$ is equal to the sum of the n-product and all possible n-products with contractions of these operators.

We can schematically write this result as follows

$$\hat{M}_1 \hat{M}_2 \dots \hat{M}_N = n[\hat{M}_1 \hat{M}_2 \dots \hat{M}_N] + \sum n[M_1 \underbrace{M_2 \dots M_N}], \qquad (2.43)$$

where the last term contains the sum of all n-products with all possible contraction schemes. More explicitly we can write

$$\begin{aligned} \hat{M}_1 \hat{M}_2 \dots \hat{M}_N = {} & n[\hat{M}_1 \hat{M}_2 \dots \hat{M}_N] \\ & + \sum_{1 \le i < j \le N} n[M_1 \dots \underbrace{M_i \dots M_j} \dots M_N] \\ & + \sum_{i,j,k,\ell}{}' n[M_1 \dots \underbrace{M_i \dots M_j} \dots \underbrace{M_k \dots M_\ell} \dots M_N] \\ & + \dots \\ & + \sum n[\underbrace{M_1 \dots M_i \dots \dots M_N}], \end{aligned} \qquad (2.44)$$

where we have collected under the same summation symbol all n-products with the same number of contractions. Thus, the second term contains all the n-products with one contraction, the third term is the sum of all possible n-products with two contractions, etc., until the last term represents the sum over all possible n-products having a maximal number of contractions (i.e., all operators contracted if N is even and all but one contracted if N is odd).
For example, in the case of $N=3$ and $N=4$ we have

$$\begin{aligned} \hat{M}_1 \hat{M}_2 \hat{M}_3 = {} & n[\hat{M}_1 \hat{M}_2 \hat{M}_3] + n[\underbrace{\hat{M}_1 \hat{M}_2} \hat{M}_3] + n[\underbrace{\hat{M}_1 \hat{M}_2 \hat{M}_3}] \\ & + n[M_1 \underbrace{M_2 M_3}] \end{aligned} \qquad (2.45)$$

and

$$\begin{aligned} \hat{M}_1 \hat{M}_2 \hat{M}_3 \hat{M}_4 = {} & n[\hat{M}_1 \hat{M}_2 \hat{M}_3 \hat{M}_4] + n[\underbrace{\hat{M}_1 \hat{M}_2} \hat{M}_3 \hat{M}_4] \\ & + n[\underbrace{\hat{M}_1 \hat{M}_2} \hat{M}_3 \hat{M}_4] + n[\hat{M}_1 \underbrace{\hat{M}_2 \hat{M}_3} \hat{M}_4] \\ & + n[\hat{M}_1 \underbrace{\hat{M}_2 \hat{M}_3 \hat{M}_4}] + n[\hat{M}_1 \hat{M}_2 \underbrace{\hat{M}_3 \hat{M}_4}] \\ & + n[\hat{M}_1 \hat{M}_2 \underbrace{\hat{M}_3 \hat{M}_4}] + n[\underbrace{\hat{M}_1 \hat{M}_2} \underbrace{\hat{M}_3 \hat{M}_4}] \\ & + n[\underbrace{\hat{M}_1 \hat{M}_2} \hat{M}_3 \hat{M}_4] + n[\hat{M}_1 \underbrace{\hat{M}_2 \hat{M}_3 \hat{M}_4}] \\ & + n[\underbrace{\hat{M}_1 \hat{M}_2 \hat{M}_3} \hat{M}_4] + n[\hat{M}_1 \hat{M}_2 \underbrace{\hat{M}_3 \hat{M}_4}] \\ & + n[\underbrace{\hat{M}_1 \hat{M}_2} \underbrace{\hat{M}_3 \hat{M}_4}] + n[\underbrace{\hat{M}_1 \hat{M}_2 \hat{M}_3 \hat{M}_4}]. \end{aligned} \qquad (2.46)$$

We shall prove this theorem in three steps. We first prove a lemma L1 which expresses an arbitrary n-product, multipled on the righ by a single operator, in terms of the n-product and n-products with pairings of all the operators involved. Next we shall generalize this lemma (L2) for n-products with contractions and, finally, we shall use the two lemmas to prove the general theorem.

L1:

$$\begin{aligned} & n[\hat{M}_1 \hat{M}_2 \dots \hat{M}_k] \hat{M}_\ell = \\ & \quad = n[\hat{M}_1 \hat{M}_2 \dots \hat{M}_k \hat{M}_\ell] + \sum_{i=1}^{k} n[\hat{M}_1 \dots \underbrace{\hat{M}_i \dots \hat{M}_k \hat{M}_\ell}]. \end{aligned} \qquad (2.47)$$

The proof of this lemma is trivial if $\hat{M}_\ell$ is an annihilator, say X_A. Indeed, all the n-products with a contraction appearing on the right hand side vanish, since

$$\underbrace{\hat{M}_i X_A} = 0, \qquad (2.48)$$

and, further

$$n[M_1 M_2 \dots M_k] X_A = n[M_1 M_2 \dots M_k X_A]. \qquad (2.49)$$

in view of (9) and (10). We can thus assume that $\hat{M}_\ell$ is a creator, say $X_B^\dagger$. Moreover, without any loss of generality, we can further assume that all the operators $\hat{M}_i$, $i=1,\dots,k$ are annihilators, since one can easily extend this special case to a general one as follows. We simply multiply from the left both sides of (47) with the product of pertinent creation operators. These being to the left of all the operators may be brought inside of all the n-products.

Then we can add to the right hand side the terms in which these added creators are contracted, one by one, with the last operator in the product, $X_B^\dagger$. All these terms vanish, since a contraction of two creators vanishes, and will not change the validity of our identity. Finally, we may rearrange the order of the operators $\hat{M}_i$, $i=1,\dots,k$, in each n-product as desired, since this will bring about the same sign change in each term of our identity. Moreover, any permutation of the operators $\hat{M}_i$, $i=1,\dots,k$ will not reverse the ordering of contracted operators, since the only contraction present is with the rightmost operator in the product, which is not affected by the permutation.

It thus remains to prove L1 for the special case in which $\hat{M}_1, \hat{M}_2, \dots, \hat{M}_k$ are annihilators, say $X_{A_1}, X_{A_2}, \dots, X_{A_k}$, and $\hat{M}_k$ is a creator, say $X_B^\dagger$, i.e.,

$$n[X_{A_1} X_{A_2} \cdots X_{A_k}] X_B^\dagger = n[X_{A_1} \cdots X_{A_k} X_B^\dagger] + \sum_{i=1}^{k} n[X_{A_1} \cdots \underbrace{X_{A_i} \cdots X_{A_k}} X_B^\dagger]. \quad (2.50)$$

We can now use the induction since (50) is clearly valid when $k=1$,

$$n[X_{A_1}] X_B^\dagger = n[X_{A_1} X_B^\dagger] + n[\underbrace{X_{A_1} X_B^\dagger}] = -X_B^\dagger X_{A_1} + \underbrace{X_{A_1} X_B^\dagger}\, n[\emptyset] = -X_B^\dagger X_{A_1} + \langle A|B\rangle \quad (2.51)$$

where we have used definitions (3), (21) and (14) and the result (19). Moreover, using (7), the left hand side is found to be equal to a simple product $X_{A_1} X_B^\dagger$ so that (51) represents the second anticommutation relation (1.96).

Suppose, thus, that (50) is valid for $k=N\geq 1$ and prove that it is also valid for $k=N+1$. To do so we multiply (50) for $k=N$ by an arbitrary

annihilator, say X_{A_0}, from the left, obtaining

$$X_{A_0}\, n[X_{A_1} \cdots X_{A_N}] X_B^\dagger = X_{A_0}\, n[X_{A_1} \cdots X_{A_N} X_B^\dagger] + \sum_{i=1}^{N} X_{A_0}\, n[X_{A_1} \cdots \underbrace{X_{A_i} \cdots X_{A_N}} X_B^\dagger]. \quad (2.52)$$

Using repeatedly property (10b) we can give the left hand side of (52) the required form

$$X_{A_0}\, n[X_{A_1} \cdots X_{A_N}] X_B^\dagger = X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger = n[X_{A_0} X_{A_1} \cdots X_{A_N}] X_B^\dagger. \quad (2.53)$$

Similarly we can rewrite all the terms under the summation symbol on the right hand side of (52) obtaining

$$\sum_{i=1}^{N} X_{A_0}\, n[X_{A_1} \cdots \underbrace{X_{A_i} \cdots X_{A_N}} X_B^\dagger] = \sum_{i=1}^{N} n[X_{A_0} X_{A_1} \cdots \underbrace{X_{A_i} \cdots X_{A_N}} X_B^\dagger]. \quad (2.54)$$

Finally, the first term on the right hand side of (52) can be rearranged using (11) to the form

$$X_{A_0}\, n[X_{A_1} \cdots X_{A_N} X_B^\dagger] = (-1)^N X_{A_0}\, n[X_B^\dagger X_{A_1} \cdots X_{A_N}], \quad (2.55)$$

so that using properties (9) and (10b) we get

$$X_{A_0}\, n[X_{A_1} \cdots X_{A_N} X_B^\dagger] = (-1)^N X_{A_0} X_B^\dagger\, n[X_{A_1} \cdots X_{A_N}] = (-1)^N \left(n[X_{A_0} X_B^\dagger] + \underbrace{X_{A_0} X_B^\dagger} \right) n[X_{A_1} \cdots X_{A_N}], \quad (2.56)$$

where in the last equation we have used definition (15). Recalling that in view of (13) we have $n[X_{A_0} X_B^\dagger] = -n[X_B^\dagger X_{A_0}]$, we get using repeatedly properties (9) and (10)

$$n[X_{A_0} X_B^\dagger]\, n[X_{A_1} \cdots X_{A_N}] = -X_B^\dagger X_{A_0} X_{A_1} \cdots X_{A_N} = -n[X_B^\dagger X_{A_0} X_{A_1} \cdots X_{A_N}], \quad (2.57)$$

and, using definition (21) we can write

$$\underline{X_{A_0} X_B^\dagger}\; n[X_{A_1} \cdots X_{A_N}] = n[\underline{X_{A_0} X_B^\dagger} X_{A_1} \cdots X_{A_N}]$$

$$= (-1)^N n[\underline{X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger}]. \qquad (2.58)$$

Substituting from (57) and (58) into (56) we get

$$X_{A_0} n[X_{A_1} \cdots X_{A_N} X_B^\dagger] = (-1)^{N+1} n[X_B^\dagger X_{A_0} X_{A_1} \cdots X_{A_N}]$$

$$+ (-1)^{2N} n[\underline{X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger}]$$

$$= n[X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger]$$

$$+ n[\underline{X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger}], \qquad (2.59)$$

which completes the proof, since substituting (53), (54) and (59) into (52) we finally obtain

$$n[X_{A_0} X_{A_1} \cdots X_{A_N}] X_B^\dagger = n[X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger]$$

$$+ n[\underline{X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger}]$$

$$+ \sum_{i=1}^{N} n[X_{A_0} X_{A_1} \cdots \underline{X_{A_i}} \cdots X_{A_N} X_B^\dagger]$$

$$= n[X_{A_0} X_{A_1} \cdots X_{A_N} X_B^\dagger]$$

$$+ \sum_{i=0}^{N} n[X_{A_0} \cdots \underline{X_{A_i} \cdots X_{A_N} X_B^\dagger}]. \qquad (2.60)$$

We now generalize L1 to the case of N-products with contractions.

L2:

$$n[\hat{M}_1 \hat{M}_2 \cdots \underline{\hat{M}_i \cdots \hat{M}_k}]\, M_\ell =$$

$$= n[\hat{M}_1 \hat{M}_2 \cdots \underline{\hat{M}_i \cdots \hat{M}_k} \hat{M}_\ell]$$

$$+ \sum_{\substack{j=1 \\ j \notin C}}^{k} n[\hat{M}_1 \hat{M}_2 \cdots \underline{\hat{M}_i \cdots \hat{M}_j \cdots \hat{M}_k} \hat{M}_\ell], \qquad (2.61)$$

where C designates the index set of those operators M_i, $i = 1, \ldots, k$ which are already contracted in the

N-product on the left hand side.

Thus in the N-product on the left hand side and in the first term on the right hand side only the operators $\hat{M}_i$, $i \in C$, are contracted, while in the last term there is an additional contraction involving the last operator $\hat{M}_\ell$ and some yet unpaired operator $\hat{M}_i$, $i = 1, \ldots, k$; $i \notin C$. The proof of this lemma is very easy when we realize that it reduces to L1 when C is empty. Thus, we can use definition (21) to move all the pairings outside the N-products and the same expression may then be obtained by multiplying (47) for the operators $\hat{M}_i$, $i = 1, \ldots, k$; $i \notin C$ from the left by all the contractions of the operators $\hat{M}_i$, $i \in C$. Again, all the sign changes must be identical for every term involved.

We are now ready to prove Wick's theorem. We shall use again the mathematical induction, since from the definition of a contraction, Eq. (15), the theorem holds for $N = 2$

$$\hat{M}_1 \hat{M}_2 = n[\hat{M}_1 \hat{M}_2] + n[\underline{\hat{M}_1 \hat{M}_2}]$$

$$= n[\hat{M}_1 \hat{M}_2] + \underline{\hat{M}_1 \hat{M}_2}, \qquad (2.62)$$

and, trivially, for $N = 1$. We thus assume its validity for $N \geq 2$ and prove that it is then also valid for $N+1$.

Indeed, multiplying Wick's theorem for N operators, written in the form (44), with an arbitrary creation or annihilation operator $\hat{M}_{N+1}$ from the right we obtain

$$\hat{M}_1 \hat{M}_2 \cdots \hat{M}_N \hat{M}_{N+1} = n[\hat{M}_1 \hat{M}_2 \cdots \hat{M}_N] \hat{M}_{N+1}$$

$$+ \sum_{1 \leq i < j \leq N} n[\hat{M}_1 \cdots \underline{\hat{M}_i \cdots \hat{M}_j} \cdots M_N] \hat{M}_{N+1}$$

$$+ \sum n[\hat{M}_1 \underline{\quad \cdots \quad} \underline{\quad \cdots \quad} \hat{M}_N] \hat{M}_{N+1}$$

$$+ \cdots$$

$$+\sum n\left[\hat{M}_1 \cdots \hat{M}_{k-1}\,\hat{M}_k\,\hat{M}_{k+1} \cdots \hat{M}_N\right]\hat{M}_{N+1}$$

$$+\left(\sum n\left[\hat{M}_1 \cdots \hat{M}_N\right]\hat{M}_{N+1}\right), \qquad (2.63)$$

where the last term (enclosed in parentheses), in which all operators inside the n-product are contracted, is only present when N is even, while the last unbracketed term, in which all but one operator are paired in the n-product, represents the last term when N is odd.

Using now L1 to express the first term on the right hand side of (63) and, similarly, L2 for all the subsequent terms, we get

$$\begin{aligned}
\hat{M}_1\hat{M}_2\cdots\hat{M}_N\hat{M}_{N+1} = {} & n\left[\hat{M}_1\hat{M}_2\cdots\hat{M}_N\hat{M}_{N+1}\right] \\
& + \sum_{i=1}^{N} n\left[\hat{M}_1\cdots\hat{M}_i\cdots\hat{M}_N\hat{M}_{N+1}\right] + \sum_{1\le i<j\le N} n\left[\hat{M}_1\cdots\hat{M}_i\cdots\hat{M}_j\cdots\hat{M}_N\hat{M}_{N+1}\right] \\
& + \sum_{1\le i<j\le N}\ \sum_{k(\neq i,j)} n\left[\hat{M}_1\cdots\hat{M}_i\cdots\hat{M}_j\cdots\hat{M}_k\cdots\hat{M}_{N+1}\right] + \sum n\left[\hat{M}_1 \cdots \hat{M}_{N+1}\right] \\
& + \sum n\left[\hat{M}_1 \cdots \hat{M}_N\hat{M}_{N+1}\right] + \cdots \qquad (2.64) \\
& + \cdots \\
& + \sum n\left[\hat{M}_1 \cdots \hat{M}_{k-1}\,\hat{M}_k\,\hat{M}_{k+1} \cdots \hat{M}_N\,\hat{M}_{N+1}\right] \\
& + \sum n\left[\hat{M}_1 \cdots \hat{M}_{k-1}\,\hat{M}_k\,\hat{M}_{k+1} \cdots \hat{M}_N\,\hat{M}_{N+1}\right] + \left(\sum n\left[\hat{M}_1 \cdots \hat{M}_N\,\hat{M}_{N+1}\right]\right).
\end{aligned}$$

The first two terms on the right hand side originate from the first term on the right hand side of (63), the third and fourth terms originate from the second term of (63), etc., while the terms containing the same number of contractions are written on the same line. Thus, both terms in the second line above contain n-products with a single contraction: the first one contains all the contractions involving the last operator $\hat{M}_{N+1}$, while the second one consists of the terms with all possible single contractions not involving the last operator, so that both terms together contain all possible n-products of $(N+1)$ operators with a single contraction. Similarly, the third row terms contain all possible n-products with two contractions, etc. Finally, the structure of the last term depends on whether N is even or odd. For N even, the terms in the last line contain all possible n-products with all but one operators contracted, since $(N+1)$ is then odd. Conversely, for odd N the bracketed terms in both expressions (63) and (64) are absent so that the last term in (64) contains all possible fully contracted terms. Consequently, Wick's theorem also hold for $(N+1)$ operators and, thus, in general.

2.5. Generalized Wick's theorem

As may be anticipated on the basis of our remarks made in preceding sections and as will be shown explicitly in the next section, Wick's theorem is extremely useful in calculations of vacuum mean values of products of annihilation and creation operators. However, we shall also need to calculate such mean values of the operator products in which some of the operators are already in the normal product form. We shall now generalize Wick's theorem to cover this case.

T2: An arbitrary product of creation and annihilation operators $\hat{M}_i$ and of n-products of such operators $n\left[\hat{M}_i\cdots\hat{M}_j\right]$ is equal to the n-product of all these operators plus the sum of n-products with all possible contraction schemes except for those which contain contraction(s) of the operators originating from the same n-product.

We can schematically illustrate this result as follows

$$\hat{M}_1\cdots\hat{M}_{k_1}\cdot n\left[\hat{M}_{k_1+1}\cdots\hat{M}_{k_2}\right]\cdot\hat{M}_{k_2+1}\cdots\hat{M}_{k_3}\cdot n\left[\hat{M}_{k_3+1}\cdots\hat{M}_{k_4}\right]\cdots\hat{M}_N =$$

$$= n[\hat{M}_1 \cdots \hat{M}_N]$$

$$+ \sum_{1 \le i < j \le N}{}' n[\hat{M}_1 \cdots \hat{M}_i \cdots \hat{M}_j \cdots \hat{M}_N]$$

$$+ \sum{}' n[\hat{M}_1 \cdots \hat{M}_i \cdots \hat{M}_j \cdots \hat{M}_k \cdots \hat{M}_e \cdots M_N]$$

$$+ \cdots$$

$$+ \sum{}' n[\hat{M}_1 \cdots \hat{M}_N], \qquad (2.65)$$

where the primes at the summation symbols indicate that we exclude all the terms in which any two operators, both appearing in the same n-product on the left hand side, are paired. For example

$$\hat{M}_1 n[\hat{M}_2 \hat{M}_3] = n[\hat{M}_1 \hat{M}_2 \hat{M}_3] + n[\hat{M}_1 \hat{M}_2 \hat{M}_3] + n[\hat{M}_1 \hat{M}_2 \hat{M}_3], \qquad (2.66)$$

or

$$n[\hat{M}_1 \hat{M}_2]\, n[\hat{M}_3 \hat{M}_4] = n[\hat{M}_1 \hat{M}_2 \hat{M}_3 \hat{M}_4] + n[M_1 M_2 M_3 M_4] + n[M_1 M_2 M_3 M_4] + n[M_1 M_2 M_3 M_4] + n[M_1 M_2 M_3 M_4] + n[M_1 M_2 M_3 M_4] + n[M_1 M_2 M_3 M_4] \qquad (2.67)$$

We can intuitively anticipate the validity of T2 when we rewrite T1 in the form

$$n[\hat{M}_1 \cdots \hat{M}_N] = M_1 \cdots M_N - \sum n[M_1 \cdots M_N], \qquad (2.68)$$

which suggests that one can reduce T2 to T1 by replacing the n-products with ordinary products while at the same time subtracting all the terms in which these operators are contracted.

A formal proof may be carried out as follows. The operators in each product, which appears as an argument of some n-product on the left hand side of (65),

can be arbitrarily permuted among themselves, since they appear inside the n-products in every term of eq. (65) and no two operators are contracted. We can thus rearrange them to the normal product order, so that we can drop the n-product symbols on the left hand side using (9) and apply T1. This yields the n-product of all the operators and n-products with all possible contractions. However, all the terms in which any two operators originating from the same n-product on the left hand side of (65) are contracted, must vanish, since the contraction must be of the type (i), (ii) or (iii), Sec. 2.2. Deleting all these vanishing terms and carrying out an inverse permutation, which brings the operators to their initial order, we obtain (65).

2.6. Vacuum mean values of operator products

We have seen in Ex. 4, Sec. 1.7, how to express an arbitrary configuration through the product of creation operators acting on a true vacuum state. We shall also see in the next Chapter how an arbitrary n-particle operator may be expressed in second quantized form as a linear combination of products of creation and annihilation operators. Consequently, the basic problem of calculating mean values or transition matrix elements of various operators for arbitrary states reduces to an evaluation of vacuum mean values of products of creation and annihilation operators. Similarly, the result of an action of various operators on arbitrary states reduces to an evaluation of expressions involving creation and annihilation operator products acting on the true vacuum state.

We have seen on a few simple examples in Sec. 1.7 that the expressions of the type

$$\langle 0|\prod_{i=1}^{N} M_i|0\rangle \quad \text{or} \quad \left(\prod_{i=1}^{N} M_i\right)|0\rangle \tag{2.69}$$

may be always evaluated by anticommuting the annihilators to the right hand side of the product and applying property (1.47). Since the Wick theorem may be regarded as a systematic and general prescription for carrying out this process, its usefulness in the evaluation of expressions of the type (69) is at hand. Recalling basic rules R1-4 of Sec. 2.3, we immediately realize that the majority of terms originating from the application of Wick's theorem will in fact vanish so that we can formulate the following simple but important rules.

R5: A vacuum mean value of a product of an odd number of creation and annihilation operators always vanishes. This is also true if some operators in the product are already in the n-product form.

This rule follows immediately from Wick's theorem (or its generalized version T2) and the corrolary R4' of Sec. 2.3. We can, therefore, restrict our attention to the products containing an even number of basic operators.

R6: A vacuum mean value of a product of an even number of creators and annihilators is equal to the sum of all fully contracted n-products, i.e.

$$\langle 0|\hat{M}_1 \cdots \hat{M}_{(2N)}|0\rangle$$

$$= \sum_{f.c.} \langle 0|n[\underbrace{\hat{M}_1 \hat{M}_2 \cdots}\ \ \underbrace{\hat{M}_{(2N)}}]|0\rangle$$

$$= \sum_{f.c.} (-1)^p \prod_{(k)} (\underbrace{\hat{M}_{i_k} \hat{M}_{j_k}}). \tag{2.70}$$

Here "f.c." under the summation symbols means that

the summation extends over the fully contracted terms only. The individual contractions in the last equation are either zero, in which case the whole term vanishes, or simple Kronecker deltas. The phase factor $(-1)^p$ originates from "unscrambling" of individual pairings and may be determined using the result of Example 3 of Sec. 2.2. Thus, up to the sign factor, the mean value (70) is given by the sum of products of Kronecker deltas originating from the contractions of type (iv), Sec. 2.2. (Eq. 19).

Let us illustrate this rule on a examples.

1. Evaluate the vacuum mean value (1.109b) of Example 5, Sec. 1.7.

Using R6 and considering only contractions between unlike operators, we get

$$\langle 0|X_A X_B^\dagger X_C X_D^\dagger|0\rangle = \underbrace{X_A X_B^\dagger}\,\underbrace{X_C X_D^\dagger} + \underbrace{X_A X_B^\dagger X_C X_D^\dagger}$$

$$= \langle A|B\rangle\langle C|D\rangle\,, \tag{2.71}$$

since the second system of contractions contains a vanishing type (iii) contraction, Eq. (18).

2. Calculate the overlap of Example 6, Sec. 1.7.

Contracting again only the operators of unlike types we get

$$\langle \Phi|\Psi\rangle = \langle 0|X_{A_2} X_{A_1} X_{B_1}^\dagger X_{B_2}^\dagger|0\rangle$$

$$= \underbrace{X_{A_2} X_{A_1} X_{B_1}^\dagger X_{B_2}^\dagger} + \underbrace{X_{A_2} X_{A_1} X_{B_1}^\dagger X_{B_2}^\dagger}$$

$$= \langle A_1|B_1\rangle\langle A_2|B_2\rangle - \langle A_1|B_2\rangle\langle A_2|B_1\rangle\,. \tag{2.72}$$

We shall see many other examples in the next Chapter. Later, we shall represent

the systems of pairing graphically, which will lead to a further significant simplification.

In a similar way we can evaluate quantities given by the second expression (69).

R7 : A result of an operator acting on the true vacuum is equal to the sum of all n-products in which all annihilation operators are contracted, i.e.

$$\hat{M}_1 \hat{M}_2 \cdots \hat{M}_N |0\rangle =$$

$$= \sum_{\text{a.f.c.}} n\left[\hat{M}_1 \hat{M}_2 \cdots \hat{M}_{i_1} \cdots \hat{M}_{j_1} \cdots \hat{M}_N\right]|0\rangle$$

$$= \sum_{\text{a.f.c.}} (-1)^p \left(\prod_{i_k j_k \in C} \hat{M}_{i_k} \hat{M}_{j_k}\right)\left(\prod_{r \notin C} X_r^\dagger\right)|0\rangle. \qquad (2.73)$$

Here "a.f.c" designates that all annihilators are fully contracted and C is the spinorbital index set of contracted operators in a given contraction scheme.

Modifications of the above rules for the case in which some operators are already in the n-product form are straightforward. We simply replace the summation symbols in (70) and (73) with primed ones indicating that all contraction schemes in which any pair of operators appearing in the same n-product on the left hand side must be deleted.

3. Representation of operators

In order to obtain the second quantized or occupation number representation of particle number preserving operators, we will examine how to express them through the basic annihilation and creation operators or, equivalently, through the field operators of Sec. 1.9.

We will find first general expressions for an N-particle component of a general particle number conserving operator, which will then be used to derive explicit expressions for one- and two-particle operators of such a system. We will assume in these derivations that the reader is familiar with the usual Slater rules for the evaluation of matrix elements of one- and two-particle operators in the basis of Slater determinants or configurations $|\Phi_K\rangle$. Once we derive the second quantized form for operators, we will also indicate an alternative proof by showing an equivalence of the occupation number and coordinate representations of a general operator. This equivalence is most easily demonstrated using Wick's theorem of the preceding section and nicely illustrate a nontrivial application of this theorem. It will also provide a motivation for a diagrammatic or graphical implementation of this theorem, which will be presented later.

3.1. Coordinate representation of operators

Any particle number conserving operator $\hat{Q}^{(N)}$ on $\mathcal{F}_N$ can be written as a sum of zero-, one-, two-, ..., N-particle operators $\hat{Q}_i$, $i=1,\ldots,N$

$$\hat{Q}^{(N)} = \sum_{i=0}^{N} \hat{Q}_i^{(N)}. \qquad (3.1)$$

In the x-representation we can write

$$\hat{Q}^{(N)}(\vec{x}_1, \dots, \vec{x}_N) = \sum_{i=0}^{N} \hat{Q}_i^{(N)}(\vec{x}_1, \dots, \vec{x}_N) \tag{3.2}$$

where

$$\begin{aligned} \hat{Q}_0^{(N)} &= q_0 \qquad (q_0 = \text{constant}) \\ \hat{Q}_i^{(N)}(\vec{x}_1, \dots, \vec{x}_N) &= \sum_{j_1 < j_2 < \cdots < j_i = 1}^{N} \hat{q}_i(\vec{x}_{j_1}, \dots, \vec{x}_{j_i}), \quad i = 1, \dots, N. \end{aligned} \tag{3.3}$$

Such an operator may be easily extended to the full Fock space $\mathcal{F}$ as we have seen in Sec. 1.5, Eqs. (1.67) and (1.68). Evidently, we have

$$\langle \Phi^{(N')} | \hat{Q} | \Psi^{(N)} \rangle = \delta_{NN'} \langle \Phi^{(N)} | \hat{Q}^{(N)} | \Psi^{(N)} \rangle, \tag{3.4}$$

where $\hat{Q}$ is defined as in Eq. (1.67) and $|\Phi^{(N)}\rangle$, $|\Psi^{(N)}\rangle$ are the N-particle components of $|\Phi\rangle$, $|\Psi\rangle$, respectively, Eq. (1.65). A general matrix element of $\hat{Q}$ in the x-representation, assuming $\hat{Q}$ is local both in space and spin, has the form

$$\begin{aligned} &\langle \vec{x}_1' \cdots \vec{x}_{N'}' | \hat{Q} | \vec{x}_1 \cdots \vec{x}_N \rangle \\ &= \delta_{NN'} \, \hat{Q}^{(N)}(\vec{x}_1, \dots, \vec{x}_N) \, \delta(\vec{x}_1 - \vec{x}_1') \cdots \delta(\vec{x}_N - \vec{x}_N'). \end{aligned} \tag{3.5}$$

All known physical interactions can be described by at most two-particle operators. For example, an N-electron spin-independent Hamiltonian of an atomic, molecular or a solid state system, has the general form

$$\hat{H}(\vec{x}_1, \dots, \vec{x}_N) \equiv \hat{H} = \hat{Z} + \hat{V} = \sum_{i=1}^{N} \hat{z}(\vec{r}_i) + \sum_{i<j=1}^{N} \hat{v}(|\vec{r}_i - \vec{r}_j|). \tag{3.6}$$

However, more than two-particle operators are often useful when various effective quasi-particle or multi-particle interactions are considered.

We shall employ in the following the abstract Dirac notation for the matrix elements of operators. Given a complete set

of spin-orbitals $|D_i\rangle$, $i = 1, 2, \dots$, (cf., Sec. 1.1), we shall write for a general matrix element of a one-particle operator $\hat{Q}_1$

$$\begin{aligned} \langle D_i | \hat{q}_1 | D_j \rangle &= \iint d\vec{x} \, d\vec{x}' \, \langle D_i | \vec{x} \rangle \langle \vec{x} | \hat{q}_1 | \vec{x}' \rangle \langle \vec{x}' | D_j \rangle \\ &= \int d\vec{x} \, \langle D_i | \vec{x} \rangle \, \hat{q}_1(\vec{x}) \, \langle \vec{x} | D_j \rangle \\ &= \int d\vec{x} \, \psi_{D_i}^*(\vec{x}) \, \hat{q}_1(\vec{x}) \, \psi_{D_j}(\vec{x}), \end{aligned} \tag{3.7}$$

and, similarly, for a two-particle operator $\hat{Q}_2$

$$\begin{aligned} \langle D_i D_j | \hat{q}_2 | D_k D_\ell \rangle &= \langle D_j D_i | \hat{q}_2 | D_\ell D_k \rangle = \\ &= \iiiint d\vec{x}_1 \, d\vec{x}_2 \, d\vec{x}_1' \, d\vec{x}_2' \, \langle D_i | \vec{x}_1 \rangle \langle D_j | \vec{x}_2 \rangle \langle \vec{x}_1 \vec{x}_2 | \hat{q}_2 | \vec{x}_1' \vec{x}_2' \rangle \\ &\qquad \times \langle \vec{x}_1' | D_k \rangle \langle \vec{x}_2' | D_\ell \rangle \\ &= \iint d\vec{x}_1 \, d\vec{x}_2 \, \langle D_i | \vec{x}_1 \rangle \langle D_j | \vec{x}_2 \rangle \, \hat{q}_2(\vec{x}_1, \vec{x}_2) \, \langle \vec{x}_1 | D_k \rangle \langle \vec{x}_2 | D_\ell \rangle \\ &= \iint d\vec{x}_1 \, d\vec{x}_2 \, \psi_{D_i}^*(\vec{x}_1) \, \psi_{D_j}^*(\vec{x}_2) \, \hat{q}_2(\vec{x}_1, \vec{x}_2) \, \psi_{D_k}(\vec{x}_1) \, \psi_{D_\ell}(\vec{x}_2), \end{aligned} \tag{3.8}$$

where we have used relations (5) and (1.1).

As in the example (6), the operators needed in the molecular electronic structure studies are often spin independent. In such a case it is usually convenient to eliminate the explicit spin dependence and to obtain a spin-free or an orbital form of the theory. We then write the spin-orbitals $|D_i\rangle$ as a product of the spatial or the orbital part $|d_i\rangle$ and the spin part $|\delta_i\rangle$,

$$|D_i\rangle = |d_i\rangle \, |\delta_i\rangle, \tag{3.9}$$

where $|\delta_i\rangle = |1/2\rangle$ or $|\delta_i\rangle = |-1/2\rangle$, and $|1/2\rangle \equiv |1/2, 1/2\rangle$, ($|-1/2\rangle \equiv |1/2, -1/2\rangle$) are the spin up (down) eigenfunctions of $\hat{S}^2$ and $\hat{S}_z$ for a single particle. We shall designate spin-orbitals with Latin capitals, as

we did so far, while the orbitals and spin functions will be designated by the lower case letters of the Latin and Greek alphabets, respectively, as in (9). In the x-representation we then have

$$\langle \vec{x} | D_i \rangle = \langle \vec{r} | d_i \rangle \langle s | \delta_i \rangle \tag{3.10}$$

or

$$\psi_{D_i}(\vec{x}) = \varphi_{d_i}(\vec{r})\, G_{\delta_i}(s) \tag{3.10'}$$

where

$$\vec{x} \equiv (\vec{r}, s) . \tag{3.11}$$

When considering matrix elements (7) and (8) for spin-independent operators, say $\hat{z}$ and $\hat{V}$ of (6), with the spin-orbitals of the form (10), we can carry out the summation(s) over the spin coordinate(s) obtaining

$$\begin{aligned}\langle D_i | \hat{z} | D_j \rangle &= \int d\vec{x} \langle D_i | \vec{x} \rangle \hat{z}(\vec{r}) \langle \vec{x} | D_j \rangle \\ &= \sum_s \langle \delta_i | s \rangle \langle s | \delta_j \rangle \int d\vec{r} \langle d_i | \vec{r} \rangle \hat{z}(\vec{r}) \langle \vec{r} | d_j \rangle \\ &= \langle \delta_i | \delta_j \rangle \langle d_i | \hat{z} | d_j \rangle , \end{aligned} \tag{3.12}$$

where we have defined the orbital matrix element

$$\begin{aligned}\langle d_i | \hat{z} | d_j \rangle &= \int d\vec{r} \langle d_i | \vec{r} \rangle \hat{z}(\vec{r}) \langle \vec{r} | d_j \rangle \\ &= \int d\vec{r}\, \varphi^*_{d_i}(\vec{r})\, \hat{z}(\vec{r})\, \varphi_{d_j}(\vec{r}) , \end{aligned} \tag{3.13}$$

and

$$\begin{aligned}&\langle D_i D_j | \hat{v} | D_k D_\ell \rangle = \\ &= \iint d\vec{x}_1 d\vec{x}_2 \langle D_i | \vec{x}_1 \rangle \langle D_j | \vec{x}_2 \rangle \hat{v}(\vec{r}_1, \vec{r}_2) \langle \vec{x}_1 | D_k \rangle \langle \vec{x}_2 | D_\ell \rangle \\ &= \sum_{s_1} \langle \delta_i | s_1 \rangle \langle s_1 | \delta_k \rangle \sum_{s_2} \langle \delta_j | s_2 \rangle \langle s_2 | \delta_\ell \rangle \\ &\quad \times \iint d\vec{r}_1 d\vec{r}_2 \langle d_i | \vec{r}_1 \rangle \langle d_j | \vec{r}_2 \rangle \hat{v}(\vec{r}_1, \vec{r}_2) \langle \vec{r}_1 | d_k \rangle \langle \vec{r}_2 | d_\ell \rangle \\ &= \langle \delta_i | \delta_k \rangle \langle \delta_j | \delta_\ell \rangle \langle d_i\, d_j | \hat{v} | d_k\, d_\ell \rangle , \end{aligned} \tag{3.14}$$

where

$$\begin{aligned}&\langle d_i\, d_j | \hat{v} | d_k\, d_\ell \rangle = \\ &= \iint d\vec{r}_1 d\vec{r}_2 \langle d_i | \vec{r}_1 \rangle \langle d_j | \vec{r}_2 \rangle \hat{v}(\vec{r}_1, \vec{r}_2) \langle \vec{r}_1 | d_k \rangle \langle \vec{r}_2 | d_\ell \rangle \\ &= \iint d\vec{r}_1 d\vec{r}_2\, \varphi^*_{d_i}(\vec{r}_1)\, \varphi^*_{d_j}(\vec{r}_2) \hat{v}(\vec{r}_1, \vec{r}_2)\, \varphi_{d_k}(\vec{r}_1)\, \varphi_{d_\ell}(\vec{r}_2) \end{aligned} \tag{3.15}$$

respectively.

3.2. Occupation number representation of operators

Let us now find a general form of quantum mechanical operators representing observables in the second quantization formalism. In view of the relationship, Eq. (4), we can restrict our attention to the N-particle component of a general operator $\hat{Q}$. We first consider the discrete representation employing the annihilation and creation operators X_A, $X_A^\dagger$ and, second, the representation through the field operators $\hat{\psi}(\vec{x})$, $\hat{\psi}^\dagger(\vec{x})$.

Using resolution of identity on the Fock space, Eq. (1.72), we can write for the desired matrix element on the right hand side of (4)

$$\begin{aligned}&\langle \Phi^{(N)} | \hat{Q}^{(N)} | \Psi^{(N)} \rangle \\ &= \bigoplus_{N'} \bigoplus_{N''} (N'!\, N''!)^{-1} \sum_{\substack{A_1 A_2 \cdots A_{N'} \\ B_1 B_2 \cdots B_{N''}}} \langle \Phi^{(N)} | \{ A_1 \cdots A_{N'} \} \rangle \\ &\quad \times \langle \{ A_1 \cdots A_{N'} \} | \hat{Q}^{(N)} | \{ B_1 \cdots B_{N''} \} \rangle \langle \{ B_1 \cdots B_{N''} \} | \Psi^{(N)} \rangle \\ &= (N!)^{-2} \sum_{\substack{A_1 \cdots A_N \\ B_1 \cdots B_N}} \langle \{ A_1 \cdots A_N \} | \hat{Q}^{(N)} | \{ B_1 \cdots B_N \} \rangle \\ &\quad \times \langle \Phi^{(N)} | \{ A_1 \cdots A_N \} \rangle \langle \{ B_1 \cdots B_N \} | \Psi^{(N)} \rangle , \end{aligned} \tag{3.16}$$

since only the N-particle configurations give a

nonvanishing contribution in each case. Expressing now the configurations through the creation operators, as in Eq. (1.105) of Example 1.4, in the last two factors, we get

$$\langle \Phi^{(N)} \mid \hat{Q}^{(N)} \mid \Psi^{(N)} \rangle$$

$$= (N!)^{-2} \sum_{\substack{A_1 \cdots A_N \\ B_1 \cdots B_N}} \langle \{A_1 \cdots A_N\} \mid \hat{Q}^{(N)} \mid \{B_1 \cdots B_N\} \rangle$$

$$\times \langle \Phi^{(N)} \mid X^\dagger_{A_1} \cdots X^\dagger_{A_N} \mid 0 \rangle \langle 0 \mid X_{B_N} \cdots X_{B_1} \mid \Psi^{(N)} \rangle . \quad (3.17)$$

We can now reconstitute the resolution of identity in the above expression, since all the matrix elements with other than vacuum component vanish,

$$\langle \Phi^{(N)} \mid X^\dagger_{A_1} \cdots X^\dagger_{A_N} \mid \Psi^{(M)}_K \rangle = 0$$

$$\langle \Psi^{(M)}_K \mid X_{B_N} \cdots X_{B_1} \mid \Psi^{(N)} \rangle = 0 , \quad M > 0, \quad (3.18)$$

so that we can formally write

$$\langle \Phi^{(N)} \mid \hat{Q}^{(N)} \mid \Psi^{(N)} \rangle$$

$$= (N!)^{-2} \sum_{\substack{A_1 \cdots A_N \\ B_1 \cdots B_N}} \langle \{A_1 \cdots A_N\} \mid \hat{Q}^{(N)} \mid \{B_1 \cdots B_N\} \rangle \quad (3.19)$$

$$\times \left\{ \bigoplus_{M=0}^{\infty} \sum_K \langle \Phi^{(N)} \mid X^\dagger_{A_1} \cdots X^\dagger_{A_N} \mid \Psi^{(M)}_K \rangle \langle \Psi^{(M)}_K \mid X_{B_N} \cdots X_{B_1} \mid \Psi^{(N)} \rangle \right\}$$

Deleting the identity operator, Eq. (1.71), in this last equation, we finally obtain

$$\langle \Phi^{(N)} \mid \hat{Q}^{(N)} \mid \Psi^{(N)} \rangle$$

$$= (N!)^{-2} \sum_{\substack{A_1 \cdots A_N \\ B_1 \cdots B_N}} \langle \{A_1 \cdots A_N\} \mid \hat{Q}^{(N)} \mid \{B_1 \cdots B_N\} \rangle$$

$$\times \langle \Phi^{(N)} \mid X^\dagger_{A_1} \cdots X^\dagger_{A_N} X_{B_N} \cdots X_{B_1} \mid \Psi^{(N)} \rangle . \quad (3.20)$$

This is the basic relationship for obtaining the desired form of the second quantized form of an arbitrary operator. For k-particle operators with $k < N$ (in particular for $k = 1$ or 2) this representation can be further simplified and expressed through the pertinent non-antisymmetrized k-particle matrix elements (e.g., (7) and (8) for $k = 1$ and 2, respectively). In deriving this form (cf., Secs. 3.2.1 and 3.2.2) the following Lemma will prove to be useful:

L3:

$$\sum_{A_{k+1} \cdots A_N} \langle \Phi^{(N)} \mid X^\dagger_{A_1} \cdots X^\dagger_{A_k} X^\dagger_{A_{k+1}} \cdots X^\dagger_{A_N} X_{A_N} \cdots X_{A_{k+1}} X_{B_k} \cdots X_{B_1} \mid \Psi^{(N)} \rangle$$

$$= (N-k)! \, \langle \Phi^{(N)} \mid X^\dagger_{A_1} \cdots X^\dagger_{A_k} X_{B_k} \cdots X_{B_1} \mid \Psi^{(N)} \rangle . \quad (3.21)$$

This relationship follows immediately when we realize that $X_{C_1} \cdots X_{C_r} \mid \Psi^{(N)} \rangle$ is an $(N-r)$-particle state if $r \leq N$ and vanishes if $r > N$, so that

$$\hat{N} \, (X_{C_1} \cdots X_{C_r}) \mid \Psi^{(N)} \rangle = (N-r) \, (X_{C_1} \cdots X_{C_r}) \mid \Psi^{(N)} \rangle , \quad (3.22)$$

where $r \leq N$, and $\hat{N}$ is the total number operator, Eq. (1.59).

The representation in terms of the field operators $\hat{\psi}(\vec{x})$ and $\hat{\psi}^\dagger(\vec{x})$, analogous to Eq. (20), is also easily obtained. We shall employ in this case the occupation number representation, Eq. (1.30), for the basic states or configurations. Consider, thus, a general matrix element

$$\langle n'_1 \, n'_2 \cdots \mid \hat{Q} \mid n_1 \, n_2 \cdots \rangle$$

$$= \int d\vec{x}_1 \cdots \int d\vec{x}_N \int d\vec{x}'_1 \cdots \int d\vec{x}'_{N'} \, \langle n'_1 \, n'_2 \cdots \mid \vec{x}'_1 \, \vec{x}'_2 \cdots x'_{N'} \rangle$$

$$\times \langle \vec{x}'_1 \, \vec{x}'_2 \cdots \vec{x}'_{N'} \mid \hat{Q} \mid \vec{x}_1 \, \vec{x}_2 \cdots \vec{x}_N \rangle \langle \vec{x}_1 \, \vec{x}_2 \cdots \vec{x}_N \mid n_1 \, n_2 \cdots \rangle \quad (3.23)$$

or, using a shorthand notation introduced in Sec. 1.1,

$$\langle n_1' n_2' \cdots | \hat{Q} | n_1 n_2 \cdots \rangle$$

$$= \int d1 \cdots \int dN \int d1' \cdots \int dN' \langle n_1' n_2' \cdots | 1'2' \cdots N' \rangle$$

$$\times \langle 1' 2' \cdots N' | \hat{Q} | 1 2 \cdots N \rangle \langle 1 2 \cdots N | n_1 n_2 \cdots \rangle. \quad (3.24)$$

Using Eq.(5) and relationship (1.145) we get

$$\langle n_1' n_2' \cdots | \hat{Q} | n_1 n_2 \cdots \rangle$$

$$= \delta_{NN'} \int d1 \cdots \int dN \langle n_1' n_2' \cdots | 1 2 \cdots N \rangle$$

$$\times \hat{Q}^{(N)}(1,2,\cdots,N) \langle 1 2 \cdots N | n_1 n_2 \cdots \rangle$$

$$= (N!)^{-1} \delta_{NN'} \int d1 \cdots \int dN \langle n_1' n_2' \cdots | \hat{\psi}^\dagger(1) \cdots \hat{\psi}^\dagger(N) | 0 \rangle$$

$$\times \hat{Q}^{(N)}(1,2,\cdots,N) \langle 0 | \hat{\psi}(N) \cdots \hat{\psi}(1) | n_1 n_2 \cdots \rangle \quad (3.25)$$

where the occupation numbers must obey the condition

$$\sum n_i = \sum n_i' = N. \quad (3.26)$$

Observing that $\hat{\psi}(N) \cdots \hat{\psi}(1) | n_1 n_2 \cdots \rangle$ has then the only non-vanishing component in $\mathcal{F}_0$, we have that

$$\langle n_1'' n_2'' \cdots | \hat{\psi}(N) \cdots \hat{\psi}(1) | n_1 n_2 \cdots \rangle = 0, \quad (3.27)$$

unless $\sum n_i'' = N'' = 0$. We can thus again reinstate formally the sum over all the states spanning $\mathcal{F}$

$$\langle n_1' n_2' \cdots | \hat{Q} | n_1 n_2 \cdots \rangle$$

$$= (N!)^{-1} \delta_{NN'} \int d1 \cdots \int dN \langle n_1' n_2' \cdots | \hat{\psi}^\dagger(1) \cdots \hat{\psi}^\dagger(N) | 0 \rangle$$

$$\times \langle 0 | \hat{Q}^{(N)}(1,\cdots,N) \hat{\psi}(N) \cdots \hat{\psi}(1) | n_1 n_2 \cdots \rangle$$

$$= (N!)^{-1} \delta_{NN'} \int d1 \cdots \int dN \sum_{n_1'', n_2'', \ldots} \langle n_1' n_2' \cdots | \hat{\psi}^\dagger(1) \cdots \hat{\psi}^\dagger(N) | n_1'' n_2'' \cdots \rangle$$

$$\times \langle n_1'' n_2'' \cdots | \hat{Q}^{(N)}(1,\cdots,N) \hat{\psi}(N) \cdots \hat{\psi}(1) | n_1 n_2 \cdots \rangle, \quad (3.28)$$

since $\hat{Q}$ is a particle number conserving operator. Deleting the resolution of identity (written in the occupation number representation) we obtain finally

$$\langle n_1' n_2' \cdots | \hat{Q} | n_1 n_2 \cdots \rangle$$

$$= \langle n_1' n_2' \cdots | (N!)^{-1} \delta_{NN'} \int d1 \cdots \int dN \; \hat{\psi}^\dagger(1) \cdots \hat{\psi}^\dagger(N)$$

$$\times \hat{Q}^{(N)}(1,\cdots,N) \psi(N) \cdots \psi(1) | n_1 n_2 \cdots \rangle. \quad (3.29)$$

3.2.1. One-particle operators

Let $\hat{Q}^{(N)}$ is a one particle operator $\hat{Q}_1^{(N)}$,

$$\hat{Q}_1^{(N)}(1,\cdots,N) = \sum_{i=1}^{N} \hat{q}_1(i). \quad (3.30)$$

The matrix elements of $\hat{Q}_1^{(N)}$ appearing in Eq.(20) may then be easily evaluated using the well know Slater rules, which give

$$\langle \{A_1 \cdots A_N\} | \hat{Q}_1^{(N)} | \{A_1 \cdots A_N\} \rangle = \sum_{i=1}^{N} \langle A_i | \hat{q}_1 | A_i \rangle, \quad (3.31a)$$

$$\langle \{A_1 A_2 \cdots A_N\} | \hat{Q}_1^{(N)} | \{B_1 A_2 \cdots A_N\} \rangle = \langle A_1 | \hat{q}_1 | B_1 \rangle \quad (3.31b)$$

$$\langle \{A_1 A_2 \cdots\cdots\} | \hat{Q}_1^{(N)} | \{B_1 B_2 \cdots\cdots\} \rangle = 0, \quad (3.31c)$$

where we assume that $A_1 \neq B_1$ and $A_2 \neq B_2$. We also note that our assumption that the spin-orbital labels are already arranged in an appropriate order in relations (31) does not restrict the generality of our considerations, since we can arbitrarily permute their order on each side of $\hat{Q}_1^{(N)}$ when the corresponding sign change is introduced.

We can now substitute from Eqs.(31) to a general relationship (20). In doing so we shall distinguish (i) the diagonal terms and (ii) the off-diagonal terms which differ in at most one spin-orbital label, since all other terms

vanish in view of (31c).

(i) The diagonal term is defined here as any term in which $B_1, B_2, \ldots, B_N$ represent some permutation of $A_1, A_2, \ldots, A_N$. Noticing that we can arbitrarily permute labels B_i among themselves (or, equivalently, A_i's) in Eq. (20), since $|\xi B_1 \cdots B_N\rangle$ is antisymmetric and X_{B_i}, $i=1,\ldots,N$ exactly anticommute, we can write the diagonal contribution as

$$\langle \Phi^{(N)} | \hat{Q}_1^{(N)} | \Psi^{(N)} \rangle_{(i)} =$$

$$= N!\,(N!)^{-2} \sum_{A_1 \cdots A_N} \left(\sum_{i=1}^{N} \langle A_i | \hat{q}_1 | A_i \rangle \right)$$

$$\times \langle \Phi^{(N)} | X_{A_1}^{\dagger} \cdots X_{A_N}^{\dagger} X_{A_N} \cdots X_{A_1} | \Psi^{(N)} \rangle, \qquad (3.32)$$

since each listed term occurs N! times in the summation over B_1 through B_N. Since, further, A_i's are the summation labels (or, equivalently, since for any A_i we can anticommute $X_{A_i}^{\dagger}$ and X_{A_i} to the first and last positions, respectively, without any sign change) all the terms represented by the second summation are in fact identical, so that

$$\langle \Phi^{(N)} | \hat{Q}_1^{(N)} | \Psi^{(N)} \rangle_{(i)} = N \cdot (N!)^{-1} \sum_{A_1 \cdots A_N} \langle A_1 | \hat{q}_1 | A_1 \rangle$$

$$\times \langle \Phi^{(N)} | X_{A_1}^{\dagger} \cdots X_{A_N}^{\dagger} X_{A_N} \cdots X_{A_1} | \Psi^{(N)} \rangle. \qquad (3.33)$$

Applying lemma 3, Eq. (21), for $k=1$ case, we get finally

$$\langle \Phi^{(N)} | \hat{Q}_1^{(N)} | \Psi^{(N)} \rangle_{(i)} = \langle \Phi^{(N)} | \sum_{A} \langle A | \hat{q} | A \rangle X_A^{\dagger} X_A | \Psi^{(N)} \rangle. \qquad (3.34)$$

(ii) The off-diagonal terms, for which the spinorbital sets $\{A_i\}$ and $\{B_i\}$, $i=1,\ldots,N$ differ in one and only one label, can be handled in the same manner. Noticing that for any such term the given labels A and B, in which the two sets differ, can occur in any position in the respective sequences $A_1, \ldots, A_N$ and $B_1, \ldots, B_N$ (giving N^2 identical terms) and that the remaining (N-1) labels $B_2, \ldots, B_N$ can equal an arbitrary permutation of $A_2, \ldots, A_N$ (yielding (N-1)! identical terms), we can write

$$\langle \Phi^{(N)} | \hat{Q}_1^{(N)} | \Psi^{(N)} \rangle_{(ii)} = N^2 (N-1)!\,(N!)^{-2} \sum_{A,B}{}' \langle A | \hat{q}_1 | B \rangle$$

$$\times \sum_{A_2 \cdots A_N} \langle \Phi^{(N)} | X_A^{\dagger} X_{A_2}^{\dagger} \cdots X_{A_N}^{\dagger} X_{A_N} \cdots X_{A_2} X_B | \Psi^{(N)} \rangle. \qquad (3.35)$$

Thus, applying again L3, Eq. (21), $k=2$ we find

$$\langle \Phi^{(N)} | \hat{Q}_1^{(N)} | \Psi^{(N)} \rangle_{(ii)} = \langle \Phi^{(N)} | \sum_{A,B}{}' \langle A | \hat{q}_1 | B \rangle X_A^{\dagger} X_B | \Psi^{(N)} \rangle. \qquad (3.36)$$

Joining, finally, the diagonal and off-diagonal cases and recalling that $|\Phi^{(N)}\rangle$ and $|\Psi^{(N)}\rangle$ are arbitrary N-electron states and that Eq. (4) must hold, we can identify the one-particle operator $\hat{Q}_1^{(N)}$ with the corresponding operator expression on the right hand sides and write

$$\hat{Q}_1^{(N)} = \sum_{A,B} \langle A | \hat{q}_1 | B \rangle X_A^{\dagger} X_B. \qquad (3.37)$$

This is the desired expression for the case of a single-particle operator.

In a similar way, we can use Eq. (29) to obtain the corresponding expression in terms of field operators. Substituting (30) into (29) we get immediately (assuming that $N=N'$)

$$\langle m_1' m_2' \cdots | \hat{Q}_1^{(N)} | m_1 m_2 \cdots \rangle$$

$$= (N!)^{-1} \sum_{i=1}^{N} \int d1 \cdots \int dN \, \langle m_1' m_2' \cdots | \hat{\psi}^{\dagger}(1) \cdots \hat{\psi}^{\dagger}(N)$$

$$\times \hat{q}_1(i)\, \hat{\psi}(N) \cdots \hat{\psi}(1) | m_1 m_2 \cdots \rangle$$

$$= (N!)^{-1} \sum_{i=1}^{N} \int d1 \cdots \int dN \, \langle m_1' m_2' \cdots | \hat{\psi}^{\dagger}(1) \cdots \hat{\psi}^{\dagger}(N)$$

$$\times \hat{q}_1(1)\, \hat{\psi}(N) \cdots \hat{\psi}(1) | m_1 m_2 \cdots \rangle$$

$$= \langle m_1' m_2' \cdots | \int d1 \, \hat{\psi}^{\dagger}(1)\, \hat{q}_1(1)\, \psi(1) | m_1 m_2 \cdots \rangle, \qquad (3.38)$$

since

$$\hat{N}\ \hat{\psi}(i_1)\,\hat{\psi}(i_2)\cdots\hat{\psi}(i_k)\,|n_1 n_2\cdots\rangle = (N-k)\ \hat{\psi}(i_1)\,\hat{\psi}(i_2)\cdots\hat{\psi}(i_k)\,|n_1 n_2\cdots\rangle, \quad (3.39)$$

in complete analogy with (22), so that an analogous lemma to that given by Eq. (21), but with the annihilators replaced by field operators, applies. Thus, for the same reasons as above, we can write

$$\hat{Q}_1 = \int d\vec{x}\ \hat{\psi}^\dagger(\vec{x})\,\hat{q}_1(\vec{x})\,\psi(\vec{x}). \quad (3.40)$$

This expression, as well as the corresponding expression (37), is in fact independent of the particle number N. Thus, in the second quantization formalism, a given one-particle operator has the same form independently of the particle number of the system considered. We can thus drop the particle number index as we did in the latter expression (40).

Both expressions (37) and (40) are very simply related: substituting the definition of field operators, Eqs. (1.136) into (40) gives immediately

$$\begin{aligned}\hat{Q}_1 &= \int d\vec{x}\ \psi^\dagger(\vec{x})\,\hat{q}_1(\vec{x})\,\psi(\vec{x}) \\ &= \sum_{A,B}\int d\vec{x}\,\langle A|\vec{x}\rangle\,\hat{q}_1(\vec{x})\,\langle\vec{x}|B\rangle\ X_A^\dagger X_B \\ &= \sum_{A,B}\langle A|\hat{q}_1|B\rangle\ X_A^\dagger X_B\,. \end{aligned} \quad (3.41)$$

We also note that (40) can be easily generalized to non-local operators*, if desirable.

3.2.2. Two-particle operators

Consider, next, a general two-particle operator $\hat{Q}_2$, whose N-particle component in the coordinate representation can be written as

$$\hat{Q}_2^{(N)}(1,\dots,N) = \sum_{i<j}\hat{q}_2(i,j) = \frac{1}{2}\sum_{i,j=1}^{N}{}'\,\hat{q}_2(i,j), \quad (3.42)$$

where we indicate by the prime the absence of diagonal self-interacting terms. The configuration

* Indeed, in the non-local case $\hat{Q}_1 = \int d\vec{x}\int d\vec{y}\ \hat{\psi}^\dagger(\vec{x})\,\hat{q}_1(\vec{x}|\vec{y})\,\hat{\psi}(\vec{y})$

matrix elements are given again by Slater rules, which we can summarize as follows

$$\langle\{A_1\cdots A_N\}|\hat{Q}_2^{(N)}|\{A_1\cdots A_N\}\rangle = \frac{1}{2}\sum_{i,j=1}^{N}\langle A_i A_j|\hat{q}_2|A_i A_j\rangle_{\mathcal{A}} \quad (3.43a)$$

$$\langle\{A_1 A_2\cdots A_N\}|\hat{Q}_2^{(N)}|\{B_1 A_2\cdots A_N\}\rangle = \sum_{i=1}^{N}\langle A_1 A_i|\hat{q}_2|B_1 A_i\rangle_{\mathcal{A}} \quad (3.43b)$$

$$\langle\{A_1 A_2 A_3\cdots A_N\}|\hat{Q}_2^{(N)}|\{B_1 B_2 A_3\cdots A_N\}\rangle = \langle A_1 A_2|\hat{q}_2|B_1 B_2\rangle_{\mathcal{A}} \quad (3.43c)$$

$$\langle\{A_1 A_2 A_3\cdots\}|\hat{Q}_2^{(N)}|\{B_1 B_2 B_3\cdots\}\rangle = 0, \quad (3.43d)$$

where we again assume that $A_i \neq B_i$, $i=1,2,3$ and where we introduced the antisymmetrized matrix element defined as

$$\langle AB|\hat{q}_2|CD\rangle_{\mathcal{A}} \equiv \langle AB|\hat{q}_2|CD\rangle - \langle AB|\hat{q}_2|DC\rangle. \quad (3.44)$$

Obviously

$$\langle AB|\hat{q}_2|CD\rangle_{\mathcal{A}} = \langle BA|\hat{q}_2|DC\rangle_{\mathcal{A}} = -\langle AB|\hat{q}_2|DC\rangle_{\mathcal{A}}, \quad (3.45)$$

so that

$$\langle AA|\hat{q}_2|CD\rangle_{\mathcal{A}} = \langle AB|\hat{q}_2|CC\rangle_{\mathcal{A}} = \langle AA|\hat{q}_2|CC\rangle_{\mathcal{A}} = 0. \quad (3.46)$$

For this reason the summations in (43a) and (43b) can be unrestricted summations over the whole range of spinorbital labels. We also recall that, just as in the one-particle case, the spin-orbital labels in configurations on the left hand side of Eqs. (43) can be arbitrarily permuted in the bra or in the ket when the corresponding sign change is introduced if the overall permutation is odd.

We now use again the above given Slater rules to simplify the N-particle matrix elements in Eq. (20), considering separately the three different cases implied by Eqs. (43a-c), since the last case [Eq. (43d)] yields a trivial vanishing contribution.

(i) In the diagonal case the summations over B_i, $i=1,\dots,N$ in Eq. (20) yield again N! identical

terms, just as in the one-particle case, so that

$$\langle \Phi^{(N)} | \hat{Q}_2^{(N)} | \Psi^{(N)} \rangle_{(i)} = (N!)^{-1} \sum_{A_1 \cdots A_N} \left(\tfrac{1}{2} \sum_{i,j=1}^{N} \langle A_i A_j | \hat{q}_2 | A_i A_j \rangle_{\mathcal{A}} \right) \times \langle \Phi^{(N)} | X_{A_1}^\dagger \cdots X_{A_N}^\dagger X_{A_N} \cdots X_{A_1} | \Psi^{(N)} \rangle . \quad (3.47)$$

Again, in complete analogy with the one-particle case, the labels A_i and A_j can be identified with A_1 and A_2 obtaining in this case $N(N-1)$ identical terms, so that

$$\langle \Phi^{(N)} | \hat{Q}_2^{(N)} | \Psi^{(N)} \rangle_{(i)} = \tfrac{1}{2}[(N-2)!]^{-1} \sum_{A_1, A_2} \langle A_1 A_2 | \hat{q}_2 | A_1 A_2 \rangle_{\mathcal{A}} \times \sum_{A_3 \cdots A_N} \langle \Phi^{(N)} | X_{A_1}^\dagger \cdots X_{A_N}^\dagger X_{A_N} \cdots X_{A_1} | \Psi^{(N)} \rangle . \quad (3.48)$$

Applying, finally, L3, Eq. (21), $k=2$, to the last sum we obtain

$$\langle \Phi^{(N)} | \hat{Q}_2^{(N)} | \Psi^{(N)} \rangle_{(i)} = \langle \Phi^{(N)} | \tfrac{1}{2} \sum_{A_1, A_2} \langle A_1 A_2 | \hat{q}_2 | A_1 A_2 \rangle_{\mathcal{A}} X_{A_1}^\dagger X_{A_2}^\dagger X_{A_2} X_{A_1} | \Psi^{(N)} \rangle . \quad (3.49)$$

(ii) Considering, next, the terms originating from configurations differing in one and only one spin-orbital label, designated as A_1 and B_1, we get again $N^2(N-1)$ identical terms as in the one-particle case (ii), so that

$$\langle \Phi^{(N)} | \hat{Q}_2^{(N)} | \Psi^{(N)} \rangle_{(ii)} = [(N-1)!]^{-1} {\sum_{A_1, B_1}}' \sum_{A_2 \cdots A_N} \left(\sum_{i=1}^{N} \langle A_1 A_i | \hat{q}_2 | B_1 A_i \rangle_{\mathcal{A}} \right) \times \langle \Phi^{(N)} | X_{A_1}^\dagger X_{A_2}^\dagger \cdots X_{A_N}^\dagger X_{A_N} \cdots X_{A_2} X_{B_1} | \Psi^{(N)} \rangle$$

$$= [(N-2)!]^{-1} {\sum_{A_1, B_1}}' \sum_{A_2} \langle A_1 A_2 | \hat{q}_2 | B_1 A_2 \rangle_{\mathcal{A}} \times \sum_{A_3 \cdots A_N} \langle \Phi^{(N)} | X_{A_1}^\dagger X_{A_2}^\dagger X_{A_3}^\dagger \cdots X_{A_N}^\dagger X_{A_N} \cdots X_{A_3} X_{A_2} X_{B_1} | \Psi^{(N)} \rangle . \quad (3.50)$$

where in the last step we have moved each A_i to the second position from the left and from the right (no sign change) and relabeled accordingly the summation indices, obtaining $(N-1)$ identical terms (since the $i=1$ term vanishes). Thus,

applying L3, Eq. (21), $k=2$ to the last term on the right hand side gives

$$\langle \Phi^{(N)} | \hat{Q}_2^{(N)} | \Psi^{(N)} \rangle_{(ii)} = \langle \Phi^{(N)} | {\sum_{A_1, B_1}}' \sum_{A_2} \langle A_1 A_2 | \hat{q}_2 | B_1 A_2 \rangle_{\mathcal{A}} \times X_{A_1}^\dagger X_{A_2}^\dagger X_{A_2} X_{B_1} | \Psi^{(N)} \rangle \quad (3.51)$$

(iii) We consider, finally, the terms originating from configurations differing in two spin-orbital labels, say A_i, A_j, B_k and B_ℓ. We can again identify these with A_1, A_2, B_1 and B_2 obtaining $\frac{1}{2}N(N-1)$ identical terms in each case (i.e., altogether $\frac{1}{4}N^2(N-1)^2$ terms), while the summations over the remaining $B_3, \ldots, B_N$ labels yield $(N-2)!$ identical terms, so that

$$\langle \Phi^{(N)} | \hat{Q}_2^{(N)} | \Psi^{(N)} \rangle_{(iii)} = \tfrac{1}{4}[(N-2)!]^{-1} {\sum_{\substack{A_1 A_2 \\ B_1 B_2}}}' \langle A_1 A_2 | \hat{q}_2 | B_1 B_2 \rangle_{\mathcal{A}} \times \sum_{A_3 \cdots A_N} \langle \Phi^{(N)} | X_{A_1}^\dagger X_{A_2}^\dagger X_{A_3}^\dagger \cdots X_{A_N}^\dagger X_{A_N} \cdots X_{A_3} X_{B_2} X_{B_1} | \Psi^{(N)} \rangle \quad (3.52)$$

and using L3, Eq. (21), $k=2$ we get

$$\langle \Phi^{(N)} | \hat{Q}_2^{(N)} | \Psi^{(N)} \rangle_{(iii)} = \langle \Phi^{(N)} | \tfrac{1}{4} {\sum_{\substack{A_1 A_2 \\ B_1 B_2}}}' \langle A_1 A_2 | \hat{q}_2 | B_1 B_2 \rangle_{\mathcal{A}} X_{A_1}^\dagger X_{A_2}^\dagger X_{B_2} X_{B_1} | \Psi^{(N)} \rangle . \quad (3.53)$$

The prime on the summation symbol indicates that A_1 and A_2 are distinct from B_1 and B_2.

The above obtained results can be re-written in several alternative forms using the symmetry properties of antisymmetrized matrix elements, Eqs. (45), the anticommutation relations (1.96), and the fact that all spinorbital labels are summation labels. For the same reasons we can also express them in terms of non-antisymmetrized matrix elements since

$$\sum_{ABCD} \langle AB | \hat{q}_2 | CD \rangle X_A^\dagger X_B^\dagger X_D X_C = \sum_{ABCD} \langle AB | \hat{q}_2 | DC \rangle X_A^\dagger X_B^\dagger X_C X_D$$

$$= -\sum_{ABCD} \langle AB|\hat{q}_2|DC\rangle\, X_A^\dagger X_B^\dagger X_D X_C\,, \qquad (3.54)$$

so that

$$\sum_{ABCD} \langle AB|\hat{q}_2|CD\rangle_{ik}\, X_A^\dagger X_B^\dagger X_D X_C$$
$$= 2\sum_{ABCD} \langle AB|\hat{q}_2|CD\rangle\, X_A^\dagger X_B^\dagger X_D X_C\,. \qquad (3.55)$$

Similar results will hold for various restricted sums as well.

Collecting, thus, the three contributions (49), (51) and (53), we can write using (55)

$$\langle \Phi^{(N)}|\hat{Q}_2^{(N)}|\Psi^{(N)}\rangle = \langle \Phi^{(N)}|\sum_{A_1A_2} \langle A_1A_2|\hat{q}_2|A_1A_2\rangle X_{A_1}^\dagger X_{A_2}^\dagger X_{A_2} X_{A_1}$$
$$+2{\sum_{A_1B_1}}' \sum_{A_2} \langle A_1A_2|\hat{q}_2|B_1A_2\rangle X_{A_1}^\dagger X_{A_2}^\dagger X_{B_2} X_{B_1}$$
$$+\tfrac{1}{2}{\sum_{\substack{A_1B_1\\A_2B_2}}}' \langle A_1A_2|\hat{q}_2|B_1B_2\rangle X_{A_1}^\dagger X_{A_2}^\dagger X_{B_2} X_{B_1}$$
$$= \langle \Phi^{(N)}|\tfrac{1}{2}\sum_{\substack{A_1A_2\\B_1B_2}} \langle A_1A_2|\hat{q}_2|B_1B_2\rangle X_{A_1}^\dagger X_{A_2}^\dagger X_{B_2} X_{B_1}|\Psi^{(N)}\rangle\,, \qquad (3.56)$$

Since, again, $|\Phi^{(N)}\rangle$ and $|\Psi^{(N)}\rangle$ are arbitrary states and the operator on the right hand side is N-independent, we can express an arbitrary two-particle operator (42) in the second quantization formalism in one of the following forms

$$\hat{Q}_2 = (2!)^{-2} \sum_{ABCD} \langle AB|\hat{q}_2|CD\rangle_{ik}\, X_A^\dagger X_B^\dagger X_D X_C$$
$$= (2!)^{-1} \sum_{ABCD} \langle AB|\hat{q}_2|CD\rangle\, X_A^\dagger X_B^\dagger X_D X_C$$
$$= (2!)^{-1} \sum_{ABCD} \langle AB|\hat{q}_2|CD\rangle\, X_B^\dagger X_A^\dagger X_C X_D$$
$$= \sum_{A<B} \sum_{C<D} \langle AB|\hat{q}_2|CD\rangle_{ik}\, X_A^\dagger X_B^\dagger X_D X_C \qquad (3.57)$$

or similar equivalent forms. We note the reversed

order for either the annihilators or the creators as compared with the corresponding order in the two-particle matrix element. This characteristic is implied already by the general result (20) or (29).

The corresponding result expressed through the field operators follows very simply from (29). Indeed, substituting into (29) from (42)

$$\langle m_1' m_2' \cdots |\hat{Q}_2^{(N)}|m_1 m_2 \cdots\rangle$$
$$= (N!)^{-1} \sum_{i<j} \int d1 \cdots \int dN\, \langle m_1' m_2' \cdots |\hat{\psi}^\dagger(1) \cdots \hat{\psi}^\dagger(N)$$
$$\times \hat{q}_2(i,j)\, \hat{\psi}(N) \cdots \hat{\psi}(1)|m_1 m_2 \cdots\rangle$$
$$= (N!)^{-1} \tfrac{1}{2}N(N-1) \int d1 \cdots \int dN\, \langle m_1' m_2' \cdots |\hat{\psi}^\dagger(1) \cdots \hat{\psi}^\dagger(N)$$
$$\times \hat{q}_2(1,2)\, \psi(N) \cdots \psi(1)|m_1 m_2 \cdots\rangle \qquad (3.58)$$
$$= \langle m_1' m_2' \cdots |\tfrac{1}{2}\int d1 \int d2\, \hat{\psi}^\dagger(1) \psi^\dagger(2) \hat{q}_2(1,2) \psi(2) \psi(1)|m_1 m_2 \cdots\rangle,$$

where we have used repeatedly Eq. (39) in the last step (similarly as in the one-particle case), we find

$$\hat{Q}_2 = \tfrac{1}{2}\int d\vec{x} \int d\vec{x}'\, \psi^\dagger(\vec{x}) \psi^\dagger(\vec{x}') \hat{q}_2(\vec{x},\vec{x}') \psi(\vec{x}') \psi(\vec{x}). \qquad (3.59)$$

Thus, the preceding result (57) could be most simply obtained by deriving first the field operator form (59) and by making a transition to the discrete form using definition (1.136), the last step being straightforward as the single-particle case, Eq. (41), indicates. In this way it is also easy to derive a general result for the r-particle operator. Clearly

$$\hat{Q}_r = \tfrac{1}{r!} \int d\vec{x}_1 \cdots \int d\vec{x}_r\, \hat{\psi}^\dagger(\vec{x}_1) \cdots \hat{\psi}^\dagger(\vec{x}_r) \hat{q}_r(\vec{x}_1, \cdots, \vec{x}_r) \psi(\vec{x}_r) \cdots \psi(\vec{x}_1)$$
$$= \tfrac{1}{r!} \sum_{\substack{A_1 \cdots A_r\\B_1 \cdots B_r}} \langle A_1 A_2 \cdots A_r|\hat{q}_r|B_1 B_2 \cdots B_r\rangle X_{A_1}^\dagger X_{A_2}^\dagger \cdots X_{A_r}^\dagger X_{B_r} \cdots X_{B_2} X_{B_1}$$

$$= (r!)^{-2} \sum_{\substack{A_1 \cdots A_r \\ B_1 \cdots B_r}} \langle A_1 \cdots A_r | \hat{q}_r | B_1 \cdots B_r \rangle_{\mathcal{A}} X^\dagger_{A_1} \cdots X^\dagger_{A_r} X_{B_r} \cdots X_{B_1}, \quad (3.60)$$

a general antisymmetrized matrix element being defined as follows

$$\langle A_1 \cdots A_r | \hat{q}_r | B_1 \cdots B_r \rangle_{\mathcal{A}} = \sum_{P \in \mathcal{S}_r} (-1)^p \langle A_1 \cdots A_r | \hat{q}_r | B_{p_1} \cdots B_{p_r} \rangle, \quad (3.61)$$

where $(-1)^p$ is the parity of the permutation $P: i \to p_i$ and the sum extends over all $r!$ permutations $P \in \mathcal{S}_r$. For a non-local operator we can write

$$\hat{Q}_r = (r!)^{-1} \int d\vec{x}_1 \cdots \int d\vec{x}_r \int d\vec{x}'_1 \cdots \int d\vec{x}'_r \, \hat{\psi}^\dagger(\vec{x}_1) \cdots \hat{\psi}^\dagger(\vec{x}_r)$$

$$\times \langle \vec{x}_1 \cdots \vec{x}_r | \hat{Q}_r | \vec{x}'_1 \cdots \vec{x}'_r \rangle \hat{\psi}(\vec{x}'_1) \cdots \hat{\psi}(\vec{x}'_r). \quad (3.62)$$

3.3. Operator representation and Wick's theorem.

As an illustration of Wick's theorem we will now briefly outline an alternative, "a posteriori", proof of the second quantized form for the operators by showing that it yields the same matrix elements as the first quantized form, given by the Slater rules.

Consider, first, some one-particle operator, say the operator $\hat{Z}$,

$$\hat{Z}^{(N)} = \sum_{i=1}^{N} \hat{z}(i), \quad (3.63)$$

whose second quantized form is

$$\hat{Z} = \sum_{AB} \langle A | \hat{z} | B \rangle \, X^\dagger_A X_B \quad (3.64)$$

with matrix elements defined by (7). In order to demonstrate the equivalence of both forms (63) and (64) we will show that they yield the same

matrix elements between arbitrary states.

Since either form can only yield a non-vanishing matrix element between states with the same particle number, we consider a general element

$$\langle \{A_1 \cdots A_N\} | \hat{Z} | \{B_1 \cdots B_N\} \rangle$$

$$= \sum_{AB} \langle A | \hat{z} | B \rangle \langle 0 | \hat{A}_N \cdots \hat{A}_1 \hat{A}^\dagger \hat{B} \hat{B}^\dagger_1 \cdots \hat{B}^\dagger_N | 0 \rangle, \quad (3.65)$$

where we introduced a shorthand notation

$$X_A \equiv \hat{A} \quad , \quad X^\dagger_A \equiv \hat{A}^\dagger. \quad (3.66)$$

Applying R6, Eq. (2.70), we get a non-vanishing contribution only if (i) all the operators in the vacuum mean value are contracted, and (ii) all the contractions are of the type (iv), Eq. (2.19) [or, in fact, Eq. (2.20)].

We thus have to contract the operator $X_B \equiv \hat{B}$ with some creation operator on its right hand side, i.e. $\hat{B}\hat{B}^\dagger_j$, while $\hat{A}^\dagger$ must be paired with some annihilator on its left hand side, say $\hat{A}_i$. The remaining $(N-1)$ annihilators $\hat{A}_k$ $(k \neq i)$ must then be paired, one by one, with the remaining creators $\hat{B}^\dagger_l$, $l \neq j$. We see immediately that, in order to get a non-vanishing result, both configurations can differ in at most one spin-orbital: this differing spin-orbital label in the bra is then contracted with $\hat{A}^\dagger$ and the one in the ket with $\hat{B}$, while each of the remaining $(N-1)$ annihilators (with fixed labels) is contracted with an identically labeled creator. We thus have

$$\langle \{A_1 \cdots A_N\} | \hat{Z} | \{B_1 B_2 \cdots B_N\} \rangle = 0 \quad (3.67)$$

if $\{A_i\}$ and $\{B_i\}$ sets differ in more than one spin-orbital.

When $\{A_i\}$ and $\{B_i\}$ differ in one and only one spinorbital, we can assume that

$A_1 \neq B_1$, while $A_i = B_i$ for $i = 2, \dots, N$, so that

$$\langle \{A_1 A_2 \cdots A_N\} | \hat{Z} | \{B_1 A_2 \cdots A_N\} \rangle$$

$$= \sum_{AB} \langle A|\hat{z}|B\rangle \langle 0| A_N \cdots A_2 A_1 A^\dagger B B_1^\dagger A_2^\dagger \cdots A_N^\dagger |0\rangle$$

$$= \sum_{AB} \langle A|\hat{z}|B\rangle \langle A_1|A\rangle \langle B|B_1\rangle \left(\prod_{i=2}^{N} \langle A_i|A_i\rangle \right)$$

$$= \langle A_1|\hat{z}|B_1\rangle . \qquad (3.68)$$

Finally, in the diagonal case, we can assume that $A_i = B_i$, $i = 1, \dots, N$, so that the operators $X_A^\dagger$ and X_B can be contracted with any identically labeled operators in the sets $\{A_i\}$ and $\{B_i\}$, so that

$$\langle \{A_1 A_2 \cdots A_N\} | \hat{Z} | \{A_1 A_2 \cdots A_N\} \rangle$$

$$= \sum_{AB} \langle A|\hat{z}|B\rangle \sum_{k=1}^{N} \langle 0| A_N \cdots A_k \cdots A_1 A^\dagger B A_1^\dagger \cdots A_k^\dagger \cdots A_N^\dagger |0\rangle$$

$$= \sum_{AB} \langle A|\hat{z}|B\rangle \sum_{k=1}^{N} \langle A_k|A\rangle \langle B|A_k\rangle = \sum_{k} \langle A_k|\hat{z}|A_k\rangle \qquad (3.69)$$

We have thus recovered the well known Slater rules (31) for the case of a one-particle operator.

In a similar way we can consider a general two-particle operator, say

$$\hat{V} = \tfrac{1}{4} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle X_A^\dagger X_B^\dagger X_D X_C , \qquad (3.70)$$

and evaluate a general matrix element

$$\langle \{A_1 \cdots A_N\} | \hat{V} | \{B_1 \cdots B_N\} \rangle$$

$$= \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \langle 0| \hat{A}_N \cdots \hat{A}_1 \hat{A}^\dagger \hat{B}^\dagger \hat{D} \hat{C} \hat{B}_1^\dagger \cdots \hat{B}_N^\dagger |0\rangle , \qquad (3.71)$$

where we use again a simplified operator notation (66). Reasoning as in the one-particle case above we see that $\hat{A}^\dagger, \hat{B}^\dagger$ ($\hat{D}, \hat{C}$) must be contracted with some annihilators A_i, A_j (creators B_k, B_e) on their left (right) hand side and the remaining (N-2)

annihilators must be contracted, one by one, with the remaining (N-2) creators, in order to get a non-vanishing fully contracted contribution. Again, in order to get a non-vanishing result, at least (N-2) spin-orbitals in the bra must be identical with (N-2) ket spinorbitals, implying the validity of (43d).

Considering, briefly, the remaining cases, we first assume, without any loss of generality, that the two configurations appearing in (71) differ in the first two spinorbitals, so that

$$\langle \{A_1 A_2 A_3 \cdots A_N\} | \hat{V} | \{B_1 B_2 A_3 \cdots A_N\} \rangle \qquad (3.72)$$

$$= \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \langle 0| A_N \cdots A_3 \boxed{A_2 A_1 A^\dagger B^\dagger D C B_1^\dagger B_2^\dagger} A_3^\dagger \cdots A_N^\dagger |0\rangle$$

where the possible non-vanishing contraction schemes for the operators enclosed in a dashed rectangle are

$$\boxed{A_2 A_1 A^\dagger B^\dagger D C B_1^\dagger B_2^\dagger} = A_2 A_1 A^\dagger B^\dagger D C B_1^\dagger B_2^\dagger$$

$$+ A_2 A_1 A^\dagger B^\dagger D C B_1^\dagger B_2^\dagger$$

$$+ A_2 A_1 A^\dagger B^\dagger D C B_1^\dagger B_2^\dagger$$

$$+ A_2 A_1 A^\dagger B^\dagger D C B_1^\dagger B_2^\dagger \qquad (3.73)$$

$$= (\langle A_1|A\rangle \langle A_2|B\rangle - \langle A_1|B\rangle \langle A_2|A\rangle) \times (\langle C|B_1\rangle \langle D|B_2\rangle - \langle C|B_2\rangle \langle D|B_1\rangle).$$

Thus

$$\langle \{A_1 A_2 A_3 \cdots A_N\} | \hat{V} | \{B_1 B_2 A_3 \cdots A_N\} \rangle$$

$$= \tfrac{1}{2} \left(\langle A_1 A_2|\hat{v}|B_1 B_2\rangle - \langle A_1 A_2|\hat{v}|B_2 B_1\rangle - \langle A_2 A_1|\hat{v}|B_1 B_2\rangle + \langle A_2 A_1|\hat{v}|B_2 B_1\rangle \right)$$

$$= \langle A_1 A_2|\hat{v}|B_1 B_2\rangle_A \qquad (3.74)$$

in agreement with (43c).

Second, assume that $A_1 \neq B_1$ and $A_i = B_i$, $i = 2, \dots, N$, so that only one of the operators $\hat{A}^\dagger, \hat{B}^\dagger$ have to

be always contracted with $\hat{A}_1$, while the second one can now be contracted with any remaining annihilator $\hat{A}_2 \cdots \hat{A}_N$ and, similarly, for the contractions involving $\hat{D}$ and $\hat{C}$ operators, so that

$$\langle \{A_1 A_2 \cdots A_N\} | \hat{V} | \{B_1 A_2 \cdots A_N\} \rangle \qquad (3.75)$$

$$= \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \sum_{k=2}^{N} \langle 0| A_N \cdots A_k \cdots A_2 A_1 \, A^\dagger B^\dagger D C \, B_1^\dagger A_2^\dagger \cdots A_k^\dagger \cdots A_N^\dagger |0\rangle$$

where for the operators originating from $\hat{V}$ we have again the same four possibilities as in the case (3.73) except that $A_2 = B_2$ are now ranging over the remaining $(N-1)$ operators. Thus, schematically

$$\underline{\underline{A^\dagger B^\dagger D C}} = \underline{\underline{A^\dagger B^\dagger D C}} + \underline{\underline{A^\dagger B^\dagger D C}} + \underline{\underline{A^\dagger B^\dagger D C}} + \underline{\underline{A^\dagger B^\dagger D C}} \qquad (3.76)$$

so that

$$\langle \{A_1 A_2 \cdots A_N\} | \hat{V} | \{B_1 A_2 \cdots A_N\} \rangle$$

$$= \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \sum_{k=2}^{N} \left(\langle A_1|A\rangle\langle A_k|B\rangle - \langle A_1|B\rangle\langle A_k|A\rangle \right)$$

$$\times \left(\langle C|B_1\rangle\langle D|A_k\rangle - \langle C|A_k\rangle\langle D|B_1\rangle \right)$$

$$= \sum_{k=2}^{N} \langle A_1 A_k | \hat{r}_{12} | B_1 A_k \rangle_A = \sum_{k=1}^{N} \langle A_1 A_k | \hat{v} | B_1 A_k \rangle_A , \qquad (3.77)$$

which is identical with (43b).

Finally, in the diagonal case $A_i = B_i$, $i = 1, \ldots, N$, even the operators A_1 and $B_1^\dagger$ in (75) can be any other corresponding operators $A_\ell, A_\ell^\dagger$, $\ell \neq k$, so that

$$\langle \{A_1 \cdots A_N\} | \hat{V} | \{A_1 \cdots A_N\} \rangle$$

$$= \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \sum_{\substack{k,\ell=1 \\ (k>\ell)}}^{N} \left(\langle A_k|B\rangle\langle A_\ell|A\rangle - \langle A_k|A\rangle\langle A_\ell|B\rangle \right)$$

$$\times \left(\langle C|A_\ell\rangle\langle D|A_k\rangle - \langle C|A_k\rangle\langle D|A_\ell\rangle \right)$$

$$= \sum_{\substack{k,\ell=1 \\ (k>\ell)}}^{N} \langle A_k A_\ell | \hat{r}_{12} | A_k A_\ell \rangle_A = \tfrac{1}{2} \sum_{k,\ell=1}^{N} \langle A_k A_\ell | \hat{v} | A_k A_\ell \rangle , \qquad (3.78)$$

in agreement with (43a).

The reader have certainly noticed that the result (76) could be easily extended to a general p-particle case, in which case we would get $(p!)^2$ contraction schemes. These would result in $p!$ antisymmetrized matrix elements for each possible choice of the p spinorbital labels from the bra and from the ket. In this way, the above considerations could be extended to a general case, if desired.

3.4. Spin-independent operators

We now examine the second quantized form of spin-independent operators. The most important example of such an operator in the atomic and molecular electronic structure problems is the Hamiltonian (6), which describes such systems when all the relativistic and magnetic effects are neglected. We consider in this section this operator. However, the same procedure may be applied to any spin-independent operator.

In 3.1 we have seen how the spin-orbital matrix elements of one- and two-particle spin independent operators simplify when the spinorbitals $|D_i\rangle$ are expressed as products of orbitals $|d_i\rangle$ and pure spin functions $|\delta_i\rangle$, Eqs. (9-11). The creation and annihilation operators associated with the spinorbital $|D_i\rangle$ may be accordingly designated by either of the following equivalent forms

$$X_{D_i}^\dagger \equiv X_{d_i \delta_i}^\dagger$$
$$X_{D_i} \equiv X_{d_i \delta_i} \qquad (3.79)$$

so that, for example

$$[X_{c\gamma}, X_{d\delta}^\dagger]_+ = \langle c\gamma | d\delta \rangle = \langle c|d\rangle\langle \gamma|\delta\rangle . \qquad (3.80)$$

Similarly, any summation of the spinorbital states $\sum_{D_i}$

represents in fact two separate summations $\sum_{d_i \delta_i}$ or $\sum_{d_i}\sum_{\delta_i}$.

Assuming that a single particle operator $\hat{Z}$, Eq. (64), is spin-independent, we can use the property (12), (13) to write it as

$$\begin{aligned}\hat{Z} &= \sum_{CD} \langle C|\hat{z}(1)|D\rangle\, X_C^\dagger X_D \\ &= \sum_{cd}\sum_{\gamma\delta} \langle\gamma|\delta\rangle\langle c|\hat{z}|d\rangle\, X_{c\gamma}^\dagger X_{d\delta} \\ &= \sum_{cd} \langle c|\hat{z}|d\rangle \sum_{\gamma} X_{c\gamma}^\dagger X_{d\gamma} \end{aligned} \tag{3.81}$$

where, following our notational convention (cf. Sec. 3.1), we set $|C\rangle = |c\rangle|\gamma\rangle$ and $|D\rangle = |d\rangle|\delta\rangle$. The last summation over γ extends over the two possible spin orientations. Using a conventional notation for spin-up and down functions,*)

$$\begin{aligned}|\alpha\rangle &= |\tfrac{1}{2}\rangle \equiv |\tfrac{1}{2},\tfrac{1}{2}\rangle \\ |\beta\rangle &= |-\tfrac{1}{2}\rangle \equiv |\tfrac{1}{2},-\tfrac{1}{2}\rangle ,\end{aligned} \tag{3.82}$$

we can write this last summation explicitly, namely

$$\hat{Z} = \sum_{cd} \langle c|\hat{z}|d\rangle \left(X_{c\alpha}^\dagger X_{d\alpha} + X_{c\beta}^\dagger X_{d\beta}\right). \tag{3.83}$$

Similarly, for a two-particle spin-independent operator $\hat{V}$, Eq. (70), we can write

$$\begin{aligned}\hat{V} &= \tfrac{1}{2}\sum_{C_1C_2D_1D_2} \langle C_1C_2|\hat{v}(\vec{r}_1,\vec{r}_2)|D_1D_2\rangle\, X_{C_1}^\dagger X_{C_2}^\dagger X_{D_2} X_{D_1} \\ &= \tfrac{1}{2}\sum_{c_1c_2d_1d_2}\sum_{\gamma_1\gamma_2\delta_1\delta_2} \langle\gamma_1|\delta_1\rangle\langle\gamma_2|\delta_2\rangle\langle c_1c_2|\hat{v}|d_1d_2\rangle\, X_{c_1\gamma_1}^\dagger X_{c_2\gamma_2}^\dagger X_{d_2\delta_2} X_{d_1\delta_1} \\ &= \tfrac{1}{2}\sum_{c_1c_2d_1d_2} \langle c_1c_2|\hat{v}|d_1d_2\rangle \sum_{\gamma\delta} X_{c_1\gamma}^\dagger X_{c_2\delta}^\dagger X_{d_2\delta} X_{d_1\gamma} \end{aligned} \tag{3.84}$$

or, explicitly

$$\begin{aligned}\hat{V} = \tfrac{1}{2}\sum_{c_1c_2d_1d_2} \langle c_1c_2|\hat{v}|d_1d_2\rangle \big(& X_{c_1\alpha}^\dagger X_{c_2\alpha}^\dagger X_{d_2\alpha} X_{d_1\alpha} + X_{c_1\alpha}^\dagger X_{c_2\beta}^\dagger X_{d_2\beta} X_{d_1\alpha} \\ & + X_{c_1\beta}^\dagger X_{c_2\alpha}^\dagger X_{d_2\alpha} X_{d_1\beta} + X_{c_1\beta}^\dagger X_{c_2\beta}^\dagger X_{d_2\beta} X_{d_1\beta}\big).\end{aligned} \tag{3.85}$$

*) For this reason we avoid in dealing with general spin-orbitals the first two letters of the alphabet.

3.5. Canonical transformation of operators

A general second quantized form (60) of an arbitrary operator is invariant under canonical transformations (1.126). Indeed, considering a transformation

$$X_A = \sum_B d_{AB}\, \tilde{X}_B , \tag{3.86}$$

where

$$\mathbb{D} = \| d_{AB} \| \tag{3.87}$$

is unitary, we get for an arbitrary one-particle operator, say $\hat{Z}$

$$\begin{aligned}\hat{Z} &= \sum_{AB} \langle A|\hat{z}|B\rangle\, X_A^\dagger X_B \\ &= \sum_{AB} \langle A|\hat{z}|B\rangle \sum_{CD} d_{AC}^*\, d_{BD}\, \tilde{X}_C^\dagger \tilde{X}_D \\ &= \sum_{CD} \Big(\sum_{AB} d_{AC}^* \langle A|\hat{z}|B\rangle\, d_{BD} \Big)\, \tilde{X}_C^\dagger \tilde{X}_D . \end{aligned} \tag{3.88}$$

The matrix elements of $\hat{Z}$, Eq. (7),

$$z_{AB} \equiv \langle A|\hat{z}|B\rangle , \tag{3.89}$$

constitute the "A"-representation of $\hat{Z}$,

$$\mathbb{Z} = \| z_{AB} \| . \tag{3.90}$$

Designating the transformed matrix elements enclosed in the parentheses in (88) as

$$\tilde{z}_{CD} \equiv \sum_{AB} d_{AC}^*\, z_{AB}\, d_{BD} = (\mathbb{D}^\dagger \mathbb{Z}\, \mathbb{D})_{CD} , \tag{3.91}$$

we see immediately that they are related to the matrix elements z_{AB} by a unitary transformation $\mathbb{D}$. Since the single particle states are transformed in the same way as the corresponding creation operators, Eqs. (1.127) and (1.133), we can regard the matrix elements (91) as the "$\tilde{A}$"-representation of $\hat{Z}$. In view of (1.134) we can also write

represents in fact two separate summations $\sum_{di\,\delta i}$ or $\sum_{di}\sum_{\delta i}$.

Assuming that a single particle operator $\hat{Z}$, Eq.(64), is spin-independent, we can use the property (12), (13) to write it as

$$\begin{aligned}\hat{Z} &= \sum_{CD} \langle c|\hat{z}(*)|D\rangle\, X_C^\dagger X_D \\ &= \sum_{cd}\sum_{\gamma\delta} \langle\gamma|\delta\rangle\langle c|\hat{z}|d\rangle\, X_{c\gamma}^\dagger X_{d\delta} \\ &= \sum_{cd} \langle c|\hat{z}|d\rangle \sum_{\gamma} X_{c\gamma}^\dagger X_{d\gamma} \end{aligned} \tag{3.81}$$

where, following our notational convention (cf, Sec. 3.1), we set $|C\rangle = |c\rangle|\gamma\rangle$ and $|D\rangle = |d\rangle|\delta\rangle$. The last summation over γ extends over the two possible spin orientations. Using a conventional notation for spin-up and down functions,*)

$$\begin{aligned}|\alpha\rangle &= |\tfrac{1}{2}\rangle \equiv |\tfrac{1}{2},\tfrac{1}{2}\rangle \\ |\beta\rangle &= |-\tfrac{1}{2}\rangle \equiv |\tfrac{1}{2},-\tfrac{1}{2}\rangle ,\end{aligned} \tag{3.82}$$

we can write this last summation explicitly, namely

$$\hat{Z} = \sum_{cd} \langle c|\hat{z}|d\rangle \left(X_{c\alpha}^\dagger X_{d\alpha} + X_{c\beta}^\dagger X_{d\beta}\right). \tag{3.83}$$

Similarly, for a two-particle spin-independent operator $\hat{V}$, Eq. (70), we can write

$$\begin{aligned}\hat{V} &= \tfrac{1}{2}\sum_{C_1C_2D_1D_2} \langle C_1C_2|\hat{v}(\vec{r}_1,\vec{r}_2)|D_1D_2\rangle\, X_{C_1}^\dagger X_{C_2}^\dagger X_{D_2} X_{D_1} \\ &= \tfrac{1}{2}\sum_{c_1c_2d_1d_2}\sum_{\gamma_1\gamma_2\delta_1\delta_2} \langle\gamma_1|\delta_1\rangle\langle\gamma_2|\delta_2\rangle\langle c_1c_2|\hat{v}|d_1d_2\rangle\, X_{c_1\gamma_1}^\dagger X_{c_2\gamma_2}^\dagger X_{d_2\delta_2} X_{d_1\delta_1} \\ &= \tfrac{1}{2}\sum_{c_1c_2d_1d_2} \langle c_1c_2|\hat{v}|d_1d_2\rangle \sum_{\gamma\delta} X_{c_1\gamma}^\dagger X_{c_2\delta}^\dagger X_{d_2\delta} X_{d_1\gamma} \end{aligned} \tag{3.84}$$

or, explicitly

$$\begin{aligned}\hat{V} = \tfrac{1}{2}\sum_{c_1c_2d_1d_2} \langle c_1c_2|\hat{v}|d_1d_2\rangle \big(& X_{c_1\alpha}^\dagger X_{c_2\alpha}^\dagger X_{d_2\alpha} X_{d_1\alpha} + X_{c_1\alpha}^\dagger X_{c_2\beta}^\dagger X_{d_2\beta} X_{d_1\alpha} \\ & + X_{c_1\beta}^\dagger X_{c_2\alpha}^\dagger X_{d_2\alpha} X_{d_1\beta} + X_{c_1\beta}^\dagger X_{c_2\beta}^\dagger X_{d_2\beta} X_{d_1\beta}\big).\end{aligned} \tag{3.85}$$

*) For this reason we avoid in dealing with general spin-orbitals the first two letters of the alphabet.

3.5. Canonical transformation of operators

A general second quantized form (60) of an arbitrary operator is invariant under canonical transformations (1.126). Indeed, considering a transformation

$$X_A = \sum_B d_{AB}\, \tilde{X}_B , \tag{3.86}$$

where

$$\mathbb{D} = \| d_{AB} \| \tag{3.87}$$

is unitary, we get for an arbitrary one-particle operator, say $\hat{Z}$

$$\begin{aligned}\hat{Z} &= \sum_{AB} \langle A|\hat{z}|B\rangle\, X_A^\dagger X_B \\ &= \sum_{AB} \langle A|\hat{z}|B\rangle \sum_{CD} d_{AC}^*\, d_{BD}\, \tilde{X}_C^\dagger \tilde{X}_D \\ &= \sum_{CD} \Big(\sum_{AB} d_{AC}^* \langle A|\hat{z}|B\rangle\, d_{BD}\Big)\, \tilde{X}_C^\dagger \tilde{X}_D . \end{aligned} \tag{3.88}$$

The matrix elements of $\hat{Z}$, Eq. (7),

$$z_{AB} \equiv \langle A|\hat{z}|B\rangle , \tag{3.89}$$

constitute the "A"-representation of $\hat{Z}$,

$$\mathbb{Z} = \| z_{AB} \| . \tag{3.90}$$

Designating the transformed matrix elements enclosed in the parentheses in (88) as

$$\tilde{z}_{CD} \equiv \sum_{AB} d_{AC}^*\, z_{AB}\, d_{BD} = \left(\mathbb{D}^\dagger \mathbb{Z}\, \mathbb{D}\right)_{CD} , \tag{3.91}$$

we see immediately that they are related to the matrix elements z_{AB} by a unitary transformation $\mathbb{D}$. Since the single particle states are transformed in the same way as the corresponding creation operators, Eqs. (1.127) and (1.133), we can regard the matrix elements (91) as the "$\tilde{A}$"-representation of $\hat{Z}$. In view of (1.134) we can also write

$$\hat{Z} = \sum_{CD} \tilde{z}_{CD}\, \tilde{X}_C^\dagger \tilde{X}_D = \sum_{\tilde{C}\tilde{D}} \langle \tilde{C}|\hat{z}|\tilde{D}\rangle\, X_{\tilde{C}}^\dagger X_{\tilde{D}}, \qquad (3.92)$$

where

$$\langle \tilde{C}|\hat{z}|\tilde{D}\rangle = \tilde{z}_{CD} = \sum_{AB} d_{AC}^* \langle A|\hat{z}|B\rangle d_{BD}, \qquad (3.93)$$

showing explicitly the invariance of the form (37) or (64) for a one-particle operator.

The higher order operators behave in a similar manner. For example, a two-particle operator, say $\hat{V}$, transforms as

$$\begin{aligned}
\hat{V} &= \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle\, X_A^\dagger X_B^\dagger X_D X_C \\
&= \tfrac{1}{2} \sum_{ABCD} v_{AB,CD}\, X_A^\dagger X_B^\dagger X_D X_C \\
&= \tfrac{1}{2} \sum_{ABCD} v_{AB,CD} \sum_{EFGH} d_{AE}^* d_{BF}^* d_{CG} d_{DH}\, \tilde{X}_E^\dagger \tilde{X}_F^\dagger \tilde{X}_H \tilde{X}_G \\
&= \tfrac{1}{2} \sum_{EFHG} \tilde{v}_{EF,HG}\, \tilde{X}_E^\dagger \tilde{X}_F^\dagger \tilde{X}_H \tilde{X}_G \\
&= \tfrac{1}{2} \sum_{\tilde{A}\tilde{B}\tilde{C}\tilde{D}} \langle \tilde{A}\tilde{B}|\hat{v}|\tilde{C}\tilde{D}\rangle\, X_{\tilde{A}}^\dagger X_{\tilde{B}}^\dagger X_{\tilde{D}} X_{\tilde{C}}, \qquad (3.94)
\end{aligned}$$

where

$$\tilde{v}_{EF,GH} = \sum_{ABCD} d_{AE}^* d_{BF}^*\, v_{AB,CD}\, d_{CG} d_{DH} \qquad (3.95)$$

constitute the $\tilde{A}$- representation of $\hat{V}$. A similar behavior is easily derived for an arbitrary operator.

This invariance is clearly in agreement with the fact that we can use an arbitrary complete set of single particle states as the basis for the formulation of the second quantization formalism. This in turn implies that there must exist a complete set of single particle states, which is particularly convenient when we examine a given one-particle operator, namely the set in which the representation of this operator is diagonal. Such a set consists of the eigenvectors of this operator (which we assume to be hermitian so that its eigenvectors form a complete orthonormal set). For example, assume that $|A\rangle$ are eigenstates of $\hat{Z}$ with the eigenvalues ϵ_A,

$$\hat{z}|A\rangle = \epsilon_A |A\rangle, \qquad (3.96)$$

so that

$$\langle A|\hat{z}|B\rangle = \langle A|\epsilon_B|B\rangle = \epsilon_B \langle A|B\rangle = \epsilon_A \langle A|B\rangle. \qquad (3.97)$$

The second quantized form of $\hat{z}$ takes then a very simple form

$$\begin{aligned}
\hat{Z} &= \sum_{AB} \langle A|\hat{z}|B\rangle X_A^\dagger X_B = \sum_A \epsilon_A X_A^\dagger X_A \\
&= \sum_A \epsilon_A \hat{N}_A, \qquad (3.98)
\end{aligned}$$

where $\hat{N}_A$ is the number operator for the spin-orbital $|A\rangle$, Eq. (1.58).

4. Hole-particle formalism

When evaluating matrix elements of various operators as vacuum mean values of appropriate operator products, the number of creators and annihilators in these products is equal to $2(N+O)$, where N is the number of particles present (i.e., the number of single-particle states in each configuration) and O is the order of the operator whose matrix elements we calculate. While O is usually not greater than two, N is very often quite large and is in fact unbounded for extended systems like an electron gas model or a molecular crystal. Thus, a direct exploitation of Wick's theorem can become very cumbersome in many-electron systems or even impossible for extended systems.

Fortunately, in a number of problems of interest, the exact state may be reasonably well approximated by a superposition of basis states or configurations, which differ one from the other in only a few single-particle states, assuming that these states are properly chosen. (We shall indicate later how to make such a choice.) Moreover, for a given state, one configuration in the above mentioned superposition will often have a much larger weight than all the remaining ones, so that it can be chosen as a reference configuration, representing the zeroth approximation to a given state. Rephrased in a more conventional terms this means that an independent particle picture (i.e., a single configuration or a single antisymmetrized product approximation) gives a good description of the state considered and, further, that the configuration interaction or shell model expansion converges reasonably well, so that only low-excited configurations (say, mono-, bi-, tri- and tetra-excited ones*) occur in the expansion, the higher excited ones

*) An n-fold excited configuration is a configuration differing in n single-particle states from the chosen reference configuration.

being negligible.

When these assumptions apply it is clearly advantageous to choose the prominent configuration as a new reference or "vacuum" state. Purely formally, one can of course always choose one of the configurations as the reference one and express all other configurations through this chosen one, similarly as in the example 1.7.7. However, when the relevant N-particle configurations differ in only a few single-particle states, say m, $m \ll N$, from the chosen reference configuration, we can express them efficiently in terms of this new "vacuum" state using $2m$ creation and annihilation operators, rather than N operators which are necessary when we use the "true" vacuum state $|0\rangle$.

However, the relationship (1.47) or (2.35), which makes the concept of the n-product and of the Wick theorem so useful through the rules R1-R6, will not hold for the new "vacuum" state. We thus have to redefine the creation and annihilation operators in such a way that an analogous relationship holds for these redefined quantities. This in turn requires that we also modify the definitions of the normal product and the contraction of two operators, so that the Wick theorem holds.

This can be indeed easily achieved by defining the new creation and annihilation operators with respect to the new "vacuum", represented by the chosen new configuration. Formally, this procedure is not unlike the early electron-positron theory of Dirac: our refelence configuration or the new "vacuum" corresponds to the Dirac sea of filled negative energy levels. Annihilating a particle in this sea (i.e., in our new "vacuum") creates an unoccupied state, a so called "hole" (in Dirac's theory a positron), whose many properties are similar to those of "particles", which are present in addition to those

forming the "vacuum". The choice of the reference configuration forming the new "vacuum" is equivalent to a division of one-particle states into the two mutually disjoint subsets, namely the subsets of spin-orbitals which are and are not occupied in the new "vacuum" state. An arbitrary configuration may then be given relative to the new "vacuum" state by specifying which particles are missing in the "vacuum" or, equivalently, which "holes" are present in the vacuum and, on the other hand, which particle state, in addition to those of the "vacuum", are occupied. The occupied states are thus often referred to as the "hole" states, while the unoccupied ones are called "particle" states.

In the case of the ground state of an N-electron system, the occupied states are usually the energetically lowest single-particle states of a chosen independent particle model, usually the ground state Hartree-Fock orbitals. In such case an occupied spinorbital with the highest orbital energy divides the set of one-particle states into the two sets mentioned above, and its energy is referred to as the <u>Fermi level</u>; the states with higher energy than the Fermi level energy are <u>unoccupied</u> or <u>particle</u> states, and may only be occupied by "true" particles (i.e. particles of a hole-particle formalism), while the remaining spin-orbitals form a so-called <u>Fermi sea</u> of <u>occupied</u> or <u>hole</u> states, which is analogous to the Dirac electron-positron sea mentioned above. Physically, particles may be excited from the occupied states into the unoccupied ones, forming a "hole" in the Fermi sea and a "particle" above the Fermi level, in analogy to the positron-electron pair creation. We thus call the new "vacuum" state $|\Phi_0\rangle$ the <u>Fermi vacuum</u> in order to distinguish it from the <u>true</u> or <u>physical vacuum</u> $|0\rangle$.

4.1. Hole-particle operators

The above qualitative introductory remarks will now be made more precise. The N-particle reference configuration, chosen as a new vacuum (Fermi vacuum) will be designated $|\Phi_0\rangle$. In order to distinguish readily the labels associated with hole (occupied in $|\Phi_0\rangle$) and particle (unoccupied in $|\Phi_0\rangle$) states, we employ the sub- and super-scripted letters, respectively, unless specified to the contrary. In case of states which may be either occupied or unoccupied we shall use letters without any sub- or super-scripts or, in case the indices are unavoidable, we shall use subscripts enclosed in parentheses. For example,

$$|0_i\rangle = |d_i\rangle |\delta_i\rangle \tag{4.1}$$

will now designate an occupied spin-orbital, and

$$|0^i\rangle = |d^i\rangle |\delta^i\rangle \tag{4.2}$$

an unoccupied spin-orbital, while a spin-orbital which may be either occupied or unoccupied will be designated in one of the following ways

$$|0\rangle = |d\rangle |\delta\rangle \tag{4.3}$$

$$|0_{(i)}\rangle = |d_{(i)}\rangle |\delta_{(i)}\rangle . \tag{4.4}$$

Thus, the Fermi vacuum state $|\Phi_0\rangle$, may be written as follows

$$|\Phi_0\rangle = |\{A_1 A_2 \cdots A_N\}\rangle = X^+_{A_1} \cdots X^+_{A_N} |0\rangle . \tag{4.5}$$

The creation and annihilation operators defined with respect to the Fermi vacuum $|\Phi_0\rangle$ will be designated with a letter Y and an appropriate spin-orbital label. We shall define them as follows

$$Y_{A_1} = X_{A_1}^\dagger \quad , \quad Y_{A_1}^\dagger = X_{A_1}$$

and

$$Y_{A'} = X_{A'} \quad , \quad Y_{A'}^\dagger = X_{A'}^\dagger \; . \tag{4.6}$$

Thus, for the unoccupied (particle) states the new operators are identical with the operators defined earlier with respect to the true vacuum, while for the occupied (hole) states the roles of creation and annihilation operators are reversed. This is easily understood on the basis of a foregoing discussion: above the Fermi level we can create and annihilate particles as before while for the hole states, which are occupied in $|\Phi_0\rangle$, the situation changes. Clearly, an annihilation of a particle in the Fermi sea (i.e., in an occupied state) is equivalent to the creation of a hole in the Fermi sea and, vice versa, creation of a particle in an occupied state is equivalent to an annihilation of a hole, assuming the hole is present in the first place. Similarly as we distinguish hole and particle states we shall refer to the corresponding operators as <u>hole</u> and <u>particle operators</u>, respectively.

Let us now convince ourselves that the annihilator Y_A, associated with an arbitrary spin-orbital $|A\rangle$, gives a vanishing result when acting on the Fermi vacuum $|\Phi_0\rangle$,

$$Y_A |\Phi_0\rangle = 0 \; , \tag{4.7}$$

in complete analogy to (1.47) or (2.35). Indeed, if A is a hole state, say $A \equiv A_i$, we have that

$$Y_{A_i} |\Phi_0\rangle = X_{A_i}^\dagger |\Phi_0\rangle = X_{A_i}^\dagger |\{A_1 A_2 \dots A_N\}\rangle = 0 \; , \tag{4.8}$$

and when A is a particle state, say $A \equiv A^i$, then again

$$Y_{A^i} |\Phi_0\rangle = X_{A^i} |\{A_1 A_2 \dots A_N\}\rangle = 0 \; . \tag{4.9}$$

We shall see in the following that indeed the properties of the new operators Y_A and $Y_A^\dagger$ are closely analogous to those of the operators X_A and $X_A^\dagger$ when the true vacuum state $|0\rangle$ is replaced by the Fermi vacuum $|\Phi_0\rangle$. However, we must always remember that while the operators X_A and $X_A^\dagger$ are completely determined by the chosen complete orthonormal set of spinorbitals $|A\rangle$, the definition of the hole-particle operators Y_A and $Y_A^\dagger$ also depends on the choice of the vacuum state $|\Phi_0\rangle$.

4.2. Anti-commutation relations

We now demonstrate that the new operators (6) satisfy the same anti-commutation relations as the old operators X_A. This is immediately obvious when both operators are particle operators, since then

$$\begin{aligned} [Y_{A'}, Y_{A^2}]_+ &= [X_{A'}, X_{A^2}]_+ = 0 \\ [Y_{A'}^\dagger, Y_{A^2}^\dagger]_+ &= [X_{A'}^\dagger, X_{A^2}^\dagger]_+ = 0 \\ [Y_{A'}, Y_{A_2}^\dagger]_+ &= [X_{A'}, X_{A^2}^\dagger]_+ = \langle A' | A^2 \rangle . \end{aligned} \tag{4.10}$$

Also, when both operators are hole operators, we have

$$\begin{aligned} [Y_{A_1}, Y_{A_2}]_+ &= [X_{A_1}^\dagger, X_{A_2}^\dagger]_+ = 0 \\ [Y_{A_1}^\dagger, Y_{A_2}^\dagger]_+ &= [X_{A_1}, X_{A_2}]_+ = 0 \\ [Y_{A_1}, Y_{A_2}^\dagger]_+ &= [X_{A_1}^\dagger, X_{A_2}]_+ = \langle A_1 | A_2 \rangle . \end{aligned} \tag{4.11}$$

Thus, we only have to examine the cases with one particle and one hole operator, when the number of creation and annihilation operators will change by one, so that for example an anti-commutator of two annihilators,

$$[Y_{A_1}, Y_{A'}]_+ = [X_{A_1}^\dagger, X_{A'}]_+ \tag{4.12}$$

corresponds to an anti-commutator of an annihilator and a creator, which in principle can give a non-vanishing result. However, we realize immediately that the anti-commutator on the right hand side of (12) must also vanish, since occupied and unoccupied states belong to mutually disjoint sets of orthonormal spinorbitals so that

$$\langle A_1 | A' \rangle = 0. \tag{4.13}$$

Thus,

$$[Y_{A_1}, Y_{A'}]_+ = 0 \tag{4.14}$$

for arbitrary A_1 and A'. This argument applies clearly to any anti-commutator of one hole and one particle operator, irrespective of their character, so that such an anti-commutator must alway vanish.

We can thus summarize all the special cases considered and write again

$$[Y_A, Y_B]_+ = [Y_A^\dagger, Y_B^\dagger]_+ = 0$$
$$[Y_A, Y_B^\dagger]_+ = \langle A | B \rangle, \tag{4.15}$$

where A and B may be any two (i.e., particle, hole or particle and hole) spinorbital states.

4.3. Normal products and contractions

We define the normal product of operators with respect to the Fermi vacuum $|\Phi_0\rangle$ in an analogous way as for the true vacuum $|0\rangle$. Thus

$$N[\hat{M}_{B(1)} \hat{M}_{B(2)} \cdots \hat{M}_{B(n)}] = (-1)^p Y^\dagger_{C(1)} \cdots Y^\dagger_{C(k)} Y_{C(k+1)} \cdots Y_{C(n)} \tag{4.16}$$

where $(-1)^p$ is the parity of the permutation $P: B_{(i)} \to C_{(i)}$ and $M_{B(i)}$ designates either a creator $Y^\dagger_{B(i)}$ an annihilator $Y_{B(i)}$. Simple properties for the n-products, Eqs. (2.7) – (2.13), are also valid for the N-products if we replace the operators X with hole-particle formalism operators Y. Also, analogously to (2.14) we define

$$N[\mathcal{O}] = 1. \tag{4.17}$$

We have introduced a different notation for the two types of normal products since we shall often need to apply the normal product defined with respect to the Fermi vacuum to the operators X defined with respect to the true vacuum. Clearly, such a normal product can only be evaluated when we know the Fermi vacuum state $|\Phi_0\rangle$, i.e. when we are able to distinguish the hole and particle states. We shall nevertheless find it convenient to use this notation without the prior knowledge of the hole or particle character of the states involved.

A difference between both types of normal products is apparent from simple examples:

1. Analogously to an example (2.5), (2.6) we can write

$$N[Y_A^\dagger Y_B Y_C^\dagger] = -Y_A^\dagger Y_C^\dagger Y_B = Y_C^\dagger Y_A^\dagger Y_B \tag{4.18}$$

where A, B and C may label either particle or hole states.

2. Calculate now

$$N[X_A^\dagger X_B X_C^\dagger], \tag{4.19}$$

i.e., the N-product of operators defined with respect to a true vacuum. Clearly, to get an explicit result, we have to know which states are occupied and which not. Assume, thus, that for example A and B are hole states and C is a particle state:

$$A = A_1 \qquad B = B_1 \qquad C = C'. \tag{4.20}$$

Then

$$N[X^\dagger_{A_1} X_{B_1} X^\dagger_{C'}] = N[Y_{A_1} Y^\dagger_{B_1} Y^\dagger_{C'}] = Y^\dagger_{B_1} Y^\dagger_{C'} Y_{A_1} = -Y^\dagger_{C'} Y^\dagger_{B_1} Y_{A_1} \quad (4.21)$$

or

$$N[X^\dagger_{A_1} X_{B_1} X^\dagger_{C'}] = X_{B_1} X^\dagger_{C'} X^\dagger_{A_1} = -X^\dagger_{C'} X_{B_1} X^\dagger_{A_1} . \quad (4.22)$$

Comparison of this result with Eqs. (2.5), (2.6) clearly demonstrates the difference between the n- and N-products and the dependence of the latter on the choice of $|\Phi_0\rangle$.

A contraction (pairing) $\overbrace{\hat{M}_{(1)} \hat{M}_{(2)}}$ of the two operators $\hat{M}_{(1)}$ and $\hat{M}_{(2)}$ with respect to the Fermi vacuum is defined

$$\overbrace{\hat{M}_{(1)} \hat{M}_{(2)}} = \hat{M}_{(1)} \hat{M}_{(2)} - N[\hat{M}_{(1)} \hat{M}_{(2)}] . \quad (4.23)$$

This is again completely analogous to the definition (2.15). In order to distinguish both kinds of contractions we now write it above the operators concerned.

For the Y operators we clearly get the same result as in Eqs. (2.16) through (2.20), namely

$$\overbrace{Y_A Y_B} = \overbrace{Y^\dagger_A Y^\dagger_B} = \overbrace{Y^\dagger_A Y_B} = 0$$
$$\overbrace{Y_A Y^\dagger_B} = \langle A|B\rangle . \quad (4.24)$$

In principle, we can always transform the operators X into the hole-particle formalism operators Y using definition (6). However, this may not always be convenient, particularly when the operator labels are summed over all possible states. For example, when we transform a one-particle operator $\hat{Z}$, Eq. (3.64), to the hole-particle form, we have to split each summation into the two separate summations which results in $2^2 = 4$ terms:

$$\hat{Z} = \Big(\sum_{A_1} + \sum_{A'}\Big)\Big(\sum_{B_1} + \sum_{B'}\Big) \langle A|\hat{z}|B\rangle X^\dagger_A X_B$$
$$= \sum_{A_1}\sum_{B_1} \langle A_1|\hat{z}|B_1\rangle Y_{A_1} Y^\dagger_{B_1}$$

$$+ \sum_{A_1}\sum_{B'} \langle A_1|\hat{z}|B'\rangle Y_{A_1} Y_{B'}$$
$$+ \sum_{A'}\sum_{B_1} \langle A'|\hat{z}|B_1\rangle Y^\dagger_{A'} Y^\dagger_{B_1}$$
$$+ \sum_{A'}\sum_{B'} \langle A'|\hat{z}|B'\rangle Y^\dagger_{A'} Y_{B'} . \quad (4.25)$$

A similar transformation for a two-particle operator $\hat{V}$ would result in $2^4 = 16$ different terms. This would be very cumbersome in practical calculations since one would have to consider all sixteen terms separately. We shall see that it is much more convenient not to transform operators like $\hat{Z}$ or $\hat{V}$ to the hole-particle form but rather to work out the expressions for contractions of mixed X and Y operators.

In order to write effectively these mixed contractions we define the following "step functions" which will enable us to distinguish hole and particle states:

$$h(A_1) = 1 , \quad h(A') = 0$$
and
$$p(A_1) = 0 , \quad p(A') = 1 . \quad (4.26)$$

Clearly, for an arbitrary state A

$$h(A) + p(A) = 1 . \quad (4.27)$$

We also note that

$$\sum_A f(A)\, p(A) = \sum_{A'} f(A')$$
and
$$\sum_A f(A)\, h(A) = \sum_{A_1} f(A_1) , \quad (4.28)$$

where $f(A)$ is an arbitrary function of A.

We can now easily derive expressions for mixed operator contractions, since clearly

$$\overbrace{X^\dagger_A Y_B} = \overbrace{X_A Y_B} = \overbrace{Y^\dagger X_B} = \overbrace{Y^\dagger X^\dagger_B} = 0 , \quad (4.29)$$

while, for example, for $\overbrace{X_A Y_B^\dagger}$ we find

$$\begin{aligned}\overbrace{X_A Y_B^\dagger} &= [p(A) + h(A)]\,\overbrace{X_A Y_B^\dagger} \\ &= p(A)\,\overbrace{X_A Y_B^\dagger} + h(A)\,\overbrace{X_A Y_B^\dagger} \\ &= p(A)\,\overbrace{Y_A Y_B^\dagger} + h(A)\,\overbrace{Y_A^\dagger Y_B^\dagger} \\ &= p(A)\langle A|B\rangle . \end{aligned} \tag{4.30}$$

Similarly, for the Fermi vacuum contractions of X operators we find

$$\begin{aligned}\overbrace{X_A^\dagger X_B^\dagger} &= [p(A) + h(A)][p(B) + h(B)]\,\overbrace{X_A^\dagger X_B^\dagger} \\ &= p(A)p(B)\,\overbrace{Y_A^\dagger Y_B^\dagger} + h(A)p(B)\,\overbrace{Y_A Y_B^\dagger} \\ &\quad + p(A)h(B)\,\overbrace{Y_A^\dagger Y_B} + h(A)h(B)\,\overbrace{Y_A Y_B} \\ &= h(A)p(B)\,\overbrace{Y_A Y_B^\dagger} = h(A)p(B)\langle A|B\rangle = 0, \end{aligned} \tag{4.31}$$

since hole particle states belong to disjoint sets. We get similarly that also $\overbrace{X_A X_B} = 0$, while

$$\begin{aligned}\overbrace{X_A^\dagger X_B} &= h(A)h(B)\,Y_A Y_B^\dagger = h(A)h(B)\langle A|B\rangle \\ &= h(A)\langle A|B\rangle = h(B)\langle A|B\rangle \end{aligned} \tag{4.32}$$

and, of course, $\overbrace{X_A X_B^\dagger} = p(B)\langle A|B\rangle = p(A)\langle A|B\rangle$. (4.32')

Summarizing all the necessary cases we get

$$\begin{array}{lll} \overbrace{X_A Y_B} = 0 & \overbrace{Y_A X_B} = h(B)\langle A|B\rangle & \overbrace{X_A X_B} = 0 \\ \overbrace{X_A^\dagger Y_B} = 0 & \overbrace{Y_A^\dagger X_B} = 0 & \overbrace{X_A^\dagger X_B} = h(A)\langle A|B\rangle \\ \overbrace{X_A Y_B^\dagger} = p(A)\langle A|B\rangle & \overbrace{Y_A X_B^\dagger} = p(B)\langle A|B\rangle & \overbrace{X_A X_B^\dagger} = p(A)\langle A|B\rangle \\ \overbrace{X_A^\dagger Y_B^\dagger} = h(A)\langle A|B\rangle & \overbrace{Y_A^\dagger X_B^\dagger} = 0 & \overbrace{X_A^\dagger X_B^\dagger} = 0 \end{array} \tag{4.33}$$

It remains, finally, to define a normal product with contractions relative to the Fermi vacuum

or, simply, N-product with contractions. The definition is again completely analogous as in the case of the n-product with contractions. To conclude this section we compare corresponding concepts in true vacuum and hole-particle formalisms and the introduced notation:

quantity	true vacuum formalism	hole-particle (Fermi vacuum) formalism
vacuum state	$\|0\rangle$	$\|\Phi_0\rangle$
creation and annihilation operators ($\hat{M}_A$)	$X_A^\dagger$, X_A	$Y_A^\dagger$, Y_A
normal product of operators	$n[\cdots \hat{M}_A \cdots]$	$N[\cdots \hat{M}_A \cdots]$
contraction of two operators	$\hat{M}_A \hat{M}_B$	$\overbrace{M_A M_B}$
normal product with contractions	$n[\cdots\cdots]$	$N[\cdots\cdots]$

4.4. Wick's theorem and examples

If we recall the proof of the Wick theorem and of its generalized version given in Secs. 2.4 and 2.5, respectively, we see immediately that the same proof will apply to the hole-particle formalism versions of these theorems. We only have to replace everywhere the true vacuum quantities with the corresponding Fermi vacuum quantities, as indicated in the table given above. We can thus write immediately the hole-particle form of Wick's theorem as follows

$$\hat{M}_{(1)}\hat{M}_{(2)}\cdots\hat{M}_{(m)} = N[\hat{M}_{(1)}\hat{M}_{(2)}\cdots\hat{M}_{(m)}]$$

$$+\sum_{i<j} N[\hat{M}_{(1)} \cdots \overbrace{\hat{M}_{(i)} \cdots \hat{M}_{(j)}} \cdots \hat{M}_{(m)}]$$

$$+ \cdots\cdots + \sum N[\overbrace{\hat{M}_{(1)} \cdots \hat{M}_{(c)} \cdots}\ \hat{M}_{(m)}], \qquad (4.34)$$

where the operators $\hat{M}_{(i)}$ are hole-particle operators defined with respect to $|\Phi_0\rangle$. These operators may also be written through the true vacuum operators X, so that the product $\hat{M}_{(1)} \cdots \hat{M}_{(m)}$ may consist of mixed operators X and Y. We must keep in mind that the X operator form is only used for convenience and that we have to use expressions (33) when we evaluate the contractions.

In the same way we can express the generalized Wick theorem in the hole-particle formalism and re-formulate the rules R5 and R6 for the case of N-products. Thus, for example, we now have that

$$\langle\Phi_0|\hat{M}_{(1)}\hat{M}_{(2)}\cdots\hat{M}_{(2m)}|\Phi_0\rangle$$

$$= \sum_{f.c.} \langle\Phi_0|N[\overbrace{\hat{M}_{(1)}\hat{M}_{(2)}\cdots}\ \hat{M}_{(2m)}]|\Phi_0\rangle$$

$$= \sum_{f.c.} (-1)^p \prod_{k=1}^{n} \overbrace{\hat{M}_{(i_k)}\hat{M}_{(j_k)}} \qquad (4.35)$$

in a complete analogy to (2.70).

We now illustrate the hole-particle formalism introduced so far on a few examples.

1. Calculate the vacuum mean value of a single particle operator $\hat{Z}$, Eq. (3.64),

$$\langle\Phi_0|\hat{Z}|\Phi_0\rangle, \qquad (4.36)$$

where $|\Phi_0\rangle \equiv |\{A_1 A_2 \cdots A_N\}\rangle$ is the Fermi vacuum reference state.

Substituting from (3.64) and using Wick's theorem or, directly, the relationship (35) we find

$$\langle\Phi_0|\hat{Z}|\Phi_0\rangle = \sum_{AB} \langle A|\hat{z}|B\rangle \langle\Phi_0|X_A^\dagger X_B|\Phi_0\rangle$$

$$= \sum_{AB} \langle A|\hat{z}|B\rangle \langle\Phi_0|N[\overbrace{X_A^\dagger X_B}]|\Phi_0\rangle$$

$$= \sum_{AB} \langle A|\hat{z}|B\rangle\ \overbrace{X_A^\dagger X_B}$$

$$= \sum_{AB} \langle A|\hat{z}|B\rangle\, \delta(A)\langle A|B\rangle$$

$$= \sum_{A_i} \langle A_i|\hat{z}|A_i\rangle, \qquad (4.37)$$

where we used (33) and (28) in the last two steps. The last summation thus extends over occupied or hole states and can be written in either of the following equivalent forms

$$\langle\Phi_0|\hat{Z}|\Phi_0\rangle = \sum_{A_i} \langle A_i|\hat{z}|A_i\rangle = \sum_{i=1}^{N} \langle A_i|\hat{z}|A_i\rangle. \qquad (4.38)$$

Thus, A_i can either be used as a label of the first occupied spinorbital, as in $|\Phi_0\rangle$, or as a generic summation label for the sum over the occupied (hole) spinorbitals, as in (38).

2. Evaluate the Fermi vacuum mean value of a two particle operator $\hat{V}$, Eq. (3.94),

$$\langle\Phi_0|\hat{V}|\Phi_0\rangle. \qquad (4.39)$$

Proceeding as in Ex. 1 we write

$$\langle\Phi_0|\hat{V}|\Phi_0\rangle = \tfrac{1}{2}\sum_{ABCD} \langle AB|\hat{v}|CD\rangle\langle\Phi_0|X_A^\dagger X_B^\dagger X_D X_C|\Phi_0\rangle, \qquad (4.40)$$

and evaluating the vacuum mean value of the operator product in (40) using (35) and (33) we get

$$\langle\Phi_0|X_A^\dagger X_B^\dagger X_D X_C|\Phi_0\rangle =$$

$$= \langle\Phi_0|\overbrace{X_A^\dagger X_B^\dagger X_D} X_C|\Phi_0\rangle + \langle\Phi_0|\overbrace{X_A^\dagger X_B^\dagger}\ \overbrace{X_D X_C}|\Phi_0\rangle$$

$$= (X_A^\dagger X_C\, X^\dagger X_D - X_A^\dagger X_D\, X_B^\dagger X_C)\langle\Phi_0|\Phi_0\rangle$$

$$= \delta(A)\delta(B)(\langle A|C\rangle\langle B|D\rangle - \langle A|D\rangle\langle B|C\rangle) \qquad (4.41)$$

so that

$$\langle \Phi_0 | \hat{V} | \Phi_0 \rangle = \frac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \left(\langle A|C\rangle\langle B|D\rangle - \langle A|D\rangle\langle B|C\rangle \right) n(A) n(B)$$

$$= \frac{1}{2} \sum_{A_1 A_2} \left(\langle A_1 A_2|\hat{v}|A_1 A_2\rangle - \langle A_1 A_2|\hat{v}|A_2 A_1\rangle \right)$$

$$= \frac{1}{2} \sum_{A_1 A_2} \langle A_1 A_2|\hat{v}|A_1 A_2\rangle_{\mathcal{A}} \,. \qquad (4.42)$$

3. Calculate the matrix element of $\hat{V}$, Eq. (3.94) between the Fermi vacuum and a singly-excited configuration

$$\left\langle \Phi_0 \middle| \hat{V} \middle| \begin{matrix} A^k \\ A_i \end{matrix} \right\rangle \equiv \langle \Phi_0 | \hat{V} | \Phi_i^k \rangle \qquad (4.43)$$

where

$$|\Phi_0\rangle = |\{A_1 A_2 \cdots A_N\}\rangle \qquad (4.44a)$$

$$\left| \begin{matrix} A^k \\ A_i \end{matrix} \right\rangle \equiv |\Phi_i^k\rangle = |\{A_1 \cdots A_{i-1} A^k A_{i+1} \cdots A_N\}\rangle$$

$$= X_{A^k}^\dagger X_{A_i} |\Phi_0\rangle$$

$$= Y_{A^k}^\dagger Y_{A_i}^\dagger |\Phi_0\rangle \,. \qquad (4.44b)$$

We first evaluate the necessary mean value

$\langle \Phi_0 | X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle$

$$= \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger} + \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger}$$

$$+ \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger} + \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger}$$

$$+ \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger} + \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger}$$

$$+ \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger} + \overbrace{X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger}$$

$$= n(A)\langle A|D\rangle\, n(B) p(C) \left[\langle B|A^k\rangle\langle C|A_i\rangle - \langle B|A_i\rangle\langle C|A^k\rangle \right]$$

$$- n(A)\langle A|C\rangle\, n(B) p(D) \left[\langle B|A^k\rangle\langle D|A_i\rangle - \langle B|A_i\rangle\langle D|A^k\rangle \right]$$

$$- n(A)\langle A|A^k\rangle\, n(B) \left[\langle B|D\rangle p(C)\langle C|A_i\rangle - \langle B|C\rangle p(D)\langle D|A_i\rangle \right]$$

$$+ n(A)\langle A|A_i\rangle\, n(B) \left[\langle B|D\rangle p(C)\langle C|A^k\rangle - \langle B|C\rangle p(D)\langle D|A^k\rangle \right]$$

$$= -n(A)\langle A|D\rangle\langle B|A_i\rangle\langle C|A^k\rangle$$
$$+ n(A)\langle A|C\rangle\langle B|A_i\rangle\langle D|A^k\rangle$$
$$+ \langle A|A_i\rangle n(B) \left[\langle B|D\rangle\langle C|A^k\rangle - \langle B|C\rangle\langle D|A^k\rangle \right], \qquad (4.45)$$

since the terms like $n(B)\langle B|A^k\rangle$ or $p(C)\langle C|A_i\rangle$ must always vanish. Substituting this result into the expansion for the matrix element (43) we get

$$\langle \Phi_0 | \hat{V} | \Phi^k \rangle = \frac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \langle \Phi_0 | X_A^\dagger X_B^\dagger X_D X_C Y_{A^k}^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle$$

$$= \frac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle \left[n(A)\langle B|A_i\rangle \left(\langle A|C\rangle\langle D|A^k\rangle - \langle A|D\rangle\langle C|A^k\rangle \right) \right.$$
$$\left. + n(B)\langle A|A_i\rangle \left(\langle B|D\rangle\langle C|A^k\rangle - \langle B|C\rangle\langle D|A^k\rangle \right) \right]$$

$$= \frac{1}{2} \sum_{A_1} \left(\langle A_1 A_i|\hat{v}|A_1 A^k\rangle - \langle A_1 A_i|\hat{v}|A^k A_i\rangle \right)$$
$$+ \frac{1}{2} \sum_{B_1} \left(\langle A_i B_1|\hat{v}|A^k B_1\rangle - \langle A_i B_1|\hat{v}|B_1 A^k\rangle \right)$$
$$= \sum_{C_1} \langle A_i C_1|\hat{v}|A^k C_1\rangle_{\mathcal{A}} = \sum_{j=1}^{N} \langle A_i A_j|\hat{v}|A^k A_j\rangle_{\mathcal{A}} \qquad (4.46)$$

4. It is instructive to consider the above examples 2 and 3 simultaneously. Thus, consider the matrix element

$$\langle \Phi_0 | \hat{V} | \Phi' \rangle \qquad (4.47)$$

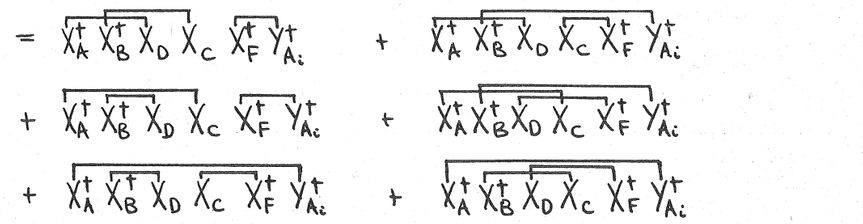
where

$$|\Phi'\rangle = X_F^\dagger Y_{A_i}^\dagger |\Phi_0\rangle = |\{A_1 \cdots A_{i-1} F A_{i+1} \cdots A_N\}\rangle \qquad (4.48)$$

Thus, when $F = A_i$, we recover example 2 while the example 3 results when $F = A^k$.

We evaluate the pertinent vacuum mean value, omitting vanishing contraction schemes and using subsequently relations (33) for of mixed type operators. This gives

$\langle \Phi_0 | X_A^\dagger X_B^\dagger X_D X_C X_F^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle =$



$$= -h(A)\langle A|D\rangle h(B)\left[\langle B|C\rangle h(F)\langle F|A_i\rangle + \langle B|A_i\rangle p(C)\langle C|F\rangle\right]$$

$$+h(A)\langle A|C\rangle h(B)\left[\langle B|D\rangle h(F)\langle F|A_i\rangle + \langle B|A_i\rangle p(D)\langle D|F\rangle\right]$$

$$+h(A)\langle A|A_i\rangle h(B)\left[\langle B|D\rangle p(C)\langle C|F\rangle - \langle B|C\rangle p(D)\langle D|F\rangle\right], \quad (4.49)$$

so that

$$\langle\Phi_0|\hat{V}|\Phi'\rangle = \tfrac{1}{2}\sum_{ABCD}\langle AB|\hat{v}|CD\rangle\langle\Phi_0|X_A^\dagger X_B^\dagger X_D X_C X_F^\dagger Y_{A_i}^\dagger|\Phi_0\rangle$$

$$= -\tfrac{1}{2}\langle F|A_i\rangle\sum_{C_1C_2}\langle C_1\,C_2|\hat{v}|C_2\,C_1\rangle$$

$$-\tfrac{1}{2}p(F)\sum_{C_1}\langle C_1\,A_i|\hat{v}|F\,C_1\rangle$$

$$+\tfrac{1}{2}\langle F|A_i\rangle\sum_{C_1C_2}\langle C_1\,C_2|\hat{v}|C_1\,C_2\rangle$$

$$+\tfrac{1}{2}p(F)\sum_{C_1}\langle C_1\,A_i|\hat{v}|C_1\,F\rangle$$

$$+\tfrac{1}{2}p(F)\sum_{D_1}\langle A_i\,D_1|\hat{v}|F\,D_1\rangle$$

$$-\tfrac{1}{2}p(F)\sum_{C_1}\langle A_i\,C_1|\hat{v}|C_1\,F\rangle \quad (4.50)$$

$$= \langle F|A_i\rangle\tfrac{1}{2}\sum_{C_1C_2}\langle C_1C_2|\hat{v}|C_1C_2\rangle_{AS} + p(F)\sum_{C_1}\langle A_iC_1|\hat{v}|FC_1\rangle_{AS},$$

which summarizes the results of previous two examples.

5. Calculate a matrix element of a one-particle operator $\hat{Z}_N$ between two mono-excited states $|\Phi'\rangle$ and $|\Phi''\rangle$, where

$$|\Phi'\rangle \equiv \left|\begin{matrix}A^j\\A_i\end{matrix}\right\rangle \equiv |\Phi_i^j\rangle = Y_{A^j}^\dagger Y_{A_i}^\dagger|\Phi_0\rangle, \quad (4.51a)$$

$$|\Phi''\rangle \equiv \left|\begin{matrix}A^l\\A_k\end{matrix}\right\rangle \equiv |\Phi_k^l\rangle = Y_{A^l}^\dagger Y_{A_k}^\dagger|\Phi_0\rangle, \quad (4.51b)$$

$$|\Phi_0\rangle = |\{A_1 \cdots A_N\}\rangle \quad (4.52)$$

and

$$\hat{Z}_N = \sum_{AB}\langle A|\hat{z}|B\rangle\, N[X_A^\dagger X_B]. \quad (4.53)$$

We shall call $\hat{Z}_N$ the <u>N-product form of Z.</u>

The desired matrix element has the form

$$\langle\Phi'|\hat{Z}_N|\Phi''\rangle \equiv \langle\Phi_i^j|\hat{Z}_N|\Phi_k^l\rangle \equiv \left\langle\begin{matrix}A^j\\A_i\end{matrix}\right|\hat{Z}_N\left|\begin{matrix}A^l\\A_k\end{matrix}\right\rangle$$

$$= \sum_{AB}\langle A|\hat{z}|B\rangle\langle\Phi_0|Y_{A_i}Y_{A^j}\,N[X_A^\dagger X_B]\,Y_{A^l}^\dagger Y_{A_k}^\dagger|\Phi_0\rangle \quad (4.54)$$

$$= \sum_{AB}\langle A|\hat{z}|B\rangle\langle\Phi_0|N[Y_{A_i}Y_{A^j}]\,N[X_A^\dagger X_B]\,N[Y_{A^l}^\dagger Y_{A_k}^\dagger]|\Phi_0\rangle,$$

so that we can exploit the generalized version of Wick's theorem in evaluating the necessary Fermi vacuum mean value. Since no two operators within the same N-product can be contracted, the two operators in each N-product in (54) must be paired with operators originating from distinct N products. Further, any contraction between hole and particle operators must vanish. Thus

$$\langle\Phi_0|N[Y_{A_i}Y_{A^j}]\,N[X_A^\dagger X_B]\,N[Y_{A^l}^\dagger Y_{A_k}^\dagger]|\Phi_0\rangle$$

$$= Y_{A_i}Y_{A^j}X_A^\dagger X_B Y_{A^l}^\dagger Y_{A_k}^\dagger + Y_{A_i}Y_{A^j}X_A^\dagger X_B Y_{A^l}^\dagger Y_{A_k}^\dagger$$

$$+ Y_{A_i}Y_{A^j}X_A^\dagger X_B Y_{A^l}^\dagger Y_{A_k}^\dagger + Y_{A_i}Y_{A^j}X_A^\dagger X_B Y_{A^l}^\dagger Y_{A_k}^\dagger$$

$$= \langle A_i|A_k\rangle\left(p(A^j)\langle A^j|A\rangle p(B)\langle B|A^l\rangle - h(A^j)\langle A^j|B\rangle h(A)\langle A|A^l\rangle\right)$$

$$+\langle A^j|A^l\rangle\left(p(A_i)\langle A_i|A\rangle p(B)\langle B|A_k\rangle - h(A_i)\langle A_i|B\rangle h(A)\langle A|A_k\rangle\right)$$

$$= \langle A_i|A_k\rangle\langle A^j|A\rangle\langle B|A^l\rangle - \langle A^j|A^l\rangle\langle A_i|B\rangle\langle A|A_k\rangle, \quad (4.55)$$

so that we find, substituting into (54), that

$$\left\langle\begin{matrix}A^j\\A_i\end{matrix}\right|\hat{Z}_N\left|\begin{matrix}A^l\\A_k\end{matrix}\right\rangle = \langle A_i|A_k\rangle\langle A^j|\hat{z}|A^l\rangle - \langle A^j|A^l\rangle\langle A_k|\hat{z}|A_i\rangle \quad (4.56)$$

The reader will notice that the fact that some groups

of operators appear in the normal product form can significantly simplify our calculations, since the number of possible contraction schemes is significantly reduced. This reduction is particularly significant for higher than one-particle operators, as will be seen in the next example.

6. Calculate a matrix element of V_N,

$$V_N = \tfrac{1}{2}\sum_{ABCD} \langle AB|\hat{v}|CD\rangle \, N\left[X_A^\dagger X_B^\dagger X_D X_C\right] \qquad (4.57)$$

between the mono-excited states (51).

We first evaluate the required Fermi vacuum mean value using generalized Wick's theorem,

$$\langle \Phi_0 | N[Y_{A_i} Y_{A^j}] \, N\left[X_A^\dagger X_B^\dagger X_D X_C\right] N\left[Y_{A^\ell}^\dagger Y_{A_k}^\dagger\right] | \Phi_0\rangle \equiv \langle \Phi_0|\hat{P}|\Phi_0\rangle$$

$$= Y_{A_i} Y_{A^j} X_A^\dagger X_B^\dagger X_D X_C Y_{A^\ell}^\dagger Y_{A_k}^\dagger + Y_{A_i} Y_{A^j} X_A^\dagger X_B^\dagger X_D X_C Y_{A^\ell}^\dagger Y_{A_k}^\dagger$$

$$+ Y_{A_i} Y_{A^j} X_A^\dagger X_B^\dagger X_D X_C Y_{A^\ell}^\dagger Y_{A_k}^\dagger + Y_{A_i} Y_{A^j} X_A^\dagger X_B^\dagger X_D X_C Y_{A^\ell}^\dagger Y_{A_k}^\dagger \qquad (4.58)$$

where we have used the fact (see Eqs. (33)) that the contractions of the type $\overbrace{YX^\dagger}$ and $\overbrace{XY^\dagger}$ must involve particles and $\overbrace{YX}$ and $\overbrace{X^\dagger Y^\dagger}$ holes, so that

$$\langle \Phi_0|\hat{P}|\Phi_0\rangle = \langle A_k|A\rangle\langle A^j|B\rangle\left(\langle C|A^\ell\rangle\langle D|A_i\rangle - \langle C|A_i\rangle\langle D|A^\ell\rangle\right)$$

$$-\langle A^j|A\rangle\langle A_k|B\rangle\left(\langle C|A^\ell\rangle\langle D|A_i\rangle - \langle C|A_i\rangle\langle D|A^\ell\rangle\right)$$

$$= \left(\langle A_k|A\rangle\langle A^j|B\rangle - \langle A^j|A\rangle\langle A_k|B\rangle\right)$$

$$\times\left(\langle C|A^\ell\rangle\langle D|A_i\rangle - \langle C|A_i\rangle\langle D|A^\ell\rangle\right). \qquad (4.59)$$

Finally,

$$\left\langle \begin{matrix} A^j \\ A_i \end{matrix} \middle| \hat{V}_N \middle| \begin{matrix} A^\ell \\ A_k \end{matrix} \right\rangle = \tfrac{1}{2}\sum_{ABCD} \langle AB|\hat{v}|CD\rangle \langle \Phi_0|\hat{P}|\Phi_0\rangle$$

$$= \langle A_k A^j|\hat{v}|A^\ell A_i\rangle_{\mathcal{A}} . \qquad (4.60)$$

4.5. Normal product form of operators

The operators in the second quantization formalism, Eqs. (3.37), (3.57) or (3.60), are automatically in the normal product form with respect to the true vacuum $|0\rangle$. This is why we did not have to consider contractions among these operators in calculating vacuum mean values, as in Sec. 3.3. In order to attain the same simplification in the hole-particle formalism, the operators must be in the N-product form as the operators (53) and (57) were assumed to be in the last two examples of the preceding Section. However, the standard second quantization operators are not automatically in the N-product form, as Examples 1-4 of the preceding Section clearly indicate. Let us thus examine if we can transform the operators to the N-product form. This is particularly important for the Hamiltonian operator $\hat{H}$, Eq. (3.6), which we write as the sum of general one- and two-particle operators (3.64) and (3.70), namely

$$\hat{H} = \hat{Z} + \hat{V} = \sum_{AB}\langle A|\hat{z}|B\rangle X_A^\dagger X_B + \tfrac{1}{2}\sum_{ABCD}\langle AB|\hat{v}|CD\rangle X_A^\dagger X_B^\dagger X_D X_C . \qquad (4.61)$$

To achieve our objective, we could, for example, rewrite $\hat{H}$ in terms of hole-particle operators (6), as we have done for the one-particle part in (25), and rearrange each operator product into the N-product form using anticommutation relations (15). This, however, would yield rather cumbersome expressions, which we tried to avoid by introducing pairings for mixed operators, Eqs. (33). Let us, therefore, transform (61) to the N-product form using Wick's theorem.

We shall start with the one-particle operator $\hat{Z}$. Applying Wick's theorem of the

hole-particle formalism to the operator product $X_A^\dagger X_B$, we get

$$X_A^\dagger X_B = N[X_A^\dagger X_B] + N[\overbrace{X_A^\dagger X_B}]$$
$$= N[X_A^\dagger X_B] + \overbrace{X_A^\dagger X_B}$$
$$= N[X_A^\dagger X_B] + h(A)\langle A|B\rangle. \tag{4.62}$$

Using this result we can write the operator $\hat{Z}$ in the form

$$\hat{Z} = \sum_{AB}\langle A|\hat{z}|B\rangle\, N[X_A^\dagger X_B] + \sum_{AB}\langle A|\hat{z}|B\rangle h(A)\langle A|B\rangle$$
$$= \sum_{AB}\langle A|\hat{z}|B\rangle\, N[X_A^\dagger X_B] + \sum_{A_1}\langle A_1|\hat{z}|A_1\rangle \tag{4.63}$$

or, using the result of Example 1 of Sec. 4.4, Eq. (38), we get

$$\hat{Z} = \langle \Phi_0|\hat{Z}|\Phi_0\rangle + Z_N, \tag{4.64}$$

where Z_N designates the N-product form of $\hat{Z}$ given by Eq. (53).

We now apply the same technique to the two-particle operator $\hat{V}$. Thus

$$X_A^\dagger X_B^\dagger X_D X_C = N[X_A^\dagger X_B^\dagger X_D X_C] + N[\overbrace{X_A^\dagger X_B^\dagger} X_D X_C]$$
$$+N[\overbrace{X_A^\dagger X_B^\dagger X_D} X_C] + N[X_A^\dagger \overbrace{X_B^\dagger X_D} X_C] + N[X_A^\dagger \overbrace{X_B^\dagger X_D X_C}]$$
$$+N[\overbrace{X_A^\dagger X_B^\dagger}\, \overbrace{X_D X_C}] + N[\overbrace{X_A^\dagger X_B^\dagger X_D X_C}]$$
$$= N[X_A^\dagger X_B^\dagger X_D X_C] - h(A)\big(\langle A|D\rangle N[X_B^\dagger X_C] - \langle A|C\rangle N[X_B^\dagger X_D]\big)$$
$$+h(B)\big(\langle B|D\rangle N[X_A^\dagger X_C] - \langle B|C\rangle N[X_A^\dagger X_D]$$
$$+h(A)h(B)\big(\langle A|C\rangle\langle B|D\rangle - \langle A|D\rangle\langle B|C\rangle\big), \tag{4.65}$$

which yields when substituted into the second quantized form of $\hat{V}$ the expression

$$\hat{V} = \tfrac{1}{2}\sum_{ABCD}\langle AB|\hat{v}|CD\rangle N[X_A^\dagger X_B^\dagger X_D X_C]$$
$$-\tfrac{1}{2}\sum_{BC}\sum_{A_1}\langle A_1 B|\hat{v}|C A_1\rangle\, N[X_B^\dagger X_C]$$
$$+\tfrac{1}{2}\sum_{BD}\sum_{A_1}\langle A_1 B|\hat{v}|A_1 D\rangle\, N[X_B^\dagger X_D]$$
$$+\tfrac{1}{2}\sum_{AC}\sum_{B_1}\langle A B_1|\hat{v}|C B_1\rangle\, N[X_A^\dagger X_C]$$
$$-\tfrac{1}{2}\sum_{AD}\sum_{B_1}\langle A B_1|\hat{v}|B_1 D\rangle\, N[X_A^\dagger X_D]$$
$$+\tfrac{1}{2}\sum_{A_1B_1}\big(\langle A_1 B_1|\hat{v}|A_1 B_1\rangle - \langle A_1 B_1|\hat{v}|B_1 A_1\rangle\big)$$
$$= \tfrac{1}{2}\sum_{ABCD}\langle AB|\hat{v}|CD\rangle\, N[X_A^\dagger X_B^\dagger X_D X_C]$$
$$+\sum_{AB}\sum_{C_1}\langle A C_1|v|B C_1\rangle_A\, N[X_A^\dagger X_B]$$
$$+\tfrac{1}{2}\sum_{A_1A_2}\langle A_1 A_2|\hat{v}|A_1 A_2\rangle_A. \tag{4.66}$$

Recalling the result of Example 2 of Sec. 4.4, we can identify the last term in (66) with the mean value $\langle \Phi_0|\hat{V}|\Phi_0\rangle$. The second term represents in turn a new one-particle operator or, rather, its N-product form. Thus, defining

$$\hat{G} = \sum_{AB}\langle A|\hat{g}|B\rangle\, X_A^\dagger X_B, \tag{4.67}$$

where

$$\langle A|\hat{g}|B\rangle = \sum_{C_1}\langle A C_1|\hat{v}|B C_1\rangle_A$$
$$= \sum_{C_1}\big(\langle A C_1|\hat{v}|B C_1\rangle - \langle A C_1|\hat{v}|C_1 B\rangle\big), \tag{4.68}$$

we can write the two-particle part of $\hat{H}$ in the form

$$\hat{V} = \langle \Phi_0|\hat{V}|\Phi_0\rangle + G_N + V_N, \tag{4.69}$$

where the normal product form of G is clearly

$$G_N = \sum_{AB}\langle A|\hat{g}|B\rangle\, N[X_A^\dagger X_B], \tag{4.70}$$

while $\hat{V}_N$ is defined by (57).

All together, the Hamiltonian (61) can be now

expressed through N-product operators using (64) and (69) as follows

$$\hat{H} = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle + \hat{Z}_N + \hat{G}_N + \hat{V}_N \tag{4.71}$$

since

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \langle \Phi_0 | \hat{z} | \Phi_0 \rangle + \langle \Phi_0 | \hat{V} | \Phi_0 \rangle, \tag{4.72}$$

representing the energy of the Fermi vacuum state. We thus define the N-product form of the Hamiltonian (61) as

$$\hat{H}_N \equiv \hat{H} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \hat{F}_N + \hat{V}_N \tag{4.73}$$

where

$$\hat{F}_N = \hat{Z}_N + \hat{G}_N \tag{4.74}$$

or, generally,

$$\hat{F} = \hat{Z} + \hat{G} = \sum_{A,B} \langle A | f | B \rangle X_A^\dagger X_B, \tag{4.75}$$

so that

$$\begin{aligned} \langle A | \hat{f} | B \rangle &= \langle A | \hat{z} | B \rangle + \langle A | \hat{g} | B \rangle \\ &= \langle A | \hat{z} | B \rangle + \sum_{C_1} \langle A C_1 | \hat{v} | B C_1 \rangle_A . \end{aligned} \tag{4.76}$$

The physical significance of this new one-particle operator $\hat{F}$ will become apparent later on.

We now reconsider some of the examples of Sec. 4.4 in the light of the definitions introduced above.

1. Calculate the Fermi vacuum mean value of the Hamiltonian $\hat{H}$, Eq. (61).

In view of the results (63) and (66) and the hole-particle version of the rule 4, Eq. (2.39), we find immediately that the mean values of the N-product terms vanish so that

$$\begin{aligned} \langle \Phi_0 | \hat{z} | \Phi_0 \rangle &= \sum_{AB} \langle A | z | B \rangle \langle \Phi_0 | N[X_A^\dagger X_B] | \Phi_0 \rangle + \sum_{C_1} \langle C_1 | \hat{z} | C_1 \rangle \\ &= \sum_{C_1} \langle C_1 | \hat{z} | C_1 \rangle \end{aligned} \tag{4.77}$$

and, similarly

$$\langle \Phi_0 | \hat{V} | \Phi_0 \rangle = \tfrac{1}{2} \sum_{C_1 C_2} \langle C_1 C_2 | \hat{v} | C_1 C_2 \rangle_A . \tag{4.78}$$

Clearly

$$\langle \Phi_0 | \hat{H}_N | \Phi_0 \rangle = 0 \tag{4.79}$$

and

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \sum_{C_1} \langle C_1 | \hat{z} | C_1 \rangle + \tfrac{1}{2} \sum_{C_1 C_2} \langle C_1 C_2 | \hat{v} | C_1 C_2 \rangle_A . \tag{4.80}$$

2. Reconsider Example 4, Sec. 4.4, using the N-product form for $\hat{V}$.

Using (69), we can write matrix element (47) as

$$\begin{aligned} \langle \Phi_0 | \hat{V} | \Phi' \rangle &= \langle \Phi_0 | \hat{V} X_F^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle \\ &= \langle \Phi_0 | \hat{V} | \Phi_0 \rangle \langle \Phi_0 | X_F^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle \\ &+ \sum_{AB} \langle A | \hat{g} | B \rangle \langle \Phi_0 | N[X_A^\dagger X_B] X_F^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle \\ &+ \tfrac{1}{2} \sum_{ABCD} \langle AB | \hat{v} | CD \rangle \langle \Phi_0 | N[X_A^\dagger X_B^\dagger X_D X_C] X_F^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle . \end{aligned} \tag{4.81}$$

On the basis of generalized Wick's theorem we now see immediately that the last term must vanish since we can never fully pair all the operators, four of them being inside the N-product and only two outside. The vacuum mean values in the remaining two terms are then easily evaluated yielding

$$\langle \Phi_0 | X_F^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle = \overbrace{X_F^\dagger Y_{A_i}^\dagger} = h(F) \langle F | A_i \rangle = \langle F | A_i \rangle \tag{4.82}$$

and

$$\begin{aligned} \langle \Phi_0 | N[X_A^\dagger X_B] X_F^\dagger Y_{A_i}^\dagger | \Phi_0 \rangle &= \overbrace{X_A^\dagger X_B \, X_F^\dagger} \, Y_{A_i}^\dagger \\ &= h(A) p(B) \langle A | A_i \rangle \langle B | F \rangle \end{aligned}$$

$$= \langle A|A_i\rangle\, p(F)\, \langle F|B\rangle, \tag{4.83}$$

so that

$$\langle \Phi_0|\hat{V}|\Phi'\rangle = \langle F|A_i\rangle\langle \Phi_0|\hat{V}|\Phi_0\rangle + p(F)\langle A_i|\hat{q}|F\rangle$$

$$= \langle F|A_i\rangle \tfrac{1}{2}\sum_{C_1C_2}\langle C_1C_2|\hat{v}|C_1C_2\rangle_{\mathcal{A}} + p(F)\sum_{C_1}\langle A_iC_1|\hat{v}|FC_1\rangle_{\mathcal{A}}, \tag{4.84}$$

where we have used (78) and (68).

5. Diagrammatic representation

In this chapter we present basic ideas and concepts of a diagrammatic or graphical representation of Wick's theorem which are then applied to a number of practical problems in the second part of this text. The purpose of this chapter is to outline general ideas of graphical representations for time-independent formalism, to derive basic rules needed in their applications and to illustrate them on simple examples.

5.1. Basic concepts

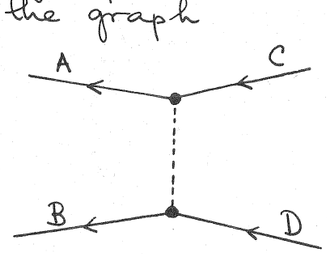
Applications of Wick's theorem can be greatly facilitated by introducing a diagrammatic representation of resulting contraction schemes. Such a representation is particularly useful if there is a one-to-one correspondence between its graphs and the algebraic expressions represented by them so that an easy transition can be made from one representation to the other. It is then possible to construct and to evaluate various expressions without any algebraic manipulations once the rules for the diagram handling are established.

To achieve such a representation we associate with each creation or annihilation operator a full oriented line, refered to as a fermion line. These lines carry the same spin-orbital labels as the operators they represent. In order to keep track of the origin of these lines we attach them to certain vertices, or sets of vertices interconnected by non-oriented lines which, <u>together</u> with the incident fermion lines*) represent various p-body operators occuring in the <u>problem</u> at hand. We note

*) When referring to these vertices we also have in mind the incident fermion lines unless stated otherwise.

that arbitrary configurations, or their linear combinations (see Sec. 5.4), may also be represented by certain excitation operators (cf., e.g., Ex. 7 of Sec. 1.7) relative to the chosen Fermi vacuum state. The non-oriented (usually dashed, dotted or wavy) lines and vertex sets (supervertices or s-vertices for short) used to represent a given operator are often referred to by the character of the operator involved: interaction, excitation, cluster, perturbation, etc. vertices or lines.

Initially we shall employ the so-called Goldstone graphs in which a p-body particle number conserving operator is represented by an s-vertex with p simple vertices, each simple vertex being incident with one entering and one leaving fermion line. These fermion lines are associated with spinorbitals having the same coordinates in the operator matrix element; the bra and ket spinorbitals corresponding to the labels on outgoing and ingoing lines, respectively. For example, a two-body operator $\hat{V}$, Eq. (3.70), is represented by the graph

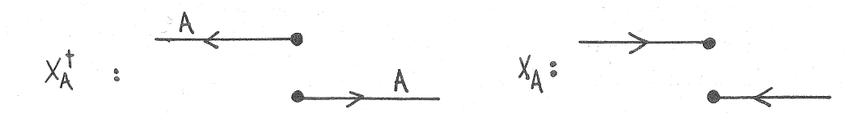


(5.1)

which we shall call a V-graph or a V-vertex. Here, each fermion line represents one creation or one annihilation operator in (3.70), and the s-vertex (called V-vertex in this case), consisting of two simple vertices joined together by the operator (interaction) dotted line, is associated with the matrix element $\langle AB|\hat{v}|CD\rangle$. This matrix element is refered to as the scalar factor d_V associated with the V-vertex (graph) (1), while the operator $\hat{D}_V$ associated with this graph is given by the product of d_V and $X_A^\dagger X_B^\dagger X_D X_C$. The operator $\hat{D}$

is then given by the sum of $\hat{D}_V$ operators over all spinorbital labels. Later, we shall also associate certain weights with each graph which will determine the numerical factor multiplying the operator matrix element ($\frac{1}{2}$ in the case of V-vertex (1)).

In general, creation operators $X_A^\dagger$ will be represented by oriented lines leaving a vertex and the annihilators X_A by fermion lines entering a vertex; schematically:



(5.2)

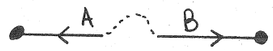
A general rule associating a scalar factor with a given s-vertex is also very simple: bra spinorbitals are given by the labels of outgoing lines and ket labels by the corresponding ingoing ones.

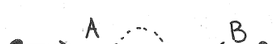
When more than one operator appears in a given algebraic expression, we assign a supervertex with each operator. Since these operators do not necessarily commute, we place their s-vertices along a hypothetical horizontal line* in the same order in which they appear in the algebraic expression. It is thus convenient to place individual vertices of a given supervertex in the vertical direction as in (1). The individual s-vertices can also be displaced vertically, as convenient for a graphical representation, as long as their left right ordering is unambiguous. The s-vertices representing commuting operators can be obviously placed arbitrarily relative one to the other, and will often be placed along a hypothetical vertical axis located at the appropriate horizontal location to stress this fact.

---

*) This convention is not followed universally. Some authors prefer to place the s-vertices along a hypothetical vertical line from the bottom upwards. Corresponding diagrams are obtained from ours by a clockwise rotation by the angle $\pi/2$.

Once we have graphically represented all the operators appearing in a given algebraic expression which we wish to evaluate, we can apply Wick's theorem or its generalized version and construct all admissible contraction schemes. This step is graphically represented by pairwise joining of the fermion lines of contracted operators. We easily find that in order to obtain a non-vanishing contraction, the orientation of fermion lines must be preserved in this process as illustrated below:

$\overbrace{X_A X_B} = 0 =$ 

$\overbrace{X_A^\dagger X_B^\dagger} = 0 =$ 

(5.3)

$\overbrace{X_A^\dagger X_B} = h(A)\langle A|B\rangle =$ 

$\overbrace{X_A X_B^\dagger} = p(A)\langle A|B\rangle =$ 

We shall also see later that this operation automatically picks up a correct hole-particle component of the operators involved (cf., e.g., Eq. (4.25)).

Thus, in contrast to external lines representing creation and annihilation operators, which are incident with only one vertex (cf. (2)), each contraction is represented by some internal fermion line connecting two vertices. Each such internal line having two spinorbital labels, say A and B, at its ends, cf. Eqs. (3), represents a Kronecker delta symbol $\langle A|B\rangle$. If at least one of the labels A and B represents a summation index, we can eliminate one summation and use a single spinorbital label for the internal line. The possible cases are listed schematically below:

$\sum_{AB}$ 

$\sum_{AB}$ 

(a)

$\sum_{B}$ 

$\sum_{B}$ 

(b)





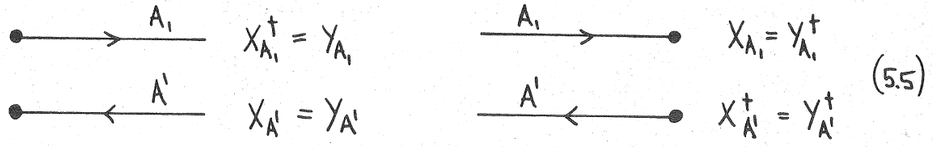
(c) (5.4)

In this context, as well as for later definition of weight factors, it is convenient to distinguish the summation labels from those which are fixed externally. We shall refer to the former as free labels and to the latter as fixed ones. For example, in expression (4.54), A, B are free labels while A_i, A^j, A_k and A^l are fixed labels.

We also see from Eqs. (3) that the left-right orientation of each fermion line determines its particle or hole character. This also enables us to eliminate the p and h step functions, Eqs. (4.26), from the resulting expressions by making the convention that the internal lines
(i) oriented from left to right represent hole lines, and those
(ii) oriented from right to left represent particle lines.
This convention is reflected in the summation conventions (4) and shows the importance of left-right ordering of s-vertices in our diagrams. It does not apply to the internal lines interconnecting the vertices of the same s-vertex, which are always hole lines since they originate from contractions of the type $\overbrace{X_A^\dagger X_B}$ as

follows from the standard form of a general operator, e.g. Eq. (3.60), which is the n-product form with respect to the true vacuum state $|0\rangle$.

Using the same convention for external lines, we can characterize the four possible cases as follows



(5.5)

We can thus conclude that the lines representing particle-hole annihilation operators extend to the right, and the creators to the left, of incident vertices.

Thus, representing each pairing by joining of corresponding fermion lines we form all admissible <u>resulting diagrams</u> (<u>R-diagrams</u>), each one representing a non-vanishing contraction scheme on the right hand side of Eq. (4.34). The type of admissible diagrams depends on the problem at hand, and for the two important cases discussed earlier (cf. rules R6 and R7 of Sec. 2.6 or their hole-particle analogues) may be characterized as follows:

(a) When evaluating a Fermi-vacuum mean value of some operator product, only the R-diagrams having no external lines (i.e., all fermion lines must be internal lines) need be considered, while

(b) when evaluating the result of an action of an operator product on the Fermi vacuum state, only the R-diagrams with no external lines extending to the right give a nonvanishing contribution.

Indeed, these rules follow immediately from the identification (5).

Once the R-diagrams are drawn we label their fermion lines as follows:

(i) all spin-orbital labels are given a hole or a particle character according to their left right orientation,

(ii) all external lines keep their original labels (if the direction of such a line is inconsistent with its particle or hole character, as the case may be, the R-diagram vanishes), and finally

(iii) all internal lines are labeled
(a) with a single summation label if both original labels were free,
(b) by a fixed label if one label was free and one fixed, and
(c) by two fixed labels (yielding a Kronecker delta in the resulting algebraic expression) if both labels were initially fixed.

We then evaluate each R-diagram by writing down the scalar factors associated with each s-vertex of the R-diagram and carry out summations over all free labels. We also assign appropriate creation and annihilation operators with all external lines, if any are present. For the time being we shall determine an appropriate sign as for the corresponding contraction scheme.

We also recall that when any operator is in the N-product form, we ignore all contractions between its operators (cf. generalized Wick theorem). In graphical language this means that all R-diagrams with internal fermion lines which are incident with vertices (or vertex) of the same s-vertex are ignored.

We can thus summarize the basic steps in our diagrammatic approach:

(i) represent all operators by s-vertices,
(ii) form all admissible R-diagrams, and
(iii) evaluate each R-diagram.

It is apparent from the above considered examples (e.g., Ex. 3, 4 and 6 of Sec. 4.4) that several distinct

contraction schemes can yield an identical algebraic expression. We shall see that such terms yield also equivalent R-diagrams so that any unnecessary duplication can be avoided in the diagrammatic approach. It will be essential, however, to formulate the rules which yield a correct numerical factor for each nonequivalent R-diagrams and a correct sign originating from the unscrambling of individual contractions. We shall formulate these rules in later sections.

5.2. Examples

To illustrate the above outlined ideas of the diagrammatic implementation of Wick's theorem we present two simple examples: (i) an evaluation of Hamiltonian matrix elements between mono-excited states and (ii) an action of a one-particle operator on mono-excited configurations. The first example closely parallels Examples 5 and 6 of Sec. 4.4 where a purely algebraic derivation was given for the matrix elements of one- and two-particle operators in the N-product form between two arbitrary monoexcited configurations.

5.2.1. Hamiltonian matrix elements between mono-excited configurations.

We write a general mono-excited configuration which results when a spin-orbital A_i, occupied in $|\Phi_o\rangle$, is replaced by an unoccupied spin-orbital A^j as

$$\left|\begin{matrix} A^j \\ A_i \end{matrix}\right\rangle \equiv |\Phi_i^j\rangle = X_{Aj}^\dagger X_{A_i} |\Phi_o\rangle = Y_{Aj}^\dagger Y_{A_i}^\dagger |\bar{\Phi}_o\rangle \tag{5.6}$$

We can also write

$$|\Phi_i^j\rangle = N\left[Y_{Ai}^\dagger Y_{A_i}^\dagger\right]|\Phi_o\rangle = N\left[X_{Aj}^\dagger X_{A_i}\right]|\Phi_o\rangle \tag{5.7}$$

or

$$\left|\begin{matrix} A^j \\ A_i \end{matrix}\right\rangle \equiv |\Phi_i^j\rangle = \hat{C}(i;j)\,|\Phi_o\rangle, \tag{5.8}$$

where

$$\hat{C}(i;j) \equiv \hat{C} = N\left[X_{Ai}^\dagger X_{A_j}\right] = X_{Aj}^\dagger X_{A_j} \tag{5.9}$$

is a <u>mono-excitation operator</u> associated with a given spin-orbital pair.

In calculating of a general matrix element of $\hat{H}$, Eq. (4.61), between arbitrary monoexcited configurations we shall exploit the N-product form of $\hat{H}$, Sec. 4.5, so that

$$\langle\Phi_i^j|\hat{H}|\Phi_k^\ell\rangle = \langle\Phi_i^j|\hat{H}_N|\Phi_k^\ell\rangle + \langle\Phi_o|\hat{H}|\Phi_o\rangle\langle\Phi_i^j|\Phi_k^\ell\rangle, \tag{5.10}$$

where the Fermi vacuum energy $\langle\Phi_o|\hat{H}|\Phi_o\rangle$ is given by Eqs. (4.38), (4.42), and

$$\langle\Phi_i^j|\hat{H}_N|\Phi_k^\ell\rangle = \langle\Phi_i^j|\hat{F}_N|\Phi_k^\ell\rangle + \langle\Phi_i^j|\hat{V}_N|\Phi_k^\ell\rangle \tag{5.11}$$

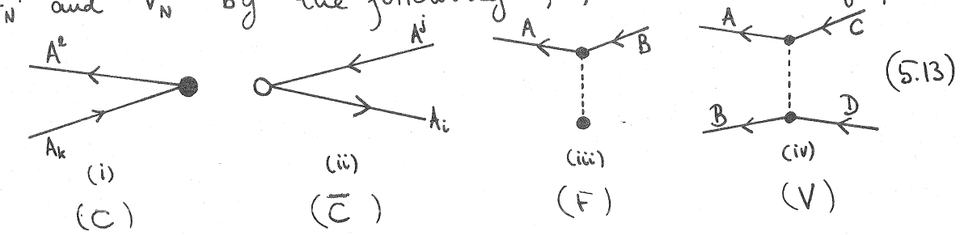
with $\hat{F}_N$ given by Eq. (4.74). The matrix element of $\hat{V}_N$ has been found algebraically in Ex. 4.4.6, while that of $\hat{F}_N$ is given by Ex. 4.4.5 with operator $\hat{Z}_N$ replaced by $\hat{F}_N$. The overlap integral $\langle\Phi_i^j|\Phi_k^\ell\rangle$ is easily found to be

$$\langle\Phi_i^j|\Phi_k^\ell\rangle = \left\langle\begin{matrix} A^j \\ A_i \end{matrix}\middle|\begin{matrix} A^\ell \\ A_k \end{matrix}\right\rangle = \langle\Phi_o|\hat{C}^\dagger(i;j)\,\hat{C}(k;\ell)\,|\Phi_o\rangle$$

$$= \langle\Phi_o|N\left[X_{A_i}^\dagger X_{A^j}\right]N\left[X_{A^\ell}^\dagger X_{A_k}\right]|\Phi_o\rangle$$

$$= \overbrace{X_{A_i}^\dagger \overbrace{X_{A^j} X_{A^\ell}^\dagger}\, X_{A_k}} = \langle A_i|A_k\rangle\langle A^j|A^\ell\rangle. \tag{5.12}$$

We shall now carry out the evaluation of matrix elements (10), (11) graphically. We thus represent the operators $\hat{C} \equiv \hat{C}(k;\ell)$, $\hat{C}^\dagger \equiv \hat{C}^\dagger(i;j)$, $\hat{F}_N$ and $\hat{V}_N$ by the following C, $\bar{C}$, F and V vertices or graphs



(5.13)

With these s-vertices or diagrams we associate the scalar factors

$$d_c \equiv d_c(A^e; A_k) = 1; \qquad d_{\bar{c}} = d_c^* \tag{5.14a}$$

$$d_F \equiv d_F(A;B) = \langle A|\hat{f}|B\rangle \tag{5.14b}$$

$$d_V \equiv d_V(AB;CD) = \langle AB|\hat{r}|CD\rangle\,, \tag{5.14c}$$

and the operators

$$\hat{D}_c \equiv \hat{D}_c(A^e; A_k) = d_c(A^e; A_k)\, N[X^+_{A^e} X_{A_k}] \tag{5.15a}$$

$$\hat{D}_{\bar{c}} \equiv \hat{D}_{\bar{c}}(A^j; A_i) = \hat{D}^+_c(A^j; A_i) = d^*_c(A^j; A_i)\, N[X^+_{A_i} X_{A^j}] \tag{5.15a'}$$

$$\hat{D}_F \equiv \hat{D}_F(A;B) = d_F(A;B)\, N[X^+_A X_B] \tag{5.15b}$$

$$\hat{D}_V \equiv \hat{D}_V(AB;CD) = d_V(AB;CD)\, N[X^+_A X^+_B X_D X_C]\,. \tag{5.15c}$$

We note that A, B, C and D are free while A_i, A^j, A_k and A^e are fixed labels, so that the relevant operators are obtained by summing the operator expressions (15) assigned to diagrams (13) over all free labels.

In general an operator $\hat{K}$ associated with the diagram $K(\mathfrak{x})$, where $\mathfrak{x}$ designates the list of spinorbital labels carried by the external lines of K, is given as

$$\hat{K} = w_K \sum_{\mathfrak{x}}{}' \hat{D}_K(\mathfrak{x})\,, \tag{5.16}$$

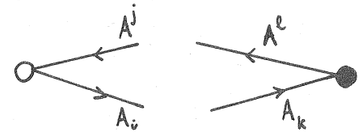
where $\hat{D}_K(\mathfrak{x}) = d_K(\mathfrak{x})\, TT(\mathfrak{x})$ is the operator associated with the diagram $K(\mathfrak{x})$, $d_K(\mathfrak{x})$ is the scalar factor and the summation extends over all free labels from $\mathfrak{x}$. $TT(\mathfrak{x})$ designates the appropriate product of creation and annihilation operators and w_K is the numerical factor. We will give precise definitions of these quantities later.

We also note that the particle-hole character of fermion lines is fully determined in the graphs (i) and (ii) for the excitation operators,

since their lines are labeled with fixed labels of a definite particle-hole character and directed accordingly (cf., associations (5)). In contrast the lines in F and V s-vertices carry free labels, corresponding to summations over all spin-orbitals, and may thus be directed either to the left or to the right of incident vertices. We shall see how the diagrammatic procedure employed picks up automatically appropriate hole-particle components of these operators.

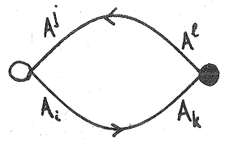
Finally, in the case considered, all operators are in the N-product form, so that the N-products of relevant operator products appear in corresponding operators (15). To be precise, we should indicate this fact in appropriate graphs (say, by the kind of operator or interaction lines) in order to distinguish them from those representing corresponding ordinary operators (cf., Eqs. (3.10) and (1)). However, this would be an unnecessary pedantry since we shall always remember whether a given operator is or is not in the N-product form. In the affirmative case, all R-diagrams in which the fermion lines of the same s-vertex are joined must be avoided.

We can now proceed with diagrammatic evaluation of various matrix elements in (10) and (11). Starting with the overlap integral (12), we represent excitation operators $\hat{C}^+$ and $\hat{C}$ by diagrams (ii) and (i) of (13), in this order, namely



(5.17)

and the only system of pairings in (12), which results from the application of (generalized) Wick's theorem, by joining of corresponding particle and hole lines, obtaining the following R-diagram:



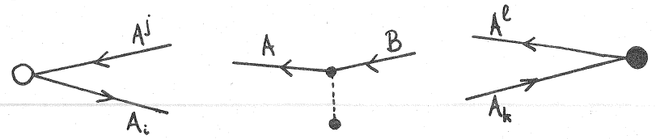
(5.18)

Equivalently, once we represent excitation operators by $\bar{C}$ and C diagrams, as in (17), we can proceed with the application of Wick's theorem, and thus with the evaluation of the desired quantity, in a graphical way: we simply construct all permissible R-diagrams which, in this case, must have only internal lines since we evaluate a Fermi vacuum mean value of $\hat{\bar{C}}^t\hat{\bar{C}}$. We see immediately that only one Rdiagram is possible in this case, namely the diagram (18), since all internal lines must interconnect different s-vertices our operators being in the N-product form. We then evaluate the R-diagram algebraically as outlined in Sec. 5.1. In the present case we only have to multiply scalar factors associated with s-vertices and write down appropriate Kronecker deltas for each fermion line:

$$d_{\bar{C}}\, d_C \langle A_i|A_k\rangle\langle A^j|A^e\rangle = \langle A_i|A_k\rangle\langle A^j|A^e\rangle, \qquad (5.19)$$

where we used (14a). Thus, we obtained the result given by Eq. (12) diagrammatically.

Consider next the one-electron matrix element $\langle \Phi_i^j|\hat{F}_N|\Phi_k^e\rangle$. We shall attempt a direct diagrammatic evaluation of this element and only subsequently will compare it with our result obtained algebraically in Example 4.4.5. We thus represent the relevant operators, namely the operators $\hat{\bar{C}}^t$, $\hat{F}_N$ and $\hat{C}$ (in this order!) by $\bar{C}$, F and C graphs (s-vertices) as follows

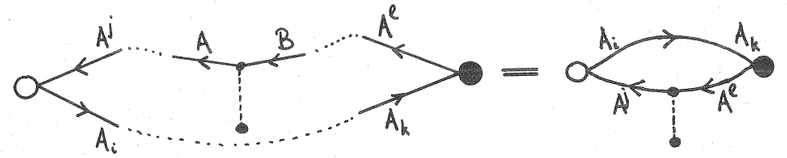


(5.20)

and construct permissible R-diagrams (no external lines, no lines issuing and ending on the same s-vertex,

and orientation of lines preserved in joining them). Considering, say, the operator line labeled A, we see immediately that we can only join it with either (i) $\bar{C}$-diagram line A^j or (ii) C-diagram line A_k. In each case there is only one possibility to join the remaining lines in order that a legitimate R-diagram is obtained. We thus get the following R-diagrams

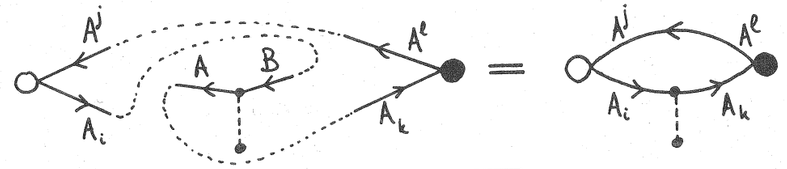
(i)



(5.21a)

$$N[\overbrace{X^t_{A_i}\overbrace{X_{A^j}X^t_A}X_B\overbrace{X^t_{A^e}X_{A_k}}}]$$

(ii)

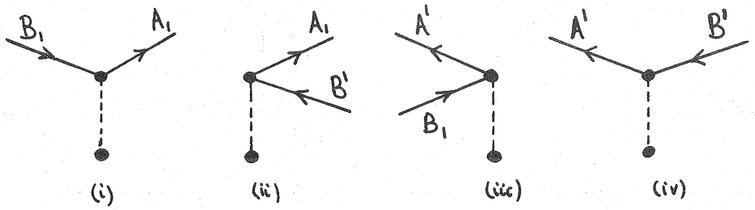


(5.21b)

$$N[\overbrace{X^t_{A_i}X_{A^j}}\overbrace{X^t_A X_B}\overbrace{X^t_{A^e}X_{A_k}}]$$

which we have re-drawn in a simpler form on the right hand side and in which we applied labeling conventions derived in Sec. 5.1, relations (4). We note that we have automatically avoided the two contraction schemes in (4.55), which yield vanishing contributions, while those giving non-zero result are shown beneath corresponding diagrams in (21).

We also observe that in constructing permissible R-diagrams we automatically select relevant hole-particle components of the operator $\hat{F}_N$. We could have, of course, transformed this operator to the hole-particle form, as we did for the operator $\hat{Z}$ in Eq. (4.25), obtaining four different terms. These terms would be represented by the following diagrams,



(5.22)

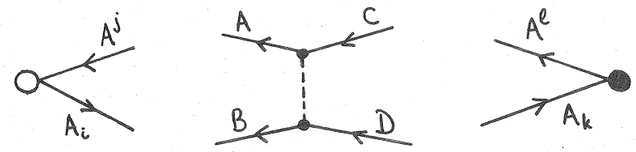
with the left-right direction of fermion lines fixed as shown. Inserting these diagrams between the $\bar{C}$ and C diagrams, as in (20), and constructing feasible R-diagrams, we realize that only the diagram (iv) contributes in case (i), Eq. (21a), and the diagram (i) in case (ii), Eq. (21b), while no legitimate R-diagram can be obtained using other components. We see, however, that the same components are automatically selected when we leave the left-right direction of fermion lines of the operator $\hat{f}_N$ unspecified and construct possible R-diagrams as in (21).

We can, finally, evaluate the R-diagrams (21). This is simply done by assigning to each diagram a product of scalar factors associated with its s-vertices and of Kronecker deltas for each doubly labeled line (there will be no operators present in this case or summations to carry out since diagrams (21) have no external lines or free labels, respectively). Thus, the following algebraic expressions may be immediately associated with the R-diagrams (21),

$$\begin{aligned} &\text{(i)} \quad \langle A^j|\hat{f}|A^\ell\rangle\langle A_k|A_i\rangle \\ &\text{(ii)} \quad \langle A_k|\hat{f}|A_i\rangle\langle A^j|A^\ell\rangle \end{aligned} \qquad (5.23)$$

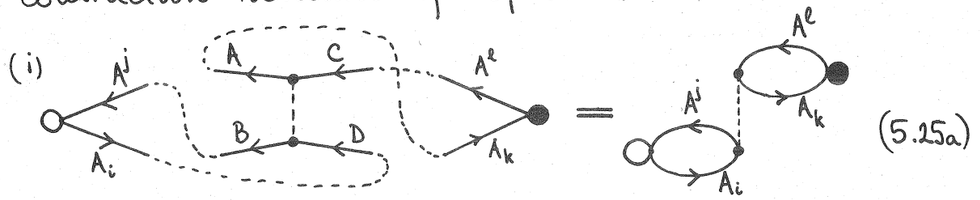
since $d_C = d_{\bar{C}} = 1$. Up to the sign factor of the second term, these are exactly the expressions obtained algebraically in Eq. (4.56) when the operator $\hat{F}_N$ is replaced by $\hat{f}_N$.

Consider, finally, the two-particle matrix element $\langle \Phi_i^j|\hat{V}_N|\Phi_k^\ell\rangle$. Representing the $\hat{C}^\dagger$, $\hat{V}_N$ and $\hat{C}$ operators by appropriate diagrams,



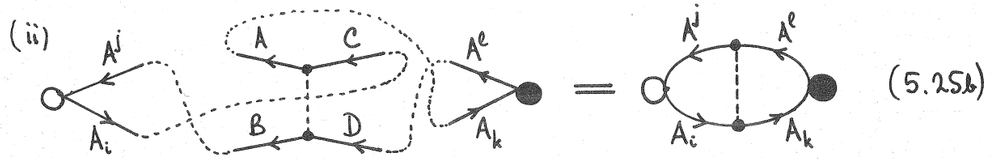
(5.24)

we can construct the following R-diagrams (recalling again the N-product form of our operators), which we may associate with corresponding contraction schemes of Eq. (4.58):



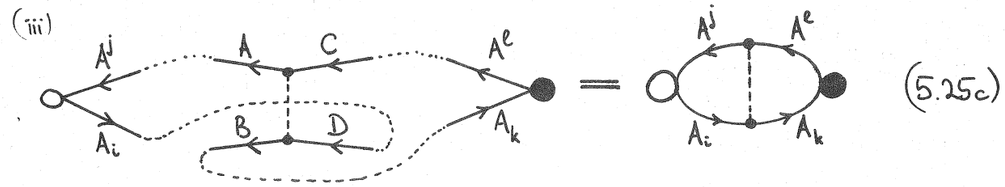
(5.25a)

$$N[X^\dagger_{A_i} X_{A^j} X^\dagger_A X^\dagger_B X_D X_C X^\dagger_{A^\ell} X_{A_k}]$$



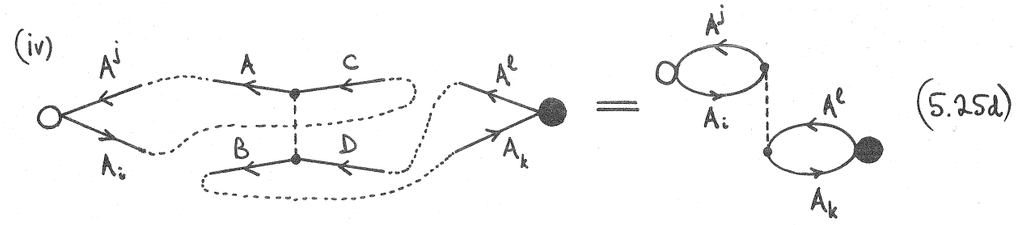
(5.25b)

$$N[X^\dagger_{A_i} X_{A^j} X^\dagger_A X^\dagger_B X_D X_C X^\dagger_{A^\ell} X_{A_k}]$$



(5.25c)

$$N[X^\dagger_{A_i} X_{A^j} X^\dagger_A X^\dagger_B X_D X_C X^\dagger_{A^\ell} X_{A_k}]$$

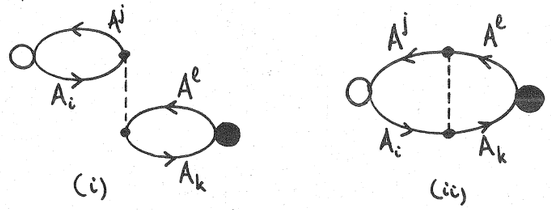


(5.25d)

$$N[X^\dagger_{A_i} X_{A^j} X^\dagger_A X^\dagger_B X_D X_C X^\dagger_{A^\ell} X_{A_k}]$$

We immediately realize that in fact two distinct contraction schemes yield identical R-diagrams,

so that we only have two distinct R-diagrams shown below



(5.26)

The algebraic expressions associated with these diagrams are given by the scalar factors

$$\begin{aligned} &\text{(i)} \quad \langle A^j A_k | \hat{v} | A_i A^e \rangle \\ &\text{(ii)} \quad \langle A^j A_k | \hat{v} | A^e A_i \rangle \end{aligned} \tag{5.27}$$

(first list outgoing lines in the bra and then corresponding ingoing lines in the ket), since all labels are fixed labels so that no summations are present. Except for the sign of the second term, this result is identical with the result of Ex. 4.4.6, Eq. (4.60). However, we must note that we have obtained the correct numerical factor by a lucky coincidence, since the factor of $\frac{1}{2}$ appearing in the operator expression for $\hat{V}_N$, Eq. (4.57), is compensated by the factor of two which results due to the fact that each diagram (26) is obtained twice when the generalized Wick theorem is applied. We shall see later how to determine a correct numerical factor for each distinct R-diagram directly from the diagram itself, without actually constructing different contraction schemes. It should also be noted that from the 16 different hole-particle versions of the diagram (iv) of (13) the diagrammatic method automatically selects the appropriate ones just as in the case of a one-particle operator discussed in greated detail above.

In order to better illustrate the rôle of free and fixed labels in the diagrammatic

approach, let us generalize the above considered example to the case of a matrix element between states $|\Phi_x^{(i)}\rangle$, $x=a,b$, given as general linear combinations of monoexcited configurations

$$\begin{aligned} |\Phi_x^{(i)}\rangle &= \sum_{i,j} c_{ji;}^{(x)} |\Phi_i^j\rangle \\ &= \sum_{A' A_1} \langle A' | \hat{c}_x | A_1 \rangle \left| \begin{smallmatrix} A' \\ A_1 \end{smallmatrix} \right\rangle . \end{aligned} \tag{5.28}$$

We can write again

$$|\Phi_x^{(i)}\rangle = \hat{C}_x^{(i)} |\Phi_0\rangle , \tag{5.29}$$

where

$$\begin{aligned} \hat{C}_x^{(i)} &= \sum_{A' A_1} \langle A' | \hat{c}_x | A_1 \rangle X_{A'}^{\dagger} X_{A_1} \\ &= \sum_{A' A_1} \langle A' | \hat{c}_x | A_1 \rangle N\left[X_{A'}^{\dagger} X_{A_1} \right] \end{aligned} \tag{5.30}$$

represents an appropriate excitation operator. Notice that the matrix elements of this operator, $\langle A' | \hat{c}_x | A_1 \rangle$, are only defined between unoccupied and occupied spinorbitals.

Using the earlier derived expressions for matrix elements between configurations, we can easily find the desired matrix elements algebraically, obtaining

$$\langle \Phi_a^{(i)} | \hat{H} | \Phi_b^{(i)} \rangle = \langle \Phi_a^{(i)} | \hat{H}_N | \Phi_b^{(i)} \rangle + \langle \Phi_0 | \hat{H} | \Phi_0 \rangle \langle \Phi_a^{(i)} | \Phi_b^{(i)} \rangle \tag{5.31}$$

where

$$\langle \Phi_a^{(i)} | \Phi_b^{(i)} \rangle = \sum_{i,j} c_{ji;}^{(a)*} \; c_{ji;}^{(b)} = \sum_{A' A_1} \langle A_1 | \hat{c}_a | A' \rangle \langle A' | \hat{c}_b | A_1 \rangle \tag{5.32}$$

and

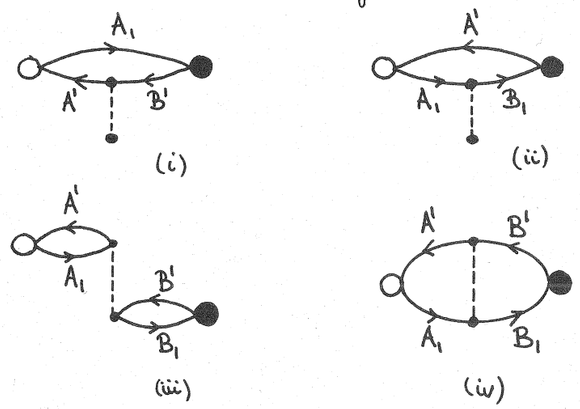
$$\begin{aligned} \langle \Phi_a^{(i)} | \hat{H}_N | \Phi_b^{(i)} \rangle = & \sum_{A' A_1} \langle A_1 | \hat{c}_a | A' \rangle \\ & \times \Big\{ \sum_{B'} \langle B' | \hat{c}_b | A_1 \rangle \langle A' | \hat{f} | B' \rangle - \sum_{B_1} \langle A' | \hat{c}_b | B_1 \rangle \langle B_1 | \hat{f} | A_1 \rangle \\ & + \sum_{B' B_1} \langle B' | \hat{c}_b | B_1 \rangle \langle A' B_1 | \hat{v} | A_1 B' \rangle_a \Big\} . \end{aligned} \tag{5.33}$$

In diagrammatic approach we can use again the representation (13) as in an earlier example, except that the scalar quantities associated with diagrams (i) and (ii) must now be re-defined as appropriate coefficients in (28) or (30), namely

$$d_c \equiv d_c(A^a; A_k) = \langle A^a | \hat{c}_b | A_k \rangle$$

$$d_{\bar{c}} \equiv d_{\bar{c}}(A_i; A^j) = d_c^*(A^j; A_i) = \langle A_i | \hat{c}_a | A^j \rangle. \quad (5.34)$$

All the spin-orbital labels will now be free labels so that the relevant R-diagrams are



(5.35)

where all the labels are again free labels. Applying the rules established in Sec. 5.1. we can immediately associate the following algebraic expressions with these diagrams:

$$\text{(i)} \sum_{A'A_1} \sum_{B'} \langle A_1|\hat{c}_a|A'\rangle\langle B'|\hat{c}_b|A_1\rangle\langle A'|\hat{f}|B'\rangle$$

$$\text{(ii)} \sum_{A'A_1} \sum_{B_1} \langle A_1|\hat{c}_a|A'\rangle\langle A'|\hat{c}_b|B_1\rangle\langle B_1|\hat{f}|A_1\rangle$$

$$\text{(iii)} \sum_{A'A_1} \sum_{B'B_1} \langle A_1|\hat{c}_a|A'\rangle\langle B'|\hat{c}_b|B_1\rangle\langle A'B_1|\hat{v}|A_1B'\rangle \quad (5.36)$$

$$\text{(iv)} \sum_{A'A_1} \sum_{B'B_1} \langle A_1|\hat{c}_a|A'\rangle\langle B'|\hat{c}_b|B_1\rangle\langle A'B_1|\hat{v}|B'A_1\rangle$$

These expressions are again identical, up to the sign, with the corresponding terms in (33).

5.2.2. Action of a one-particle operator on monoexcited configurations.

Consider a general single particle operator, say $\hat{Z}_N$, Eq. (4.53), which we assume to be in the N-product form for simplicity, to act on a general mono-excited configuration (6),

$$\hat{Z}_N \left|{}^{A^j}_{A_i}\right\rangle = \hat{Z}_N \hat{C}(i;j) \,|\Phi_o\rangle, \quad (5.37)$$

where $\hat{C}(i;j) \equiv \hat{C}$ is given by (9).

We shall first derive this result algebraically. We write

$$\hat{Z}_N \left|{}^{A^j}_{A_i}\right\rangle = \sum_{AB} \langle A|\hat{z}|B\rangle \, N[X_A^\dagger X_B] \, N[X_{A^j}^\dagger X_{A_i}] \,|\Phi_o\rangle, \quad (5.38)$$

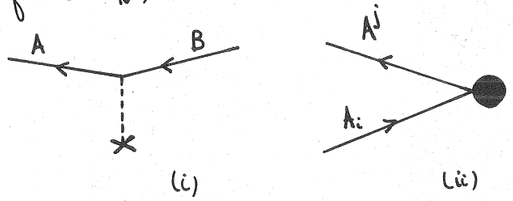
and evaluate the product of two N-products using generalized Wick's theorem

$$N[X_A^\dagger X_B] N[X_{A^j}^\dagger X_{A_i}] = N[X_A^\dagger X_B X_{A^j}^\dagger X_{A_i}]$$
$$+ N[X_A^\dagger \overbrace{X_B X_{A^j}^\dagger} X_{A_i}] + N[\overbrace{X_A^\dagger X_B} X_{A^j}^\dagger X_{A_i}]$$
$$+ N[\overbrace{X_A^\dagger X_B X_{A^j}^\dagger} X_{A_i}]. \quad (5.39)$$

Substituting this result into (38) and recalling the hole-particle version of rule R3, Eq. (38), we can conclude that only those terms from the summation over the spin-orbital states A and B will contribute, in which the operators $X_A^\dagger$ and X_B are converted to creators in the hole-particle representation. Thus, whenever these operators, originating from $\hat{Z}_N$, are uncontracted, we have to restrict the summation over A to the particle states and the summation over B to the hole states. Using, further, formulas (4.33), we get finally

$$\hat{Z}_N \left|{}^{A^j}_{A_i}\right\rangle = \Big\{ \sum_{C'} \sum_{C_1} \langle C'|\hat{z}|C_1\rangle N[X_{C'}^\dagger X_{C_1} X_{A^j}^\dagger X_{A_i}]$$
$$+ \sum_{C'} \langle C'|\hat{z}|A^j\rangle N[X_{C'}^\dagger X_{A_i}] - \sum_{C_1} \langle A_i|\hat{z}|C_1\rangle N[X_{A^j}^\dagger X_{C_1}]$$
$$+ \langle A_i|\hat{z}|A^j\rangle \Big\} |\Phi_o\rangle. \quad (5.40)$$

In the diagrammatic version of this derivation we first represent the operators $\hat{Z}_N$ and $\hat{C}$ graphically, as in Ex. 5.2.1 (we use a cross-type vertex for the one-particle operator in order to distinguish it from $\hat{F}_N$)



(i) (ii) (5.41)

With these Z and C graphs we associate scalar factors

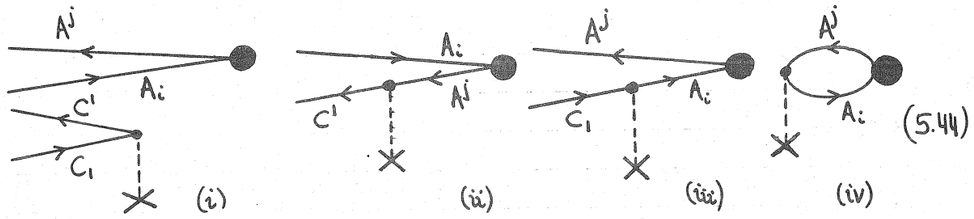
$$d_Z \equiv d_Z(A;B) = \langle A|\hat{z}|B\rangle$$
$$d_C \equiv d_C(A^j;A_i) = 1 \qquad (5.42)$$

and operators

$$\hat{D}_Z \equiv \hat{D}_Z(A;B) = d_Z(A;B)\, N[X_A^\dagger X_B]$$
$$\hat{D}_C \equiv \hat{D}_C(A^j;A_i) = d_C(A^j;A_i)\, N[X_{A^j}^\dagger X_{A_i}], \qquad (5.43)$$

so that the operators $\hat{Z}_N$ and $\hat{C}$ are given by the general formula (16) where $w_Z = w_C = 1$, and A,B are free and A^j, A_i fixed labels.

In this problem we do not evaluate a Fermi vacuum mean value of an operator product, so that we also have to admit R-diagrams having external lines, since also not fully contracted terms may contribute when we apply Wick's theorem. Since the Fermi vacuum occurs on the right hand side of the considered expression, Eq. (37), all external lines must be located on the left hand side of their incident vertices (since all annihilators Y_A must be contracted; cf. rule and representation (5)). We can thus obtain an R-diagram without joining any lines, as in (i) below, or by interconnecting a single line as in (ii) and (iii) and, finally, "contracting" all the lines as in (iv):



(i) (ii) (iii) (iv) (5.44)

The order of vertices in (i) is immaterial since both operators commute (being within the same N-product). We also labeled fermion lines in R-diagrams (44) following our conventions. We can, therefore, immediately write down the algebraic expressions represented by these diagrams:

$$\text{(i)} \quad \sum_{C'C_1} \langle C'|\hat{z}|C_1\rangle\, N[X_{A^j}^\dagger X_{A_i} X_{C'}^\dagger X_{C_1}]$$
$$\text{(ii)} \quad \sum_{C'} \langle C'|\hat{z}|A^j\rangle\, N[X_{C'}^\dagger X_{A_i}]$$
$$\text{(iii)} \quad \sum_{C_1} \langle A_i|\hat{z}|C_1\rangle\, N[X_{A^j}^\dagger X_{C_1}] \qquad (5.45)$$
$$\text{(iv)} \quad \langle A_i|\hat{z}|A^j\rangle$$

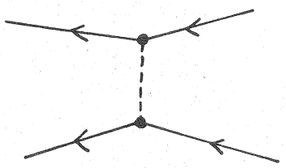
These expressions, acting on $|\Phi_0\rangle$, are easily recognized to be identical with individual terms in (40), except for the sign of the third one. We again observe how all algebraic manipulations leading to Eq. (40), including the selection of correct particle-hole components of the operator $\hat{Z}$, are automatically carried out by constructing pertinent R-diagrams.

5.3. Skeletons and weights

In order that our diagrammatic approach is complete, we must be able to determine correct numerical factors multiplying an algebraic contribution of each R-diagram (as given by the product of scalar factors and operators associated with its s-vertices and external lines, respectively). We shall

now outline a necessary formalism which will enable us to determine absolute values of these numerical factors. The signs, given by the parity of permutations "unscrambling" corresponding contraction schemes into a simple product of individual pairings, will be determined later (see Sec. 5.5).

We have seen in examples of Sec. 2 that free labels on fermion lines can be deleted when forming R-diagrams and re-inserted as needed, following the rules (4a) and (4b). It is thus helpful to distinguish <u>diagrams</u> (graphs in which every fermion line carries a spin-orbital label) from <u>skeletons</u>, which we define as diagrams stripped of all free labels. These skeletons possess a necessary combinatorial information, which determines required numerical factors, when we use appropriate graphs for s-vertices. For example, the skeleton which is associated with diagram (1), representing the operator V, Eq. (3.70), is



(5.46)

This skeleton has two equivalent vertices (or, equivalently, two pairs of equivalent fermion lines attached to them), which may be transformed one into the other (for example, by an in-plane rotation by π) without changing the skeleton. This symmetry reflects the symmetry of matrix elements of this operator as expressed by the first Eq (3.8)* and, together with the fact that the summations over spin-orbital labels in (3.70) are unrestricted, is closely related with the factor $(1/2)$ appearing in this expression for $\hat{V}$. We shall now formulate these observations more precisely.

*) Or by Eqs. (3.55) and (3.45).

We define a <u>topological transformation</u>*) (or <u>deformation</u>) of a skeleton as an arbitrary geometrical transformation in the three-dimensional space, in which we can imagine ourselves the skeleton to be "made" from elastic lines (both oriented and non-oriented), which can be arbitrarily moved, bent or stretched as long as no breaking and re-connecting of lines, no reordering of s-vertices along a hypothetical horizontal axis, and no change in the hole-particle character of fermion lines carrying fixed labels will take place.

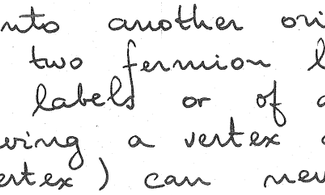
Let us remark that the restriction to preserve the ordering of s-vertices in topological transformations is typical of time-independent theories. There is no such restriction in time-dependent theories and the skeletons or diagrams which differ only in the order of their s-vertices are referred to as different time versions of the same diagram. This is also why the right-left direction, chosen to represent particle lines, is referred to as the time axis: particles "move" along with the time while holes "move" against the time. We may use this observation as a simple mnemonic rule in drawing our diagrams.

(9) Each topological transformation has the inverse. A topological transformation which transforms a skeleton S into itself is called an <u>automorphism</u> of S. Among the possible automorphisms of S we also include a trivial <u>identity automorphism</u> e

*) We use this term for the lack of a better one, implying that aspect of the theory of topological spaces which is sometimes refered to as a "rubber geometry".

which corresponds to a trivial topological transformation leaving the skeleton unchanged. Consequently, in a non-trivial automorphism, at least one oriented line of S must be transformed into another oriented line of S. Of course, any two fermion lines carrying distinct fixed labels or of different types (i.e., one leaving a vertex and the other entering a vertex) can never be transformed one into the other by any non-trivial automorphism.

When we regard the two automorphisms which transform all lines and vertices of a given skeleton S in the same way as <u>identical</u> (or <u>equivalent</u>), we immediately realize that all non-equivalent automorphisms form a group. Equivalence of two automorphisms g and h of S can be formally defined as follows: number arbitrarily all fermion lines by consecutive integers*. Then we can associate with each automorphism g a certain permutation P_g of integers carried by fermion lines. We say that g and h are equivalent, $g \sim h$, if $P_g = P_h$. For example, numbering the lines and vertices of skeleton (46) as follows



(5.47)

we see immediately that any automorphism of (46) is equivalent to either the identity automorphism e or the automorphism g, with

$$P_e = (1), \quad P_g = (12)(34), \tag{5.48}$$

where (1) is the identity permutation and $(i_1, i_2, \ldots, i_k)$ is a k-cycle. We can also define a product gh of two automorphisms of S and an inverse g^{-1} of g in an obvious way (gh is the result of h followed by g, so that $P_{gh} = P_g P_h$, while g^{-1} is an automorphism such that $gg^{-1} = g^{-1}g = e$ so that $P_{g^{-1}} = P_g^{-1}$). Since the product is clearly associative, we realize that indeed all non-equivalent automorphisms of S, including e, form a group: <u>the group Γ_S of automorphisms of S.</u>

The order $g_S \equiv g(S)$ of this group will be called a <u>topological factor</u> of S (or of a corresponding diagram D) and its inverse,

$$w(S) \equiv w_S = [g(S)]^{-1} \tag{5.49}$$

will be refered to as the <u>weight</u> of S. For example, $w_V = 1/2$ for the V-skeleton (46) or the V-diagram (1), since $g_V = 2$, Γ_V having two elements e and g (cf., Eq. (48)). Thus, the s-vertex (1) is appropriate for a representation of a general two-particle operator $\hat{V}$ when expressed using non-antisymmetrized matrix elements, Eq. (3.70).

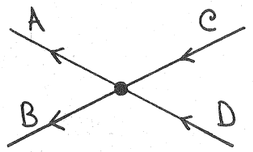
p.70a

In general, when choosing a diagrammatic representation for some operator we shall require that the numerical factor appearing in an algebraic expression for this operator is given by the weight of a corresponding skeleton so that Eq. (16) holds. This in turn requires that the chosen graph or s-vertex expresses the symmetry of the matrix elements used in the algebraic representation of a given operator. For example, when we wish to use the representation for a two-particle operator V using antisymmetrized matrix elements, i.e.

$$\hat{V} = \frac{1}{4} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle_A \, X_A^\dagger X_B^\dagger X_D X_C, \tag{5.50}$$

we shall employ a so-called <u>Hugenholtz</u> (or <u>degenerate</u>) vertex

* We could equivalently label the vertices as long as each simple vertex is incident with one entering and one leaving fermion line. We could also label separately the ingoing and outgoing lines.



(5.51)

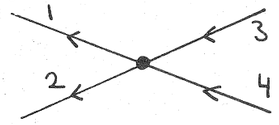
with which we associate the scalar factor

$$d_V^H = \langle AB \,|v|\, CD \rangle_A \qquad (5.52)$$

and the operator

$$\hat{D}_V^H = d_V \, X_A^\dagger X_B^\dagger X_D X_C \,. \qquad (5.53)$$

The weight w_V^H of the corresponding skeleton



(5.54)

in which we numbered fermion lines with consecutive integers, is $w_V^H = 1/4$ since Γ_V has now four elements characterized by permutations

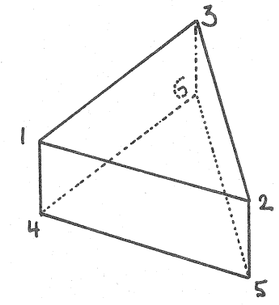
$$P_1 = (1), \; P_2 = (12), \; P_3 = (34), \; P_4 = (12)(34). \qquad (5.55)$$

Note that the sign of (52) or (53) is not fully determined by the diagram (51). The usefulness of such degenerate representations will become apparent later.

We shall conclude our discussion of topological factors with a few definitions and rules which will facilitate the determination of weights in later applications. We define two fermion <u>lines</u> or two <u>vertices</u> of a graph G to be <u>equivalent</u> if there exists an automorphism of the corresponding skeleton which transforms them one into the other. This equivalence relation makes it possible to partition all fermion lines and all vertices into disjoint <u>equivalence classes</u>. The number of lines or vertices in each class is a

⊛ (to be inserted on p.70):

In order to better elucidate the concept of a group of automorphisms of a given skeleton S, let us draw a parallel with the concept of a symmetry group of a rigid body, for example, the symmetry group of a prism with an equilateral triangular basis:



(5.)

Applying to this body operations (transformations, elements) of a three-dimensional Euclidean group (rotations about a fixed point, reflections, etc.) we observe that some of these operations transform the prism into itself (e.g., rotations by $k(\frac{2\pi}{3})$; $k = 0, 1, 2$ about the three-fold axis of the pyramid). We refer to such transformations as the symmetry operations or elements of a given body. The set of these operations, together with the identity transformation which leaves the body unchanged, form the symmetry group (in this case the finite point group D_{3h}) of the body. Thus, the operations of the Euclidean group play here the same role as topological transformations of skeletons while the symmetry operations of the body correspond to automorphisms of our graphs. Similarly, in order to distinguish different symmetry operations we can number individual vertices (or edges or faces) of our body as shown above

and assign permutations to various symmetry operations, just as we did for automorphisms. Again, different symmetry operations will be represented by distinct permutations, which will form a subgroup of the symmetric group S_N, where N is the number of labels on vertices (faces, edges, etc).

divisor of g_G. For Goldstone type diagrams the number of simple vertices in each class is the same as the number of outgoing (ingoing) fermion lines which are incident with these vertices. If such a diagram G possesses r equivalent classes, each one containing n_i equivalent vertices, $(i = 1, \dots, r)$, the topological factor $g(G)$ is given as

$$g(G) = \prod_{i=1}^{r} (n_i!) \quad . \tag{5.56}$$

Any two lines which are incident with the same vertex are called <u>adjacent.</u> A sequence of adjacent lines (oriented or non-oriented) is called a <u>path.</u> If the path contains only oriented lines it is called an <u>oriented</u> or <u>fermion path.</u> In Goldstone diagrams orientation must be preserved along each fermion path, irrespectively of how many vertices are involved or how many times the left-right direction of the path changes. Each oriented path has either two external lines or none. The latter is refered to as a <u>closed loop</u> (or, simply, a <u>loop</u>), while the former one is called an <u>open path</u>.

Any two loops or open paths of a Goldstone diagram have either no vertices or fermion lines in common or are identical. This property follows immediately from the fact that only two fermion lines (one entering and one leaving) are incident with each vertex of a Goldstone diagram.

A graph (diagram or skeleton) is <u>connected</u> if there is at least one path between any two vertices. A <u>disconnected</u> graph G (a graph which is not connected) consists of at least two connected parts and we write

$$G = \sum_i G_i \, , \tag{5.57}$$

where G_i are connected. In general, several connected parts of a disconnected diagram G may be equivalent, so that

$$G = \sum_{i=1}^{k} n_i G_i \quad , \tag{5.58}$$

where G_i are non-equivalent connected parts of G. The topological factor of G is then easily found to be

$$g(G) = \prod_{i=1}^{k} [n_i!\, g(G_i)^{n_i}] , \tag{5.59}$$

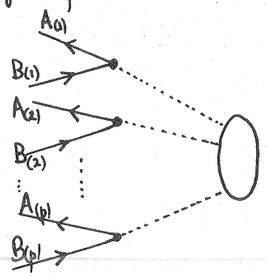
assuming that topological transformations interchanging equivalent disconnected parts are permitted.

5.4. Graphical representation of operators

We now formulate general rules for the diagrammatic representation of an arbitrary operator. We shall employ Goldstone diagrams in which each simple vertex is associated with one electron and has one entering and one leaving fermion line. Extension to other types of diagrams will be given later (Sec. 5.6). A p-body operator $\hat{Q}_p$, Eq. (3.60),

$$\hat{Q}_p = \frac{1}{p!} \sum_{\substack{A_{(1)} \cdots A_{(p)} \\ B_{(1)} \cdots B_{(p)}}} \langle A_{(1)} \cdots A_{(p)} | \hat{q}_p | B_{(1)} \cdots B_{(p)} \rangle X^{\dagger}_{A_{(1)}} \cdots X^{\dagger}_{A_{(p)}} X_{B_{(p)}} \cdots X_{B_{(1)}} , \tag{5.60}$$

will be generally represented by an s-vertex having p equivalent simple vertices, which we can symbolically represent as follows



(5.61)

The scalar factor associated with such an s-vertex is

$$d_Q \equiv d_Q (A_{(1)} \cdots A_{(p)} ; B_{(1)} \cdots B_{(p)})$$

$$= \langle A_{(1)} \cdots A_{(p)} | \hat{q}_p | B_{(1)} \cdots B_{(p)} \rangle \tag{5.62}$$

and the operator is

$$\hat{D}_Q \equiv \hat{D}_Q (A_{(1)} \cdots A_{(p)} ; B_{(1)} \cdots B_{(p)})$$

$$= d_Q\, X^{\dagger}_{A_{(1)}} \cdots X^{\dagger}_{A_{(p)}} X_{B_{(p)}} \cdots X_{B_{(1)}} . \tag{5.63}$$

When this operator is in the N-product form, we can also write

$$\hat{D}_Q = d_Q\, N\Big[\prod_{i=1}^{p} (X^{\dagger}_{A_{(i)}} X_{B_{(i)}}) \Big] , \tag{5.64}$$

since the operator pairs $X^{\dagger}_{A_{(i)}} X_{B_{(i)}}$, which are associated with the i-th open path in the diagram (61), commute within the N-product. This is in fact always the case for R-diagrams, once Wick's theorem has been applied (see Sec. 5.5).

If all spin-orbital labels in (61) are free labels, the weight w_Q is

$$w_Q = (p!)^{-1} , \tag{5.65}$$

so that

$$\hat{Q}_p = w_Q \sum_{\chi_Q}{}' \hat{D}(\chi_Q) , \tag{5.66}$$

where $\chi_Q \equiv \{A_{(1)} \cdots A_{(p)} ; B_{(1)} \cdots B_{(p)}\}$ is the set of labels carried by external lines of open paths of Q, and the prime designates that the summation extends over free labels only. These rules can be easily generalized to an arbitrary R diagram (see Sec. 5.5) when the numerical factor is properly chosen.

For one- and two-particle operators we can thus employ s-vertices of types (13 iii) or (41 i) and (13 iv) introduced above. Different operators can be distinguished by different types of operator or interaction lines used (dashed, wavy, dotted, solid, etc) and by different apexes*) attached to them if needed

*) P.T.O.

[cf., diagrams (13iii) and (41i) for $\hat{F}$ and $\hat{Z}$ operators, respectively, where dot and cross apexes* were chosen, or a general diagram (61) where a general apex*, which is incident with all operator lines, is shown as an open circle or ellipse].

Higher than two-body operators will usually arise as excitation operators. A general r-times excited state relative to $|\Phi_0\rangle$,

$$|\Phi^{(r)}\rangle = \sum_{\substack{A^1<A^2<\cdots<A^r \\ A_1<A_2<\cdots<A_r}} C^{A^1\cdots A^r}_{A_1\cdots A_r} \left|\begin{matrix} A^1\cdots A^r \\ A_1\cdots A_r \end{matrix}\right\rangle , \tag{5.67}$$

where r-excited configurations are given as

$$\left|\begin{matrix} A^1\cdots A^r \\ A_1\cdots A_r \end{matrix}\right\rangle = \left[\prod_{i=1}^{r} (X^\dagger_{A^i} X_{A_i})\right] |\Phi_0\rangle , \tag{5.68}$$

can be represented by an excitation operator $\hat{C}_r$,

$$|\Phi^{(r)}\rangle = \hat{C}_r |\Phi_0\rangle \tag{5.69}$$

defined as follows

$$\hat{C}_r = (r!)^{-1} \sum_{\substack{A^1\cdots A^r \\ A_1\cdots A_r}} \langle A^1 A^2\cdots A^r|\hat{C}_r|A_1 A_2\cdots A_r\rangle \prod_{i=1}^{r} (X^\dagger_{A^i} X_{A_i}) \tag{5.70a}$$

$$= (r!)^{-2} \sum_{\substack{A^1\cdots A^r \\ A_1\cdots A_r}} \langle A^1 A^2\cdots A^r|\hat{C}_r|A_1 A_2\cdots A_r\rangle_{|A} \prod_{i=1}^{r} (X^\dagger_{A^i} X_{A_i}), \tag{5.70b}$$

where

$$\langle A^1 A^2\cdots A^r|\hat{C}_r|A_1 A_2\cdots A_r\rangle_{|A} = \sum_{P\in\mathcal{S}_r} (-1)^p \langle A^1 A^2\cdots A^r|\hat{C}_r|A_{p_1} A_{p_2}\cdots A_{p_r}\rangle , \tag{5.71}$$

the sum extending over all permutations $P: i\to p_i$ of $\mathcal{S}_r$. The coefficients in (67), which must be antisymmetric in both its superscripts and subscripts, can thus be identified with antisymmetrized matrix elements of the excitation operator $\hat{C}_r$,

$$C^{A^1 A^2\cdots A^r}_{A_1 A_2\cdots A_r} = \langle A^1 A^2\cdots A^r|\hat{C}_r|A_1 A_2\cdots A_r\rangle_{|A} , \tag{5.72}$$

so that

$$C^{A^1 A^2\cdots A^r}_{A_1 A_2\cdots A_r} = (-1)^p C^{A^1 A^2\cdots A^r}_{A_p A_{p_2}\cdots A_{p_r}} = (-1)^p C^{A^{p_1} A^{p_2}\cdots A^{p_r}}_{A_1 A_2\cdots A_r} \tag{5.73}$$

for any permutation $P: i\to p_i$ of $\mathcal{S}_r$.

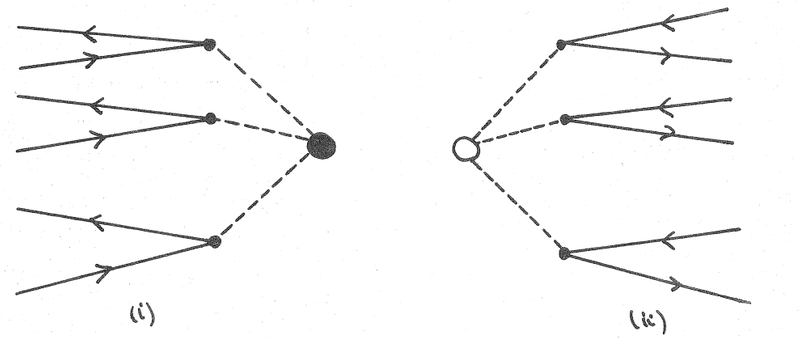
The operator $\hat{C}_r$, Eq.(70), is in fact automatically in the N-product form since

$$\prod_{i=1}^{r} (X^\dagger_{A^i} X_{A_i}) = \prod_{i=1}^{r} (Y^\dagger_{A^i} Y^\dagger_{A_i}) = N\Big[\prod_{i=1}^{r} Y^\dagger_{A^i} Y^\dagger_{A_i}\Big]$$

$$= N\Big[\prod_{i=1}^{r} (X^\dagger_{A^i} X_{A_i})\Big] = N\left[X^\dagger_{A^1}\cdots X^\dagger_{A^r} X_{A_r}\cdots X_{A_1}\right]$$

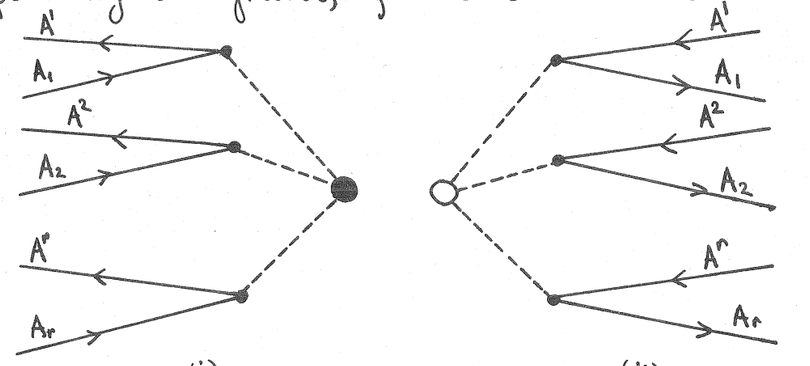
$$= X^\dagger_{A^1}\cdots X^\dagger_{A^r} X_{A_r}\cdots X_{A_1} . \tag{5.74}$$

The skeletons which we can associate with the operator $\hat{C}_r$, Eq.(70a), or with its hermitean conjugate $\hat{C}^\dagger_r$, are



(5.75)

so that $w_{C_r} = (r!)^{-1}$ in each case. With corresponding diagrams, referred to as C and $\bar{C}$ diagrams,



(5.76)

*) Vertices which are incident with only non-fermion lines are called apexes.

we associate the following scalar quantities

$$d_c(\chi_r) = \langle A^1 A^2 \cdots A^r | \hat{C}_r | A_1 A_2 \cdots A_r \rangle \qquad (5.77\text{i})$$

$$d_{\bar{c}}(\chi_r) = d_c^*(\chi_r), \qquad (5.77\text{ii})$$

where

$$\chi_r \equiv \{ A^1 A^2 \cdots A^r ; A_1 A_2 \cdots A_r \} \qquad (5.78)$$

is the set of labels carried by fermion lines, and the operators:

$$\hat{D}_c(\chi_r) = d_c(\chi_r) \prod_{i=1}^{r} (X_{A^i}^\dagger X_{A_i}) \qquad (5.79\text{i})$$

$$\hat{D}_{\bar{c}}(\chi_r) = \hat{D}_c^\dagger(\chi_r), \qquad (5.79\text{ii})$$

so that in agreement with a general formula (66) we can write

$$\hat{C}_r = w_c \sum_{\chi_r} \hat{D}_c(\chi_r) \qquad (5.80\text{i})$$

and

$$\hat{C}_r^\dagger = w_c \sum_{\chi_r} \hat{D}_{\bar{c}}(\chi_r). \qquad (5.80\text{ii})$$

Later, we shall use a similar representation for cluster operators $\hat{T}_r$ (cf. Sec. 10.2) except that we shall choose full operator lines. We shall also introduce other representations for special type of operators as, for example, the projection operators of perturbation theory (See, Sec. 9.4).

5.5. General rules for resulting diagrams

Having introduced a diagrammatic representation for operators, we can graphically evaluate quantities given as products of operators

$$\hat{Q} = \hat{Q}^{(r)} \hat{Q}^{(r-1)} \cdots \hat{Q}^{(1)}, \qquad (5.81)$$

in particular a Fermi vacuum mean value of such products

$$\langle \hat{Q} \rangle \equiv \langle \Phi_0 | \hat{Q}^{(r)} \cdots \hat{Q}^{(1)} | \Phi_0 \rangle, \qquad (5.82)$$

and the states which $\hat{Q}$ produces when acting on a Fermi vacuum,

$$|\Xi\rangle = \hat{Q} |\Phi_0\rangle. \qquad (5.83)$$

Since one- and two-body operators may be always easily expressed in the N-product form (see Sec. 4.5) and the excitation operators, Eq. (74), are automatically in this form, we shall assume that all operators $\hat{Q}^{(i)}$ in the product (81) are in the N-product form. Generalization to operators which are not in the N-product form is straightforward and will be given later (cf. Sec. 5.6).

We have seen in Secs. 1 and 2 above that, up to a numerical factor, we can obtain desired expressions directly from pertinent R-diagrams which can be formed from component graphs representing operators $Q^{(i)}$ in the product (1). These rules will be summarized below and we now turn our attention to the determination of correct numerical factors.

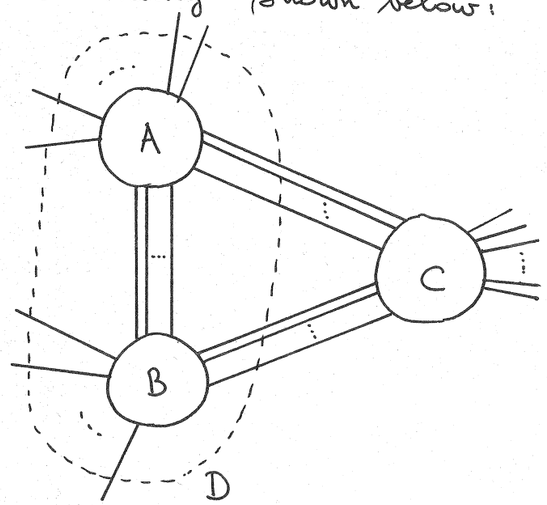
In order to find the absolute value of the correct numerical factor for a given R-diagram, we have to know how many different contraction schemes yield this diagram or, in other words, in how many distinct ways we can form a given R-diagram from the component skeletons representing individual $Q^{(i)}$ operators in the product (1). In this task the following general combinatorial result is very helpful:

T3: Let A, B and R be connected skeletons with weights w_A, w_B and w_R, respectively. Let R be the resulting skeleton obtained by joining pairwise a certain

number of oriented lines of A and B. Then the number N of distinct ways in which the skeleton R may be obtained from skeletons A and B is

$$N = w_R (w_A w_B)^{-1} . \tag{5.84}$$

This theorem may be easily generalized to an arbitrary number of components. Suppose that R is obtained by connecting a certain number of lines of skeletons A, B and C, as schematically shown below:



(5.85)

Proceeding in stages, we first form D skeletons from A and B by interconnecting appropriate lines. This can be done in $N_{AB} \equiv N_D$ distinct ways

$$N_D \equiv N_{AB} = w_D (w_A w_B)^{-1} . \tag{5.86}$$

The desired R skeleton may then be obtained from D and C skeletons in

$$N_{(AB)C} \equiv N_{DC} = w_R (w_D w_C)^{-1} \tag{5.87}$$

different ways. Thus, all together, R can be obtained from A, B and C in

$$N = N_{(AB)C} \cdot N_{AB} = \frac{w_R}{w_D w_C} \cdot \frac{w_D}{w_A w_B} = w_R (w_A w_B w_C)^{-1} \tag{5.88}$$

distinct ways. This result is clearly independent of the manner in which we partition the R diagram: the same result is obtained when we form an intermediate skeleton D from either B, C or A, C skeletons.

Let us illustrate this theorem on Ex. 5.2.1, where $w_C = w_{\bar{C}} = 1$, $w_V = 1/2$ and $w_R = 1$ for each diagram (26). Thus, according to T3 each R-diagram (26) can be obtained in

$$N = w_R (w_{\bar{C}} w_C w_V)^{-1} = 1 \cdot (1 \cdot 1 \cdot 1/2)^{-1} = 2 \tag{5.89}$$

different ways. Indeed, each R diagram represents in this case two distinct contraction schemes as follows from Eqs. (26). A proof of T3, at least for the type of diagrams considered, is sketched in Appendix A.

It is now easily seen that the desired numerical factor (up to the sign) is given by the weight of each R-diagram. Indeed, consider a set of diagrams (skeletons) representing operators $\hat{Q}^{(i)}$ in (81) and having weights $w_{Q^{(i)}}$ from which we form some R-diagram (skeleton) with weight w_R. The desired numerical factor is then given by the product of weight factors $w_{Q^{(i)}}$, appearing in expressions for corresponding operators, Eq. (66), times the multiplicity N, Eq. (84),

$$N = w_R \left(\prod_{i=1}^{*} w_{Q^{(i)}} \right)^{-1} , \tag{5.90}$$

which yields simply w_R.

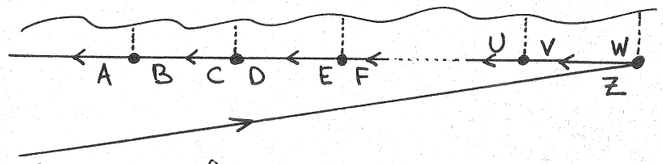
It thus remains to find a correct sign for each R-diagram, which is given by the parity of a permutation "unscrambling" a given system of pairings represented by this diagram, into a simple product of individual contractions.

We can first "unscramble" any system of pairings into separate factors corresponding to individual closed loops and open

paths. This can be done without any phase change since (74) holds for any operator in the N-product form,

$$N[X^\dagger_{A(1)} \cdots X^\dagger_{A(k)} X_{B(k)} \cdots X_{B(1)}] = N\Big[\prod_{i=1}^{k} (X^\dagger_{A(i)} X_{B(i)})\Big], \qquad (5.91)$$

so that the operator pairs $(X^\dagger_{A(i)} X_{B(i)})$, represented by fermion lines which are incident with the same vertex, can be arbitrarily permuted inside the N-product (as long as the order in each pair of contracted operators is preserved). Consider, then, the factor associated with a single open path containing only internal particle lines, schematically



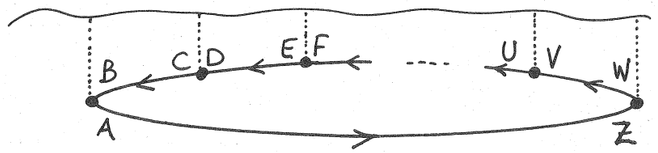
(5.92a)

which has the form

$$N[\cdots(\overbrace{X^\dagger_A X_B}\ \overbrace{X^\dagger_C X_D}\ \overbrace{X^\dagger_E X_F}\ \cdots\ \overbrace{X^\dagger_U X_V}\ \overbrace{X^\dagger_W X_Z})\cdots] = \overbrace{X_B X^\dagger_C}\ \overbrace{X_D X^\dagger_E}\ \overbrace{X_F}\ \cdots\ \overbrace{X^\dagger_U X_V}\ \overbrace{X^\dagger_W}\ N[\cdots(X^\dagger_A X_Z)\cdots], \qquad (5.92b)$$

so that no sign change is necessary when we assign the operator $N[(X^\dagger_A X_Z)\cdots]$ to this path. This result is independent of the number of particle lines in the chain.

Similarly, for a loop with the single hole line, schematically,



(5.93a)

we have the factor

$$\cdots(\overbrace{X^\dagger_A X_B}\ \overbrace{X^\dagger_C X_D}\ \overbrace{X^\dagger_E X_F}\ \cdots\ \overbrace{X^\dagger_U X_V}\ \overbrace{X^\dagger_W X_Z})\cdots =$$

Appendix A: Proof of T3.

We shall sketch the proof of T3 for the case of Goldstone-type diagrams considered.

Let diagram A (B) has q (r) equivalence classes each containing n_i (m_i) equivalent vertices (or pairs of fermion lines incident with them), so that

$$g(A) = \prod_{i=1}^{q} (n_i!) \quad , \quad g(B) = \prod_{i=1}^{r} (m_i!) \qquad (A.1)$$

according to (5.56). For a given R-diagram the number of distinct contraction schemes will be determined by, say, the lines outgoing from A and ingoing in B; the other type of lines being fixed by the choice of a given R-diagram. Let the number of these lines interconnecting A and B is t,

$$t = \sum_{i,j} t_{ij}, \qquad (A.2)$$

where t_{ij} designates the number of these lines between equivalence classes i of A and j of B, ($i = 1, \cdots, q$; $j = 1, \cdots, r$).

The number of possible interconnections for lines from the i-th equivalence class of A is

$$N_i = \binom{n_i}{t_{i1}} \binom{n_i - t_{i1}}{t_{i2}} \cdots \binom{n_i - (t_{i1} + t_{i2} + \cdots + t_{i,r-1})}{t_{ir}}$$

and, similarly, for the j-th equivalence class of B

$$M_j = \binom{m_j}{t_{1j}} \binom{m_j - t_{1j}}{t_{2j}} \cdots \binom{m_j - (t_{1j} + \cdots + t_{q-1,j})}{t_{qj}}.$$

Note that

$$\binom{n}{t_1}\binom{n-t_1}{t_2} = \binom{n}{t_2}\binom{n-t_2}{t_1}$$

or, in general

$$\binom{n}{t_1}\binom{n-t_1}{t_2}\cdots\binom{n-(t_1+t_2+\cdots+t_{k-1})}{t_k}$$

$$=\binom{n}{t_{p_1}}\binom{n-t_{p_1}}{t_{p_2}}\cdots\binom{n-(t_{p_1}+t_{p_2}+\cdots+t_{p_{k-1}})}{t_{p_k}}$$

for any permutation $P: i \to p_i$ from $\mathfrak{S}_k$. Designating the partial sums of t_{ij}'s as follows

$$t_i^A = \sum_{j=1}^{r} t_{ij} \quad , \quad t_j^B = \sum_{i=1}^{q} t_{ij}$$

we can write N_i and M_j as follows

$$N_i = \frac{n_i!}{(n_i - t_i^A)!\, t_{i1}!\, t_{i2}! \cdots t_{ir}!} ,$$

$$M_j = \frac{m_j!}{(m_j - t_j^B)!\, t_{1j}!\, t_{2j}! \cdots t_{qj}!}$$

The number of possible contraction schemes is then

$$N = \left(\prod_{i=1}^{q} N_i\right)\left(\prod_{j=1}^{r} M_j\right)\prod_{ij}(t_{ij}!) ,$$

where the last factor originates from the fact that the t_{ij} contraction lines between each pair of equivalence classes of A and B can be interconnected in $t_{ij}!$ different ways. When we realize that

$$g(R) = \prod_{i=1}^{q}(n_i - t_i^A)! \cdot \prod_{j=1}^{r}(m_j - t_j^B)! \cdot \prod_{ij}(t_{ij})! ,$$

since we have $(n_i - t_i^A)$ or $(m_j - t_j^B)$ equivalent vertices left in respective equivalence classes and t_{ij} equivalent internal lines, we find that

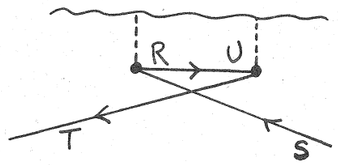
$$N = \frac{g(A)\, g(B)}{g(R)} = \frac{w_R}{w_A \cdot w_B} ,$$

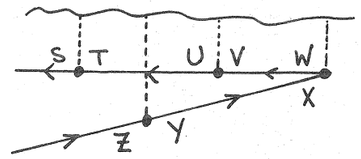
proving the theorem.

$$= \overrightarrow{X_A^\dagger X_Z}\ \overrightarrow{X_B X_C^\dagger}\ \overrightarrow{X_D X_E^\dagger}\ \overrightarrow{X_F} \cdots \overrightarrow{X_U^\dagger}\ \overrightarrow{X_V X_W^\dagger} , \qquad (5.93b)$$

so that again no sign change is required.

Introducing, however, any additional hole line, either into the open chain (92) or into the loop (9), brings the factor of (-1) as illustrated below:





(5.94)

$$N\Big[\cdots \overrightarrow{X_R^\dagger X_S}\ \overrightarrow{X_T^\dagger X_U} \cdots\Big]$$

$$= -\overrightarrow{X_R^\dagger X_U}\, N\Big[\cdots \overrightarrow{X_T^\dagger X_S} \cdots\Big]$$

$$N\Big[\cdots \overrightarrow{X_S^\dagger X_T}\ \overrightarrow{X_Y^\dagger X_Z}\ \overrightarrow{X_U^\dagger X_V}\ \overrightarrow{X_W^\dagger X_X} \cdots\Big]$$

$$= -\overrightarrow{X_T X_U^\dagger}\ \overrightarrow{X_V X_W^\dagger}\ \overrightarrow{X_Y^\dagger X_X}\, N\Big[\cdots \overrightarrow{X_S^\dagger X_Z} \cdots\Big]$$

We can thus associate the factor $(-1)^{h_i}$ with each open path, and $(-1)^{h_i+1}$ with each closed loop, with h_i designating the number of internal hole lines in i-th loop or open path. For an R-diagram containing l loops and k open paths we have the sign factor

$$\prod_{i=1}^{l}(-1)^{h_i+1} \cdot \prod_{j=1}^{k}(-1)^{h_j} = (-1)^{l+h} , \qquad (5.95)$$

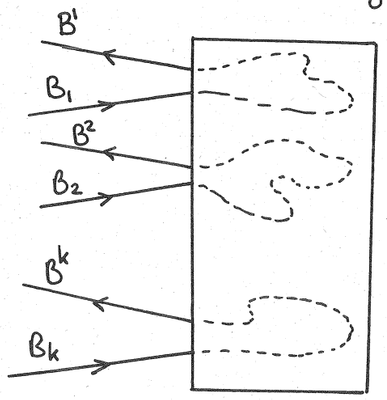
where

$$h = \sum_{i=1}^{l} h_i + \sum_{j=1}^{k} h_j \qquad (5.96)$$

designates the total number of internal hole lines.

We can now summarize the rules which assign appropriate algebraic expressions to non-equivalent R-diagrams. Consider an R-diagram with l closed loops and k open paths. When evaluating the Fermi vacuum mean value (82) we have clearly $k=0$, while for the expression of the type given by Eq. (83) all external lines extend to the left of incident vertices. Let the in-

and out-going external lines of the ith open path carry labels B^i and B_i, respectively, as shown schematically below



(5.97)

and let ξ designates the set of spin-orbital labels carried by these external lines,

$$\xi \equiv \{ B' B^2 \cdots B^k ; B_1 B_2 \cdots B_k \}, \qquad (5.98)$$

and $\tilde{\xi}$ the set of labels on internal lines. Designate, further, the product of scalar factors associated with individual s-vertices by $d_R(\xi,\tilde{\xi})$ and define the operator

$$\hat{D}_R(\xi,\tilde{\xi}) = d_R(\xi,\tilde{\xi})\, N\Big[\prod_{i=1}^{k} (X^\dagger_{B^i} X_{B_i})\Big]. \qquad (5.99)$$

Clearly, when ξ is empty (i.e., $k=0$), we set $\hat{D}_R = d_R$ in view of convention (4.17). The algebraic expression associated with this R-diagram is thus

$$w_R\,(-1)^{h+l} \sum_{\xi,\tilde{\xi}}{}' \hat{D}_R(\xi,\tilde{\xi}), \qquad (5.100)$$

where the prime indicates that the summation extends over free labels. In case (82) this is a scalar while for (83) expression (100) gives the operator acting on $|\Phi_0\rangle$. The sum of expressions (100) for all non-equivalent R-diagrams gives the desired resulting quantity (82) or (83).

5.6. Standard and normal product operator forms.

So far we have assumed that the operators defining a given problem, Eq. (81), are in the N-product form. This in turn enabled us to exploit generalized Wick's theorem and thus to avoid all R-diagrams with internal fermion lines connecting vertices of the same s-vertex (such lines would represent contractions between operators appearing in the same N-product).

Since the excitation-type operators, Eq. (74), are automatically in the N-product form, and the one- and two-body operators representing various observables can be easily transformed to such a form as shown in Sec. 4.5, we can always use the N-product form of operators and thus eliminate a number of repetitively occuring R-diagrams. However, we can also easily generalize our diagrammatic approach so that some or all the operators appearing in the evaluated expression, Eqs. (81)-(83), are in the standard n-product form, rather than in the N-product form. We shall see that this generalization requires only a minor addition to the existing rules. We shall also see that the additional R-diagrams which must be considered in such a case are easily recognized and represent the terms which occur in bringing the operators to the N-product form, as expected. This is particularly apparent when we interpret diagrammatically the transformation of standard n-product form operators, Eq. (3.60), to the N-product form (cf., Sec. 4.5).

A standard form of a general r-body operator, as given by Eq. (3.60), contains products of creation and annihilation operators which are automatically in the n-product form, namely the products

$$X^\dagger_{A(1)} \cdots X^\dagger_{A(m)} X_{B(m)} \cdots X_{B(1)} . \qquad (5.101)$$

Since contractions (with respect to the Fermi vacuum) of like operators always vanish, cf. Eqs. (4.33) or (4.31), we see that the only additional contractions which have to be considered when operators are not in the N-product form and an ordinary Wick's theorem is applied are of the hole type, namely

$$\overbrace{X_A^{\dagger}\, X_B} = h(A)\langle A|B\rangle . \qquad (5.102)$$

Consequently, all internal fermion lines which are incident with vertices of the same s-vertex (i.e., oriented lines originating and ending on the same vertex or connecting two vertices which are also connected by a path of non-oriented operator lines) are always hole lines irrespective of their direction relative to the "time axis". In fact, the right-left direction of these lines is always ill-defined since individual vertices of a given s-vertex are usually displaced in vertical direction (i.e., have the same "time").

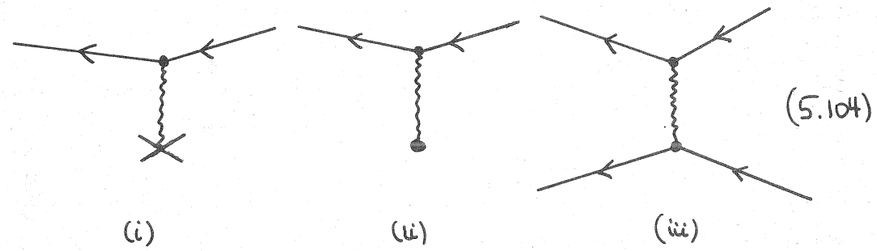
Thus, when considering the operators in standard rather than N-product form, we have to:
(i) allow R-diagrams with internal lines which are incident with vertices of the same s-vertex, and
(ii) label all such internal lines with hole labels.

Otherwise, all the rules for the evaluation of R-diagrams, which we derived earlier, remain in force, as may be easily ascertained. In particular the sign rule is not changed, since our derivation did not use the fact that the operators $Q^{(i)}$ in Eq(81) were assumed to be in the N-product form. Indeed, we have only used the fact that all the resulting terms were in the N-product form, as given by Wick's theorem. The only extension which we need is to realize that the theorem T3, Eq.(84), also applies when the R-diagram is formed from a single initial

diagram A by interconnecting some of its oriented lines, so that

$$N = W_R (W_A)^{-1} . \qquad (5.103)$$

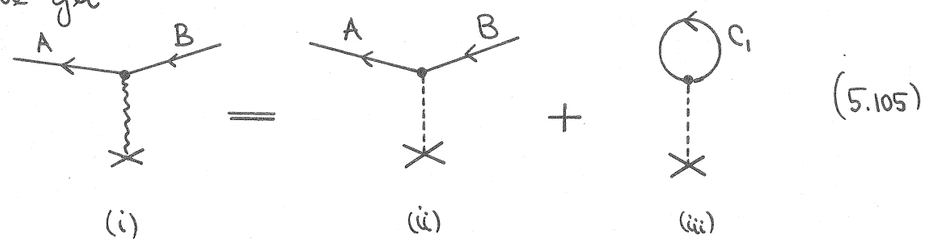
We now illustrate on a few simple examples the type of additional diagrams which result when we use operators in a standard N-product form, Eq. (3.60). In order to distinguish diagrammatically the operators in a standard and an N-product forms, we use wavy operator lines for the former. Thus, for example, the one- and two-particle operators $\hat{Z}$, $\hat{F}$ and $\hat{V}$ in their standard form, Eqs. (3.64), (4.75) and (3.70), respectively, are represented by skeletons



(i) (ii) (iii) (5.104)

and corresponding diagrams obtained be labeling fermion lines with free spin-orbital indices.

Since all the terms which we obtain when applying Wick's theorem are in the N-product form, we can use again diagrams with dashed interaction lines introduced earlier to represent operators in the N-product form. We can thus interpret graphically the N-product form for one- and two-body operators, which we derived algebraically in Sec. 4.5. Applying Wick's theorem to the operator $\hat{Z}$, Eq. (3.64), we get



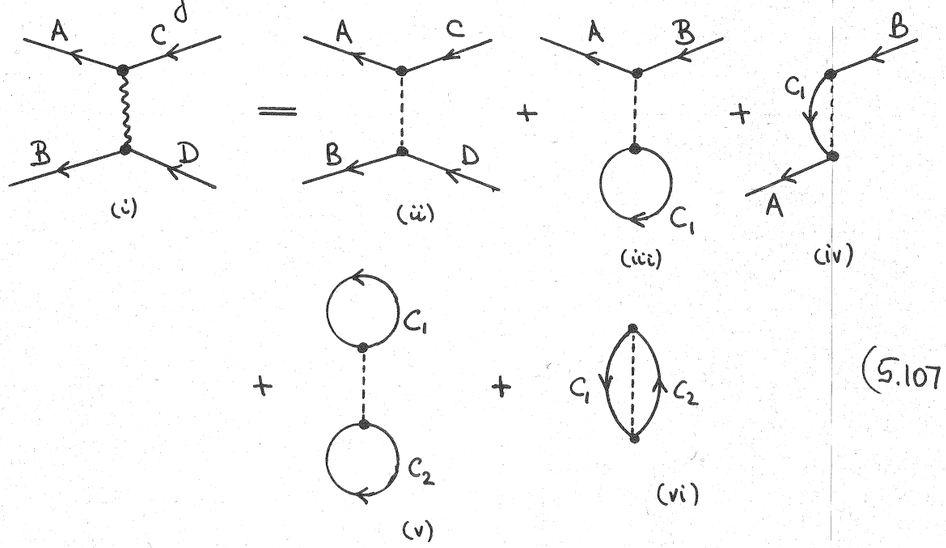
(i) (ii) (iii) (5.105)

which yields immediately (4.63) : the first term on the right hand side, (105 ii), represents $\hat{Z}_N$, while the last term, (105 iii), represents the expression

$$\sum_{C_1} \langle C_1 | \hat{z} | C_1 \rangle , \tag{5.106}$$

since the internal hole line labeled C_1 both leaves and enters the z-vertex.

In a similar way, applying Wick's theorem to a two-particle operator, say the operator $\hat{V}$, Eq. (3.70), we form all permissible R-diagrams from (104 iii) obtaining



(5.107)

Again, all the internal lines in (107) must be hole lines and all labels are free. The diagram (ii) represents the N-product form of V, Eq. (4.57), diagrams (iii) and (iv) define some new one-particle operator which we called $\hat{G}$, Eq. (4.67) [this operator is, of course, in the N-product form on the right hand side of (107), Eq. (4.70)];

$$\hat{G}_N = \sum_{A,B} \langle A | \hat{g} | B \rangle \, N[X_A^\dagger X_B] , \tag{5.108}$$

$$\langle A | \hat{g} | B \rangle = \sum_{C_1} \left(\langle A C_1 | \hat{v} | B C_1 \rangle - \langle A C_1 | \hat{v} | C_1 B \rangle \right) , \tag{5.109}$$

and the last two diagrams (v) and (vi), corresponding to the fully contracted terms in (4.65), must represent the Fermi vacuum mean value of $\hat{V}$,

$$\langle V \rangle = \sum_{C_1, C_2} \left(\langle C_1 C_2 | \hat{v} | C_1 C_2 \rangle - \langle C_1 C_2 | \hat{v} | C_2 C_1 \rangle \right) . \tag{5.110}$$

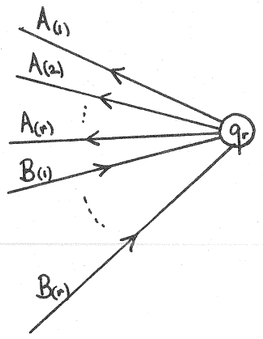
We recall that the sign factor $(-1)^{l+h}$ for diagrams (iv) and (vi) equals (-1) since $l+h$ is odd in both cases. Further, the weights of corresponding skeletons are ½ for diagrams (i), (ii), (v) and (vi) and 1 for diagrams (iii) and (iv). This case illustrates the special case of T3, Eq. (103), since $w_V = \frac{1}{2}$ while for both skeletons (iii) and (iv) we have $w_R = 1$ so that $N = 2$ in agreement with Eq. (4.65) [i.e., each diagram (iii) and (iv) represents two systems of pairings in Eq. (4.65)].

We thus conclude that it is immaterial whether we use the N-product or standard n-product form of operators when we modify the rules for the construction of R-diagrams accordingly: While no R-diagrams with internal lines which are incident with the vertices of the same s-vertex are allowed when the N-product form is used, these diagrams must be considered when a given operator is in its standard n-product form. Moreover, all such internal lines must be regarded as hole lines. Otherwise, the same rules for the evaluation of R-diagrams apply as given earlier (Sec. 5.5).

5.7. Hugenholtz and Brandow diagrams

So far we have almost exclusively used Goldstone type diagrams in which there is only one ingoing and one outgoing fermion line attached to each simple vertex. These diagrams are appropriate for a diagrammatic representation of operators defined via non-antisymmetrized matrix elements, in which each spinorbital label, in both bra and the ket part, is associated with coordinates of only one given electron. These diagrams represent a special case of general Feynman diagrams used in quantum field theories and were initially introduced by Goldstone in his famous study of the many-body perturbation theory (MBPT).

A different, though closely related, type of diagrams was introduced by Hugenholtz, who employed a convolution formalism in his formulation of the MBPT. An example of a Hugenholtz-type diagram was given in Sec. 5.3, Eqs. (50)-(54). These diagrams, also called degenerate diagrams, are generally associated with antisymmetrized form of operator matrix elements, and a general r-body operator is represented by a single vertex with r outgoing and r ingoing lines. Thus, a general operator given by the last equation (3.60) is represented by the diagram of the following type



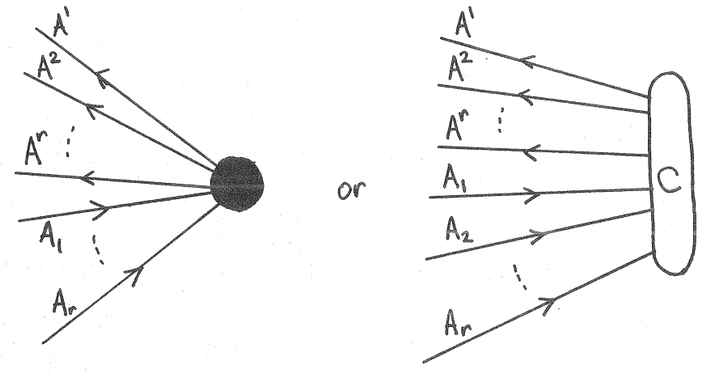
(5.111)

The weight of the corresponding skeleton is clearly

$w_Q^H = (r!)^{-2}$, since an arbitrary permutation of the outgoing and, separately, the ingoing fermion lines is possible. Generally, the weight of a Hugenholtz-type of diagram with k classes of equivalent fermion lines, each containing r_i lines, has the weight

$$w = \prod_{i=1}^{k} (r_i!)^{-1}. \qquad (5.112)$$

For example, a two-particle operator $\hat{V}$, Eq. (50), is represented by the diagram (51). Similarly, the excitation operator $\hat{C}_r$, Eq. (70b), will be represented by the diagram



(5.113)

The scalar factor associated with this diagram is given by the corresponding antisymmetrized matrix element

$$d_C^H = \langle A' A^2 \cdots A^r | \hat{C}_r | A_1 A_2 \cdots A_r \rangle_A , \qquad (5.114)$$

and the weight is

$$w_C^H = (r!)^{-2} . \qquad (5.115)$$

The operator associated with diagram (113) is defined analogously as for the Goldstone diagrams,

$$D_C^H = d_C^H \; X_{A'}^+ \cdots X_{A^r}^+ X_{A_r} \cdots X_{A_1} = d_C^H \prod_{i=1}^{r} X_{A^i}^+ X_{A_i} . \qquad (5.116)$$

However, just as in the two-body case, Eqs. (51)-(53), the sign of the scalar factor and of the operator

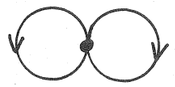
is not uniquely defined by the Hugenholtz diagram.

Hugenholtz diagrams represent a homomorphic mapping of Goldstone diagrams, which is simply obtained by "collapsing" each s-vertex into a single Hugenholtz vertex. Thus, for example, both diagrams (26) become



(5.117)

or, diagrams (107 v) and (107 vi) become



(5.118)

so that in each case both coulomb and exchange type Goldstone diagrams are represented by the same Hugenholtz diagram. The Hugenholtz diagrams are consequently less numerous than Goldstone type diagrams. This is the main merit of Hugenholtz diagrams since this reduction may be very significant when either a large number of operators is present in the evaluated product (as, for example, in higher orders of perturbation theory, see Sec. 9) or if individual operators themselves are of a high order.

Unfortunately, Hugenholtz diagrams do not determine uniquely the sign of algebraic quantities they represent as we have seen above. It is thus convenient to use them concurrently with Goldstone diagrams, for which a unique one-to-one mapping with algebraic expressions holds, so that the advantages of both representations can be exploited. This may be either achieved by "expanding" each Hugenholtz vertex into a corresponding Goldstone vertex (and considering all possible permutations of fermion lines on Goldstone vertices)

thus obtaining all non-equivalent Goldstone diagrams (in this way Hugenholtz diagrams serve as an intermediate step which helps to find all possible Goldstone R-diagrams) or, more effectively, by representing each Hugenholtz diagram by one Goldstone version. This is possible since the sign of the component given by the chosen Goldstone version will determine correctly the sign of the whole antisymmetrized matrix element. It is therefore irrelevant which Goldstone diagram is used to represent each Hugenholtz diagram.

For example, we can use either diagram (107 v) or diagram (107 vi) to represent Hugenholtz diagram (118), both representations yielding obviously the same result

$$\langle V \rangle = \tfrac{1}{2} \sum_{c_1 c_2} \langle c_1 c_2 | \hat{v} | c_1 c_2 \rangle_{\mathcal{A}} = -\tfrac{1}{2} \sum_{c_1 c_2} \langle c_1 c_2 | \hat{v} | c_2 c_1 \rangle_{\mathcal{A}}. \qquad (5.119)$$

This possibility of exploiting one Goldstone version to determine the correct sign of the Hugenholtz diagram contribution was first explicitly formulated for perturbation theory diagrams by Brandow. We shall thus refer to a Goldstone diagram representing a Hugenholtz diagram as a Brandow diagram.

We can now summarize essential steps in this procedure as follows:

(i) represent the operators which are defined through antisymmetrized matrix elements by Hugenholtz diagrams

(ii) form all permissible resulting diagrams

(iii) represent each Hugenholtz R-diagram by one of its Goldstone versions (i.e., form Brandow diagrams)

(iv) when evaluating Brandow R-diagrams:

(a) assign Hugenholtz' scalar factor to each s-vertex of the Brandow diagram

(b) determine the weight from a corresponding Hugenholtz diagram, and

(c) determine the sign from a Brandow diagram using the same rule as for Goldstone diagrams.

In this way we have both the advantage of having a small number of R-diagrams to handle while preserving a one-to-one correspondence between diagrams and algebraic expressions.

Let us mention, finally, that Goldstone type diagrams are also indispensable for a simple transition to spin-free formalism when spin-independent operators are considered, as we shall briefly indicate in the next section. However, one can always define appropriate spin-adapted quantities through antisymmetrized matrix elements and combine accordingly Brandow diagrams with spin diagrams (i.e., special case of Jucys diagrams for ½ angular momentum), which are then evaluated independently using graphical methods of spin algebras. We will return to this general approach, which can be also used for central systems when angular momentum adapted theory is desired, in Sec. . This approach may also be of value even for spin-dependent operators. Sometimes it may also be preferable not to symmetry adapt all the operators in the evaluated expression (say, the two-electron integrals) and exploit a mixed Goldstone-Brandow representation, in which certain s-vertices are genuine Goldstone vertices associated with non-antisymmetrized matrix elements, while the remaining vertices are Goldstone representatives of Hugenholtz vertices (i.e., Brandow vertices). These remarks are intended to indicate the versatility of the graphical approach which we have outlined above. They will be examined more closely and exploited in actual applications in the following Chapters.

5.8. Spin-independent operators

A general technique of spin-adaptation, exploiting Jucys spin diagrams in conjunction with (spin) orbital Goldstone or Brandow diagrams will be given later (Ch. 6.). Here we only wish to mention the possibility of a simple transition to spin-free orbital formalism in a special case of spin-independent operators and closed-shell singlet reference states. This simple method is particularly useful in the case of non-degenerate perturbation theory.

Consider a closed shell system and a singlet reference state $|\Phi_0\rangle$ which consists of doubly occupied orbitals $|a_i\rangle$, so that

$$|A_i\rangle = |a_i\rangle\,|\eta_i\rangle\,, \tag{5.120a}$$

where $|\eta_i\rangle = |1/2\rangle$ or $|-1/2\rangle$ are spin-up or down eigenfunctions of $\hat{S}^2$ and $\hat{S}_z$. The unoccupied (particle) states are similarly labeled with superscripted letters,

$$|A^i\rangle = |a^i\rangle\,|\eta^i\rangle\,. \tag{5.120b}$$

Every summation over the spin-orbitals may then be written as two independent summations over the orbitals and spin states, respectively. For spin-independent operators, as for example the one- and two-electron parts of a molecular electronic Hamiltonian, Eqs. (3.81) and (3.84), the matrix elements have the property given by Eqs. (3.12) and (3.14), respectively*). Consequently, both spinorbital labels associated with fermion lines which are incident with a given simple vertex of a corresponding Goldstone diagram must have identical spin parts (both up or both down) in order that the matrix element (i.e., scalar factor) associated with such a diagram does not vanish. Thus, the spin state must be preserved along each

*) Recall that subscripts in Eq. (3.9), (3.12), (3.14) etc. do not have the meaning of hole labels but simply distinguish general spinorbital labels.

oriented path in such diagrams. Consequently, for each closed loop of oriented lines there are two possibilities for spin states, each one yielding the same result in view of the spin independence of the operators considered. Thus, a transition from spin-orbital to orbital labels we simply have to add a factor of 2 for each closed loop of oriented lines (assuming, of course, that all the labels on this loop are free ones).

We can thus summarize the above considerations for the special case of closed shell singlet reference state: to obtain spin-independent expressions (i) replace the spin-orbital labels with orbital labels and, (ii) associate with each R-diagram an additional factor f,

$$f = (2^{\ell}) . \qquad (5.121)$$

Clearly, it is essential that we use Goldstone type diagrams in this case.

For example, evaluating the vacuum mean value of a two-particle operator $\hat{V}$, Eq. (3.84), we have to consider the Goldstone diagrams (107 v) and (107 vi), which yield

$$\langle \hat{V} \rangle = \tfrac{1}{2} \sum_{c_1 c_2} \left(4 \langle c_1 c_2 | \hat{v} | c_1 c_2 \rangle - 2 \langle c_1 c_2 | \hat{v} | c_2 c_1 \rangle \right), \qquad (5.122)$$

where the factors 4 and 2 result from the above rule (ii), Eq. (121). Thus

$$\langle \hat{V} \rangle = \sum_{c_1 c_2} \left(2 \langle c_1 c_2 | \hat{v} | c_1 c_2 \rangle - \langle c_1 c_2 | \hat{v} | c_2 c_1 \rangle \right), \qquad (5.123)$$

in agreement with the well known result.

5.9. Summary of rules

We now summarize basic rules for an implementation of the diagrammatic approach assuming that either Goldstone or Brandow type diagrams are used (in the latter case the antisymmetrized scalar factors and Hugenholtz weights must be used):

(i) Assign appropriate basic diagrams to individual operators in a calculated product (81) (i.e., define appropriate scalar factors, free and fixed labels, weights and operators for each basic diagram).

(ii) Form all permissible distinct R-skeletons from corresponding basic skeletons, determine their weights and form corresponding diagrams by labelling unlabeled lines with appropriate (i.e., hole or particle) free labels. The character of permissible R-diagrams depends on the type of an expression we evaluate (no external lines for vacuum mean values, etc.) and on the fact whether or not the operators are in the N-product form (see Sec. 5.6).

(iii) Assign to each distinct R-diagram the scalar factor d_R,

$$d_R(\rho_1, \rho_2) = (-1)^{\ell + h} \, \Delta(\rho_1) \prod_{G \in R} d_G(\rho_1, \rho_2), \qquad (5.124)$$

where ρ_1 (ρ_2) designate the set of spin-orbital labels carried by the internal (external) lines, ℓ is the number of closed loops, h is the number of <u>internal</u> hole lines and $\Delta(\rho_1)$ represents the product of Kronecker delta symbols $\langle A|B \rangle$ assigned to internal lines carrying two fixed labels A and B. The product includes all scalar factors associated with basic diagrams (or s-vertices) G of R (see (i) above) which result when all internal lines of R are dissected.

Assign, further, the operator quantity $\hat{D}_R$ to each R-diagram (with external lines),

$$\hat{D}_R(\mathcal{P}_1,\mathcal{P}_2) = d_R(\mathcal{P}_1,\mathcal{P}_2)\,\hat{\mathcal{X}}_R(\mathcal{P}_2), \qquad (5.125)$$

where d_R is the scalar factor (124) and

$$\hat{\mathcal{X}}_R(\mathcal{P}_2) = N\left[X^+_{A(1)} X_{B(1)} X^+_{A(2)} X_{B(2)} \cdots X^+_{A(p)} X_{B(p)} \right],$$

where $A(i), B(i) \in \mathcal{P}_2$, $(i=1,\dots,p)$ are labels carried by external lines of R, which is assumed to have p open paths, and $A(i)$ and $B(i)$ label the outgoing and the ingoing external lines of the i-th open path, respectively. Note that when $\mathcal{P}_2$ is empty ($p=0$), we have $N[\varnothing] = 1$, Eq. (4.17).

(iv) Evaluate, finally, each R-diagram by assigning to it the algebraic expression $Val(R)$,

$$Val(R) = w_R \sum_{\mathcal{P}_1}{}' \sum_{\mathcal{P}_2}{}' \hat{D}_R(\mathcal{P}_1,\mathcal{P}_2), \qquad (5.126)$$

where the summations extend over all free labels. The desired algebraic expression for the product (81) is then obtained as a sum of algebraic quantities (126) from all permissible R-diagrams.

These rules will be amply illustrated on various practical applications in subsequent Chapters.

8. Independent Particle Model Approximation

In introducing the hole-particle formalism we have mentioned that, in a number of cases, one can reasonably well approximate molecular ground states with a single configuration wavefunction. This configuration can then be chosen as a new vacuum, the Fermi vacuum of the hole-particle formalism. This, of course, assumes that the one-particle states or (spin)-orbitals, which are used to define the occupation number representation, have been properly selected. We shall now examine some possible choices for such (spin)orbitals.

We must stress that, formally, the occupation number representation, and the hole-particle formalism, can be used with an arbitrary orthonormal spin-orbital basis. However, it is important for practical applications that the spin-orbital basis is selected in such a way that $|\Phi_0\rangle$ gives the "best possible" single configuration reference state.

The two most common criteria which define such a "best" single-configuration state or, equivalently, the spin-orbitals from which it is constructed, are:

(i) the lowest energy criterion, and
(ii) the criterion of the largest overlap with the exact wavefunction.

These two criteria lead to the Hartree-Fock and Brueckner (or Maximum Overlap) orbitals or approximations, respectively. Since single configuration or a single antisymmetrized product (ASP) wavefunction represent an exact solution for a separable Hamiltonian, given as a sum of one-body terms, the approximations using this form of the wavefunction are referred to as Independent Particle Model (IPM) approximations.

The concept of Brueckner orbitals which maximize the overlap with an exact wavefunction is very important theoretically. However, since this concept requires in principle the knowledge

of the exact wavefunction (cf., Section 8.4 for a more detailed discussion of Brueckner orbitals) one cannot formulate general equations for such orbitals even though one can device techniques leading to approximate Brueckner orbitals.

On the other hand the first criterion of the lowest energy (which is still an upper bound to the exact energy), leading to the Hartree-Fock (HF) approximation, can be easily implemented using the variational principle. In practice this is done by applying the variational principle to a single configuration wavefunction and requiring that the first variation of the energy vanishes for the desired single particle states. This yields the necessary but not the sufficient conditions for the given state to represent the single configuration lowest energy state. These conditions are usually given in the form of Hartree-Fock equations for the single particle states or in an equivalent form of the Brillouin theorem. The corresponding energy

$$E_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle \,, \quad \langle \Phi_0 | \Phi_0 \rangle = 1 \qquad (8.1)$$

then represents a stationary point on the energy hypersurface of a single-configuration variational space.

It is usually assumed that this stationary point represents the lowest upper bound within the IPM. This indeed seems to be the case in a large number of cases, particularly for molecular systems near their equilibrium geometries. However, the fact that E_0 represents a stationary point on the energy hypersurface is no guarantee that it represents an absolute minimum or, in fact, even a local minimum.

In order to find out at least the local character of such a stationary point we can examine the second variation of the energy at this point. This leads to Thouless' stability conditions, which are sufficient conditions for an HF solution to represent a local minimum on the energy hypersurface. They are also necessary but not sufficient conditions for an HF solution to represent the global minimum. For Hamiltonians which are invariant with respect to various symmetry group operations (most importantly for the spin-independent Hamiltonians which are invariant with respect to the SU(2) operations) one can then formulate various symmetry-restricted stability conditions. Such singlet and non-singlet stability conditions for closed-shell states and doublet stability conditions for simple open-shell states, for systems described by a spin-independent Hamiltonian, were first formulated by Čížek and Paldus.

The global stability problem, refered to as a symmetry dilemma, was first clearly stated by Löwdin. He has shown that in contrast to the exact solutions which must be automatically symmetry adapted (i.e., invariant with respect to the symmetry operations possessed by the Hamiltonian), the approximate solutions, determined on the basis of the variational principle, will not be automatically symmetry adapted. In fact, requiring the symmetry adaptation of these approximate solutions imposes constraints on the variational problem which, generally, can only raise the energy. Thus, relaxing the symmetry adaptation requirement may often lead to the IPM broken-symmetry solutions with lower energy than that of the symmetry adapted (also called restricted) solution. Clearly, a local instability of a solution implies the global instability. However, a local stability is no guarantee of a global stability. Unfortunately, it is very difficult to formulate global stability

conditions which would not require a detailed investigation of the whole variational space.

8.1. Exponential form of the IPM wavefunction

Consider an N-electron system and an orthonormal set of spin-orbitals $|A\rangle$, arranged in some definite order and such that the state $|\Phi_0\rangle$, in which the first N spinorbitals are occupied,

$$|\Phi_0\rangle = \left(\prod_{A_1=1}^{N} X_{A_1}^\dagger\right)|0\rangle, \tag{8.2}$$

represents an approximation to the ground state wavefunction of a given system. We also choose $|\Phi_0\rangle$ as the Fermi vacuum of our hole-particle formalism.

An arbitrary N-electron single ASP trial wavefunction $|\Phi\rangle$ has the form

$$|\Phi\rangle = \left[\prod_{A=1}^{N}\left(\sum_{B=1}^{\infty} a_{AB}\, X_B^\dagger\right)\right]|0\rangle. \tag{8.3}$$

We shall be interested in states $|\Phi\rangle$ which are in the "neighborhood" of $|\Phi_0\rangle$. We can thus restrict ourselves to states $|\Phi\rangle$, Eq. (3), which are not orthogonal to $|\Phi_0\rangle$, Eq. (2), and which we can write as

$$|\Phi\rangle = \left(\prod_{B_1}^{N} \tilde{X}_{B_1}^\dagger\right)|0\rangle \tag{8.4}$$

where

$$\tilde{X}_{B_1}^\dagger = \sum_{C_1=1}^{N} f_{B_1C_1}\, X_{C_1}^\dagger + \sum_{C'=N+1}^{\infty} g_{B_1C'}\, X_{C'}^\dagger. \tag{8.4'}$$

In the following derivation we shall drop the summation ranges since they are implied by the type of the summation label. Since $|\Phi\rangle$ is not orthogonal to $|\Phi_0\rangle$, we can also assume an intermediate normalization for $|\Phi\rangle$,

$$\langle\Phi|\Phi_0\rangle = \det\| f_{A_1B_1}\| = 1, \quad \langle\bar{\Phi}_0|\Phi_0\rangle = 1, \tag{8.5}$$

which in turn guarantees a non-singularity of the transformation $\mathbf{f} := \|f_{A_1B_1}\|$ among the occupied spinorbitals.

The state $|\Phi\rangle$, Eq. (4), will not be clearly changed (except for the phase or a normalization factor) if we subject the operators $\tilde{X}_A^\dagger$ to a non-singular transformation $\mathbf{f}^{-1}$,

$$\begin{aligned}\tilde{\tilde{X}}_{B_1}^\dagger &= \sum_{C_1}\left(\mathbf{f}^{-1}\right)_{B_1C_1}\tilde{X}_{C_1}^\dagger \\ &= X_{B_1}^\dagger + \sum_{B'} c_{B_1B'}\, X_{B'}^\dagger,\end{aligned} \tag{8.6}$$

where

$$c_{B_1B'} = \sum_{C_1}\left(\mathbf{f}^{-1}\right)_{B_1C_1} g_{C_1B'}, \tag{8.6'}$$

so that

$$\begin{aligned}|\Phi\rangle &= \left\{\prod_{B_1}\left(X_{B_1}^\dagger + \sum_{B'} c_{B_1B'}\, X_{B'}^\dagger\right)\right\}|0\rangle \\ &= \left\{\prod_{B_1}\left(1 + \sum_{B'} c_{B_1B'}\, X_{B'}^\dagger X_{B_1}\right)X_{B_1}^\dagger\right\}|0\rangle\end{aligned} \tag{8.7}$$

since

$$X_{B_1}X_{B_1}^\dagger|0\rangle = \left(1 - X_{B_1}^\dagger X_{B_1}\right)|0\rangle = |0\rangle. \tag{8.8}$$

We can now exploit the excitation operators

$$E_{B'B_1} := X_{B'}^\dagger X_{B_1}, \tag{8.9}$$

which behave as the generators of $U(\infty)$ given by $E_{AB} = X_A^\dagger X_B$ and which have the properties (the reader may wish to derive these, Cf. Sec. 1.6)

$$\left[E_{B'B_1}^\dagger, X_{C_1}^\dagger\right]_- = X_{B'}^\dagger\langle B_1|C_1\rangle \tag{8.10a}$$

$$\left[E_{B'B_1}, E_{C'C_1}\right]_- = 0 \tag{8.10b}$$

and

$$\left(E_{B'B_1}\right)^2 = E_{B'B_1}E_{C'B_1} = 0. \tag{8.10c}$$

Now, since all the spinorbitals B_1 in the product (7) are distinct, we can use Eq. (10a)

to write

$$|\Phi\rangle = \left\{\prod_{B_1=1}^{N}\left(1+\sum_{B'} c_{B_1B'}\,E_{B'B_1}\right)\right\}\left(\prod_{B_1=1}^{N} X_{B_1}^{\dagger}\right)|0\rangle$$

$$= \left\{\prod_{B_1}\left(1+\sum_{B'} c_{B_1B'}\,E_{B'B_1}\right)\right\}|\Phi_0\rangle. \qquad (8.11)$$

The last product is not necessarily ordered since all operators $E_{B'B_1}$ mutually commute as Eq. (10b) shows. Designating, thus

$$\tilde{E}_{B_1} = \sum_{B'} c_{B_1B'}\,E_{B'B_1}, \qquad (8.12)$$

we have that

$$[\tilde{E}_{B_1}, \tilde{E}_{B_2}]_- = 0\,, \quad (\tilde{E}_{B_1})^2 = 0, \qquad (8.13)$$

so that

$$\prod_{B_1}(1+\tilde{E}_{B_1}) = 1 + \sum_{B_1}\tilde{E}_{B_1} + \sum_{B_1<B_2}\tilde{E}_{B_1}\tilde{E}_{B_2} + \cdots + \sum_{B_1<\cdots<B_N}\tilde{E}_{B_1}\cdots\tilde{E}_{B_N}. \qquad (8.14)$$

In view of the property given by Eq. (10c) we can also write

$$\prod_{B_1}(1+\tilde{E}_{B_1}) = 1 + \hat{C}_1 + \tfrac{1}{2}\hat{C}_1^2 + \cdots + (N!)^{-1}\hat{C}_1^N \qquad (8.15)$$

where

$$\hat{C}_1 = \sum_{B_1}\tilde{E}_{B_1} = \sum_{B_1B'} c_{B_1B'}\,E_{B'B_1}. \qquad (8.16)$$

Moreover, since there are only N distinct occupied spinorbital labels, we have that

$$\hat{C}_1^M = 0 \quad \text{if} \quad M > N, \qquad (8.17)$$

so that we can also write

$$\prod_{B_1}(1+\tilde{E}_{B_1}) = \sum_{M=0}^{\infty}(M!)^{-1}\hat{C}_1^M = e^{\hat{C}_1} \qquad (8.18)$$

and thus

$$|\Phi\rangle = e^{\hat{C}_1}|\Phi_0\rangle, \qquad (8.19)$$

with $\hat{C}_1$ defined by Eq. (16). Clearly, $\hat{C}_1$ is a general mono-excitation operator [cf. Sec. 5.4, Eq. (70a)]

$$\hat{C}_1 = \sum_{B'B_1}\langle B'|\hat{c}_1|B_1\rangle\, X_{B'}^{\dagger}X_{B_1}, \qquad (8.20)$$

where

$$\langle B'|\hat{c}_1|B_1\rangle = c_{B_1B'}. \qquad (8.21)$$

For those readers who are familiar with the basic facts of Lie group theory we would like to remark that the above result could have been anticipated since the operators E_{AB} transform as generators of a general linear or a unitary group on the space spanned by our spin-orbitals. Since any element in the neighborhood of the identity can be written in the exponential form

$$e^{\sum_{AB} c_{AB}E_{AB}} \qquad (8.22)$$

it thus remains to show that in the case considered (a single ASP state) only the generators between the occupied and unoccupied spinorbital states are needed.

8.2. Hartree-Fock Equations

We now investigate the necessary conditions for $|\Phi_0\rangle$ to represent a stationary point on the energy hypersurface. We thus require that the first variation of the energy functional

$$E(\Phi) = \langle\Phi|\hat{H}|\Phi\rangle/\langle\Phi|\Phi\rangle \qquad (8.23)$$

vanishes when $|\Phi\rangle = |\Phi_0\rangle$.

We have seen above that a general ASP wavefunction $|\Phi\rangle$, which is in the "neighborhood" of $|\Phi_0\rangle$, has the form

$$|\Phi\rangle = e^{\hat{C}_1}|\Phi_0\rangle = (1+\hat{C}_1+\tfrac{1}{2}\hat{C}_1^2+\cdots)|\Phi_0\rangle, \qquad (8.24)$$

where the matrix elements of the rectangular

matrix $C = \| c_{B,B'} \|$ can be interpreted as infinitesimal variational parameters which give the admixture of unoccupied spin-orbitals $|B'\rangle$ into the occupied spin-orbitals $|B_1\rangle$, the latter characterizing the desired stationary state $|\Phi_0\rangle$. We note that $|\Phi\rangle$ is normalized to the second order in these variational parameters since

$$\langle \Phi | \Phi \rangle = 1 + \langle \Phi_0 | \hat{C}_1^\dagger \hat{C} | \Phi_0 \rangle + \cdots , \tag{8.25}$$

since $\langle \Phi_0 | \hat{C}_1 | \Phi_0 \rangle$ and similar terms are easily seen to vanish.

Thus, to the first order in variational parameters $c_{B,B'}$ the variation of (23) is given by

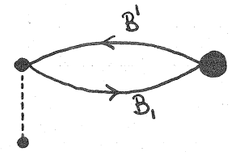
$$\delta^{(1)} E = \langle \Phi_0 | \hat{H} \hat{C}_1 | \Phi_0 \rangle + \langle \Phi_0 | \hat{C}_1^\dagger \hat{H} | \Phi_0 \rangle . \tag{8.26}$$

Since the second term on the right hand side of (26) is complex conjugate of the first one, it is sufficient to consider one of these terms, since they are linearly independent and must vanish independently if $\delta^{(1)} E$ is to vanish. The conditions implied by the vanishing of the other term may then be obtained by a complex conjugation.

Using the normal product of the Hamiltonian (cf., Sec. 4.5),

$$H = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle + \hat{F}_N + \hat{V}_N , \tag{8.27}$$

we see immediately that the only terms which will not automatically vanish for an arbitrary $\hat{C}_1$ involve the one-particle component $\hat{F}_N$. Indeed, using a diagrammatic representation for $\hat{H}$ and $\hat{C}_1$ (cf, Sec. 5.4), we find immediately that the only possible R-diagrams without the external lines representing the Fermi vacuum mean value of $\hat{H}\hat{C}_1$ (or $\hat{F}_N\hat{C}_1$ and $\hat{V}_N\hat{C}_1$) have the form (note also that both operators in the product are in the N-product form)



(8.28)

so that the condition for the first variation $\delta^{(1)} E$ to vanish has the form

$$\sum_{B' B_1} \langle B_1 | \hat{f} | B' \rangle \langle B' | \hat{c}_1 | B_1 \rangle = 0 \tag{8.29}$$

Since the variational parameters $\langle B' | \hat{c}_1 | B_1 \rangle = c_{B,B'}$ are linearly independent, the necessary and sufficient conditions for the first variation to vanish is that the f-matrix elements in (28) or (29) vanish, i.e. that

$$\langle B_1 | \hat{f} | B' \rangle = 0 \tag{8.30}$$

for any B_1 and B'.

These matrix elements represent in fact the matrix elements of the Hamiltonian between $|\Phi_0\rangle$ and corresponding singly excited states, since

$$\langle \Phi_0 | \hat{H} \hat{C}_1 | \Phi_0 \rangle = \sum_{B,B'} c_{B,B'} \langle \Phi_0 | \hat{H} | {}^{B'}_{B_1} \rangle \tag{8.31}$$

and

$$\begin{aligned} \langle \Phi_0 | \hat{H} | {}^{B'}_{B_1} \rangle &= \langle \Phi_0 | \hat{H} \hat{E}_{B'B_1} | \Phi_0 \rangle \\ &= \langle \Phi_0 | \hat{H} \hat{C}_1(B', B_1) | \Phi_0 \rangle \\ &= \langle B_1 | \hat{f} | B' \rangle , \end{aligned} \tag{8.32}$$

where

$$\hat{E}_{B'B_1} = \hat{C}_1(B', B_1) = X_{B'}^\dagger X_{B_1} . \tag{8.33}$$

Thus, the conditions (30) express the so-called Brillouin theorem which states that the matrix elements between the Hartree-Fock ground state and an arbitrary mono-excited state vanish.

The Hartree-Fock spin-orbitals $|B_1\rangle$ as

defined by conditions (30) are only determined up to an arbitrary unitary (or canonical, cf. Sec. 1.8) transformation of the occupied (and, separately, of the unoccupied) spin-orbitals. It is easy to see that $\hat{F}$ is invariant to such a transformation. Indeed, considering a canonical transformation which factors into the separate canonical transformations within the occupied and, separately, within the unoccupied subspaces

$$|B_1\rangle = \sum_{C_1} d_{B_1C_1} |\tilde{C}_1\rangle$$
$$|B'\rangle = \sum_{C'} d_{B'C'} |\tilde{C}'\rangle, \tag{8.34}$$

where $\mathbb{D}_1 := \| d_{B_1C_1} \|$, $\mathbb{D}' := \| d'_{B'C'} \|$ and $\mathbb{D} = \mathbb{D}_1 \oplus \mathbb{D}'$ are unitary, we find that the two-body dependent part of $\hat{F}$ transforms in the same way as the one-body part $\hat{z}$, since

$$\begin{aligned} g_{AB} &= \sum_{B_1} \langle AB_1 | \hat{v} | BB_1 \rangle_{\mathcal{A}} \\ &= \sum_{CD} \sum_{C_1C_2} \sum_{B_1} d^*_{AC} d^*_{B_1C_1} \langle \tilde{C}\tilde{C}_1 | \hat{v} | \tilde{D}\tilde{C}_2 \rangle_{\mathcal{A}} d_{BD} d_{B_1C_2} \\ &= \sum_{CD} \sum_{C_1C_2} \Big(\sum_{B_1} d^*_{B_1C_1} d_{B_1C_2} \Big) d^*_{AC} \langle \tilde{C}\tilde{C}_1 | \hat{v} | \tilde{D}\tilde{C}_2 \rangle_{\mathcal{A}} d_{BD} \\ &= \sum_{CD} \sum_{C_1} d^*_{AC} \langle \tilde{C}\tilde{C}_1 | \hat{v} | \tilde{D}\tilde{C} \rangle_{\mathcal{A}} d_{BD} \\ &= \sum_{CD} d^*_{AC} \tilde{g}_{CD} d_{BD} . \end{aligned} \tag{8.35}$$

Thus

$$\hat{F}_N = \sum_{AB} \langle A | \hat{f} | B \rangle N[X^\dagger_A X_B] = \sum_{\tilde{A}\tilde{B}} \langle \tilde{A} | \hat{f} | \tilde{B} \rangle N[X^\dagger_{\tilde{A}} X_{\tilde{B}}]. \tag{8.36}$$

It is, therefore, most convenient to choose such a unitary transformation which brings $\hat{F}_N$ to a diagonal form (that is, separately within the occupied and unoccupied subspaces; cf. Sec. 3.5), so that

$$\langle B | \hat{f} | A \rangle = \omega_A \langle B | A \rangle , \tag{8.37}$$

for an arbitrary B and A. The ω'_As must be real in view of the hermiticity of $\hat{F}$ and may be interpreted as Hartree-Fock orbital energies

$$\omega_A = \langle A | \hat{f} | A \rangle . \tag{8.38}$$

Since $|B\rangle$ in Eq. (37) is arbitrary, we can also write it as

$$\hat{f} |A\rangle = \sum_C |C\rangle \langle C | \hat{f} | A \rangle = \omega_A |A\rangle \tag{8.39}$$

which is the usual form of <u>Hartree-Fock equations</u> (sometimes called canonical Hartree-Fock equations). Clearly, the operator $\hat{f}$ depends itself on the occupied spinorbitals $|A_i\rangle$ so that (39) represents a pseudo-eigenvalue problem. The equations (39) and their solution in a truncated finite-dimensional subspace spanned by a chosen set of atomic orbitals is usually referred to as the <u>self-consistent field (SCF)</u> equations and solutions, respectively.

8.3. Stability conditions

In order that $|\Phi_0\rangle$ as determined by the Hartree-Fock equations (39) represents a local minimum we have to show that the second variation of $E(\Phi)$, Eq. (23), is positive definite in that stationary point.

Calculate, thus, the energy difference δE,

$$\delta E = E(\Phi) - E(\Phi_0) = \frac{\langle \Phi | \hat{H}_N | \Phi \rangle}{\langle \Phi | \Phi \rangle}, \tag{8.40}$$

where

$$\hat{H}_N = \hat{H} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \hat{F}_N + \hat{V}_N , \tag{8.41}$$

to the second order in the variational parameters $c_{BB'}$ using again the exponential form (19) for the trial wavefunction. We find that

$$\delta E = \delta^{(1)}E + \delta^{(2)}E + \cdots , \tag{8.42}$$

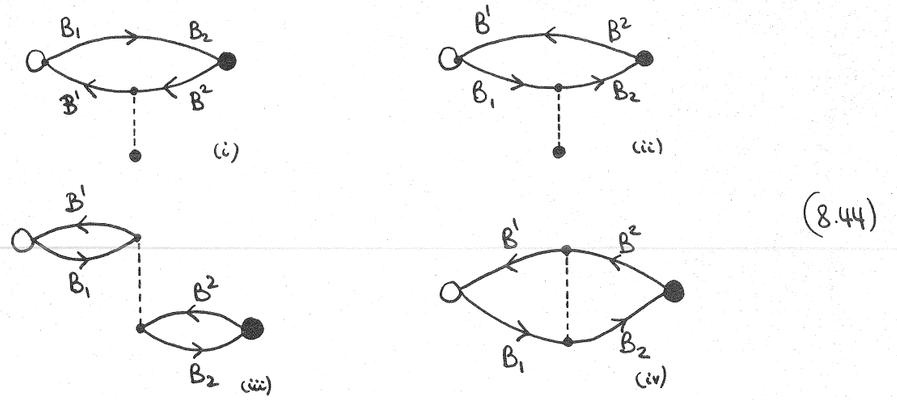
where $\delta^{(1)}E$ is given by Eq. (26) [in which the $\hat{H}$ can be replaced by $\hat{H}_N$] and

$$\delta^{(2)}E = \langle\Phi_0|\hat{C}_1^\dagger \hat{H}_N \hat{C}_1|\Phi_0\rangle + \tfrac{1}{2}\langle\Phi_0|H_N \hat{C}_1^2|\Phi_0\rangle + \tfrac{1}{2}\langle\Phi_0|(\hat{C}_1^\dagger)^2\hat{H}_N|\Phi_0\rangle, \tag{8.43}$$

since $\langle\Phi_0|\hat{H}_N|\Phi_0\rangle = 0$. We also note that $|\Phi\rangle$ is normalized to the second order, so that expanding the denominator in (40) using Eq. (25), the first correction would be of the fourth order, since we assume $|\Phi_0\rangle$ stationary so that $\delta^{(1)}E = 0$.

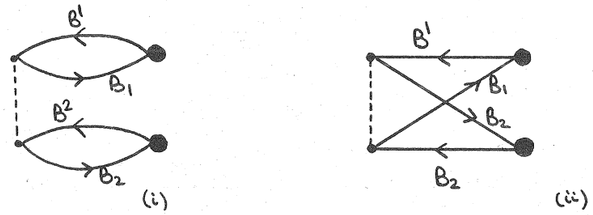
We use again the diagrammatic method to evaluate the required matrix elements. To represent the operator $\frac{1}{2}\hat{C}_1^2$ we shall employ pair of diagrams representing $\hat{C}_1$ (see Chap. 10 for a general representation of the exponential cluster expansion). We thus easily find the following R-diagrams for the relevant matrix elements (note the similarity with the evaluation of matrix elements in Chap. 7):

$\langle\Phi_0|\hat{C}_1^\dagger \hat{H}\hat{C}_1|\Phi_0\rangle$:



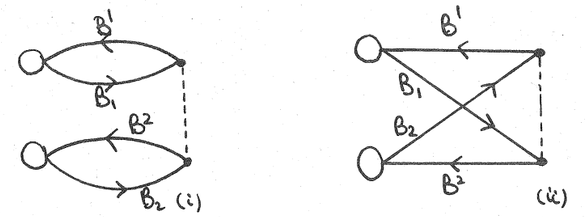
(8.44)

$\frac{1}{2}\langle\Phi_0|\hat{H}_N\hat{C}_1^2|\Phi_0\rangle$:



(8.45)

$\frac{1}{2}\langle\Phi_0|(\hat{C}_1^\dagger)^2\hat{H}_N|\Phi_0\rangle$:



(8.46)

These give the following algebraic expression (note that we could have also conveniently used Brandow diagrams in which case we can disregard the exchange diagrams (44iv), (45ii) and (46ii) or, equivalently, the corresponding Coulomb (direct) ones):

$$\delta^{(2)}E = \sum_{B_1B_2B'B^2}{}' \Big\{ \big[\langle B'|\hat{f}|B^2\rangle\langle B_2|B_1\rangle - \langle B_2|\hat{f}|B_1\rangle\langle B'|B^2\rangle + \langle B'B_2|\hat{v}|B_1B^2\rangle_{AS}\big]\langle B_1|\hat{c}_1|B'\rangle\langle B^2|\hat{c}_1|B_2\rangle + \tfrac{1}{2}\langle B_1B_2|\hat{v}|B'B^2\rangle_{AS}\langle B'|\hat{c}_1|B_1\rangle\langle B^2|\hat{c}_1|B_2\rangle + \tfrac{1}{2}\langle B'B^2|\hat{v}|B_1B_2\rangle_{AS}\langle B_1|\hat{c}_1|B'\rangle\langle B_2|\hat{c}_1|B^2\rangle \Big\} \tag{8.47}$$

Define the matrices $\mathbb{A}$, $\mathbb{B}$ and $\mathbb{C}$, whose rows (and columns) are labeled by mono-excitations $\binom{B'}{B_1}$ arranged in some arbitrary but fixed order, and whose matrix elements are defined as follows

$$A\binom{B'\ B^2}{B_1\ B_2} = \langle B'|\hat{f}|B^2\rangle\langle B_2|B_1\rangle - \langle B_2|\hat{f}|B_1\rangle\langle B'|B^2\rangle + \langle B'B_2|\hat{v}|B_1B^2\rangle_{AS} = \left\langle \begin{matrix} B' \\ B_1 \end{matrix} \middle| \hat{H}_N \middle| \begin{matrix} B^2 \\ B_2 \end{matrix} \right\rangle , \tag{8.48a}$$

$$B\begin{pmatrix} B^1 & B^2 \\ B_1 & B_2 \end{pmatrix} = \langle B_1 B_2 | \hat{r} | B^1 B^2 \rangle_{\boxtimes} = \langle \Phi_0 | \hat{H}_N | {}^{B^1 B^2}_{B_1 B_2} \rangle \qquad (8.48b)$$

and
$$C\begin{pmatrix} B^1 \\ B_1 \end{pmatrix} = \langle B_1 | \hat{c}_1 | B^1 \rangle = c_{B_1 B^1} , \qquad (8.48c)$$

the latter being the column matrix. Clearly, the second variation (47) can then be expressed in a more compact form

$$\delta^{(2)} E = \tfrac{1}{2}\left[\mathbb{C}^* \mathbb{A} \mathbb{C} + \mathbb{C} \mathbb{A}^* \mathbb{C}^* + \mathbb{C}^* \mathbb{B} \mathbb{C}^* + \mathbb{C} \mathbb{B}^* \mathbb{C} \right]$$

$$= \tfrac{1}{2} \begin{pmatrix} \mathbb{C} \\ \mathbb{C}^* \end{pmatrix}^\dagger \begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \overline{\mathbb{B}} & \overline{\mathbb{A}} \end{pmatrix} \begin{pmatrix} \mathbb{C} \\ \mathbb{C}^* \end{pmatrix} \qquad (8.49)$$

since $\mathbb{C}^* \mathbb{A} \mathbb{C} = \mathbb{C} \mathbb{A}^* \mathbb{C}^*$ in view of hermicity of $\mathbb{A}$. Further, since

$$\mathbb{A}^\dagger = \mathbb{A} , \quad \mathbb{B}^\dagger = \mathbb{B} \qquad (8.50)$$

we see that the matrix $\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \overline{\mathbb{B}} & \overline{\mathbb{A}} \end{pmatrix}$ is hermitian.

Thus, in order that the hermitian quadratic form (49) is positive definite it is necessary and sufficient that all the eigenvalues of the matrix

$$\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \overline{\mathbb{B}} & \overline{\mathbb{A}} \end{pmatrix} \qquad (8.51)$$

are positive.

Note also that the eigenvalue problem for (8.51) factorizes into the eigenvalue problems for $(\mathbb{A} \pm \mathbb{B})$ matrices if $\mathbb{A}, \mathbb{B}$ are real. One can also show that for a negative eigenvalue, the corresponding eigenvector gives the direction of steepest descent on the energy hypersurface at the stationary point defined by $|\Phi_0\rangle$.

8.4. Brueckner (maximum overlap) orbitals

As mentioned earlier, these orbitals are of a considerable theoretical significance because of their interesting properties (cf., e.g., Sec. 10.3), even though they can only be determined approximately in actual applications (except for exactly solvable simple models).

The Brueckner independent particle model (IPM) wavefunction $|\Phi_0^B\rangle$ is a single configuration (i.e., a single Slater determinant) built from the maximum overlap spin-orbitals $|A_i\rangle$, $i=1,\dots,N$, which we assume to be orthonormal. Designating again the unoccupied spin-orbitals by superscripted capitals, i.e. $|A^i\rangle$, $i=1,2,\dots$, and assuming that $\{|A_i\rangle, |A^j\rangle\}$ form a complete*) ortho-normal set of single-particle states, we can use them to define the hole-particle formalism creation and annihilation operators with $|\Phi_0^B\rangle$ representing the Fermi vacuum. Then the Brueckner wavefunction (or, more precisely, the corresponding abstract ket state) $|\Phi_0^B\rangle$ can be defined as an IPM state which represents a stationary point (hopefully a maximum)**) of the functional

$$S(\Phi) = \mathrm{Re}\, \langle \Psi_0 | \Phi \rangle / \sqrt{\langle \Phi | \Phi \rangle} , \qquad (8.52)$$

where Ψ_0 is the exact wavefunction of the system and Φ is the (unnormalized) IPM trial function. The state $|\Phi_0^B\rangle$, which corresponds to a global maximum of the functional (52) [and which is not necessarily symmetry adapted as in the HF case**)], represents clearly an IPM state which is closest to the exact one in a given variational space, i.e., the state which has a minimal

*) In all practical applications one considers only a certain finite dimensional one-particle subspace spanned by occupied spinorbitals $|A_i\rangle$ and unoccupied or virtual spinorbitals $|A^i\rangle$, $i=1,\dots,M$, similarly as in the HF case.

**) We can formulate stability conditions for Brueckner orbitals analogously as in the HF case (cf., e.g., Ref.). Similarly, Löwdin's symmetry dilemma arises just as in any other IPM problem.

distance from the exact state since

$$\| \Psi_0' - \Phi' \|^2 = \langle \Psi_0' | \Psi_0' \rangle + \langle \Phi' | \Phi' \rangle - \langle \Psi_0' | \Phi' \rangle - \langle \Phi' | \Psi_0' \rangle$$

$$= 2\left(1 - S(\Phi)/\|\Psi_0\|\right), \qquad (8.53)$$

where $|\Psi_0'\rangle$ and $|\Phi'\rangle$ designate the states $|\Psi_0\rangle$ and $|\Phi\rangle$ normaliged to unity,

$$|\Psi_0'\rangle = |\Psi_0\rangle / \sqrt{\langle \Psi_0 | \Psi_0 \rangle} = |\Psi_0\rangle \cdot \|\Psi_0\|^{-1/2} \qquad (8.54a)$$

$$|\Phi'\rangle = |\Phi\rangle / \sqrt{\langle \Phi | \Phi \rangle} = |\Phi\rangle \cdot \|\Phi\|^{-1/2}. \qquad (8.54b)$$

Writing again the trial state $|\Phi\rangle$ in the exponential form (19),

$$|\Phi\rangle = e^{\hat{C}_1} |\Phi_0^B\rangle, \qquad (8.55)$$

and assuming $|\Phi_0^B\rangle$ to be non-degenerate, we can find the necessary and sufficient conditions for the stationary points of the functional (52). To do so, we have to know of course the exact wavefunction Ψ_0, which we assume to be expanded in the N-electron basis constructed from the configurations built from the maximum overlap spin-orbitals $|A_i\rangle$ and $|A^i\rangle$. We thus write

$$|\Psi_0\rangle = \sum_{k=0}^{N} \hat{C}_r^B |\Phi_0^B\rangle, \qquad (8.56)$$

where

$$\hat{C}_0^B = \hat{1} \qquad \text{(intermediate normalization)}, \qquad (8.57)$$

and

$$\hat{C}_r^B = (r!)^{-2} \sum_{\substack{D' \cdots D^r \\ D_1 \cdots D_r}} \langle D' \cdots D^r | \hat{C}_r^B | D_1 \cdots D_r \rangle_{AS} X_{D'}^\dagger \cdots X_{D^r}^\dagger X_{D_r} \cdots X_{D_1}, \qquad (8.58)$$

as in Eq. (5.70b). The r-times excited component of $|\Psi_0\rangle$ can then be written as

$$\hat{C}_r^B |\Phi_0\rangle = \sum_{\substack{D'<\cdots<D^r \\ D_1<\cdots<D_r}} \langle D' \cdots D^r | \hat{C}_r^B | D_1 \cdots D_r \rangle_{AS} \begin{vmatrix} D' \cdots D^r \\ D_1 \cdots D_r \end{vmatrix}$$

$$= \sum_{\substack{D'<\cdots<D^r \\ D_1<\cdots<D_r}} C_r^B \begin{pmatrix} D' \cdots D^r \\ D_1 \cdots D_r \end{pmatrix} \begin{vmatrix} D' \cdots D^r \\ D_1 \cdots D_r \end{vmatrix}, \qquad (8.59)$$

where the superscript B indicates that the Brueckner (maximum overlap) orbitals has been employed in the configuration-interaction-type expansion of $|\Psi_0\rangle$ and $C_r^B \begin{pmatrix} D' \cdots D^r \\ D_1 \cdots D_r \end{pmatrix}$ designate the appropriate configuration interaction coefficients.

In view of the orthonormality of the configurations we find that the variation of $S(\Phi)$ to the first order in the infinitesimal variational parameters, defined by matrix elements of $\hat{C}_1$, Eq. (55), is

$$\delta S(\Phi) = \mathrm{Re} \langle \Phi_0^B | \hat{C}_1^{B\dagger} \hat{C}_1 | \Phi_0^B \rangle + O(\hat{C}_1^2). \qquad (8.60)$$

We recall that $|\Phi_0^B\rangle$ is normalized to the second order in variational parameters, Eq. (25). Consequently, the necessary and sufficient conditions for the first variation (60) to vanish is that $\hat{C}_1^B$ itself vanishes, i.e.

$$C_1^B \begin{pmatrix} D' \\ D_1 \end{pmatrix} = 0 \qquad \text{for any } D_1 \text{ and } D'. \qquad (8.61)$$

In other words, in the configuration interaction expansion of the exact wavefunction in terms of the configurations built from Brueckner orbitals, all mono-excited contributions must be absent. This property is sometimes used to define Brueckner orbitals: it then follows that the corresponding IPM state is closest to the exact state, i.e., has a maximum overlap with $|\Psi_0\rangle$.

The conditions (61) should be paralleled with the Hartree-Fock conditions as expressed through the Brillouin theorem, Eqs. (36), (32). The latter require that the Hamiltonian matrix elements (32) between the HF ground state and all possible mono-excited states vanish. Consequently, there is no first order contribution of the mono-excited states in

the exact wavefunction, and only higher than the first order contributions can contribute through the interaction of the ground state Hartree-Fock state and mono-excited states via doubly, triply, etc.-excited configurations. In contrast, with the Brueckner orbitals, the ground state and mono-excited state interaction matrix elements summed to all orders exactly vanish, so that mono-excited states do not appear in the exact wavefunction. The difference between HF and Brueckner orbitals are usually very small, since even with the HF wavefunction used as a reference state the mono-excited contribution, appearing only through higher than the first order interactions, is very small. Consequently, only a very small difference* was found between HF and Brueckner molecular orbitals in model studies, when the latter can be determined.

*) The atomic orbital expansion coefficients for the PPP model of butadiene differed only in the sixth or seventh decimal place.

**) Note that HF and Brueckner molecular orbitals are identical in case of minimum basis set when the molecular orbitals are given by symmetry (e.g., cyclicpolyene PPP model).

9. Perturbation Theory

In this chapter we shall develop the formalism of a time-independent Rayleigh-Schrödinger perturbation theory (RSPT) for a non-degenerate level, which is often refered to as many-body perturbation theory (MBPT) due to its size extensive properties. We shall start with an elementary derivation of general formulas and will investigate the structure of various terms appearing in an arbitrary order using diagrammatic approach. The main characteristics of this structure are embodied in the so called linked cluster theorem, first postulated by Brueckner and latter proved generally by Goldstone, Hubbard and Hugenholtz, who introduced for the first time the mathematical techniques of quantum field theory into the study of a many-fermion correlation problem. These results were later extended by Bethe, Brandow, Lindgren and others and the first applications of the MBPT to the atomic correlation problem were given by Kelly. Nowadays numerous texts exist which give a detail account of MBPT. Most of these texts follow the original time-dependent derivations of Goldstone or Hubbard. We wish to present a completely time-independent derivation following basically our earlier expository article.

9.1. Rayleigh-Schrödinger perturbation formula

We wish to find the eigenvalues and corresponding eigenvectors of the perturbed problem

$$\hat{K}\,|\Psi_i\rangle = k_i\,|\Psi_i\rangle, \qquad (9.1)$$

whose Hamilton operator $\hat{K}$ is given as a sum of the unperturbed operator $\hat{K}_0$ and of the perturbation $\hat{W}$,

$$\hat{K} = \hat{K}_0 + \hat{W}, \tag{9.2}$$

assuming that we know the exact eigenvalues $\varkappa_i$ and eigenstates $|\Phi_i\rangle$ of $\hat{K}_0$,

$$\hat{K}_0 |\Phi_i\rangle = \varkappa_i |\Phi_i\rangle. \tag{9.3}$$

We shall further assume that the states $|\Phi_i\rangle$ are non-degenerate, orthonormal

$$\langle \Phi_i | \Phi_j \rangle = \delta_{ij}, \tag{9.4}$$

and that they go over into the corresponding perturbed states $|\Psi_i\rangle$ when the perturbation $\hat{W}$ is "switched on".

In deriving basic perturbation theoretical expressions for the perturbed eigenstates and eigenvalues we shall exploit the so-called asymmetric formula for the energy,

$$k_i = \frac{\langle \Phi_i | \hat{K} | \Psi_i \rangle}{\langle \Phi_i | \Psi_i \rangle} = \varkappa_i + \frac{\langle \Phi_i | \hat{W} | \Psi_i \rangle}{\langle \Phi_i | \Psi_i \rangle}, \tag{9.5}$$

which follows immediately from (1) when we project it onto $|\Phi_i\rangle$,

$$\begin{aligned} \langle \Phi_i | \hat{K} | \Psi_i \rangle &= \varkappa_i \langle \Phi_i | \Psi_i \rangle + \langle \Phi_i | \hat{W} | \Psi_i \rangle \\ &= k_i \langle \Phi_i | \Psi_i \rangle. \end{aligned} \tag{9.6}$$

We also define a projection operator $\hat{R}_i$ onto the orthogonal complement of an unperturbed eigenstate $|\Phi_i\rangle$,

$$\hat{R}_i = \hat{1} - |\Phi_i\rangle\langle\Phi_i| = \sum_{j(\neq i)} |\Phi_j\rangle\langle\Phi_j|. \tag{9.7}$$

We can easily verify that this operator is both idempotent and Hermitian,

$$\hat{R}_i^2 = \hat{R}_i, \quad \hat{R}_i^\dagger = \hat{R}_i \tag{9.8}$$

and that

$$\hat{R}_i |\Phi_i\rangle = 0. \tag{9.9}$$

It is also easy to see that $\hat{R}_i$ commutes with an unperturbed operator $\hat{K}_0$,

$$[\hat{R}_i, \hat{K}_0]_- = 0, \tag{9.10}$$

since

$$(\hat{1} - |\Phi_i\rangle\langle\Phi_i|)\hat{K}_0 = \hat{K}_0 - \varkappa_i |\Phi_i\rangle\langle\Phi_i| = \hat{K}_0(\hat{1} - |\Phi_i\rangle\langle\Phi_i|). \tag{9.11}$$

Further, since $\hat{R}_i$ obviously commutes with any scalar operator, we have also that

$$[\hat{R}_i, \varkappa_i - \hat{K}_0]_- = 0, \tag{9.12}$$

which implies that

$$\begin{aligned} (\varkappa_i - \hat{K}_0)\hat{R}_i |\Psi_i\rangle &= \hat{R}_i (\varkappa_i - \hat{K}_0)|\Psi_i\rangle \\ &= \hat{R}_i (\varkappa_i - \hat{K} + \hat{W})|\Psi_i\rangle \\ &= \hat{R}_i (\varkappa_i - k_i + \hat{W})|\Psi_i\rangle, \end{aligned} \tag{9.13}$$

where we have used Eqs. (2) and (1) in the last two steps, respectively.

Let us assume that $(\varkappa_i - \hat{K}_0)$ is invertible*), so that we can write (13) as

$$\begin{aligned} \hat{R}_i |\Psi_i\rangle &= (\varkappa_i - \hat{K}_0)^{-1} \hat{R}_i (\varkappa_i - k_i + \hat{W})|\Psi_i\rangle \\ &= \hat{P}_i (\varkappa_i - k_i + \hat{W})|\Psi_i\rangle, \end{aligned} \tag{9.14}$$

where we defined

$$\hat{P}_i = (\varkappa_i - \hat{K}_0)^{-1}\hat{R}_i = \hat{R}_i (\varkappa_i - \hat{K}_0)^{-1} = \frac{\hat{R}_i}{\varkappa_i - \hat{K}_0} \tag{9.15}$$

in view of (12). Writing, formally, the inverse of

*) Assuming non-degeneracy of $\varkappa_i$ eigenvalues, this operator is singular on $|\Phi_i\rangle$ but non-singular on its complement, i.e. on $\hat{R}_i|f\rangle$, $|f\rangle$ arbitrary.

$(x_i - \hat{K}_0)$ in the form of an infinite series expansion

$$(x_i - K_0)^{-1} = x_i^{-1}\,(1 - \frac{\hat{K}_0}{x_i})^{-1} = x_i^{-1} \sum_{j=0}^{\infty} \left(\frac{\hat{K}_0}{x_i}\right)^j , \qquad (9.16)$$

we see immediately that (3) implies that

$$(x_i - K_0)^{-1} |\Phi_j\rangle = (x_i - x_j)^{-1} |\Phi_j\rangle , \quad (j \neq i), \qquad (9.17)$$

so that we can also write operator $\hat{P}_i$, Eq. (15), in the form

$$\hat{P}_i = \sum_{j(\neq i)} \frac{|\Phi_j\rangle\langle\Phi_j|}{x_i - x_j} . \qquad (9.18)$$

This operator must also annihilate $|\Phi_i\rangle$, i.e.

$$\hat{P}_i |\Phi_i\rangle = 0 . \qquad (9.19)$$

Recalling now definition of the projector $\hat{R}_i$, Eq. (7), which implies that

$$\hat{R}_i |\Psi_i\rangle = |\Psi_i\rangle - |\Phi_i\rangle\langle\Phi_i|\Psi_i\rangle$$

$$= |\Psi_i\rangle - N_i |\Phi_i\rangle , \qquad (9.20)$$

where

$$N_i = \langle\Phi_i|\Psi_i\rangle \qquad (9.21)$$

is an intermediate normalization factor for the i-th eigenstate, we obtain from Eq. (14) the following recursion relationship for $|\Psi_i\rangle$,

$$|\Psi_i\rangle = N_i |\Phi_i\rangle + \hat{R}_i |\Psi_i\rangle$$

$$= N_i |\Phi_i\rangle + \hat{P}_i (x_i - k_i + \hat{W}) |\Psi_i\rangle . \qquad (9.22)$$

Iterating this relationship we get

$$|\Psi_i\rangle = N_i \sum_{m=0}^{\infty} [\hat{P}_i (x_i - k_i + W)]^m |\Phi_i\rangle , \qquad (9.23)$$

and using the asymmetric formula for the eigen-

value, Eq. (5), we get the desired RS perturbation formula

$$k_i = x_i + \sum_{m=0}^{\infty} \langle\Phi_i|\hat{W}\,[\hat{P}_i (x_i - k_i + \hat{W})]^m |\Phi_i\rangle . \qquad (9.24)$$

Thus, once we determine the exact eigenvalue k_i from this equation, the corresponding eigenvector $|\Psi_i\rangle$ is given by (23) assuming an intermediate normalization

$$N_i = \langle\Phi_i|\Psi_i\rangle = 1 . \qquad (9.25)$$

An equivalent result could have been derived by expanding the denominators, containing the exact eigenvalue k_i, of the Brillouin-Wigner perturbation expansion.

9.2. Renormalization terms and bracketing technique*)

The RS formulas for the perturbed eigenvalues, Eq. (24), and eigenvectors, Eq. (23), do not contain the exact eigenvalue k_i in the denominators in contrast to the Brillouin-Wigner perturbation theory. Nevertheless, these formulas are not explicit since k_i appears on the rhs of Eq. (24) in terms with $n > 0$. Approximating eigenvalues k_i appearing on the rhs of Eqs (24) and (23) by the same expansion (24) leads to the occurence of the so called "renormalization terms" when the terms which are of the same order in the perturbation W are grouped together. We thus write

$$k_i = \sum_{j=0}^{\infty} k_i^{(j)} \qquad (9.26)$$

and

* To avoid any misunderstanding we note that this technique has no relationship to the "bracketing function" technique of Löwdin (in P.T. and its Applications in Q.M., C.H. Wilcox, Ed., p. 255, Wiley, N.Y., 1966).

$$|\Psi_i\rangle = \sum_{j=0}^{\infty} |\Psi_i^{(j)}\rangle \,, \tag{9.27}$$

where the superscript j indicates the jth order component of k_i or $|\Psi_i\rangle$, i.e. the terms which are of the j-th order in $\hat{W}$.

The zero order term $k_i^{(0)}$ is given by the unperturbed eigenvalue,

$$k_i^{(0)} = \varkappa_i = \langle \Phi_i | \hat{K}_0 | \Phi_i \rangle \tag{9.28}$$

and, assuming an intermediate normalization of $|\Psi_i\rangle$, Eq. (25), the zero order component of $|\Psi_i\rangle$ equals the unperturbed state,

$$|\Psi_i^{(0)}\rangle = |\Phi_i\rangle \tag{9.29}$$

as may have been expected. Let us now find a general form of the k-th order terms considering first the eigenvalues.

9.2.1. Eigenvalue expansion

The first and the second order contributions $k_i^{(1)}$ and $k_i^{(2)}$ are also given directly by the corresponding $n=0$ and $n=1$ terms on the rhs of Eq. (24), just as the zero order term (28) above,

$$k_i^{(1)} = \langle \Phi_i | \hat{W} | \Phi_i \rangle \tag{9.30}$$

$$k_i^{(2)} = \langle \Phi_i | \hat{W} \hat{P}_i (\hat{W} + \varkappa_i - \hat{K}_i) | \Phi_i \rangle = \langle \Phi_i | \hat{W} \hat{P}_i \hat{W} | \Phi_i \rangle, \tag{9.31}$$

since (19) holds. For the sake of simplicity we now drop the subscripts i and use shorthand notation for corresponding mean values, so that Eqs. (28), (30) and (31) take the following simple form

$$k^{(0)} = \langle \hat{K}_0 \rangle = \varkappa \tag{9.32}$$
$$k^{(1)} = \langle W \rangle \tag{9.33}$$
$$k^{(2)} = \langle WPW \rangle \,, \tag{9.34}$$

where we also deleted the operator symbols over $\hat{P}$ and $\hat{W}$.

We now see that the terms in (24) with $n \geq 2$ will yield in addition to a principal term of the type $\langle W(PW)^n \rangle$ also various "renormalization" terms involving products of lower order components. Thus, from the $n=2$ term we get

$$\begin{aligned} &\langle WP(W+\varkappa-k)P(W+\varkappa-k)\rangle \\ &= \langle WP(W+\varkappa-\sum_{j=0}^{\infty} k^{(j)})PW\rangle \\ &= \langle WP(W-\sum_{j=1}^{\infty} k^{(j)})PW\rangle \\ &= \langle W(PW)^2\rangle - \sum_{j=1}^{\infty} k^{(j)} \langle WP^2W\rangle . \end{aligned} \tag{9.35}$$

The third order term $k^{(3)}$ will thus contain a renormalization term involving the first order contribution ($j=1$) from the infinite sum in (35),

$$\begin{aligned} k^{(3)} &= \langle W(PW)^2\rangle - k^{(1)} \langle WP^2W\rangle \\ &= \langle W(PW)^2\rangle - \langle W\rangle\langle WP^2W\rangle, \end{aligned} \tag{9.36}$$

while the remaining terms of the form

$$k^{(j)} \langle WP^2W\rangle \,, \qquad j \geq 2 \tag{9.37}$$

from the rhs of (35) have to be included in $(j+2)$-th order terms $k^{(j+2)}$, respectively.

Thus, each term on the rhs of Eq. (24), $n \geq 2$, brings new renormalization terms which contribute in all subsequent orders. A general form of these terms in each order is easily found, since the sum of all powers to which W appears must equal the given order. In the fourth order, for example, we shall have the renormalization term $k^{(2)}\langle WP^2W\rangle$, Eq. (37), from the $n=2$ term (35), in addition to the principal and renormalization terms arising

from the $n=3$ term of (24), which are of the type

$$\langle WP(W-k^{(j)})P(W-k^{(\ell)})PW\rangle$$

$$= \langle W(PW)^3\rangle - k^{(j)}\langle WP^2WPW\rangle - k^{(\ell)}\langle WPWP^2W\rangle + k^{(j)}k^{(\ell)}\langle WP^3W\rangle \quad ; \quad (j,\ell \geq 1). \tag{9.38}$$

The terms which contribute in the same order as the principal term are those for which $j=\ell=1$. This is true for any order of perturbation theory. The renormalization terms may thus be simply obtained by removing one or more "internal" (i.e., other than the first and the last) W operators from the principal term and replacing it by $k_i^{(i)} = \langle W\rangle$. This operation may be interpreted as a so-called "bracketing". For the fourth order term (38) this can be expressed as follows

$$\langle WPWPWPW\rangle - \langle WP\langle W\rangle PWPW\rangle - \langle WPWP\langle W\rangle PW\rangle + \langle WP\langle W\rangle P\langle W\rangle PW\rangle$$

$$= \langle W(PW)^3\rangle - \langle W\rangle(\langle WP^2WPW\rangle + \langle WPWP^2W\rangle) + \langle W\rangle^2\langle WP^3W\rangle. \tag{9.39}$$

We also see that the sign of each term is simply determined by the number of bracketings r; i.e. $(-1)^r$.

The terms which propagate from the lower order terms, $n<4$ in the example considered, may be similarly obtained if bracketings over an arbitrary number of "internal" W operators are included. In the example of the fourth order contribution considered, the term from the $n=2$ term of (24), given by Eq. (37) with $j=2$,

$$-k^{(2)}\langle WP^2W\rangle = -\langle WPW\rangle\langle WP^2W\rangle \tag{9.40}$$

is obtained by the following bracketing of

the principal term

$$\langle WP\langle WPW\rangle PW\rangle, \tag{9.41}$$

and its sign is again given by the rule $(-1)^r$, where r is the number of bracketings. The fourth order contribution is thus given by the principal term and all the possible bracketings of the internal W operators (bracketings of the external, i.e. the first and the last $\hat{W}$ operators, are not permitted in view of the property (19)) taken with the signs given by the $(-1)^r$ rule:

$$k^{(4)} = \langle W(PW)^3\rangle - \langle W\rangle(\langle WP(PW)^2\rangle + \langle(WP)^2PW\rangle) + \langle W\rangle^2\langle WP^3W\rangle - \langle WPW\rangle\langle WP^2W\rangle. \tag{9.42}$$

The same rule applied to the fifth order term yields

$$\begin{aligned} k^{(5)} = {} & \langle W(PW)^4\rangle && \text{(i)}\\ & - \langle W\rangle(\langle WP(PW)^3\rangle + \langle(WP)^2(PW)^2\rangle + \langle(WP)^3PW\rangle) && \text{(ii)}\\ & + \langle W\rangle^2(\langle WP^2(PW)^2\rangle + \langle(WP^2)^2W\rangle + \langle(WP)^2P^2W\rangle) && \text{(iii)}\\ & - \langle W\rangle^3\langle WP^4W\rangle && \text{(iv)}\\ & - \langle WPW\rangle(\langle WP(PW)^2\rangle + \langle(WP)^2PW\rangle) && \text{(v)}\\ & + 2\langle W\rangle\langle WPW\rangle\langle WP^3W\rangle && \text{(vi)}\\ & - (\langle(WP)^2W\rangle - \langle W\rangle\langle WP^2W\rangle)\langle WP^2W\rangle, && \text{(vii)} \end{aligned} \tag{9.43}$$

which can be directly verified. The terms (ii)–(iv) originate from simple bracketings in the principal term (i) over one, two or three

"isolated" W operators, the terms (v)(vi) come from the $n=3$ term, Eq. (38), and the last term (vii) from the $n=2$ term, Eq. (35). In this last term, which is of the form (37) for $j=3$, we have substituted the expression (36) for $k^{(3)}$, which itself contains the renormalization term. Such terms may be again formally obtained in the explicit form if we allow bracketings within a bracketing. In the fifth order case there is only one such possible bracketing, namely

$$\langle WP\langle WP\langle W\rangle PW\rangle PW\rangle$$

$$= \langle W\rangle\langle WP^2W\rangle^2 \qquad (9.44)$$

which is contained in (vii). In the sixth order, we have several possibilities for a bracketing within a bracketing, for example

$$\langle WP\langle W\rangle P\langle WP\langle W\rangle PW\rangle PW\rangle$$

$$= -\langle W\rangle^2\langle WP^2W\rangle\langle WP^3W\rangle, \qquad (9.45)$$

where there are two brackets inserted in the principal one and, moreover, in the second bracket there is inserted another bracket, or

$$\langle WP\langle WP\langle W\rangle P\langle W\rangle PW\rangle PW\rangle = -\langle W\rangle^2\langle WP^2W\rangle\langle WP^3W\rangle, \qquad (9.46)$$

where we have inserted two brackets inside the bracket, which is itself already an inserted bracket in the principal one. Note that the different bracketings (45) and (46) yield identical terms as in the case (vi), Eq. (43).

The general bracketing rules were first formulated by Hubby and can be summarized as follows:

R: The nth order contribution $k^{(n)}$ is given by the sum of the principal term $\langle W(PW)^{n-1}\rangle$ and of the renormalization

terms, which are obtained by inserting any number of non-straddling bra-ket pairs into the principal term in all possible ways so that no two brackets touch, and assigning each term the sign $(-1)^r$, where r is the number of inserted bracketings.

The total number of terms in the n-th order, including the principal term, was given by Tong

$$\frac{1}{n-1}\binom{2(n-1)}{n} = \frac{(2n-2)!}{n!\,(n-1)!}. \qquad (9.47)$$

The number of bracketings obviously drastically increases with an increasing order of perturbation theory. This is partially caused by the fact that the lower order perturbation terms $k^{(m)}$, $m<n$, are always expressed explicitly through the principle m-th order and relevant renormalization m-th order terms. This occurs for the first time in the fifth order, Eq. (43), where the last term (vii) originates from expression (37) with $j=3$. In higher orders, more and more complex expressions with a large number of renormalization terms will be so inserted. Therefore, if we shall keep the lower order contributions $k^{(m)}$ in our formulas in the implicit form, we shall obtain a considerably more compact recursive-type formulas. This can be achieved formally by defining "super-bracketings" as the m-th order terms $\bar{k}^{(m)}$,

$$\bar{k}^{(m)} = \{W(PW)^{m-1}\}, \qquad (9.48)$$

where we use curly rather than angular brackets.

The rules for super-bracketings are then obviously similar to those for the usual bracketings with the exception that it is now forbidden to insert a superbracket into another superbracket. Thus, $k^{(n)}$ is given by the principal (or leading) term and the renormalization terms given by all possible superbracketings, using the same sign rule as for bracketings. Thus, the last term (vii) in (43) would originate from a single third order superbracket rather than from a similar bracketing and a bracketing within this bracketing. For example, the following single superbracketing appearing in the ninth order,

$$\langle WP\{W(PW)^4\}P\{W\}PW\rangle = k^{(1)}k^{(5)}\langle WP^3W\rangle, \tag{9.49}$$

represents 14 bracketings (given by the $k^{(5)}$ term, Eq.(43)).

The superbracketing technique is probably the best compromise between the completely explicit Brueckner-Hubby bracketing rules and the usual implicit but asymmetric scheme.

Another possibility is to introduce "symmetrized" quantities

$$\langle W\rangle_s^r = \sum_{\{t_i\}}{}' \langle WP^{t_1}WP^{t_2}W\cdots WP^{t_{r-1}}W\rangle, \tag{9.50}$$

where the sum extends over all positive integers t_i such that

$$\sum_{i=1}^{r-1} t_i = s \quad , \quad t_i \geq 1 \, . \tag{9.51}$$

Thus, if $r=1$, then $s=0$, so that

$$\langle W\rangle_0^1 = \langle W\rangle, \; \langle W\rangle_s^1 = 0 \text{ if } s>0, \tag{9.52}$$

while for $r \geq 2$ we must have that $s \geq r-1$. Introduction of these quantities is suggested, for example, by the general structure of the fifth order term, Eq. (43). Generally, we can write the n-th order term explicitly as follows

$$k^{(n)} = \sum_{i=1}^{n-1} (-1)^{i+1} \sum_{m_1\leq m_2\leq\cdots\leq m_i}{}' \sum_{m_1 m_2\cdots m_i}{}' \prod_{j=1}^{i} \langle W\rangle_{m_i}^{m_i} , \tag{9.53}$$

where the second and the third summations extend over the terms satisfying the following conditions:

$$m_j \geq n_j - 1 > 0 \tag{9.54a}$$

$$m_j = 0 \quad \text{if} \quad n_j = 1 \tag{9.54b}$$

and

$$\sum_{j=1}^{i} n_j = n = 1 + \sum_{j=1}^{i} m_j \, . \tag{9.54c}$$

This explicit expression nicely reveals the internal "symmetry" of a general n-th order term, and the number of bracketings is given by $(i-1)$.

9.2.2. Eigenfunction expansion

Once the eigenvalue contributions $k^{(n)}$ are known it is straightforward to obtain explicit expressions for the eigenfunction components, Eqs. (23) and (27). In view of the property (19) it is convenient to rewrite Eq. (23) as follows (dropping again the subscripts for simplicity)

$$\begin{aligned} |\Psi\rangle &= \sum_{n=0}^{\infty} [P(W+\varepsilon-k)]^n |\Phi\rangle \\ &= |\Phi\rangle + \sum_{n=0}^{\infty} [P(W+\varepsilon-k)]^n P(W+\varepsilon-k)|\Phi\rangle \\ &= |\Phi\rangle + \sum_{n=0}^{\infty} [P(W-\sum_{j=1}^{\infty} k^{(j)})]^n PW\,|\Phi\rangle, \end{aligned} \tag{9.55}$$

assuming an intermediate normalization, Eq. (25).

For the first few components $|\Psi^{(n)}\rangle$ we thus get immediately the following expressions (which, of course, are also simply obtained by removing the bra state $\langle\Phi|$ and the leftmost W interaction from the corresponding

eigenvalue contributions, as suggested by the asymmetric energy formula, Eq. (5))

$$|\Psi^{(0)}\rangle = |\Phi\rangle \tag{9.56a}$$

$$|\Psi^{(1)}\rangle = PW|\Phi\rangle \tag{9.56b}$$

$$|\Psi^{(2)}\rangle = [(PW)^2 - \langle W\rangle P^2 W]|\Phi\rangle \tag{9.56c}$$

$$|\Psi^{(3)}\rangle = [(PW)^3 - \langle W\rangle (P^2WPW + PWP^2W) + \langle W\rangle^2 P^3W - \langle WPW\rangle P^2W]|\Phi\rangle \tag{9.56d}$$

We can obviously use again the bracketing technique if we wish to obtain these terms directly. Writing the principal term in the same shorthand notation as above, i.e. as $(PW)^n\rangle$, the same bracketing rules,*) as formulated for the eigenvalues, will apply.

The asymmetric energy formula, Eq. (5), implies the following simple rule for obtaining a corresponding eigenvalue contribution (and vice versa as mentioned above)

$$k^{(n+1)} = \langle W|\Psi^{(n)}\rangle . \tag{9.57}$$

It must be noted, however, that the n-th order eigenfunction contribution determins in fact the (2n+1)st energy contribution when the symmetric energy expression is used (so called Wigner's (2n+1) rule).

*) The rule that brackets cannot touch implies that we cannot bracket the rightmost W in the eigenfunction case considered. Recall that in the eigenvalue case this same rule eliminated both "external" (i.e., the leftmost and the rightmost) W's from bracketings.

We also note the original compact expression for $|\Psi^{(n)}\rangle$ due to Bloch, which is equivalent to the expression obtained via bracketing technique as demonstrated by Hubby; namely

$$|\Psi^{(n)}\rangle = \sum{}' S^{k_1} W S^{k_2} W \cdots W S^{k_n} W |\Phi\rangle \tag{9.58}$$

where

$$S^0 = -|\Phi\rangle\langle\Phi| \tag{9.59a}$$

and

$$S^k = P^k , \quad k > 0 , \tag{9.59b}$$

and where the summation extends over all non-negative integers k_i for which

$$\sum_{i=1}^{n} k_i = n , \quad \sum_{i=1}^{p} k_i \geq p \quad (1 \leq p < n) . \tag{9.60}$$

Let us mention, finally, that all the renormalization terms, either for an eigenvalue or an eigenfunction expansion, are in fact cancelled, in each order, by the so-called unlinked terms originating from the principal term. Consequently, these terms will not appear in final expressions for the n-th order eigenvalue or eigenfunction contributions once formulated using diagrammatic rules.

9.3. Ground State Perturbation Theory

Consider, now, the above developed RSPT for the case of a non-degenerate ground state of some Hamiltonian $\hat{H}$ containing at most two-particle interactions, Eq. (4.61),

$$\hat{H} = \hat{Z} + \hat{V} , \quad \hat{Z} = \sum_{AB} \langle A|\hat{z}|B\rangle X_A^\dagger X_B$$

$$\hat{V} = \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle X_A^\dagger X_B^\dagger X_D X_C . \tag{9.61}$$

We shall approximate this Hamiltonian by a separable unperturbed Hamiltonian $\hat{H}_o$,

$$\hat{H}_o = \hat{Z} + \hat{U}\,, \qquad \hat{U} = \sum_{AB} \langle A|\hat{u}|B\rangle\, X_A^+ X_B \tag{9.62}$$

where $\hat{U}$ represents some one-electron approximation to V. For example, we can choose for the operator $\hat{U}$ the operator $\hat{F}$, Eq. (4.75), i.e. we can employ the Hartree-Fock approximation. However, we do not wish to make this choice at this stage in order to keep our derivations as general as possible. We shall only assume that we can find exact eigensolutions of $\hat{H}_o$,

$$\hat{H}_o\,|\Phi_i\rangle = \varepsilon_i\,|\Phi_i\rangle\,, \tag{9.63}$$

and that $|\Phi_o\rangle$ represents an approximation to the exact ground state $|\Psi_o\rangle$,

$$\hat{H}\,|\Psi_i\rangle = E_i\,|\Psi_i\rangle\,, \tag{9.64}$$

of the full Hamiltonian (61).

In view of the separability of $\hat{H}_o$, the problem (63) reduces to the solution of a one-particle problem*)

$$(\hat{z} + \hat{u})\,|A\rangle = \omega_A\,|A\rangle\,. \tag{9.65}$$

We assume that the one-particle states $|A\rangle$ are ordered according to increasing one-particle energies ω_A and that the unperturbed N-particle ground state $|\Phi_o\rangle$ is given as an anti-symmetrized product of the first N eigenstates*) $|A\rangle$. We can then choose $|\Phi_o\rangle$ as the Fermi vacuum of our hole-particle formalism and label the

*) Note that $\hat{u}$ may depend on the solutions $|A\rangle$, as in the Hartree-Fock theory, so that (65) is a pseudo-eigenvalue problem.

eigenstates occupied in $|\Phi_o\rangle$, representing the hole states, by subscripted capitals, i.e.

$$|\Phi_o\rangle = |\{A_1 A_2 \cdots A_N\}\rangle\,, \tag{9.66}$$

while the remaining "virtual" eigenstates of (65), representing particle states, will be labeled by superscripted capitals, $|A^i\rangle$, in agreement with our earlier convention.

In order to simplify our derivations we first transform pertinent operators to the N-product form. We thus define the operators to which we shall apply the above developed perturbation theory as

$$\hat{K} = \hat{H} - \langle\Phi_o|\hat{H}|\Phi_o\rangle \equiv \hat{H} - \langle\hat{H}\rangle\,, \tag{9.67}$$

and

$$\hat{K}_o = \hat{H}_o - \langle\Phi_o|\hat{H}_o|\Phi_o\rangle \equiv \hat{H}_o - \langle H_o\rangle\,, \tag{9.68}$$

so that we have again (cf., Eqs. (1) and (3))

$$\hat{K}\,|\Psi_i\rangle = k_i\,|\Psi_i\rangle\,, \tag{9.69}$$

$$K_o\,|\Phi_i\rangle = \varkappa_i\,|\Phi_i\rangle\,, \tag{9.70}$$

where now

$$k_i = E_i - \langle H\rangle\,, \tag{9.71}$$

$$\varkappa_i = \varepsilon_i - \langle H_o\rangle = \varepsilon_i - \varepsilon_o\,, \tag{9.72}$$

since

$$\varepsilon_o = \langle\Phi_o|\hat{H}_o|\Phi_o\rangle \equiv \langle H_o\rangle = \sum_{Bi} \omega_{Bi}\,. \tag{9.73}$$

Since $|\Phi_o\rangle$ was chosen as a Fermi vacuum, the angular brackets $\langle\hat{O}\rangle$ designate the Fermi vacuum mean value of $\hat{O}$. Other unperturbed eigenstates $|\Phi_i\rangle$, $i \neq o$, may be conveniently expressed in terms of the Fermi vacuum state $|\Phi_o\rangle$ as mono-, bi-, etc. excited configurations. Labelling these states

accordingly we can write

$$|\Phi_K^{(k)}\rangle \equiv \left|\begin{matrix} B^1 & B^2 & \cdots & B^k \\ B_1 & B_2 & \cdots & B_k \end{matrix}\right\rangle = \left\{\prod_{i=1}^{k} (X_{B^i}^\dagger X_{B_i})\right\}|\Phi_0\rangle$$

$$= \hat{C}_K^{(k)}|\Phi_0\rangle \qquad (9.74)$$

where

$$\hat{C}_K^{(k)} = \prod_{i=1}^{k}(X_{B^i}^\dagger X_{B_i}) = N\left[\prod_{i=1}^{k}(X_{B^i}^\dagger X_{B_i})\right] \qquad (9.75)$$

and

$$K \equiv \{B^1 B^2 \cdots B^k \,;\, B_1 B_2 \cdots B_k\} \qquad (9.76)$$

is an ordered index set representing a given ordered configuration. The order of the operator products $(X_{B^i}^\dagger X_{B_i})$ is immaterial since they commute (cf., Eq. (8.10b)) and the excitation operator $\hat{C}_K^{(k)}$ is automatically in the N-product form. Using this notation we can re-write eigenvalue problem (70) as follows

$$\hat{K}_0|\Phi_K^{(k)}\rangle = \varkappa_K^{(k)}|\Phi_K^{(k)}\rangle, \qquad (0\leq k\leq N) \qquad (9.77)$$

where

$$|\Phi^{(0)}\rangle \equiv |\Phi_0\rangle, \qquad (9.78)$$

$|\Phi_K^{(k)}\rangle$ is given by (74), and

$$\varkappa_K^{(k)} = \mathcal{E}_K^{(k)} - \mathcal{E}_0 = \sum_{i=1}^{k}(\omega_{B^i} - \omega_{B_i}) \;;\; k\geqslant 1,$$
$$\varkappa^{(0)} \equiv \varkappa_0 = 0. \qquad (9.79)$$

The operator $\hat{P}_0$, Eq. (18), may be similarly written in the form

$$\hat{P}_0 = \sum_{k=1}^{N}\hat{P}_0^{(k)}, \qquad (9.80)$$

where

$$\hat{P}_0^{(k)} = \sum_K \frac{|\Phi_K^{(k)}\rangle\langle\Phi_K^{(k)}|}{\varkappa_0 - \varkappa_K^{(k)}} = \sum_K (-\varkappa_K^{(k)})^{-1}\hat{C}_K^{(k)}|\Phi_0\rangle\langle\Phi_0|\hat{C}_K^{(k)\dagger}, \qquad (9.81)$$

the sum extending over all k-times excited ordered configurations K, Eq. (76). We can also write $P_0^{(k)}$ in terms of unordered configurations, since

$$P_0^{(k)} = \sum_{\substack{B^1<\cdots<B^k \\ B_1<\cdots<B_k}} (-\varkappa_K^{(k)})^{-1}\left|\begin{matrix} B^1 & B^2 & \cdots & B^k \\ B_1 & B_2 & \cdots & B_k \end{matrix}\right\rangle\left\langle\begin{matrix} B^1 & B^2 & \cdots & B^k \\ B_1 & B_2 & \cdots & B_k \end{matrix}\right|$$
$$= (k!)^{-2}\sum_{\substack{B^1\cdots B^k \\ B_1\cdots B_k}} (-\varkappa_K^{(k)})^{-1}\left|\begin{matrix} B^1 & B^2 & \cdots & B^k \\ B_1 & B_2 & \cdots & B_k \end{matrix}\right\rangle\left\langle\begin{matrix} B^1 & B^2 & \cdots & B^k \\ B_1 & B_2 & \cdots & B_k \end{matrix}\right| \qquad (9.82)$$
$$= (k!)^{-2}\sum_{\substack{B^1\cdots B^k \\ B_1\cdots B_k}} (-\varkappa_K^{(k)})^{-1}\prod_{i=1}^{k}(Y_{B^i}^\dagger Y_{B_i}^\dagger)\,|\Phi_0\rangle\langle\Phi_0|\prod_{j=1}^{k}(Y_{B_i}Y_{B^i}),$$

where $\varkappa_K^{(k)}$ is the energy of the configuration $|\Phi_K^{(k)}\rangle$ given by Eq. (79).

Consider, finally, the perturbation $\hat{W}$,

$$\hat{W} = \hat{K} - \hat{K}_0 = \hat{H} - \hat{H}_0 - \langle\hat{H} - \hat{H}_0\rangle$$
$$= \hat{V} - \hat{U} - \langle\hat{V} - \hat{U}\rangle, \qquad (9.83)$$

which we shall also express in the N-product form. Recalling (cf., Sec. 4.5, Eq. (4.69)) that

$$\hat{V} - \langle\hat{V}\rangle = \hat{G}_N + \hat{V}_N, \qquad (9.84)$$

where $\hat{G}_N$ is defined in Eq. (4.70) and $\hat{V}_N$ designates the N-product form of $\hat{V}$, Eq. (4.57), and expressing one-particle operator $\hat{U}$ in the same way as a general operator $\hat{Z}$, Eq. (4.64),

$$\hat{U} - \langle\hat{U}\rangle = \hat{U}_N, \qquad (9.85)$$

we get

$$\hat{W} = \hat{G}_N - \hat{U}_N + \hat{V}_N = \hat{W}_1 + \hat{W}_2, \qquad (9.86)$$

where

$$\hat{W}_1 = \hat{G}_N - \hat{U}_N \equiv \hat{Q}_N \tag{9.87}$$

$$\hat{W}_2 = \hat{V}_N \tag{9.88}$$

and where we defined

$$\langle A|\hat{q}|\hat{B}\rangle \equiv \langle A|\hat{w}_2|B\rangle = \langle A|\hat{g}|B\rangle - \langle A|\hat{u}|B\rangle, \tag{9.89}$$

with $\hat{g}$ matrix element defined by Eq. (4.68). The perturbation $\hat{W}$ thus consists of a one- and two-particle parts $\hat{W}_1 \equiv \hat{Q}_N$ and $\hat{W}_2 \equiv \hat{V}_N$, respectively, both in the N-product form.

The form of the perturbation $\hat{W}$ as given by Eqs. (86)-(88) immediately suggests that we choose the one-particle potential $\hat{U}$, approximating the two-particle operator $\hat{V}$, to be the operator $\hat{G}$, Eq. (4.67), since then

$$\hat{W}_1 \equiv \hat{Q}_N = 0. \tag{9.90}$$

This choice significantly simplifies the perturbation theory expansion since it eliminates a large number of terms involving the one-particle part of the perturbation $\hat{W}$. However, in order to better understand this cancellation, and since $\hat{W}_1$ can also contain additional one-particle perturbations (as for example in double perturbation theories), we shall not eliminate the one-particle part from $\hat{W}$ at this stage.

The above outlined choice of the imperturbed and perturbed problems, Eqs. (67), (68) considerably simplifies the perturbation expansions since both the zero and the first order energy contributions vanish,

$$k_0^{(0)} = \varkappa_0 = 0, \tag{9.91a}$$

$$k_0^{(1)} = \langle W\rangle = 0, \tag{9.91b}$$

as follows immediately from Eqs. (79) and (83). (The latter also follows immediately from the fact that $\hat{W}$, Eq. (86), is in the N-product form.) Consequently, the renormalization terms will appear for the first time in the fourth order energy and the third order wavefunction terms. Thus, in the first few orders we get

$$k_0^{(2)} = \langle W P_0 W\rangle \tag{9.91c}$$

$$k_0^{(3)} = \langle W P_0 W P_0 W\rangle \tag{9.91d}$$

$$k_0^{(4)} = \langle W (P_0 W)^3\rangle - \langle W P_0 W\rangle \langle W P_0^2 W\rangle \tag{9.91e}$$

$$k_0^{(5)} = \langle W (P_0 W)^4\rangle - \langle (W P_0)^2 W\rangle \langle W P_0^2 W\rangle - \langle W P_0 W\rangle \left(\langle W P_0 (P_0 W)^2\rangle + \langle (W P_0)^2 P_0 W\rangle\right), \tag{9.91f}$$

and

$$|\Psi_0^{(0)}\rangle = |\Phi_0\rangle \quad \text{(intermediate normalization)} \tag{9.92a}$$

$$|\Psi_0^{(1)}\rangle = P_0 W |\Phi_0\rangle \tag{9.92b}$$

$$|\Psi_0^{(2)}\rangle = (P_0 W)^2 |\Phi_0\rangle \tag{9.92c}$$

$$|\Psi_0^{(3)}\rangle = \left[(P_0 W)^3 - \langle W P_0 W\rangle P_0^2 W\right] |\Phi_0\rangle \tag{9.92d}$$

$$|\Psi_0^{(4)}\rangle = \left[(P_0 W)^4 - \langle W P_0 W\rangle \{P_0 (P_0 W)^2 + P_0 W P_0^2 W\} - \langle W (P_0 W)^2\rangle P_0^2 W\right] |\Phi_0\rangle. \tag{9.92e}$$

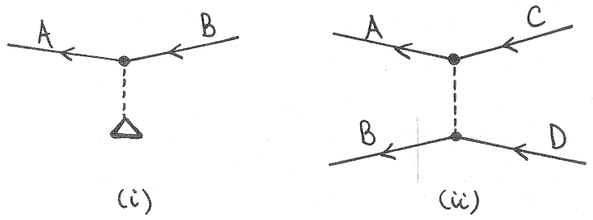
9.4. Diagrammatic representation

We now wish to evaluate the individual terms in perturbation theory expressions as given by the bracketing (or superbracketing) technique, e.g. Eqs. (91), (92). We thus have to consider the ordered operator products of the type

$$\prod_{i=1}^{n} (P_0^{m_i} W_k), \tag{9.93}$$

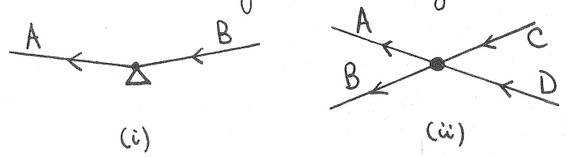
where $l_i = 1$ or $l_i = 2$ and $n_1 \geq 0$, $n_j \geq 1$ $(j>1)$, which act on $|\Phi_0\rangle$, $(n_1 > 0)$, or their Fermi vacuum mean values, $(n_1 = 0)$. In order to exploit the above outlined diagrammatic technique we first choose appropriate graphical representation for the basic operators P_0^k and W_e.

The perturbation operators $\hat{W}_i$ $(i=1,2)$ will be represented by the following Goldstone diagrams



(9.94)

or, if convenient, by their Hugenholtz equivalents



(9.95)

The scalar factors, weights and operators associated with these diagrams are summarized in Table I.

Table 9.I

Index set χ, scalar factor $d_G(\chi)$, operator $\hat{D}_G(\chi)$ and weight w_G associated with the diagrams G, Eqs. (94), (95).

G	χ	$d_G(\chi)$	$\hat{D}_G(\chi)$	w_G
(94 i)	{A; B}	$\langle A\|\hat{q}\|B\rangle$	$d_G N[X_A^\dagger X_B]$	1
(94 ii)	{A,B; C,D}	$\langle AB\|\hat{v}\|CD\rangle$	$d_G N[X_A^\dagger X_B^\dagger X_D X_C]$	$\frac{1}{2}$
(95 i)	{A; B}	$\langle A\|\hat{q}\|B\rangle$	$d_G N[X_A^\dagger X_B]$	1
(95 ii)	{A,B; C,D}	$\langle AB\|\hat{v}\|CD\rangle_{IA}$	$d_G N[X_A^\dagger X_B^\dagger X_D X_C]$	$\frac{1}{4}$

With these definitions the operators $\hat{W}_1 = \hat{Q}_N$ and $\hat{W}_2 = \hat{V}_N$ are given by the general formula, Eq. (5.66) or (5.16).

Consider now the "energy denominator" operator $\hat{P}_0$, Eq. (80), and its integer powers. Since the unperturbed configuration states (74) are orthonormal,

$$\langle \Phi_K^{(k)} | \Phi_L^{(l)} \rangle = \delta_{kl}\, \delta_{KL}\,, \tag{9.96}$$

where K, L are ordered configurations, we find easily from Eqs. (80), (81) that

$$\hat{P}_0^m = \sum_{k=1}^{N} [\hat{P}_0^{(k)}]^m, \tag{9.97}$$

and

$$[\hat{P}_0^{(k)}]^m = (k!)^{-2} \sum_K \hat{C}_K^{(k)} |\Phi_0\rangle\langle\Phi_0| \hat{C}_K^{(k)\dagger} (-\varepsilon_K^{(k)})^{-m}, \tag{9.98}$$

where $K = \{B^1, B^2, \cdots, B^k;\ B_1, B_2, \cdots, B_k\}$ is an unordered configuration and

$$\begin{aligned} \hat{C}_K^{(k)} &= Y_{B^1}^\dagger \cdots Y_{B^k}^\dagger\, Y_{B_k}^\dagger \cdots Y_{B_1}^\dagger \\ &= N[X_{B^1}^\dagger \cdots X_{B^k}^\dagger\, X_{B_k} \cdots X_{B_1}]. \end{aligned} \tag{9.99}$$

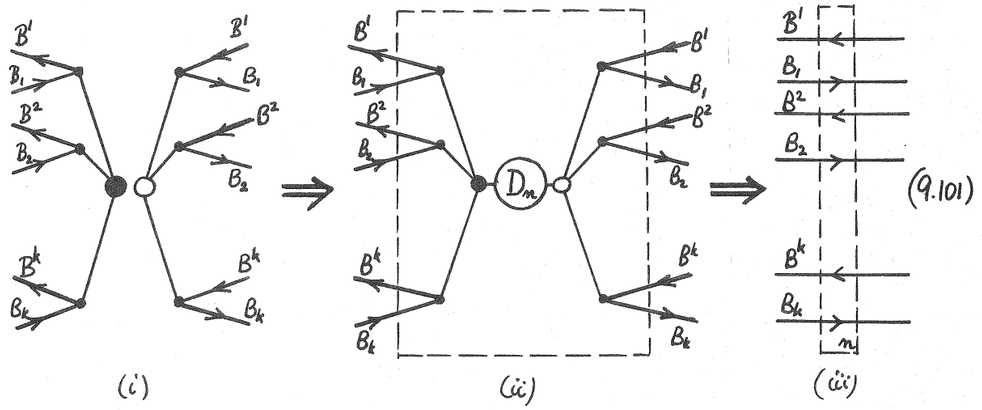
The operators (98) are given as linear combinations of projection operators,

$$|\Phi_K^{(k)}\rangle\langle\Phi_K^{(k)}| = \hat{C}_K^{(k)} |\Phi_0\rangle\langle\Phi_0| \hat{C}_K^{(k)\dagger} = \left|\begin{matrix} B^1 \cdots B^k \\ B_1 \cdots B_k \end{matrix}\right\rangle\left\langle\begin{matrix} B^1 \cdots B^k \\ B_1 \cdots B_k \end{matrix}\right|, \tag{9.100}$$

but they are not themselves idempotent due to the denominator factors multiplying individual projectors (100). In order that the phase of each factor is preserved the same ordering of the hole and the particle states must be used in both the ket and the bra parts. Further, the operators $\hat{C}_K^{(k)}$ and $\hat{C}_K^{(k)\dagger}$ must be represented by fermion lines extending to the left and to the right of incident vertices, respectively, as follows immediately

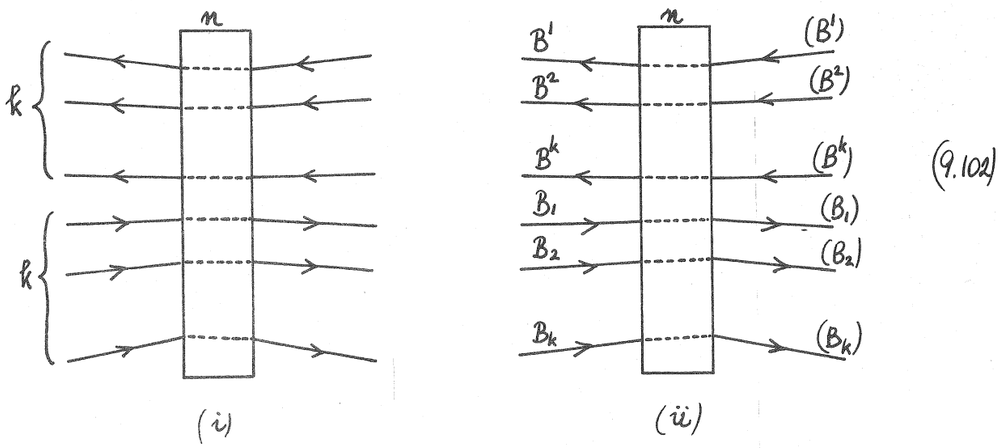
from the first Eq. (99), [cf., also, Sec. 5.4, Eqs. (75), (76)].

We can graphically represent operators (98) by joining together corresponding C and $\bar{C}$ diagrams, Eqs. (5.76), in such a way that the order of identically labeled lines is the same on either side and that it enables to determine the scalar factor $(-x_K^{(k)})^{-n}$ which we have to associate with it. This can be schematically illustrated as follows



(9.101)

We thus introduce the following D_m-skeletons and D_m-diagrams,



(9.102)

with which we associate the scalar factor

$$d_{D_m}(k;K) \equiv d_D = \left[\sum_{i=1}^{k} w_{B_i} - w_{B^i}\right]^{-n}, \quad k \geq 1, \qquad (9.103)$$

and the operator

$$\hat{D}_{D_m}(k;K) \equiv \hat{D}_D = d_{D_m}(k;K)\, \hat{C}_K^{(k)} |\Phi_0\rangle\langle\Phi_0| \hat{C}_K^{(k)\dagger} \qquad (9.104)$$

with K and $\hat{C}_K^{(k)}$ defined by Eqs. (76) and (75), respectively.

The rectangular vertex with label n indicates the scalar factor (103), while the dotted lines assure the same ordering of operators in the ket and in the bra of each projector. This also makes the second spin-orbital label on each line (which is enclosed in parentheses in diagram (102 ii)) superfluous. We will thus delete it in subsequent considerations. The permissible topological deformations are defined as transformations which arbitrarily permute the particle, and the hole lines, so that the weight of D_m-skeleton (102 i) is

$$w_{D_m}(k) \equiv w_D = (k!)^{-2}. \qquad (9.105)$$

With these definitions we have again

$$[P_0^{(k)}]^n = w_D(k) \sum_K \hat{D}_{D_m}(k;K). \qquad (9.106)$$

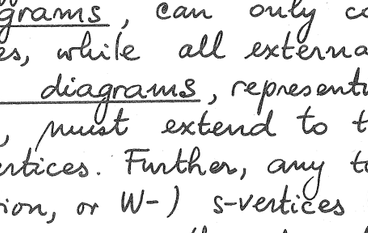
The reader have noticed that 2k vertices of C and $\bar{C}$ diagrams, diagram (101 i), have been replaced by 2k dotted lines inside the rectangular vertex. If desired, these dotted lines can be shrunk into simple vertices, so that D-diagram (102 ii) can be regarded as a Goldstone-type diagram. In view of its simple structure, we shall be able to replace it by a simple "denominator rule" and thus eliminate it entirely from our diagrams.

9.5. Perturbation theory diagrams

We now examine the structure of R-diagrams which we obtain when evaluating the perturbation expansion terms, such as given for the first few orders

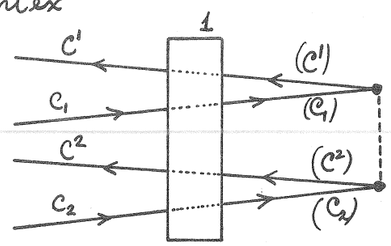
in Eqs. (91) and (92). We see immediately from our general rules (cf., Chapter 5) that the R-diagrams for the energy contributions, which we shall call energy diagrams, can only contain internal fermion lines, while all external lines in wavefunction diagrams, representing the wavefunction contributions, must extend to the left of their incident vertices. Further, any two interaction (or perturbation, or W-) s-vertices (94) or (95) will be separated by one n-th order denominator (or D_n-) s-vertex (102) and, in wavefunction Rdiagrams, there will also be one D_n-vertex to the left of the last (i.e., the leftmost) W-vertex. We also note that all spinorbital labels are free and all the operators are in the N-product form, so that all internal fermion lines must be incident with distinct s-vertices. Thus, in order to obtain the desired R-diagrams, we consider the relevant skeletons, form all permissible R-skeletons, and label all fermion lines with appropriate particle or hole summation labels.

A typical energy diagram obtained in the first non-vanishing order (i.e., the second order), Eq. (91c), has the form



(9.107)

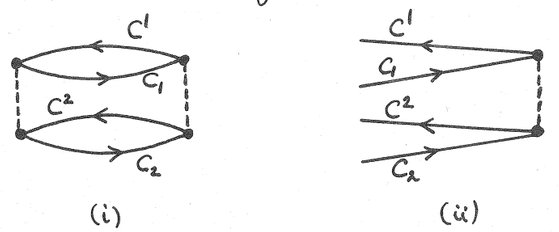
while a corresponding first order wavefunction diagram, Eq. (92b), is obtained by removing the leftmost interaction vertex



(9.108)

In both cases, only the $P_0^{(2)}$ component of $\hat{P}_0$, Eq. (80), gives a non-vanishing R-diagram.

These simple examples already suggest that we can in fact omit the D-vertices from our R-diagrams, and interconnect identically labeled fermion lines, carrying the same spinorbital label, and connected by dotted lines in the D-vertex. Using a single spin-orbital label for each line, we obtain simpler diagrams



(9.109)

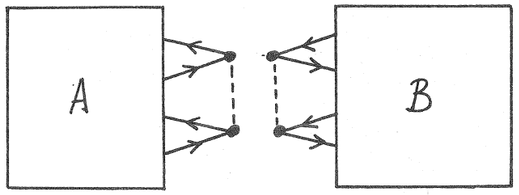
when we simultaneously introduce the following denominator convention:

With each pair of neighboring interaction vertices (and with the fermion lines extending to the left of the leftmost interaction vertex in the wavefunction diagrams) we associate an energy denominator (103), which we obtain by assigning the orbital energies ω_A to each fermion line carrying the spin-orbital label A and crossing a hypothetical vertical line separating these neighboring vertices (and an additional line to the left of all vertices in wavefunction diagrams), and taking those associated with the hole lines with a positive sign and those associated with the particle lines with a negative sign. In the renormalization terms we have to remember the power of each denominator operator $\hat{P}_0$ and to raise the denominator to the same power.

It is clearly irrelevant that the individual fermion lines may be arbitrarily distorted and intersect a hypothetical vertical line more than once (but, of course, an odd number of

times) since any multiple contributions must cancel out. It is also easily seen that the numbers of particle and of hole lines, which are crossing any hypothetical vertical line, are equal. In other words, the number of particle lines equals the number of hole lines at any instant.

We also note that our diagrammatic representation will automatically pick up appropriate "denominator vertices" from the full sums in Eqs. (80) and (81), and will automatically eliminate the ground state projector $|\Phi_0\rangle\langle\Phi_0|$ from the sum in Eq. (81) [or Eq. (18), $i=0$] if we exclude all R-diagrams with "dangerous denominators", i.e. all disconnected R-diagrams of the type

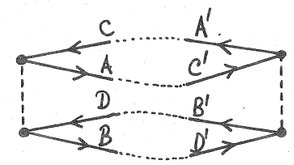


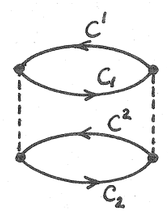
(9.110)

which would yield a vanishing denominator following the above given rule. This does not mean, of course, that legitimate R-diagrams must be connected: we must only eliminate those disconnected diagrams which would yield vanishing ("dangerous") denominators.

The reader may find it instructive to work out several examples algebraically, applying generalized Wick's theorem and representing graphically each system of pairings. For example, in the case of the second order energy contribution involving two V vertices (operators), we can form four non-vanishing contraction schemes and represent them diagrammatically as shown below. We note that we can ignore the projection operators $(\hat{P}_0)^k$ in an algebraic derivation as well, replacing each contraction pair $\overbrace{M_1 (Y_A^\dagger}\ \overbrace{Y_A)}\ M_2$, where $A \in K$, Eq. (76), with a single contraction $\overbrace{M_1\ M_2}$ of the operators originating from different $\hat{W}_i$ operators. We thus find the following contraction schemes and corresponding R-diagrams:

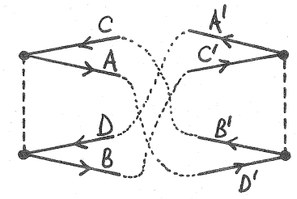
(i) $N[\overbrace{X_A^\dagger X_B^\dagger X_D X_C}\ \overbrace{X_{A'}^\dagger X_{B'}^\dagger X_{D'} X_{C'}^\dagger}]$





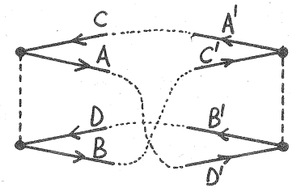
(i)

(ii) $N[\overbrace{X_A^\dagger X_B^\dagger X_D X_C}\ \overbrace{X_{A'}^\dagger X_{B'}^\dagger X_{D'} X_{C'}^\dagger}]$

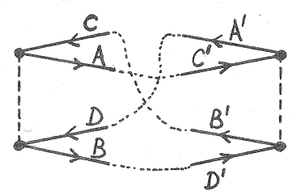


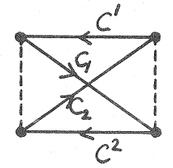
(9.111)

(iii) $N[\overbrace{X_A^\dagger X_B^\dagger X_D X_C}\ \overbrace{X_{A'}^\dagger X_{B'}^\dagger X_{D'} X_{C'}^\dagger}]$



(iv) $N[\overbrace{X_A^\dagger X_B^\dagger X_D X_C}\ \overbrace{X_{A'}^\dagger X_{B'}^\dagger X_{D'} X_{C'}^\dagger}]$





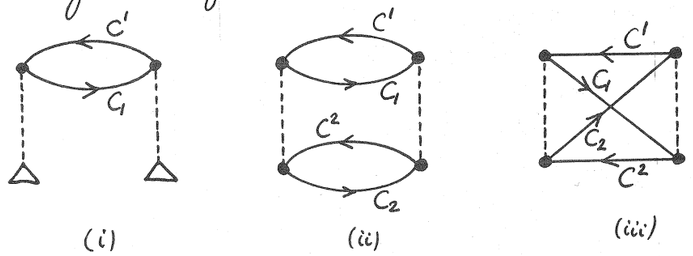
(ii)

Since $W_V = \frac{1}{2}$ and also $W_R = \frac{1}{2}$ in each case, we see that according to theorem T3 we have $N = \frac{1}{2}(\frac{1}{2}\cdot\frac{1}{2})^{-1} = 2$,

so that each R diagram represents two contraction schemes as we have just seen.

9.5.1. Second and third order energy contributions

Applying our diagrammatic rules (Ch. 5) to the second order case we obtain immediately the following R-diagrams



(i) (ii) (iii) (9.112)

which give the following algebraic contributions

$$k_0^{(2)\prime} = \sum_{C'C_0} \langle C_1|\hat{q}|C'\rangle\langle C'|\hat{q}|C_1\rangle\, \Delta'(C_1;C') \tag{9.113}$$

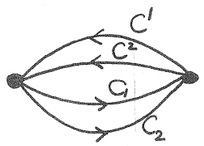
$$k_0^{(2)\prime\prime} = \tfrac{1}{2}\sum_{C'C^2C_1C_2} \langle C_1C_2|\hat{v}|C'C^2\rangle\left(\langle C'C^2|\hat{v}|C_1C_2\rangle - \langle C'C^2|\hat{v}|C_2C_1\rangle\right)\Delta'(C_1C_2;C'C^2),$$

$$k_0^{(2)} = k_0^{(2)\prime} + k_0^{(2)\prime\prime},$$

where we defined a general denominator as

$$\Delta^n(C_1\,C_2\,\dots\,C_p\,;C'\,C^2\,\dots\,C^l) = \left[\sum_{i=1}^{p}(\omega_{C_i} - \omega_{C^i})\right]^{-n}. \tag{9.114}$$

Note that both diagrams (ii) and (iii), Eq. (112), correspond to the same Hugenholtz diagram



(9.115)

while, of course, the Hugenholtz version of diagram (112 i) is essentially the same (except for the disappearance of dashed lines). The weight of diagram (115) is clearly $w_{\bar{D}_2} = 1/4$. Let us illustrate in this

case that using one Goldstone representative of the Hugenholtz diagram (115), i.e. the so called Brandow diagrams (cf., Sec. 5.7), we obtain the same result given by Eq. (113). Since this is obviously the case for the one-particle contribution $k_0^{(2)\prime}$, we only examine the two-particle term $k_0^{(2)\prime\prime}$. Consider, first, diagram (112 ii) as a Brandow diagram, so that

$$k_0^{(2)\prime\prime} = \tfrac{1}{4}\sum_{C'C^2C_1C_2} \langle C_1C_2|\hat{v}|C'C^2\rangle_{\bar{D}_2}\langle C'C^2|\hat{v}|C_1C_2\rangle_{\bar{D}_2}\,\Delta'(C_1C_2;C'C^2)$$

$$= \tfrac{1}{4}\sum_{C'C^2C_1C_2}\left(\langle C_1C_2|\hat{v}|C'C^2\rangle - \langle C_1C_2|\hat{v}|C^2C'\rangle\right)$$

$$\times\left(\langle C'C^2|\hat{v}|C_1C_2\rangle - \langle C'C^2|\hat{v}|C_2C_1\rangle\right)\Delta'(C_1C_2;C'C^2)$$

$$= \tfrac{1}{2}\sum_{C'C^2C_1C_2} \langle C_1C_2|\hat{v}|C'C^2\rangle\left(\langle C'C^2|\hat{v}|C_1C_2\rangle - \langle C'C^2|\hat{v}|C_2C_1\rangle\right)$$

$$\times\,\Delta'(C_1C_2;C'C^2), \tag{9.116}$$

which agrees with Eq. (113). Second, using diagram (112 iii) as a Brandow diagram, we get

$$k_0^{(2)\prime\prime} = -\tfrac{1}{4}\sum_{CC^2C_1C_2} \langle C_1C_2|\hat{v}|C'C^2\rangle_{\bar{D}_2}\langle C'C^2|\hat{v}|C_2C_1\rangle_{\bar{D}_2}\,\Delta'(C_1C_2;C'C^2), \tag{9.117}$$

which again is identical with Eq. (113) since

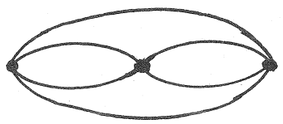
$$\langle C'C^2|\hat{v}|C_1C_2\rangle_{\bar{D}_2} = -\langle C'C^2|\hat{v}|C_2C_1\rangle_{\bar{D}_2}. \tag{9.118}$$

In deriving the third order contribution, it is already very useful to exploit Hugenholtz diagrams, either as an intermediate step or evaluating them directly using Brandow diagrams. Consider, first, the case that all three operators $\hat{W}_i$ are two-particle operators. Let us ignore initially the orientation of fermion lines, so that we have to consider R-diagrams formed from three Hugenholtz type vertices, i.e. schematically



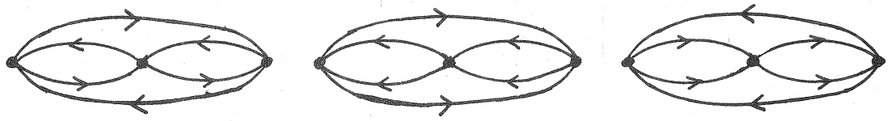
(9.119)

Since we must not connect lines incident with the same vertex we can proceed systematically as follows: Connect, first, vertices (I) and (II) with $k = 0, 1, 2, 3$ or 4 lines and, second, try to connect the remaining lines in all possible ways. The problem is clearly symmetric so that $k=6$ and 1 cases are obtained from $k=3,4$ cases by a reflection (I) $\leftrightarrow$ (III). Further, $k=0$ ($k=4$) case must be disregarded since it would require joining lines belonging to the same vertex (and would result in a diagram with a dangerous denominator). Also the case $k=1$ ($k=3$) cannot be realized without joining lines incident with the same vertex so that only $k=2$ case yields a possible prototype of the third order two-particle Hugenholtz energy diagram



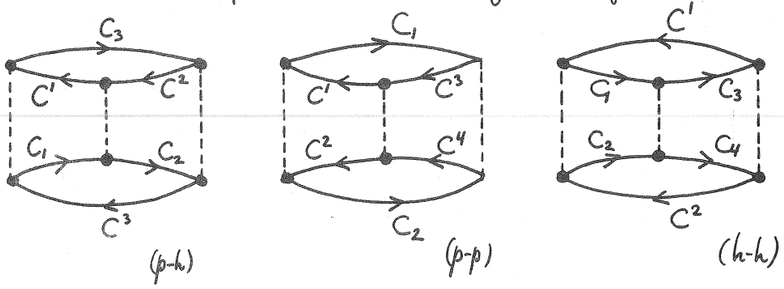
(9.120)

Considering, next, all possible orientations of fermion lines, we find the following three Hugenholtz diagrams



(i) (ii) (iii) (9.121)

According to the interaction on the central vertex we shall refer to these diagrams as (i) particle-hole (p-h), (ii) particle-particle (p-p) and (iii) hole-hole (h-h) diagrams. As corresponding Brandow diagrams we can choose the following Goldstone representatives of diagrams (121):



(9.123)

These diagrams yield the following algebraic contributions:

$$k^{(2)}_{ph} = -\sum_{C'C^2C^3}\sum_{C_1C_2C_3} \langle C_1C_3|\hat{v}|C^3C'\rangle_{as}\langle C'C_2|\hat{v}|C_1C^2\rangle_{as}\langle C^2C^3|\hat{v}|C_3C_2\rangle_{as}\,\Delta_{ph}$$

$$k^{(3)}_{pp} = \frac{1}{8}\sum_{C'C^2C^3C^4}\sum_{C_1C_2}\langle C_1C_2|\hat{v}|C'C^2\rangle_{as}\langle C'C^2|\hat{v}|C^3C^4\rangle_{as}\langle C^3C^4|\hat{v}|C_1C_2\rangle_{as}\,\Delta_{pp} \qquad (9.124)$$

$$k^{(3)}_{hh} = \frac{1}{8}\sum_{C'C^2}\sum_{C_1C_2C_3C_4}\langle C_1C_2|\hat{v}|C'C^2\rangle_{as}\langle C_3C_4|\hat{v}|C_1C_2\rangle_{as}\langle C'C^2|\hat{v}|C_3C_4\rangle_{as}\,\Delta_{hh}$$

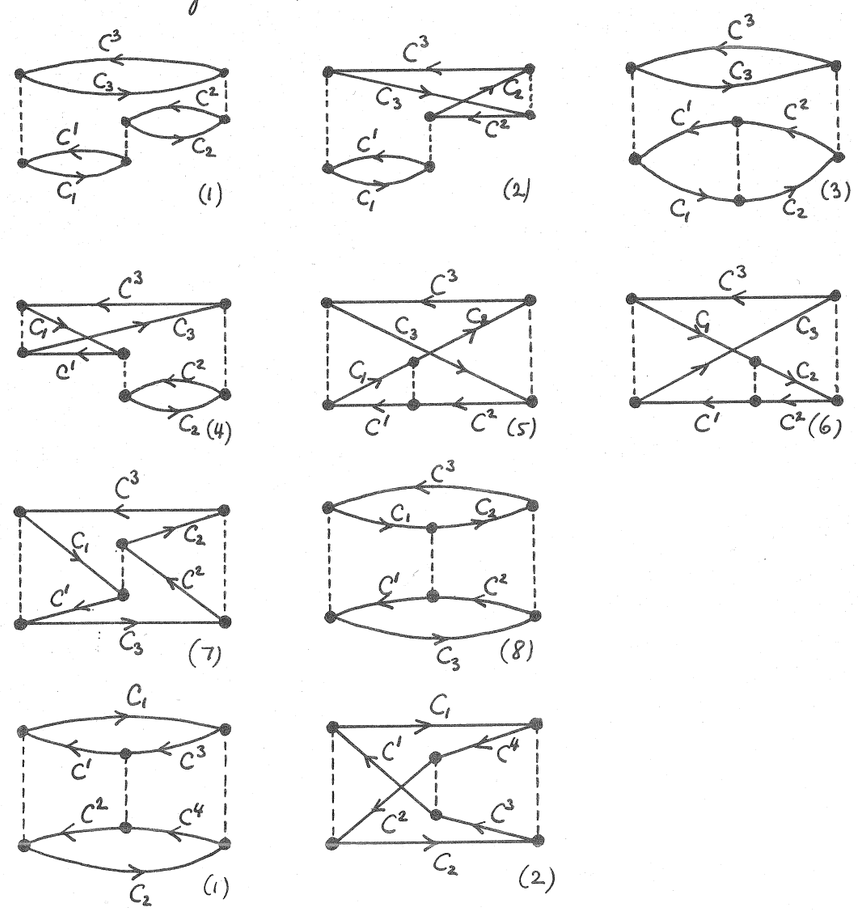
where

$$\Delta_{ph} = \Delta'(C_1C_3;\,C'C^3)\;\Delta'(C_2C_3;\,C^2C^3)$$

$$\Delta_{pp} = \Delta'(C_1C_2;\,C'C^2)\;\Delta'(C_1C_2;\,C^3C^4) \qquad (9.125)$$

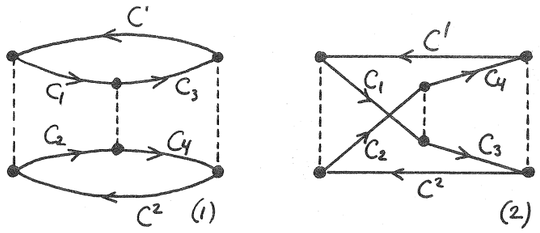
$$\Delta_{hh} = \Delta'(C_1C_2;\,C'C^2)\;\Delta'(C_3C_4;\,C'C^2).$$

The reader should convince himself that an identical result is obtained using Goldstone diagrams, which are given below:



(9.126a)

(9.126b)



(9.126c)

The p-h diagram (121 i) thus represents eight Goldstone diagrams (126a) while each pp, (121 ii) and h-h, (121 iii) diagrams represent only two Goldstone diagrams given in (126 b) and (126 c), respectively. In a general n-th order case each Hugenholtz two-particle perturbation theory energy diagram represents up to 2^n Goldstone diagrams. A convenient way to obtain all Goldstone diagrams is to draw the Goldstone diagram with a maximum number of closed loops and to carry out an exchange on each vertex independently. For the third order case considered these operations are summarized in Table II, where we also list the Goldstone weights, the number of closed loops l and the number of (internal) hole lines h. Designating relevant matrix elements as follows:

$$A = \langle C^2C^3|\hat{v}|C_2C_3\rangle \qquad \tilde{A} = \langle C^2C^3|\hat{v}|C_3C_2\rangle$$
$$B = \langle C'C_2|\hat{v}|C_1C^2\rangle \qquad \tilde{B} = \langle C'C_2|\hat{v}|C^2C_1\rangle$$
$$C = \langle C_1C_3|\hat{v}|C'C^3\rangle \qquad \tilde{C} = \langle C_1C_3|\hat{v}|C^3C'\rangle, \qquad (9.127a)$$

$$A' = \langle C^3C^4|\hat{v}|C_1C_2\rangle \qquad \tilde{A}' = \langle C^3C^4|\hat{v}|C_2C_1\rangle$$
$$B' = \langle C'C^2|\hat{v}|C^3C^4\rangle$$
$$C' = \langle C_1C_2|\hat{v}|C'C^2\rangle, \qquad (9.127b)$$

and

$$A'' = \langle C'C^2|\hat{v}|C_3C_4\rangle \qquad \tilde{A}'' = \langle C'C^2|\hat{v}|C_4C_3\rangle$$
$$B'' = \langle C_3C_4|\hat{v}|C_1C_2\rangle$$
$$C'' = \langle C_1C_2|\hat{v}|C'C^2\rangle = C', \qquad (9.127c)$$

where unprimed, singly and doubly primed matrix elements correspond to the p-h, pp and h-h diagrams and tilda designates corresponding exchange matrix element, and using the same denominators, Eqs. (125), as in the Hugenholtz case, we also give a contribution of each Goldstone diagram in the last column of Table II. The total contribution is then obtained by summing over all spin-orbital labels.

Table 9.II.
Contributions from the third-order two-particle Goldstone energy diagrams.

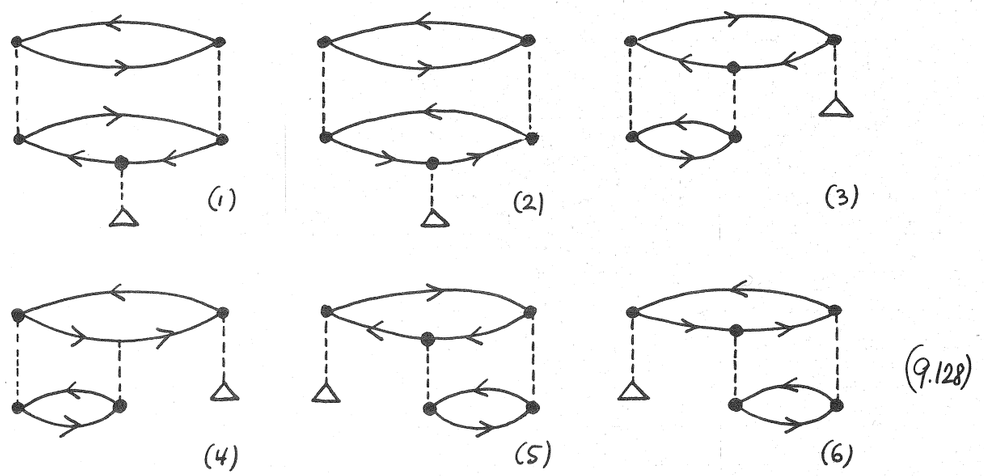
Hugenholtz diagram Eq.(121)	Goldstone diagram Eq.(126)	Exchange at vertices [cf. Eq. (119)]	W_G	l	h	Algebraic contribution (sums implied)
(i)	a1	–	1	3	3	$ABC\,\Delta_{ph}$
	a2	I	1	2	3	$-\tilde{A}BC\,\Delta_{ph}$
	a3	II	1	2	3	$-A\tilde{B}C\,\Delta_{ph}$
	a4	III	1	2	3	$-AB\tilde{C}\,\Delta_{ph}$
	a5	I, II	1	1	3	$\tilde{A}\tilde{B}C\,\Delta_{ph}$
	a6	II, III	1	1	3	$A\tilde{B}\tilde{C}\,\Delta_{ph}$
	a7	I, III	1	1	3	$\tilde{A}B\tilde{C}\,\Delta_{ph}$
	a8	I, II, III	1	2	3	$-\tilde{A}\tilde{B}\tilde{C}\,\Delta_{ph}$
(ii)	b1	–; I, II; II, III; I, III	½	2	2	$A'B'C'\,\Delta_{pp}$
	b2	I; II; III; I, II, III	½	1	2	$-\tilde{A}'B'C'\,\Delta_{pp}$
(iii)	c1	–; I, II; II, III; I, III	½	2	4	$A''B''C''\,\Delta_{hh}$
	c2	I; II; III; I, II, III	½	1	4	$-\tilde{A}''B''C''\,\Delta_{hh}$

Thus, in the p-p and h-h cases, there are only two distinct terms out of the total of eight terms contained in expressions given by the last two equations, Eqs. (124). Consequently, the factor 1/8 in Eq.(124), given by the Hugenholtz diagram weight, reduces to the factor ½, given by the Goldstone weight.

This example nicely illustrates the advantages of Brandow-type formalism. However, we shall see in Section 9.12 that Goldstone diagrams are needed if a simple elimination of spin dependence (for spin-free interactions) is to be carried out.

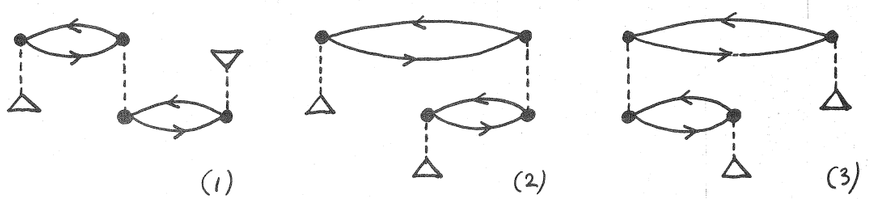
For completeness, we now list Brandow

diagrams for the remaining third-order terms which involve one, two or three one-particle operators (we drop for simplicity line labels, i.e., present pertinent skeletons).
Diagrams with one Q vertex:



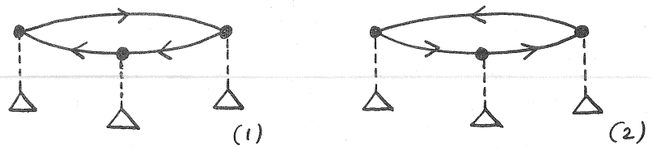
(9.128)

Corresponding Goldstone diagrams are obtained by a single exchange on one of the two-particle vertices, so that there are twelve Goldstone diagrams of this type.
We further present:
Diagrams with two Q vertices:



(9.129)

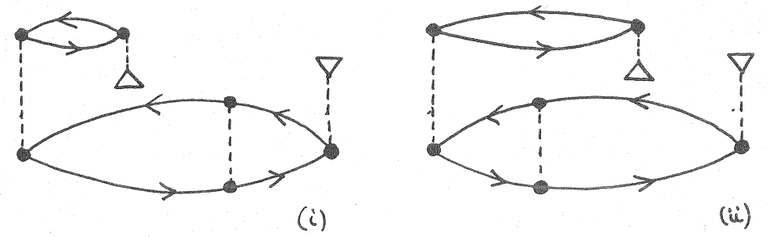
which again yield six Goldstone diagrams by an exchange on the two-particle vertex.
Finally, we have two
Diagrams with three Q vertices:



(9.130)

It should be noted that in fact all the diagrams in each group, Eqs. (128), (129), (130), are closely related and could be obtained one from the other if we would allow the permutation of vertices along the horizontal time axis. Such groupings of diagrams, which differ only by the permutation of their vertices, follow naturally when using the time-dependent formalism of Goldstone. The diagrams which may be transformed one into the other by topological deformations (transformations) which do not preserve the order of operators along the time axis are referred to as <u>time versions</u> of the same diagrams. Later, we shall distinguish time versions of two kinds:

Time versions of the <u>first kind</u> may be obtained one from another by the permutation of vertices (or s-vertices), which do not change the particle-hole character of any fermion line in the diagram. For example, the following diagrams are time versions of the first kind of a fourth order energy diagram with two Q vertices:

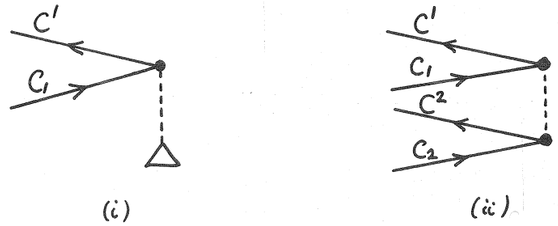


(9.131)

When the vertex permutation changes the hole-particle character of at least one line, we obtain different time versions of the <u>second kind</u>. Thus, all diagrams (128) are different time versions of the second kind of the first diagram (1), Eq. (128), and similarly for diagrams (129) and (130). The reader can also identify different time versions among the two-particle diagrams (126).

9.5.2. First and second order wavefunction contributions

In the first order, Eq. (92b), we have only two Goldstone (or Hugenholtz) diagrams, namely



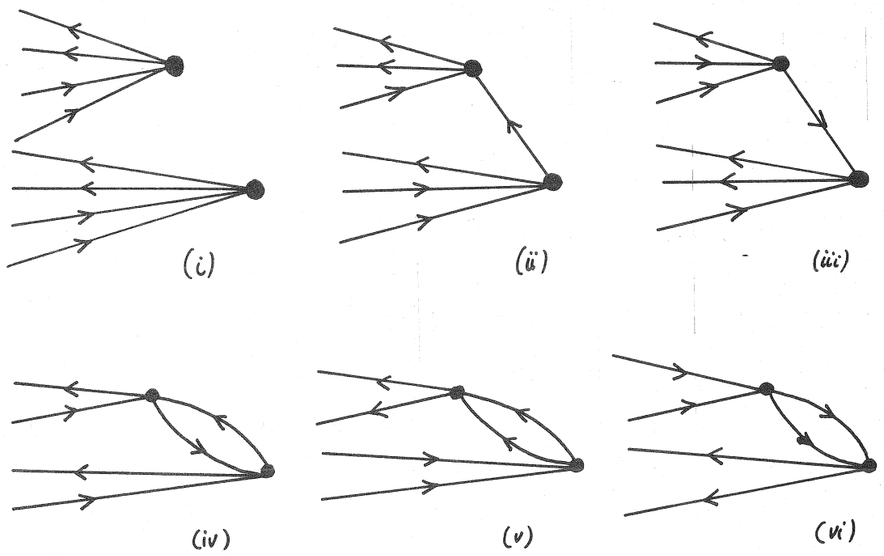
(i) (ii)

(9.132)

which give

$$|\Psi_0^{(1)}\rangle = \sum_{C'C_1} \langle C'|\hat{q}|C_1\rangle \Delta'(C_1;C')\, N[X_{C'}^{\dagger} X_{C_1}]\,|\Phi_0\rangle + \frac{1}{2}\sum_{C'C^2C_2} \langle C'C^2|\hat{v}|C_1C_2\rangle \Delta'(C_1C_2;C'C^2)\, N[X_{C'}^{\dagger} X_{C_1} X_{C^2}^{\dagger} X_{C_2}]\,|\Phi_0\rangle. \tag{9.133}$$

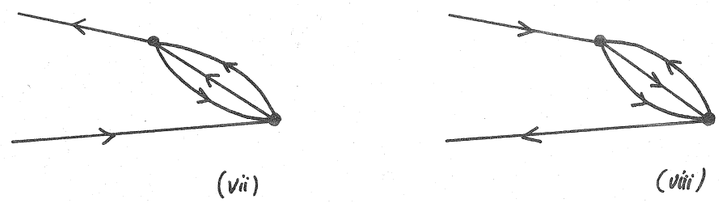
In the second order, Eq. (92c), we can have either disconnected or connected diagrams. We present below all possible Hugenholtz diagrams for the second order wavefunction contribution:



(i) (ii) (iii)

(iv) (v) (vi)

(9.134)



(vii) (viii)

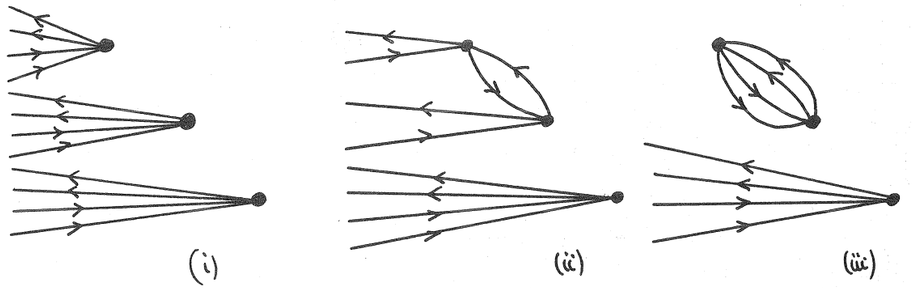
Thus, we have one disconnected diagram (i), which yields a quadruply excited contribution, while the remaining diagrams are connected and correspond to triply (ii), (iii), doubly (iv), (v), (vi) and singly (vii), (viii) excited contributions. No contribution to the vacuum can arise, since any diagram having only internal lines (i.e., energy diagram) would lead to a dangerous denominator. The reader should write down corresponding Brandow and Goldstone diagrams and carry out their evaluation as an exercise. We also observe that these diagrams not only yield the third order energy diagrams if the interaction vertex on the left hand side is added and permissibly energy diagrams formed but that in fact we can obtain the fourth order energy contribution when calculating a mean value of $\hat{W}$ in $|\Psi_0^{(2)}\rangle$ state.

We also note that with our choice of the unperturbed and perturbed Hamiltonians, no renormalization terms arise in the orders considered above. We shall examine the role of these terms in detail in subsequent sections.

9.6. Linked and unlinked diagrams

We have seen in preceding section that for wavefunction contribution we obtain disconnected diagrams already in the second order. In higher orders of perturbation theory disconnected diagrams of another type will occur. For example, a few possible third order wavefunction diagrams,

which are disconnected, are shown below.



(9.135)

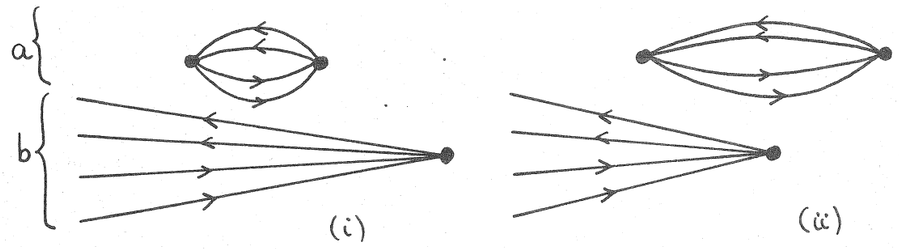
Even though all these diagrams are bona-fide wave-function diagrams (i.e., no dangerous denominators, all external lines to the left of incident vertices, etc.) we shall see that the latter diagram (iii) has a very different character than the former two diagrams (i) and (ii), since it contains an energy diagram as a disconnected part. Notice, however, that due to our denominator rule, the contribution from this diagram does not simply factor into a product of contributions of disconnected parts.

We shall refer to energy diagrams, which have no external lines, as <u>vacuum diagrams</u> (or <u>vacuum parts</u> when they form a disconnected part of some diagram), since they represent Fermi vacuum mean values. Further, a disconnected diagram is <u>unlinked</u> if it has at least one disconnected vacuum part, and <u>linked</u>, if it has no disconnected vacuum part. Thus, diagrams (135 i) and (135 ii) are linked while the diagram (135 iii) is unlinked.

Any unlinked diagram is by definition a disconnected diagram, while a linked diagram can be either connected or disconnected. In the latter case, however, none of its disconnected parts can be a vacuum diagram. On the other hand, a connected diagram is always linked (even if it is a vacuum diagram), while a disconnected diagram can be either linked or unlinked, depending on whether all of its disconnected parts are of a non-vacuum

type, or not, respectively.

We shall now examine the role of unlinked diagrams. Obviously, each unlinked diagram has a number of time versions of the first kind, since its disconnected parts can be positioned relative one another in all distinct ways which do not introduce dangerous denominators. Thus, in the case of diagram (135 iii), there are two possible time versions



(9.136)

The contributions from either of these timeversions differ only in the denominator part since all the scalar factors associated with s-vertices and all the operators associated with external lines are clearly identical. Designating the denominator of the vacuum part, considered as a separate diagram, by $\underline{a}$ and, similarly, the denominator of the part involving external lines (considered separately) as $\underline{b}$, as indicated schematically on the left hand side of Eqs. (136), the contribution from both time versions is

$$N \frac{1}{b\,(a+b)\,b} + N \frac{1}{a\,(a+b)\,b}\,, \qquad (9.137)$$

where N designates the identical numerator part (given by the product of three v-matrix elements, N-product of four $Y_A^\dagger$ operators, sign and weight factors). Carrying out the sum in Eq. (137) we get

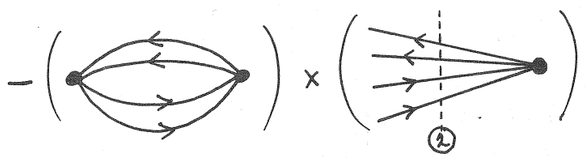
$$N \frac{1}{b(a+b)}\left(\frac{1}{b}+\frac{1}{a}\right) = N \frac{a+b}{b(a+b)ab} = N \frac{1}{ab^2}\,. \qquad (9.138)$$

The result (138) is easily seen to be precisely the

contribution, except for the sign, from the third order renormalization term, (cf. Eq. (92d)),

$$-\langle W P_0 W\rangle\, P_0^2 W\, |\Phi_0\rangle\,, \tag{9.139}$$

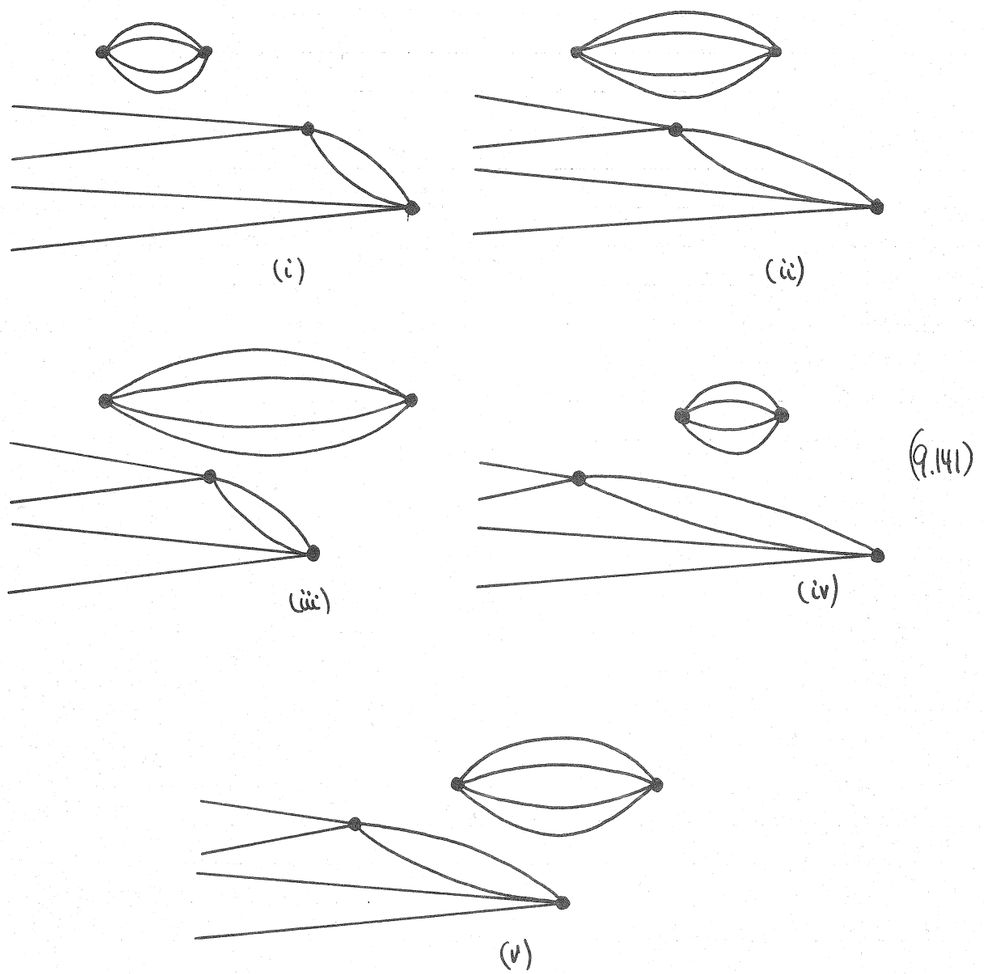
which is given, up to the sign, by the product of the second order energy contribution and the first order wavefunction contribution taken with the second order denominator vertex (which we indicate by the dashed line with index (2) below)



(9.140)

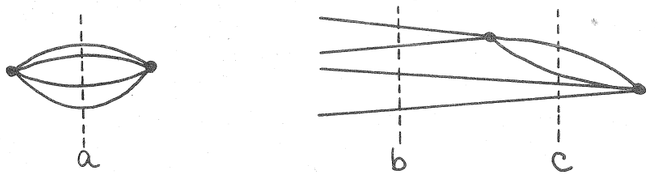
and thus equals $-N\frac{1}{a}\cdot\frac{b}{a^2}$. Therefore, the renormalization term (139) exactly cancells the contribution from unlinked diagrams (136) originating from the principal third order term, given by Eq. (138). In a completely analogous way we can show the cancellation of unlinked terms involving one-particle W_1 vertices. We can therefore conclude that, at least in the first three orders of perturbation theory, the wavefunction is given by all possible linked diagrams.

For higher than the third orders, the direct verification is possible but it becomes more and more laborious since different types of unlinked diagrams appear and each type may involve a large number of time versions. We recommend to the reader that he verifies this cancellation in the fourth order of the wavefunction contribution in full detail. For example, one of the unlinked diagrams with the second order two-particle vacuum part, which originate from the fourth order principal term, will occur in the following five time versions:
(we drop line orientation for simplicity)



(9.141)

Designating the denominator of the vacuum part, taken as a separate diagram, by $\underline{a}$, and the denominators of the second order operator part by $\underline{b}$ and $\underline{c}$, as schematically shown below:



(9.142)

we get for the contribution from the five time versions (141) the expression $N.D$, where N designates the numerator part which is the same for all the time versions and D the denominator part given by the expression

$$D = \frac{1}{b(a+b)bc} + \frac{1}{b(a+b)(a+c)c} + \frac{1}{b(a+b)(a+c)a}$$

$$+ \frac{1}{bc(a+c)c} + \frac{1}{bc(a+c)a}$$

$$= \frac{1}{b(a+b)bc} + \frac{1}{b(a+b)ac} + \frac{1}{bc \cdot ac}$$

$$= \frac{1}{ab^2c} + \frac{1}{abc^2} . \qquad (9.143)$$

This contribution is again easily seen to be cancelled by the appropriate terms originating from the renormalization terms, (cf. Eq. (92e)),

$$-\langle W P_0 W \rangle \left(P_0^2 W P_0 W + P_0 W P_0^2 W \right) | \Phi_0 \rangle . \qquad (9.144)$$

It is thus plausible to postulate that the contributions of all the renormalization terms, in every order of perturbation theory, are cancelled by the contributions from unlinked diagrams, which originate from the principal term. Consequently, the wavefunction is given in every order of perturbation theory by linked diagrams. These diagrams, which of course may be disconnected, can be divided into disjoint classes characterized by their disconnected parts. Thus, all members of a given class have the same external lines and the same interaction vertices interconnected in the same way and only differ in a relative ordering of vertices. In other words, they consist of all permissible (i.e., no dangerous denominators) time versions of the first kind formed from a given set of disconnected diagrams. For example, (i) and (ii) of (135) represent two such classes in the third order.

We shall verify this hypothesis concerning a general structure of the wavefunction in an

arbitrary order of perturbation theory after we prove general factorization properties for these linked diagrams in Sec. 8. However, before doing so, we wish to point out the importance and the implications of unrestricted summations in our derivations.

9.7. Exclusion Principle Violating Diagrams

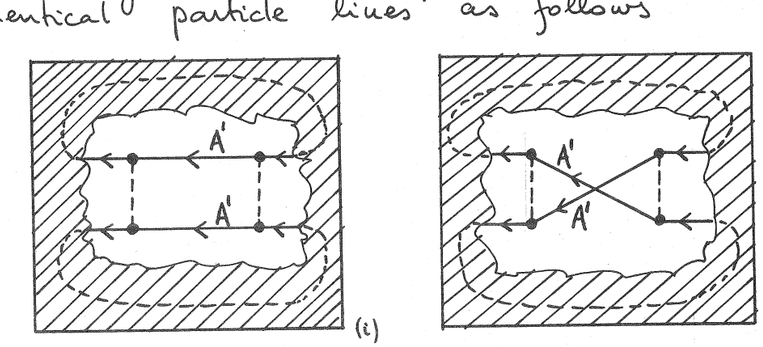
In order to obtain simple rules for the cancellation of renormalization and unlinked terms or for the factorization of disconnected diagram contributions (see the next section) it is essential that the summations are not restricted in the factorized form.

As we have seen earlier, it is irrelevant whether we restrict the summations or not since the added terms vanish anyway. For example, in the projection (or denominator) operator representation, Eq. (82), all the terms in which either $B^i = B^i$ or $B_i = B_j$ $(i,j = 1, \cdots, k)$ vanish in view of the antisymmetry properties, Eqs. (4.15), or Eqs. (1.24), (1.85), etc. For the same reason all the terms in any operator expression, Eq. (3.60), which have two identical spin-orbital labels in the bra or in the ket will vanish. Consequently, it makes no difference whether we eliminate these terms from our expressions by imposing appropriate restrictions on the summation symbols or not.

However, once we go over to more complex quantities we must exercise more care. Of course, when we keep all the terms, which result from any derivation, it is again irrelevant whether these terms, originating from the vanishing terms in basic operators, are included or not. Even though these terms will no longer necessarily vanish (as in various fully contracted terms) they will nevertheless mutually cancel out. We thus

have to be particularly carefull with these terms once we allow various partial cancellation of different terms as, for example, the cancellation of unlinked diagrams with renormalization terms contributions. Since these terms originate from vanishing components of operators or states in which the Pauli or exclusion principle is violated (i.e., a given spin-orbital state is occupied more than once) they are generally refered to as the <u>exclusion principle violating (EPV) terms.</u> In the diagrammatical form these terms are easily recognized as diagrams which have two identically labeled hole or particle lines at a given time.

Generally, the contributions from EPV diagrams cancel out. This may be schematically illustrated for the case of two identical particle lines as follows

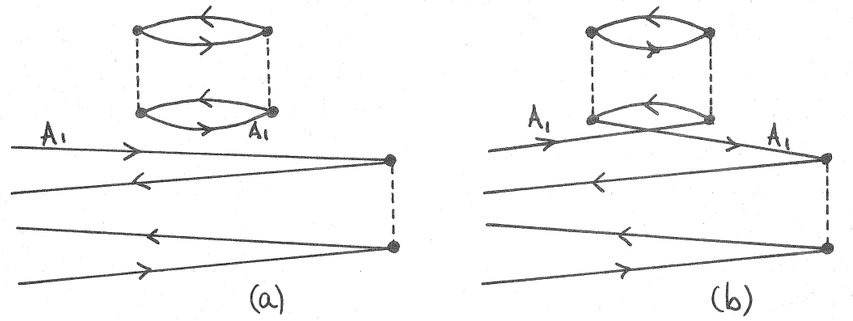


(9.145)

where the dashed lines, which are continuations of fermion lines, represent schematically the paths of oriented lines. (Note that a similar demonstration can be carried out when one or both paths represented by dashed lines are open.) Contributions from both diagrams (i) and (ii) of (145) are clearly identical except for the sign since the number of closed loops in (i) and (ii) differs by one. Thus, both contributions mutually cancel.

It may happen, however, that one of the diagrams which mutually cancel, such as (i) and (ii) of (145), is linked while the other

one is not. Consider, for example, one Goldstone version of the EPV diagram (with, say, two identical hole lines) of type (135 iii). Such an unlinked diagram (a) shown below is cancelled by the corresponding linked EPV diagram (b) [only relevant hole lines are labeled].



(9.146)

It is, clearly, again irrelevant whether or not we consider these diagrams, even though they do not vanish individually but cancel one another. However, in order to achieve the cancellation of unlinked diagrams with renormalization terms as shown in the preceding section, we have to include diagram (a) of (146) among the unlinked diagrams, since an analogous term clearly occurs in the renormalization term, which is given by the product of two disconnected parts (cf., Eqs. (136) – (140)), the summations in each disconnected part being independent of other parts. Consequently, we then have to keep also the EPV diagram (b) among the linked diagrams in our final expression for the third order contribution.

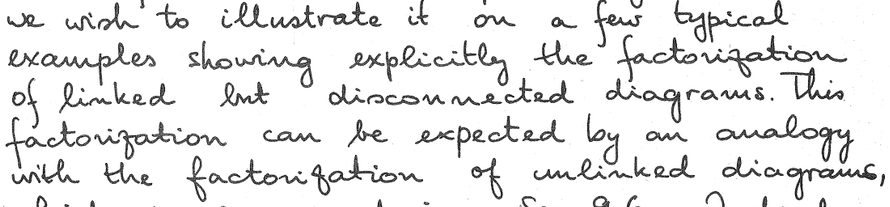
We shall see in the next section that EPV diagrams are also essential for the validity of the factorization lemma, and for the wavefunction factorization considered later. We can also make the postulate, stated toward the end of the preceding section, more precise by

requiring that all linked diagrams, including the EPV ones, be considered when evaluating the wavefunction in any order of perturbation theory.

9.8. Factorization Lemma

Before we state and prove a general result we wish to illustrate it on a few typical examples showing explicitly the factorization of linked but disconnected diagrams. This factorization can be expected by an analogy with the factorization of unlinked diagrams, which we examined in Sec. 9.6. Indeed, removing the last (i.e., the leftmost) interaction on the vacuum part, we obtain a linked diagram for the wavefunction (this assumes, of course, that there is only one vacuum part in the unlinked diagram considered so that this argument does not apply to a general unlinked diagram) and we can carry out the same procedure as for the unlinked diagrams. The only complicating factor is the possibility that the disconnected parts can be topologically equivalent. (In fact this same complication will arise in unlinked diagram factorization as well when the diagram contains several identical vacuum parts.)

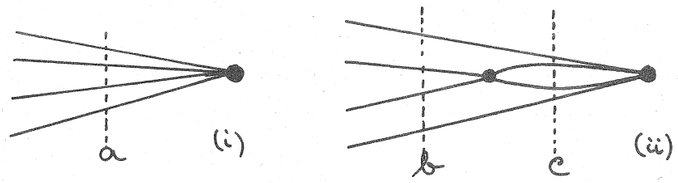
We first consider the non-equivalent components case, say, the diagrams of type (ii), Eq. (135). In this case there are obviously possible three time versions of the first kind, namely



(9.147)

Note that the orientation of fermion lines is not essential in this factorization so that the

considered example applies to any one of the three possible cases (i.e., p-h, p-p or h-h interaction on the leftmost vertex in (ii) or (iii)). The matrix elements, weights or operators associated with time versions (147) are clearly identical so that it suffices to consider the denominator parts. If we designate the denominators, which are associated with separate disconnected parts, as indicated below



(9.148)

we can write the denominator part for the time versions (i)-(iii) of (147) as

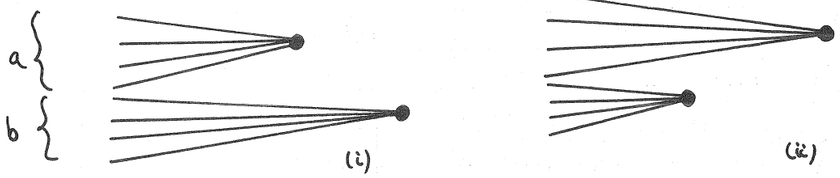
$$\frac{1}{(a+b)\,bc}+\frac{1}{(a+b)(a+c)\,c}+\frac{1}{(a+b)(a+c)\,a} \tag{9.149}$$

$$=\frac{1}{(a+b)\,bc}+\frac{1}{(a+b)\,ac}=\frac{1}{abc}=\left(\frac{1}{a}\right)\cdot\left(\frac{1}{bc}\right).$$

Thus, the contribution from all the time versions yields an expression which factorizes into the product of expressions associated separately with its disconnected parts (148).

We next consider the second order case, which involves two identical diagrams, Eq. (134i). This diagram does not involve any contractions and is clearly associated with the first term on the right hand side of the generalized Wick theorem, and thus with a single system of pairings (namely the one with no contractions whatsoever). Moreover, the denominators in both equivalent parts may be either identical or distinct. Assuming distinct denominators, we can take the R-diagram (134i) twice, in two equivalent time versions (note that all spin-orbital labels

are free)



(9.150)

and we obtain again

$$\frac{1}{(a+b)\,b} - \frac{1}{(a+b)\,a} = \frac{1}{ab} \,. \tag{9.151}$$

Since we have taken this term twice, we get in fact

$$\tfrac{1}{2}\cdot\frac{1}{ab} = \tfrac{1}{2}\left(\frac{1}{a}\right)\left(\frac{1}{b}\right). \tag{9.152}$$

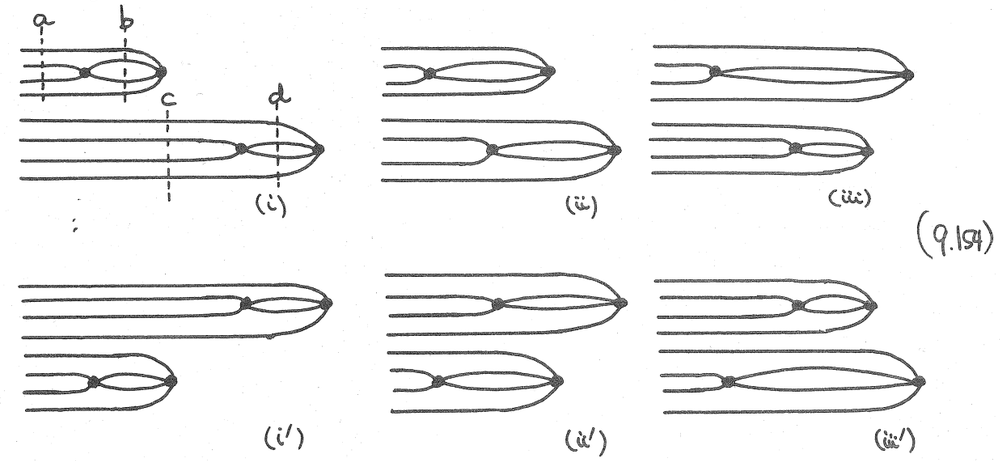
If both denominators are identical ($a=b$), then the above two contributions (or time versions) are identical and we obtain for each separately

$$\frac{1}{2a}\cdot\frac{1}{a} = \tfrac{1}{2}\left(\frac{1}{a}\right)\left(\frac{1}{b}\right). \tag{9.153}$$

Thus, in each case the factor ($\frac{1}{2}$) results, which is associated with the fact that both disconnected diagrams are equivalent. This is consistent with our rules for topological factors, since the disconnected equivalent parts may be arbitrarily permuted among themselves <u>when factored</u>, and considered thus as independent (cf., the rule given by Eq. (5.59)).

Since this latter example did not involve any contractions we now consider another example of the second case, taking the fourth order linked disconnected diagram consisting of two equivalent parts of the type (iv), (v) or (vi) of (134). We suppress again the line orientation for simplicity so that our example applies again to either of the three possible cases (note, however, that the mixed cases, say pp and ph, belong to the first category of

diagrams with non-equivalent parts). The three possible systems of pairings obtained from the Wick theorem can be represented by equivalent time versions shown above one another below



(9.154)

Designating the denominators as shown in (154i) we get

$$\frac{1}{(a+c)(b+c)\,c\,d} + \frac{1}{(a+c)(b+c)(b+d)\,d} + \frac{1}{(a+c)(b+c)(b+d)\,b}$$

$$+ \frac{1}{(a+c)(a+d)\,a\,b} + \frac{1}{(a+c)(a+d)(b+d)\,b} + \frac{1}{(a+c)(a+d)(b+d)\,d}$$

$$= \frac{1}{ab}\cdot\frac{1}{cd} \tag{9.155}$$

so that the final contribution is ($\frac{1}{2}$) of that given by (155) since each time version was considered twice. Considering the case of identical denominators ($a=c$, $b=d$), only the first three diagrams are needed, yielding

$$\frac{1}{2a\,(a+b)\,ab} + \frac{1}{2a\,(a+b)\,2b\,.\,b} + \frac{1}{2a\,(a+b)\,2b\,.\,b}$$

$$= \tfrac{1}{2}\left(\frac{1}{ab}\right)\cdot\left(\frac{1}{ab}\right). \tag{9.156}$$

We also recommend to the reader that he

examins other examples involving more than two disconnected parts.

We will now present a general proof of the factorization theorem following the work of Frantz and Mills.

Consider all possible time versions for a linked diagram consisting of two disconnected parts called a and b. These parts are not necessarily connected so that the theorem is completely general. We designate the set of energy denominators for the part a alone by Δ^a_i, $i=1,\dots,m$, and for the part b, Δ^b_i, $i=1,\dots,n$. The denominators are numbered along the time axis, i.e., the rightmost denominators are Δ^a_1 and Δ^b_1 in parts a and b, respectively.

The denominator contribution from all possible time versions of the first kind, corresponding to all possible orderings of energy vertices in both disconnected parts relative one another (the numerator part being identical for all time versions) can be written as

$$D^{ab}_{mn} = \sum_{\{\alpha,\beta\}} \prod_{p=1}^{m+n} \left(\Delta^a_{\alpha(p)} + \Delta^b_{\beta(p)}\right)^{-1}, \tag{9.157}$$

where the summation extends over all sets of $(m+n)$ integer pairs $\Gamma_p := (\alpha(p), \beta(p))$ defined as follows

(i) either $\Gamma_1 = (1,0)$ or $\Gamma_1 = (0,1)$

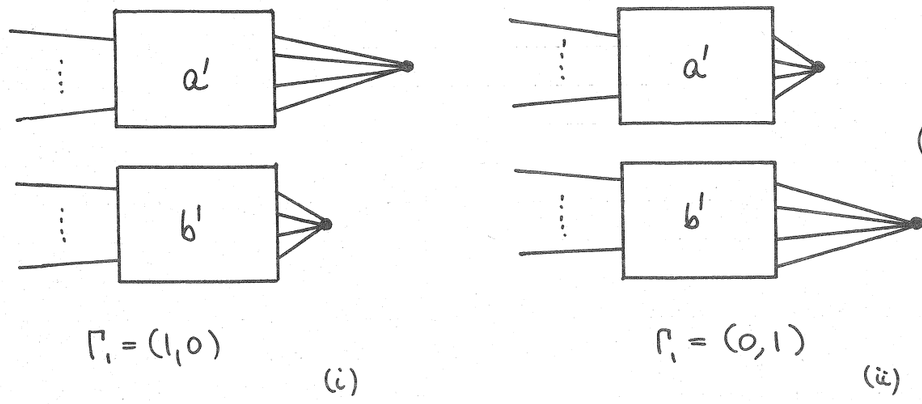
(ii) either $\Gamma_{p+1} = (\alpha(p)+1, \beta(p))$ or $\Gamma_{p+1} = (\alpha(p), \beta(p)+1)$

(iii) $\Gamma_{m+n} = (m, n)$

and where we also define

$$\Delta^a_0 = \Delta^b_0 := 0. \tag{9.158}$$

The first condition (i) corresponds to the two possible time orderings for the rightmost vertex in each disconnected part, i.e. schematically



(9.159)

$\Gamma_1 = (1,0)$ (i) $\Gamma_1 = (0,1)$ (ii)

where a' and b' symbolize the disconnected parts a and b, respectively, after the rightmost vertex has been removed. Thus, the first denominator in case (i) is Δ^a_1 and in the case (ii) it is Δ^b_1. The last (i.e., the leftmost) denominator is clearly $(\Delta^a_m + \Delta^b_n)$ in either case, as required by condition (iii) above. The second condition (ii) then expresses the possibility that, after going through the first p vertices in the whole diagram, the next vertex can be in either the a or in the b part, respectively.

For the separate disconnected parts the denominators are given by the products of Δ^a_i or Δ^b_i, respectively, which we can also write using a general expression, Eq. (157), as

$$D^a_m := D^{ab}_{mo} = \prod_{i=1}^{m} (\Delta^a_i)^{-1}, \tag{9.160a}$$

$$D^b_n := D^{ab}_{on} = \prod_{j=1}^{n} (\Delta^b_j)^{-1}. \tag{9.160b}$$

It is also useful to define

$$D^a_0 = D^b_0 = D^{ab}_{00} := 1. \tag{9.160c}$$

The desired factorization lemma can now be simply stated as

L: $$D^{ab}_{mn} = D^a_m \cdot D^b_n. \tag{9.161}$$

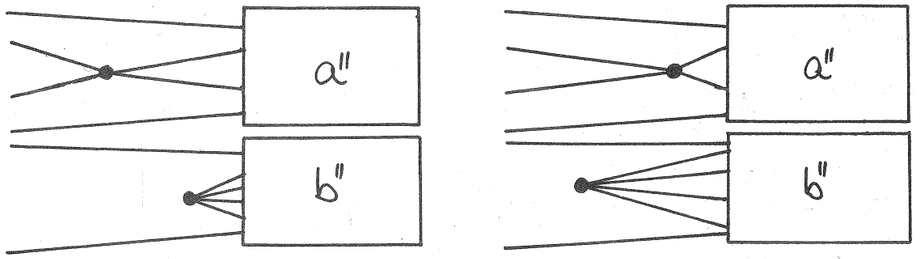
The proof is easily carried out using mathematical induction. The lemma holds when $m=0$ or $n=0$, since

$$D^{ab}_{mo} = D^a_m D^b_o = D^a_m \tag{9.162a}$$

or

$$D^{ab}_{om} = D^a_o D^b_m = D^b_m \tag{9.162b}$$

in agreement with Eqs. (160). Assume, thus, that the lemma holds for $M = m-1$, $N=n$ and $M=m$ and $N = n-1$, $m, n \geq 1$. Clearly, all the terms in D^{ab}_{mn} can be divided into the two disjoint classes according to whether the leftmost interaction occurs in a or in b subgraph, respectively, as schematically shown below



(i) (ii) (9.163)

The last denominator factor being always the same as required by (iii), namely $(\Delta^a_m + \Delta^b_n)$, we can write

$$D^{ab}_{mn} = (\Delta^a_m + \Delta^b_n)^{-1} \left(D^{ab}_{m-1,n} + D^{ab}_{m,n-1} \right), \tag{9.164}$$

since all the remaining factors are identical with those characterizing the disconnected diagrams which result when one leftmost vertex is deleted: either in the a part ($M=m-1, N=n$) or in the b part ($M=m, N=n-1$). This results holds even when m or n equals 1 as may be easily verified.

Since the lemma (161) holds for $M=m-1, N=n$ and $M=m, N=n-1$ by assumption, we can write Eq. (164) as

$$D^{ab}_{mn} = (\Delta^a_m + \Delta^b_n)^{-1} \left(D^a_{m-1} D^b_n + D^a_m D^b_{n-1} \right), \tag{9.165}$$

and since the denominators of separate parts are given by a single product, Eqs. (160 a) and (6), so that

$$D^a_m = D^a_{m-1} \cdot (\Delta^a_m)^{-1} \tag{9.166}$$

or

$$D^a_{m-1} = \Delta^a_m D^a_m , \tag{9.166'}$$

and similarly for the b part, we get finally that

$$D^{ab}_{mn} = (\Delta^a_m + \Delta^b_n)^{-1} \left(\Delta^a_m D^a_m D^b_n + \Delta^b_n D^a_m D^b_n \right)$$

$$= D^a_m D^b_n , \tag{9.167}$$

proving the lemma.

9.9. Linked Cluster Theorem

When we examined the structure of the first few orders of perturbation theory expressions we have seen that all the renormalization terms are cancelled by the unlinked terms originating from the principle term, so that the net contribution in each order was given by all the linked diagrams (including the EPV ones, cf. Sec. 9.7.).

We now postulate this result for an arbitrary order, so that

$$|\Psi_0\rangle = \sum_{N=0}^{\infty L} (P_0 W)^N |\Phi_0\rangle , \tag{9.168}$$

where L indicates that the summation extends over linked diagrams only. Equivalently, we can also use the following notational convention, if convenient:

$$|\Psi_0\rangle = \sum_{N=0}^{\infty} \{(P_0 W)^N\}_L |\Phi_0\rangle . \tag{9.168'}$$

This implies that the exact energy, k_0, is given by

$$k_0 - x_0 = \langle \Phi_0 | \hat{W} \sum_{m=0}^{\infty L} (\hat{P}_0 \hat{W})^m | \Phi_0 \rangle = \sum_{m=0}^{\infty} \langle \Phi_0 | \hat{W} | \Psi_0^{(m)} \rangle$$
$$= \sum_{m=0}^{\infty} \langle \Phi_0 | \hat{W} \{ (\hat{P}_0 \hat{W})^m \}_L | \Phi_0 \rangle$$
$$= \sum_{m=0}^{\infty L} \langle \hat{W} (\hat{P}_0 \hat{W})^m \rangle \quad , \tag{9.169}$$

i.e., by the sum of connected (linked) vacuum diagrams, since all the external lines from each $|\Psi_0^{(m)}\rangle$, $m>0$ (note that $m=0$ term vanishes) must terminate on the added leftmost $\hat{W}$ interaction vertex in order to obtain a non-vanishing Fermi vacuum mean value.

In order to prove the linked cluster theorem as expressed by Eq. (168), we apply to it the operator $(x_0 - \hat{K}_0)$ obtaining

$$(x_0 - \hat{K}_0) |\Psi_0\rangle = (x_0 - \hat{K}_0) \Big[|\Phi_0\rangle + \sum_{m=1}^{\infty L} (\hat{P}_0 \hat{W})^m |\Phi_0\rangle \Big]$$
$$= (x_0 - \hat{K}_0) \sum_{m=1}^{\infty L} (\hat{P}_0 \hat{W})^m |\Phi_0\rangle$$
$$= (x_0 - \hat{K}_0) \hat{P}_0 \sum_{m=1}^{\infty L} \hat{W} (\hat{P}_0 \hat{W})^{m-1} |\Phi_0\rangle, \tag{9.170}$$

where in each term $(P_0 \hat{W})^m |\Phi_0\rangle$ with $m \geq 1$ there is at least one external fermion line extending to the left of its incident vertex. Recalling that (cf., Eqs. (15), (7))

$$\hat{P}_0 = \frac{1 - |\Phi_0\rangle\langle\Phi_0|}{x_0 - \hat{K}_0} \quad , \tag{9.171}$$

we can write Eq. (170) as

$$(x_0 - \hat{K}_0) |\Psi_0\rangle = (1 - |\Phi_0\rangle\langle\Phi_0|) \sum_{m=1}^{\infty L} \hat{W} (\hat{P}_0 \hat{W})^{m-1} |\Phi_0\rangle$$
$$= \sum_{m=1}^{\infty L} \hat{W} (\hat{P}_0 \hat{W})^{m-1} |\Phi_0\rangle$$
$$- |\Phi_0\rangle \sum_{m=1}^{\infty L} \langle \Phi_0 | \hat{W} (\hat{P}_0 \hat{W})^{m-1} |\Phi_0\rangle \, . \tag{9.172}$$

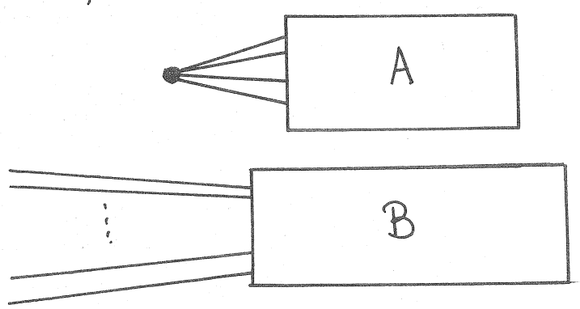
Now, the last term in Eq. (172) must vanish, since in each diagram there is at least one fermion line extending to the left as noted above, and thus at least one creation operator $Y_A^\dagger$ acting on $\langle \Phi_0 |$ from the right (note that the multiplication by $(x_0 - \hat{K}_0)$ only removes the leftmost denominator leaving the operator part unchanged). Thus

$$(x_0 - \hat{K}_0) |\Psi_0\rangle = \sum_{m=0}^{\infty L} \hat{W} (P_0 W)^m |\Phi_0\rangle, \tag{9.173}$$

and the leftmost $\hat{W}$ operator can be taken outside of the summation symbol if we compensate for the possible unlinked terms, i.e.

$$(x_0 - \hat{K}_0) |\Psi_0\rangle = \hat{W} \sum_{m=0}^{\infty L} (P_0 W)^m |\Phi_0\rangle - \Big\{ W \sum_{m=0}^{\infty L} (P_0 W)^m |\Phi_0\rangle \Big\}_{UL}, \tag{9.174}$$

since any diagram which has become unlinked at any point cannot be linked again by subsequent interactions. Thus, the leftmost $\hat{W}$ operator cannot influence the linked property of the part preceeding it, so that the last term in Eq. (174) represents diagrams which are unlinked due to the leftmost $\hat{W}$ interaction. This can be schematically illustrated as follows

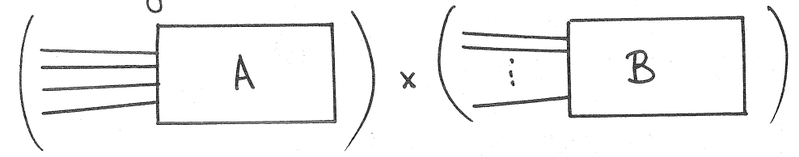


(9.175)

since this diagram must be linked when we remove the leftmost interaction.

Once the last vertex is removed in (175), we can apply the factorization lemma obtaining the product of contributions from separate diagrams,

schematically



(9.176)

Addition of last interaction vertex as in (175) will not change this factorization since it will only replace the operators associated with the external lines of A by an appropriate matrix element associated with added $\hat{W}$ vertex.*)

The subgraph B can have obviously any number of external lines and both A and B diagrams in (175) or (176) must be internally linked. Therefore, the factor corresponding to the B diagram in (176) or (175) contains all possible linked diagrams of the wavefunction type and is thus identical with (168), while the A factor in (175) can be identified with the total energy $(k_0 - \varkappa_0)$, Eq. (169). With these identifications we can thus write the last term in Eq. (174) as $(k_0 - \varkappa_0)|\Psi_0\rangle$ so that we obtain

$$(\varkappa_0 - \hat{K}_0)|\Psi_0\rangle = \hat{W}|\Psi_0\rangle - (k_0 - \varkappa_0)|\Psi_0\rangle \qquad (9.177)$$

or, equivalently

$$(\hat{K}_0 + \hat{W})|\Psi_0\rangle = \hat{K}|\Psi_0\rangle = k_0|\Psi_0\rangle, \qquad (9.178)$$

proving our hypothesis of the linked cluster theorem as expressed by Eq. (168), implying Eq. (169) for the energy.

*) Note that the left hand side vertex of the vacuum part in (175) must be always the <u>leftmost</u> vertex of the unlinked diagram (175) representing the second term on the right hand side of Eq. (174).

9.10 Connected cluster structure of the wavefunction

We have seen that the exact wavefunction $|\Psi_0\rangle$, Eq. (168), is given in each order of perturbation theory by the sum of all linked diagrams, including EPV diagrams. We shall now show that we can write this wavefunction in yet simpler form, which also reveals its structure which is of basic importance for the so called coupled cluster approach, considered in the next chapter.

Let us classify the linked diagrams representing the terms in Eq. (168) by the number r of disconnected parts they contain. The class $r = 1$ will thus include all possible <u>connected diagrams</u> of the wavefunction type. This class can be represented formally by some operator $\hat{T}$ defined as follows

$$\hat{T}|\Phi_0\rangle = \sum_{m=1}^{\infty}{}^{C} (P_0 W)^m |\Phi_0\rangle, \qquad (9.179)$$

where the sum extends in each order over all connected wavefunction diagrams. The connected diagrams appearing in (179) may be classified by the perturbation theory order n as well as by the number $(2j)$ of external lines they contain. This number must be positive and characterizes the excitation order. We can thus write

$$\hat{T} = \sum_{j=1}^{N} \hat{T}_j, \qquad (9.180)$$

since we can restrict ourselves to the N-times excited states when we consider a system with N-electrons. The j-times excited components of $\hat{T}$ are thus given as follows

$$\hat{T}_i = \sum_{m=1}^{\infty} T_i^{(m)} \qquad (9.181)$$

where

$$\hat{T}_j^{(n)} |\Phi_0\rangle = \{(P_0 W)^n |\Phi_0\rangle\}_{c,jj} \qquad (9.182)$$

designates all connected diagrams with n one- and two-particle vertices and with $2j$ (j particle and j hole) external lines.

Since $\hat{T}$ represents some operator defined on the subspace spanned by the unperturbed wavefunction $|\Phi_0\rangle$, which generally will consist from one-, two-, etc. body components, we can also write using a general formula, Eq. (3.60),

$$\hat{T}_j = (j!)^{-1} \sum_{\{A^i, A_i\}} \langle A^1 \cdots A^j | \hat{t}_j | A_1 \cdots A_j \rangle \, N[X_{A^1}^\dagger \cdots X_{A^j}^\dagger X_{A_j} \cdots X_{A_1}]$$
$$= (j!)^{-2} \sum_{\{A^i, A_i\}} \langle A^1 \cdots A^j | \hat{t}_j | A_1 \cdots A_j \rangle_{\mathcal{A}} \, N[X_{A^1}^\dagger \cdots X_{A^j}^\dagger X_{A_j} \cdots X_{A_1}], \qquad (9.183)$$

where $\{A^i, A_i\}$ designates that the summations extend independently over all A^i and A_i, $i=1,\ldots,j$. The antisymmetrized and ordinary $\hat{t}_j$ matrix elements are related as usual

$$\langle A^1 \cdots A^j | \hat{t}_j | A_1 \cdots A_j \rangle_{\mathcal{A}} = \sum_{P \in \mathcal{P}_j} (-1)^P \langle A^1 \cdots A^j | \hat{t}_j | A_{P_1} \cdots A_{P_j} \rangle. \qquad (9.184)$$

These matrix elements are determined by the appropriate scalar part associated with connected diagrams representing $\hat{T}_j$, Eq. (182).

All possible linked diagrams which contain two disconnected parts ($r=2$) can then be expressed in view of the factorization lemma, Sec. 9.8, as

$$\tfrac{1}{2} \hat{T}^2 |\Phi_0\rangle, \qquad (9.185)$$

the factor $(1/2)$ appearing in view of a double counting of diagrams with non-equivalent disconnected parts and as required by the diagram factorization with equivalent parts. This factor is also consistent with our general rules for topological factors: in the factorized form the contributions from disconnect parts commute so that we have to consider independent disconnect parts. Thus, just as in the algebraic form, we get the factor $(1/2)$ if both disconnected parts are equivalent and we also need the factor $(1/2)$ to correct for the double counting of non-equivalent parts when the summations are carried out without restrictions.

We can illustrate this property by exploiting a partial classification of connected diagrams as expressed by Eq. (180) which yields

$$\tfrac{1}{2}\hat{T}^2 = \tfrac{1}{2}\Big(\sum_j \hat{T}_j\Big)^2 = \tfrac{1}{2}\sum_i \hat{T}_i^2 + \sum_{i<j} \hat{T}_i \hat{T}_j. \qquad (9.186)$$

The same property will hold when $\hat{T}_i$ is expressed as a sum of contributions given by individual connected diagrams. ⑨ It is then easily seen that a general component with r disconnected parts can be represented by the term

$$(r!)^{-1} \hat{T}^r |\Phi_0\rangle \qquad (9.187)$$

so that we can write $|\Psi_0\rangle$, Eq. (168), as

$$|\Psi_0\rangle = |\Phi_0\rangle + \hat{T}|\Phi_0\rangle + \tfrac{1}{2}\hat{T}^2|\Phi_0\rangle + \cdots + (r!)^{-1}\hat{T}^r|\Phi_0\rangle + \cdots$$
$$= (1 + \hat{T} + \tfrac{1}{2}\hat{T}^2 + \cdots + \tfrac{1}{r!}\hat{T}^r + \cdots)|\Phi_0\rangle$$
$$= e^{\hat{T}} |\Phi_0\rangle. \qquad (9.188)$$

We thus find that the exact wavefunction $|\Psi_0\rangle$ may be written through the connected diagrams defining the operator $\hat{T}$, Eq. (179). This result suggests, in turn, that we write the so-called wave operator $\hat{W}$, which transforms the unperturbed IPM wavefunction $|\Phi_0\rangle$ into the exact wavefunction,

$$|\Psi_0\rangle = \hat{W} |\Phi_0\rangle, \qquad (9.189)$$

as an exponential of the so-called cluster operator $\hat{T}$,

$$\hat{W} = e^{\hat{T}}, \qquad (9.190)$$

which in turn is defined in terms of connected wavefunction diagrams.

The expansion (188) or (190) in terms of the connected diagram contributions is usually refered to as a <u>cluster expansion</u> in view of its similarity with an analogous exponential expansion exploited in statistical mechanics. The operator $\hat{T}$ is then called a <u>cluster operator</u> and its j-particle components T_j are refered to as <u>j-times excited</u> or <u>j-body clusters</u>. The form of the wave operator given by Eq. (190) is also called a <u>cluster expansion ansatz</u>.

One of the basic advantages of this ansatz stems from the basic property of the exponential mapping which transforms the additive structures into the multiplicative ones. Thus, considering two independent non-interacting systems A and B,

$$|\Psi_o^X\rangle = e^{T^{(X)}} |\Phi_o^X\rangle \quad , \quad X = A, B \tag{9.191}$$

$$[T^{(A)}, T^{(B)}] = 0$$

as a single system, we find that

$$|\Psi_o^{AB}\rangle = e^{T^{(A)} + T^{(B)}} |\Phi_o^{AB}\rangle \tag{9.192}$$

since

$$|\Phi_o^{AB}\rangle = |\Phi_o^A\rangle |\Phi_o^B\rangle . \tag{9.192'}$$

We shall exploit this form of the wavefunction in Chapter 10 where we will formulate equations for a direct determinations of cluster components.

9.11 Hartree-Fock Diagrams

The reader will have noticed that by appropriate definitions of the unperturbed and the perturbed operators $\hat{K}_o$ and $\hat{K}$, respectively, we have achieved a considerable simplification of the perturbation formulas. The key to this simplification was the fact that we have transformed all the operators (in particular, the perturbation operator $\hat{W}$, Eqs. (86)-(88)), to the N-product form. This in turn enabled us to apply generalized form of the Wick theorem and thus to eliminate all the diagrams with internal lines incident with vertices of the same s-vertex (cf., Sec. 5.6). In this way the number of diagrams which must be explicitly considered was significantly reduced. This transformation to the N-product form also eliminated the first order contribution, Eq. (91b), and consequently a large number of renormalization terms [cf., Eqs. (32)-(34), (36), (42) and (43) vs. Eqs. (91) and Eqs. (56) vs. Eqs. (92)].

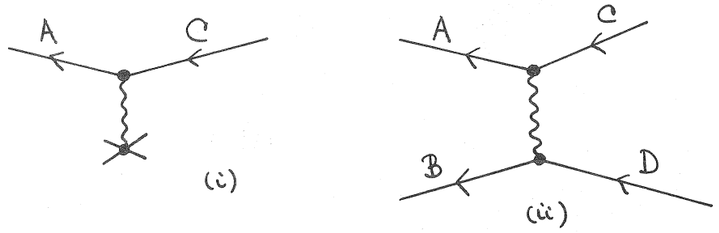
Should we exploit directly the Hamiltonians H_o and H, Eqs. (61), (62), in our perturbation theory, the perturbation operator $\tilde{W}$,

$$\tilde{W} = \hat{H} - \hat{H}_o = \tilde{V} - \hat{U} = \tilde{W}_1 + \tilde{W}_2 \tag{9.193}$$

where

$$\tilde{W}_1 = -\hat{U} \quad , \quad \tilde{W}_2 = \hat{V} \quad , \tag{9.193'}$$

will not be in the N-product form. Thus, the perturbation formulas (32)-(34), (36), etc. must be employed as well as an ordinary rather than generalized Wick's theorem. We can then represent the one- and two-particle components of the perturbation operator (193) by the diagrams with wavy interaction lines as we did in Sec. 5.6, namely by the following $\tilde{U}$ and $\tilde{V}$ diagrams



(9.194)

The scalar factors associated with these diagrams are

$$d_{\tilde{U}} = -\langle A|\hat{u}|C\rangle = d_U, \tag{9.195a}$$

$$d_{\tilde{V}} = \langle AB|\hat{v}|CD\rangle = d_V, \tag{9.195b}$$

and the operators

$$\hat{D}_{\tilde{U}} = d_{\tilde{U}}\, X_A^\dagger X_C, \tag{9.196a}$$

$$\hat{D}_{\tilde{V}} = d_V\, X_A X_B X_D X_C. \tag{9.196b}$$

We note, however, that once we apply Wick's theorem (that is, once we form the R-diagrams) we must replace the wavy lines by the dashed ones since all the terms resulting from Wick's theorem are in the N-product form.

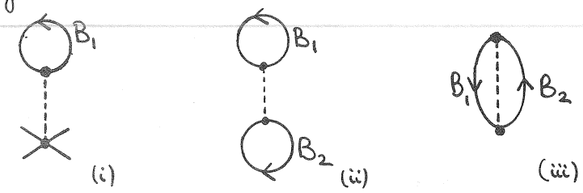
The zero order contribution to pertubation theory is then given by ε_0, Eq. (73),

$$E_0^{(0)} = \varepsilon_0 = \sum_{B_1}\langle B_1|\hat{z}+\hat{u}|B_1\rangle = \sum_{B_1}\omega_{B_1}, \tag{9.197}$$

while the first order contribution $E_0^{(1)}$ is

$$E_0^{(1)} = \langle \tilde{W}\rangle, \tag{9.198}$$

and is given by the following R-diagrams involving a single $\tilde{U}$ or $\tilde{V}$ vertex



(9.199)

where we use the dashed lines for the reasons explained above (in fact, for the vacuum type diagrams it is irrelevant which type of vertices is used since all the operators have been contracted and the scalar factors assigned to the wavy and dashed type vertices are identical). We also recall the rule (cf. Sec. 5.6) that the internal lines which are incident with the vertices of the same s-vertex are always the hole lines. The algebraic form of the first order contribution (198), as given by the diagrams (199), is

$$E_0^{(1)} = -\sum_{B_1}\langle B_1|\hat{u}|B_1\rangle + \tfrac{1}{2}\sum_{B_1B_2}\left(\langle B_1B_2|\hat{v}|B_1B_2\rangle - \langle B_1B_2|\hat{v}|B_2B_1\rangle\right)$$

$$= \sum_{B_1}\left(\tfrac{1}{2}\langle B_1|\hat{q}|B_1\rangle - \langle B_1|\hat{u}|B_1\rangle\right). \tag{9.200}$$

The overall contribution from the zero and the first orders of the perturbation theory, which vanished when we used transformed Hamiltonians in the N-product form, is now

$$E_0^{(0)} + E_0^{(1)} = \sum_{B_1}\langle B_1|\hat{z}+\tfrac{1}{2}\hat{q}|B_1\rangle = \langle\hat{H}\rangle, \tag{9.201}$$

that is the vacuum mean value of $\hat{H}$, which we have used as the zero of energy in our approach [cf., Eq. (67) or (71)]. Note that both of these contributions are present even when we choose the Hartree-Fock reference state when

$$\hat{U} = \hat{G}, \tag{9.202}$$

so that

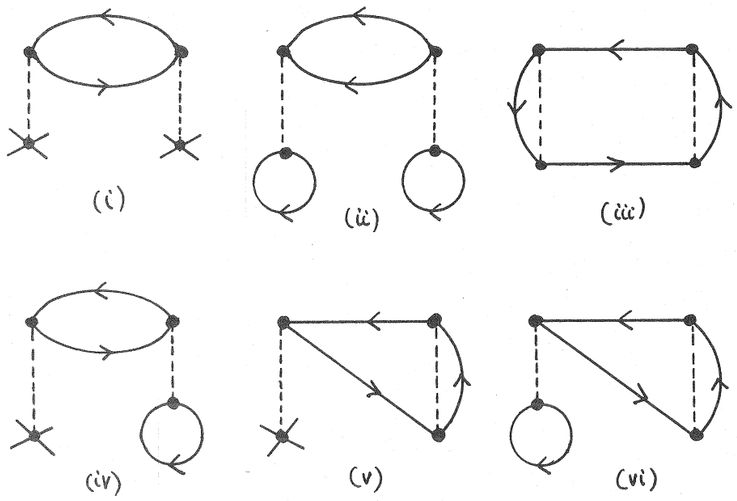
$$E_0^{(0)} = \sum_{B_1}\langle B_1|\hat{z}+\hat{q}|B_1\rangle \tag{9.203}$$

and

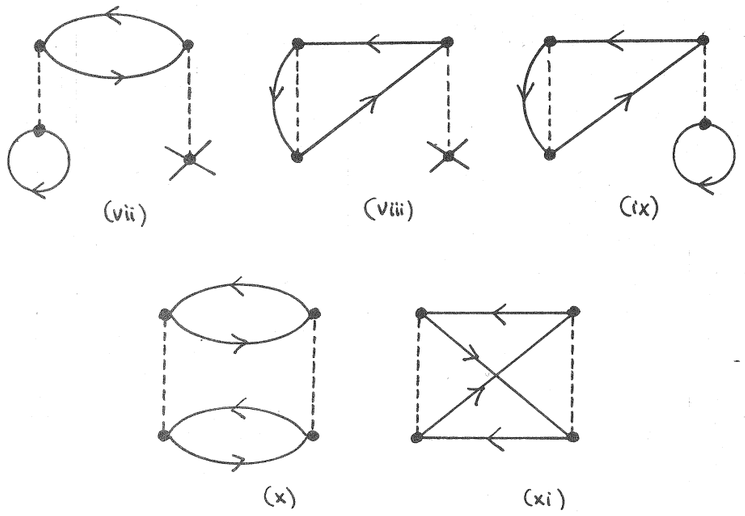
$$E_0^{(1)} = -\tfrac{1}{2}\sum_{B_1}\langle B_1|\hat{q}|B_1\rangle. \tag{9.204}$$

In this case $E_0^{(1)}$ corrects a simple sum of orbital energies, Eq. (73) or (197), for the double-counting of the two-electron interactions.

In the second order of perturbation theory we then obtain the following Goldstone-type R-diagrams (resp. skeletons)



(9.205)

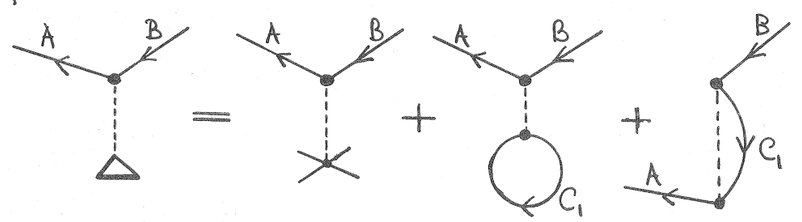


The last two diagrams [(x) and (xi)] are clearly identical with diagrams (ii) and (iii) of (112), while the first nine diagrams [(i) – (ix)] are represented in our earlier formalism by the single diagram (i) of (112). This is easily seen if we interpret diagrammatically the transition of the operator $\tilde{W}$, Eq. (193),

to its N-product form, Eq. (86), as in Sec. 5.6, Eqs. (5.105) and (5.107) [Note, however, that the single-particle cross-type vertex, diagram (104i) or (194i), represents now the operator $(-\hat{U})$, Eqs. (195a), (196a), rather than the operator $\hat{Z}$ as in Sec. 5.6.].

The vacuum diagrams in Eqs. (5.105) and (5.107) [i.e., diagram (iii) of Eq. (5.105) and diagrams (v) and (vi) of Eq. (5.107)] represent the first order contribution given by Eq. (200). In the third and higher order terms they will also yield disconnected diagrams which cancel corresponding renormalization terms (note, however, that they cannot contribute in the left- and right-most operators in the Fermi vacuum mean values).

The two-electron N-product operator given by the diagram (ii) of Eq. (5.107) will then yield the same diagrams as considered in our earlier approach based on W in the N-product form. Finally, all the remaining diagrams [i.e., diagram (ii) of Eq. (5.105) and diagrams (iii) and (iv) of Eq. (5.107)] represent one-particle operators in the N-product form and can thus be collected into a single one-body operator, which must be equivalent to the $\hat{W}_1 \equiv \hat{Q}_N$ operator, Eq. (87). In diagrammatic form we thus have



(9.206)

which is consistent with the algebraic definition of $\hat{Q}_N$ used earlier [cf., Eqs. (87) and (4.70)]. In an algebraic form Eq. (206) can be written as

$$\hat{Q}_N = \sum_{AB} \langle A|\hat{q}|B\rangle \, N[X_A^\dagger X_B]$$

$$= -\sum_{AB} \langle A|\hat{u}|B\rangle N[X_A^\dagger X_B] + \sum_{AB}\Big(\sum_{C_1}(\langle AC_1|\hat{v}|BC_1\rangle - \langle AC_1|\hat{v}|C_1B\rangle)\Big) N[X_A^\dagger X_B]$$

$$= \sum_{AB}\big(\langle A|\hat{q}|B\rangle - \langle A|\hat{u}|B\rangle\big) N[X_A^\dagger X_B] = \hat{G}_N - \hat{U}_N\,, \qquad (9.207)$$

which is equivalent with Eq. (87).

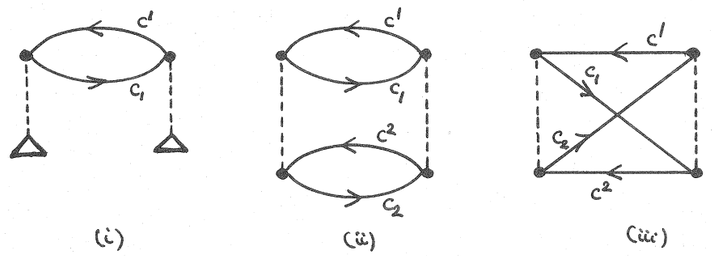
We thus find that all the nine diagrams (i)–(ix) of (205) are represented in our formalism by a single diagram (i) of (112). Moreover, all these diagrams mutually cancel out if we choose the potential $\hat{U}$ to be given by $\hat{G}$, Eq. (202). This is automatically the case if we use Hartree-Fock (spin)-orbitals. In fact, it can be shown (Tolmachev) that the sum of contributions of the diagrams of the type (i) – (ix) of (205) to the infinite order of perturbation theory is equivalent to the use of the Hartree-Fock ground state wavefunction as an unperturbed Hamiltonian (and the Hartree-Fock orbitals as the basis of the second quantization formalism). For these reasons the diagrams containing internal fermion lines which are incident with the same $\hat{v}$-vertex are often called Hartree-Fock diagrams.

We must stress, however, that the above given formalism can be employed with an arbitrary reference state and corresponding spin-orbital basis. The graphs containing k $\hat{Q}$-vertices then represent 3^k Hartree-Fock type diagrams, whose contribution vanishes if the one particle potential $\hat{U}$ is chosen as in Eq. (202). However, an introduction of the operator $\hat{Q}$, Eq. (87), significantly simplifies the perturbation theory formalism even when Hartree-Fock orbitals are not used and one-particle q-matrix elements are non-vanishing.

9.12 Spin-independent formalism

For the closed-shell singlet states we can exploit a simple formalism outlined in Sec. 5.8 to eliminate an explicit spin dependence. This is easily done by replacing the spin-orbital labels with orbital ones in all Goldstone R-diagrams and by assigning an additional factor 2^l to each Goldstone diagram, where l is the number of closed loops of oriented lines (cf, Sec. 5.8, Eq. (5.121)).

Thus, in the second order, the R-diagrams (112) become



(i) (ii) (iii) (9.208)

which give the following second-order spin-free algebraic expression for the energy

$$k_0^{(2)} = 2\sum_{c_1 c'} \langle c_1|\hat{q}|c'\rangle\langle c'|\hat{q}|c_1\rangle\,\Delta(c_1;c') + \sum_{c_1c_2c'c^2}\langle c_1c_2|\hat{v}|c'c^2\rangle\big(2\langle c'c^2|\hat{v}|c_1c_2\rangle - \langle c'c^2|\hat{v}|c_2c_1\rangle\big)\Delta(c_1c_2;c'c^2), \qquad (9.209)$$

where again

$$\Delta(c_1c_2\cdots c_p;\, c'c^2\cdots c^p) = \Big[\sum_{i=1}^{p}(\omega_{c_i} - \omega_{c^i})\Big]^{-1}. \qquad (9.210)$$

Similarly, the contribution of the third order two-particle diagrams is easily found from the last column of Table II, Sec. 9.5.1, if we replace all spin-orbital labels by the orbital ones and multiply each contribution by the factor 2^l where l is given in the fifth column in the same Table II. In order to obtain the contributions involving one-particle

matrix elements we form Goldstone diagrams as indicated in Sec. 9.5.1. following Brandow diagram lists (128)–(130) and evaluate these following the same rules as above (see also the Summary of these rules in the next Section).

A similar procedure can also be devised for the wavefunction contributions: we can eliminate the explicit spin dependence from the scalar factor part though not from the operator part. The scalar part is therefore obtained from a corresponding spin-orbital expression when we replace the spin-orbital labels by the orbital labels and introduce an additional factor 2^l. The operator part, where explicit spin and orbital labels must be used, is then modified in such a way that the same spin summation label is used for the operator pair associated with each open path, since the spin-projection must be preserved along each oriented fermion path when all interactions are spin-independent. We can also include the summations over these spin variables in the operator part.

Thus, the general rule given by Eq. (5.99) takes the form

$$\hat{D}_R = d_R \sum_{\eta_1 \cdots \eta_k} N\left[\prod_{i=1}^{k} \left(X^{\dagger}_{b^i \eta_i} X_{b_i \eta_i}\right)\right], \qquad (9.211)$$

where now d_R is given by the same expression as in the spin-orbital case (cf, Sec. 9.13 for the summary of rules) but with spin-orbital labels B^i (B_i) replaced by orbital labels b^i (b_i). The overall contribution from a given Goldston R-diagram R is then

$$w_R (-1)^{h+l} 2^l \sum_{\substack{b^1 b^2 \cdots b^k \\ b_1 b_2 \cdots b_k}} \sum_{\{c^i c_i\}} \hat{D}_R(b_i, b^i, c_i, c^i) \, |\Phi_0\rangle, \qquad (9.212)$$

where the second sum extends over the orbital labels associated with internal fermion lines.

This procedure is, clearly, not satisfactory if we wish to obtain the wavefunction contributions expressed in terms of spin-adapted states with the spin-coupling of our choice. For example, in the first order case [cf., diagrams (132)] we obtain

$$|\Psi_0^{(1)}\rangle = \sum_{c' c_1} \langle c'|\hat{q}|c_1\rangle \Delta'(c_1; c') \sum_{\eta} N\left[X^{\dagger}_{c'\eta} X_{c_1\eta}\right] |\Phi_0\rangle \qquad (9.213)$$

$$+ \tfrac{1}{2} \sum_{c' c^2 c_1 c_2} \langle c' c^2|\hat{v}|c_1 c_2\rangle \Delta'(c_1 c_2; c' c^2) \sum_{\eta_1 \eta_2} N\left[X^{\dagger}_{c'\eta_1} X^{\dagger}_{c^2\eta_2} X_{c_2\eta_2} X_{c_1\eta_1}\right] |\Phi_0\rangle$$

We note that this result follows immediately from Eq. (133) when we implement the spin-independence of $\hat{Q}$ and $\hat{V}$ operators as expressed by relations (3.12) and (3.14). However, such an algebraic elimination of the spin-dependence in higher orders of perturbation theory can be very laborious.

Designating the resulting "spin-adapted" configurations in $|\Psi_0^{(1)}\rangle$, Eq. (213), as

$$\sqrt{2} \begin{vmatrix} c' \\ c_1 \end{vmatrix} := \sum_{\eta} N\left[X^{\dagger}_{c'\eta} X_{c_1\eta}\right] |\Phi_0\rangle = \begin{vmatrix} c'\alpha \\ c_1\alpha \end{vmatrix} + \begin{vmatrix} c'\beta \\ c_1\beta \end{vmatrix} \qquad (9.214)$$

and

$$2 \begin{vmatrix} c' & c^2 \\ c_1 & c_2 \end{vmatrix} := \sum_{\eta_1 \eta_2} N\left[X^{\dagger}_{c'\eta_1} X^{\dagger}_{c^2\eta_2} X_{c_2\eta_2} X_{c_1\eta_1}\right] |\Phi_0\rangle \qquad (9.215)$$

$$= \begin{vmatrix} c'\alpha & c^2\alpha \\ c_1\alpha & c_2\alpha \end{vmatrix} + \begin{vmatrix} c'\alpha & c^2\beta \\ c_1\alpha & c_2\beta \end{vmatrix} + \begin{vmatrix} c'\beta & c^2\alpha \\ c_1\beta & c_2\alpha \end{vmatrix} + \begin{vmatrix} c'\beta & c^2\beta \\ c_1\beta & c_2\beta \end{vmatrix}$$

we can rewrite (213) as follows

$$|\Psi_0^{(1)}\rangle = \sqrt{2} \sum_{c' c_1} \langle c'|\hat{q}|c_1\rangle \Delta'(c_1; c') \begin{vmatrix} c' \\ c_1 \end{vmatrix} \qquad (9.216)$$

$$+ \sum_{c' c^2 c_1 c_2} \langle c' c^2|\hat{v}|c_1 c_2\rangle \Delta'(c_1 c_2; c' c^2) \begin{vmatrix} c' & c^2 \\ c_1 & c_2 \end{vmatrix}.$$

While Eq. (214) defines the only possible singlet spin function in the monoexcited state, we see that the doubly excited states defined by Eq. (215) correspond to non-orthogonal spin-bonded functions since, for example

$$\left\langle \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \middle| \begin{matrix} c^2 c^1 \\ c_1 c_2 \end{matrix} \right\rangle = -\tfrac{1}{2} , \qquad (9.217)$$

assuming $c^1 \neq c^2$ and $c_1 \neq c_2$.

Should we wish to express the doubly excited part of $|\Psi_0^{(1)}\rangle$, Eq. (216), in terms of ortho-normal pp-hh coupled states with desirable transformation properties, i.e.,

$$\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle_0 = \tfrac{1}{2}\left(\left| \begin{matrix} c^1\beta\, c^2\alpha \\ c_1\alpha\, c_2\beta \end{matrix} \right\rangle + \left| \begin{matrix} c^1\alpha\, c^2\beta \\ c_1\beta\, c_2\alpha \end{matrix} \right\rangle - \left| \begin{matrix} c^1\alpha\, c^2\beta \\ c_1\alpha\, c_2\beta \end{matrix} \right\rangle - \left| \begin{matrix} c^1\beta\, c^2\alpha \\ c_1\beta\, c_2\alpha \end{matrix} \right\rangle \right)$$

$$= -\left(\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle + \left| \begin{matrix} c^2 c^1 \\ c_1 c_2 \end{matrix} \right\rangle \right) \qquad (9.218a)$$

$$\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle_1 = \tfrac{1}{2\sqrt{3}}\left(2\left| \begin{matrix} c^1\beta\, c^2\beta \\ c_1\alpha\, c_2\alpha \end{matrix} \right\rangle + 2\left| \begin{matrix} c^1\beta\, c^2\beta \\ c_1\beta\, c_2\beta \end{matrix} \right\rangle \right.$$

$$\left. + \left| \begin{matrix} c^1\beta\, c^2\alpha \\ c_1\alpha\, c_2\beta \end{matrix} \right\rangle + \left| \begin{matrix} c^1\alpha\, c^2\beta \\ c_1\beta\, c_2\alpha \end{matrix} \right\rangle + \left| \begin{matrix} c^1\alpha\, c^2\beta \\ c_1\alpha\, c_2\beta \end{matrix} \right\rangle + \left| \begin{matrix} c^1\beta\, c^2\alpha \\ c_1\beta\, c_2\alpha \end{matrix} \right\rangle \right)$$

$$= \tfrac{1}{\sqrt{3}}\left(\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle - \left| \begin{matrix} c^2 c^1 \\ c_1 c_2 \end{matrix} \right\rangle \right) \qquad (9.218b)$$

we have to replace the functions (215) in Eq. (216) by

$$\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle = \tfrac{1}{2}\left(-\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle_0 + \sqrt{3}\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle_1 \right). \qquad (9.219)$$

Recall that

$$\left| \begin{matrix} c^2 c^1 \\ c_1 c_2 \end{matrix} \right\rangle_0 = \left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle_0 , \qquad (9.220a)$$

while

$$\left| \begin{matrix} c^2 c^1 \\ c_1 c_2 \end{matrix} \right\rangle_1 = -\left| \begin{matrix} c^1 c^2 \\ c_1 c_2 \end{matrix} \right\rangle_1 . \qquad (9.220b)$$

Such an ad hoc procedure of transforming the perturbation contributions to the desired spin-adapted form is clearly unsatisfactory, particularly when higher than doubly excited states are involved, and a proper spin-adaptation formalism, as given in Chapter 6, is again called for.

9.13 Summary of rules for perturbation theory diagrams

We now briefly summarize the rules for obtaining explicit perturbation theory expressions in the n-th order when we employ Goldstone diagrams:

(i) Draw all possible linked R-skeletons having n_1 Q-vertices and n_2 V-vertices, where $n_i \geq 0$ $(i=1,2)$ and $n_1 + n_2 = n$, and determine their weights w_R. For the energy contribution vacuum diagrams have to be considered while for the wavefunction contribution all linked diagrams with $(2k)$ external lines extending to the left of their incident vertices, where $k = 1, 2, \ldots, 2n$, must be used.

(ii) Label the fermion lines with appropriate particle or hole spinorbital labels or, in case of a closed-shell singlet state when spin-free formulas are desired, with corresponding orbital labels, thus forming all possible n-th order perturbation theory R-diagrams.

(iii) Associate with each R-diagram R, obtained in step (ii), the following

algebraic contribution

$$w_R f(R) \sum_{\{\chi\}} d_R\, o(R) , \qquad d_R = n(R)\, d(R) ,$$

where the sum extends over all (spin)-orbital labels and

(a) w_R is the weight of the corresponding R-skeleton,
(b) the numerical factor $f(R)$ is

$$f(R) = (-1)^{l+h}$$

(or

$$f(R) = (-1)^{l+h}\, 2^l$$

in the case of spin-free formalism),
(c) the numerator $n(R)$ is given by the product of matrix elements $\langle A|\hat{q}|C\rangle$ and $\langle AB|\hat{v}|CD\rangle$ assigned to each Q- and V-vertex, respectively, where A, B (C, D) are labels carried by outgoing (ingoing) lines and A and C, and B and D, label lines which are incident with the same simple vertex of the Goldstone R-diagram R,
(d) the denominator $d(R)$ is given by the product of n [or (n+1) for the wavefunction diagrams] denominator factors*) $\Delta(\wp_r)$, $\wp_r = \{B_1 \cdots B_r ; B^1 \cdots B^r\}$,

$$\Delta(\wp_r) = \Delta(B_1 \cdots B_r ; B^1 \cdots B^r) = \left[\sum_{i=1}^{r} (\omega_{B_i} - \omega_{B^i})\right]^{-1} ,$$

where the index sets $\wp_r$ are determined by the labels of fermion lines, which are crossing a hypothetical verticle line separating any two neighboring vertices [and, for the wavefunction diagrams, also by an additional hypothetical vertical line which is to the left of all vertices), and, finally
(e) the "operator" factor $o(R)$ which equals unity for the energy diagrams,

$$o(R) = 1$$

and which yields the appropriate k-excited states for the wavefunction diagrams,

$$o(R) = N[X^\dagger_{B^1} \cdots X^\dagger_{B^k} X_{B_k} \cdots X_{B_1}]\,|\Phi_0\rangle$$

(or

$$o(R) = \sum_{\eta_1 \cdots \eta_k} N[X^\dagger_{b_1\eta_1} \cdots X^\dagger_{b_k\eta_k} X_{b_k\eta_k} \cdots X_{b_1\eta_1}]\,|\Phi_0\rangle$$

for spin-free formalism),
where the ingoing external line of the i-th open path carries the hole label B_i (b_i) and the corresponding outgoing particle line is labelled with B^i (b^i), $i = 1, \ldots, k$.

If only the spin-orbital form of the perturbation formulas is required we can exploit Brandow diagrams (one Goldstone representative for each distinct Hugenholtz diagrams). The above rules are then simply modified as follows:

(i) the weights w^H_R of the corresponding Hugenholtz diagram must be used in place of Goldstone weights, and
(ii) antisymmetrized matrix elements $\langle AB|\hat{v}|CD\rangle_A$ must be used instead of the non-antisymmetrized ones.

The above rules establish a one-to-one correspondence between the perturbation theory diagrams and explicit algebraic expressions (in terms of basic one- and two-electron matrix elements and independent particle model orbital energies). The diagrams represent thus not only a very

*) Note that we only define the first order denominators (similarly in Sec. 9.12) since all the linked terms originate from the principal term where only the first power of $\hat{P}_0$ can appear. For the spin-free formalism an analogous expression is used, obtained by replacing the spinorbital labels by the orbital ones (cf. Eq. (210)).

convenient book-keeping device when a large number of perturbation theory terms have to be considered, but most importantly they also represent a very efficient computational tool, which enables one to obtain explicit algebraic expressions without any algebraic manipulations and which also provides an insight into the algebraic structure of various terms on the first sight. The latter is particularly important when we wish to carry out summations of certain classes of diagrams to an infinite order of perturbation theory or to inter-relate various distinct approaches.

10. Coupled Cluster Approach

We have seen in Sec. 9.10 that the exact wavefunction can be written in the form of the cluster expansion, where the clusters represent appropriate sums of connected wavefunction diagrams. We shall now exploit this structure of the wavefunction by assuming the cluster ansatz, Eqs. (9.188 – 190), for our wavefunction and we shall attempt to derive a system of equations which determine directly various cluster components. Such a theory should have the advantage that it is formulated in terms of extensive quantities, which are additive as illustrated by Eqs. (9.191 – 192').

It would be ideal if we could exploit the cluster ansatz, Eq. (9.188), in the variational approach and obtain a system of equations which determine cluster components $\hat{T}_i$ variationally. In fact, in a certain sense, we can regard the Hartree-Fock equations, derived in Sec. 8.2, as a special case of such an approach: we truncate $\hat{T}$ to the mono-excited clusters $\hat{T}_1$ only and make $\hat{T}_1$ to vanish in the first order. This is of course possible since we have the freedom in the choice of our one-particle basis. We could therefore attempt to truncate the cluster operator $\hat{T}$ up to and including the biexcited component $\hat{T}_2$ and find $\hat{T}_2$ which will minimize the energy expectation value. Unfortunately, the energy mean value expression for the trial wavefunction $\exp(\hat{T}_2)|\Phi_0\rangle$ will contain all positive powers *) of the pair cluster operator $\hat{T}_2$, so that we cannot simply apply the variational principle. It would also make little sense to truncate the energy mean value expression to, say, the quadratic or cubic terms in $\hat{T}_2$ since we would loose

*) This is easily seen if we evaluate this quantity diagrammatically, cf. Sec. 10.2.

the variational character (and, thus, the upper bound property) anyway and since the $\hat{T}_2$ components are very important and thus large as we shall see in Sec. 10.5. Indeed, there seems to be a basic incompatibility between the variational and size-extensive approaches: the terms which must be included in the energy in order to warrant its upper bound character may be shown to possess the unlinked characted, thus making the energy size-inconsistent.*)

We shall thus directly exploit the cluster ansatz in the Schrödinger equation and will determine the cluster components by requiring that this equation is satisfied in a certain subspace spanned by low-excited configurations. We shall see that such an approach yields, in general, a non-linear system of algebraic equations for the cluster components. In fact, these equations can be regarded as recursive formulas for the construction of higher order perturbation contributions of certain kind. Thus, by solving such a system of equations we not only construct all the required diagrams of a given type to the infinite order of perturbation theory but, at the same time, evaluate this sum. It is also interesting that the system of equations determining the cluster components is energy independent. The energy is then given by the $\hat{T}_1$ and $\hat{T}_2$ cluster components only and is linear in these components, which again indicates the size-extensive character of the procedure employed.

We shall outline a general procedure for the construction of coupled cluster equations, which was first presented by Čížek. We will then

*) By the size consistency or size extensivity we understand here a proper dependence on the particle number for $N \to \infty$ in the non-interacting limit.

apply this general procedure to the most important case when the cluster operator $\hat{T}$ is approximated by its pair component $\hat{T}_2$, which leads to the so called coupled-pair many-electron theory (CPMET). We will also briefly discuss possible extensions of this approach when also $\hat{T}_1$ and $\hat{T}_3$ clusters are taken into account. We will also examine the relationship of this approach with both variational and perturbative approaches to the electron correlation problem. Finally, we shall obtain the orthogonally spin-adapted version of the CCA and will discuss its properties and characteristics.

10.1. Coupled Cluster Ansatz

We shall consider an N-electron system, characterized by the Hamiltonian (3.6) with a non-degenerate ground state. We shall employ the second quantization and diagrammatic formalisms based on the spin-orbitals provided by some IPM, for example, the restricted Hartree-Fock model. We shall then exploit a general structure of the exact wavefunction as given by the perturbation theory, and shall write the exact wavefunction for the ground state in the form

$$|\Psi_0\rangle = e^{\hat{T}} |\Phi_0\rangle , \qquad (10.1)$$

referred to as the cluster expansion of $|\Psi_0\rangle$ or the cluster ansatz for $|\Psi_0\rangle$. Here $|\Phi_0\rangle$ designates the ground state wavefunction for our system in the chosen IPM.

The i-particle components of the cluster operator $\hat{T}$,

$$\hat{T} = \sum_{i=1}^{N} \hat{T}_i , \qquad (10.2)$$

defined on the subspace spanned by $|\Phi_0\rangle$, are given by their matrix elements as follows

$$\hat{T}_i = (i!)^{-1} \sum_{\substack{B' \cdots B^i \\ B_1 \cdots B_i}} \langle B' \cdots B^i | \hat{t}_i | B_1 \cdots B_i \rangle X^\dagger_{B'} \cdots X^\dagger_{B^i} X_{B_i} \cdots X_{B_1} \quad (10.3a)$$

$$= (i!)^{-2} \sum_{\substack{B' \cdots B^i \\ B_1 \cdots B_i}} \langle B' \cdots B^i | \hat{t}_i | B_1 \cdots B_i \rangle_{\mathcal{A}} X^\dagger_{B'} \cdots X^\dagger_{B^i} X_{B_i} \cdots X_{B_1}, \quad (10.3b)$$

and may thus be regarded as excitation operators. These operators have a general form of the i-particle (or i-body) operators, as given by Eqs. (3.60), in which all matrix elements between the unoccupied and the occupied spin-orbitals vanish, since the domain of $\hat{T}$ or $\hat{T}_i$ is the subspace $\mathcal{D}_0$ spanned by $|\Phi_0\rangle$ and the range is the othogonal complement to $\mathcal{D}_0$ in the space considered (note that we are interested in correlating $|\Phi_0\rangle$!). These operators are automatically in the N-product form if $|\Phi_0\rangle$ is chosen as the Fermi vacuum in our hole-particle formalism. In fact, they are given as linear combinations of products containing an even number of hole-particle creation operators $Y^\dagger_{B^i}$ and $Y^\dagger_{B_i}$, so that they mutually commute

$$[\hat{T}_i, \hat{T}_j]_- = 0. \qquad (\forall_{i,j}) \qquad (10.4)$$

Consequently, there is no ambiguity in writing various products of these operators and in expanding the exponential in (1) as a power series.

Since cluster operators are excitation operators, the states $\hat{T}_i|\Phi_0\rangle$ [or, generally, $(\prod_j \hat{T}_j^{n_j})|\Phi_0\rangle$] represent linear combinations of i-times [or, generally, $\sum_j (j \cdot n_j)$-times] excited configurations, e.g.

$$\hat{T}_i|\Phi_0\rangle = (i!)^{-1} \sum_{\substack{B' \cdots B^i \\ B_1 \cdots B_i}} \langle B' \cdots B^i | \hat{t}_i | B_1 \cdots B_i \rangle \left| \begin{matrix} B' \cdots B^i \\ B_1 \cdots B_i \end{matrix} \right\rangle \quad (10.5a)$$

$$= (i!)^{-2} \sum_{\substack{B' \cdots B^i \\ B_1 \cdots B_i}} \langle B' \cdots B^i | \hat{t}_i | B_1 \cdots B_i \rangle_{\mathcal{A}} \left| \begin{matrix} B' \cdots B^i \\ B_1 \cdots B_i \end{matrix} \right\rangle, \quad (10.5b)$$

where the i-times excited configurations are defined as in Eq. (5.68) [cf., also Eqs. (5.74). Since the latter expansion, Eq. (5b), can be re-written in the following restricted form

$$\hat{T}_i|\Phi_0\rangle = \sum_{\substack{B'<B''<\cdots<B^i \\ B_1<B_2<\cdots<B_i}} \langle B'B'' \cdots B^i | \hat{t}_i | B_1 B_2 \cdots B_i \rangle_{\mathcal{A}} \left| \begin{matrix} B'B'' \cdots B^i \\ B_1 B_2 \cdots B_i \end{matrix} \right\rangle, \quad (10.6)$$

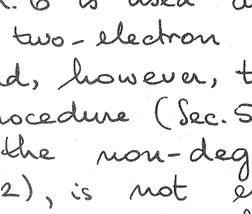
we see that in a spin-orbital formalism it is appropriate to employ the anti-symmetrized t-matrix elements, since the non-anti-symmetrized ones are in fact linearly dependent. Thus, the anti-symmetrized t-matrix elements, which are related in the usual way with the non-antisymmetriged ones [cf., eg., Eq. (5.71) or Eq. (9.184)] represent the coefficients in the CI-type expansion in the i-times excited subspace. We shall see how these coefficients are related with the CI coefficients in Sec. 10.3. The relationship (6) also implies how to proceed when we desire to formulate the theory in terms of chosen orthonormal spin-adapted states in case that $\hat{H}$ is spin-independent, or if we wish to relate cluster components defined with respect to different systems of spin-adapted states.

10.2. Diagrammatic representation

We shall represent the cluster operators $\hat{T}_i$, Eq. (3), by the Goldstone or the Hugenholtz excitation diagrams which are similar to those given in Sec. 5.4, Eqs. (5.75), (5.76). In view of the remark made at the end of Sec. 10.1 it will be particularly convenient to use Brandow representation. Only when a spin-free formulation is required, we have to employ either Goldstone diagrams (when a simple spin-adaptation procedure of Sec. 5.8 is used) or a combination of Brandow and Goldstone diagrams (when a general spin-adaptation procedure of Ch. 6 is used and we do not wish to spin-adapt two-electron matrix elements). We must add, however, that a simple spin-adaptation procedure (Sec. 5.8), as used for example in the non-degenerate perturbation theory (Sec. 9.12), is not entirely satisfactory for the same reasons which limited its usefulness in the wavefunction case, since it leads to the cluster components which correspond to non-orthogonal states, and in higher than doubly-excited case are in fact linearly dependent.

We shall see that in order to obtain orthogonally spin-adapted version of the coupled cluster approach it is most convenient to employ Brandow diagrams for cluster operator components and Goldstone diagrams for spin-independent Hamiltonian (namely its two-electron part). However, in the spin-orbital formalism we can employ the Brandow representation throughout.

We shall thus represent a general $\hat{T}_i$ component by the following Brandow diagram



(10.7)

The weight of this diagram (considered as a Hugenholtz diagram) is

$$W_T^H(i) = (i!)^{-2} \tag{10.8}$$

and we associate with it the following scalar factor and operator

$$d_T^H \equiv d_T^H(\rho_k) = \langle B' \cdots B^k | \hat{t}_k | B_1 \cdots B_k \rangle_{id} \tag{10.9}$$

$$\hat{D}_T^X \equiv \hat{D}_T^X(\rho_k) := d_T^X(\rho_k)\, N\Big[\prod_{i=1}^{k} (X_{B'i}^{\dagger} X_{B_i})\Big], \quad X = H, \tag{10.10}$$

where

$$\rho_k = \{ B' \cdots B^k \,;\, B_1 \cdots B_k \} \tag{10.11}$$

is the relevant index set, so that again

$$\hat{T}_k = W_T^X \sum_{\rho_k} \hat{D}_T^X(\rho_k) \,, \quad X = H. \tag{10.12}$$

When employing the Goldstone diagrams we shall represent $\hat{T}_i$ by standard excitation i-th order diagrams as in (5.76) or, simply, again by the diagrams of type (7). In this case

$$W_T^G(i) \equiv W_T(i) = (i!)^{-1}, \tag{10.13}$$

$$d_T^G \equiv d_T^G(\rho_k) = \langle B' \cdots B^k | \hat{t}_k | B_1 \cdots B_k \rangle, \tag{10.14}$$

while Eqs. (10) and (12) hold again with $X = G$.

In order to represent diagrammatically the cluster expansion operator $\exp(\hat{T})$,

$$e^{\hat{T}} = \sum_{n=0}^{\infty} (n!)^{-1} \hat{T}^n = 1 + \hat{T} + \tfrac{1}{2}\hat{T}^2 + \cdots, \tag{10.15}$$

we also have to represent products of cluster components arising from powers of $\hat{T}$. Since cluster components mutually commute, Eq. (4), we can represent a general product

$$\prod_{i=1}^{k} (n_i!)^{-1} \hat{T}_i^{n_i} \tag{10.16}$$

by k disconnected groups, each containing n_i disconnected $\hat{T}_i$ diagrams. In view of the commutation properties, Eq. (4), the left right order of the individual $\hat{T}_i$ diagrams is immaterial and we thus place them along a hypothetical vertical axis. Since these disconnected parts may then be arbitrarily interchanged, the weight of such a <u>multiple</u> (or M for short) <u>diagram</u> is

$$W_M^X = \prod_{i=1}^{k} (n_i!)^{-1} \left[w_T^X(i) \right]^{n_i} , \qquad X = G \text{ or } H. \tag{10.17}$$

We thus obtain a correct numerical factor arising from the exponential expansion (15) when substituting from Eq. (2) for the cluster operator $\hat{T}$. We can therefore define the scalar factor (and the operator) associated with such an M-diagram as the product of corresponding factors for its disconnected parts

$$d_M^X = \prod_i d_T^X(\rho_i), \tag{10.18}$$

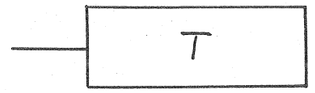
$$\hat{D}_M^X = \prod_i \hat{D}_T^X(\rho_i), \tag{10.19}$$

where the product extends over all $\sum_{i=1}^{k} n_i = n$ disconnected components. We then get

$$e^{\hat{T}} = \sum_M w_M^X \hat{D}_M^X , \tag{10.20}$$

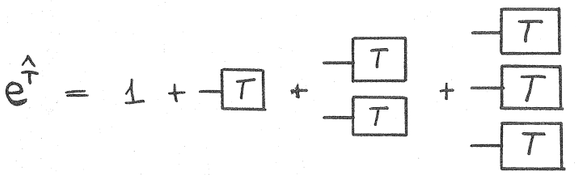
with the sum extending over all possible M diagrams, including an empty diagram representing unity by definition*)

Thus, in the expansion (15), the unity is represented by an empty M-diagram, the operator $\hat{T}$ by all connected (i.e., single component) M diagrams, $\frac{1}{2}T^2$ by all possible two-component M-diagrams, etc. Representing schematically the set of all possible T-diagrams (i.e., the set of all one-component or connected M-diagrams) by a rectangle



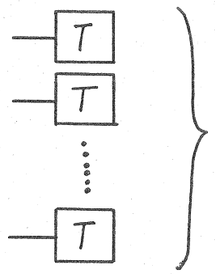
, (10.21)

we can represent schematically the cluster expansion, Eq. (15), as follows



, (10.22)

where



(10.23)

represents all possible M diagrams with n components. The line attached to the left of each rectangle indicates schematically that all fermion lines, of which there is at least one, must extend to the left of their incident vertices. We also assume that the weights of individual $\hat{T}_i$ diagrams, appearing in (21), are associated with their explicit representation, such as (7), so that the weight of (23) equals $(n!)^{-1}$ [since we have n identical and disconnected parts] in agreement with Eq. (15).

*) In the diagrammatic representation we shall represent this empty M diagram, containing no $\hat{T}_i$ components, by its algebraic value, i.e. 1.

10.3 Connected and disconnected clusters

In order to illustrate the significance of connected and disconnected cluster components we compare the cluster expansion, Eq. (1), with an ordinary CI expansion of the exact wavefunction Ψ_0. We shall write the latter in terms of the excitation operators defined in Sec. 5.4, Eqs. (5.70), so that

$$|\Psi_0\rangle = \sum_{r=0}^{N} \hat{C}_r |\Phi_0\rangle , \qquad (10.24)$$

where

$$\hat{C}_0 = \hat{1} \quad \text{(intermediate normalization)}, \qquad (10.25)$$

while $\hat{C}_r$, $r \geq 1$, are defined by Eqs. (5.70). The antisymmetrized matrix elements $\langle B' \cdots B' | \hat{C}_r | B_1 \cdots B_r \rangle_A$ represent the coefficients of appropriate spin-orbital configurations, Eq. (5.72), and are antisymmetric (independently) in both sets of indices, Eq. (5.73).

When we now compare the expansion (24),

$$|\Psi_0\rangle = (1 + \hat{C}_1 + \hat{C}_2 + \hat{C}_3 + \cdots + \hat{C}_N)|\Phi_0\rangle , \qquad (10.26)$$

with the cluster expansion (1),

$$|\Psi_0\rangle = (1 + \hat{T} + \tfrac{1}{2}\hat{T}^2 + \tfrac{1}{3!}\hat{T}^3 + \cdots)|\Phi_0\rangle , \qquad (10.27)$$

where

$$\hat{T} = \sum_{i=1}^{N} \hat{T}_i , \qquad (10.2)$$

we find immediately the following relationships

$$\begin{aligned} \hat{C}_1 &= \hat{T}_1 \\ \hat{C}_2 &= \hat{T}_2 + \tfrac{1}{2}\hat{T}_1^2 \\ \hat{C}_3 &= \hat{T}_3 + \hat{T}_1\hat{T}_2 + \tfrac{1}{3!}\hat{T}_1^3 \\ \hat{C}_4 &= \hat{T}_4 + \tfrac{1}{2}\hat{T}_2^2 + \hat{T}_1\hat{T}_3 + \tfrac{1}{2}\hat{T}_1^2\hat{T}_2 + \tfrac{1}{4!}\hat{T}_1^4 \end{aligned} \qquad (10.28)$$

or, generally

$$\hat{C}_r = \hat{T}_r + \sum_{\mathcal{P}_r} \prod_{j=1}^{p} (n_j!)^{-1} \hat{T}_j^{n_j} , \qquad (10.29)$$

where the sum extends over all (nontrivial) partitions $\mathcal{P}_r$ of r, $r = \sum_{j=1}^{p} j n_j$, $1 \leq n_j \leq r$, $p < r$. The first term on the right hand side, $\hat{T}_r$, represents then the r-times excited connected cluster component while the remaining terms represent disconnected r-times excited clusters.

Should we choose Brueckner (maximum overlap) orbitals as the basis of our formalism (cf., Sec. 8.4) then the mono-excited component vanishes

$$T_1 = 0 , \qquad (10.30)$$

and formulas (28) considerably simplify, namely

$$\begin{aligned} \hat{C}_1 &= \hat{T}_1 = 0 \\ \hat{C}_2 &= \hat{T}_2 \\ \hat{C}_3 &= \hat{T}_3 \\ \hat{C}_4 &= \hat{T}_4 + \tfrac{1}{2}\hat{T}_2^2 . \end{aligned} \qquad (10.31)$$

Thus, the first disconnected cluster contribution will occur in the tetraexcited component.

For molecular systems with a non-metallic closed shell ground state the $\hat{T}_1$ component is usually small even when Hartree-Fock orbitals are employed (cf., also the remark at the end of Sec. 8.4), and by far the largest contribution to the correlation energy is given by the doubly excited component $\hat{C}_2 \approx \hat{T}_2$. This is approximately true even for highly delocalized (i.e., metallic in character) systems, such as the π-electron system of benzene, if the number of electrons is not too large (say, less than 10 or 15).

The next in importance is usually the tetraexcited component $\hat{C}_4$. It was shown by Sinanoğlu that in this case the major contribution comes from the disconnected component, so that we can write formally that

$$\hat{T}_4 \ll \tfrac{1}{2}\hat{T}_2^2 . \qquad (10.32)$$

This approximation is indeed very well satisfied for the ground states of medium-sized non-metallic systems and holds also reasonably well for metallic ones (in the benzene model, the T_4 component represents less than 3% of the total quadruply excited component when measured by corresponding probability amplitudes). Consequently, the approximation

$$\hat{T} \approx \hat{T}_2 \tag{10.33}$$

represents usually an excellent approximation for closed-shell molecular ground states of medium sized systems. This approximation leads to the so-called coupled-pair many electron theory (CPMET) or CCD (coupled cluster approach with doubles) approximations. When also $\hat{T}_1$ and $\hat{T}_3$ clusters are considered we then talk about various versions of the extended CPMET (ECPMET). We examine these approaches in greater detail in Sec. 10.5 as special cases of the general coupled cluster approach (CCA).

10.4. Connected Cluster Theorem

We shall now demonstrate that when exploiting directly the cluster ansatz, Eq. (1), in the Schrödinger equation for the non-degenerate ground state,

$$\hat{H} |\Psi_0\rangle = E_0 |\Psi_0\rangle \ , \tag{10.34}$$

all disconnected diagram contributions cancel out. Rewriting the Schrödinger equation in the N-product form

$$\hat{H}_N |\Psi_0\rangle = \Delta E |\Psi_0\rangle \ , \tag{10.35}$$

where

$$\hat{H}_N = \hat{F}_N + \hat{V}_N \tag{10.36}$$

and

$$\Delta E = E_0 - \langle \hat{H} \rangle \tag{10.37}$$

represents the correlation energy when we take the restricted Hartree-Fock solution $|\Phi_0\rangle$ as the Fermi vacuum of our hole-particle formalism, so that

$$\langle \hat{H} \rangle \equiv \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = E_0^{HF} . \tag{10.38}$$

Using the cluster ansatz for $|\Psi_0\rangle$, Eq. (1), and substituting into (35) we get

$$\hat{H}_N e^{\hat{T}} |\Phi_0\rangle = \Delta E \, e^{\hat{T}} |\Phi_0\rangle . \tag{10.39}$$

We can now use our diagrammatic technique to evaluate the left-hand-side of this equation to any desired order in $\hat{T}$. We note that both $\hat{H}_N$ and $e^{\hat{T}}$ (or any of its connected or disconnected component) are in the N-product form so that no internal fermion lines, which are incident with the vertices of the same s-vertex (i.e., F, V, T or M), are permitted in the R-diagrams.

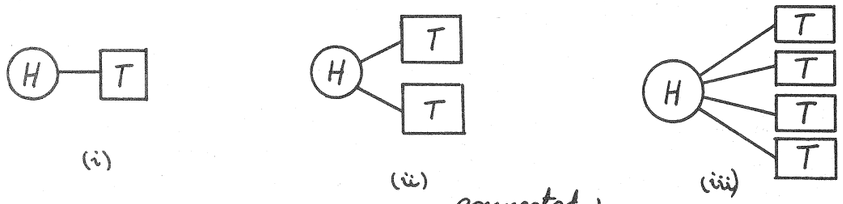
Let us carry out the evaluation of the left-hand-side of (39) in a general schematic form in order to find out its overall structure. We thus represent schematically the F or V vertices as

(H) (10.40)

with attached fermion lines (we shall suppress arrows for simplicity) as needed. Clearly, there may be at most four fermion lines attached to this vertex.

We shall similarly represent the cluster expansion operator $\exp(\hat{T})$ as in Eq. (22), where each component (21) represents an

arbitrary connected cluster component, whose fermion lines must always extend to the left since $\hat{T}$ contains only particle-hole creation operators. We further make a convention that each explicitly shown line indicates at least one fermion line extending in that direction (cf., Eqs. (21), (22)). When no lines are shown, as in (40), they may extend either to the right or to the left. Similarly, a line interconnecting the representatives (21) and (40) indicates that there is at least one internal line between these components, i.e., that there is at least one contraction between the operators $\hat{F}_N$ or $\hat{V}_N$ and $\hat{T}_i$. Thus, for example,

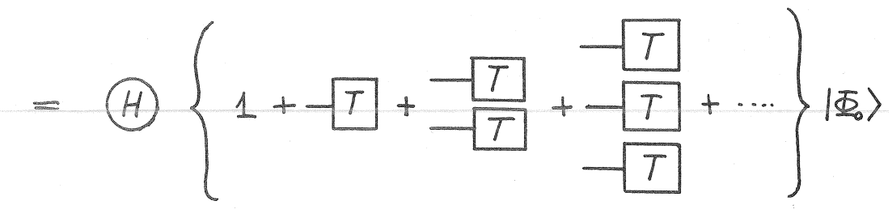


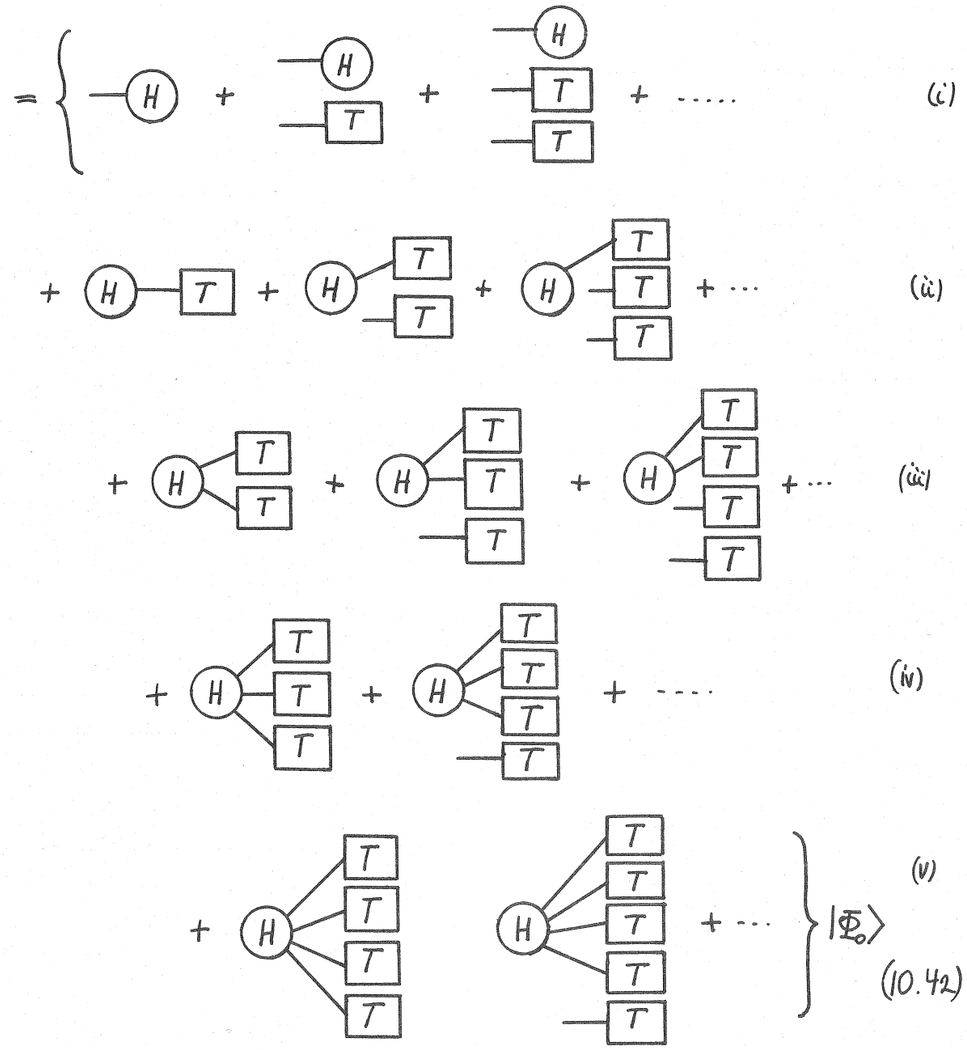
(10.41)

diagram (41i) represents any connected R-diagram containing up to four internal lines between the F or V vertex and a single $\hat{T}$ vertex. Similarly, diagram (41ii) represents any connected R-diagram in which there is at least one internal fermion line between one F or one V vertex and each of the two $\hat{T}$ vertices forming a two-component M vertex. Clearly, at most four T components can be attached to the operator vertex as shown by the diagram (41iii).

Using these conventions we can schematically represent the left hand side of Eq. (39) by the following R-diagrams, which result when the generalized Wick theorem is applied:

$\hat{H}_N e^{\hat{T}} |\Phi_0\rangle =$



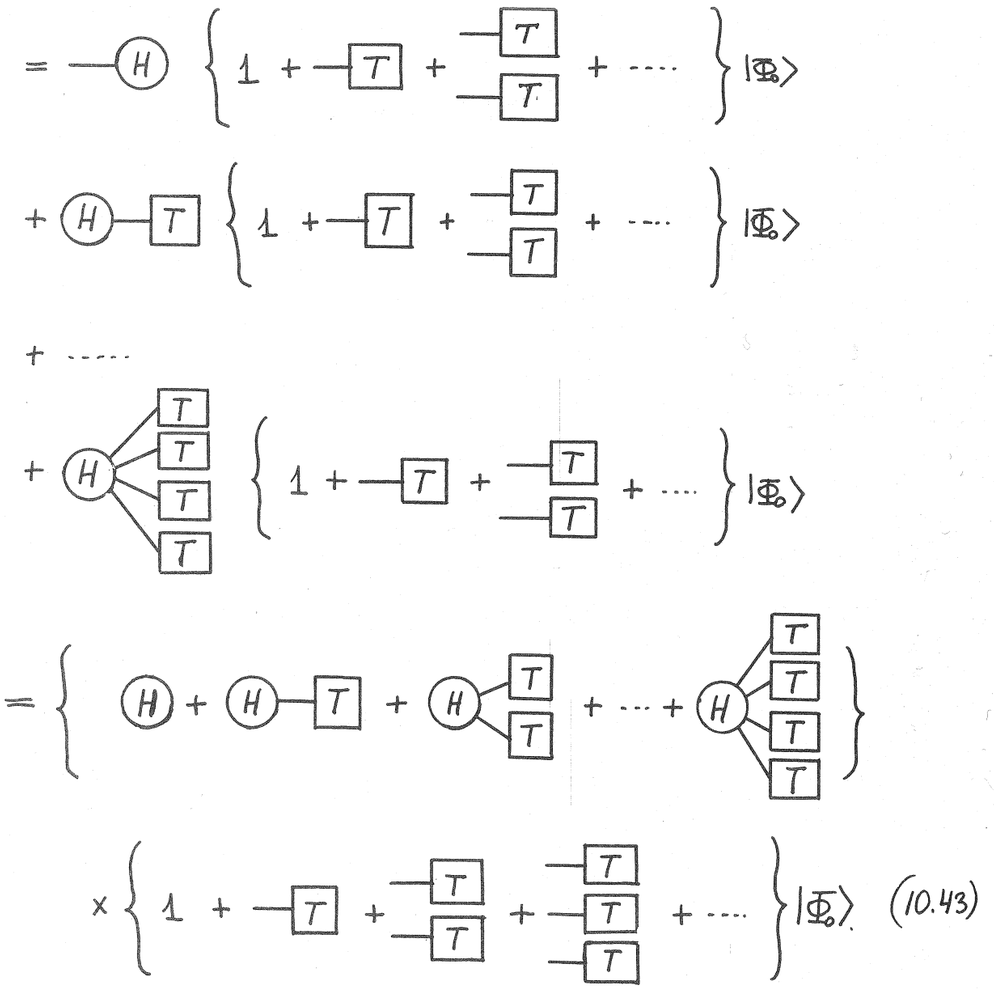


(10.42)

where the first line, Eq. (42i), corresponds to the terms with no contractions, the second line, Eq. (42ii), to the terms with at least one contraction beween the operator part (40) and a single (connected) T component, etc., until the last row, Eq. (42v), contains the terms in which each fermion line of the V vertex is joined with a different T component. Since the individual disconnected components in each term on the right hand side of Eq. (42) mutually commute, (note that each disconnected component contains an even number of hole-particle creation operators, and all these operators which are associated with a given term are within the

same N-product), we can rewrite the above result as follows

$\hat{H}_N e^{\hat{T}} |\Phi_0\rangle =$



The last factor in Eq. (43) is clearly the cluster expansion of the wavefunction $|\Psi_0\rangle$, Eq. (22), while the first factor on the right hand side of the last Eq. (43) represents all possible connected R-diagrams which can be formed from one F or one V diagram and one M-diagram, i.e. zero, one, two, ..., four $\hat{T}$ diagrams. Moreover, both factors mutually commute since all disconnected components mutually commute.

Algebraically we can write this result as

follows

$$(\hat{H}_N e^{\hat{T}}) |\Phi_0\rangle = (\hat{H}_N e^{\hat{T}})_C \, e^{\hat{T}} |\Phi_0\rangle = e^{\hat{T}} (\hat{H}_N e^{T})_C |\Phi_0\rangle, \quad (10.44)$$

where the subscript C indicates that only the component given by connected diagrams is to be considered. We shall refer to the factorization property as expressed by Eq. (44) as a <u>connected cluster theorem</u>. Since the connected component $(\hat{H}_N e^{\hat{T}})_C$ can contain at most four T components, as we have seen above, we can also write the first factor in the last Eq. (43) in the following more explicit form

$$(\hat{H}_N e^{\hat{T}})_C = \sum_{n=0}^{4} \{(n!)^{-1} \; \hat{H}_N \hat{T}^n\}_C . \quad (10.45)$$

Exploiting now the connected cluster theorem, Eq. (44), in the Schrödinger equation (39), we obtain after multiplying each side with $e^{-\hat{T}}$

$$(\hat{H}_N e^{\hat{T}})_C |\Phi_0\rangle = \Delta E |\Phi_0\rangle, \quad (10.46)$$

which is the desired relationship which can be regarded as the connected cluster form of the Schrödinger Equation (34) or (35). Projecting this relationship on the low-excited subspaces we obtain required system of equations which determine the selected cluster components. In fact, we see immediately that energy is given by the vacuum component of Eq. (46). Indeed, projecting Eq. (46) onto the Fermi vacuum $|\Phi_0\rangle$ we get

$$\Delta E = \langle (\hat{H}_N e^{\hat{T}})_C \rangle , \quad (10.47)$$

while projecting on any excited configuration states $|\Phi_K^{(k)}\rangle$, which a orthogonal to $|\Phi_0\rangle$, we obtain the <u>energy independent</u> equations for the cluster components

$$\langle \Phi_K^{(k)} | (\hat{H}_N e^{\hat{T}})_C |\Phi_0\rangle = 0 . \quad (10.48)$$

Since the matrix elements of the connected cluster components represent the coefficients at the linearly independent configurations in an appropriate k-times excited subspace, we obtain the same number of equations of the type (48) as there are unknown cluster components. We shall also see that neglecting higher excited cluster components $\hat{T}_i$, $i > k_0$ than certain prescribed excitation level k_0 will automatically decouple the chain of equations (48) without any additional approximations.

Before deriving an explicit form of these equations for important approximations (i.e., CPMET and ECPMET equations), we shall quickly comment on one alternative way of deriving the connected cluster form of the Schrödinger equation, Eq. (46), which has been used by several authors (Kümmel et al., Harris and Monkhorst). These authors pre-multiply Eq. (39) by $e^{-\hat{T}}$ before the factorization, Eq. (44), has been achieved, obtaining

$$\left(e^{-\hat{T}}\,\hat{H}_N\,e^{\hat{T}}\right)|\Phi_0\rangle = \Delta E\,|\Phi_0\rangle \qquad (10.49)$$

and evaluate the operator product on the left hand side of Eq. (49) using a well known expansion*)

$$e^{-\hat{A}}\,\hat{B}\,e^{\hat{A}} = \sum_{n=0}^{\infty} [[\cdots[[\hat{B},\underbrace{A]_-,A]_-\cdots,A]_-,A]_-}_{n}\,(n!)^{-1}, \qquad (10.50)$$

obtaining

$$e^{-\hat{T}}\,\hat{H}_N\,e^{\hat{T}}\,|\Phi_0\rangle =$$

$$= \Big(\hat{H}_N + [\hat{H}_N,\hat{T}]_- + \tfrac{1}{2}[[\hat{H}_N,\hat{T}],\hat{T}]_- + \tfrac{1}{3!}[[[\hat{H}_N,\hat{T}]_-,\hat{T}]_-,\hat{T}]_-$$
$$+ \tfrac{1}{4!}[[[[\hat{H}_N,\hat{T}]_-,\hat{T}]_-,\hat{T}]_-,\hat{T}]_-\Big)|\Phi_0\rangle. \qquad (10.51)$$

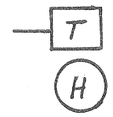
*) This result is often erroneously referred to as the Campbell-Baker-Hausdorff (CBH) theorem. In fact, it represents only one lemma needed in the proof of the CBH formula.

The termination of this expansion occurs due to the fact that $\hat{H}_N$ contains at most two-particle operators, i.e. the products of at most four creation and annihilation operators. Since each commutator $[\hat{A},\hat{B}]_-$ "eliminates" one creation or annihilation operator from each of the arguments $\hat{A}$ and $\hat{B}$, replacing them by the Kronecker delta symbol in view of the anticommutation relations (1.96) or (4.15), the termination of the expansion (50) will occur after four such commutators are performed with $\hat{B} = \hat{H}_N$. However, each term on the right hand side of (51) (except the first term which does not contain any commutators) contains disconnected diagram contributions which mutually cancel out. We can see this easily by carrying out the evaluation of these commutators diagrammatically. This will also provide a link with our earlier results.

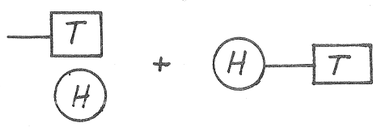
Consider, thus, a commutator of $\hat{H}_N$ and $\hat{T}$,

$$[\hat{H}_N,\hat{T}]_- = \hat{H}_N\hat{T} - \hat{T}\hat{H}_N. \qquad (10.52)$$

The last term $\hat{T}\hat{H}_N$ is automatically in the N-product form since $\hat{T}$ contains only particle-hole creation operators. This is also immediately apparent from our diagrammatic representation since all the fermion lines issuing from a T-vertex extend to the left and cannot be contracted with any of the lines of the operator vertices. We can thus represent the last term in Eq. (52) as

$\hat{T}\hat{H}_N =$  . (10.53)

The first term on the right hand side of Eq. (52) is identical with the terms evaluated earlier (cf. e.g. Eq. (43)) so that we can represent it as follows

$\hat{H}_N\hat{T} =$  , (10.54)

where again the first and second terms on the right hand side represent the uncontracted and contracted (with one up to four contractions) terms, which result by applying generalized Wick theorem. The disconnected terms will thus cancel out and we can write that

$$[\hat{H}_N, \hat{T}]_- = \text{(H)—[T]} = (H_N T)_C . \qquad (10.55)$$

Similarly we obtain

$$\tfrac{1}{2}[[H_N, T]_-, T]_- = \tfrac{1}{2}[\text{(H)—[T]}, \text{—[T]}]_-$$

$$= \text{(H) with [T] [T]} = (\tfrac{1}{2}\hat{H}_N \hat{T}^2)_C , \qquad (10.56)$$

and, generally

$$(n!)^{-1}[[\hat{H}_N, \hat{T}]_-, \ldots, \hat{T}]_- = \{(n!)^{-1}\hat{H}_N \hat{T}^n\}_C \qquad (10.57)$$

so that

$$e^{-\hat{T}} \hat{H}_N e^{\hat{T}} = \sum_{n=0}^{4} \{(n!)^{-1} H_N T^n\}_C , \qquad (10.58)$$

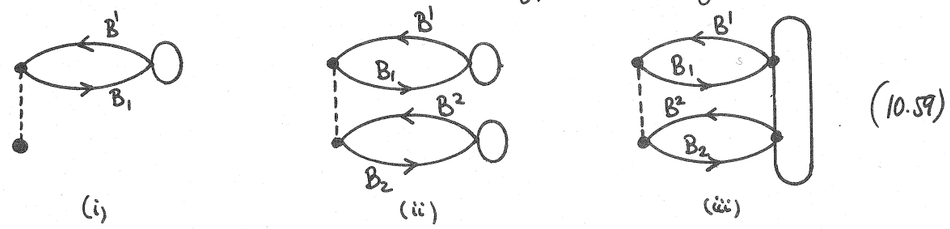
which is identical with previously obtained result, Eqs. (45,46).

We can thus summarize the general procedure for obtaining the explicit form of the CCA equations: We simply consider all <u>connected</u> R-diagrams which can be formed from one interaction (F or V) vertex and from none, one, etc., up to four T_i vertices, representing the cluster components considered, and which posses $2k$ external lines.

10.5. Spin-orbital form of CCA Equations

We shall now find an explicit form of the basic equations of the CCA for the most important case of pair clusters and will also briefly outline the derivation for an extended case involving mono- and tri-excited clusters.

The correlation energy is always given by Eq. (47), which involves at most doubly excited clusters. Thus, the relevant Brandow-type R-diagrams are



(10.59)

which give

$$\Delta E = \sum_{B_1 B'} \langle B_1 | \hat{f} | B' \rangle \langle B' | t_1 | B_1 \rangle$$

$$+ \tfrac{1}{2} \sum_{B_1 B_2 B' B^2} \langle B_1 B_2 | \hat{v} | B' B^2 \rangle_{AS} \left(\langle B' | t_1 | B_1 \rangle \langle B^2 | t_1 | B_2 \rangle \right.$$

$$\left. + \tfrac{1}{2} \langle B' B^2 | \hat{t}_2 | B_1 B_2 \rangle_{AS} \right) . \qquad (10.60)$$

Consequently, at most doubly excited clusters contribute <u>directly</u> to the energy, the higher excited ones contributing only indirectly via their interaction with mono-excited and pair clusters. In fact, even the mono-excited clusters contribute primarily indirectly (Brueckner orbitals are used their contribution vanishes altogether) since the matrix elements $\langle B' | t_1 | B_1 \rangle$ are usually very small and occur quadratically with the two-electron integrals. Further, when HF orbitals are used,

$$\langle B_1 | \hat{f} | B' \rangle = 0 , \qquad (10.60')$$

so that the one-particle contribution, diagram (59 i),

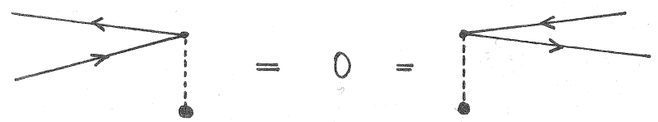
vanishes. Thus, for all practical purposes, the correlation energy is given by the expression

$$\Delta E \approx \tfrac{1}{4} \sum_{B_1 B_2 B' B'} \langle B_1 B_2 | \hat{v} | B' B^2 \rangle_{\mathcal{A}} \langle B' B^2 | \hat{t}_2 | B_1 B_2 \rangle_{\mathcal{A}} \qquad (10.61)$$

$$= \tfrac{1}{2} \sum_{B_1 B_2 B' B^2} \langle B_1 B_2 | \hat{v} | B' B^2 \rangle \left(\langle B' B^2 | t_2 | B_1 B_2 \rangle - \langle B' B^2 | v | B_2 B_1 \rangle \right),$$

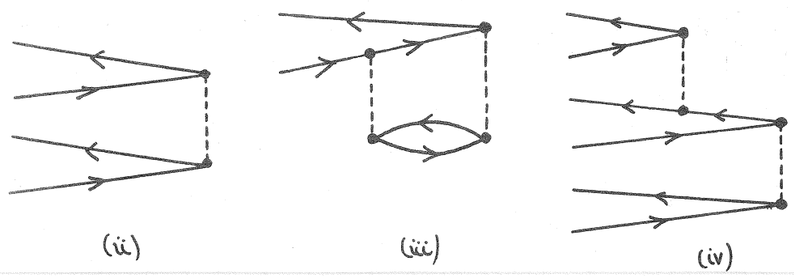
assuming that the HF orbitals are employed. The second expression (61) corresponds to the Goldstone type R-diagrams.

We now briefly examine the relative importance of various cluster components for the correlation energy. Assuming that the reference state $|\Phi_0\rangle$ is the RHF state, so that the off-diagonal f- matrix elements between unoccupied and occupied spinorbitals vanish [Eq. (8.30) and diagram (62i)], we see immediately that the T_2 clusters appear already in the first order perturbation theory contribution to the wavefunction, diagram (62ii), while the T_1 and T_3 clusters occur for the first time only in the second order contribution as illustrated by the typical wavefunction diagrams (62 iii) and (62 iv), respectively.



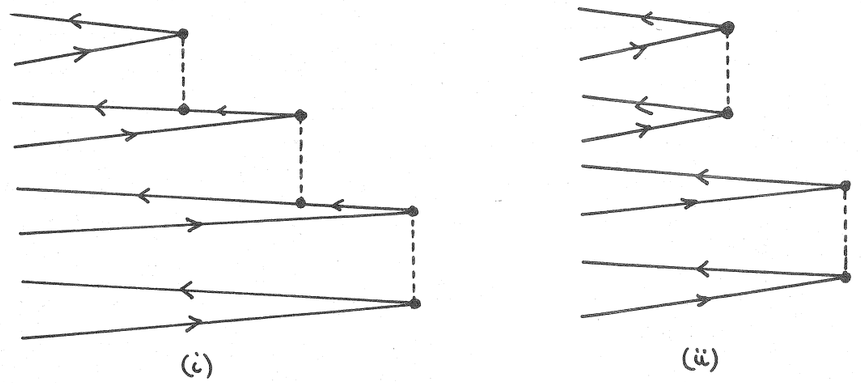
(i)

(10.62)



(ii) (iii) (iv)

For tetraexcited clusters, the connected component T_4 contributes for the first time in the third order of perturbation theory while the disconnected component $\frac{1}{2} T_2^2$, Eq. (31), contributes already in the second order as illustrated by the following typical diagrams:



(i) (ii)

(10.63)

We can thus summarize the orders of perturbation theory in which a given wavefunction contribution occurs for the first time as follows:

Table 10.I.
The lowest order of perturbation theory in which a given cluster contributes for the first time in the wavefunction.

Cluster	Order
T_1	2
T_2	1
T_3	2
T_4	3
$\frac{1}{2} T_2^2$	2

These simple considerations show clearly the great importance of the T_2 clusters: their connected contribution, corresponding to the following doubly excited state contribution, [cf., Eq. (6)],

$$\hat{T}_2 |\Phi_0\rangle = \sum_{\substack{B'<B^2 \\ B_1<B_2}} \langle B' B^2 | \hat{t}_2 | B_1 B_2 \rangle_{\mathcal{A}} \left| \begin{matrix} B' & B^2 \\ B_1 & B_2 \end{matrix} \right\rangle, \qquad (10.64)$$

yields usually well over 90% of the correlation energy for small systems (say, with less than ten electrons) and over 80% even for medium sized systems unless their ground state $|\Phi_0\rangle$ is quasi-degenerate with some of the doubly excited configuration(s). However, even for the Be atom, in which the well known $2s^2 - 2p^2$ quasi-degeneracy is very strong, the doubly excited contribution yields well over 90% of the correlation energy.

The next important contribution according to the classification given in Table I comes from the T_1, T_3 and $\frac{1}{2}T_2^2$ clusters. The T_1 contribution depends very significantly on the orbitals employed [it exactly vanishes for the Brueckner orbitals, cf. Sec. 8.4 and Eq. (30)]. With the RHF orbitals this contribution is usually non-negligible though rather small. The T_3 contribution has been usually heuristically neglected since it requires a considerable computational effort in comparison with the pair contribution and only recently its effect is being studied in more detail. It seems that this contribution can play quite a significant rôle in some cases even though in all the cases examined so far it is always considerably less important than the disconnected tetraexcited contribution from the $\frac{1}{2}T_2^2$ term. We can thus say that while the T_1, T_3 and $\frac{1}{2}T_2^2$ clusters all appear for the first time in the second order perturbation theory contribution to the wavefunction, and are much less important for the correlation energy than the T_2 clusters appearing already in the first order, their relative importance is roughly proportional to their numbers.

The relative importance of disconnected tetraexcited clusters when compared with the connected ones (which occur for the first time in the third order contribution to the wavefunction, cf. Table I) is easily understandable in simple physical terms. The disconnected terms may be associated with the correlations between two independent electronic pairs while the connected terms correspond to the true four-"particle" correlations. This was also the key idea which lead to earlier independent pair theories (Sinanoğlu) and to various forms of approximate coupled pair theories (Meyer, Kutzelnigg, et. al.). The approximation (33) was then fully implemented in the CPMET formalism given by Čížek as a special case of his general CCA formalism.

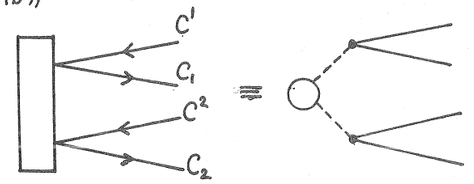
We shall now investigate the equations which determine the cluster components, Eq. (48). We first assume that the approximation (33) holds and obtain the CPMET equations. Subsequently we shall briefly consider the contributions due to $\hat{T}_1$ and $\hat{T}_3$ operators leading to the ECPMET formalism. In the subsequent sections we shall employ the spin-orbital formalism and only later we will present the spin-adapted form of the theory.

10.5.1. CPMET Equations (Spin-orbital Form).

Consider the pair approximation, Eq. (33), and find explicit equations determining the cluster components $\langle B'B^2|\hat{t}_2|B_1B_2\rangle_d$, Eq. (64). A compact form of these equations is given by Eq. (48) with $k=2$. These equations are obtained when projecting the general CCA equations (46) onto the doubly excited subspace spanned by configurations $|\Phi_K^{(2)}\rangle$,

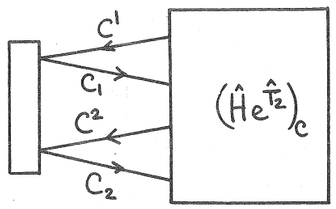
$$|\Phi_K^{(2)}\rangle = \left|\begin{matrix} C' & C^2 \\ C_1 & C_2 \end{matrix}\right\rangle . \qquad (10.65)$$

To find the desired explicit form of these equations we shall employ the diagrammatic approach as outlined in Chap. V. Representing bra state associated with the configuration (65) by the following Brandow-type diagram (cf. Dgms(5.76))



(10.66)

to which we assign the scalar factor 1, we can schematically represent a general $\bar{R}$-diagram contributing to Eq. (48) with $|\Phi_K^{(k)}\rangle$ given by Eq. (33) as follows

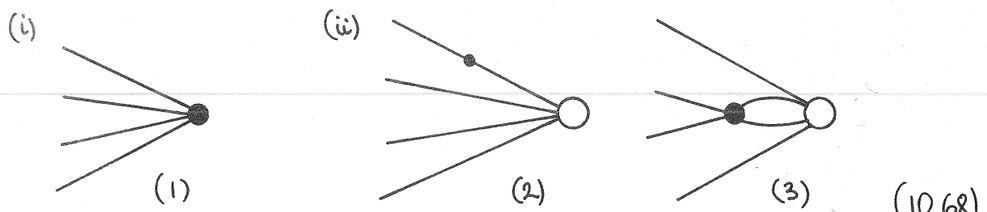


(10.67)

where the right hand side rectangle represents a connected Brandow-type R-diagram of the kind represented schematically by Dgms (40) or (41) with four external (two hole and two particle) fermion lines. These fermion lines carry the fixed labels characterizing the configuration (65) onto which we project. It also follows immediately that the R-diagrams of the type (40) or (41) with four external lines can only be formed from one H vertex and (i) none, (ii) one or (iii) two T_2 vertices [Dgm (7) with $i=2$].

To obtain these Brandow R-diagrams we can employ the same procedure which we used in constructing the third order perturbation theory diagrams in Sec. 9.5.1. In summary, this procedure involves the following steps: (i) Construct all Hugenholtz R-skeletons containing relevant vertices but ignoring the orientation of fermion lines. (ii) Introduce the orientation on fermion lines in all possible but distinct ways. (iii) Represent each Hugenholtz skeleton by one Goldstone representative and, finally, (iv) label all fermion lines with appropriate labels in all permissible ways.

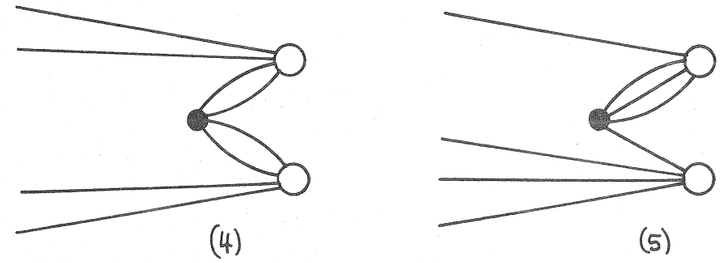
In the case considered, the first step yields the following five "pre-skeletons" [we can use the same reasoning as in Sec. 9.5.1 following Eq. (9.119)]:



(10.68)

(10.68)

(iii)

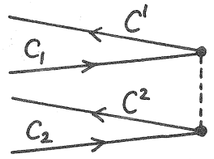


The fermion line orientation can now be introduced in only one way in skeleton (1), in two different ways in skeletons (2), (4) and (5) and in three different ways in skeleton (3). Thus, the second step will yield ten Hugenholtz skeletons, which are prototypes of the Goldstone representatives given below, Dgms. (69) [these skeletons may be easily obtained by deleting all labels and by merging the Goldstone vertices into the Hugenholtz ones in Dgms. (69)]. This step is straightforward to carry out even though it is not unambigous since any Goldstone version may be chosen. The reader should convince himself that the final result is independent of this choice.

In the final step we must recall that the external lines carry fixed labels. Consequently, each set of equivalent external lines can be labeled arbitrarily but each set of n non-equivalent external lines must be labeled in all possible $n!$ distinct ways. This is conveniently done by using generally indexed labels and by considering all permutations of the indices. In the case considered we have at most two non-equivalent external lines in each skeleton, so that we can simply use the indices κ and $\bar{\kappa}$, say, with $\kappa = 1, 2$ and $\bar{\kappa} = 3 - \kappa$; i.e. $\bar{1} = 2$ and $\bar{2} = 1$. The internal lines are then labeled with appropriate (i.e., hole or particle) summation labels. We thus obtain eighteen Brandow-type R-diagrams for the three types mentioned above, which are conveniently represented by the following

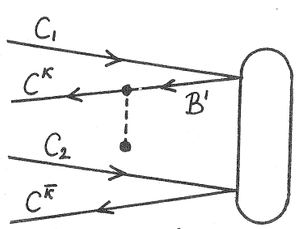
ten diagrams with $\kappa, \lambda = 1, 2$.

(i)

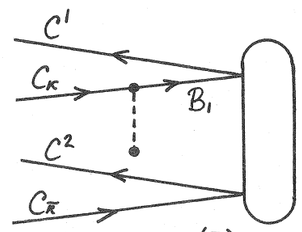


(1)

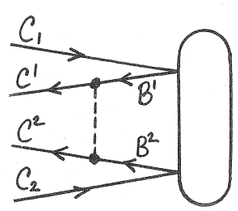
(ii)



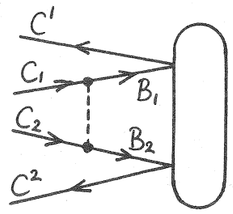
(2)



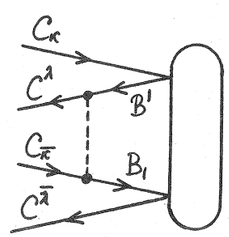
(3)



(4)



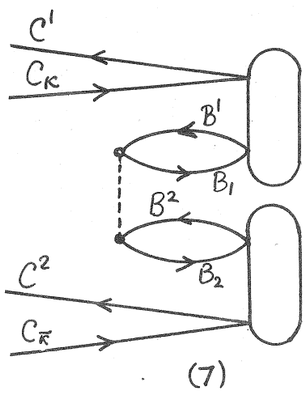
(5)



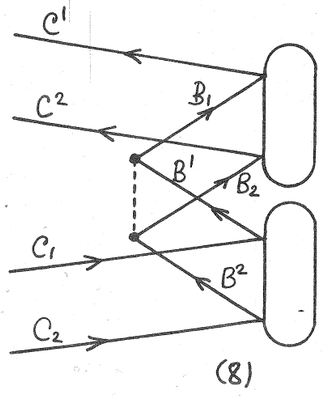
(6)

(10.69)

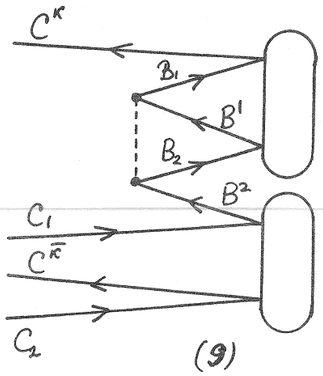
(iii)



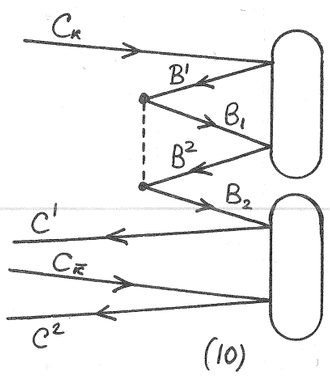
(7)



(8)



(9)



(10)

The Hugenholtz weights w_H, the number of internal hole lines l and the number of closed loops l are easily determined for each R-diagram type (69) and are listed in Table II.

Table 10.II.

Hugenholtz weights w_H, number of internal hole lines l, number of closed loops l, and sign factor f due to the projection for each CPMET R-diagram type, Dgms. (69).

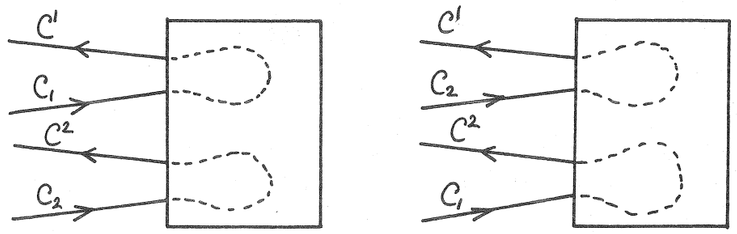
	Diagram	w_H	h	l	f
(i)	1	1	0	0	1
(ii)	2	1	0	0	$(-1)^{\kappa+1}$
	3	1	1	0	$(-1)^{\kappa+1}$
	4	1/2	0	0	1
	5	1/2	2	0	1
	6	1	1	0	$(-1)^{\kappa+\lambda}$
(iii)	7	1	2	2	$(-1)^{\kappa+1}$
	8	1/4	2	0	1
	9	1/2	2	0	$(-1)^{\kappa+1}$
	10	1/2	2	0	$(-1)^{\kappa+1}$

Designating now the contributions from the absolute [Dgms (69i)], linear [Dgms (69ii)] and quadratic [Dgms (69iii)] terms in the T_2 clusters to the equation (48) with $|\Phi_\kappa^{(k)}\rangle$ given by Eq. (65) by the symbol $\Lambda^{(j)}(C'C^2; C_1C_2)$ where $j = 0, 1$ and 2, respectively, we can write Eq. (48) as

$$\sum_{j=0}^{2} \Lambda^{(j)}(C'C^2; C_1C_2) = 0 . \qquad (10.70)$$

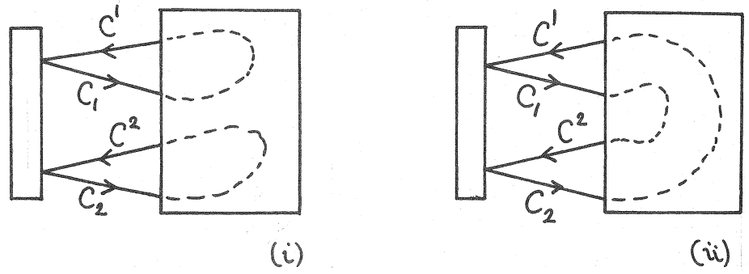
An explicit form of the individual terms in this equation can now be immediately found by evaluating the $\bar{R}$-diagrams (67) with the right hand side rectangle representing individual R-diagrams (69). Each of the seventeen R-diagrams (69); $\kappa, \lambda = 1, 2$, belongs to one of the following two types [using a schematic representation

of oriented paths as in (5.97)]



(i) (ii) (10.71)

These types can be characterized as diagrams with canonical (i) and non-canonical labeling of open paths [cf. Dgm (5.97)]. The $\bar{R}$-diagrams for the projected components, Dgm (67), have thus the following general form for the canonically and non-canonically labeled $\bar{R}$ diagrams (71):



(i) (ii) (10.72)

and their value is given by the scalar factor of the corresponding R diagram (69), represented in (72) by the right hand side block, times the phase factor (-1) in the case of the diagram (ii), Dgm (72). This additional phase factor f is also given in Table II for each diagram (69). It is easy to see that, generally, the phase factor f resulting from a projection onto the fully antisymmetric canonical states (5.68) is given by the parity of the permutation $P(i \to p_i)$ assuming that the i-th open path of the unprojected diagram carries the labels C^i and C_{p_i}.

Using this simple phase rule we can easily find an explicit form of the terms $\Lambda^{(i)}$, Eq. (70), by

evaluating the R-diagrams (69), $\kappa,\lambda = 1,2$. This gives

(i) the absolute term:

$$\Lambda^{(0)}(C'C^2; C_1C_2) = \langle C'C^2|\hat{v}|C_1C_2\rangle_{\mathcal{A}} \tag{10.73i}$$

(ii) the linear term

$$\Lambda^{(1)}(C'C^2; C_1C_2) = \sum_{\kappa=1}^{2}(-1)^{\kappa+1}\Big[\sum_{B'}\langle C^\kappa|f|B'\rangle\langle B'C^{\bar\kappa}|\hat{t}_2|C_1C_2\rangle_{\mathcal{A}}$$

$$-\sum_{B_1}\langle B_1|\hat{f}^\dagger|C_\kappa\rangle\langle C'C^2|\hat{t}_2|B_1C_{\bar\kappa}\rangle_{\mathcal{A}}\Big]$$

$$+\tfrac{1}{2}\sum_{B'B^2}\langle C'C^2|\hat{v}|B'B^2\rangle_{\mathcal{A}}\langle B'B^2|\hat{t}_2|C_1C_2\rangle_{\mathcal{A}} \tag{10.73ii}$$

$$+\tfrac{1}{2}\sum_{B_1B_2}\langle B_1B_2|\hat{v}|C_1C_2\rangle_{\mathcal{A}}\langle C'C^2|\hat{t}_2|B_1B_2\rangle_{\mathcal{A}}$$

$$-\sum_{\kappa,\lambda=1}^{2}(-1)^{\kappa+\lambda}\sum_{B'B_1}\langle B_1C^\lambda|\hat{v}|C_{\bar\kappa}B'\rangle_{\mathcal{A}}\langle B'C^{\bar\lambda}|\hat{t}_2|C_\kappa B_1\rangle_{\mathcal{A}}$$

(iii) the quadratic term

$$\Lambda^{(2)}(C'C^2; C_1C_2) = \sum_{B'B^2B_1B_2}\langle B_1B_2|\hat{v}|B'B^2\rangle_{\mathcal{A}}$$

$$\times\Big\{\sum_{\kappa=1}^{2}(-1)^{\kappa+1}\Big(\langle C'B'|\hat{t}_2|C_\kappa B_1\rangle_{\mathcal{A}}\langle B^2C^2|\hat{t}_2|B_2C_{\bar\kappa}\rangle_{\mathcal{A}} \tag{10.73iii}$$

$$+\tfrac{1}{2}\langle C^\kappa B'|\hat{t}_2|B_1B_2\rangle_{\mathcal{A}}\langle B^2C^{\bar\kappa}|\hat{t}_2|C_1C_2\rangle_{\mathcal{A}}$$

$$+\tfrac{1}{2}\langle B'B^2|\hat{t}_2|C_\kappa B_1\rangle_{\mathcal{A}}\langle C'C^2|\hat{t}_2|B_2C_{\bar\kappa}\rangle_{\mathcal{A}}\Big)$$

$$+\tfrac{1}{4}\langle C'C^2|\hat{t}_2|B_1B_2\rangle_{\mathcal{A}}\langle B'B^2|\hat{t}_2|C_1C_2\rangle_{\mathcal{A}}\Big\}.$$

The five terms on the right hand side of Eq. (73iii) are given by diagrams (2) through (6), Dgms. (69) while the four terms on the right hand side of Eq. (73iii) correspond to diagrams (7), (9), (10) and

(8), Dgms (69), respectively.

The CPMET equations determining the T_2 cluster components are thus non-linear algebraic equations, which are independent of energy. Arranging the $\hat{F}_2$ matrix elements in some arbitrary but fixed order and designating them simply as the unknowns x_j, the system (70) has the form

$$a_i + \sum_j b_{ij} x_j + \sum_{j \le k} c_{ijk} x_j x_k = 0, \qquad (10.74)$$

where the absolute, linear and quadratic coefficients a_i, b_{ij} and c_{ijk}, respectively, are given by the matrix elements of the Hamiltonian $\hat{H}_N$,

$$a_i = \langle i | \hat{H}_N | \Phi_0 \rangle$$

$$b_{ij} = \langle i | \hat{H}_N | j \rangle \qquad (10.75)$$

$$c_{ijk} = \langle i | \hat{H}_N | j,k \rangle$$

where $|i\rangle$ or $|j\rangle$ designate appropriate biexcited configurations, Eq. (65), and $|j,k\rangle = |j\rangle|k\rangle$.

Neglecting completely the inter-pair interaction terms, Eq. (73iii), we obtain a linear system of equations. This approximation has been termed as the linear CPMET (L-CPMET) or as L-CCD approximation. We shall examine its relationship with a corresponding doubly excited CI approximation, which employs the same information about the Hamiltonian, in Sec. 10.6.

10.5.2. ECPMET Equations (Spin-orbital Form).

We now briefly outline an extension of the CPMET to a more general case in which also T_1 and T_3 clusters are explicitly considered. In order to get an idea about the importance of contributions from various disconnected terms, which appear in the cluster expansion

(1) [cf. e.g. Eqs. (28)], we list in Table III the order of perturbation theory in which they contribute for the first time to the wavefunction and to the energy [assuming that we use a HF reference state so that Eq. (60') holds]. In the last two columns of this table we then list the excitation level of the cluster and the range of excitation levels in which the cluster can give a connected contribution through a single two-particle interaction.

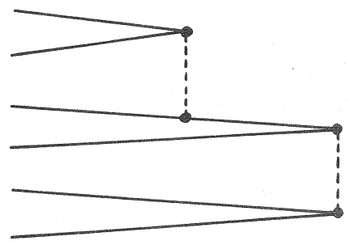
Table 10.III.

Perturbation theory orders $o(W)$ and $o(E)$ in which various connected and disconnected clusters contribute for the first time to the wavefunction and to the energy, respectively, their excitation levels and excitation level ranges for connected contributions to the left hand side of Eq. (46).

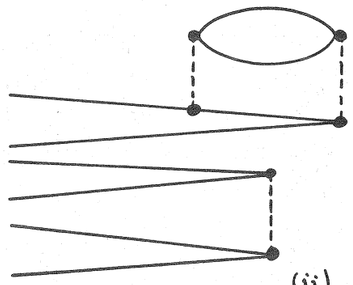
Cluster	$o(W)$	$o(E)$	Excitation level	Excitation range
T_1	2	4	1	0-2
T_2	1	2	2	0-3
T_3	2	4	3	1-4
T_4	3	5	4	2-5
T_5	4	7	5	3-6
T_1^2	4	5	2	0-2
T_1T_2	3	5	3	1-3
T_1T_3	4	6	4	2-4
T_1T_4	5	8	5	3-5
T_2^2	2	4	4	2-4
T_2T_3	3	6	5	3-5
T_2T_4, T_3^2	4	7	6	4-6
T_3T_4	5	9	7	5-7
T_1^3	6	8	3	1-2
$T_1T_2^2$	4	7	5	3-4
$T_3T_2^2$	4	8	7	5-6
T_2^3	3	6	6	4-5
T_2^4	4	8	8	6

Thus, considering the T_1 and T_3 clusters in addition to the T_2 clusters, we do not have to add any new disconnected clusters if we limit ourselves

to those which contribute for the first time in the fourth order to the energy. In this respect we find a reversed behavior for the triexcited clusters to that found for the tetraexcited ones, Eq.(32) and Dgms.(63), as indicated by typical diagrams shown below



(i)



(ii)

(10.76)

Thus, while connected triexcited clusters $\hat{T}_3$ (e.g., Dgm (76i)) contribute already in the second order to the wavefunction, the disconnected $\hat{T}_1\hat{T}_2$ clusters (e.g., Dgm (76ii)) contribute only in the third order. We shall thus ignore any contribution from disconnected clusters except those appearing in the CPMET (i.e., T_2^2 clusters). This approximation seems to be well justified in view of the existing studies considering triexcited clusters.

We shall now give the spin-orbital form of the ECPMET equations in a diagrammatic form. The algebraic form can be easily written down by evaluating pertinent diagrams using the same rules as in the preceding section. We write the general ECPMET equations in the form

$$\sum_j \Lambda^{(k,j)} (C'C^2\cdots C^k ; C_1C_2\cdots C_k) = 0 , \qquad (10.77)$$

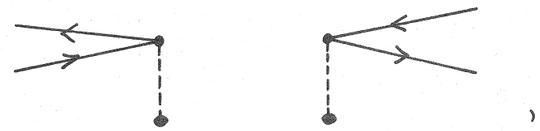
where $\Lambda^{(k,j)}$ designates the component originating from clusters of type j in the projection of Eq.(46) onto the k-times excited subspace, i.e. in Eq. (48), and the summation extends over all connected and disconnected cluster components considered.

Thus, for the pair clusters ($k=2$), in addition to the components arising in the CPMET approach (see Sec. 10.5.1, Eq. (70)), which are now designated as follows

$$\begin{aligned} \Lambda^{(2,0)} &:= \Lambda^{(0)} \\ \Lambda^{(2,2)} &:= \Lambda^{(1)} \\ \Lambda^{(2,[2,2])} &:= \Lambda^{(2)} , \end{aligned} \qquad (10.78)$$

we also consider the components $\Lambda^{(2,1)}$ and $\Lambda^{(2,3)}$, which give the interaction of pair clusters with the mono and tri-excited clusters T_1 and T_3, respectively, while neglecting interactions with other disconnected clusters, such as T_1^2, T_1^3, T_1T_2, etc., for the reasons explained above.

We now list the Brandow-type R-diagrams for the ECPMET approach assuming again that the HF reference state is used. Should this not be the case, the additional R-diagrams, involving the following one-particle components



(i) (ii)

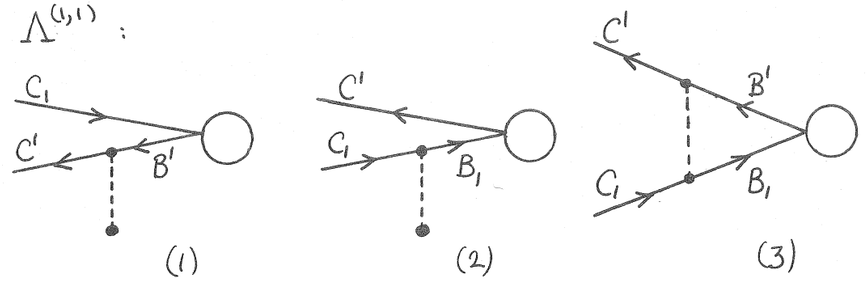
(10.79)

must be also included. The s-o labels for triexcited components are understood to be evaluated modulo 3 (i.e. $4=1$, $5=2$ and $6=3$).

(i) Mono-excited components

$\Lambda^{(1,0)} = 0$ (Diagram (79i) in general)

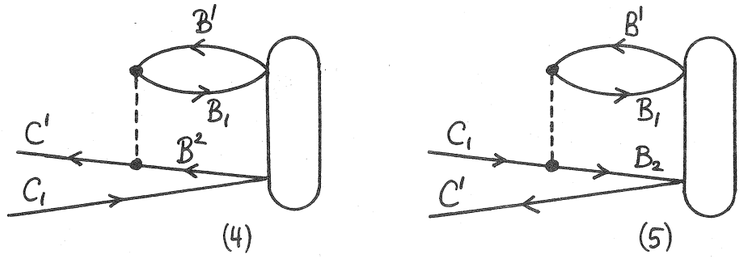
$\Lambda^{(1,1)}$:



(1) (2) (3)

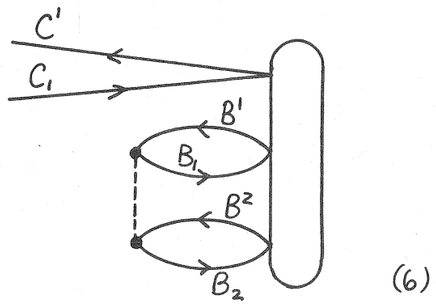
(10.80i)

$\Lambda^{(1,2)}$:



(10.80i)

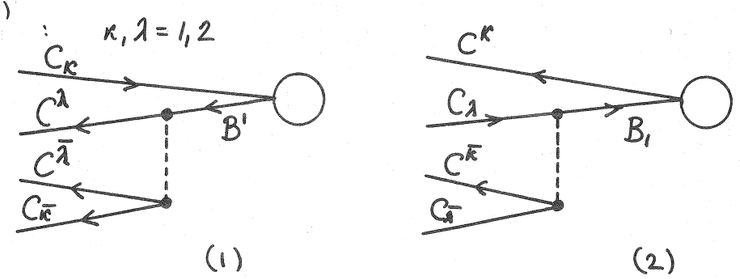
$\Lambda^{(1,3)}$:



(ii) Bi-excited components

$\Lambda^{(2,0)} \equiv \Lambda^{(0)}$ (Sec. 10.5.1, Dgm (1), (69i))

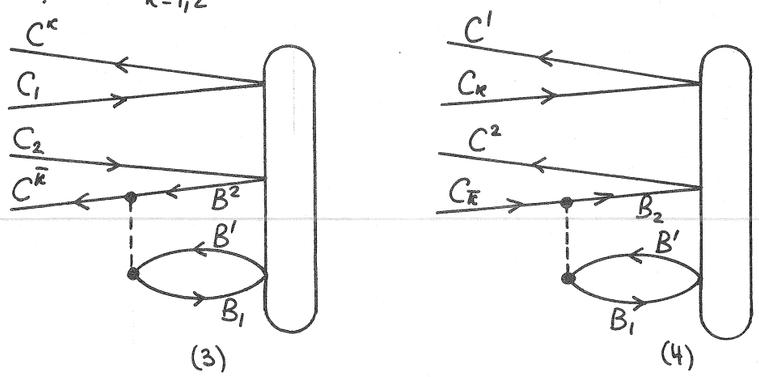
$\Lambda^{(2,1)}$: $\kappa, \lambda = 1,2$



$\Lambda^{(2,2)} \equiv \Lambda^{(1)}$ (Sec. 10.5.1, Dgms. (2)-(6), (69ii))

(10.80ii)

$\Lambda^{(2,3)}$: $\kappa = 1,2$

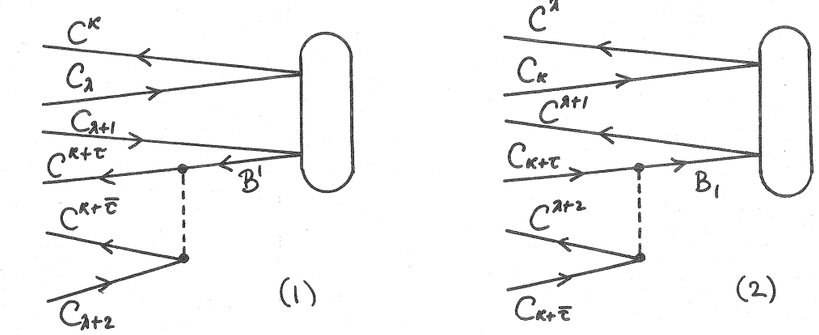


$\Lambda^{(2,[2,2])} \equiv \Lambda^{(2)}$ (Sec 10.5.1, Dgms. (7)-(10), (69iii))

(iii) Tri-excited components

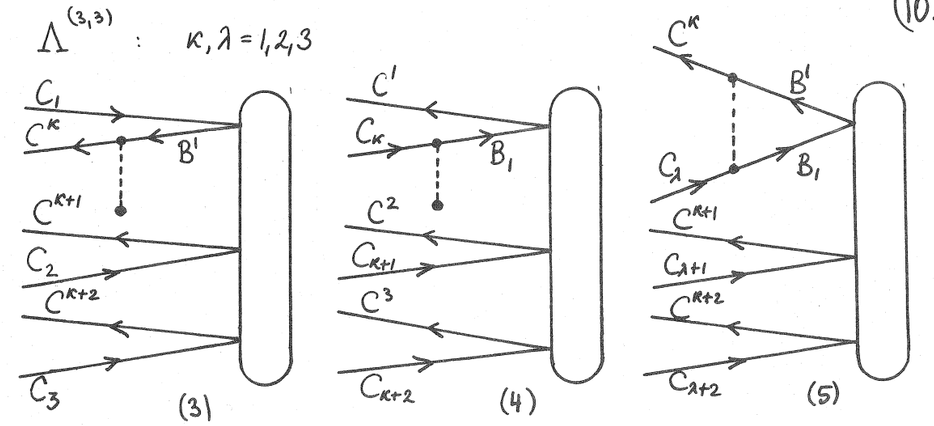
$\Lambda^{(3,0)} = \Lambda^{(3,1)} \equiv 0$

$\Lambda^{(3,2)}$: $\kappa, \lambda = 1,2,3$; $\tau = 1,2$



(10.80iii)

$\Lambda^{(3,3)}$: $\kappa, \lambda = 1,2,3$



When determining the Hugenholtz weights of

these diagrams we must remember that C_i and C^i are fixed labels while B_i and B^i are free summation labels. Thus, all Hugenholtz weighs for diagrams (80) are one except for diagrams (4), (5) of Dgm (80i) and (3),(4) of Dgm. (80ii) having the weight 1/2 and the diagram (6), Dgm. (80i), having the weight 1/4.

10.6. Relationship with other approaches

We shall now investigate a relationship of the CCA, and in particular of the CPMET and the L-CPMET approaches, with both variational (i.e., configuration interaction) and perturbative approaches. This comparison provides a better understanding of the approximations made in the CCA and gives us some clues as to the limits of the applicability of the CCA.

10.6.1. Variational approach (CI)

Let us first compare the truncation schemes which lead to limited CI approaches on the one hand and to the corresponding cluster approaches on the other hand. We shall examine the most important case of the D-CI (CI limited to doubly excited configurations) and DQ-CI (CI limited to doubly and quadruply excited configurations) approaches and the corresponding L-CPMET and CPMET approaches. We thus write the CI expansion of $|\Psi_0\rangle$ in the intermediate normalization, Eq. (26),

$$|\Psi_0\rangle = \sum_{i=0}^{N} \hat{C}_i |\Phi_0\rangle , \tag{10.81}$$

where

$$\hat{C}_0 = 1 \tag{10.82i}$$

and

$$\hat{C}_k |\Phi_0\rangle = \sum_K c_K^{(k)} |\Phi_K^{(k)}\rangle , \tag{10.82ii}$$

with the sum in Eq. (82ii) extending over all ordered k-times excited orthonormal configurations $|\Phi_K^{(k)}\rangle$. In fact, these may be either the spin-orbital configurations, Eq. (5.68), or chosen spin-adapted configurations which will be used in the orthogonally spin-adapted approaches (cf., Secs.).

As it is well known, the variational CI equations can be simply obtained by projecting the Schrödinger equation, Eq. (34) or (35), onto the pertinent configuration states $|\Phi_K^{(k)}\rangle$. Thus, in general

$$\langle \Phi_K^{(k)} | \hat{H}_N | \Psi_0 \rangle = \Delta E \langle \Phi_K^{(k)} | \Psi_0 \rangle = \Delta E \, c_K^{(k)} . \tag{10.83}$$

Since $\hat{H}_N$ contains at most two-body operators, we find

$$\langle \Phi_K^{(k)} | \hat{H}_N | \Psi_0 \rangle = \sum_\ell \sum_L c_L^{(\ell)} \langle \Phi_K^{(k)} | \hat{H}_N | \Phi_L^{(\ell)} \rangle$$

$$= \sum_{\ell=k-2}^{k+2} \sum_L c_L^{(\ell)} H_{KL}^{(k\ell)} , \tag{10.84}$$

where we defined

$$c_L^{(\ell)} = 0 \qquad \text{if} \quad \ell < 0 \tag{10.85}$$

and

$$H_{KL}^{(k,\ell)} := \langle \Phi_K^{(k)} | \hat{H}_N | \Phi_L^{(\ell)} \rangle ,$$
$$H_{00}^{(0,0)} := \langle \Phi_0 | \hat{H}_N | \Phi_0 \rangle \equiv 0 . \tag{10.86}$$

Thus

$$\sum_{\ell=k-2}^{k+2} \sum_L c_L^{(\ell)} H_{KL}^{(\ell L)} = \Delta E \, c_K^{(k)} , \tag{10.87}$$

which yields the well known chain of CI equations

$$\sum_{\ell=k-2}^{k+2} \sum_L \left(H_{KL}^{(k\ell)} - \Delta E \, \delta_{KL}^{(k\ell)} \right) c_L^{(\ell)} = 0 , \tag{10.88}$$

$$\ell = 0, 1, \ldots, N .$$

To simplify our considerations we assume that the odd-number-of-times excited clusters, particularly the $T_1 = C_1$ and C_3 clusters can be neglected, so that the chain (87) or (88) takes on the form

$$\sum_L H_{0L}^{(0,2)} C_L^{(2)} = \Delta E$$

$$H_{K0}^{(2,0)} + \sum_L H_{KL}^{(2,2)} C_L^{(2)} + \sum_L H_{KL}^{(2,4)} C_L^{(4)} = \Delta E\, C_K^{(2)} \qquad (10.89)$$

$$\sum_L H_{KL}^{(4,2)} C_L^{(2)} + \sum_L H_{KL}^{(4,4)} C_L^{(4)} + \sum H_{KL}^{(4,6)} C_L^{(6)} = \Delta E\, C_K^{(4)}, \text{ etc.}$$

Each Eq.(89) for $k>0$ represents a system of equations for all possible ordered configurations K of the given excitation order. When all equations for all $k = 0, \dots, N(=2n)$ are considered, we obtain the full CI approximation (representing the exact solution for a chosen model space). To obtain a limited CI (L-CI) approximation we have to de-couple the above chain. For a given k this is done in the L-CI approach by neglecting all $H_{KL}^{(k,l)}$ matrix elements with $l > k$ (in our case, $l = k+2$).

Thus, for example, the D-CI equations are obtained neglecting all $H_{KL}^{(2,4)}$ matrix elements in the second equation (89) so that we get

$$\sum_L H_{0L}^{(0,2)} C_L^{(2)} = \Delta E$$

$$H_{K0}^{(2,0)} + \sum_L H_{KL}^{(2,2)} C_L^{(2)} = \Delta E\, C_K^{(2)} \qquad (10.90)$$

or simply

$$\sum_L a_L^* C_L = \Delta E$$

$$a_K + \sum_L b_{KL} C_L = \Delta E\, C_K\,, \qquad (10.91)$$

designating

$$a_K := H_{K0}^{(2,0)}\,, \quad b_{KL} := H_{KL}^{(2,2)} \quad \text{and} \quad C_K := C_K^{(2)}\,, \qquad (10.92)$$

where $K = 1, \dots, M$ and M designates the number of biexcited configurations. We can write (91) in the matrix form as

$$\begin{bmatrix} 0 & a_1^* & a_2^* & \cdots & a_M^* \\ a_1 & b_{11} & b_{12} & \cdots & b_{1M} \\ a_2 & b_{21} & b_{22} & \cdots & b_{2M} \\ \vdots & \vdots & \vdots & & \vdots \\ a_M & b_{M1} & b_{M2} & \cdots & b_{MM} \end{bmatrix} \begin{bmatrix} 1 \\ C_1 \\ C_2 \\ \vdots \\ C_M \end{bmatrix} = \Delta E \begin{bmatrix} 1 \\ C_1 \\ C_2 \\ \vdots \\ C_M \end{bmatrix} \qquad (10.93)$$

or, shortly

$$\begin{bmatrix} 0 & \mathbb{a}^\dagger \\ \mathbb{a} & \mathbb{b} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbb{C} \end{bmatrix} = \Delta E \begin{bmatrix} 1 \\ \mathbb{C} \end{bmatrix} \qquad (10.94)$$

where

$$\mathbb{a} = [a_K]_{M\times 1}\,, \quad \mathbb{b} = [b_{KL}]_{M\times M}\,, \quad \mathbb{C} = [C_K]_{M\times 1}\,. \qquad (10.95)$$

In a similar way we can truncate at the $k=4$ level by neglecting all $H_{KL}^{(4,6)}$ matrix elements obtaining the DQ-CI equations which will have the following general form

$$\begin{bmatrix} 0 & \mathbb{a}^\dagger & \mathbb{0} \\ \mathbb{a} & \mathbb{b} & \mathbb{p}^\dagger \\ 0 & \mathbb{p} & \mathbb{q} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbb{C}^{(2)} \\ \mathbb{C}^{(4)} \end{bmatrix} = \Delta E \begin{bmatrix} 1 \\ \mathbb{C}^{(2)} \\ \mathbb{C}^{(4)} \end{bmatrix}, \qquad (10.96)$$

where $\mathbb{p}$ and $\mathbb{q}$ are matrices of tetra-bi and tetra-tetra-excited CI matrix elements.

Let us now compare this decoupling with that used in the CPMET approach. Using the $\hat{T} \approx \hat{T}_2$ approximation, we see that we obtain the same system of equations as that given by Eq. (89) in the CI case, except that all higher than doubly excited configurations are now given by the $\hat{T}_2^m$ disconnected components, i.e.

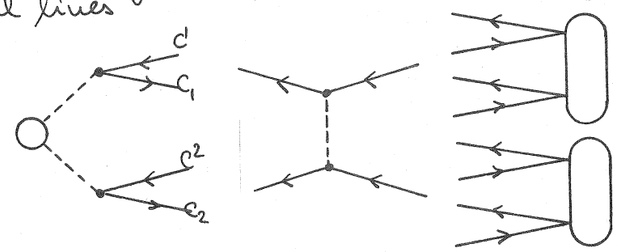
$$\hat{C}_{2j}\,|\Phi_0\rangle \approx (j!)^{-1}\ \hat{T}_2^{\,j}\,|\Phi_0\rangle \qquad (10.97)$$

since $\hat{T}_j = 0$ unless $j=2$ in this case. Since also $\hat{C}_2 = \hat{T}_2$ in this case, we can also identify the corresponding matrix elements

$$c_K^{(2)} = t_K^{(2)}. \qquad (10.98)$$

Consequently, the matrix elements constituting the matrices a and b, Eqs. (95), are identical with those of the CI approach (cf., also, Eqs. (75)).

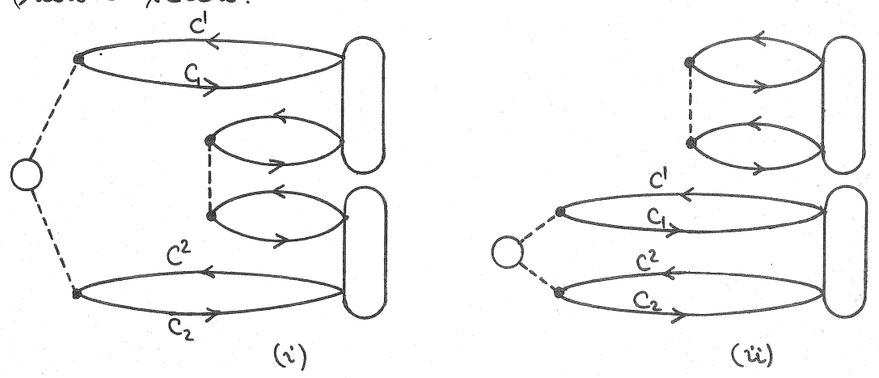
We shall thus examine the last term on the left hand side of the second Eq. (89) involving the $H_{KL}^{(2,4)}$ matrix elements, which had to be neglected in the L-CI approach in order to achieve the desired decoupling of the chain (89). Let us examine the structure of this term using a diagrammatic approach. Since only the two-electronic part of $\hat{H}_N$ can contribute in this case, the value of a general $H_{KL}^{(2,4)}$ matrix element in the CPMET approach is given by all R-diagrams which can be formed from the two T_2 diagrams, one V diagram (or vertex) and one C_2 diagram with which we associate the scalar factor 1 (see Sec. 10.5.1, Dgm (67), and the discussion there concerning the diagrammatic representation of the projections), i.e. all R-diagrams formed from the following components and having no external lines



(10.99)

In all these R-diagrams the fermion lines associated with the interaction V vertex must be clearly connected with the lines issuing

from the T_2 vertices. (This implies in turn that in a non-vanishing $H_{KL}^{(2,4)}$ matrix element the (spin)-orbital labels characterizing the configuration K must form a subset of those characterizing L.) We also see immediately that these R-diagrams can be of two basic types, namely the connected and the disconnected (or linked and unlinked in this case) ones. A typical diagram for each case is shown below.



(10.100)

Contributions from the disconnected diagrams of the type (100 ii) clearly factorize and equal

$$\Delta E \cdot c_K^{(2)}, \qquad (10.101)$$

thus exactly cancelling the right hand side term in the second equation (89). Consequently, the approximation $T \approx T_2$ automatically decouples the chain of equations (89), yielding the energy independent CPMET equations

$$a_K + \sum_L b_{KL}\, t_L^{(2)} + \sum_{L\le M} d_{KLM}\, t_L^{(2)} t_M^{(2)} = 0, \qquad (10.102)$$

the last term originating from connected diagrams of the kind (100 i). The energy is then given by the first equation (89) so that

$$\Delta E_{CPMET} = \sum_K a_K^*\, t_K^{(2)} = a^\dagger t^{(2)}. \qquad (10.103)$$

Thus, the approximation $\hat{T} \approx \hat{T}_2$ of the CPMET approach leads automatically to the decoupling of the chain of Eqs. (89) or (87). Moreover, the most

important effect of tetraexcited configurations is included by introducing the pair-pair interaction non-linear terms. This decoupling procedure is therefore more satisfactory from the physical viewpoint, even though the upper bound property for the energy of the variational approaches is lost. More importantly, the CCA approaches are size consistent as warranted by the exponential ansatz (1) [see, also, Sec. 10.6.2].

We should remark that the CPMET (or corresponding L-CPMET) system of algebraic equations (102) is equivalent with any system obtained by multiplying each equation by a non-vanishing scalar. Thus, in general, $\mathbb{b}$ need not be hermitean and a simple relationship between the absolute and linear coefficients of the CPMET system of equations and the submatrices of the corresponding D-CI matrix need not hold (particularly in certain spin-adapted versions as we shall see later). However, we can always multiply the CPMET Eqs. (102) by appropriate non-vanishing scalars (usually given as simple normalization factors for the configurations $|\Phi_K^{(2)}\rangle$ on which one projects) such that this relationship is restored. Consequently, we can assume without any restriction of the generality of our considerations that the absolute and linear coefficients in the CPMET systems (102) or (74) are identical with corresponding D-CI matrix elements between biexcited states and the ground states and between pairs of biexcited states, respectively [cf, also Eqs. (75)].

When solving the CPMET system (102) or (74) it is often convenient to start with the solution of the corresponding linear system

$$a_K + \sum_L b_{KL} \tilde{t}_L^{(2)} = 0 , \qquad (10.104)$$

defining the L-CPMET approximation with the energy

$$\Delta E_L = a^\dagger \tilde{t} . \qquad (10.105)$$

The L-CPMET energies ΔE_L are usually very close to the full CPMET ones (the difference is usually less than 1% of ΔE) unless the HF ground state is quasi-degenerate with some low-lying doubly-excited configuration. In fact, the L-CPMET energies are usually closer to the exact full-CI energies than the CPMET energies are, since the linear approximation overcorrects the D-CI result for unlinked contributions and this overcorrection is in the direction in which the energy is shifted when also T3 clusters are considered.

In the context of approximate CEPA-type approaches (coupled electron-pair approximation), which represent certain approximations to the CPMET (see later), the L-CPMET is refered to as the CEPA(0) approach. We shall also see a simple perturbation theory interpretation of the L-CPMET in Sec. 10.6.2. Let us therefore examine this approximation in greater detail, particularly its relationship with the corresponding D-CI approach.

Both, the D-CI and the L-CPMET approaches employ the same information about the Hamiltonian as given by the submatrices $\mathbb{a}$ and $\mathbb{b}$ of the D-CI matrix, Eq. (94). However, while the correlation energy is given by the lowest eigenvalue of the D-CI matrix, the L-CPMET energy is obtained by first solving a linear system of energy independent inhomogeneous algebraic equations (104) and, second, by evaluating the scalar product (105), i.e.

$$\mathbb{a} + \mathbb{b}\tilde{\mathbb{t}} = 0 \; ; \quad \Delta E_L = a^\dagger \tilde{t} . \qquad (10.106)$$

This approach will thus break down if $\mathbb{b}$ is singular. This will be the case if one o more eigenvalues of $\mathbb{b}$ will vanish. However, in

such case there exists a state in the doubly excited subspace which is degenerate with the ground state $|\Phi_0\rangle$ so that our basic assumption breaks down and the degenerate version of the theory should be used.

To see this fact more clearly, we consider a unitary transformation v which diagonalizes b,

$$b v = v \beta \quad , \quad v^\dagger v = v v^\dagger = \mathbb{1} \, , \qquad \beta = [\beta_i \delta_{ij}] \, . \tag{10.107}$$

This transformation may be simply regarded as a canonical transformation of the D-CI matrix to the "principle axes" in the doubly-excited subspace. Transforming, thus, correspondingly the D-CI matrix, we find

$$\begin{bmatrix} 1 & 0 \\ 0 & v \end{bmatrix}^\dagger \begin{bmatrix} 0 & a^\dagger \\ a & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & v \end{bmatrix} = \begin{bmatrix} 0 & \alpha^\dagger \\ \alpha & \beta \end{bmatrix} , \tag{10.108}$$

where

$$\alpha = v^\dagger a \, . \tag{10.109}$$

The resulting matrix (108) has obviously the same eigenvalues as the original D-CI matrix.

It is easy to see that also a corresponding L-CPMET problem

$$\alpha + \beta \tau = 0 \, , \quad \Delta\tilde{E}_L = \alpha^\dagger \tau \, , \tag{10.110}$$

yields the same L-CPMET energy as the original one, Eq. (106),

$$\begin{aligned} \Delta\tilde{E}_L = \alpha^\dagger \tau &= (v^\dagger a)^\dagger (-\beta^{-1} \alpha) \\ &= - a^\dagger v \beta^{-1} v^\dagger a = - a^\dagger b^{-1} a \\ &= a^\dagger \tilde{t} = \Delta E_L \, , \end{aligned} \tag{10.111}$$

since Eq.(107) implies that

$$v \beta^{-1} v^\dagger = b^{-1} \tag{10.112}$$

assuming that b is regular. In fact, the energy (111) can be written in the form

$$\Delta E_L = - \sum_{i=1}^{M} \alpha_i^2 / \beta_i \, , \tag{10.113}$$

which reveals immediately the fact that ΔE_L will have a singularity as $\beta_i \to 0$.

In a well behaved problem, the IPM energies of doubly excited states will be much larger than corresponding interaction matrix elements with the ground state so that b will be diagonally dominant, positive definite matrix. Only in quasi-degenerate cases we observe that the lowest eigenvalue of b may approach zero in which case the L-CPMET approximation breaks down.

The main effect of the L-CPMET approach as compared to the D-CI one lies in its size-consistent (or, better yet, size-extensive) character, which is a direct consequence of the cluster ansatz (1). This ansatz transforms multiplicative structures into additive ones and leads to the cancellation of all disconnected clusters. In order to see this cancellation in more detail, let us consider the appropriate moments (or projections) of the Schrödinger equation associated with the D-CI and the L-CPMET approaches.

In the D-CI approach, Eq. (83) can be written in the form

$$\langle \Phi_K^{(k)} | (\hat{H}_N - \Delta E_D)(1 + \hat{C}_2) | \Phi_0 \rangle = 0 \tag{10.114}$$

while the corresponding L-CPMET equations follow from Eq. (46) or (48),

$$\langle \Phi_K^{(k)} | \{ \hat{H}_N (1 + \hat{T}_2) \}_C - \Delta E_L | \Phi_0 \rangle = 0 , \qquad (10.115)$$

where $k \leq 2$ and C designates a connected component. The basic difference between these two systems of equations (114) and (115) lies in the presence of the term

$$\langle \Phi_K^{(2)} | \Delta E_D \hat{C}_2 | \Phi_0 \rangle = \Delta E_D \cdot C_I^{(2)} \qquad (10.116)$$

in the D-CI system (114). This term corresponds to the unlinked diagrams in the perturbation theory language and is therefore responsible for the size-inconsistency of the D-CI approach. In the L-CPMET (or, in fact, any CCA) approach these unlinked terms have been cancelled by the unlinked (and, therefore, disconnected) component

$$\langle \Phi_K^{(2)} | \left(\tfrac{1}{2} H_N \hat{T}_2^2 \right)_{UL} | \Phi_0 \rangle . \qquad (10.117)$$

This cancellation follows immediately from the connected structure of the CPMET (or CCA) equations [cf., Sec. 10.4] and is therefore implicit in the connected cluster form of the Schrödinger equation, Eq. (46).

It must be noted, however, that the cancelled term (116) contains both the EPV and the non-EPV diagrams since the summations over the (spin)-orbital labels in each factor are independent. Therefore, the L-CPMET "overcorrects" the D-CI result and the EPV terms are then again corrected for through the non-linear EPV terms in the CPMET approach. The most important terms in this respect are those represented by diagrams which can be separated over one or two hole lines [i.e., diagrams (8) and (10) of Dgms (69)]. These terms play the most important role when going from the L-CPMET approach to the CPMET and form the basis of various CEPA(k), $k \neq 0$,

approaches.

In this respect it is interesting to compare the D-CI and the L-CPMET approaches for the case of two-electron systems. Obviously, in this special case D-CI yields the exact energy within the model space considered, since it is equivalent to the full-CI approach, while the L-CPMET approach will only yield an approximate energy. This is due to the fact that the latter approach includes a correction for the unlinked tetraexcited clusters (which is implicit in the connected form of the Schrödinger equation (46) as explained above). However, all tetraexcited components are unphysical in this case and thus of the EPV type. These terms are therefore "spoiling" the full-CI wavefunction and would be subsequently cancelled by the non-linear CPMET terms, which are also always of the EPV type in the two-electron case considered, thus restoring the D-CI result. Clearly, the CPMET, and its simple version the L-CPMET, are designed for many-electron rather than two-electron systems. For these systems the number of non-EPV unlinked type contributions will be much larger than the number of similar EPV type terms. Consequently, the corrections for the unlinked tetraexcited terms, which make the D-CI approach size inconsistent, will be much more important than the "overcorrection" due to the inclusion of the EPV terms, which are necessary to achieve the factorization and which are corrected for when the proper "coupling" of pair clusters is considered in the CPMET approach.

(Recently, an explicit relationship between the D-CI and the L-CPMET energies and wavefunctions was found and is described in an attached preprint, Appendix 2).

10.6.2. Perturbative approach

Using our diagrammatic representation of the CCA it is easy to find its relationship with the perturbation theory as developed in Ch. 9. To see this relationship let us first separate the diagonal terms b_{ii} from the remaining terms, since they must clearly contain the differences between one-particle orbital energies forming the denominators of the perturbation theory contributions (we use the form of the CPMET equations as given in (74):

$$-b_{ii}x_i = a_i + \sum_j{}' b_{ij}x_j + \sum_{j\le k} c_{ijk}x_j x_k , \qquad (10.118)$$

where the prime on the first summation symbol indicates the omission of the diagonal terms. Designate, further, the one-particle orbital energy contributions to the diagonal element b_{ii}, given by diagrams (2) and (3) of (69ii) for $B' = C^k$ and $B_1 = C_k$, $k=1,2$, by Ω_i (note that these are the only contributions of these diagrams if canonical HF orbitals are used, since Eq. (8.37) holds and the sum over B' or B_1 reduces to a single term)

$$\Omega_i = \sum_{B'\in K_i} \omega_{B'} - \sum_{B_1\in K_i} \omega_{B_1} , \qquad (10.119)$$

where the sums extend over the particle and hole labels characterizing the i-th configuration K_i, i.e.

$$\Omega_i = \omega_{C^1} + \omega_{C^2} - \omega_{C_1} - \omega_{C_2} \qquad (10.120)$$

when $K_i \equiv \{C^1C^2; C_1C_2\}$. We thus write

$$b_{ii} = \Omega_i + \Delta_i , \qquad (10.121)$$

where Δ_i designates the two-electron part of b_{ii} given by diagrams (4)–(6), Dgms. (69ii).

In order to achieve some flexibility in the choice of the perturbation theory denominators (or, equivalently, in the separation of the Hamiltonian into the unperturbed one and the perturbation) we further define

$$b_{ii} = b_{ii}^{(1)} + b_{ii}^{(2)} \qquad (10.122)$$

with

$$\begin{aligned} b_{ii}^{(1)} &= \Omega_i + (1-\lambda)\Delta_i \\ b_{ii}^{(2)} &= \lambda\Delta_i \end{aligned} \qquad , \lambda\in[0,1]. \qquad (10.123)$$

The $\lambda=1$ limit corresponds to the RSPT developed in Ch. 9, also referred to in this context as the Møller-Plesset perturbation theory when the HF reference state is used. On the other hand, one can build the perturbation theory directly on the CI matrix using the diagonal as the unperturbed Hamiltonian and the off-diagonal part as the perturbation, so that the denominators are given directly by the b_{ii} matrix elements (recall that we use Hamiltonian in the normal product form thus relating all energies relative to the chosen HF reference state). This form of the perturbation theory is known as the Epstein-Nesbet perturbation theory and corresponds to the choice $\lambda=0$ in Eqs. (123). Thus, defining the splitting of the diagonal matrix elements b_{ii} by Eqs. (122) and (123), we are able to go continuously from the Møller-Plesset form ($\lambda=1$) to the Epstein-Nesbet form ($\lambda=0$).

Using now the partitioning (122) in Eq. (118) we can write

$$-b_{ii}^{(1)}x_i = a_i + b_{ii}^{(2)}x_i + \sum_j{}' b_{ij}x_j + \sum_{j\le k} c_{ijk}x_j x_k . \qquad (10.124)$$

Similarly as in the Jacobi algorithm for the iterative solution of systems of linear algebraic

equations, we can write Eq. (124) in a recursive form

$$x_i^{(n+1)} = -\left[b_{ii}^{(1)}\right]^{-1}\left\{a_i + b_{ii}^{(2)} x_i^{(n)} + \sum_j{}' b_{ij} x_j^{(n)} + \sum_{j\le k} c_{ijk} x_j^{(n)} x_k^{(n)}\right\}, \quad (10.125)$$

where the superscript (n) designates the n-th iterate.

In order to start with the chosen reference state, we take vanishing components for the initial approximation,

$$x_i^{(0)} := 0, \quad (10.126)$$

so that

$$x_i^{(1)} = -a_i / b_{ii}^{(1)}, \quad (10.127)$$

and

$$\Delta E_L^{(1)} = -\sum_i |a_i|^2 / b_{ii}^{(1)}, \quad (10.128)$$

obtaining the well known second order perturbation theory result. In subsequent iterations we shall be obtaining higher and higher orders of the perturbation theory. It must be noted, however, that subsequent iterants do not simply correspond to the next perturbation theory order but due to the presence of the non-linear terms in Eq. (125) various orders of perturbation theory become intermixed.

(*) See p. 26a →

In order to obtain a more precise idea about the perturbation theory structure of the CPMET we can carry out the above outlined iterative scheme diagrammatically. When we assume that $\lambda = 1$, we can simply ignore the one-particle diagrams (2) and (3), Dgms (69 ii), [by incorporating them via Eq. (125) into the denominator rule of the perturbation theory], and interpret the remaining diagrams as perturbation theory wavefunction diagrams [cf. Ch. 9]. In this way Eq. (7) can be rewritten in the following manner

(*)

(¶) It is interesting to note that employing the quasidiagonal form of the D-CI matrix, Eq. (108) [i.e., the L-CPMET system, Eqs. (110)], we obtain the whole L-CPMET energy already in the second order, since we can write Eq. (111) in the form

$$\Delta E_L = -\alpha^\dagger \beta^{-1} \alpha = -\sum_{i=1}^{M} \frac{|\alpha_i|^2}{\beta_i}. \quad (10.128a)$$

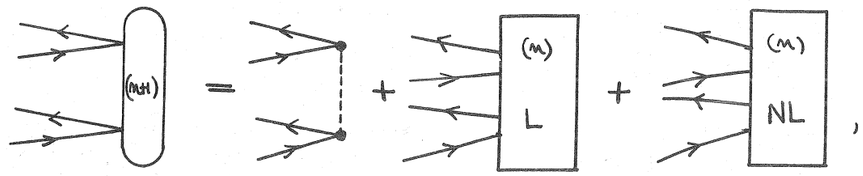
It is also worthwhile to note that the corresponding variational problem, which is characterized by the following secular equation

$$\det \begin{bmatrix} -\Delta E_D & \alpha^\dagger \\ \alpha & \beta - \Delta E_D \cdot \mathbb{1} \end{bmatrix} = 0, \quad (10.128b)$$

can be expressed in the following form

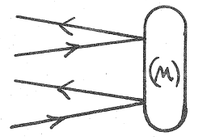
$$\Delta E_D = -\sum_{i=1}^{M} \frac{|\alpha_i|^2}{\beta_i - \Delta E_D}. \quad (10.128c)$$

We thus see that the corresponding D-CI energy ΔE_D is given by the second-order Brillouin-Wigner perturbation theory, rather than by the Rayleigh-Schrödinger perturbation theory yielding the L-CPMET energy ΔE_L, Eq. (128a). This is also consistent with the fact that the L-CPMET approach is size-extensive while the D-CI approach is not.



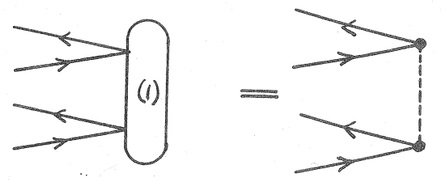
(10.129)

where the factor $[-b_{ii}^{(1)}]^{-1}$ is now implicitly included in the right hand side diagrams as the leftmost denominator factor and the second and third terms on the right hand side are obtained from the linear (L) diagrams (4)–(6), Dgms. (69ii), and non-linear (NL) diagrams (7)–(10), Dgms. (69iii), respectively, by replacing general t_2 vertices with the m-th iterate doubly-excited contribution to the exact wavefunction represented by the diagram (we ignore explicit line labeling):



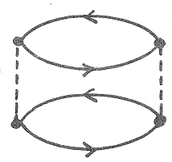
(10.130)

Thus, taking again Eq. (126) as the starting approximation, we obtain for the first order iterate from Eq. (129)



(10.131)

which yields the first iterate to the energy, $\Delta E_L^{(1)}$, when "substituted" into the energy diagram (59iii)

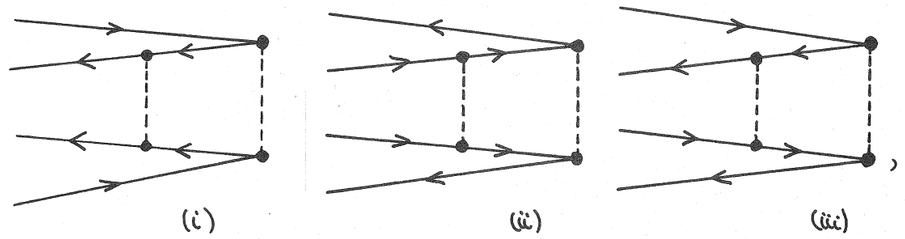
$\Delta E_L^{(1)}$ = 

(10.132)

We thus recover the second order perturbation theory energy (see Ch. 9) assuming that we are using

the HF reference state, corresponding to Eq. (128).

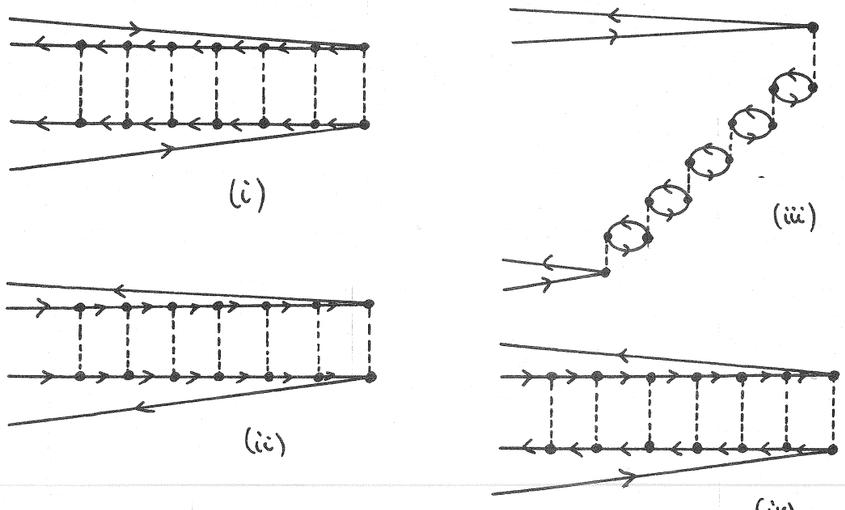
Consider now, at least very schematically, the next iteration. From the linear terms in Eq. (129) we obtain the following contributions



(10.133)

which yield the third order energy contribution as given by Brandow-type diagrams (9.123). Thus, with the HF reference state, already the first non-trivial Jacobi iteration of the L-CPMET system yields the energy which is correct up to the third order of perturbation theory.

In higher orders of perturbation theory these linear terms yield in fact all possible ladder-type diagrams [Dgm (134 i, ii)], ring diagrams [Dgm (134 iii)] and ring diagrams with exchange [Dgm (134 iv)] as well as their various combinations.



(10.134)

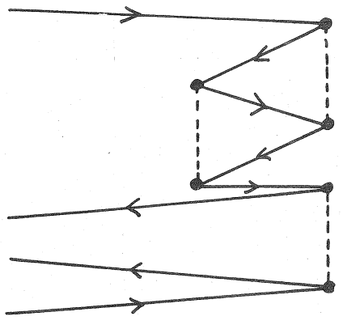
(Diagrams (134) are Goldstone-type versions.)

We remark that both ladders and rings (with or without exchange) can be summed to infinite order of perturbation theory by solving a Dyson-like equation of the general propagator theory and yield, respectively, the so called Brueckner perturbation theory and the random phase approximation (RPA). It is noteworthy that these both approximations represent special cases of the L-CPMET approach, which includes not only all ladders and rings but also their various combinations. It must be added, however, that the above mentioned approximations have been developed to treat various extended systems (nuclear matter, electron gas) and that it is very difficult to solve the L-CPMET equations for these models, since they become a system of integro-differential equations (see e.g., Bishop and Lührmann, Freeman).

We also see immediately from Dgms. (133) that the subsequent iterations of the linear part always yield doubly excited wavefunction diagrams. Consequently, by a systematic iteration of the linear part of Eq. (129) we obtain all possible perturbation theory diagrams to all orders in which all intermediate states are doubly excited configurations. Such an approximation was independently carried out by Bartlett and Shavitt from the many-body perturbation theory viewpoint and was termed the DMBPT(∞) [or, simply, DMBPT] approximation. We thus find that this approximation is equivalent to the L-CPMET or CEPA(0) approximations.

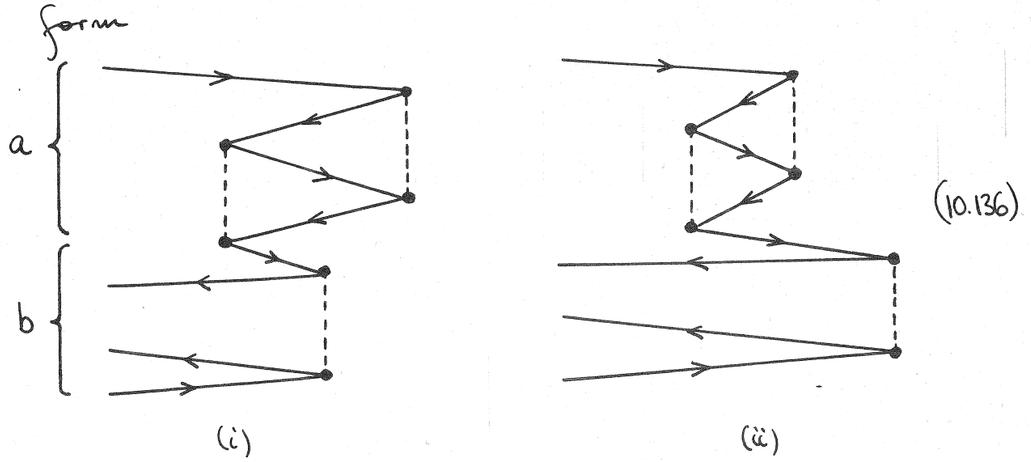
Consider, next, the contributions from non-linear terms. Using the first non-trivial approximation for the t_2 clusters, Eq. (131), on the right hand side of Eq. (129), the non-linear terms will give the third-order contribution to the wavefunction (the fourth-order contribution to the energy). For example, the contribution from the last diagram, Dgms. (69 iii), can be formally described by the following diagram

which results when we substitute from Eq. (131) into the diagram (10), Dgms. (69iii)



(10.135)

The other diagrams (69iii) will yield analogous third order contributions. If we recall that the time ordering of the individual T_2 vertices (in general of any T_i vertices) among themselves was immaterial, we realize immediately that the above diagram (135) represents in fact the sum of all possible time versions of the first kind. Indeed, these time versions have the form



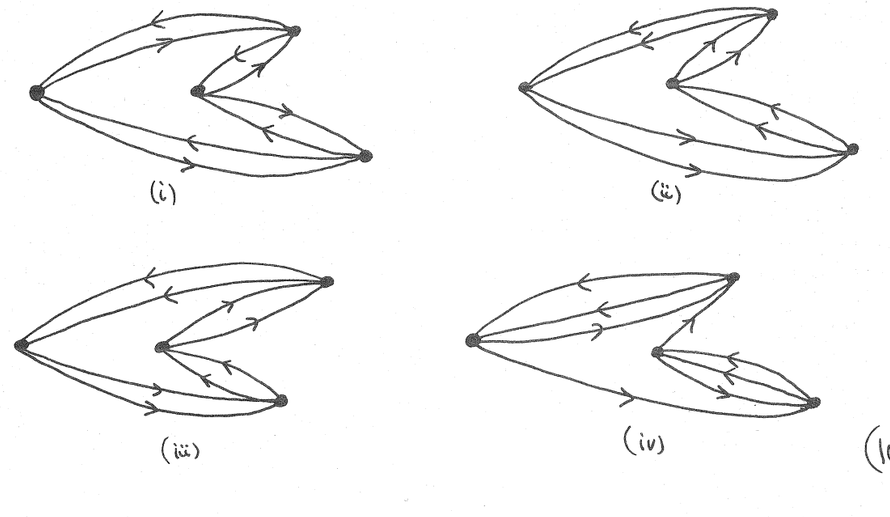
(i) (ii)

(10.136)

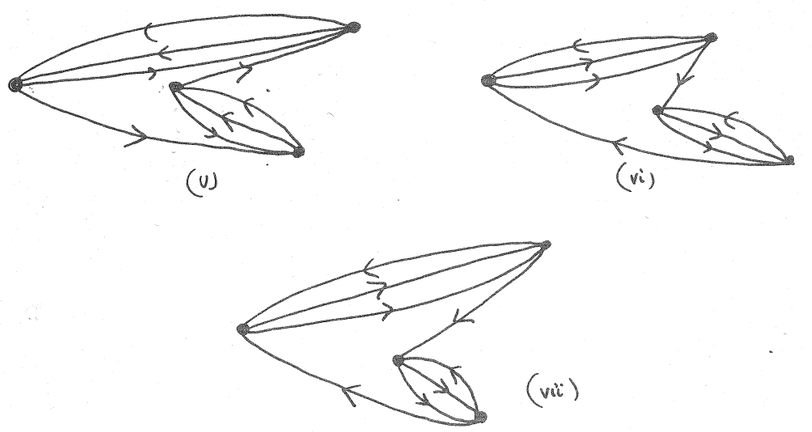
Designating the right most denominators associated with the first order approximants (131) by a and b, as schematically indicated on the left hand side of Dgms (136), the contribution from both time versions (136) is

$$N \frac{1}{d} \left\{ \frac{1}{(a+b)a} + \frac{1}{(a+b)b} \right\} = N \frac{1}{d} \cdot \frac{1}{ab} \,, \qquad (10.137)$$

where N designates the common numerator and d the common leftmost denominator. Thus, in the second iteration, the non-linear terms in Eq. (129) yield the following fourth order contributions to the energy, which we present for simplicity in the Hugenholtz form



(10.138)

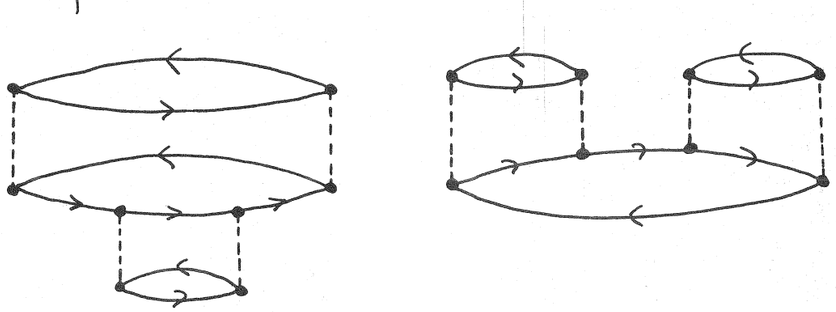


In the subsequent iterations we use all the diagrams (13) and (133) and all possible diagrams of the type (135) or (136) on the right hand side of Eq. (129). The linear terms simply increase the perturbation theory order by one in each iteration. However, using the k and l-th order diagrams in the nonlinear part we obtain the (k+l+1)-th order contribution.

We can thus regard the CPMET (or, in general,

the CCA) equations as a convenient recursive scheme for a systematic generation of higher and higher order perturbation theory diagrams. Solving these equations we not only implicitly construct all these diagrams but we also simultaneously evaluate them and sum to the infinite order of perturbation theory.

In fact, we can use the CCA as a simple way to formulate the many-body perturbation theory and to obtain the linked cluster expansion for the energy. When all cluster components are considered, the result is completely equivalent to the many-body perturbation theory. When approximations are introduced (such as the pair approximation $\hat{T} \approx \hat{T}_2$, for example) certain diagrams are omitted from the perturbation theory expansion. In the CPMET approach, for example, the following fourth order diagrams will not be present



(10.139)

since they involve the connected triexcited and monoexcited intermediate states. Thus, in order to include these diagrams, we have to use the ECPMET. However, except for the diagrams of the type (139) all fourth order diagrams are included in the CPMET approach, as may be seen from Table III.

Appendix 1. Solved Examples

Ex. 1. Show that $\hat{N}_A$ is idempotent; i.e. $\hat{N}_A^2 = \hat{N}_A$.

Sol^{n}: Recall that $\hat{N}_A = \hat{X}_A^\dagger \hat{X}_A$; thus

$$\hat{N}_A^2 = (\hat{X}_A^\dagger \hat{X}_A)(\hat{X}_A^\dagger \hat{X}_A) = \hat{X}_A^\dagger (\hat{1} - \hat{X}_A^\dagger \hat{X}_A)\hat{X}_A \qquad (\text{using: } [\hat{X}_A^\dagger, \hat{X}_A]_+ = \hat{1})$$
$$= \hat{X}_A^\dagger \hat{X}_A - (X_A^\dagger)^2 (X_A)^2 \qquad (\text{using: } (\hat{X}_A^\dagger)^2 = (\hat{X}_A)^2 = 0)$$
$$= \hat{X}_A^\dagger \hat{X}_A = \hat{N}_A$$

Ex. 2. Show that the number operators commute, ie. $[\hat{N}_A, \hat{N}_B]_- = 0$.

Sol^{n}: Using the anticommutation relations we get:

$$\hat{N}_A \hat{N}_B = (\hat{X}_A^\dagger \hat{X}_A)(\hat{X}_B^\dagger \hat{X}_B) = \hat{X}_A^\dagger (\langle A|B\rangle - \hat{X}_B^\dagger \hat{X}_A)\hat{X}_B$$
$$= \langle A|B\rangle \hat{X}_A^\dagger \hat{X}_B - \hat{X}_A^\dagger \hat{X}_B^\dagger \hat{X}_A \hat{X}_B$$
$$= \langle A|B\rangle \hat{X}_A^\dagger \hat{X}_B - (-\hat{X}_B^\dagger \hat{X}_A^\dagger)(-\hat{X}_B \hat{X}_A)$$
$$= \langle A|B\rangle \hat{X}_A^\dagger \hat{X}_B - \hat{X}_B^\dagger (\langle A|B\rangle - \hat{X}_B \hat{X}_A^\dagger)\hat{X}_A$$
$$= \langle A|B\rangle (\hat{X}_A^\dagger \hat{X}_B - \hat{X}_B^\dagger \hat{X}_A) + \hat{X}_B^\dagger \hat{X}_B \hat{X}_A^\dagger \hat{X}_A$$

The first term on the r.h.s. of the last equation always vanishes, since either $A \neq B$ in which case $\langle A|B\rangle = 0$ or $A = B$ in which case $(X_A^\dagger X_B - X_B^\dagger X_A) = X_A^\dagger X_A - X_A^\dagger X_A = 0$. The last term equals $\hat{N}_B \hat{N}_A$, so that we have

$$\hat{N}_A \hat{N}_B = \hat{N}_B \hat{N}_A \quad \text{or} \quad [\hat{N}_A, \hat{N}_B]_- = 0.$$

Ex. 3. Show that

$$[\hat{X}_A, \hat{N}_B]_- = \langle B|A\rangle X_A$$
$$[\hat{X}_A^\dagger, \hat{N}_B]_- = \langle B|A\rangle X_A^\dagger$$

and find a sign error in one of these relations.

Sol^{n}: Let's start with the first relationship: we simply "anticommute" the operator $\hat{X}_A$ "through" the operator $\hat{N}_B$ and see what remains.

$$\hat{X}_A \hat{N}_B = \hat{X}_A \hat{X}_B^\dagger \hat{X}_B = (\langle B|A\rangle - \hat{X}_B^\dagger \hat{X}_A)\hat{X}_B$$
$$= \langle B|A\rangle \hat{X}_B - \hat{X}_B^\dagger (-\hat{X}_B \hat{X}_A)$$
$$= \langle B|A\rangle \hat{X}_B + \hat{N}_B X_A$$

Thus $$[\hat{X}_A, \hat{N}_B]_- = \hat{X}_A \hat{N}_B - \hat{N}_B \hat{X}_A = \langle B|A\rangle \hat{X}_B = \langle B|A\rangle \hat{X}_A.$$

The second relationship may be proved in an analogous way or, we can remember that $\hat{N}_A$ is Hermitian, and take simply the Hermitian conjugate of the first equation.

Recalling that $[\hat{A},\hat{B}]_-^\dagger = [\hat{B}^\dagger,\hat{A}^\dagger]_-$ we get:

$$[\hat{X}_A,\hat{N}_B]_-^\dagger = [N_B^\dagger, X_A^\dagger]_- = -[\hat{X}_A^\dagger,\hat{N}_B]_- =$$
$$= \overline{\langle B|A\rangle}\, X_A^\dagger .$$

Assuming that $\langle B|A\rangle$ is real (in fact in our case this is the Kronecker delta symbol) we have finally that $[\hat{X}_A^\dagger,\hat{N}_B]_- = -\langle B|A\rangle X_A^\dagger$.

Thus, the sign is wrong in the second identity.

Ex. 4. Note that an arbitrary basis ket $|\{A_1 A_2 \cdots A_N\}\rangle$ can be expressed in the form

$$|\{A_1 A_2 \cdots A_N\}\rangle = \Big(\prod_{i=1}^{N} \hat{X}_{A_i}^\dagger\Big)|0\rangle \equiv \hat{X}_{A_1}^\dagger \hat{X}_{A_2}^\dagger \cdots \hat{X}_{A_N}^\dagger |0\rangle . \quad (\S)$$

Solⁿ: Recall that $|A\rangle = X_A^\dagger|0\rangle$. Thus $|A_N\rangle = X_{A_N}^\dagger|0\rangle$. Using now repeatedly the definition of the creation operator we get

$$|\{A_{N-1} A_N\}\rangle = \hat{X}_{A_{N-1}}^\dagger |A_N\rangle = \hat{X}_{A_{N-1}}^\dagger X_{A_N}^\dagger |0\rangle$$

and, generally

$$|\{A_i A_{i+1} \cdots A_N\}\rangle = \hat{X}_{A_i}^\dagger |\{A_{i+1} \cdots A_N\}\rangle$$
$$= \hat{X}_{A_i}^\dagger \hat{X}_{A_{i+1}}^\dagger \cdots \hat{X}_{A_N}^\dagger |0\rangle ,$$

which is equivalent to $(\S)$ when $i=1$.

Ex. 5. Calculate overlap of the following two state vectors $|\Phi\rangle$ and $|\Psi\rangle$,

$$|\Phi\rangle = |\{A_1 A_2\}\rangle$$
$$|\Psi\rangle = |\{B_1 B_2\}\rangle$$

(assuming again that $|A_i\rangle$ & $|B_i\rangle$ are orthonormal; or, at least that the pertinent overlap integrals are real).

Solⁿ: Using the result of Ex. 4, we can write

$$|\Phi\rangle = \hat{X}_{A_1}^\dagger \hat{X}_{A_2}^\dagger |0\rangle \; ; \text{ so that } \langle\Phi| = \langle 0|\hat{X}_{A_2}\hat{X}_{A_1} \; ;$$

and

$$|\Psi\rangle = \hat{X}_{B_1}^\dagger \hat{X}_{B_2}^\dagger |0\rangle .$$

Thus

$$\langle\Phi|\Psi\rangle = \langle 0|\hat{X}_{A_2}\hat{X}_{A_1}\hat{X}_{B_1}^\dagger \hat{X}_{B_2}^\dagger |0\rangle ,$$

i.e. a vacuum mean value of some product of the creation and annihilation operators.

Recall that $\hat{X}_A|0\rangle = 0 = \langle 0|\hat{X}_A^\dagger$. Thus, we shall use the anticommutation relations to move the annihilation operators to the rhs and the creators to the lhs:

$$\langle\Phi|\Psi\rangle = \langle 0|\hat{X}_{A_2}\big(\langle A_1|B_1\rangle - \hat{X}_{B_1}^\dagger \hat{X}_{A_1}\big)\hat{X}_{B_2}^\dagger|0\rangle$$
$$= \langle 0|\hat{X}_{A_2}\hat{X}_{B_2}^\dagger|0\rangle\langle A_1|B_1\rangle - \langle 0|\hat{X}_{A_2}X_{B_1}^\dagger X_{A_1}X_{B_2}^\dagger|0\rangle$$
$$= \langle 0|\big(\langle A_2|B_2\rangle - \hat{X}_{B_2}^\dagger\hat{X}_{A_2}\big)|0\rangle\langle A_1|B_1\rangle -$$
$$-\langle 0|\hat{X}_{A_2}\hat{X}_{B_1}^\dagger\big(\langle A_1|B_2\rangle - \hat{X}_{B_2}^\dagger\hat{X}_{A_1}\big)|0\rangle$$
$$= \langle A_1|B_1\rangle\langle A_2|B_2\rangle\langle 0|0\rangle - \langle A_1|B_1\rangle\langle 0|\hat{X}_{B_2}^\dagger\hat{X}_{A_2}|0\rangle$$
$$-\langle A_1|B_2\rangle\langle 0|\hat{X}_{A_2}\hat{X}_{B_1}^\dagger|0\rangle + \langle 0|\hat{X}_{A_2}\hat{X}_{B_1}^\dagger\hat{X}_{B_2}^\dagger\hat{X}_{A_1}|0\rangle$$
$$= \langle A_1|B_1\rangle\langle A_2|B_2\rangle - \langle A_1|B_2\rangle\langle 0|\hat{X}_{A_2}\hat{X}_{B_1}^\dagger|0\rangle$$

since $\langle 0|0\rangle = 1$ and $\hat{X}_{A_2}|0\rangle = \hat{X}_{A_1}|0\rangle = 0$.

Finally:

$$\langle\Phi|\Psi\rangle = \langle A_1|B_1\rangle\langle A_2|B_2\rangle -$$
$$-\langle A_1|B_2\rangle\langle 0|\big(\langle A_2|B_1\rangle - \hat{X}_{B_1}^\dagger\hat{X}_{A_2}\big)|0\rangle$$
$$= \langle A_1|B_1\rangle\langle A_2|B_2\rangle - \langle A_1|B_2\rangle\langle A_2|B_1\rangle$$
$$+ \cancel{\langle A_1|B_2\rangle\langle 0|\hat{X}_{B_1}^\dagger\hat{X}_{A_2}|0\rangle}$$
$$= \begin{vmatrix} \langle A_1|B_1\rangle & \langle A_1|B_2\rangle \\ \langle A_2|B_1\rangle & \langle A_2|B_2\rangle \end{vmatrix} ,$$

a well known result.

Ex. 6: Simplify (i) $\hat{X}_A \hat{X}_B^+ \hat{X}_C^+ |0\rangle$

(ii) $\langle 0| \hat{X}_A \hat{X}_B^+ \hat{X}_C \hat{X}_D^+ |0\rangle$

(iii) $\langle 0| \hat{X}_A \hat{X}_B \hat{X}_C^+ \hat{X}_D \hat{X}_E^+ |0\rangle$

Sol^n: (i) Anticommute $\hat{X}_A$ to the rhs:

$$\hat{X}_A \hat{X}_B^+ \hat{X}_C^+ |0\rangle = (\langle A|B\rangle - X_B^+ X_A) X_C^+ |0\rangle$$
$$= \langle A|B\rangle \hat{X}_C^+ |0\rangle - \hat{X}_B^+ (\langle A|C\rangle - \hat{X}_C^+ \hat{X}_A)|0\rangle$$

Since $\hat{X}_A|0\rangle = 0$

we get $\hat{X}_A \hat{X}_B^+ \hat{X}_C^+ |0\rangle = \langle A|B\rangle \hat{X}_C^+ |0\rangle - \langle A|C\rangle X_B^+ |0\rangle$

$$= \underline{\langle A|B\rangle |C\rangle - \langle A|C\rangle |B\rangle}.$$

(ii) Again, anticommute annihilators to the rhs:

$$\langle 0| \hat{X}_A \hat{X}_B^+ \hat{X}_C \hat{X}_D^+ |0\rangle = \langle 0| \hat{X}_A \hat{X}_B^+ (\langle C|D\rangle - \hat{X}_D^+ \hat{X}_C)|0\rangle$$
$$= \langle C|D\rangle \langle 0| \hat{X}_A \hat{X}_B^+ |0\rangle - \langle 0| \hat{X}_A \hat{X}_B^+ X_D^+ X_C |0\rangle$$
$$= \langle C|D\rangle \langle 0| (\langle A|B\rangle - \hat{X}_B^+ \hat{X}_A)|0\rangle$$
$$= \langle C|D\rangle \langle A|B\rangle \langle 0|0\rangle = \underline{\langle A|B\rangle \langle C|D\rangle}$$

(iii) Using the same technique as in (ii) we get

$$\langle 0| \hat{X}_A \hat{X}_B \hat{X}_C^+ \hat{X}_D \hat{X}_E^+ |0\rangle = \langle 0| \hat{X}_A \hat{X}_B \hat{X}_C^+ (\langle D|E\rangle - \hat{X}_E^+ \hat{X}_D)|0\rangle$$
$$= \langle D|E\rangle \langle 0| \hat{X}_A \hat{X}_B \hat{X}_C^+ |0\rangle$$
$$= \langle D|E\rangle \langle 0| \hat{X}_A (\langle B|C\rangle - \hat{X}_C^+ \hat{X}_B)|0\rangle$$
$$= \underline{0}$$

We shall see how the same result is obtained using Wick's thm in a much easier way.

Ex 7: Show that the two states which differ in one spinorbital only, say $|\Phi_1\rangle = |\{A\, A_2 A_3 \cdots A_N\}\rangle$ and $|\Phi_2\rangle = |\{B\, A_2 A_3 \cdots A_N\}\rangle$ may be related as follows $|\Phi_2\rangle = \hat{X}_B^+ \hat{X}_A |\Phi_1\rangle$ or $|\Phi_1\rangle = \hat{X}_A^+ \hat{X}_B |\Phi_2\rangle$.

Sol^n: This follows immediately from the definitions of the creation and annihilation operators. Or, more formally, we write both states using Ex. 4

$$|\Phi_1\rangle = \hat{X}_A^+ \hat{X}_{A_2}^+ \cdots X_{A_N}^+ |0\rangle$$
$$|\Phi_2\rangle = \hat{X}_B^+ X_{A_2}^+ \cdots X_{A_N}^+ |0\rangle.$$

Substituting now for $|\Phi_1\rangle$ into the rhs of our first relationship we get

$$|\Phi_2\rangle = \hat{X}_B^+ \hat{X}_A |\Phi_1\rangle$$
$$= (\hat{X}_B^+ \hat{X}_A)\, \hat{X}_A^+ \hat{X}_{A_2}^+ \cdots \hat{X}_{A_N}^+ |0\rangle$$
$$= \hat{X}_B^+ (\hat{X}_A \hat{X}_A^+)\, \hat{X}_{A_2}^+ \cdots X_{A_N}^+ |0\rangle$$
$$= \hat{X}_B^+ (1 - \hat{X}_A^+ \hat{X}_A) |\{A_2 \cdots A_N\}\rangle = \hat{X}_B^+ |\{A_2 \cdots A_N\}\rangle - \hat{N}_A |\{A_2 \cdots A_N\}\rangle$$

Now, since A cannot equal any of the A_i's (otherwise $|\Phi_1\rangle$ vanishes), we have $\hat{N}_A |\{A_2 \cdots A_N\}\rangle = 0 |\{A_2 \cdots A_N\}\rangle$ (i.e., occ. number of A is 0)

and, thus $|\Phi_2\rangle = \hat{X}_B^+ |\{A_2 \cdots A_N\}\rangle = |\{B A_2 \cdots A_N\}\rangle = |\Phi_2\rangle$. Q.E.D.

Clearly, the operators of the type $(X_B^+ X_A)$ may be interpreted as the "excitation operators".

Ex 8: Calculate (i) $n[\hat{X}_A \hat{X}_B^+ \hat{X}_C^+]$

(ii) $n[\hat{X}_A \hat{X}_B \hat{X}_C \hat{X}_D^+ \hat{X}_E^+ \hat{X}_F^+ \hat{X}_G \hat{X}_H]$

(iii) $\langle 0| n[\hat{X}_A \hat{X}_B \hat{X}_C \hat{X}_D^+ \hat{X}_E^+ \hat{X}_F^+]|0\rangle$

Sol^n: Using the definition of the n-product & of the n-product with contractions we find easily:

(i) $n[\hat{X}_A \hat{X}_B^+ \hat{X}_C^+] = \underline{\hat{X}_B^+ \hat{X}_C^+ \hat{X}_A = -\hat{X}_C^+ \hat{X}_B^+ \hat{X}_A}$

(ii) $n[\hat{X}_A \hat{X}_B \hat{X}_C \hat{X}_D^+ \hat{X}_E^+ \hat{X}_F^+ \hat{X}_G \hat{X}_H] = -X_A X_E^+ X_B X_F^+ X_C X_D^+ n[\hat{X}_G \hat{X}_H]$

$$= \underline{-\langle A|E\rangle \langle B|F\rangle \langle C|D\rangle \hat{X}_G \hat{X}_H}$$

(iii) $\langle 0| n[\hat{X}_A \hat{X}_B \hat{X}_C \hat{X}_D^+ \hat{X}_E^+ \hat{X}_F^+]|0\rangle = -\langle A|E\rangle \langle B|F\rangle \langle C|D\rangle \times$

$$\times \underbrace{\langle 0| \underbrace{n[\phi]}_{1} |0\rangle}_{1} = \underline{-\langle A|E\rangle \langle B|F\rangle \langle C|D\rangle}$$

Ex. 9 Simplify expressions of Ex.6. using Wick's theorem.

Remark: Let us use the following shorthand notation: $\hat{X}_C \equiv \hat{C}$; $\hat{X}_C^+ \equiv \hat{C}^+$

Sol^n: (i) $\hat{X}_A \hat{X}_B^+ \hat{X}_C^+ |0\rangle$ i.e. $\hat{A}\hat{B}^+\hat{C}^+|0\rangle$

Wick's theorem gives

$$\hat{A}\hat{B}^+\hat{C}^+ = n[\hat{A}\hat{B}^+\hat{C}^+] + n[\underbrace{\hat{A}\hat{B}^+}\hat{C}^+] + n[\hat{A}\underbrace{\hat{B}^+\hat{C}^+}] + n[\underbrace{\hat{A}\hat{B}^+\hat{C}^+}]$$

Since $n[\hat{A}\hat{B}^+\hat{C}^+]|0\rangle = 0$ and $\underbrace{\hat{B}^+\hat{C}^+} = 0$ we have

$$\underline{\hat{X}_A \hat{X}_B^+ \hat{X}_C^+ |0\rangle = \langle A|B\rangle \hat{X}_C^+|0\rangle - \langle A|C\rangle \hat{X}_B^+|0\rangle}$$

(ii) $\langle 0|\hat{X}_A \hat{X}_B^+ \hat{X}_C \hat{X}_D^+|0\rangle$ i.e. $\langle 0|\hat{A}\hat{B}^+\hat{C}\hat{D}^+|0\rangle$

Recall that only fully contracted terms can contribute to the vacuum mean value and that the only nonvanishing contractions are of the type $\underbrace{\hat{X}_A \hat{X}_B^+} = \langle A|B\rangle$.

Thus

$$\langle 0|\hat{A}\hat{B}^+\hat{C}\hat{D}^+|0\rangle = \langle 0|n[\underbrace{\hat{A}\hat{B}^+}\underbrace{\hat{C}\hat{D}^+}]|0\rangle = \langle A|B\rangle\langle C|D\rangle .$$

(iii) $\langle 0|\hat{X}_A \hat{X}_B \hat{X}_C^+ \hat{X}_D \hat{X}_E^+|0\rangle = 0$

Recalling that the vacuum mean value of an *odd* number of creation and annihilation operators must always vanish.

Ex. 10 Calculate one & two particle operator matrix element & show that you get Slater rules.

Sol^n: (i) *One particle operators*; say $\hat{Z} = \sum_{A,B} \langle A|\hat{z}|B\rangle \hat{X}_A^+ \hat{X}_B$.

Cousiden a general matrix element between two N-particle states, i.e.

$$\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{Z}|\{A_{j_1} \cdots A_{j_N}\}\rangle =$$

$$= \sum_{A,B} \langle A|\hat{z}|B\rangle \langle 0|\hat{X}_{A_{i_N}} \cdots \hat{X}_{A_{i_1}} \hat{X}_A^+ \hat{X}_B \hat{X}_{A_{j_1}}^+ \cdots \hat{X}_{A_{j_N}}^+|0\rangle$$

We shall use again the shorthand notation: $\hat{X}_{A_i} \equiv \hat{A}_i$, $\hat{X}_{A_i}^+ \equiv \hat{A}_i^+$

Thus, we have to evaluate the following matrix element

$$\langle 0|\hat{A}_{i_N} \cdots \hat{A}_{i_2} \hat{A}^+ \hat{B} \hat{A}_{j_1}^+ \cdots \hat{A}_{j_N}^+|0\rangle$$

In order to get a nonvanishing contribution to this mtx elmt using Wick's thm: (i) only fully contracted terms need be considered

(ii) all contractions must be of the type $\underbrace{\hat{X}_M \hat{X}_N^+} = \langle M|N\rangle$.

Thus, $\hat{X}_B$ must be contracted with some (creation) operator on its right hand side (rhs), i.e. $\underbrace{\hat{B}\hat{A}_{j_k}^+}$, while $\hat{X}_A^+$ must be contracted with some annihilation operator on its lhs, i.e. $\underbrace{\hat{A}_{i_\ell}\hat{A}^+}$.

Further, the remaining (N-1) annihilators $\hat{X}_{A_i}$ must be contracted (one by one, of course) with the remaining (N-1) creators $\hat{X}_{A_j}^+$.

Clearly, we get a nonvanishing result only if both states differ in at most one spinorbital: say $A_{i_1} \neq A_{j_1}$, but $A_{i_k} = A_{j_k}$ $(k = 2, \cdots, N)$. Then $\hat{A}_{i_1}$ and $\hat{A}_{j_1}^+$ may be contracted with $\hat{A}^+$ and $\hat{B}$, respectively, and the remaining (N-1) operators on each side are contracted one by one: $\underbrace{\hat{A}_{i_k}\hat{A}_{i_k}^+}$, each contraction yielding a factor 1.

Thus, (a) $\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{Z}|\{A_{j_1} \cdots A_{j_N}\}\rangle = 0$ if the N-el. states differ in more than one spinorbital, and

(b) $\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{Z}|\{A_{j_1} A_{i_2} \cdots A_{i_N}\}\rangle =$

$$= \sum_{A,B} \langle A|\hat{z}|B\rangle \langle 0|\hat{A}_{i_N} \cdots \hat{A}_{i_2} \hat{A}_{i_1} \hat{A}^+ \hat{B} \hat{A}_{j_1}^+ \hat{A}_{i_2}^+ \cdots \hat{A}_{i_N}^+|0\rangle$$

$$= \sum_{A,B} \langle A|\hat{z}|B\rangle \langle A_{i_1}|A\rangle\langle B|A_{j_1}\rangle \underbrace{\prod_{k=2}^{N} \langle A_{i_k}|A_{i_k}\rangle}_{1}$$

$$= \underline{\langle A_{i_1}|\hat{z}|A_{j_1}\rangle}$$ if the N-el. states differ in *one* spinorbital (assumed to be the first spinorbital in each list).

Finally, we have to cousider the diagonal case

(c) $\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{Z}|\{A_{i_1} \cdots A_{i_N}\}\rangle =$

$$= \sum_{A,B} \langle A|\hat{z}|B\rangle \langle 0|\hat{A}_{i_N} \cdots \hat{A}_{i_1} \hat{A}^+ \hat{B} \hat{A}_{i_1}^+ \cdots \hat{A}_{i_N}^+|0\rangle =$$

In this case any (N-1) operators on each side may be contracted, one by one, to give a nonvanishing result. Thus, the operators $\hat{A}^+$ & $\hat{B}$ may be contracted with any operators $\hat{A}_{i_k}$ and $\hat{A}_{i_k}^+$ $(k = 1, \cdots, N)$. Thus

$$= \sum_{A,B} \langle A|\hat{z}|B\rangle \sum_{k=1}^{N} \langle 0|\hat{A}_{i_N} \cdots \hat{A}_{i_k} \cdots \hat{A}_{i_1} \hat{A}^+ \hat{B} \hat{A}_{i_1}^+ \cdots \hat{A}_{i_k}^+ \cdots \hat{A}_{i_N}^+|0\rangle$$

$$= \sum_{A,B} \langle A|\hat{z}|B\rangle \sum_{k=1}^{N} \langle A_{i_k}|A\rangle\langle B|A_{i_k}\rangle = \underline{\sum_{k=1}^{N} \langle A_{i_k}|\hat{z}|A_{i_k}\rangle}$$

(ii) Two-particle operators; say $\hat{V} = \frac{1}{2}\sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \hat{X}_A^+ \hat{X}_B^+ \hat{X}_D \hat{X}_C$.

Consider a general mtx elmt

$$\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{V}|\{A_{j_1} \cdots A_{j_N}\}\rangle =$$

$$= \frac{1}{2}\sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \langle 0|\hat{A}_{i_N} \cdots \hat{A}_{i_1} \hat{A}^+ \hat{B}^+ \hat{D}\hat{C} \hat{A}_{j_1}^+ \cdots \hat{A}_{j_N}^+|0\rangle$$

Using analogous reasoning as in the one-particle case we find immediately that the vacuum mean value mtx elmts above will vanish unless $\hat{A}^+, \hat{B}^+$ are contracted with some annihilators $\hat{A}_{i_k}, \hat{A}_{i_\ell}$ on their rhs and the operators $\hat{D}, \hat{C}$ are contracted with some creators $\hat{X}_{A_{j_m}}^+, \hat{X}_{A_{j_n}}^+$ on their rhs. The remaining $(N-2)$ annihilators must be contracted one-by-one with the remaining $(N-2)$ creators. Thus, to get a nonvanishing result, at least $(N-2)$ single particle states on each side of the operator $\hat{V}$ must be identical. Thus

(a) $\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{V}|\{A_{j_1} \cdots A_{j_m}\}\rangle = 0$ if both states differ in more than two spinorbitals

(b) $\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{V}|\{A_{j_1} A_{j_2} A_{i_3} \cdots A_{i_N}\}\rangle =$, i.e. we assume that the two states differ in two spinorbitals (& write them as the first two states without any loss of generality)

$$= \frac{1}{2}\sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \times$$

$$\times \langle 0|\hat{A}_{i_N} \cdots \hat{A}_{i_3} \boxed{\hat{A}_{i_2} \hat{A}_{i_1} \hat{A}^+ \hat{B}^+ \hat{D} \hat{C} \hat{A}_{j_1}^+ \hat{A}_{j_2}^+} \hat{A}_{i_3}^+ \cdots \hat{A}_{i_N}^+|0\rangle$$

Consider the possible contraction schemes for the operators enclosed in a dashed rectangle, or more generally

$$\boxed{A_{i_k} A_{i_\ell} A^+ B^+ D C A_{j_m}^+ A_{j_n}^+} =$$

$$= A_{i_k} A_{i_\ell} A^+ B^+ D C A_{j_m}^+ A_{j_n}^+$$
$$+ A_{i_k} A_{i_\ell} A^+ B^+ D C A_{j_m}^+ A_{j_n}^+$$
$$+ A_{i_k} A_{i_\ell} A^+ B^+ D C A_{j_m}^+ A_{j_n}^+$$
$$+ A_{i_k} A_{i_\ell} A^+ B^+ D C A_{j_m}^+ A_{j_n}^+$$

$$= \langle A_{i_\ell}|A\rangle\langle A_{i_k}|B\rangle\left(\langle C|A_{j_m}\rangle\langle D|A_{j_n}\rangle - \langle C|A_{j_n}\rangle\langle D|A_{j_m}\rangle\right)$$
$$- \langle A_{i_k}|A\rangle\langle A_{i_\ell}|B\rangle\left(\langle C|A_{j_m}\rangle\langle D|A_{j_n}\rangle - \langle C|A_{j_n}\rangle\langle D|A_{j_m}\rangle\right)$$

$$= \left(\langle A_{i_k}|B\rangle\langle A_{i_\ell}|A\rangle - \langle A_{i_\ell}|B\rangle\langle A_{i_k}|A\rangle\right)\left(\langle C|A_{j_m}\rangle\langle D|A_{j_n}\rangle - \langle C|A_{j_n}\rangle\langle D|A_{j_m}\rangle\right).$$

Thus, in case (b)

$$\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{V}|\{A_{j_1} A_{j_2} A_{i_3} \cdots A_{i_N}\}\rangle =$$

$$= \frac{1}{2}\sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \left(\langle A_{i_1}|A\rangle\langle A_{i_2}|B\rangle - \langle A_{i_1}|B\rangle\langle A_{i_2}|A\rangle\right) \times$$
$$\times \left(\langle A_{j_1}|C\rangle\langle A_{j_2}|D\rangle - \langle A_{j_1}|D\rangle\langle A_{j_2}|C\rangle\right)$$

$$= \frac{1}{2}\left\{\langle A_{i_1} A_{i_2}|\hat{v}|A_{j_1} A_{j_2}\rangle - \langle A_{i_2} A_{i_1}|\hat{v}|A_{j_1} A_{j_2}\rangle \right.$$
$$\left. - \langle A_{i_1} A_{i_2}|\hat{v}|A_{j_2} A_{j_1}\rangle + \langle A_{i_2} A_{i_1}|\hat{v}|A_{j_2} A_{j_1}\rangle\right\}$$

$$= \langle A_{i_1} A_{i_2}|\hat{v}|A_{j_1} A_{j_2}\rangle - \langle A_{i_1} A_{i_2}|\hat{v}|A_{j_2} A_{j_1}\rangle$$

$$= \underline{\langle A_{i_1} A_{i_2}|\hat{v}|A_{j_1} A_{j_2}\rangle_A}.$$

Consider next case

(c) $\langle \{A_{i_1} \cdots A_{i_N}\}|\hat{V}|\{A_{j_1} A_{i_2} \cdots A_{i_N}\}\rangle$, i.e. assume that the two states differ in one spinorbital.

$$= \frac{1}{2}\sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \langle 0|\hat{A}_{i_N} \cdots \hat{A}_{i_2} \hat{A}_{i_1} \hat{A}^+ \hat{B}^+ \hat{D} \hat{C} \hat{A}_{j_1}^+ \hat{A}_{i_2}^+ \cdots \hat{A}_{i_N}^+|0\rangle$$

Now, we shall have that $i_\ell = i_2$, $j_m = j_2$

and i_k and $j_n \equiv i_k$ $(k \neq 1)$.

In other words, A_{i_1} and $A_{j_1}^+$ must be always contracted with one of the operators A^+, B^+ and $\hat{O}, \hat{C}$, respectively. The second operator coming from $\hat{V}$ (i.e., one creator & one annihilator) may now be contracted with any of the $(N-1)$ particle states, which are assumed to be identical, on each side of $\hat{V}$.

Thus, we can write

$$\langle \{A_{i_1} \cdots A_{i_N}\} | \hat{V} | \{A_{j_1} A_{i_2} \cdots A_{i_N}\} \rangle =$$

$$= \tfrac{1}{2} \sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \sum_{k=2}^{N} \left(\langle A_{i_1}|A\rangle\langle A_{i_k}|B\rangle - \langle A_{i_1}|B\rangle\langle A_{i_k}|A\rangle\right) \times$$
$$\times \left(\langle C|A_{j_1}\rangle\langle D|A_{i_k}\rangle - \langle C|A_{i_k}\rangle\langle D|A_{j_1}\rangle\right)$$

$$= \tfrac{1}{2} \sum_{k=2}^{N} \{ \langle A_{i_1} A_{i_k}|\hat{v}|A_{j_1} A_{i_k}\rangle - \langle A_{i_1} A_{i_k}|\hat{v}|A_{i_k} A_{j_1}\rangle$$
$$- \langle A_{i_k} A_{i_1}|\hat{v}|A_{j_1} A_{i_k}\rangle + \langle A_{i_k} A_{i_1}|\hat{v}|A_{i_k} A_{j_1}\rangle \}$$

$$= \sum_{k=2}^{N} \langle A_{i_1} A_{i_k}|\hat{v}|A_{j_1} A_{i_k}\rangle_A = \sum_{k=1}^{N} \langle A_{i_1} A_{i_k}|\hat{v}|A_{j_1} A_{i_k}\rangle_A$$

Consider, finally, the diagonal case

(d) $\langle \{A_{i_1} \cdots A_{i_N}\} | \hat{V} | \{A_{i_1} \cdots A_{i_N}\} \rangle$

Here we can have: i_ℓ, i_k any indices, and $j_m = i_\ell$, $j_n = i_k$.

Thus

$$\langle \{A_{i_1} \cdots A_{i_N}\} | \hat{V} | \{A_{i_1} \cdots A_{i_N}\} \rangle =$$

$$= \tfrac{1}{2} \sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \sum_{\substack{k,\ell=1 \\ (k<\ell)}}^{N} \{ \left(\langle A_{i_k}|B\rangle\langle A_{i_\ell}|A\rangle - \langle A_{i_k}|A\rangle\langle A_{i_\ell}|B\rangle\right) \times$$
$$\times \left(\langle A_{i_\ell}|C\rangle\langle A_{i_k}|D\rangle - \langle A_{i_\ell}|D\rangle\langle A_{i_\ell}|C\rangle\right) \}$$

$$= \sum_{\substack{k,\ell=1 \\ (k<\ell)}}^{N} \langle A_{i_k} A_{i_\ell}|\hat{v}|A_{i_k} A_{i_\ell}\rangle_A$$

$$= \tfrac{1}{2} \sum_{k,\ell=1}^{N} \langle A_{i_k} A_{i_\ell}|\hat{v}|A_{i_k} A_{i_\ell}\rangle_A .$$

Ex. 11: Show that $\langle \{AB\}|\hat{v}|\{CD\}\rangle = \langle AB|\hat{v}|CD\rangle_A$.

Sol^n: We can write, using our shorthand notation

$$\langle \{AB\}|\hat{v}|\{CD\}\rangle = \tfrac{1}{2} \sum_{E,F,G,H} \langle EF|\hat{v}|GH\rangle \langle 0|\hat{B}\hat{A}\hat{E}^\dagger\hat{F}^\dagger\hat{H}\hat{G}\hat{C}^\dagger\hat{D}^\dagger|0\rangle$$

and, applying Wick's theorem and dropping the vanishing terms we get, further

$$= \tfrac{1}{2} \sum_{E,F,G,H} \langle EF|\hat{v}|GH\rangle \{ \langle 0|\hat{B}\hat{A}\hat{E}^\dagger\hat{F}^\dagger\hat{H}\hat{G}\hat{C}^\dagger\hat{D}^\dagger|0\rangle$$
$$+ \langle 0|\hat{B}\hat{A}\hat{E}^\dagger\hat{F}^\dagger\hat{H}\hat{G}\hat{C}^\dagger\hat{D}^\dagger|0\rangle$$
$$+ \langle 0|\hat{B}\hat{A}\hat{E}^\dagger\hat{F}^\dagger\hat{H}\hat{G}\hat{C}^\dagger\hat{D}^\dagger|0\rangle$$
$$+ \langle 0|\hat{B}\hat{A}\hat{E}^\dagger\hat{F}^\dagger\hat{H}\hat{G}\hat{C}^\dagger\hat{D}^\dagger|0\rangle$$

$$= \tfrac{1}{2} \sum_{E,F,G,H} \langle EF|\hat{v}|GH\rangle \{ \langle B|F\rangle\langle A|E\rangle\langle H|D\rangle\langle G|C\rangle$$
$$- \langle B|F\rangle\langle A|E\rangle\langle H|C\rangle\langle G|D\rangle$$
$$- \langle B|E\rangle\langle A|F\rangle\langle H|D\rangle\langle G|C\rangle$$
$$+ \langle B|E\rangle\langle A|F\rangle\langle H|C\rangle\langle G|D\rangle \}$$

$$= \tfrac{1}{2} [\langle AB|\hat{v}|CD\rangle - \langle AB|\hat{v}|DC\rangle - \langle BA|\hat{v}|CD\rangle + \langle BA|\hat{v}|DC\rangle]$$
$$- \langle AB|\hat{v}|CD\rangle - \langle AB|\hat{v}|DC\rangle = \langle AB|\hat{v}|CD\rangle_A .$$

Ex. 12: Write the 2-particle operator $\hat{V}$ in a h-p formalism.

Sol^n:

$$\hat{V} = \tfrac{1}{2} \sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \hat{X}_A^\dagger \hat{X}_B^\dagger \hat{X}_D \hat{X}_C$$

$$= \tfrac{1}{2} \left(\sum_{A'} + \sum_{A''}\right)\left(\sum_{B'} + \sum_{B''}\right)\left(\sum_{C'} + \sum_{C''}\right)\left(\sum_{D'} + \sum_{D''}\right) \langle AB|\hat{v}|CD\rangle \hat{X}_A^\dagger \hat{X}_B^\dagger \hat{X}_D \hat{X}_C$$

$$= \tfrac{1}{2} \sum_{A''} \sum_{B''} \sum_{C''} \sum_{D''} \langle A''B''|\hat{v}|C''D''\rangle \hat{Y}_{A''}^\dagger \hat{Y}_{B''}^\dagger \hat{Y}_{D''} \hat{Y}_{C''} \quad \cdots\cdot (a)$$

$$+ \tfrac{1}{2} \sum_{A''} \sum_{B'} \sum_{C''} \sum_{D'} \langle A''B'|\hat{v}|C''D'\rangle \hat{Y}_{A''}^\dagger \hat{Y}_{B'} \hat{Y}_{D'}^\dagger \hat{Y}_{C''} \quad \cdots\cdot (b)$$

$$+ \tfrac{1}{2} \sum_{A'} \sum_{B''} \sum_{C'} \sum_{D''} \langle A'B''|\hat{v}|C'D''\rangle \hat{Y}_{A'} \hat{Y}_{B''}^\dagger \hat{Y}_{D''} \hat{Y}_{C'}^\dagger \quad \cdots\cdot (c)$$

$$+ \tfrac{1}{2} \sum_{A'} \sum_{B'} \sum_{C'} \sum_{D'} \langle A'B'|\hat{v}|C'D'\rangle \hat{Y}_{A'} \hat{Y}_{B'} \hat{Y}_{D'}^\dagger \hat{Y}_{C'}^\dagger \quad \cdots\cdot (d)$$

$$+\tfrac{1}{2}\sum_{A''}\sum_{B'}\sum_{C'}\sum_{D''}\langle A''B'|\hat{v}|C'D''\rangle\ \hat{Y}^+_{A''}\hat{Y}_{B'}\hat{Y}_{D''}\hat{Y}^+_{C'} \quad \text{---- (e)}$$

$$+\tfrac{1}{2}\sum_{A'}\sum_{B''}\sum_{C''}\sum_{D'}\langle A'B''|\hat{v}|C''D'\rangle\ \hat{Y}_{A'}\hat{Y}^+_{B''}\hat{Y}^+_{D'}\hat{Y}_{C''} \quad \text{--- (f)}$$

$$+\tfrac{1}{2}\sum_{A'}\sum_{B'}\sum_{C''}\sum_{D''}\langle A'B'|\hat{v}|C''D''\rangle\ \hat{Y}_{A'}\hat{Y}_{B'}\hat{Y}_{D''}\hat{Y}_{C''} \quad \text{---- (g)}$$

$$+\tfrac{1}{2}\sum_{A''}\sum_{B''}\sum_{C'}\sum_{D'}\langle A''B''|\hat{v}|C'D'\rangle\ \hat{Y}^+_{A''}\hat{Y}^+_{B''}\hat{Y}^+_{D'}\hat{Y}^+_{C'} \quad \text{---- (h)}$$

$$+\tfrac{1}{2}\sum_{A''}\sum_{B''}\sum_{C''}\sum_{D'}\langle A''B''|\hat{v}|C''D'\rangle\ \hat{Y}^+_{A''}\hat{Y}^+_{B''}\hat{Y}^+_{D'}\hat{Y}_{C''} \quad \text{--- (i)}$$

$$+\tfrac{1}{2}\sum_{A''}\sum_{B'}\sum_{C''}\sum_{D''}\langle A''B'|\hat{v}|C''D''\rangle\ \hat{Y}^+_{A''}\hat{Y}_{B'}\hat{Y}_{D''}\hat{Y}_{C''} \quad \text{--- (j)}$$

$$+\tfrac{1}{2}\sum_{A'}\sum_{B''}\sum_{C''}\sum_{D'}\langle A'B''|\hat{v}|C''D'\rangle\ \hat{Y}_{A'}\hat{Y}^+_{B''}\hat{Y}^+_{D'}\hat{Y}_{C''} \quad \text{--- (k)}$$

$$+\tfrac{1}{2}\sum_{A'}\sum_{B'}\sum_{C'}\sum_{D''}\langle A'B'|\hat{v}|C'D''\rangle\ \hat{Y}_{A'}\hat{Y}_{B'}\hat{Y}_{D''}\hat{Y}^+_{C'} \quad \text{--- (l)}$$

$$+\tfrac{1}{2}\sum_{A''}\sum_{B''}\sum_{C'}\sum_{D''}\langle A''B''|\hat{v}|C'D''\rangle\ \hat{Y}^+_{A''}\hat{Y}^+_{B''}\hat{Y}_{D''}\hat{Y}^+_{C'} \quad \text{---- (m)}$$

$$+\tfrac{1}{2}\sum_{A'}\sum_{B''}\sum_{C''}\sum_{D''}\langle A'B''|\hat{v}|C''D''\rangle\ \hat{Y}_{A'}\hat{Y}^+_{B''}\hat{Y}_{D''}\hat{Y}_{C''} \quad \text{--- (n)}$$

$$+\tfrac{1}{2}\sum_{A''}\sum_{B'}\sum_{C'}\sum_{D'}\langle A''B'|\hat{v}|C'D'\rangle\ \hat{Y}^+_{A''}\hat{Y}_{B'}\hat{Y}^+_{D'}\hat{Y}^+_{C'} \quad \text{--- (o)}$$

$$+\tfrac{1}{2}\sum_{A'}\sum_{B'}\sum_{C''}\sum_{D'}\langle A'B'|\hat{v}|C''D'\rangle\ \hat{Y}_{A'}\hat{Y}_{B'}\hat{Y}^+_{D'}\hat{Y}_{C''} \quad \text{---- (p)}$$

Ex. 13: Derive the expressions for contractions of mixed products of cr. & annih. operators w.r.t. Fermi vacuum.

Sol^{n}: We find easily

$$\overbrace{X_A\ Y_B} = 0 \qquad \overbrace{Y_A\ X_B} = h(B)\langle A|B\rangle$$

$$\overbrace{X^+_A\ Y_B} = 0 \qquad \overbrace{Y^+_A\ X_B} = 0$$

$$\overbrace{X_A\ Y^+_B} = p(A)\langle A|B\rangle \qquad \overbrace{Y_A\ X^+_B} = p(B)\langle A|B\rangle$$

$$\overbrace{X^+_A\ Y^+_B} = h(A)\langle A|B\rangle \qquad \overbrace{Y^+_A\ X^+_B} = 0$$

Ex. 14: Calculate the mtx elmnt

$$\langle \Phi_0|\hat{V}|\Phi_1\rangle$$

where

$$|\Phi_0\rangle = |\{A'_1 \dots A'_N\}\rangle$$

$$|\Phi_1\rangle = |\{B_1 A'_2 \dots A'_N\}\rangle = \hat{X}^+_{B_1}\hat{X}_{A'_1}|\Phi_0\rangle = \hat{X}^+_{B_1}\hat{Y}^+_{A'_1}|\Phi_0\rangle$$

and (a) $B_1 \neq A'_1$, (b) $B_1 = A'_1$ (and, of course, $B_1 \neq A'_2, \dots, A'_N$)

Sol^{n}: We have

$$\langle \Phi_0|\hat{V}|\Phi_1\rangle = \tfrac{1}{2}\sum'_{AB,CD}\langle AB|\hat{v}|CD\rangle\langle \Phi_0|\hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{Y}^+_{A'_1}|\Phi_0\rangle$$

Ommiting the vanishing contractions like $\overbrace{X^+_A X^+_B} = \overbrace{X_A X_B} = 0$, we get

$$\langle \Phi_0|\hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{X}^+_{A'_1}|\Phi_0\rangle =$$

$$= \hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{Y}^+_{A'_1} + \hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{Y}^+_{A'_1}$$

$$+ \hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{Y}^+_{A'_1} + \hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{Y}^+_{A'_1}$$

$$+ \hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{Y}^+_{A'_1} + \hat{X}^+_A\hat{X}^+_B\hat{X}_D\hat{X}_C\hat{X}^+_{B_1}\hat{Y}^+_{A'_1}$$

$$= -h(A)\langle A|D\rangle\big[h(B)\langle B|C\rangle\, h(B_1)\langle B_1|A'_1\rangle + h(B)\langle B|A'_1\rangle\, p(C)\langle C|B_1\rangle\big]$$

$$+ h(A)\langle A|C\rangle\big[h(B)\langle B|D\rangle\, h(B_1)\langle B_1|A'_1\rangle + h(B)\langle B|A'_1\rangle\, p(D)\langle D|B_1\rangle\big]$$

$$+ h(A)\langle A|A'_1\rangle\big[h(B)\langle B|D\rangle\, p(C)\langle C|B_1\rangle - h(B)\langle B|C\rangle\, p(D)\langle D|B_1\rangle\big].$$

Thus

$$\langle \Phi_0|\hat{V}|\Phi_1\rangle = -\tfrac{1}{2}\langle B_1|A'_1\rangle\sum_{A',B'}\langle A'B'|\hat{v}|B'A'\rangle - \tfrac{1}{2}p(B_1)\sum_{A'}\langle A'A'_1|\hat{v}|B_1A'\rangle$$

$$+\tfrac{1}{2}\langle B_1|A'_1\rangle\sum_{A',B'}\langle A'B'|\hat{v}|A'B'\rangle + \tfrac{1}{2}p(B_1)\sum_{A'}\langle A'A'_1|\hat{v}|A'B_1\rangle$$

$$+\tfrac{1}{2}p(B_1)\sum_{B'}\langle A'_1B'|\hat{v}|B_1B'\rangle - \tfrac{1}{2}p(B_1)\sum_{B'}\langle A'_1B'|\hat{v}|B'B_1\rangle$$

or,

$$\langle \Phi_0|\hat{V}|\Phi_1\rangle = \langle B_1|A_1'\rangle \tfrac{1}{2}\sum_{A',B'}\langle A'B'|\hat{v}|A'B'\rangle_A + p(B_1)\sum_{A'}\langle A_1'A'|\hat{v}|B_1A'\rangle_A$$

Indeed, this is a general result for the given matrix element.

If (a) $B_1 \neq A_1' \Rightarrow B_1$ is some particle state, say $B_1 = B''$

Thus $p(B'') = 1$ and $\langle B''|A_1'\rangle = 0$

So that $\langle \Phi_0|\hat{V}|\Phi_1\rangle = \sum_{A'}\langle A_1'A'|\hat{v}|B''A'\rangle_A \equiv \langle A_1'|\hat{q}|B''\rangle$.

On the other hand, when

(b) $B_1 = A_1' \Rightarrow |\Phi_1\rangle = |\Phi_0\rangle$ and $\langle A_1'|A_1'\rangle = 1$, $p(A_1') = 0$

$$\langle \Phi_0|\hat{V}|\Phi_0\rangle = \tfrac{1}{2}\sum_{A',B'}\langle A'B'|\hat{v}|A'B'\rangle_A \equiv \langle \hat{V}\rangle .$$

Ex. 15: Reconsider Ex. 14 using the N-product form for $\hat{V}$.

Sol'n: Recall that

$$\hat{V} = \langle \hat{V}\rangle + \sum_{A,B}\langle A|\hat{q}|B\rangle N[\hat{X}_A^+\hat{X}_B] + \hat{V}_N$$

$$\hat{V}_N = \tfrac{1}{2}\sum_{A,B,C,D}\langle AB|\hat{v}|CD\rangle N[\hat{X}_A^+\hat{X}_B^+\hat{X}_D\hat{X}_C] .$$

Thus

$$\langle \Phi_0|\hat{V}|\Phi_1\rangle = \langle \Phi_0|\hat{V}\hat{X}_{B_1}^+\hat{Y}_{A_1'}^+|\Phi_0\rangle$$

$$= \langle \hat{V}\rangle\langle \Phi_0|\hat{X}_{B_1}^+\hat{Y}_{A_1'}^+|\Phi_0\rangle$$

$$+ \sum_{A,B}\langle A|\hat{q}|B\rangle\langle \Phi_0|N[\hat{X}_A^+\hat{X}_B]\hat{X}_{B_1}^+\hat{Y}_{A_1'}^+|\Phi_0\rangle$$

$$+ \tfrac{1}{2}\sum_{A,B,C,D}\langle AB|\hat{v}|CD\rangle\langle \Phi_0|N[\hat{X}_A^+\hat{X}_B^+\hat{X}_D\hat{X}_C]\hat{X}_{B_1}^+\hat{Y}_{A_1'}^+|\Phi_0\rangle .$$

Using generalized Wick's theorem, we find:

(i) the first term gives: $\langle \hat{V}\rangle h(B_1)\langle B_1|A_1'\rangle = \langle B_1|A_1'\rangle\langle \hat{V}\rangle$

(ii) the second term gives:

$$\sum_{A,B}\langle A|\hat{q}|B\rangle\Big\{\langle \Phi_0|N[\overbrace{\hat{X}_A^+\hat{X}_B\hat{X}_{B_1}^+}\hat{Y}_{A_1'}^+]|\Phi_0\rangle +$$

$$+\langle \Phi_0|N[\overbrace{\hat{X}_A^+\hat{X}_B\hat{X}_{B_1}^+}\hat{Y}_{A_1'}^+]|\Phi_0\rangle\Big\}$$

$$= \sum_{A,B}\langle A|\hat{q}|B\rangle h(A)p(B)\left[\langle A|A_1'\rangle\langle B|B_1\rangle - \langle A|B_1\rangle\langle B|A_1'\rangle\right]$$

$$= \sum_{A,B}\langle A|\hat{q}|B\rangle\langle A|A_1'\rangle\langle B|B_1\rangle p(B_1)$$

since $p(B)\langle B|A_1'\rangle \equiv 0$.

Thus, the second term gives

$$p(B_1)\langle A_1'|\hat{q}|B_1\rangle$$

finally,

(iii) the third (last) term always vanishes since no fully contracted terms may appear (i.e., using generalized Wick's theorem, the only possible contractions are between the operators which were not originally in the same N-product; here, however, four operators appear in the same N-product & only two outside — thus, no fully contracted terms can arise)

We see that the same result as in Ex. 14 is obtained

$$\langle \Phi_0|\hat{V}|\Phi_1\rangle = \langle B_1|A_1'\rangle\langle V\rangle + p(B_1)\langle A_1'|\hat{q}|B_1\rangle$$

where

$$\langle A|\hat{q}|B\rangle = \sum_{C'}\langle AC'|\hat{v}|BC'\rangle_A .$$

Clearly, considerably less work is involved using the N-product form for the operators.

Ex. 16 Use diagrammatic technique to evaluate the matrix elements of the Hamiltonian

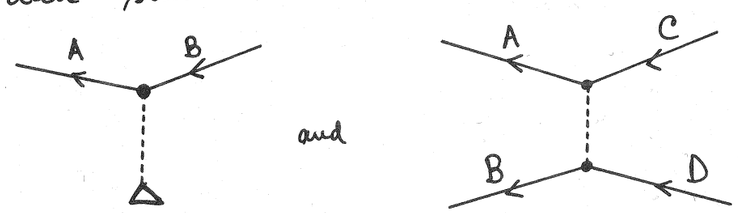
$$\hat{H} = \sum_{A,B} \langle A|\hat{t}|B\rangle \hat{X}_A^\dagger \hat{X}_B + \tfrac{1}{4} \sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle \hat{X}_A^\dagger \hat{X}_B^\dagger \hat{X}_D \hat{X}_C \qquad (1)$$

between the ground, mono- and bi-excited (spinorbital) configurations.

Solⁿ: Denote a general k-times excited (spinorbital) configuration by

$$\left| \begin{matrix} A^1 A^2 \cdots A^k \\ A_1 A_2 \cdots A_k \end{matrix} \right\rangle = \prod_{i=1}^{k} (\hat{X}_{A^i}^\dagger \hat{X}_{A_i}) \, |\Phi_0\rangle , \qquad (2)$$

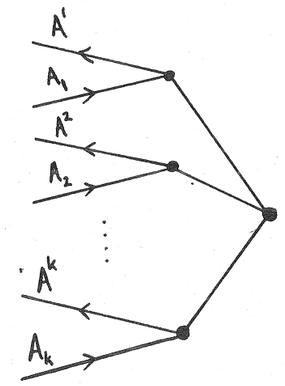
where $|\Phi_0\rangle$ is the ground state configuration, which we take as the Fermi vacuum for the particle-hole formalism, and where we designated the particle and hole state labels by superscripted and subscripted variables, respectively.
We associate standard one- and two-particle vertices



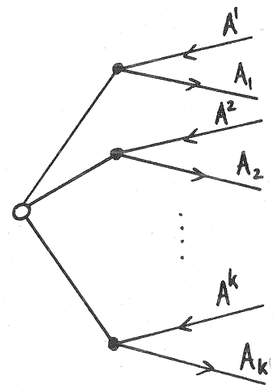
with the one- and two-particle parts, respectively, of the Hamiltonian (1) in the N-product form, i.e.

$$\hat{H} - \langle \hat{H} \rangle = \sum_{A,B} \langle A|\hat{f}|B\rangle N[\hat{X}_A^\dagger \hat{X}_B] + \tfrac{1}{4} \sum_{A,B,C,D} \langle AB|\hat{v}|CD\rangle N[\hat{X}_A^\dagger \hat{X}_B^\dagger \hat{X}_D \hat{X}_C] .$$

Finally, we associate the following T_k and $\overline{T}_k$ diagrams (skeletons) with the k-times excited configuration function (2) and its hermitian conjugate, respectively:



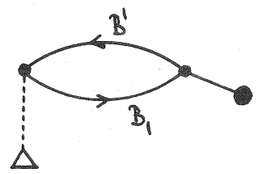
T_k



$\overline{T}_k$

We now construct for each particular case the pertinent resulting diagrams and write down the desired expressions for the matrix elements.

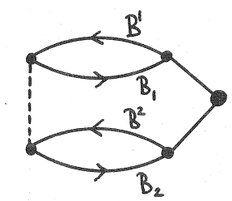
Ground-Mono

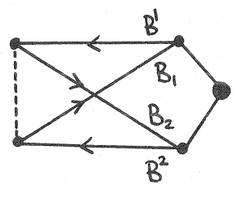


$$\left\langle \Phi_0 \middle| \hat{H} \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle = \left\langle \Phi_0 \middle| \hat{H} - \langle \hat{H} \rangle \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle =$$

$$= \langle B_1 | \hat{f} | B' \rangle$$

Ground-Bi





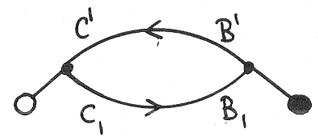
$$\left\langle \Phi_0 \middle| \hat{H} \middle| \begin{matrix} B' B^2 \\ B_1 B_2 \end{matrix} \right\rangle = \left\langle \Phi_0 \middle| \hat{H} - \langle H \rangle \middle| \begin{matrix} B' B^2 \\ B_1 B_2 \end{matrix} \right\rangle =$$

$$= \langle B_1 B_2 | \hat{v} | B' B^2 \rangle_A$$

Mono-Mono

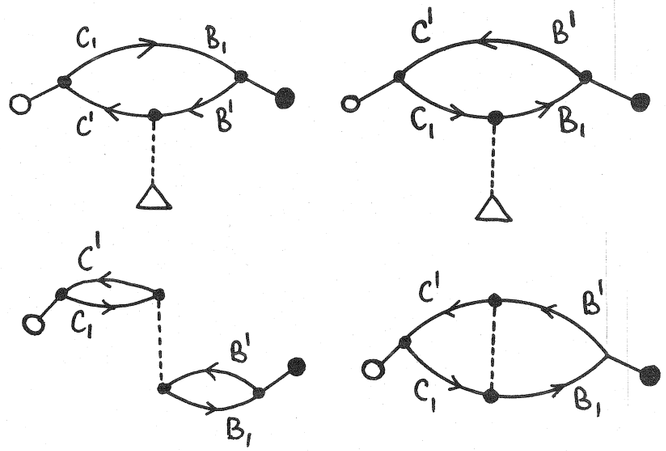
$$\left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \hat{H} - \langle H \rangle \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle = \left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \hat{H} \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle - \langle H \rangle \left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle$$

Overlap:



$$\left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle = \langle c'|B'\rangle\langle c_1|B_1\rangle$$

energy matrix elements:



Thus:

$$\left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \hat{H} - \langle H \rangle \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle = \langle c'|\hat{f}|B'\rangle\langle c_1|B_1\rangle - \langle B_1|\hat{f}|c_1\rangle\langle c'|B'\rangle + \langle c'B_1|\hat{v}|c_1B'\rangle_A$$

Designating $E_0 \equiv \langle \hat{H} \rangle = \langle \Phi_0|\hat{H}|\Phi_0\rangle$

$\Delta^{\mu\nu} = \langle c^\mu|B^\nu\rangle$ and $\Delta_{\mu\nu} = \langle c_\mu|B_\nu\rangle$

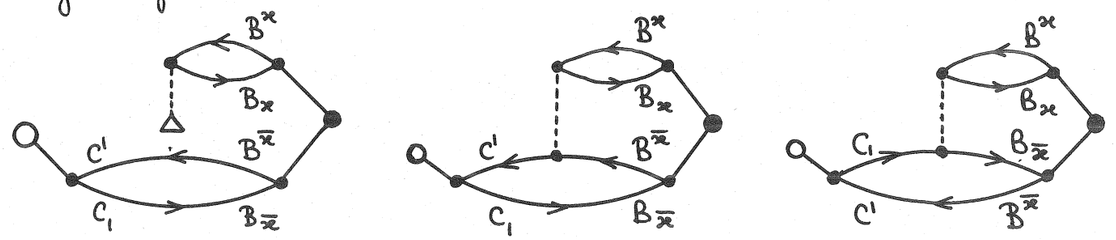
we can write:

$$\left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \hat{H} \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle = E_0\Delta^{11}\Delta_{22} + \langle c'|\hat{f}|B'\rangle\Delta_{11} - \langle B_1|\hat{f}|c_1\rangle\Delta^{11} + \langle c'B_1|\hat{v}|c_1B'\rangle_A .$$

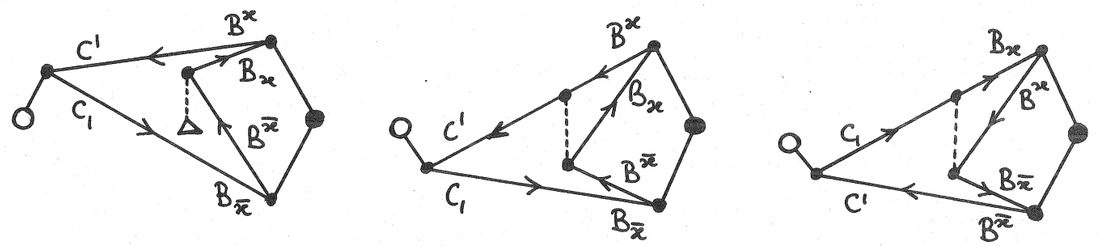
Mono-Bi

$$\left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \hat{H} - \langle \hat{H} \rangle \middle| \begin{matrix} B'B^2 \\ B_1B_2 \end{matrix} \right\rangle = \left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \hat{H} \middle| \begin{matrix} B'B^2 \\ B_1B_2 \end{matrix} \right\rangle$$

We get diagrams:



where $x = 1, 2$ and $\bar{x} = 1 + \delta_{1x}$ (i.e., $\bar{1} = 2$, $\bar{2} = 1$); and the corresponding exchange diagrams



Thus, we have

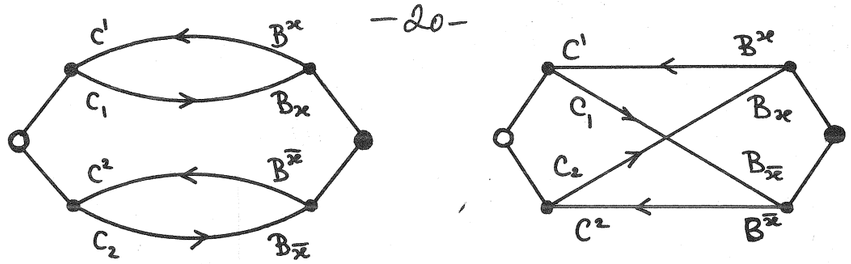
$$\left\langle \begin{matrix} c' \\ c_1 \end{matrix} \middle| \hat{H} \middle| \begin{matrix} B'B^2 \\ B_1B_2 \end{matrix} \right\rangle = \sum_{x=1}^{2}{}' \Big[\Delta^{1\bar{x}}\Delta_{1\bar{x}}\langle B_x|\hat{f}|B^x\rangle - \Delta^{1x}\Delta_{1\bar{x}}\langle B_x|\hat{f}|B^{\bar{x}}\rangle + \Delta_{1\bar{x}}\langle c'B_x|\hat{v}|B^{\bar{x}}B^x\rangle_A - \Delta^{1\bar{x}}\langle B_xB_{\bar{x}}|\hat{v}|B^xc_1\rangle_A \Big]$$

$$= \sum_{x=1}^{2}\Big[\sum_{\lambda=1}^{2} \Delta^{1\lambda}\Delta_{1\bar{x}}\langle B_x|\hat{f}|B^\lambda\rangle(-1)^{x+\lambda} + \Delta_{1\bar{x}}\langle c'B_x|\hat{v}|B^{\bar{x}}B^x\rangle - \Delta^{1\bar{x}}\langle B_xB_{\bar{x}}|\hat{v}|B^xc_1\rangle_A \Big]$$

Bi-Bi

$$\left\langle \begin{matrix} c'c^2 \\ c_1c_2 \end{matrix} \middle| \hat{H} \middle| \begin{matrix} B'B^2 \\ B_1B_2 \end{matrix} \right\rangle = E_0 \left\langle \begin{matrix} c'c^2 \\ c_1c_2 \end{matrix} \middle| \begin{matrix} B'B^2 \\ B_1B_2 \end{matrix} \right\rangle + \left\langle \begin{matrix} c'c^2 \\ c_1c_2 \end{matrix} \middle| \hat{H} - \langle \hat{H} \rangle \middle| \begin{matrix} B'B^2 \\ B_1B_2 \end{matrix} \right\rangle$$

Overlap:



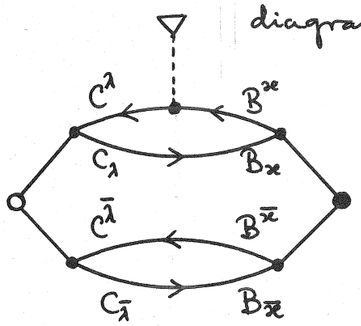
$x = 1,2$

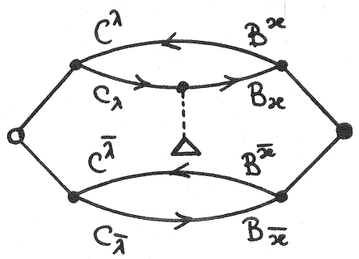
$$\left\langle \begin{matrix} C' C^2 \\ C_1 C_2 \end{matrix} \middle| \begin{matrix} B' B^2 \\ B_1 B_2 \end{matrix} \right\rangle = \sum_{x=1}^{2} \Delta^{1x}\, \Delta^{2\bar{x}} \left(\Delta_{1x}\, \Delta_{2\bar{x}} - \Delta_{1\bar{x}}\, \Delta_{2x} \right)$$

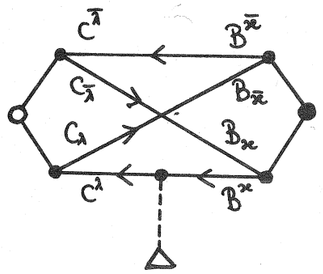
$$= \sum_{x,\lambda=1}^{2} (-1)^{x+\lambda}\, \Delta^{1x}\, \Delta^{2\bar{x}}\, \Delta_{1\lambda}\, \Delta_{2\bar{\lambda}}$$

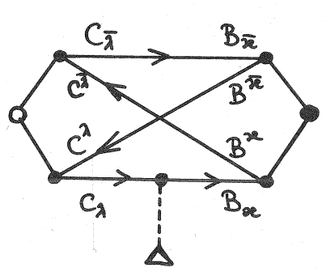
Energy diagrams:

<u>1-particle part</u> (we can obtain all graphs of this type by inserting a 1-part. vertex into all possible lines in the above overlap diagrams).









where

$\lambda = 1,2$

$x = 1,2$

and

$\bar{\lambda} = 1 + \delta_{1,\lambda}$

$\bar{x} = 1 + \delta_{1,x}$

($\Rightarrow$ 16 diagrams)

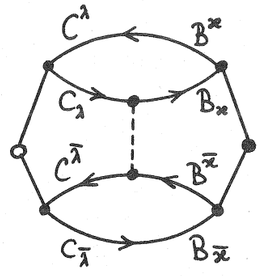
$$E_1 = \sum_{x,\lambda=1}^{2} \Big[\langle C^\lambda | \hat{f} | B^x \rangle\, \Delta^{\bar{\lambda}\bar{x}} \left(\Delta_{\lambda x}\, \Delta_{\bar{\lambda}\bar{x}} - \Delta_{\lambda\bar{x}}\, \Delta_{\bar{\lambda}x} \right)$$

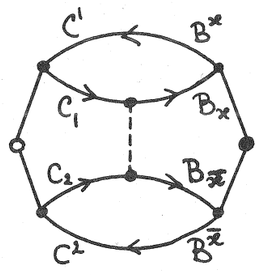
$$- \langle B_x | \hat{f} | C_\lambda \rangle\, \Delta_{\bar{\lambda}\bar{x}} \left(\Delta^{\lambda x}\, \Delta^{\bar{\lambda}\bar{x}} - \Delta^{\lambda\bar{x}}\, \Delta^{\bar{\lambda}x} \right)$$

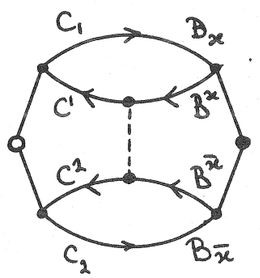
$$= \sum_{x,\lambda,\mu=1}^{2} (-1)^{x+\mu} \Big[\Delta^{\bar{\lambda}\bar{x}}\, \Delta_{\lambda\mu}\, \Delta_{\bar{\lambda}\bar{\mu}}\, \langle C^\lambda | \hat{f} | B^x \rangle -$$

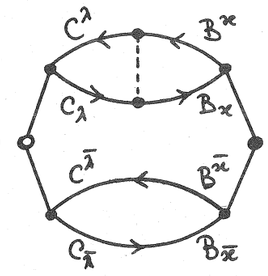
$$- \Delta_{\bar{\lambda}\bar{x}}\, \Delta^{\lambda\mu}\, \Delta^{\bar{\lambda}\bar{\mu}}\, \langle B_x | \hat{f} | C_\lambda \rangle \Big]$$

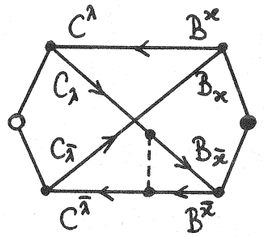
2-particle part (we draw first all the diagrams obtained by inserting a 2-particle vertex into the overlap diagrams: clearly, for each overlap diagram (having 4 oriented lines) this can be done in $\frac{1}{2}4(4-1) = 6$ possible ways: thus, we get $4 \times 6 = 24$ diagrams of this type; again, it suffices to draw only 8 of them explicitly.)

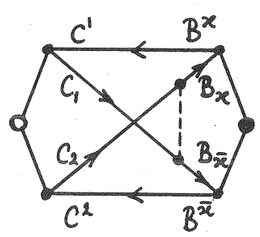


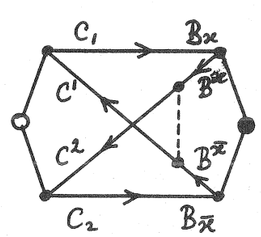


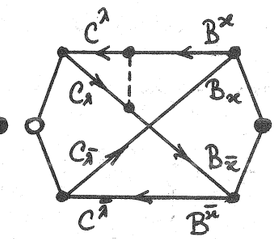








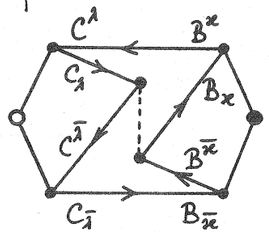




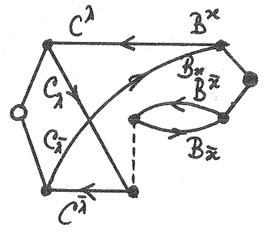
$x, \lambda = 1,2$ | $x = 1,2$ | $x = 1,2$ | $x, \lambda = 1,2$

4+4 (I) | 2+2 (II) | 2+2 (III) | 4+4 (IV)

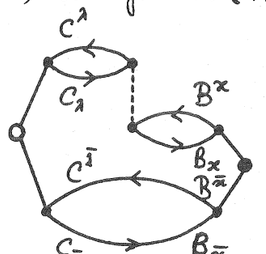
The above written diagrams are also called "forward" dgms (the p or h lines involved in the interaction proceed all the way from the T to $\bar{T}$ or vice versa); the only other diagrams may be obtained by a p-h exchange on the interaction vertex; this gives "backward" diagrams (the oriented lines involved in the interaction returns back to the T or $\bar{T}$ vertex where they started) - this p-h exchange is possible for 16 p-h diagrams above: thus, we find 16 p-h exchange (or backward) diagrams ($4 \times 4 = 16$):



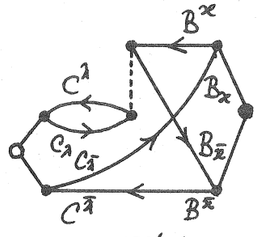
(I'a)



(I'b)



(IV'a)



(IV'b)

$x, \lambda = 1,2$

Thus, we get for the 2-particle contribution

$$E_2 = \sum_{\lambda,\varkappa=1}^{2} \Delta^{\lambda\varkappa}\left(-\Delta_{\bar\lambda\bar\varkappa}\langle C^{\bar\lambda} B_\varkappa|\hat v|B^{\bar\varkappa} C_\lambda\rangle + \Delta_{\bar\lambda\varkappa}\langle C^{\bar\lambda} B_{\bar\varkappa}|\hat v|B^{\bar\varkappa} C_\lambda\rangle\right) \quad \text{..... (I)}$$

$$+ \sum_{\lambda,\varkappa=1}^{2} \Delta^{\lambda\varkappa}\left(\Delta_{\bar\lambda\bar\varkappa}\langle C^{\bar\lambda} B_\varkappa|\hat v|C_\lambda B^{\bar\varkappa}\rangle - \Delta_{\bar\lambda\varkappa}\langle C^{\bar\lambda} B_{\bar\varkappa}|\hat v|C_\lambda B^{\bar\varkappa}\rangle\right) \quad \text{..... (I'a) \& (I'b)}$$

$$+ \sum_{\lambda,\varkappa=1}^{2} \Delta^{\bar\lambda\bar\varkappa}\left(-\Delta_{\bar\lambda\bar\varkappa}\langle C^{\lambda} B_\varkappa|\hat v|B^{\varkappa} C_\lambda\rangle + \Delta_{\bar\lambda\varkappa}\langle C^{\lambda} B_{\bar\varkappa}|\hat v|B^{\varkappa} C_\lambda\rangle\right) \quad \text{..... (II)}$$

$$+ \sum_{\lambda,\varkappa=1}^{2} \Delta^{\bar\lambda\bar\varkappa}\left(\Delta_{\bar\lambda\bar\varkappa}\langle C^{\lambda} B_\varkappa|\hat v|C_\lambda B^{\varkappa}\rangle - \Delta_{\bar\lambda\varkappa}\langle C^{\lambda} B_{\bar\varkappa}|\hat v|C_\lambda B^{\varkappa}\rangle\right) \quad \text{..... (II'a) \& (II'b)}$$

$$+ \sum_{\varkappa=1}^{2}\left(\Delta^{1\varkappa}\Delta^{2\bar\varkappa}\langle B_\varkappa B_{\bar\varkappa}|\hat v|C_1 C_2\rangle_A + \Delta_{1\varkappa}\Delta_{2\bar\varkappa}\langle C^1 C^2|\hat v|B^\varkappa B^{\bar\varkappa}\rangle_A\right) \quad \text{..... (II) \& (III)}$$

Notice also that: $\langle C^1 C^2|\hat v|B^\varkappa B^{\bar\varkappa}\rangle_A = (-1)^{\bar\varkappa}\langle C^1 C^2|\hat v|B^1 B^2\rangle_A$, $(\varkappa=1,2)$, for example.

Thus, writing the lines (I), (I'a), (I'b) and (II), (II'a), (II'b) together, we have

$$E_2 = \sum_{\varkappa,\lambda=1}^{2}\Big[\Delta^{\lambda\varkappa}\left(\Delta_{\bar\lambda\bar\varkappa}\langle C^{\bar\lambda} B_\varkappa|\hat v|C_\lambda B^{\bar\varkappa}\rangle_A - \Delta_{\bar\lambda\varkappa}\langle C^{\bar\lambda} B_{\bar\varkappa}|\hat v|C_\lambda B^{\bar\varkappa}\rangle_A\right)$$

$$+ \Delta^{\bar\lambda\bar\varkappa}\left(\Delta_{\bar\lambda\bar\varkappa}\langle C^{\lambda} B_\varkappa|\hat v|C_\lambda B^{\varkappa}\rangle_A - \Delta_{\bar\lambda\varkappa}\langle C^{\lambda} B_{\bar\varkappa}|\hat v|C_\lambda B^{\varkappa}\rangle_A\right)\Big]$$

$$+ \sum_{\varkappa=1}^{2}\left(\Delta^{1\varkappa}\Delta^{2\bar\varkappa}\langle B_1 B_2|\hat v|C_1 C_2\rangle_A + \Delta_{1\varkappa}\Delta_{2\bar\varkappa}\langle C^1 C^2|\hat v|B^1 B^2\rangle_A\right)(-1)^{\bar\varkappa}$$

$$= \sum_{\varkappa,\lambda,\mu=1}^{2}\Delta_{\bar\lambda\bar\mu}\left(\Delta^{\lambda\varkappa}\langle C^{\bar\lambda} B_\mu|\hat v|C_\lambda B^{\bar\varkappa}\rangle_A + \Delta^{\bar\lambda\bar\varkappa}\langle C^{\lambda} B_\mu|\hat v|C_\lambda B^{\varkappa}\rangle\right)(-1)^{\varkappa+\mu}$$

$$+ \sum_{\varkappa=1}^{2}\left(\Delta^{1\varkappa}\Delta^{2\bar\varkappa}\langle B_1 B_2|\hat v|C_1 C_2\rangle_A + \Delta_{1\varkappa}\Delta_{2\bar\varkappa}\langle C^1 C^2|\hat v|B^1 B^2\rangle_A\right)(-1)^{\bar\varkappa}$$

Thus, finally, we have: $\left\langle \begin{matrix} C^1 C^2 \\ C_1 C_2 \end{matrix}\middle|\, \hat H - \langle\hat H\rangle \,\middle| \begin{matrix} B^1 B^2 \\ B_1 B_2 \end{matrix}\right\rangle =$

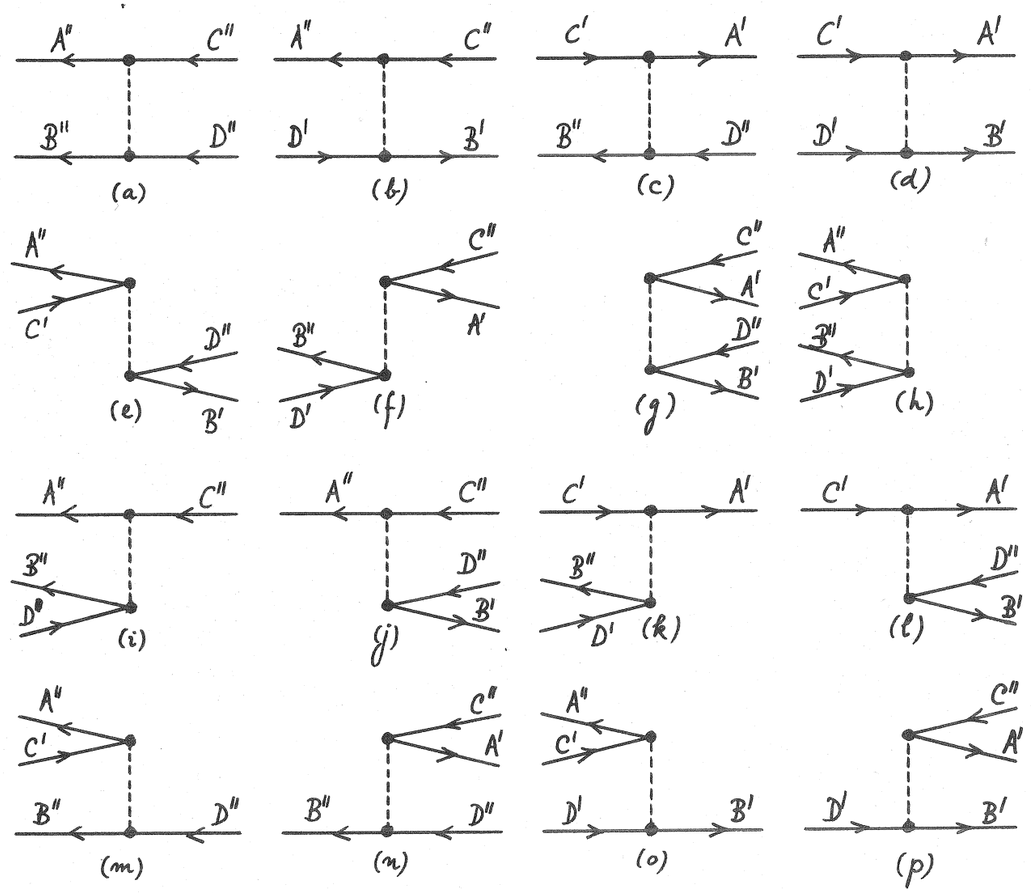
$$= \sum_{\varkappa,\lambda,\mu=1}^{2}(-1)^{\varkappa+\mu}\left[\Delta^{\bar\lambda\bar\varkappa}\Delta_{\lambda\mu}\Delta_{\bar\lambda\bar\mu}\langle C^\lambda|\hat f|B^\varkappa\rangle - \Delta_{\bar\lambda\bar\varkappa}\Delta^{\lambda\mu}\Delta^{\bar\lambda\bar\mu}\langle B_\varkappa|\hat f|C_\lambda\rangle\right]$$

$$+ \sum_{\varkappa,\lambda,\mu,\nu=1}^{2}\Delta^{\nu\varkappa}\Delta_{\bar\lambda\bar\mu}(-1)^{\varkappa+\lambda+\mu+\nu}\langle C^{\bar\nu} B_\mu|\hat v|C_\lambda B^{\bar\varkappa}\rangle_A$$

$$+ \sum_{\varkappa=1}^{2}(-1)^{\bar\varkappa}\left[\Delta^{1\varkappa}\Delta^{2\bar\varkappa}\langle B_1 B_2|\hat v|C_1 C_2\rangle_A + \Delta_{1\varkappa}\Delta_{2\bar\varkappa}\langle C^1 C^2|\hat v|B^1 B^2\rangle_A\right].$$

Ex. 17 Express graphically the two-particle interaction potential $\hat V$ in the particle-hole formalism (i.e., draw a graphical representation of individual terms from Ex. 12 using the usual conventions).

Solution: We write down the graphical representation of terms of $\hat V$ as they appear in Ex. 12:

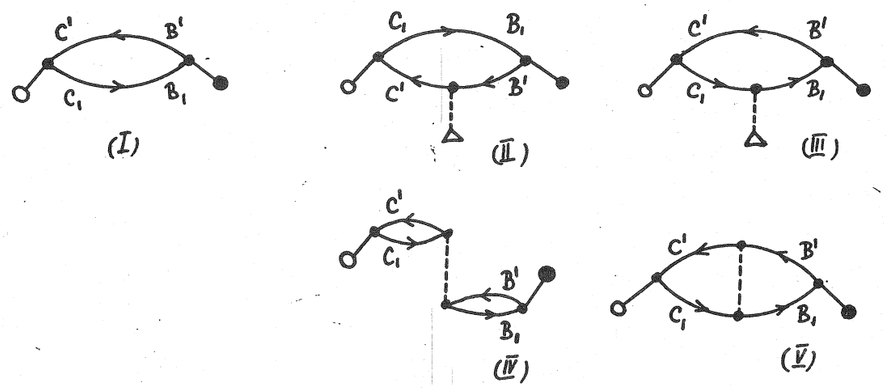


Observe, that in actual calculations of matrix elements of $\hat V$ most of these 16 terms yields a vanishing result. Thus, it is much easier to use the true-vacuum representation for $\hat V$, representing it by a single graph whose oriented lines may be arbitrarily distorted (preserving, of course, their orientation). Then the appropriate p-h components are picked out automatically in drawing the resulting diagrams.

Ex. 18 Compare a calculation of the matrix elements between the singly excited configurations using the N-product form for $\hat{H}$ (as carried out in Ex. 16) and using an ordinary form for $\hat{H}$.

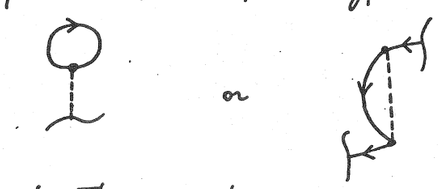
Sol^{n}: Recall that with the N-product form for $\hat{H}$ we got the following diagrams

(1)



since in view of the generalized Wick theorems all diagrams, which have internal lines between the vertices of the same supervertex, i.e. of the type

(2)



or (*)

must vanish. Thus we have

$$\left\langle \begin{matrix} C' \\ C_1 \end{matrix} \middle| \hat{H} \middle| \begin{matrix} B' \\ B_1 \end{matrix} \right\rangle = E_0 \times \vartheta(\mathrm{I}) + \vartheta(\mathrm{II}) + \vartheta(\mathrm{III}) + \vartheta(\mathrm{IV}) + \vartheta(\mathrm{V})$$

$$= E_0 \Delta'' \Delta_{22} + \langle C' | \hat{f} | B' \rangle \Delta_{11} - \langle B_1 | \hat{f} | C_1 \rangle \Delta'' + \langle C' B_1 | \hat{v} | C_1 B' \rangle_A \qquad (3)$$

where

$$\Delta^{\mu\nu} = \langle C^\mu | B^\nu \rangle \ , \quad \Delta_{\mu\nu} = \langle C_\mu | B_\nu \rangle$$

and $\vartheta(K)$ is the algebraic equivalent of the diagram K, $K = \mathrm{I}, \cdots, \mathrm{V}$.

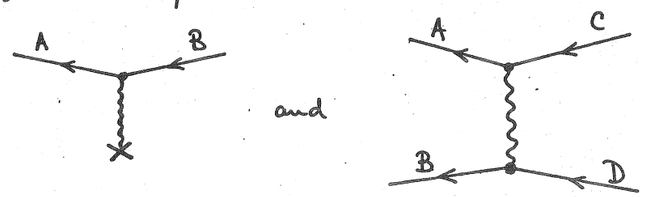
Now, using the ordinary form for $\hat{H}$,

$$\hat{H} = \sum_{A,B} \langle A | \hat{z} | B \rangle \hat{X}_A^+ \hat{X}_B + \frac{1}{2} \sum_{A,B,C,D} \langle AB | \hat{v} | CD \rangle \hat{X}_A^+ \hat{X}_B^+ \hat{X}_D \hat{X}_C \qquad (4)$$

the operator products originating in the Hamiltonian

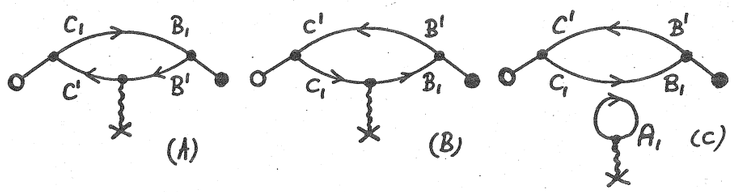
are not in the N-product form. Consequently, all nonvanishing contractions of these (and other) operators must be considered. Thus, even the contractions between these operators themselves, leading to the internal lines of the type (*), Eq.(2), must be included. In order to keep clear distinction between the two cases considered, we now introduce the following one- and two-particle vertices:

(5)

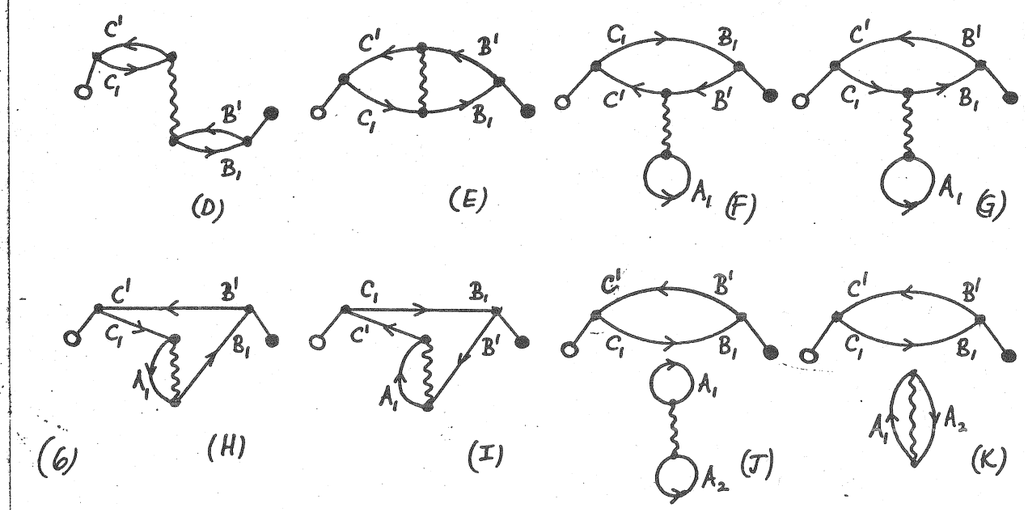


representing the one- and two-particle parts of $\hat{H}$, Eq. (4).

We now find the following resulting diagrams for the mtx elements between monoexcited states



(6)



Let us now compare both results. We recall first the definition of E_0 in Eq. (3)

$$E_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \sum_{A_1} \langle A_1 | \hat{z} | A_1 \rangle + \frac{1}{2} \sum_{A_1, A_2} \langle A_1 A_2 | \hat{v} | A_1 A_2 \rangle_A \qquad (7)$$

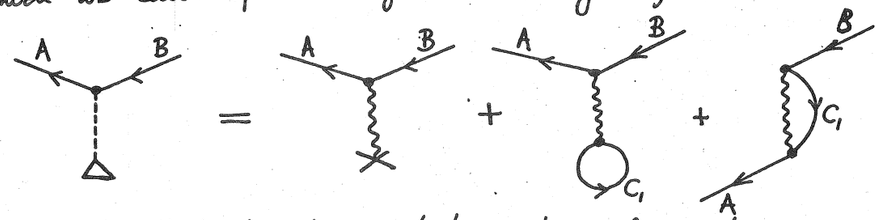
$$= \quad A_1 \quad + \quad A_1 \ A_2 \quad + \quad A_1 \ A_2$$

Further, we have that

(8) $$\langle A|\hat{f}|B\rangle = \langle A|\hat{v}|B\rangle + \sum_{C_1} \langle AC_1|\hat{v}|BC_1\rangle_A$$

which we can express diagrammatically as follows

(9)

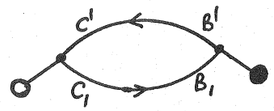


Notice also that the two-particle matrix elements are identical in both N-product and ordinary forms of the Hamiltonian, so that the two particle dashed and wavy lines may be interchanged, representing the same matrix elements. We must remember, however, that in the N-product form the diagrams with internal lines of the type (*), Eq.(2), are forbidden.

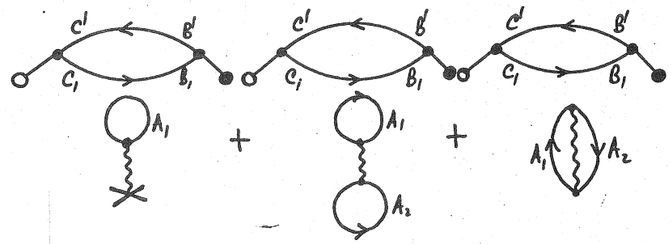
Using, now Eqs.(7-9) we can make a term by term contribution of the diagrams (1), Eq.(3) and of the diagrams (6). We find

(i)

$$E_0\, V(I) = E_0 \left\langle \begin{matrix} C' & B' \\ C_1 & B_1 \end{matrix} \right\rangle = E_0 \times$$

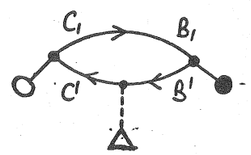


=

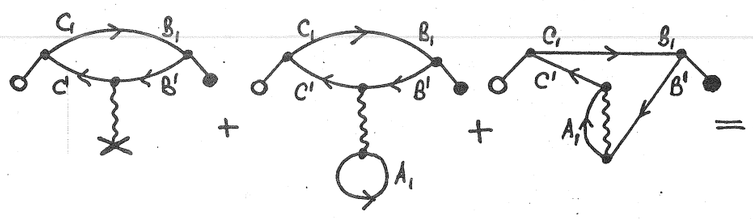


$$= V(C) + V(J) + V(K)$$

(ii) $V(II) = \langle C'|\hat{f}|B'\rangle \Delta_{II} =$



=



$$= V(A) + V(F) + V(I)$$

Similarly,

(iii) $$V(III) = -\langle B_1|\hat{f}|C_1\rangle \Delta'' =$$

$$= V(B) + V(G) + V(H)$$

Finally, we have

(iv) $V(IV) = V(D)$ and $V(V) = V(E)$

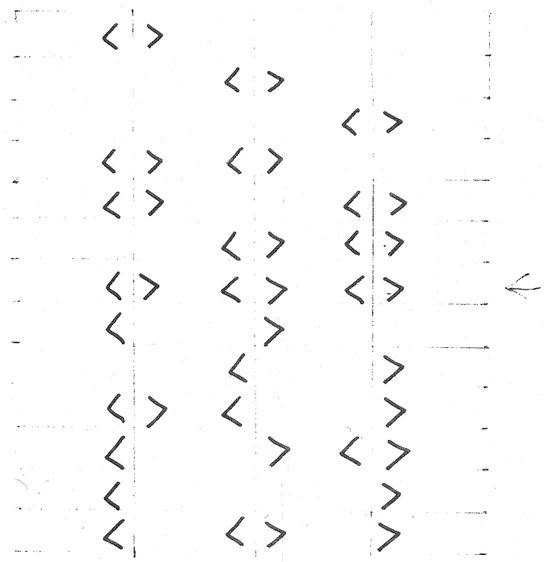
Thus, we have established the relationship of both forms for the mono-mono matrix elements.

Ex. 19 Use bracketing technique to obtain the 5th order Rayleigh-Schrödinger perturbation theory contribution in terms of N-electron unperturbed states.

Sol^{n}: We must get $\frac{(2\cdot5-2)!}{5!(5-1)!} = \frac{8!}{5!\,4!} = \frac{8\cdot7\cdot6}{4\cdot3\cdot2} = 14$ terms.

Let's write down the principal term and underneath all the pertinent bracketings (in a schematic way):

$\langle W\ P\ W\ P\ W\ P\ W\ P\ W \rangle$



Thus

$$E^{(5)} = \langle W(PW)^4 \rangle \qquad W\,P|W\,P|W\,P^2\,W$$
$$- \langle W \rangle \left(\langle WP^2W(PW)^2 \rangle + \langle (WP)^2(PW)^2 \rangle + \langle W(PW)^2P^2W \rangle \right)$$
$$+ \langle W \rangle^2 \left(\langle WP^3WPW \rangle + \langle WP^2WP^2W \rangle + \langle WPWP^3W \rangle \right)$$
$$- \langle W \rangle^3 \langle WP^4W \rangle - \langle WPW \rangle \left(\langle WP^2WPW \rangle + \langle WPWP^2W \rangle \right)$$
$$+ 2\langle W \rangle \langle WPW \rangle \langle WP^3W \rangle$$
$$- \langle WPWPW \rangle \langle WP^2W \rangle + \langle W \rangle \langle WP^2W \rangle^2$$

Exercise 20: Obtain explicit form of the second order P.T. contributions using both an algebraic and diagrammatic method. Compare the two.

Sol^{n}:

$$k_0^{(2)} = \langle \Phi_0 | \hat{W}\hat{P}_0 W | \Phi_0 \rangle$$
$$= \langle \Phi_0 | \hat{Q}_N \hat{P}_0 \hat{Q}_N | \Phi_0 \rangle + \langle \Phi_0 | \hat{V}_N \hat{P}_0 \hat{V}_N | \Phi_0 \rangle$$
$$+ \langle \Phi_0 | \hat{Q}_N \hat{P}_0 \hat{V}_N | \Phi_0 \rangle + \langle \Phi_0 | \hat{V}_N \hat{P}_0 \hat{Q}_N | \Phi_0 \rangle$$

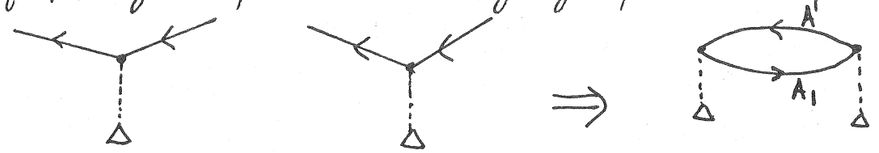
Note first that the last two terms vanish, since

$$\langle \Phi_0 | \hat{Q}_N \hat{P}_0 \hat{V}_N | \Phi_0 \rangle = \sum_{A,B} \langle A|\hat{q}|B\rangle \sum_{CDEF} \langle CD|\hat{v}|EF\rangle \sum_{k=1}^{N} \sum_{K} \frac{1}{-\omega_K^{(k)}} \times$$
$$\times \langle \Phi_0 | N[X_A^+ X_B] | \Phi_K^{(k)} \rangle \langle \Phi_K^{(k)} | N[X_C^+ X_D^+ X_F X_E] | \Phi_0 \rangle$$

and no $|\Phi_K^{(k)}\rangle$ exists which yields both elements simultaneously different from zero (in order that the 1st elmt does not vanish, we must have $k=1$ while for the second one $k=2$ is required). Consider, thus, the first two terms:

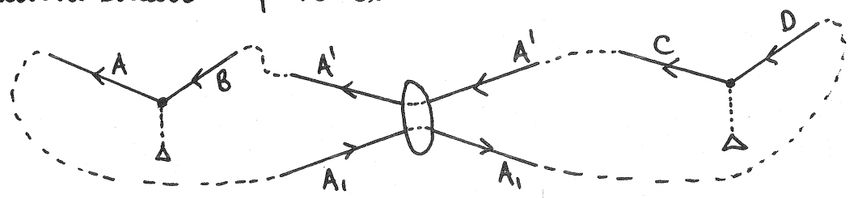
(i) $$k_0^{(2)1} = \langle \Phi_0 | \hat{Q}_N \hat{P}_0 \hat{Q}_N | \Phi_0 \rangle$$
$$= \sum_{k=1}^{N} \sum_{K} \frac{1}{-\omega_K^{(k)}} \langle \Phi_0 | \hat{Q}_N | \Phi_K^{(k)} \rangle \langle \Phi_K^{(k)} | \hat{Q}_N | \Phi_0 \rangle$$
$$= \sum_{A_1', A_1} \frac{1}{\omega_{A_1} - \omega_{A'}} \langle \Phi_0 | \hat{Q}_N | {}^{A'}_{A_1} \rangle \langle {}^{A'}_{A_1} | \hat{Q}_N | \Phi_0 \rangle$$
$$= \sum_{A_1', A_1} \frac{1}{\omega_{A_1} - \omega_{A'}} \sum_{A_1 B_1 C_1 D} \langle A|\hat{q}|B\rangle \langle C|\hat{q}|D\rangle \langle \Phi_0 | N[\overrightarrow{X_A^+ X_B}] N[\overrightarrow{X_{A'}^+ X_{A_1}}] | \Phi_0 \rangle \times$$
$$\times \langle \Phi_0 | N[\overrightarrow{X_{A_1}^+ X_{A'}}] N[\overrightarrow{X_C^+ X_D}] | \Phi_0 \rangle$$
$$= \sum_{A_1', A_1} \frac{1}{\omega_{A_1} - \omega_{A'}} \sum_{ABCD} \langle A|\hat{q}|B\rangle \langle C|\hat{q}|D\rangle \langle A|A_1\rangle \langle B|A'\rangle \langle A_1|D\rangle \langle A'|C\rangle$$
$$= \sum_{A_1', A_1} \langle A_1|\hat{q}|A'\rangle \langle A'|\hat{q}|A_1\rangle \left(\omega_{A_1} - \omega_{A'}\right)^{-1}$$

Graphically, we form all resulting dgms from two q-vertices



which gives the same result.

We can also symbolize the contractions using the intermediate p-vertex :



Clearly, this yields the above diagram using the "denominator convention".

We also observe here that in fact $\hat{P}_0$ may be ignored when deriving the contraction schemes, since it will always contain the pertinent state we need as an intermediate state (we only have to eliminate $|\Phi_0\rangle\langle\Phi_0|$ case). We use this fact in working out the second term contribution:

(ii) $$k_0^{(2)\prime\prime} = \langle\Phi_0|\hat{V}_N\,\hat{P}_0\,\hat{V}_N|\Phi_0\rangle$$

We can thus consider a simple mtx element

$$k'' = \langle\Phi_0|\hat{V}_N\,\hat{V}_N|\Phi_0\rangle$$

and avoid disconnected terms while using the "denominator convention".

Thus

$$k'' = \langle\Phi_0|\hat{V}_N\,\hat{V}_N|\Phi_0\rangle$$

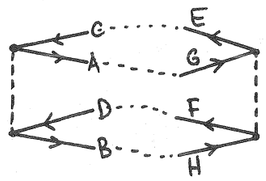
$$= \frac{1}{4}\sum_{A,B,C,D}\ \sum_{E,F,G,H}\langle AB|\hat{v}|CD\rangle\langle EF|\hat{v}|GH\rangle \times$$

$$\times\langle\Phi_0|N[A^+B^+DC]\,N[E^+F^+HG]|\Phi_0\rangle \quad (*)$$

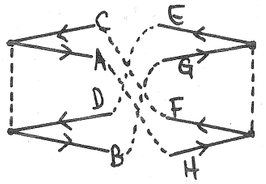
using a shorthand convention: $X_A^+ = A^+$, $X_B = B$ etc. We shall apply generalized Wick's theorem to the product of N-products keeping only fully contracted terms & we shall represent the contractions graphically on the rhs :

$$\langle\Phi_0|N[A^+B^+DC]\,N[E^+F^+HG]|\Phi_0\rangle$$

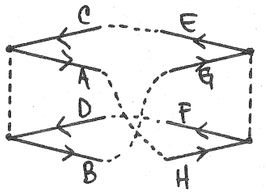
$$= \langle\Phi_0|N[\overbrace{A^+\,\overbrace{B^+\,\overbrace{D\,\overbrace{C\,\overbrace{E^+\,\overbrace{F^+\,H\,G}}}}}}]|\Phi_0\rangle$$



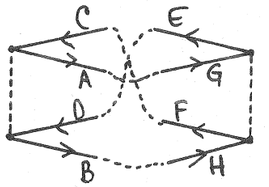
$$+\langle\Phi_0|N[\overbrace{A^+\,\overbrace{B^+\,D\,\overbrace{C\,\overbrace{E^+\,\overbrace{F^+\,H\,G}}}}}]|\Phi_0\rangle$$



$$+\langle\Phi_0|N[\overbrace{A^+\,B^+\,\overbrace{D\,C\,\overbrace{E^+\,\overbrace{F^+\,H\,G}}}}]|\Phi_0\rangle$$



$$+\langle\Phi_0|N[\overbrace{A^+\,B^+\,\overbrace{D\,C\,\overbrace{E^+\,\overbrace{F^+\,H\,G}}}}]|\Phi_0\rangle =$$



$$= h(A)h(B)p(C)p(D)\Big\{\langle A|G\rangle\langle B|H\rangle\langle D|F\rangle\langle C|E\rangle + \langle A|H\rangle\langle B|G\rangle\langle D|E\rangle\langle C|F\rangle - \langle A|H\rangle\langle B|G\rangle\langle D|F\rangle\langle C|E\rangle - \langle A|G\rangle\langle B|H\rangle\langle D|E\rangle\langle C|F\rangle\Big\}$$

Substituting into (*) we find

$$k'' = \frac{1}{4}\sum_{A_1B_1}\ \sum_{C'D'}\langle A_1B_1|\hat{v}|C'D'\rangle\big(\langle C'D'|\hat{v}|A_1B_1\rangle + \langle D'C'|\hat{v}|B_1A_1\rangle - \langle C'D'|\hat{v}|B_1A_1\rangle - \langle D'C'|\hat{v}|A_1B_1\rangle\big)$$

$$= \frac{1}{2}\sum_{\substack{A_1A_2\\A'A^2}}\langle A_1A_2|\hat{v}|A'A^2\rangle\big(\langle A'A^2|\hat{v}|A_1A_2\rangle - \langle A'A^2|\hat{v}|A_2A_1\rangle\big)$$

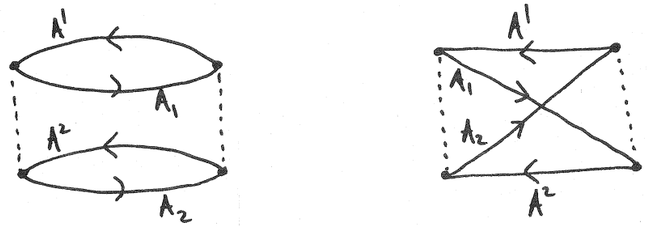
and using denominator rule (resp. going back to $k_0^{(2)\prime\prime}$)

$$k_0^{(2)\prime\prime} = \frac{1}{2}\sum_{\substack{A_1A_2\\A'A^2}}\frac{\langle A_1A_2|\hat{v}|A'A^2\rangle\big(\langle A'A^2|\hat{v}|A_1A_2\rangle - \langle A'A^2|\hat{v}|A_2A_1\rangle\big)}{\omega_{A_1}+\omega_{A_2}-\omega_{A'}-\omega_{A^2}}$$

We observe that the terms yielding identical result correspond in fact to the same diagrams, this is due to the fact that each diagram must be obtained $N_R = \frac{w_R}{w_a w_a} = \frac{\frac{1}{2}}{\frac{1}{2}\cdot\frac{1}{2}} = 2$ times

(note that the weight of resulting skeletons is also $\frac{1}{2}$).

Using diagrams, we find two resulting diagrams from the two V- vertices:



which yield the same formula as obtained above since:

$w = \frac{1}{2}$, $l = 2$, $h = 2$, $(-1)^{l+h} = +1$

$\langle A_1 A_2 | \hat{G} | A' A^2 \rangle \langle A' A^2 | \hat{v} | A_1 A_2 \rangle$

$w = \frac{1}{2}$, $l = 1$, $h = 2$, $(-1)^{l+h} = -1$

$\langle A_1 A_2 | \hat{v} | A' A^2 \rangle \langle A' A^2 | \hat{G} | A_2 A_1 \rangle$

denominator:

$$(\omega_{A_1} + \omega_{A_2} - \omega_{A'} - \omega_{A^2})$$

Ex. 21. Using the diagrammatic approach obtain the 3rd order contribution to P.T. for the spin-independent Hamiltonian.

Sol^n: Since in our convention we have $k_0^{(1)} = 0$, only principal term in the third order has to be considered, namely

$$k_0^{(3)} = \langle \Phi_0 | \hat{W} \hat{P}_0 \hat{W} \hat{P}_0 \hat{W} | \Phi_0 \rangle \quad \ldots\ldots \text{Eq. (78)}$$

G. Example: Third-Order Contribution

We shall first illustrate the above exposed theory on the derivation of the explicit spin-independent formulas for the third-order contribution $k_0^{(3)}$, Eq. (78), assuming that the Hartree–Fock solution is used as the unperturbed state, so that $\hat{Q}_N = 0$ and $\hat{W} = \hat{V}_N$.

We shall follow the rules of Section IV.D and first determine the pertinent Hugenholtz diagrams. To do so we must consider all possible ways of obtaining connected diagrams from three two-particle vertices, shown schematically in Fig. 13a. To simplify our problem we first consider a simpler case ignoring the orientation of lines. We also recall that we must not connect lines of the same vertex since we eliminated the HF diagrams. We can proceed in a systematic way considering that we connect vertices (I) and (II) of Fig. 13a with k lines, $k = 0, 1, 2, 3, 4$. We can immediately discard cases $k = 0$ and $k = 4$, since they would require that we join lines belonging to the same vertex. The same is also true for cases $k = 1$ and $k = 3$ [which are obtained from each other by a reflection

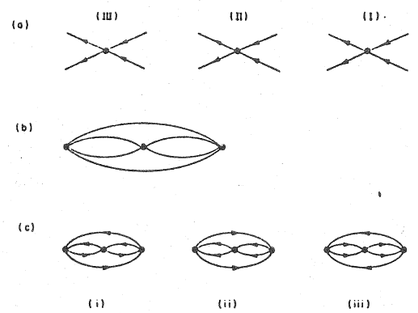


Fig. 13. Diagrammatic evaluation of the two-particle third-order contribution to the ground state energy. (a) Diagrammatic representation of the operators involved (in Hugenholtz form) and vertex numbering used. (b) Intermediate resulting diagram, in which the orientation of particle and hole lines is ignored. (c) Resulting Hugenholtz diagrams.

(I) ↔ (III)], so that only the case $k = 2$ yields a possible prototype of a third-order Hugenholtz diagram, shown in Fig. 13b. Considering now all possible orientations of lines, we find the three Hugenholtz diagrams (skeletons) shown in Fig. 13c. According to the type of interaction on the central vertex (II), we shall refer to them and to the corresponding terms as (i) particle-hole (p-h), (ii) particle-particle (p-p), and (iii) hole-hole (h-h) diagrams (terms).

We shall next find the corresponding Goldstone diagrams. A convenient way to effectuate this transition is (i) to first draw the Goldstone diagram having the maximum number of closed loops, (ii) to form exchange diagrams by "switching" two oriented lines at single V-vertices, pairs of V-vertices, etc., and (iii) to eliminate equivalent skeletons (diagrams). The resulting diagrams are shown in Fig. 14a. In the third column of Table I we have indicated at which vertices we carried out step (ii) above. We see that for the p-h diagram (i) of Fig. 13c all $2^3 = 8$ corresponding Goldstone diagrams are distinct, while only two are distinct in the p-p and h-h cases.

In order to write down the corresponding algebraic expressions, we first observe that we shall have certain matrix elements repeated a number of times in different diagrams. We therefore define

$$A_0 = \langle a''b'' | \hat{v} | a'b' \rangle \qquad A_1 = \langle a''b'' | \hat{v} | b'a' \rangle$$
$$B_0 = \langle b'c'' | \hat{v} | b''c' \rangle \qquad B_1 = \langle b'c'' | \hat{v} | c'b'' \rangle$$
$$C_0 = \langle a'c' | \hat{v} | a''c'' \rangle \qquad C_1 = \langle a'c' | \hat{v} | c''a'' \rangle \qquad (81)$$

for p-h terms and, similarly,

$$A_0' = \langle b_1'' b_2'' | \hat{v} | a_1' a_2' \rangle$$
$$B_0' = \langle c_1'' c_2'' | \hat{v} | b_1'' b_2'' \rangle$$
$$C_0' = \langle a_1' a_2' | \hat{v} | c_1'' c_2'' \rangle \qquad C_1' = \langle a_1' a_2' | \hat{v} | c_2'' c_1'' \rangle \qquad (82)$$

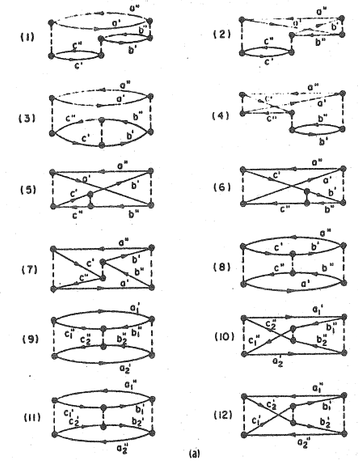


Fig. 14a. Goldstone form of the resulting diagrams for the two-particle (type A) third-order contribution to the ground state energy.

TABLE I

CONTRIBUTIONS FROM THE NON-HF THIRD-ORDER PT DIAGRAMS TO THE GROUND STATE ENERGY[a]

Hugenholtz diagram (Fig. 13c)	Goldstone diagram (Fig. 14a)	Exchange at vertices (cf. Figs. 13a,c)	w	l	h	Algebraic contribution of the diagram
(i)	1	NONE	1	3	3	$8R_{ph}A_0B_0C_0$
	2	I	1	2	3	$-4R_{ph}A_1B_0C_0$
	3	II	1	2	3	$-4R_{ph}A_0B_1C_0$
	4	III	1	2	3	$-4R_{ph}A_0B_0C_1$
	5	I, II	1	1	3	$2R_{ph}A_1B_1C_0$
	6	II, III	1	1	3	$2R_{ph}A_0B_1C_1$
	7	I, III	1	1	3	$2R_{ph}A_1B_0C_1$
	8	I, II, III	1	2	3	$-4R_{ph}A_1B_1C_1$
(ii)	9	NONE; I, II; I, III; II, III	½	2	2	$2R_{pp}A_0B_0C_0$
	10	I; II; III; I, II, III	½	1	2	$-R_{pp}A_0B_0C_1$
(iii)	11	NONE; I, II; I, III; II, III	½	2	4	$2R_{hh}A_0'B_0'C_0'$
	12	I; II; III; I, II, III	½	1	4	$-R_{hh}A_0'B_0'C_1'$

[a] The first and the second columns give the pertinent Hugenholtz and Goldstone diagrams, respectively; the third column lists the vertices at which the exchange has to be performed in the Goldstone diagram, containing maximum number of loops, in order to get a given diagram (cf. text for details); the next three columns list the weights (w), number of closed loops of oriented lines (l), and number of (internal) hole lines (h), and the last column gives the algebraic form of the given contribution [cf. Eqs. (81)-(84)].

and

$$A_0'' = \langle a_1'' a_2'' | \hat{v} | b_1' b_2' \rangle$$
$$B_0'' = \langle b_1' b_2' | \hat{v} | c_1' c_2' \rangle$$
$$C_0'' = \langle c_1' c_2' | \hat{v} | a_1'' a_2'' \rangle \qquad C_1'' = \langle c_1' c_2' | \hat{v} | a_2'' a_1'' \rangle \qquad (83)$$

for p-p and h-h terms, respectively.

Next we observe that the denominators of all Goldstone diagrams originating in the same Hugenholtz diagram are identical. Thus, the denominator factors will be

$$R_{ph} = [(\omega_{a'} + \omega_{c''} - \omega_{a''} - \omega_{c'})(\omega_{a'} + \omega_{b''} - \omega_{a''} - \omega_{b'})]^{-1}$$
$$R_{pp} = [(\omega_{a_1'} + \omega_{a_2'} - \omega_{c_1''} - \omega_{c_2''})(\omega_{a_1'} + \omega_{a_2'} - \omega_{b_1''} - \omega_{b_2''})]^{-1}$$
$$R_{hh} = [(\omega_{c_1'} + \omega_{c_2'} - \omega_{a_1''} - \omega_{a_2''})(\omega_{b_1'} + \omega_{b_2'} - \omega_{a_1''} - \omega_{a_2''})]^{-1}. \qquad (84)$$

Finally, the topological factor w, the number of closed loops l, and the number of hole lines h are easily determined for each Goldstone diagram of Fig. 14a and are given in Table I. The resulting algebraic expression is then given in the last column of Table I in terms of the quantities just defined. It should be noted how the product of matrix elements is related to the exchanges we have carried out in order to generate all Goldstone diagrams. It is not difficult to see that all these contributions may be conveniently expressed as follows

$$\begin{aligned} k_0^{(3)} = & \sum_{\substack{a',b',c' \\ a'',b'',c''}} R_{\mathrm{ph}}[(2A_0 - A_1)(2B_0 - B_1) \\ & \times (2C_0 - C_1) - 3A_1 B_1 C_1] \\ & + \sum_{\substack{a_1',a_2' \\ b_1'',b_2'',c_1'',c_2''}} R_{\mathrm{pp}} A_0' B_0' (2C_0' - C_1') \\ & + \sum_{\substack{b_1',b_2',c_1',c_2' \\ a_1'',a_2''}} R_{\mathrm{hh}} A_0'' B_0'' (2C_0'' - C_1''). \end{aligned} \tag{85}$$

Defining the antisymmetrized orbital matrix elements as follows

$$\langle ab \,|\, \bar{v} \,|\, cd \rangle_a = 2\langle ab \,|\, \bar{v} \,|\, cd \rangle - \langle ab \,|\, \bar{v} \,|\, dc \rangle \tag{86}$$

i.e.

$$X_a = 2X_0 - X_1, \; X = A, B, C, C', C'', \tag{87}$$

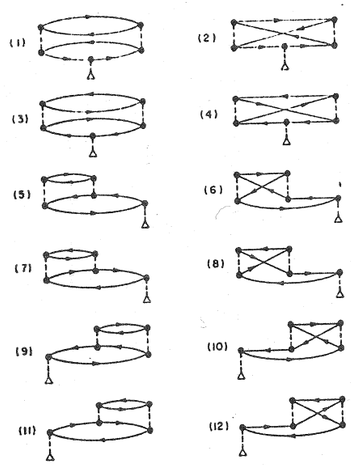
we get finally

$$k_0^{(3)} = \sum R_{\mathrm{ph}}(A_a B_a C_a - 3A_1 B_1 C_1) + \sum R_{\mathrm{pp}} A_0' B_0 C_a' + \sum R_{\mathrm{hh}} A_0'' B_0' C_a'', \tag{88}$$

where the summations extend over the same hole and particles index sets as in (85).

In a completely analogous way we can find the explicit spin-independent expressions for the nonHartree-Fock basis. Clearly, in this case, the two-particle (type A) contribution is again given by (88) and we only have to find the additional mixed contributions involving both Q- and V-vertices as well as the pure one-particle contribution.

Proceeding along the same lines as in the case of the two-particle contribution, we first find all pertinent resulting Goldstone diagrams and then write down immediately the explicit algebraic expressions using the rules given in Sections IV.D and F. The latter step being straightforward we only present here the relevant diagrams. We observe that the mixed

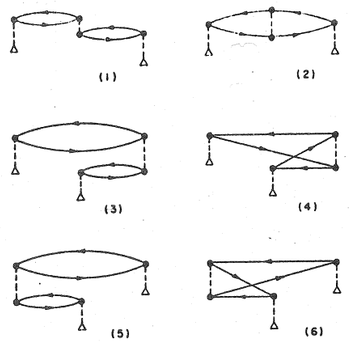


(b)

Fig. 14b. Resulting Goldstone diagrams for the mixed (type B, i.e. one Q- and two V-vertices) third-order contribution to the ground state energy.

terms can be of two basic types, depending on whether they contain one Q-vertex and two V-vertices (B type) or two Q- and one V-vertices (type C). The resulting Goldstone diagrams of both types are given in Figs. 14b and c, respectively. Finally, the one-particle contribution (type D, three Q-vertices) is given by the diagrams shown in Fig. 14d.

We would like to illustrate with this example one interesting feature of these resulting diagrams, which we shall conveniently exploit in the more complex excitation energy case treated in the next section. Consider, for example, a B-type mixed contribution given by the diagrams of Fig. 14b. We first observe that all the odd numbered diagrams are in fact the



(c)

Fig. 14c. Resulting Goldstone diagrams for the mixed (type C, i.e. two Q- and one V-vertices) third-order contribution to the ground state energy.

exchange diagrams corresponding to the direct (even numbered) ones. Moreover, all these odd (even) numbered diagrams are easily obtained from the first (second) listed diagram by permuting the left–right ordering of vertices in all possible 3! = 6 ways. Such grouping of diagrams, which differ only by a vertex permutation, follows naturally when using the time-dependent formalism of Goldstone (1957). The diagrams related in this way are referred to as different *time versions* of the same diagrams.

This grouping is particularly important in higher orders and in the excitation energy case, where we have to deal with large sets (several



(d)

Fig. 14d. Resulting Goldstone diagrams for the one-particle (type D) third-order contribution to the ground state energy.

hundreds or even thousands) of Goldstone diagrams. Indeed, this is one of the homomorphic "projections" used in our "double projection scheme" (Paldus and Čížek, 1974; Čížek and Paldus, 1972), which enables one not only to write down the compact explicit expressions for higher orders of the PT, but also makes possible their efficient numerical evaluation. For example, in the considered case of B-type diagrams of Fig. 14b we see that all these diagrams are easily generated from a single Hugenholtz diagram, which will be referred to as the *essentially distinct diagram* (cf. Paldus and Wong, 1973; also Section V).

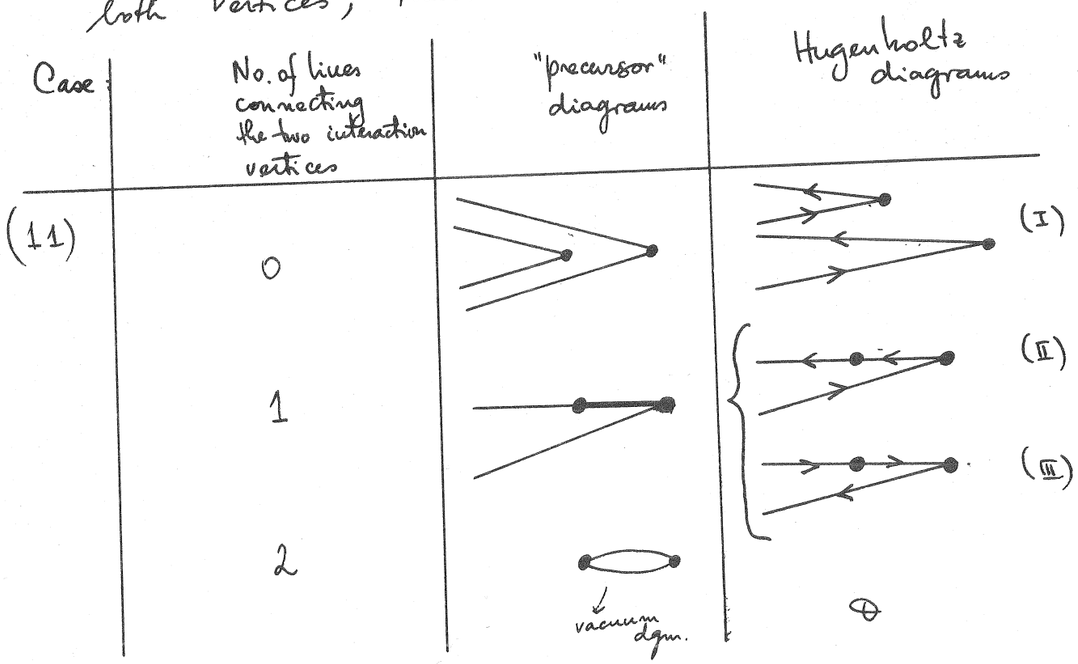
Ex. 22 Find an explicit form of the second order perturbation theory contribution to the wavefunction using diagrammatic approach.

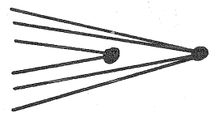
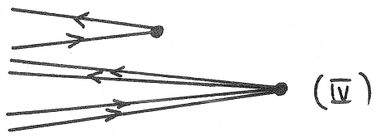
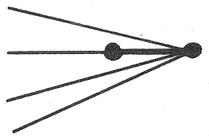
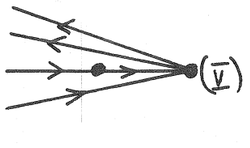
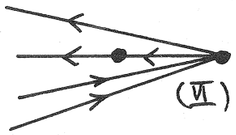
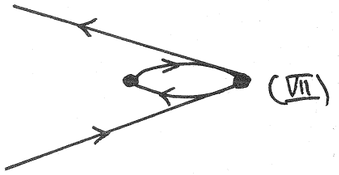
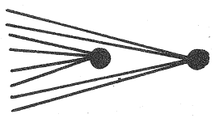
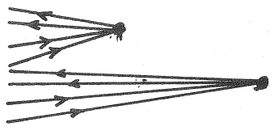
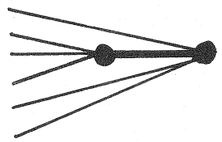
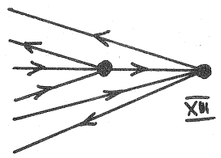
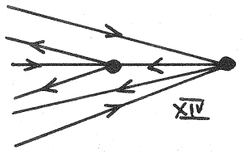
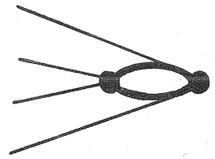
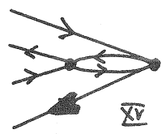
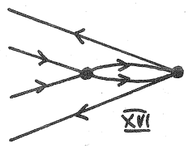
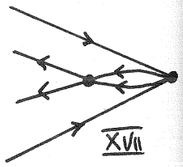
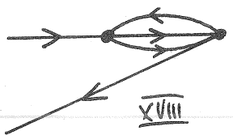
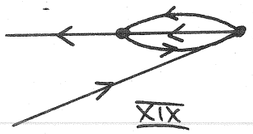
Sol^{n}: Using the N-product form for the perturbation, we have to find all the topologically distinct nonvacuum diagrams having two interaction vertices (recall that $|\Psi_0\rangle^{(2)} = \hat{P}_0 \hat{W} \hat{P}_0 \hat{W} |\Phi_0\rangle$ since $\langle \hat{W} \rangle = 0$ in our formulation).

We'll distinguish four basic types of diagrams depending on how many single particle vertices they involve:

Term:	No. of 1-p. vertices:	term represented:
(11)	2	$(\hat{P}_0 \hat{Q}_N)^2 \|\Phi_0\rangle$
(12)	1	$\hat{P}_0 \hat{Q}_N \hat{P}_0 \hat{V}_N \|\Phi_0\rangle$
(21)	1	$\hat{P}_0 \hat{V}_N \hat{P}_0 \hat{Q}_N \|\Phi_0\rangle$
(22)	0	$(\hat{P}_0 \hat{V}_N)^2 \|\Phi_0\rangle$.

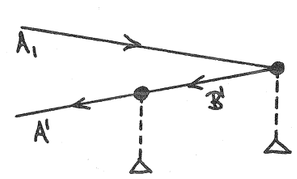
In line with previous considerations (cf. Ex. 21) we shall first draw the pertinent Hugenholtz diagrams; first we ignore the orientation of full lines and classify diagrams by the number of lines connecting both vertices; thus:

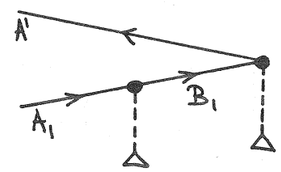


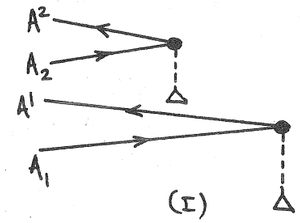
Case	No. of linear connected vertices	"precursors"	Hugenholtz diags
12	0		 (IV)
	1		 (V)  (VI)
	2		 (VII)
21	time versions of (12) ⇒ 0 1 2		VIII IX, X XI
22	0		 XII
	1		 XIII  XIV
	2		 XV  XVI  XVII
	3		 XVIII  XIX
	4	 vacuum	

We now form corresponding Goldstone diagrams and write down their contributions: (we order them by excitation order)

Case (11)


(II)

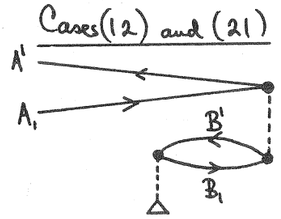

(III)

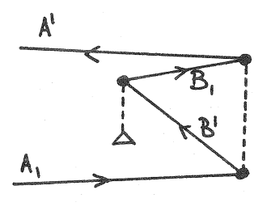

(I)

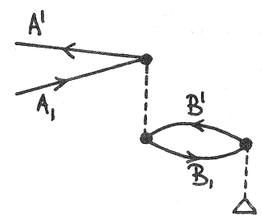
$$|\Psi_0\rangle^{(2)}_{(11)} = \sum_{A_1,A'} \frac{1}{\omega_{A_1}-\omega_{A'}} \left[\sum_{B'} \frac{\langle A'|\hat{q}|B'\rangle\langle B'|\hat{q}|A_1\rangle}{\omega_{A_1}-\omega_{B'}} - \sum_{B_1} \frac{\langle B_1|\hat{q}|A_1\rangle\langle A'|\hat{q}|B_1\rangle}{\omega_{B_1}-\omega_{A'}} \right] \left| \begin{matrix} A' \\ A_1 \end{matrix} \right\rangle +$$

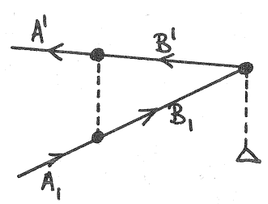
$$+ \sum_{A_1,A_2,A',A^2} \frac{\langle A^2|\hat{q}|A_2\rangle\langle A'|\hat{q}|A_1\rangle}{(\omega_{A_1}+\omega_{A_2}-\omega_{A'}-\omega_{A^2})(\omega_{A_1}-\omega_{A'})} \left| \begin{matrix} A' & A^2 \\ A_1 & A_2 \end{matrix} \right\rangle$$

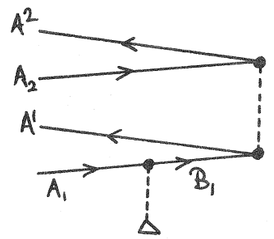
Cases (12) and (21)

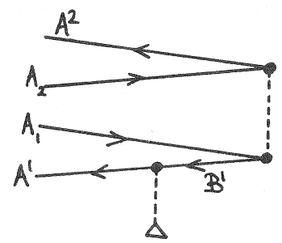

(VII-1)

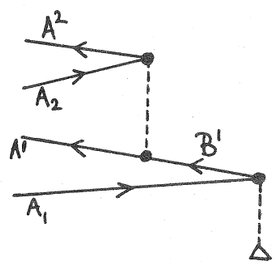

(VII-2)

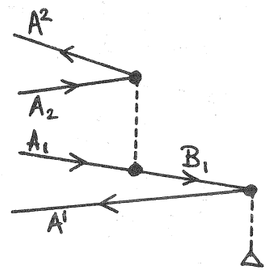

(XI-1)

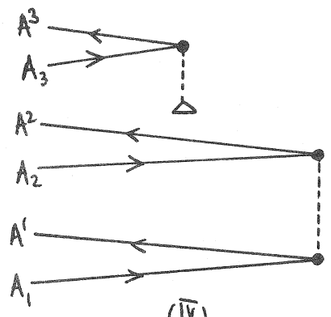

(XI-2)

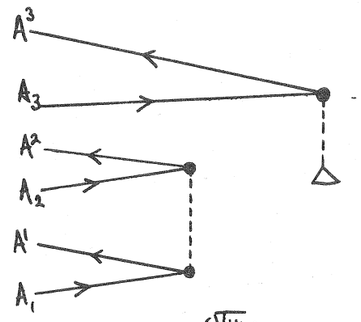

(V)


(VI)


(IX)


(X)


(IV)


(VIII)

$$|\Psi_0\rangle^{(2)}_{(1,2)} + |\Psi_0\rangle^{(2)}_{(2,1)} =$$

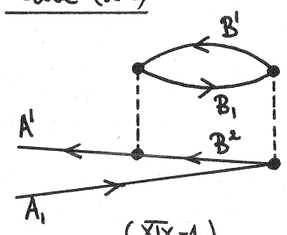
$$= \sum_{A_1,A'} \frac{1}{\omega_{A_1}-\omega_{A'}} \sum_{B_1,B'} \left[\frac{\langle B_1|\hat{q}|B'\rangle\langle A B'|\hat{v}|A_1 B_1\rangle_A}{\omega_{A_1}+\omega_{B_1}-\omega_{A'}-\omega_{B'}} + \frac{\langle A' B_1|\hat{v}|A_1 B'\rangle_A \langle B'|\hat{q}|B_1\rangle}{\omega_{B_1}-\omega_{B'}} \right] \left| \begin{matrix} A' \\ A_1 \end{matrix} \right\rangle +$$

$$+ \sum_{A_1,A_2,A',A^2} \frac{1}{\omega_{A_1}+\omega_{A_2}-\omega_{A'}-\omega_{A^2}} \left[\sum_{B'} \left(\frac{\langle A'|\hat{q}|B'\rangle\langle B' A^2|\hat{v}|A_1 A_2\rangle}{\omega_{A_1}+\omega_{A_2}-\omega_{B'}-\omega_{A^2}} + \frac{\langle A' A^2|\hat{v}|B' A_2\rangle\langle B'|\hat{q}|A_1\rangle}{\omega_{A_1}-\omega_{B'}} \right) \right.$$

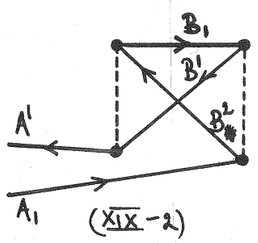
$$\left. - \sum_{B_1} \left(\frac{\langle B_1|\hat{q}|A_1\rangle\langle A' A^2|\hat{v}|B_1 A_2\rangle}{\omega_{B_1}+\omega_{A_2}-\omega_{A'}-\omega_{A^2}} + \frac{\langle B_1 A^2|\hat{v}|A_1 A_2\rangle\langle A'|\hat{q}|B_1\rangle}{\omega_{B_1}-\omega_{A'}} \right) \right] \left| \begin{matrix} A' & A^2 \\ A_1 & A_2 \end{matrix} \right\rangle +$$

$$+ \sum_{A_1,A_2,A_3,A',A^2,A^3} \frac{\langle A^3|\hat{q}|A_3\rangle\langle A' A^2|\hat{v}|A_1 A_2\rangle}{\omega_{A_1}+\omega_{A_2}+\omega_{A_3}-\omega_{A'}-\omega_{A^2}-\omega_{A^3}} \left(\frac{1}{\omega_{A_1}+\omega_{A_2}-\omega_{A'}-\omega_{A^2}} + \frac{1}{\omega_{A_3}-\omega_{A^3}} \right) \left| \begin{matrix} A' & A^2 & A^3 \\ A_1 & A_2 & A_3 \end{matrix} \right\rangle$$

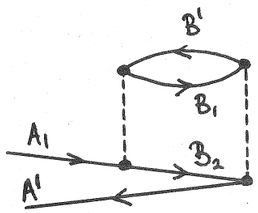
Case (22)



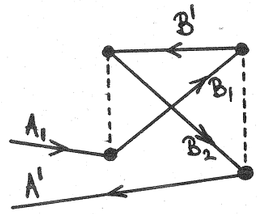
(XIX-1)



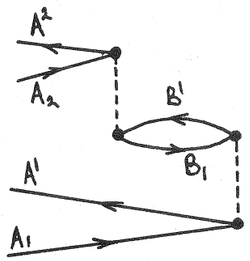
(XIX-2)



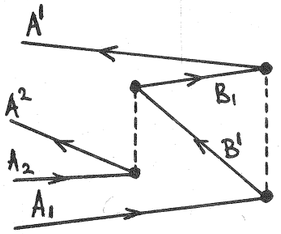
(XVIII-1)



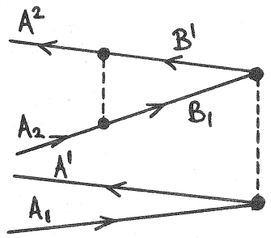
(XVIII-2)



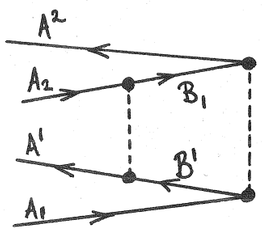
(XV-1)



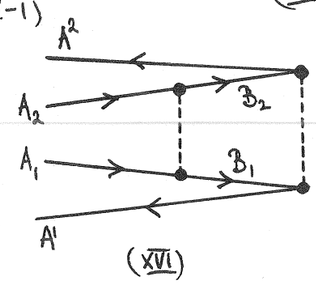
(XV-2)



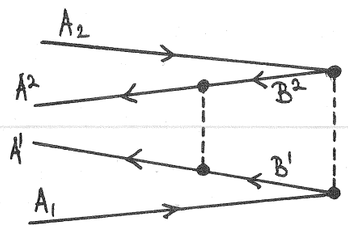
(XV-3)



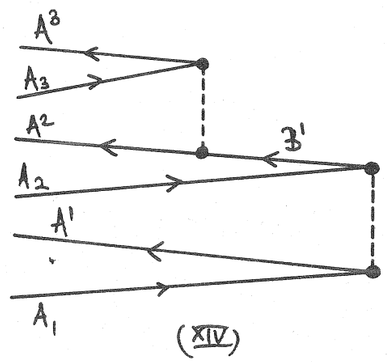
(XV-4)



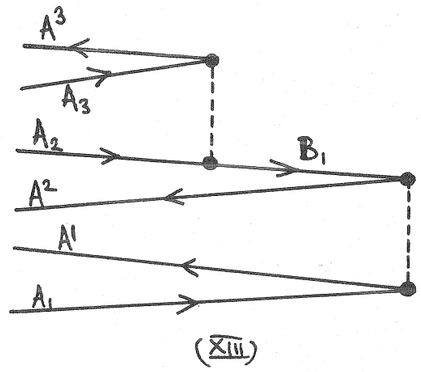
(XVI)



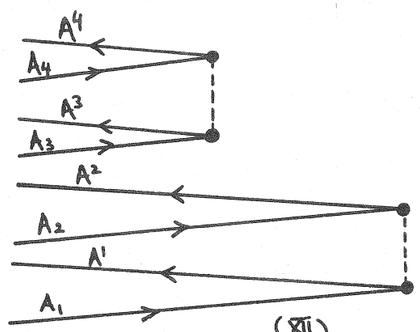
(XVII)



(XIV)



(XIII)



(XII)

$$|\Psi_0\rangle^{(2)}_{(2,2)} = \sum_{A_1,A'} \frac{1}{\omega_{A_1}-\omega_{A'}} \sum_{B_1,B'} \left[\sum_{B^2} \frac{\langle A' B_1|\hat{v}|B^2 B'\rangle_A \langle B' B^2|\hat{v}|B_1 A_1\rangle}{\omega_{A_1}+\omega_{B_1}-\omega_{B'}-\omega_{B^2}} - \right.$$

$$\left. - \sum_{B_2} \frac{\langle B_1 B_2|\hat{v}|B' A'\rangle_A \langle A' B'|\hat{v}|B_2 B_1\rangle}{\omega_{B_1}+\omega_{B_2}-\omega_{A'}-\omega_{B'}} \right] \left| \begin{matrix} A' \\ A_1 \end{matrix} \right\rangle +$$

$$+ \sum_{A_1,A_2,A',A^2} \frac{1}{\omega_{A_1}+\omega_{A_2}-\omega_{A'}-\omega_{A^2}} \left[\sum_{B_1,B'} \frac{\langle B_1 A^2|\hat{v}|B' A_2\rangle_A \langle A' B'|\hat{v}|A_1 B_1\rangle_A}{\omega_{A_1}+\omega_{B_1}-\omega_{A'}-\omega_{B'}} \right.$$

$$\left. + \frac{1}{2} \sum_{B_1,B_2} \frac{\langle B_1 B_2|\hat{v}|A_1 A_2\rangle\langle A' A^2|\hat{v}|B_1 B_2\rangle}{\omega_{B_1}+\omega_{B_2}-\omega_{A'}-\omega_{A^2}} + \frac{1}{2} \sum_{B',B^2} \frac{\langle A' A^2|\hat{v}|B' B^2\rangle\langle B' B^2|\hat{v}|A_1 A_2\rangle}{\omega_{A_1}+\omega_{A_2}-\omega_{B'}-\omega_{B^2}} \right] \left| \begin{matrix} A' & A^2 \\ A_1 & A_2 \end{matrix} \right\rangle$$

$$+ \sum_{A_1,A_2,A_3,A',A^2,A^3} \frac{1}{\omega_{A_1}+\omega_{A_2}+\omega_{A_3}-\omega_{A'}-\omega_{A^2}-\omega_{A^3}} \left[\sum_{B'} \frac{\langle A_2 A_3|\hat{v}|B' A_3\rangle\langle A' B'|\hat{v}|A_1 A_2\rangle}{\omega_{A_1}+\omega_{A_2}-\omega_{A'}-\omega_{B'}} - \right.$$

$$-\sum_{B_1}\frac{\langle B_1 A^3|\hat{v}|A_2 A_3\rangle\langle A^1 A^2|\hat{v}|A_1 B_1\rangle}{\omega_{A_1}+\omega_{B_1}-\omega_{A^1}-\omega_{A^2}}\Big]\left|\begin{matrix}A^1 A^2 A^3\\ A_1 A_2 A_3\end{matrix}\right\rangle +$$

$$+\frac{1}{4}\sum_{A_1,A_2,A_3,A_4,A^1,A^2,A^3,A^4}\frac{\langle A^1 A^2|\hat{v}|A_1 A_2\rangle\langle A^3 A^4|\hat{v}|A_3 A_4\rangle}{(\omega_{A_1}+\omega_{A_2}+\omega_{A_3}+\omega_{A_4}-\omega_{A^1}-\omega_{A^2}-\omega_{A^3}-\omega_{A^4})(\omega_{A_1}+\omega_{A_2}-\omega_{A^1}-\omega_{A^2})}\left|\begin{matrix}A^1 A^2 A^3 A^4\\ A_1 A_2 A_3 A_4\end{matrix}\right\rangle$$

I hope I did not miss any terms -
Should you find some terms missing, please let me know (I worked it out in a hurry.)

Clearly $|\Psi_0\rangle^{(2)} = \sum_{i,j=1}^{2} |\Psi_0\rangle^{(2)}_{(i,j)}$

Can you work out the rules for getting directly the spin-independent form?
It is not so easy in this case, but for singly & doubly excited, say, it is not so difficult.
General rules for spin-independent case would require angular-momentum diagrams (or, equivalently, graphical methods for spin algebras).

Relationship between configuration interaction and coupled cluster approaches

J. Paldus*, P.E.S. Wormer, F. Visser and A. van der Avoird
Institute of Theoretical Chemistry
Catholic University
Toernooiveld, Nijmegen
The Netherlands

Abstract

A variational principle for the linear coupled pair many electron theory (L-CPMET) is given and its implications for the relationship between certain configuration interaction (CI) and coupled cluster approaches (CCA), and for the computational aspects of solving the algebraic systems occurring in CCA, are discussed. An exact relationship between the correlated energies and wavefunctions as obtained with the L-CPMET and with the CI limited to at most doubly excited configurations (D-CI) is derived and used to provide a new viewpoint on the origins of Davidson's correction for unlinked cluster contributions in the D-CI energy. The results are illustrated on the CO molecule.

1. INTRODUCTION

The coupled cluster approach (CCA) to the many-electron correlation problem is based on the cluster expansion of the exact non-relativistic electronic wavefunction. It was used for the first time in the nuclear many fermion correlation problem by Coester and Kümmel[1]. The form of the CCA wavefunction guarantees the additivity of cluster components for non-interacting systems. However, in contrast to an ordinary configuration interaction (CI) linear expansion, the cluster expansion is not amenable to a variational approach. Even though the equations of the CCA are obtained by considering various moments of the Schrödinger equation, these are not associated with the usual variational principle of quantum mechanics and, consequently, the upper bound property for the energy is lost (cf., e.g., Ref. 2).

A general method of deriving the explicit form of the CCA equations was given by Čížek[3]. He has also derived these equations for the most important case of pair clusters thus formulating a coupled pair many electron theory (CPMET)[3]. Presently, a number of reviews and derivations of the CPMET equations and of their applications is available[4-13]. The orthogonally spin-adapted version of the CCA was also formulated[14,15] and, very recently, this form of the CPMET equations was used by Chiles and Dykstra[16] to obtain a computationally very promising SCEP (self-consistent electron pair) form[17] of these equations (essentially, by transforming the virtual orbital part back into the atomic orbital basis). The most recent survey of the CCA and its applications to various molecular systems was given by Bartlett[18].

The attractiveness of the CCA lies particularly in its size extensivity (cf., e.g., Ref. 11), which is especially important when studying various associative or dissociative processes. This is also the case for various perturbative approaches based on the Rayleigh-Schrödinger (RS) many body perturbation theory (MBPT) (cf., e.g., Ref. 18). In fact, the special version of this approach, in which all the MBPT diagrams with at most doubly excited intermediate states are summed to all orders (the so-called DMBPT[18]) is equivalent to the linear version of the CPMET[3] (the so-called L-CPMET). This approach yields correlation energies which are very close to the full CPMET correlation energies, unless the studied ground state is quasi-degenerate, as for example in the case of the beryllium atom. In the latter case the assumption of non-degeneracy represents too severe an approximation and one should in fact use the CCA for a (nearly) degenerate ground state recently developed by Lindgren[19]. The L-CPMET approach is also known as CEPA(0)[7,13], since it involves no explicit coupling of pair clusters (all non-linear terms in the CPMET approach are neglected).

Since the L-CPMET approach is closely related to perturbation theory it is not surprising that one can associate with it a variational functional of the Hylleraas type (cf. Ref. 2, p. 187) as we shall show in this paper. This viewpoint suggests in turn a very simple and useful relationship between the L-CPMET and the corresponding variational approach, the doubly excited CI (referred to as D-CI). This relationship elucidates from a new viewpoint the often used Davidson corrections[20-24] to the D-CI energies and enables us to both improve upon this correction and to obtain the corresponding approximate wavefunction. It also suggests the suitability of certain numerical methods for solving the CCA equations by providing a better understanding of the relationship between variational and coupled cluster approaches in general.

II. BASIC FORMALISM

We briefly introduce the necessary notation and the basic formalism for the coupled cluster and variational CI approaches. We consider a non-degenerate ground state of an N-electron system described by the non-relativistic electronic Hamiltonian $\hat{H}$. We designate the exact and the independent particle model (IPM) ground states by $|\Psi_0\rangle$ and $|\Phi_0\rangle$, respectively. We shall use the time-independent Schrödinger equation in the normal product form (relative to the IPM ground state $|\Phi_0\rangle$ taken as the Fermi vacuum of the hole-particle formalism, cf., e.g., Ref. 25)

$$\hat{H}_N|\Psi_0\rangle = \Delta E|\Psi_0\rangle, \tag{1}$$

where ΔE is the correlation energy when we choose the restricted Hartree-Fock (RHF) solution for $|\Phi_0\rangle$,

$$\Delta E = E_0 - \langle\Phi_0|\hat{H}|\Phi_0\rangle, \tag{2}$$

with E_0 being the lowest energy eigenvalue of the Hamiltonian $\hat{H}$. Imposing the intermediate normalization condition on $|\Psi_0\rangle$,

$$\langle\Psi_0|\Phi_0\rangle = 1, \tag{3}$$

the correlation energy is also given by the asymmetric energy formula

$$\Delta E = \langle\Phi_0|\hat{H}_N|\Psi_0\rangle. \tag{4}$$

1. CI approach

In the variational CI approach the exact state $|\Psi_0\rangle$ is expanded through $|\Phi_0\rangle$ by promoting one, two, three, etc. electrons from the spinorbitals $|B_i\rangle$ occupied in $|\Phi_0\rangle$ into the virtual spinorbitals $|B^i\rangle$. Using the intermediate normalization, Eq. (3), we can thus write

$$|\Psi_0\rangle = (1 + \hat{C})|\Phi_0\rangle, \tag{5}$$

where $\hat{C}$ is the sum of excitation operators $\hat{C}_i$,

$$\hat{C} = \sum_{i=1}^{N} \hat{C}_i. \tag{6}$$

The i-th component $\hat{C}_i$ of $\hat{C}$ will thus generate from $|\Phi_0\rangle$ an appropriate vector from the i-times excited subspace

$$\hat{C}_i|\Phi_0\rangle = \sum_I c_I^{(i)}|\Phi_I^{(i)}\rangle, \tag{7}$$

where $|\Phi_I^{(i)}\rangle$ is the i-times excited configuration state characterized by an ordered configuration I,

$$I = \{B^1B^2\ldots B^i;B_1B_2\ldots B_i\}, \tag{8}$$

obtained by promoting i electrons from the state $|B_1\rangle$, $|B_2\rangle$, ... , $|B_i\rangle$ occupied in the reference IPM state $|\Phi_0\rangle$ into the virtual states $|B^1\rangle$, $|B^2\rangle$, ... , $|B^i\rangle$, respectively, while $c_I^{(i)}$ designates the corresponding coefficient in the resulting CI wavefunction (5).

For a spin-independent Hamiltonian $\hat{H}$ we can choose an appropriate coupling scheme and consider the expansion (7) in terms of orthonormal spin-adapted configurations. Our conclusions will hold equally for the spinorbital formalism and for the spin-adapted formalism so that $|\Phi_I^{(i)}\rangle$ can in the following represent either an ordered spinorbital configuration state with I given by (8) or an orthonormal spin-adapted configuration, in which case I designates appropriate orbital labels and spin coupling quantum numbers (cf., e.g., Refs. 14,15 or 26).

The basic equations of the variational approach may be simply obtained by projecting the Schrödinger equation (1), with $|\Psi_0\rangle$ given by Eq. (5), onto the configuration states $|\Phi_I^{(i)}\rangle$ appearing linearly in $|\Psi_0\rangle$ (cf., e.g., Ref. 2)

$$\langle\Phi_I^{(i)}|\hat{H}_N|\Psi_0\rangle = \Delta E\,\langle\Phi_I^{(i)}|\Psi_0\rangle = \Delta E\, c_I^{(i)}. \tag{9}$$

Since $\hat{H}_N$ contains at most two-body operators, the left hand side of Eq. (9) has the form

$$\langle\Phi_I^{(i)}|\hat{H}_N|\Psi_0\rangle = \sum_{j=i-2}^{i+2} \sum_J c_J^{(j)} H_{IJ}^{(ij)}, \tag{10}$$

where

$$c_J^{(j)} = 0 \text{ if } j < 0; \; c_J^{(0)} \equiv c_0^{(0)} = 1 \tag{11}$$

and

$$H_{IJ}^{(ij)} = \langle\Phi_I^{(i)}|\hat{H}_N|\Phi_J^{(j)}\rangle. \tag{12}$$

Clearly,

$$H_{00}^{(00)} = \langle\Phi_0|\hat{H}_N|\Phi_0\rangle \equiv 0. \tag{13}$$

We thus obtain the well known chain of CI equations (i = 0,1,...,N; I ordered)

$$\sum_{j=i-2}^{i+2} \sum_J (H_{IJ}^{(ij)} - \Delta E \, \delta_{IJ}^{(ij)}) c_J^{(j)} = 0, \tag{14}$$

where

$$\delta_{IJ}^{(ij)} = \delta_{ij} \, \delta_{IJ}. \tag{15}$$

Considering, for the sake of simplicity, only the even-number-of-times excited configurations, the chain (14) has the form (note that $c_0^{(0)} = 1$)

$$\begin{aligned}
&\sum_J H_{0J}^{(02)} \, c_J^{(2)} = \Delta E, \\
&H_{I0}^{(20)} + \sum_J H_{IJ}^{(22)} c_J^{(2)} + \sum_J H_{IJ}^{(24)} \, c_J^{(4)} = \Delta E \, c_I^{(2)}, \\
&\sum_J H_{IJ}^{(42)} \, c_J^{(2)} + \sum_J H_{IJ}^{(44)} \, c_J^{(4)} + \sum_J H_{IJ}^{(46)} \, c_J^{(6)} = \Delta E \, c_I^{(4)}, \\
&\text{etc.},
\end{aligned} \tag{16}$$

where the equation for each i > 0 represents a system of equations for all possible configurations I. With some given finite set of (spin)orbital states, the full

system (14) or (16) will yield the full-CI solution. To obtain a limited-CI approximation, we decouple the above chain (14) or (16) at a certain excitation level i = m by neglecting all $H_{IJ}^{(ij)}$ matrix elements with j > m. Thus, for example, the doubly-excited CI equations (D-CI) are obtained by neglecting all $H_{IJ}^{(24)}$ matrix elements in the second equation (16), which gives

$$\begin{aligned}
&\sum_J a_J^* \, c_J = E_D \\
&a_I + \sum_J b_{IJ} c_J = E_D \, c_I,
\end{aligned} \tag{17}$$

where we designated

$$a_I = H_{I0}^{(20)}, \; b_{IJ} = H_{IJ}^{(22)} \text{ and } c_I = c_I^{(2)}, \tag{18}$$

with I, J = 1, ..., M, where M is the number of biexcited configurations $|\Phi_I^{(2)}\rangle$, m = 2, E_D is the D-CI approximation to ΔE, Eq. (2), and a* is the complex conjugate of a. In matrix form we can write (17) as

$$\underset{\sim}{H} \, \underset{\sim}{C} = E_D \, \underset{\sim}{C}, \tag{19}$$

where

$$\underset{\sim}{H} = \begin{bmatrix} 0 & \underset{\sim}{a}^\dagger \\ \underset{\sim}{a} & \underset{\sim}{b} \end{bmatrix}, \quad \underset{\sim}{C} = \begin{bmatrix} 1 \\ \underset{\sim}{c} \end{bmatrix}, \tag{20}$$

with $\underset{\sim}{a}$ and $\underset{\sim}{c}$ being column matrices with M entries a_I and c_I, respectively, and $\underset{\sim}{b}$ is the M×M matrix with entries b_{IJ}.

2. CC approach

In the coupled cluster approach the state $|\Psi_0\rangle$ is expanded through the same IPM configuration states $|\Phi_I^{(i)}\rangle$ as above, however, with the cluster expansion ansatz for the wave operator

$$|\Psi_0\rangle = e^{\hat{T}} \, |\Phi_0\rangle, \tag{21}$$

where again $\hat{T}$ is the sum of the i-fold excitation operators $\hat{T}_i$. Analogously to Eq. (7) we have that

$$\hat{T}_i|\Phi_0\rangle = \sum_I t_I^{(i)} |\Phi_I^{(i)}\rangle. \tag{22}$$

The relationship between both excitation operators, or between the corresponding expansion coefficients in Eqs. (7) and (22), follows immediately from a comparison of both expansions, Eqs. (5) and (21) (cf., e.g.,Refs. 27,28), yielding

$$\hat{C}_i = \sum_{\mathcal{P}_i} \prod_{j=1}^{i} (n_j!)^{-1} \hat{T}_j^{n_j} , \tag{23}$$

where the sum extends over all partitions $\mathcal{P}_i$ of i, i.e.

$$i = \sum_{j=1}^{i} (n_j)j \ , \ 0 \leq n_j \leq i; \quad \hat{T}_j^0 = \hat{1}. \tag{24}$$

The right hand side of Eq. (23) contains one term consisting of a single factor $\hat{T}_i$ which represents a connected i-times excited cluster component while the remaining terms, involving products of at least two lower order components, represent disconnected i-times excited clusters. In the MBPT language, all the terms in the expansions (5) and (21), except for the Fermi vacuum component $|\Phi_0>$, are represented by linked wavefunction type diagrams (i.e., the diagrams with at least one hole-particle pair of fermion lines in each disconnected part extending in the positive "time" direction). Further, the term $\hat{T}_i|\Phi_0>$ is represented by a connected diagram while the product terms $(T_1^{n_1}T_2^{n_2}\ldots)|\Phi_0>$ are represented by disconnected diagrams.

Assuming that $\hat{T}_1 = 0$ (which is exactly the case if Brueckner or maximum overlap (spin)orbitals are used), the first disconnected term appears in the quadruply excited component, namely

$$\begin{aligned} \hat{C}_1 &= \hat{T}_1 = 0 \\ \hat{C}_i &= \hat{T}_i \quad (i = 2,3) \\ \hat{C}_4 &= \hat{T}_4 + \tfrac{1}{2} \hat{T}_2^2. \end{aligned} \tag{25}$$

Since, further, $\hat{T}_4$ is usually negligible[29] relative to the corresponding disconnected component $\hat{T}_2^2$, the approximation $\hat{T} \approx \hat{T}_2$ represents in most cases (i.e. when $\hat{T}_3$ is not very important, cf., e.g., Refs. 18 and 30) an excellent approximation[3].

With the coupled cluster ansatz, Eq. (21), we can reduce the Schrödinger equation (1) to the following connected cluster form[3,14,15]

$$(\hat{H}_N \, e^{\hat{T}})_C \, |\Phi_0> = \Delta E \, |\Phi_0>, \tag{26}$$

where the subscript C indicates that only connected diagrams (or terms) are to

be considered. The equations which determine various cluster components (i.e., the coefficients $t_I^{(i)}$) are again obtained[14,15] by projecting Eq. (26) onto the configuration states $|\Phi_I^{(i)}>$, (i.e., by considering moments of Eq. (26) relative to the configuration states $|\Phi_I^{(i)}>$), even though the resulting equations are not variational equations in the sense of providing an upper bound to the energy. In this approach, the energy is given by the projection of Eq. (26) onto the reference state $|\Phi_0>$, i.e. by the asymmetric formula, Eq. (4),

$$\Delta E = <\Phi_0|(\hat{H}_N \, e^{\hat{T}})_C \, |\Phi_0>, \tag{27}$$

while the equations determining the cluster coefficients, which are obtained by projecting Eq. (26) onto the excited configurations,

$$<\Phi_I^{(i)}|(\hat{H}_N \, e^{\hat{T}})_C \, |\Phi_0> = 0, \tag{28}$$

are energy independent.

An explicit form of these equations is best obtained through the use of diagrammatic techniques (cf., e.g., Refs. 3,14,15,25,26). One sees immediately from Eq. (27) that ΔE depends at most on the first two components $\hat{T}_1$ and $\hat{T}_2$. Assuming that $\hat{T} \approx \hat{T}_2$ only the linear term survives so that the CPMET correlation energy E_{CP} is given by

$$E_{CP} = <\Phi_0|(\hat{H}_N \, \hat{T}_2)_C \, |\Phi_0>. \tag{29}$$

We also find that the system (28) decouples automatically obtaining

$$<\Phi_I^{(i)}|(\hat{H}_N + \hat{H}_N \, \hat{T}_2 + \tfrac{1}{2}\hat{H}_N \, \hat{T}_2^2)_C \, |\Phi_0> = 0, \tag{30}$$

or, more explicitly

$$a_I + \sum_J b_{IJ} \, t_J + \sum_{J \leq K} d_{IJK} \, t_J t_K = 0, \tag{31}$$

where a_I and b_{IJ} are given again by Eqs. (18), $t_J = t_J^{(2)}$, and d_{IJK} represents the connected part of the matrix element between the doubly excited configuration $|\Phi_I^{(2)}>$ and the quadruply excited configuration consisting of double excitations given by the configurations J and K. For an explicit form of these coefficients, see Ref. 3, or in the orthogonally spin-adapted case, Ref. 14.

Designating the column matrix of t_J coefficients as $\underline{t}$, we can also write Eq. (29) as

$$E_{CP} = \underset{\sim}{a}^{\dagger}\ \underset{\sim}{t}\ . \tag{32}$$

Neglecting the non-linear terms in the CPMET equations (31), we obtain the corresponding L-CPMET equations

$$\underset{\sim}{a} + \underset{\sim}{b}\ \underset{\sim}{t}_L = \underset{\sim}{0}, \tag{33}$$

$$E_L = \underset{\sim}{a}^{\dagger}\ \underset{\sim}{t}_L = -\underset{\sim}{a}^{\dagger}\ \underset{\sim}{b}^{-1}\ \underset{\sim}{a}, \tag{34}$$

where the subscript L indicates the linear approximation and where we have assumed that $\underset{\sim}{b}$ is non-singular in the last equation.

Let us mention, finally, that the same simple relationship between the D-CI matrix $\underset{\sim}{H}$, Eq. (20), and the corresponding L-CPMET equations, Eqs. (33), (34), holds even when both mono- and bi-excited configurations are considered and an RHF reference state is used (cf., e.g., Ref. 15). However, this is not the case when also triexcited configurations are taken into account (cf., diagram 7(c) of Ref. 15). We thus assume in the following that $\underset{\sim}{H}$, Eq. (20), and the corresponding CCA equations, involve either doubly excited configurations (D-CI) or both singly- and doubly excited configurations (sometimes referred to as the SD-CI), even though we shall always use the D-CI matrix $\underset{\sim}{H}$, Eq. (20), for simplicity's sake.

III. COMPARISON OF VARIATIONAL AND COUPLED CLUSTER APPROACHES

We now briefly outline the relationship between the variational limited CI approaches and the corresponding CCA's in order to understand better the derivations and the discussion in the subsequent sections. We shall in particular examine the simplest CCA, namely the L-CPMET and the corresponding D-CI approach.

Both D-CI and L-CPMET exploit the same information about the system considered as given by the D-CI matrix $\underset{\sim}{H}$, Eq. (20), or, equivalently, by its submatrices $\underset{\sim}{a}$ and $\underset{\sim}{b}$ with matrix elements defined by Eqs. (18). It should be noted that $\underset{\sim}{b}$ in Eq. (33) need not be hermitean, since we can multiply each equation of this linear system by an arbitrary non-vanishing scalar. This is in fact the case in the usual formulation of the CPMET and L-CPMET equations (cf., e.g., Refs. 3,14,15), where renormalized coefficients $t_I^{(i)}$ are often used. However, it is very easy to make a transition from these un-normalized t-matrix elements or coefficients to the normalized ones as is shown in Refs. 14,15, so that the assumption that the matrices $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are identical in the D-CI problem, Eq. (19), and in the CPMET, Eqs. (31), or the L-CPMET, Eqs. (33), problems represents no restriction on the generality of our considerations.

In order to see the conditions for the breakdown of the L-CPMET approach in more physical terms, as well as to clarify the relationship of this method with both variational and perturbative approaches, it is convenient to transform the D-CI matrix $\underset{\sim}{H}$, Eq. (20), to a form in which its doubly excited part is diagonal (cf., Refs. 31,32). Consider, thus, a unitarily equivalent matrix

$$\underset{\sim}{\tilde{H}} = \underset{\sim}{V}^{\dagger}\ \underset{\sim}{H}\ \underset{\sim}{V}, \tag{35}$$

where $\underset{\sim}{V}$ has the same block structure as $\underset{\sim}{H}$, Eq. (20),

$$\underset{\sim}{V} = \begin{bmatrix} 1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{v} \end{bmatrix}, \tag{36}$$

and where $\underset{\sim}{v}$ is the unitary matrix which diagonalizes $\underset{\sim}{b}$,

$$\underset{\sim}{b}\ \underset{\sim}{v} = \underset{\sim}{v}\ \underset{\sim}{\beta}, \tag{37}$$

with

$$\underset{\sim}{\beta} = [\beta_I\ \delta_{IJ}]. \tag{38}$$

Thus

$$\underset{\sim}{\tilde{H}} = \begin{bmatrix} 0 & \underset{\sim}{\alpha}^{\dagger} \\ \underset{\sim}{\alpha} & \underset{\sim}{\beta} \end{bmatrix}, \tag{39}$$

where

$$\underset{\sim}{\alpha} = \underset{\sim}{v}^{\dagger}\ \underset{\sim}{a}. \tag{40}$$

The transformation $\underset{\sim}{V}$ thus introduces a new basis in the doubly excited subspace.

Clearly, $\underset{\sim}{\tilde{H}}$, Eq. (39), has the same eigenvalues as the original D-CI matrix $\underset{\sim}{H}$, Eq. (20). It is easy to see that also the corresponding L-CPMET problem

$$\underset{\sim}{\alpha} + \underset{\sim}{\beta}\underset{\sim}{\tau} = \underset{\sim}{0}\ , \quad \tilde{E}_L = \underset{\sim}{\alpha}^{\dagger}\ \underset{\sim}{\tau} \tag{41}$$

yields the same energy as the original problem, Eqs. (33), (34),

$$\tilde{E}_L = \underset{\sim}{\alpha}^{\dagger}\ \underset{\sim}{\tau} = -\ \underset{\sim}{\alpha}^{\dagger}\ \underset{\sim}{\beta}^{-1}\ \underset{\sim}{\alpha} = -\ \underset{\sim}{a}^{\dagger}\ \underset{\sim}{b}^{-1}\ \underset{\sim}{a} = E_L, \tag{42}$$

since Eq. (37) implies that

$$\underset{\sim}{v}\ \underset{\sim}{\beta}^{-1}\ \underset{\sim}{v}^{\dagger} = \underset{\sim}{b}^{-1}, \tag{43}$$

assuming b is regular. In fact, the L-CPMET correlation energy, Eq. (42),can be written in the form

$$E_L = -\sum_{I=1}^{M} |\alpha_I|^2/\beta_I, \tag{44}$$

which shows immediately that E_L will have a singularity if $\beta_I \to 0$ for some I.

In a well behaved problem the matrix $\underset{\sim}{b}$ will be diagonally dominant and positive definite. Only when some of the doubly-excited configurations will be quasi-degenerate with the reference RHF state can we expect the lowest eigenvalue of $\underset{\sim}{b}$ to be very small or even negative. We have investigated the behaviour of the L-CPMET approximation in such quasi-degenerate cases elsewhere[31-34].

Using the canonical form (39) of the D-CI matrix we can also write the corresponding secular problem in a simple form. Expanding the associated secular determinant about the first row or column, or using simply the partitioning technique, we obtain

$$E_D = -\sum_{I=1}^{M} |\alpha_I|^2/(\beta_I - E_D), \tag{45}$$

the well known formula of the escalator diagonalization method.

Equations (44) and (45) also indicate a connection with corresponding perturbative approaches. Solving the L-CPMET equations (33) iteratively, using basically an algorithm due to Jacobi, we see immediately that we obtain an equivalent result to the DMBPT.[11,18] In fact, partitioning properly the diagonal terms, we can go continuously from an Epstein-Nesbet type perturbation theory to the Møller-Plesset type (cf., Ref. 32). Moreover, when the canonical form (39) is used, we obtain the desired L-CPMET or DMBPT energy already in the second order of Rayleigh-Schrödinger perturbation theory (RSPT), as Eq. (44) indicates. On the other hand, the D-CI energy can be obtained from the corresponding second order Brillouin-Wigner perturbation theory (BWPT), as follows from Eq. (45). It is well known that the RSPT is size extensive while BWPT is not, which is consistent with the characteristics of the D-CI and L-CPMET approaches.

The size extensivity of the L-CPMET (or, in fact, of the CCA in general) follows immediately from the cluster ansatz, Eq. (21), which transforms multiplicative structures into additive ones and which leads to the cancellation of all disconnected clusters, Eq. (26). In order to see this cancellation in more detail, let us consider the appropriate moments of the Schrödinger equation associated with the D-CI and L-CPMET cases.

In the D-CI case Eq. (9) can be written in the form

$$\langle \Phi_I^{(i)} | (\hat{H}_N - E_D)(1 + \hat{C}_2) | \Phi_0 \rangle = 0, \tag{46}$$

while the corresponding L-CPMET equations follow from Eq. (26),

$$\langle \Phi_I^{(i)} | \{\hat{H}_N(1 + \hat{T}_2)\}_C - E_L\ | \Phi_0 \rangle = 0, \tag{47}$$

where $i \leq 2$. The basic difference between these two systems of equations lies in the presence of the term $\langle \Phi_I^{(2)} | E_D\ \hat{C}_2 | \Phi_0 \rangle = E_D\ c_I^{(2)}$ in the D-CI system (46). This term corresponds to the unlinked diagrams in the perturbation theory approach and is therefore responsible for the size-inconsistency of the D-CI approach. In the L-CPMET (or CPMET) approach, these unlinked terms have been cancelled by the unlinked (and thus disconnected) component $\langle \Phi_I^{(2)} | (\frac{1}{2}\hat{H}_N\ \hat{T}_2^2)_{UL} | \Phi_0 \rangle$. This cancellation is implicit in the connected cluster form of the Schrödinger equation, Eq. (26).

It should be noted, however, that the cancelled term $\langle \Phi_I^{(2)} | E_D\ \hat{C}_2 | \Phi_0 \rangle$ contains both the EPV (exclusion principle violating) and the non-EPV components since the summations over the (spin)orbital labels in each factor are independent. From this viewpoint the L-CPMET "overcorrects" the D-CI result and the EPV terms are reintroduced through certain non-linear EPV connected terms in the CPMET approach. The most important terms in this respect are those represented by diagrams which can be separated over one or two hole lines: these terms play the most important role when going from the L-CPMET to the CPMET approach (cf., Refs. 31, 34, 35) and form the basis of various CEPA(k), $k \neq 0$, approaches (cf., Refs. 31,34).

In this respect it is interesting to compare the D-CI and the L-CPMET approaches for the case of two-electron systems. Clearly, the D-CI approach yields the exact energy within the model space considered while the L-CPMET does not. This is due to the fact that the latter approach includes a correction for the unlinked tetra-excited clusters (which is implicit in the connected form given by Eq. (26) as explained above). However, all tetra-excited components are unphysical in this case and thus of the EPV type. These terms would be subsequently cancelled by the non-linear CPMET terms, which are also always of the EPV type in the two-electron case considered, and the D-CI result would be restored. Clearly, the CPMET, or its simple version the L-CPMET, are designed for many-electron rather than two-electron systems, for which the number of non-EPV unlinked type contributions will be much larger than the number of similar EPV type terms. Therefore, the corrections for the unlinked tetra-excited terms, present in the D-CI approach, will be much more important than the "overcorrection" due to the inclusion of the EPV terms, which are necessary to achieve the factorization and which are corrected for when the proper "coupling" of pair clusters is considered in the CPMET approach.

IV. HYLLERAAS-TYPE FUNCTIONAL FOR THE L-CPMET

Since the L-CPMET can be regarded as a special case of perturbation theory as explained above, it is worthwhile to look for the Hylleraas-type functional (cf., Ref 2, p. 187) whose optimization yields an equivalent result. Such a functional must clearly contain the same information about the system as contained in the D-CI matrix $\underset{\sim}{H}$, Eq. (20). Such a functional might thus provide, at least in principle, a useful insight into the relationship of CI and CC approaches.

We recall that in the D-CI approach the correlation energy is given by the minimum of the following Rayleigh quotient

$$F_D(\underset{\sim}{Y}) = \underset{\sim}{Y}^\dagger \underset{\sim}{H}\, \underset{\sim}{Y}/(\underset{\sim}{Y}^\dagger \underset{\sim}{Y}), \tag{48}$$

where $\underset{\sim}{H}$ is the D-CI matrix (20) (with the diagonal entries relative to the reference RHF state energy) and $\underset{\sim}{Y}$ is an arbitrary trial vector. Since we wish to relate this variational approach with the L-CPMET approach, we shall use an intermediate normalization for $\underset{\sim}{Y}$ and write

$$\underset{\sim}{Y}_L = \begin{bmatrix} 1 \\ \underset{\sim}{x} \end{bmatrix}, \tag{49}$$

without restricting generality in any way.

The minimum condition for $F_D(\underset{\sim}{Y})$, Eq. (48), is given by the eigenvalue equation

$$\underset{\sim}{H}\, \underset{\sim}{U} = \underset{\sim}{U}\, \underset{\sim}{\Lambda}, \tag{50}$$

where $\underset{\sim}{\Lambda}$ is diagonal,

$$\underset{\sim}{\Lambda} = [\lambda_i\, \delta_{ij}], \tag{51}$$

with the lowest eigenvalue λ_0 giving the minimum of $F_D(\underset{\sim}{Y})$ which defines the desired correlation energy E_D,

$$E_D = \min F_D(\underset{\sim}{Y}) = F_D(\underset{\sim}{Y}^{(e)}) = \lambda_0, \qquad \underset{\sim}{Y}^{(e)} \sim \underset{\sim}{u}_0, \tag{52}$$

and $\underset{\sim}{U}$ is a unitary matrix whose i-th column $\underset{\sim}{u}_i$ represents an eigenvalue of $\underset{\sim}{H}$ which is associated with an eigenvalue λ_i,

$$\underset{\sim}{U}^\dagger \underset{\sim}{U} = \underset{\sim}{U}\, \underset{\sim}{U}^\dagger = \underset{\sim}{1}. \tag{53}$$

Recalling elementary properties of the diagonal form of a hermitean matrix, we can write

$$\underset{\sim}{H} = \underset{\sim}{U}\, \underset{\sim}{\Lambda}\, \underset{\sim}{U}^\dagger \tag{54}$$

or, equivalently, in terms of matrix elements

$$h_{ij} = \sum_k \lambda_k\, u_{ik}\, u_{jk}^*, \tag{55}$$

where we designated the matrix elements of $\underset{\sim}{H}$ and $\underset{\sim}{U}$ by corresponding lower case letters

$$\underset{\sim}{H} = [h_{ij}], \quad \underset{\sim}{U} = [u_{ij}]. \tag{56}$$

After recalling these elementary facts let us consider the problem outlined at the outset. Just as in perturbation theory we can construct the desired Hylleraas functionals in each order by inspection[36], it is easy to see that the L-CPMET equations (33) are obtained as a condition for an extremum of the functional

$$F_L(\underset{\sim}{Y}_L) \equiv f_L(\underset{\sim}{x}) = \underset{\sim}{Y}_L^\dagger \underset{\sim}{H}\, \underset{\sim}{Y}_L, \tag{57}$$

with $\underset{\sim}{Y}_L$ given by Eq. (49). Using an explicit form of the D-CI matrix $\underset{\sim}{H}$, Eq. (20), we obtain the following quadratic function

$$f_L(\underset{\sim}{x}) = \begin{bmatrix} 1 \\ \underset{\sim}{x} \end{bmatrix}^\dagger \begin{bmatrix} 0 & \underset{\sim}{a}^\dagger \\ \underset{\sim}{a} & \underset{\sim}{b} \end{bmatrix} \begin{bmatrix} 1 \\ \underset{\sim}{x} \end{bmatrix} = \underset{\sim}{x}^\dagger \underset{\sim}{a} + \underset{\sim}{a}^\dagger \underset{\sim}{x} + \underset{\sim}{x}^\dagger \underset{\sim}{b}\underset{\sim}{x}, \tag{58}$$

so that the conditions for an extremum of $f_L(\underset{\sim}{x})$ become

$$\underset{\sim}{a} + \underset{\sim}{b}\, \underset{\sim}{x} = \underset{\sim}{0}. \tag{59}$$

The conditions are identical with the L-CPMET equations (33) when we identify the unknowns $\underset{\sim}{x}$ and $\underset{\sim}{t}_L$.

The necessary and sufficient condition for an existence of a unique extremum is the definiteness of the matrix $\underset{\sim}{b}$. In this case $\underset{\sim}{b}^{-1}$ exists and we can easily verify that the L-CPMET energy E_L, as given by Eq. (34), can be obtained as an extremal value of the functional (57) or (58),

$$E_L = \underset{\sim}{a}^{\dagger}\underset{\sim}{x}_e = -\underset{\sim}{a}^{\dagger}\underset{\sim}{b}^{-1}\underset{\sim}{a} = f_L(\underset{\sim}{x}_e), \tag{60}$$

with $\underset{\sim}{x}_e$ designating the solution of the system (59).

Both functionals (48) and (57) represent clearly special cases of the same general functional G, given by the quadratic form

$$G(\underset{\sim}{X}) = \underset{\sim}{X}^{\dagger}\ \underset{\sim}{H}\ \underset{\sim}{X}. \tag{61}$$

Indeed, we have

$$F_D(\underset{\sim}{Y}) = G(\underset{\sim}{X}) = G(\underset{\sim}{Y}_D) \tag{62}$$

when

$$\underset{\sim}{X} = \underset{\sim}{Y}_D \equiv \underset{\sim}{Y}(\underset{\sim}{Y}^{\dagger}\underset{\sim}{Y})^{-\frac{1}{2}} = \underset{\sim}{Y}/\|\underset{\sim}{Y}\|, \tag{63}$$

and

$$F_L(\underset{\sim}{Y}_L) \equiv f_L(\underset{\sim}{x}) = G(\underset{\sim}{X}) = G(\underset{\sim}{Y}_L) \tag{64}$$

when

$$\underset{\sim}{X} = \underset{\sim}{Y}_L = \begin{bmatrix} 1 \\ \underset{\sim}{x} \end{bmatrix}. \tag{65}$$

This formulation provides an interesting geometrical viewpoint on the two approaches considered. Both the D-CI and the L-CPMET energies can be regarded as the extrema of the same functional but under different constraints: the trial vectors are restricted to the unit sphere $\|\underset{\sim}{X}\| = 1$ in the D-CI case and to the tangent plane $x_0 = 1$ to this sphere at the point corresponding to the reference RHF state in the L-CPMET case.

V. MODEL EXAMPLE

Before we discuss the implications of the above given functionals we wish to illustrate them in more detail on the simplest possible example: a two-electron, two-level model system with only one doubly excited configuration. This system will be described by the following D-CI matrix

$$\underset{\sim}{H}_1 = \begin{bmatrix} 0 & a \\ a & b \end{bmatrix}, \tag{66}$$

which we assume to be real in order to simplify our notation as much as possible. In this case, the general functional (61) is given by the following quadratic form in two variables

$$G_1(\underset{\sim}{X}) = G_1\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}^{\dagger}\begin{bmatrix} 0 & a \\ a & b \end{bmatrix}\begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

$$= x_1(2ax_0 + bx_1). \tag{67}$$

When restricting the variables $x_0 = y_0^D$, $x_1 = y_1^D$ to the unit circle,

$$(y_0^D)^2 + (y_1^D)^2 = 1, \tag{68}$$

we obtain the D-CI functional

$$F_D(\underset{\sim}{Y}) = G(\underset{\sim}{Y}_D) = G_1\begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}$$

$$= \sin\beta(2a\cos\beta + b\sin\beta)$$

$$= a\sin 2\beta + \tfrac{1}{2}b(1 - \cos 2\beta)$$

$$= \tfrac{1}{2}b\{1 - (1 + \tan^2 2\beta)^{-\frac{1}{2}}(1 - \tfrac{2a}{b}\tan 2\beta)\} \qquad [\text{assuming } |\beta| < \pi/4]$$

$$=: g_D(\beta) \tag{69}$$

where we have set

$$x_0 = y_0^D = \cos\beta = y_0(y_0^2 + y_1^2)^{-\frac{1}{2}}$$

$$x_1 = y_1^D = \sin\beta = y_1(y_0^2 + y_1^2)^{-\frac{1}{2}}. \tag{70}$$

Similarly, setting $x_0 = y_0^L = 1$ and $x_1 = y_1^L = x$, following Eq. (49), we obtain the L-CPMET functional

$$f_L(x) = G_1\begin{pmatrix} 1 \\ x \end{pmatrix} = x(2a + bx). \tag{71}$$

Both sets of variational parameters are related as follows

$$x = \frac{x_1}{x_0} = \frac{y_1}{y_0} = \frac{y_1^D}{y_0^D} = \tan\beta \tag{72a}$$

and

$$y_0^D = \cos\beta = (1 + x^2)^{-\frac{1}{2}}$$

$$y_1^D = \sin\beta = x(1 + x^2)^{-\frac{1}{2}}, \tag{72b}$$

so that we can also write the above functionals in the following equivalent forms

$$F_L \equiv f_L(x) \equiv g_L(\beta) = \tan\beta(2a + b\tan\beta) \tag{73a}$$

and

$$\begin{aligned} F_D \equiv g_D(\beta) &= x(2a + bx)/(1 + x^2) \\ &= F_L(\tan\beta)/(1 + \tan^2\beta) \\ &= g_L(\beta)\cos^2\beta. \end{aligned} \tag{73b}$$

The relationship between various variational parameters used above is illustrated in Fig. 1.

Thus, while in the D-CI approach we are searching for a minimum of the functional (67) along the unit circle (68) (cf., Fig. 1), in the L-CPMET approach we search for the minimum of the same functional along the tangent at unity ($x_0 = 1$, $x_1 = 0$), which corresponds to the RHF reference state used. Since the functional (67), or (61) in general, is a homogeneous quadratic form, we have for any scalar k that

$$G(k\,\underset{\sim}{X}) = k^2\, G(\underset{\sim}{X}). \tag{74}$$

We can also see immediately that both functionals (73a) and (73b) have the same nodes, at least in the physically relevant interval $-\pi/2 < \beta < \pi/2$.

The extremum of $f_L(x)$ will occur for $x = x_e^L$ given by the condition

$$\left.\frac{d\, f_L(x)}{dx}\right|_{x=x_e^L} = 2(a + bx)\Big|_{x=x_e^L} = 0, \tag{75}$$

which is equivalent to the corresponding L-CPMET equation, Eq. (33) or (59), for x_e^L

$$a + bx_e^L = 0 \tag{76}$$

so that

$$x_e^L = \tan\beta_e^L = -\,a/b,\ -\pi/2 < \beta_e^L < \pi/2 \tag{77}$$

and

$$E_L = a\, x_e^L = f_L(x_e^L) = -\,a^2/b = -\,b\tan^2\beta_e^L = g_L(\beta_e^L). \tag{78}$$

Similarly the extremum of $F_D(\underset{\sim}{Y}) = g_D(\beta)$ occurs at $\beta = \beta_e^D$ given by the condition

$$\left.\frac{dg_D(\beta)}{d\beta}\right|_{\beta=\beta_e^D} = 2a\cos 2\beta + b\sin 2\beta\Big|_{\beta=\beta_e^D} = 0, \tag{79}$$

which yields

$$\tan 2\beta_e^D = -\,2a/b, \tag{80}$$

so that we find immediately from the last Eq. (69) that

$$E_D = g_D(\beta_e^D) = \tfrac{1}{2}b\{1 - [1 + (2a/b)^2]^{\frac{1}{2}}\}, \tag{81}$$

which is clearly identical with the lowest eigenvalue of $\underset{\sim}{H}_1$, Eq. (66), assuming that $b > 0$.

The relationship between both extremal parameter values follows immediately from Eqs. (77) and (80), yielding (cf., also, Fig. 1)

$$\tan 2\beta_e^D = 2\tan\beta_e^L \tag{82}$$

or

$$\tan\beta_e^L = \tan\beta_e^D(1 - \tan^2\beta_e^D)^{-1}. \tag{83}$$

Geometrically, the L-CPMET extremum occurs when the gradient of $G_1(\underset{\sim}{X})$, Eq. (67), is perpendicular to the line $x_0 = 1$, while the D-CI minimum occurs when the gradient is normal to the unit circle, Eq. (68).

The above simple model may be easily extended by considering N non-interacting systems of this kind in which case the D-CI matrix, Eq. (20), has the form

$$\underset{\sim}{H}_N = \begin{bmatrix} 0 & \underset{\sim}{a}^T \\ \underset{\sim}{a} & \underset{\sim}{b}_1 \end{bmatrix},\quad \underset{\sim}{a}^T = a\,\underset{\sim}{\psi}_N \tag{84}$$

where

$$\underset{\sim}{\psi}_N = [\underbrace{1\ 1\ \ldots\ 1}_{N}]. \tag{85}$$

This model has been often employed[11,18,21,37-39] to examine the size-extensivity of the D-CI approach. This was done by comparing the D-CI energies with the exact one given by NE_D. We shall thus only concentrate on a comparison of the D-CI and the L-CPMET energies for this model, which are relevant to this paper.

In view of the symmetry of this model we have that

$$\underset{\sim}{X}^T = [x_0 \; \underset{\sim}{x}_1^T], \quad \underset{\sim}{x}_1^T = x_1 \; \underset{\sim}{\psi}_N \tag{86}$$

with $x_0 = 1$, $x_1 = x$ in the L-CPMET case and $\| \underset{\sim}{X} \| = \| \underset{\sim}{Y}_D \| = 1$ in the D-CI case, so that

$$x_0 = y_0^D = \cos\beta, \quad x_1 = y_1^D = N^{-\frac{1}{2}} \sin\beta. \tag{87}$$

The general functional, Eq. (61), is now simply a multiple of the corresponding functional G_1, Eq. (67), for one isolated system,

$$G_N(\underset{\sim}{X}) = N \, G_1\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}. \tag{88}$$

Correspondingly, the L-CPMET functional, Eq. (57), (58) is also additive

$$f_L^{(N)}(\underset{\sim}{x}) = N \, f_L(x) = N \, G_1\begin{pmatrix} 1 \\ x \end{pmatrix}, \tag{89}$$

while the D-CI functional will depend non-linearly on the particle number N,

$$F_D^{(N)}(\underset{\sim}{Y}) = N \, G_1\begin{pmatrix} \cos\beta \\ N^{-\frac{1}{2}} \sin\beta \end{pmatrix}. \tag{90}$$

We also note that the Hamiltonian matrix $\underset{\sim}{H}_N$, Eq. (84), is automatically in the canonical form (39) used in Sec. III so that

$$E_L^{(N)} = N \, E_L = - N \, a^2/b, \tag{91}$$

while the D-CI energy $E_D^{(N)}$ is given by Eq. (45), which in this case becomes

$$E_D^{(N)} = - N \frac{a^2}{b - E_D^{(N)}}, \tag{92}$$

so that

$$E_D^{(N)} = \tfrac{1}{2} b \{1 - [1 + N(2a/b)^2]^{\frac{1}{2}}\}, \tag{93}$$

which is also easily seen to minimize $F_D^{(N)}$, Eq. (90).

This result also nicely illustrates the size extensivity of the L-CPMET procedure in contrast to the D-CI method. Expanding the D-CI energy $E_D^{(N)}$, Eq. (93), about the L-CPMET value $E_L^{(N)}$, Eq. (91), we obtain

$$E_D^{(N)} = - N(a^2/b)\{1 - N(a/b)^2 + 2N^2(a/b)^4 - 5N^3(a/b)^6 + 14N^4(a/b)^8 - \ldots\}$$

$$= E_L^{(N)}\{1 - N(x_e^L)^2 + 2N^2(x_e^L)^4 - 5N^3(x_e^L)^6 + 14N^4(x_e^L)^8 - \ldots.\}. \tag{94}$$

The reason behind this very different behaviour of the D-CI and the L-CPMET (and CPMET) approaches may be easily understood from the geometrical properties of the variational spaces characterizing them. As we have seen above, the variational space in the D-CI approach is compact, namely a unit sphere in the (M+1)-dimensional Euclidean space, while in the L-CPMET case, the variational space is a linear space given by the tangent hyperplane to the unit sphere just mentioned, at the point characterizing the RHF reference state (cf., also, the above example). Thus, if one would regard the D-CI variational space as a Lie group then the corresponding L-CPMET variational space is given by the corresponding Lie algebra. Consequently, the independent systems behave additively in the linear L-CPMET space, so that this approach is size-extensive, while this is not the case for the D-CI space, which has a multiplicative rather than additive structure. The additive character of the L-CPMET space will be preserved in the CPMET approach even though the respective functional will be given by a form which is higher than quadratic.

This viewpoint nicely illustrates the basic incompatibility between the requirements of size extensivity and the upper bound property, as warranted by the use of the variational principle, as long as the trial functions are selected from the subspaces characterized by a rather limited excitation level.

VI. RELATIONSHIP BETWEEN THE D-CI AND THE L-CPMET APPROACHES

1. Exact relationship

We shall exploit the fact that both approaches and the corresponding functionals are closely related / and that they are defined through the same matrix representative of the Hamiltonian, as a comparison of Eqs. (48) and (57) indicates. To obtain the desired relationship, we thus could express the L-CPMET functional (57) in terms of the D-CI eigenvalues and eigenvectors by substituting for $\underset{\sim}{H}$ using Eq. (54), and then attempt to find an extremum of this functional explicitly. Indeed, this procedure would yield a system of linear equations which can be

easily solved in terms of λ_i and u_{ij}. Substituting this solution into the L-CPMET energy expression, Eq. (60), we would obtain the L-CPMET energy in terms of the variational energies λ_i and corresponding eigenvectors. Even though this procedure is not difficult to carry out, it is a bit laborious, and one can simply obtain the same result directly using the fact that the L-CPMET equations are related with the D-CI matrix as Eqs. (20) and (33) indicate.

Assume, thus, that none of the eigenvalues λ_i of $\underset{\sim}{H}$, Eq. (20), vanishes so that $\underset{\sim}{H}$ is invertible. Since $\underset{\sim}{H}^{-1}$ must also be hermitean, we can write it in the same block structure as the matrix $\underset{\sim}{H}$, namely

$$\underset{\sim}{H}^{-1} = \begin{bmatrix} p & \underset{\sim}{q}^{\dagger} \\ \underset{\sim}{q} & \underset{\sim}{r} \end{bmatrix}, \tag{95}$$

with $\underset{\sim}{r}$ hermitean,

$$\underset{\sim}{r}^{\dagger} = \underset{\sim}{r}. \tag{96}$$

The condition that $\underset{\sim}{H}^{-1}$, Eq. (95), is the inverse of $\underset{\sim}{H}$, Eq. (20), implies the following conditions for the component blocks of $\underset{\sim}{H}^{-1}$,

$$\begin{bmatrix} 0 & \underset{\sim}{a}^{\dagger} \\ \underset{\sim}{a} & \underset{\sim}{b} \end{bmatrix} \begin{bmatrix} p & \underset{\sim}{q}^{\dagger} \\ \underset{\sim}{q} & \underset{\sim}{r} \end{bmatrix} = \begin{bmatrix} 1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{1} \end{bmatrix}. \tag{97}$$

An explicit form of these conditions for the first column is clearly

$$\underset{\sim}{a}^{\dagger}\ \underset{\sim}{q} = 1 \tag{98}$$

$$\underset{\sim}{a}\ p + \underset{\sim}{b}\ \underset{\sim}{q} = \underset{\sim}{0}. \tag{99}$$

Since p is a non-zero scalar, this last equation is in fact identical with the L-CPMET system, Eq. (59), when we set

$$\underset{\sim}{x}_e = \underset{\sim}{t}_L = p^{-1}\ \underset{\sim}{q}. \tag{100}$$

We thus find for the L-CPMET energy that

$$E_L = \underset{\sim}{a}^{\dagger}\ \underset{\sim}{x}_e = p^{-1}\ \underset{\sim}{a}^{\dagger}\ \underset{\sim}{q} = p^{-1}, \tag{101}$$

where we have used Eq. (98) in the last step. Knowing the eigenvectors $\underset{\sim}{u}_i$ and the corresponding eigenvalues λ_i of $\underset{\sim}{H}$, Eq. (50), we can write for the elements of $\underset{\sim}{H}^{-1}$

$$(\underset{\sim}{H}^{-1})_{ij} = \sum_k \lambda_k^{-1}\ u_{ik}\ u_{jk}^*, \tag{102}$$

so that in the special case needed we have

$$p = (\underset{\sim}{H}^{-1})_{00} = \sum_k \lambda_k^{-1} |u_{0k}|^2. \tag{103}$$

We thus find that

$$E_L = \left[\sum_k |u_{0k}|^2/\lambda_k\right]^{-1}. \tag{104}$$

Further, using again Eq. (100),

$$\underset{\sim}{x}_e = p^{-1}\ \underset{\sim}{q} = E_L\ \underset{\sim}{q}, \tag{105}$$

and Eq. (102) for $\underset{\sim}{q}$,

$$q_i = (\underset{\sim}{H}^{-1})_{0i} = \sum_k u_{0k} u_{ik}^*/\lambda_k, \tag{106}$$

we find easily the components x_i, where x_i and q_i are components of $\underset{\sim}{x}$ and $\underset{\sim}{q}$ column vectors, respectively

$$\underset{\sim}{q} = [q_i],\ \underset{\sim}{x}_e = \underset{\sim}{t}_L = [x_i];\ i = 1, \ldots, M, \tag{107}$$

and M designates the number of doubly excited configurations considered.

We can thus summarize

$$E_D = \lambda_0 \tag{108}$$

$$E_L = \left(\sum_{j=0}^{M} \kappa_j\right)^{-1} \tag{109}$$

and

$$x_i = \sum_{j=0}^{M} (u_{ij}/u_{0j})^* \kappa_j \Big/ \sum_{j=0}^{M} \kappa_j, \tag{110}$$

where

$$\kappa_j = |u_{0j}|^2/\lambda_j. \tag{111}$$

Finally, we also note that with identifications given by Eqs. (100) and (101), the system of linear equations (98) and (99) is equivalent to the L-CPMET problem. This fact could be exploited in finding actual L-CPMET solutions when a D-CI matrix is available in either an implicit or an explicit form.

2. Approximate relationships

Let us now examine some simplifications and applications of the derived relationships, Eqs. (109) and (110), between the D-CI and the L-CPMET approaches. In turn, this analysis will enable us to see in a new light the very often employed simple approximations which correct the D-CI results for the size inconsistency due to the unlinked terms, particularly the well known Davidson correction[20].

We first note that the coefficients κ_j, Eq. (111), are given by the ratio of the probability factor $|u_{0j}|^2$, representing the weight of the RHF reference configuration in the j-th D-CI state, $j \geq 0$, and the j-th D-CI energy eigenvalue λ_j. Except in quasi-degenerate cases, the RHF reference configuration gives clearly the largest contribution to the D-CI ground state. We can thus assume that

$$|u_{00}|^2 >> |u_{0j}|^2 , \quad j > 0. \tag{112}$$

We also assume that the D-CI eigenstates are ordered according to their energies,

$$\lambda_j < \lambda_{j+1} , \; j \geq 0, \tag{113}$$

and, if no quasi-degeneracy is present, that there is a large energy gap between the ground state and the excited D-CI states, namely

$$\lambda_0 << \lambda_j , \; j > 0. \tag{114}$$

Then

$$|\kappa_0| = |u_{00}|^2/|\lambda_0| >> |u_{0j}|^2/|\lambda_j| = |\kappa_j|, \; j > 0, \tag{115}$$

so that retaining only the first term in the sum on the right hand side of Eq. (109) will yield a very good approximation $E_L^{(0)}$ to E_L. Thus,

$$E_L^{(0)} = \kappa_0^{-1} = \lambda_0/|u_{00}|^2, \tag{116}$$

which we can also write as

$$\begin{aligned} E_L^{(0)} &= \lambda_0 + \lambda_0(|u_{00}|^{-2} - 1) \\ &= E_D + E_D(1 - |u_{00}|^2)/|u_{00}|^2. \end{aligned} \tag{117}$$

This is exactly the approximation derived by Siegbahn[23], who used a perturbation theory estimate for the unlinked cluster contribution (cf., also Eq. (1') of Ref. 21 and Eq . (24) of Ref. 24). In fact this approximation is implicitly contained in the paper by Bartlett and Shavitt[22] (when Eq. (20) of Ref. 22 is solved exactly for S rather than expanded to the first order in S), who first showed explicitly the origin of this correction. When also $|u_{00}|^2 \approx 1$, the approximation (117) is very close to the well known Davidson correction[20]

$$E_L^{(D)} = E_D + E_D(1 - |u_{00}|^2) \ . \tag{118}$$

Thus both the original[20] and renormalized[21-24] Davidson corrections can also be regarded as the first approximation to the L-CPMET energy based on the D-CI results.

Using higher than the ground state D-CI eigenvalues and eigenvectors (or, more precisely, the reference state components of these eigenvectors) we can improve the approximation to E_L, which in the absence of quasi-degeneracy is always very close to the full CPMET result. In fact, the L-CPMET approach yields often better energies, and corresponding properties derived from the potential energy hypersurface near equilibrium geometry, than the CPMET approach (e.g., Ref. 40), since it overestimates the correction due to the unlinked tetra-excited clusters and this overestimate is in the same direction and of roughly the same magnitude as the correction to CPMET due to the connected triexcited clusters. Thus, employing other than the ground state D-CI eigenvalues λ_j $(j > 0)$ and the corresponding weights of the RHF ground state $|u_{0j}|^2$, the relationship (109) enables one to introduce further systematic corrections to D-CI. With increasing λ_j , the terms κ_j , Eq. (111), will generally decrease (though not monotonically) so that using a few terms in the relationships (109) and (110) might yield a result which is very close to the L-CPMET one. We illustrate this behaviour on an example in Sec. VIII.

Another very simple improvement may be obtained using an Unsöld type approximation. Thus, assuming that

$$\lambda_j \approx \overline{\lambda} \ , \ j > 0 \tag{119}$$

where $\overline{\lambda}$ is the average doubly excited state energy

$$\overline{\lambda} = \mathrm{Tr}(\underset{\sim}{b})/M, \ \mathrm{Tr}(\underset{\sim}{b}) = \sum_{K=1}^{M} b_{KK}, \tag{120}$$

and M is the number of doubly excited states used, we obtain

$$E_L^{(U)} = \{|u_{00}|^2/\lambda_0 + (1 - |u_{00}|^2)/\overline{\lambda}\}^{-1}. \tag{121}$$

We note that $E_L^{(U)}$ will tend to $E_L^{(0)}$ as the basis set used is extended since, in this case, $\overline{\lambda} \to \infty$. Obviously, in this case the assumption (120) is not valid and it should be replaced by a more realistic way of obtaining $\overline{\lambda}$.

Finally, we observe that the various approximate energies discussed above will be generally ordered as follows

$$E_L < E_L^{(U)} < E_L^{(0)} < E_L^{(D)} < E_D. \tag{122}$$

VII. COMPUTATIONAL IMPLICATIONS FOR THE SOLUTION OF CCA EQUATIONS

We outline briefly the implications of the above given results for the appropriateness of various numerical algorithms for the solution of algebraic systems of equations arising in the CCA.

For the L-CPMET approach, which may be regarded as the first approximation to the CPMET, the problem is equivalent to that of finding an appropriate stationary point of the hermitean quadratic function $f_L(\underset{\sim}{x})$, Eq. (58). This is most easily seen if we redefine the function $f_L(\underset{\sim}{x})$ in terms of the new variable $\underset{\sim}{z}$,

$$\underset{\sim}{z} = \underset{\sim}{v}^{\dagger} (\underset{\sim}{x} - \underset{\sim}{x}_e), \tag{123}$$

with $\underset{\sim}{v}$ defined by Eq. (37) and $\underset{\sim}{x}_e$ is the solution of the system (59). Substituting $\underset{\sim}{x}$ from Eq. (123), i.e.

$$\underset{\sim}{x} = \underset{\sim}{x}_e + \underset{\sim}{v} \underset{\sim}{z}, \tag{124}$$

we find easily that

$$\tilde{f}_L(\underset{\sim}{z}) \equiv f_L(\underset{\sim}{x}) = \underset{\sim}{z}^{\dagger} \underset{\sim}{\beta} \underset{\sim}{z} + f_L(\underset{\sim}{x}_e) \tag{125}$$

or, explicitly

$$\tilde{f}_L(\underset{\sim}{z}) = E_L + \sum_{K=1}^{M} \beta_K |z_K|^2. \tag{126}$$

Considering, for simplicity, the most often occurring case of a real D-CI matrix so that both the $\underset{\sim}{a}$ and $\underset{\sim}{b}$ submatrices of $\underset{\sim}{H}$ are real and $\underset{\sim}{b}$ is symmetric, we see immediately that $f_L(\underset{\sim}{x})$ is a quadratic function with gradient $\underset{\sim}{g}$,

$$\underset{\sim}{g}(\underset{\sim}{x}) = 2(\underset{\sim}{b} \underset{\sim}{x} + \underset{\sim}{a}), \tag{127}$$

and with the Hessian $\underset{\sim}{h}$,

$$\underset{\sim}{h} = 2\,\underset{\sim}{b}. \tag{128}$$

We also see immediately from Eq. (126) that $\underset{\sim}{z} = \underset{\sim}{0}$ or $\underset{\sim}{x} = \underset{\sim}{x}_e$ is a unique minimum point (or a strong global minimizer) of $\tilde{f}_L(\underset{\sim}{z}) \equiv f_L(\underset{\sim}{x})$ when $\beta_K > 0$ for all K, i.e., when $\underset{\sim}{b}$ is positive definite. This is usually the case unless the RHF ground state is quasi-degenerate (cf. the text following Eq. (44)).

Thus, when $\underset{\sim}{b}$ is positive definite the solution of the L-CPMET equations is equivalent to the problem of minimization of the functional $f_L(\underset{\sim}{x})$ suggesting the use of conjugate direction methods. Recently, these techniques for solving linear algebraic systems have been extensively studied[41], particularly in connection with large sparse systems.

It is well known that the convergence of the conjugate gradient method, as measured by the number of iterations which are required to lower the error of the initial approximation $\underset{\sim}{x}^{(0)}$ by a prescribed factor, is proportional to the square root of the spectral condition number $\mathcal{K}(\underset{\sim}{b})$ of the coefficient matrix $\underset{\sim}{b}$. The error is measured in the so called energy norm so that one minimizes the value of the functional $f_L(\underset{\sim}{x})$ rather than the distance $\|\underset{\sim}{x} - \underset{\sim}{x}_e\|$. The spectral condition number is defined as the ratio of the largest and the smallest eigenvalue of $\underset{\sim}{b}$,

$$\mathcal{K}(\underset{\sim}{b}) = \beta_m/\beta_1. \tag{129}$$

Thus, the more clustered the eigenvalues of $\underset{\sim}{b}$ are, the faster is the convergence of the conjugate gradient method.

The convergence characteristics as represented by the spectral condition number $\mathcal{K}(\underset{\sim}{b})$ can be improved by a so-called preconditioning. Generally, a preconditioning is achieved through a transformation of the unknowns

$$\underset{\sim}{y} = \underset{\sim}{E}^T\,\underset{\sim}{x}, \tag{130}$$

where the preconditioning matrix $\underset{\sim}{C}$,

$$\underset{\sim}{C} = \underset{\sim}{E}\,\underset{\sim}{E}^T, \tag{131}$$

is a positive definite matrix and $\underset{\sim}{E}^T$ designates the transpose of $\underset{\sim}{E}$. For simplicity we assume that $\underset{\sim}{b}$ is a positive definite real symmetric matrix. Since

$$\underset{\sim}{x}^T\,\underset{\sim}{b}\,\underset{\sim}{x} = \underset{\sim}{y}^T\,\tilde{\underset{\sim}{b}}\,\underset{\sim}{y}, \tag{132}$$

with

$$\tilde{\underset{\sim}{b}} = \underset{\sim}{E}^{-1}\,\underset{\sim}{b}\,\underset{\sim}{E}^{-T}, \tag{133}$$

the system (33) or (59) becomes

$$\tilde{\underset{\sim}{a}} + \tilde{\underset{\sim}{b}}\,\underset{\sim}{y} = \underset{\sim}{0}, \tag{134}$$

where

$$\tilde{\underset{\sim}{a}} = \underset{\sim}{E}^{-1}\,\underset{\sim}{a}. \tag{135}$$

Since, further, $\tilde{\underset{\sim}{b}}$ is similar (with similarity transformation $\underset{\sim}{E}^T$) to $\underset{\sim}{C}^{-1}\,\underset{\sim}{b}$, it has the same eigenvalues as $\underset{\sim}{C}^{-1}\,\underset{\sim}{b}$. Thus, choosing an appropriate preconditioning matrix $\underset{\sim}{C}$ the spectral condition number of $\tilde{\underset{\sim}{b}}$ can be lower than that of $\underset{\sim}{b}$ and a faster convergence can be achieved for the system (134).

Since the transformation (130) must be performed at least once when using the preconditioning, it is of practical importance that $\underset{\sim}{E}$ or $\underset{\sim}{C}$ is so chosen that the system (130) can be easily solved. It is thus usually assumed that $\underset{\sim}{E}$ is a sparse lower triangular matrix. Assuming that $\underset{\sim}{b}$ is diagonally dominant so that its eigenvalues lie not very far away from the diagonal entries of $\underset{\sim}{b}$ we can apply the following simple diagonal preconditioning matrix

$$C = \underset{\sim}{b}^{(d)} = \underset{\sim}{e}\,\underset{\sim}{e}^T, \tag{136}$$

where $\underset{\sim}{b}^{(d)}$ designates the diagonal part of $\underset{\sim}{b}$, so that

$$\underset{\sim}{e} = \left[(b_{II})^{\frac{1}{2}}\,\delta_{IJ}\right]. \tag{137}$$

Assuming that the off-diagonal matrix elements are small compared to 1, the majority of the eigenvalues of $\tilde{\underset{\sim}{b}}$ should be clustered around one.

We have programmed the conjugate gradient method using the algorithm and the code as given by Axelsson and Barker[42]. This program will be referred to as the CG (conjugate gradient) program. We have also written the preconditioned version using the simple preconditioning matrix (136). This version is referred to as the PCG (preconditioned CG) program.

For comparison we also programmed and used a very recently proposed approach by Purvis and Bartlett, called the reduced linear equation (RLE) method[43]. This method is very similar to that proposed earlier by Pople et al[44]. In fact, both these approaches can be shown to fall into the category of conjugate gradient algorithms[41,42]. Moreover, since the RLE space is built up as a span of the

iterants of the usual Jacobi algorithm,

$$\underset{\sim}{x}^{(n+1)} = -(\underset{\sim}{b}^{(d)})^{-1} \left[\underset{\sim}{a} + (\underset{\sim}{b} - \underset{\sim}{b}^{(d)})\underset{\sim}{x}^{(n)}\right], \tag{138}$$

the RLE procedure should perform similarly as the above mentioned PCG method. This indeed turns out to be the case and the work on a detailed elucidation of this relationship is in progress.

It should also be noted that the basic computational step in all the algorithms based on the steepest descent or conjugate gradient methods (with or without preconditioning) is that of forming the matrix product $\underset{\sim}{b}\,\underset{\sim}{c}^{(n)}$ of a sparse coefficient matrix $\underset{\sim}{b}$ with some vector $\underset{\sim}{c}^{(n)}$ characterizing the n-th iteration. This is precisely the same step which is required in most modern diagonalization algorithms for large sparse matrices. Moreover, this step is also amenable to a so called direct approach and may be handled in the same fashion as in the corresponding direct D-CI approach[45]. This enables one to exploit considerable computational experience. We would also like to note that a very close similarity of algorithms which can be employed to obtain the D-CI and DMBPT solutions has been both noted and exploited by Blomberg and Siegbahn[46]. Further, a suitable variant of Newton's method or the so-called self-correcting conjugate gradient algorithm[47] should provide solutions for the general large scale CPMET or ECPMET approaches, which are based on higher than quadratic functionals.

VIII. NUMERICAL EXAMPLE

In this section we will illustrate the theory of the previous sections by some results on the CO molecule.

An RHF solution for CO is obtained at the experimental internuclear distance of 2.132 a.u. using the ATMOL3 suite of programs[48]. The basis of 58 atomic orbitals used is the same as the one employed by Amos[49] and consists of a set of (11,6,3/5,4,2) contracted Gaussians on each nucleus. The SCF energy of -112.78098 a.u. obtained with this basis lies 0.01 a.u. above the HF limit of Ref. 50.

In the ensuing L-CPMET and CI calculations we keep the lowest two molecular orbitals (the core MO's) doubly occupied. Apart from this restriction all possible single and double excitations from the RHF reference state are included. Using the spatial symmetry group C_{2v} one obtains 7272 spin-projected states, which in our case are symmetrically orthogonalized spin bonded functions[51].

Before considering the 7272-dimensional L-CPMET and CI problems we first examine the usefulness of the relationship (109). To study how rapidly the sum in this equation approaches the exact L-CPMET energy we need all eigenvectors and eigenvalues of the D-CI matrix $\underset{\sim}{H}$. In order to obtain these, we have to truncate our basis of singly and doubly excited configurations. To that end we have examined separately the interaction of each of the 7271 configurations with the

RHF ground configuration and selected those which produced an energy lowering of more than 3.10^{-4} a.u. In this manner a 225-dimensional D-CI matrix $\underset{\sim}{H}$ was obtained and subsequently completely diagonalized.

In order to investigate the "convergence" of the sum in Eq. (109) we define

$$E_L^{(n)} := \left(\sum_{j=0}^{n} \kappa_j\right)^{-1}, \tag{139}$$

where κ_j has been defined in Eq. (111). The values of $E_L^{(n)}$ as a function of n are given in Table I together with the corresponding eigenvalues λ_n of $\underset{\sim}{H}$ (cf., Eq. (50)). Recall that λ_0 gives the D-CI correlation energy, Eq. (108). Note also that $E_L^{(0)}$ gives the renormalized Davidson correction[21-23] as shown by Eq. (117) and that $E_L^{(n)}$ with n = 224 gives the exact L-CPMET energy for this model. The different approximations to the L-CPMET energy E_L, which were introduced in Sec. VI.2, are summarized in Table II. Note that they satisfy the inequalities (122).

The main conclusion which can be drawn from this model calculation is that the Davidson corrections account for the major part of the difference between the D-CI and the L-CPMET energies. We observe that the approximants (138) following $E_L^{(0)}$ approach very slowly the L-CPMET energy E_L. This is due to an irregular distribution of the ground state components u_{0j} as a function of j, which causes κ_j not to decrease monotonically although λ_j and $E_L^{(n)}$ are monotonic. The Unsöld-type approximation is not far from the exact result in this case.

Since the $E_L^{(n)}$ values, Eq. (138), approached E_L very slowly in the model 225-dimensional case, we consider only the eigenvector with the lowest eigenvalue in the full 7272-dimensional case. The energy corrections which can be derived from this lowest eigenstate are given in the last column of Table II. The L-CPMET energy E_L is obtained in this case using the PCG program as outlined in Sec. VII and, alternatively, using the RLE method of Purvis and Bartlett[43]. Starting with the zero vector one needs 7 iterations for a 7 digit accuracy.

We find again that the Davidson $E_L^{(D)}$ and renormalized Davidson $E_L^{(0)}$ corrections for the unlinked clusters bring the D-CI result close to the L-CPMET energy E_L. The Unsöld-type approximation $E_L^{(U)}$ is again closest to E_L as predicted by the inequalities (122). However, the difference between $E_L^{(0)}$ and $E_L^{(U)}$ energies is small due to the fact that $\bar{\lambda}$ is large ($\bar{\lambda}$ = 9.20318 a.u.). It is also worthwhile to note that the relative importance of the correction for unlinked clusters as given by the L-CPMET approach is significantly larger for the 7272-dimensional model (namely 8.8% of the D-CI energy) than for the 225-dimensional case (5.1%).

ACKNOWLEDGEMENTS

The authors are very much obliged to Dr. I. Gustafsson of the Mathematics Institute of the Catholic University of Nijmegen for a very illuminating discussion on the conjugate gradient methods.

One of the authors (J.P.) wishes to express his gratitude to the Institute of Theoretical Chemistry of the Catholic University of Nijmegen for making his stay in the Netherlands possible and extend his most sincere thanks to his co-authors for their kind hospitality and a stimulating milieu which they provided to him during the tenure of his visiting professorship in Nijmegen. Continued support by NSERC is also acknowledged (J.P.).

REFERENCES

* On leave of absence from: Department of Applied Mathematics, Faculty of Mathematics and $(GWC)^2$, Faculty of Science, Waterloo Campus, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1.

1. F. Coester, Nucl. Phys. 7, 421 (1958);
   F. Coester and H. Kümmel, Nucl. Phys. 17, 477 (1960);
   H. Kümmel, in Lectures on the Many-Body Problem, edited by E.R. Caianiello (Academic, New York, 1962), p. 265.
2. S.T. Epstein, The Variational Method in Quantum Chemistry (Academic, New York, 1974), p. 19.
3. J. Čížek, J. Chem. Phys. 45, 4256 (1966); Adv. Chem. Phys. 14, 35 (1969).
4. J. Čížek and J. Paldus, Physica Scripta 21, 251 (1980) and references therein.
5. H . Kümmel, K.H. Lührmann and J.G. Zabolitzky, Phys. Reports 36, 1 (1978) and references therein.
6. A.C. Hurley, Electron Correlation in Small Molecules (Academic, New York, 1976).
7. W. Kutzelnigg, in Methods of Electronic Structure Theory, edited by H.F. Schaefer, III (Plenum Press, New York, 1977), p. 129.
8. F.E. Harris, in Electrons in Finite and Infinite Structures, edited by P. Phariseau and L. Scheire (Plenum Press, New York, 1977), p. 274.
9. J. Čížek and J. Paldus, Int. J. Quantum Chem. 5. 359 (1971).
10. J. Paldus and J. Čížek, in Energy, Structure and Reactivity, edited by D.W. Smith and W.B. McRae (Wiley, New York, 1973), p. 198.
11. R.J. Bartlett and G.D. Purvis, Physica Scripta 21, 255 (1980).
12. I. Lindgren and J. Morrison, Atomic Many Body Theory (Springer Verlag, New York, in press).
13. R. Ahlrichs, Comp. Phys. Commun. 17, 31 (1979).
14. J. Paldus, J. Chem. Phys. 67, 303 (1977).
15. B.G. Adams and J. Paldus, Phys. Rev. A20, 1 (1979).
16. R.A. Chiles and C.E. Dykstra, J. Chem. Phys. 74, 4544 (1981).

17. W. Meyer, J. Chem. Phys. 64, 2901 (1976);
    C.E. Dykstra, H.F. Schaefer and W. Meyer, J. Chem. Phys. 65, 2740 (1976).
18. R.J. Bartlett, Annu. Rev. Phys. Chem. 32, 359 (1981).
19. I. Lindgren, Int. J. Quantum Chem. S12, 33 (1978).
20. E.R. Davidson, The World of Quantum Chemistry, edited by R. Daudel and B. Pullman (Reidel, Dordrecht, 1974);
    S.R. Langhoff and E.R. Davidson, Int. J. Quantum Chem. 8, 61 (1974).
21. E.R. Davidson and D.W. Silver, Chem. Phys. Lett. 52, 403 (1977).
22. R.J. Bartlett and I. Shavitt, Int. J. Quantum Chem. S11, 165 (1977), S12, 543 (1978), (erratum). [Note that the summation symbol is missing in the second Eq. (20)].
23. P.E.M. Siegbahn, Chem. Phys. Lett. 55, 386 (1978).
24. W.L. Luken, Chem. Phys. Lett. 58, 421 (1978).
25. J. Paldus and J. Čížek, Adv. Quantum Chem. 9, 105 (1975).
26. J. Paldus, B.G. Adams and J. Čížek, Int. J. Quantum Chem. 11, 813 (1977).
27. H. Primas, in Modern Quantum Chemistry, edited by O. Sinanoğlu (Academic, New York, 1965), Vol. 2, p. 45.
28. J. Čížek, J. Paldus and L. Šroubková, Int. J. Quantum Chem. 3, 149 (1969).
29. O. Sinanoğlu, J. Chem. Phys. 36, 706 (1962);
    Adv. Chem. Phys. 6, 315 (1964), 14, 237 (1969).
30. J. Paldus, J. Čížek and I. Shavitt, Phys. Rev. A5, 50 (1972).
31. K. Jankowski and J. Paldus, Int. J. Quantum Chem. 18, 1243 (1980).
32. J. Paldus, K. Jankowski, B.G. Adamas and J. Čížek,
    in Proceedings of the Daresbury Study Weekend, edited by M.F. Guest and S. Wilson (Science Research Council, Daresbury, Warrington, 1980), p. 92.
33. B.G. Adams, K. Jankowski and J. Paldus, Chem. Phys. Lett. 67, 144 (1979).
34. B.G. Adams, K. Jankowski and J. Paldus, Phys. Rev. A (in press).
35. R.A. Chiles and C.E. Dykstra, Chem. Phys. Lett. 80, 69 (1981).
36. Ref. 2, p. 189.
37. J.A. Pople, R. Seeger and R. Krishnan, Int. J. Quantum Chem. S11, 149 (1977).
38. A. Meunier, B. Lévy and G. Berthier, Int. J. Quantum Chem. 10, 1061 (1976);
    W. Kutzelnigg, A. Meunier, B. Lévy and G. Berthier, Int. J. Quantum Chem. 12, 777 (1977); A. Meunier and B. Lévy, Int. J. Quantum Chem. 16, 955 (1979).
39. P.J.A. Ruttink, Chem. Phys. Lett. 79, 253 (1981).
40. R.J. Bartlett, I. Shavitt and G.D. Purvis, J. Chem. Phys. 71, 281 (1979).
41. M. Hestenes, Conjugate Direction Methods in Optimization (Springer, New York, 1980).
42. O. Axelsson and V.A. Barker, Finite Element Methods, (Lecture Notes, Mathematics Institute, University of Nijmegen, 1980).
43. G.D. Purvis and R.J. Bartlett, J. Chem. Phys. 75, 1284 (1981).
44. J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, Int. J. Quantum Chem. S13, 225 (1979).
45. B.O. Roos and P.E.M. Siegbahn, in Methods of Electronic Structure Theory, edited by H.F. Schaefer, III (Plenum, New York, 1977), p. 277 and references therein.
46. M.R.A. Blomberg and P.E.M. Siegbahn, Int. J. Quantum Chem. 14, 583 (1978).
47. A. Perry, Int. J. Comput. Math. 6, 327 (1978).

48. V.R. Saunders and M.F. Guest, The ATMOL 3 suite of programs, Daresbury Laboratory, Warrington, U.K.

49. R.D. Amos, Chem. Phys. Lett. 68, 536 (1979).

50. P.A. Christiansen and E.A. McCullough, Chem. Phys. Lett. 51, 468 (1977).

51. R.M. Berns and P.E.S. Wormer, Spin-bonded CI program, University of Nijmegen (1979); and R.M. Berns, Ph.D. Thesis, University of Nijmegen (1981).

TABLE I

Convergence of L-CPMET energies (in a.u.) in a 225-dimensional model space for CO. See Eqs. (50), (111) and (139) for definition of entries.

n	λ_n	κ_n	$E_L^{(n)}$
0	-0.131454	-0.725843	-0.137771
1	+0.714645	0.0	-0.137771
2	0.744336	0.000399	-0.137778
3	0.807597	0.000357	0.137785
4	0.897953	0.0	-0.137785
5	0.932770	0.004717	-0.137875
6	0.992352	0.000136	-0.137877
7	1.091888	0.000001	-0.137877
8	1.222070	0.0	-0.137877
50	2.141860	0.000029	-0.138064
100	3.119776	0.000017	-0.138143
150	4.052877	0.000143	-0.138170
200	5.417043	0.0	-0.138189
224	18.100258	0.000001	-0.138202

TABLE II

Different approximations to the L-CPMET energy (in a.u.) of CO in 225 and 7272-dimensional model spaces. See Sec. VII for definition of energies.

Energy	Dim=225	Dim=7272
E_D	-.13145	-.25329
$E_L^{(D)}$	-.13748	-.27127
$E_L^{(0)}$	-.13777	-.27264
$E_L^{(U)}$	-.13802	-.27320
E_L	-.13820	-.27561

(Caption to Fig. 1)

Relationship of D-CI and L-CPMET variational parameters in a simple two-dimensional model case. The relationship given by Eq. (82), which connects extremal values of variational parameters, is shown below the x_0 axis, since $x_e^L = -a/b < 0$ assuming that $a,b > 0$. The points marked by full circles indicate the absolute minimum of the functional $G_1(\underset{\sim}{X})$, Eq. (67), on the unit circle and on its tangent at $x_0 = 1$, representing D-CI and L-CPMET solutions, respectively.

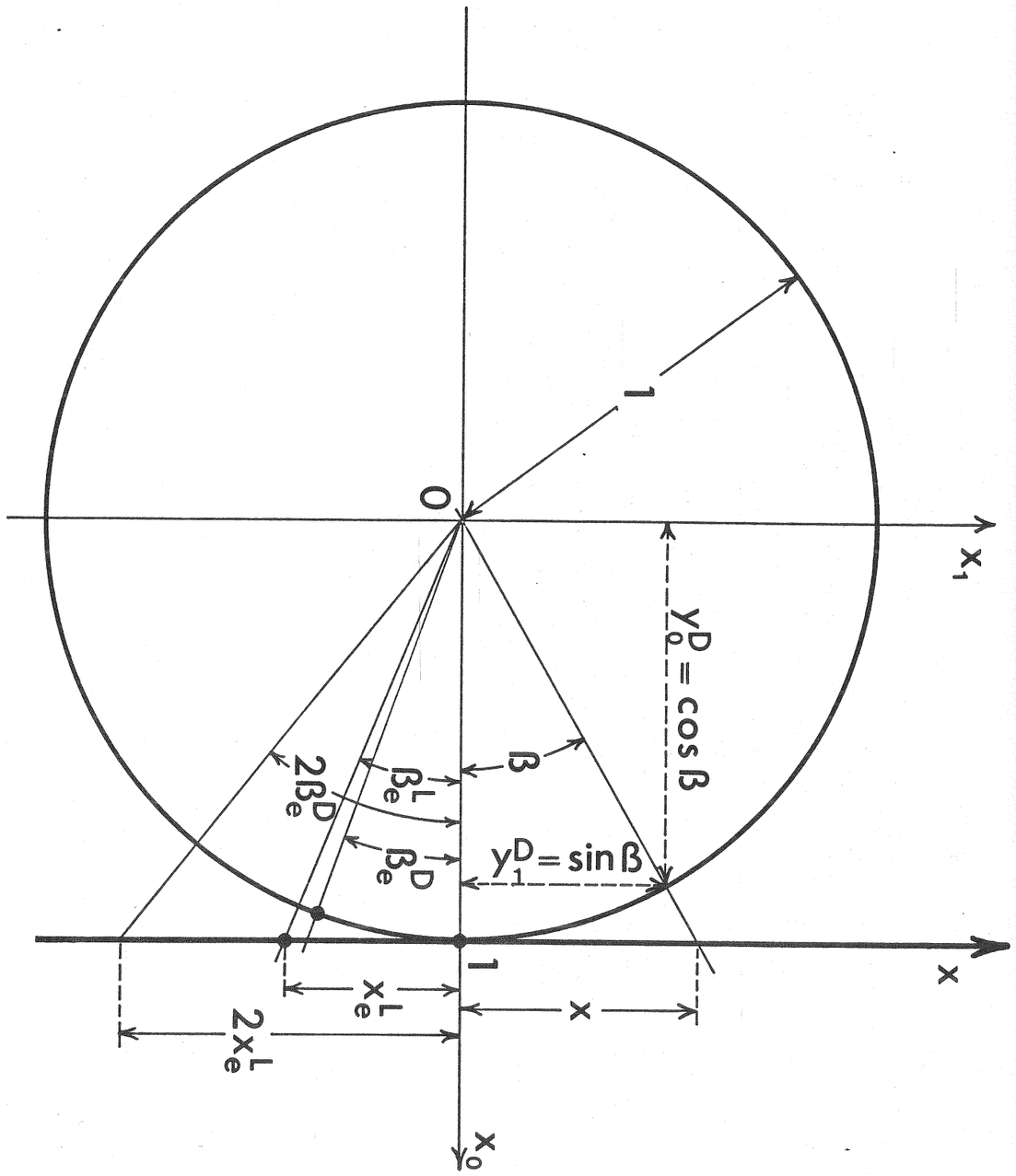


Fig. 1

Appendix 3
Summary of definitions and formulas for P.T.

$\hat{K}|\Psi_i\rangle = k_i|\Psi_i\rangle$
$\hat{K}_o|\Phi_i\rangle = \varkappa_i|\Phi_i\rangle$
$\hat{K} = \hat{K}_o + \hat{W}$

then

$$k_i = \varkappa_i + \sum_{M=0}^{\infty} \langle\Phi_i|\hat{W}[\hat{P}_i(\hat{W}+\varkappa_i-k_i)]^M|\Phi_i\rangle$$

where

$$\hat{P}_i = \sum_{j(\neq i)} \frac{|\Phi_j\rangle\langle\Phi_j|}{\varkappa_i - \varkappa_j} \quad , \quad \hat{P}_i|\Phi_i\rangle = 0 .$$

Let

$$k_i = \sum_{M=0}^{\infty} k_i^{(n)}$$

where $k_i^{(n)}$ contains the terms with n operators $\hat{W}$ (n^{th} order term) then for a few lowest orders

$$k_i^{(0)} = \varkappa_i = \langle\Phi_i|K_o|\Phi_i\rangle$$
$$k_i^{(1)} = \langle\Phi_i|\hat{W}|\Phi_i\rangle$$
$$k_i^{(2)} = \langle\Phi_i|\hat{W}\hat{P}_i\hat{W}|\Phi_i\rangle$$
$$k_i^{(3)} = \langle\Phi_i|\hat{W}\hat{P}_i\hat{W}\hat{P}_i\hat{W}|\Phi_i\rangle - k_i^{(1)}\langle\Phi_i|\hat{W}\hat{P}_i^2\hat{W}|\Phi_i\rangle$$
etc. (using bracketing technique).

Consider next:
Ground state (nondegenerate) of some Hamiltonian $\hat{H}$:

$\hat{H} = \hat{Z} + \hat{V}$
$\hat{H}_o = \hat{Z} + \hat{U}$

$$\hat{Z} = \sum_i \hat{z}(i) \quad , \quad \hat{V} = \sum_{i<j} \hat{v}(i,j) \quad , \quad \hat{U} = \sum_i u(i)$$

such that we can solve: $(\hat{z} + \hat{u})|A\rangle = \omega_A|A\rangle$.

Let

$$\hat{H}|\Psi_i\rangle = E_i|\Psi_i\rangle$$
$$H_o|\Phi_i\rangle = \varepsilon_i|\Phi_i\rangle \qquad \varepsilon_i = \sum_{A\in K_i} \omega_A$$

(sum over 1-particle states occupied in $|\Phi_i\rangle$.)

Consider now the ground state and choose all operators in the normal product form (this will simplify the P.T. expansion so that the first order will identically vanish):

$$\hat{K} = \hat{H} - \langle\Phi_0|\hat{H}|\Phi_0\rangle \qquad K|\Psi_i\rangle = k_i|\Psi_i\rangle$$
$$K_0 = \hat{H}_0 - \langle\Phi_0|\hat{H}_0|\Phi_0\rangle \qquad \hat{K}_0|\Phi_i\rangle = \varkappa_i|\Phi_i\rangle$$

$$k_i = E_i - \langle\Phi_0|\hat{H}|\Phi_0\rangle$$
$$\varkappa_i = \varepsilon_i - \varepsilon_0 \;, \quad \varepsilon_0 = \langle\Phi_0|H_0|\Phi_0\rangle.$$

Then also the perturbation is in the N-product form:

$$\hat{W} = V - U - \langle\Phi_0|V-U|\Phi_0\rangle = \hat{W}^{(1)} + \hat{W}^{(2)} = \hat{Q}_N + \hat{V}_N$$

$$\hat{Q}_N = \hat{G}_N - \hat{U}_N \qquad \langle A|g|B\rangle = \sum_{C_1} \langle AC_1|\hat{v}|BC_1\rangle_A \;.$$

Designate the solutions of $\hat{K}_0$ (i.e., eigenstates of $\hat{K}_0$) as:

$$|\Phi_K^{(k)}\rangle \equiv \left|\begin{matrix} A^1 & A^2 & \cdots & A^k \\ A_1 & A_2 & \cdots & A_k \end{matrix}\right\rangle \;, \quad |\Phi^{(0)}\rangle \equiv |\Phi_0\rangle$$

using the h-p formalism and indicating explicitly in the ket symbol which hole states $(A_1, \ldots, A_k)$ are replaced by which particle (unoccupied) states $(A^1, \ldots, A^k)$. Then

$$\hat{K}_0 |\Phi_K^{(k)}\rangle = \varkappa_K^{(k)} |\Phi_K^{(k)}\rangle \qquad (0 \le k \le N)$$

and

$$\varkappa_K^{(k)} = \varepsilon_K^{(k)} - \varepsilon_0 = \sum_{i=1}^{k} (\omega_{A^i} - \omega_{A_i}) \;, \quad \varkappa_0 = 0 \;.$$

Thus, $\hat{P}_0$ may be expressed as follows

$$\hat{P}_0 = \sum_{k=1}^{N} \hat{P}_0^{(k)} \;, \qquad \hat{P}_0^{(k)} = \sum_K \frac{|\Phi_K^{(k)}\rangle\langle\Phi_K^{(k)}|}{\varkappa_0 - \varkappa_K^{(k)}}$$

the summation extending over ordered configurations K. Equivalently we can write:

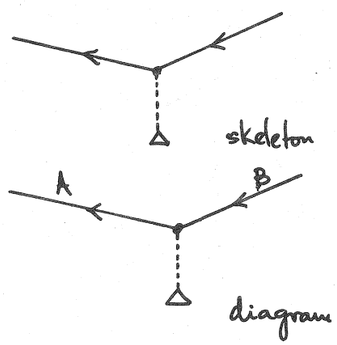
$$P_0^{(k)} = \sum_{\substack{A_1 < A_2 < \cdots < A_k \\ A^1 < A^2 < \cdots < A^k}} \frac{\left|\begin{matrix} A^1 & A^2 & \cdots & A^k \\ A_1 & A_2 & \cdots & A_k \end{matrix}\right\rangle\left\langle\begin{matrix} A^1 & A^2 & \cdots & A^k \\ A_1 & A_2 & \cdots & A_k \end{matrix}\right|}{-\varkappa_K^{(k)}}$$

$$= \frac{1}{(k!)^2} \sum_{\substack{A_1, A_2, \ldots, A_k \\ A^1, A^2, \ldots, A^k}} \prod_{i=1}^{k} (Y_{A^i}^{+} Y_{A_i}^{+}) |\Phi_0\rangle\langle\Phi_0| \prod_{j=1}^{k} (Y_{A_j} Y_{A^j}) \left[-\varkappa_K^{(k)}\right]^{-1} .$$

Diagram assignment

1-particle operator part of $\hat{W}$

(i) $$\hat{Q}_N = \sum_{A,B} \langle A|\hat{q}|B\rangle\, N[X_A^{+} X_B]$$



$\chi = (A;B)$

$d_Q(\chi) = \langle A|\hat{q}|B\rangle$

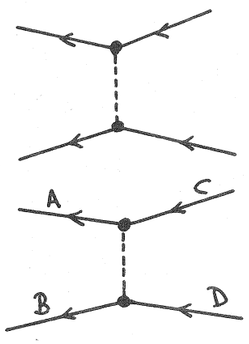
$\hat{D}_Q^N(\chi) = d_Q(\chi)\, N[\hat{X}_A^{+} \hat{X}_B]$

$w_Q = 1$

thus $$\hat{Q}_N = w_Q \sum_{\chi} \hat{D}_Q^N(\chi)$$

(ii) two-particle part of $\hat{W}$

$$\hat{V}_N = \tfrac{1}{2} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle\, N[\hat{X}_A^{+} \hat{X}_B^{+} \hat{X}_D \hat{X}_C]$$



$\chi = (A,B;CD)$

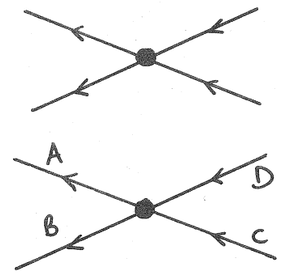
$d_V(\chi) = \langle AB|\hat{v}|CD\rangle$

$\hat{D}_V^N(\chi) = d_V(\chi)\, N[\hat{X}_A^{+} \hat{X}_B^{+} \hat{X}_D \hat{X}_C]$

$w_V = \tfrac{1}{2}$

$$\hat{V}_N = w_V \sum_{\chi} \hat{D}_V^N(\chi)$$

OR: $$V_N = \tfrac{1}{4} \sum_{ABCD} \langle AB|\hat{v}|CD\rangle_A\, N[\hat{X}_A^{+} \hat{X}_B^{+} \hat{X}_D \hat{X}_C]$$



$\tilde{\chi} = (A,B;C,D)$ or $(A,B;D,C)$

$d_{\tilde{V}}(\tilde{\chi}) = \langle AB|\hat{v}|CD\rangle_A$ or $\langle AB|\hat{v}|DC\rangle_A$

$\hat{D}_{\tilde{V}}(\tilde{\chi}) = \langle AB|\hat{v}|CD\rangle_A\, N[\hat{X}_A^{+} \hat{X}_B^{+} \hat{X}_D \hat{X}_C]$

$= \langle AB|\hat{v}|DC\rangle_A\, N[\hat{X}_A^{+} \hat{X}_B^{+} \hat{X}_C \hat{X}_D]$

$w_{\tilde{V}} = \tfrac{1}{4}$

$$\hat{V}_N = w_{\tilde{V}} \sum_{\tilde{\chi}} \hat{D}_{\tilde{V}}(\tilde{\chi})$$

(iii) <u>single configuration</u>

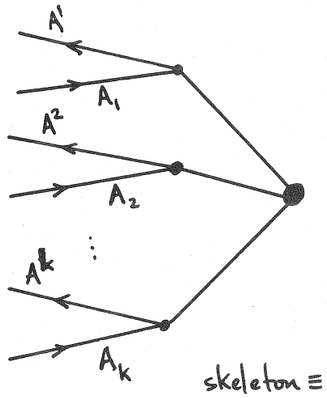
$$|\Phi_K^{(k)}\rangle \equiv \left|\begin{matrix} A^1 A^2 \cdots A^k \\ A_1 A_2 \cdots A_k \end{matrix}\right\rangle = \prod_{i=1}^{k} (Y_{A^i}^+ Y_{A_i}^+)\,|\Phi_0\rangle = \hat{C}_k^{(k)}\,|\Phi_0\rangle$$

$$K = (A^1 A^2 \cdots A^k\,;\, A_1 A_2 \cdots A_k)$$

$$\hat{C}_K^{(k)} = (\hat{Y}_{A^1}^+ \hat{Y}_{A_1}^+)(\hat{Y}_{A^2}^+ \hat{Y}_{A_2}^+) \cdots (\hat{Y}_{A^k}^+ \hat{Y}_{A_k}^+)$$

$$= \hat{Y}_{A^1}^+ \hat{Y}_{A^2}^+ \cdots \hat{Y}_{A^k}^+ \hat{Y}_{A_k}^+ \cdots \hat{Y}_{A_2}^+ \hat{Y}_{A_1}^+$$

$$= \hat{X}_{A^1}^+ \hat{X}_{A^2}^+ \cdots \hat{X}_{A^k}^+ \hat{X}_{A_k} \cdots \hat{X}_{A_2} \hat{X}_{A_1}$$



skeleton ≡ diagram

(fixed labels; no summation over A^i, A_i)

$$w_T^{(k)} = 1$$

$$d_T^{(k)}(K) = 1$$

$$\hat{D}_T^{(k)}(K) = d_T^{(k)}(K) \prod_{i=1}^{k} (\hat{Y}_{A^i}^+ \hat{Y}_{A_i}^+)$$

$$\hat{C}_K^{(k)} = w_T^{(k)}\, \hat{D}_T^{(k)}(K)$$

(iv) <u>General wavefunction in the occup. number rep^n.</u>

Let's write

$$|\Psi\rangle = \hat{C}\,|\Phi_0\rangle = \sum_k \sum_K c_K^{(k)}\, \hat{C}_K^{(k)}\,|\Phi_0\rangle$$

$$= \sum_{k=0}^{N} \sum_{\substack{A^1 < \cdots < A^k \\ A_1 < \cdots < A_k}} C\begin{bmatrix} A^1 \cdots A^k \\ A_1 \cdots A_k \end{bmatrix} \left|\begin{matrix} A^1 \cdots A^k \\ A_1 \cdots A_k \end{matrix}\right\rangle$$

$$= \sum_{k=0}^{N} \frac{1}{(k!)^2} \sum_{\substack{A^1, \cdots, A^k \\ A_1, \cdots, A_k}} C\begin{Bmatrix} A^1 \cdots A^k \\ A_1 \cdots A_k \end{Bmatrix} \left|\begin{matrix} A^1 \cdots A^k \\ A_1 \cdots A_k \end{matrix}\right\rangle$$

where $C\{\cdots\}$ designate antisymmetrized coefs, i.e.

$$C\begin{Bmatrix} A^1 A^2 \cdots A^k \\ A_1 A_2 \cdots A_k \end{Bmatrix} = \sum_{P \in \mathcal{S}_k} (-1)^p\, C\begin{bmatrix} A^1 A^2 \cdots A^k \\ A_{p_1} A_{p_2} \cdots A_{p_k} \end{bmatrix},$$

$$P = \begin{pmatrix} 1 & 2 & \cdots & k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix} \in \mathcal{S}_k \,.$$

Equivalently, we can use operator notation

$$\hat{C} = \sum_{k=0}^{N} \hat{C}_k \quad ; \qquad \hat{C}_0 = \hat{I}$$

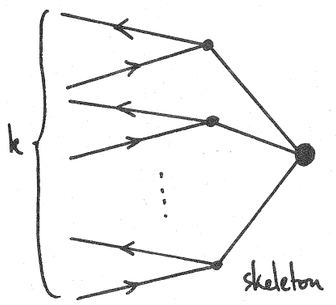
$$\hat{C}_k = \frac{1}{k!} \sum_{\substack{A_1 \cdots A_k \\ A^1 \cdots A^k}} \langle A^1 \cdots A^k\,|\,\hat{c}^{(k)}\,|\,A_1 \cdots A_k\rangle\, \hat{X}_{A^1}^+ \cdots \hat{X}_{A^k}^+ \hat{X}_{A_k} \cdots \hat{X}_{A_1}$$

$$= \frac{1}{(k!)^2} \sum_{\substack{A^1 \cdots A^k \\ A_1 \cdots A_k}} \langle A^1 \cdots A^k\,|\,\hat{c}^{(k)}\,|\,A_1 \cdots A_k\rangle_A\, \hat{X}_{A^1}^+ \cdots \hat{X}_{A^k}^+ \hat{X}_{A_k} \cdots \hat{X}_{A_1}$$

where

$$\langle A^1 \cdots A^k\,|\,\hat{c}^{(k)}\,|\,A_1 \cdots A_k\rangle_A = \sum_{P \in \mathcal{S}_k}{}' (-1)^p \langle A^1 \cdots A^k\,|\,\hat{c}^{(k)}\,|\,A_{p_1} \cdots A_{p_k}\rangle$$

Now, we can use either of the flwg representations:



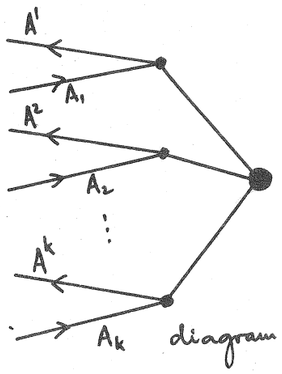
$$K \equiv \chi = (A^1 A^2 \cdots A^k\,;\, A_1 A_2 \cdots A_k)$$

$$d_C^{(k)}(\chi) = \langle A^1 \cdots A^k\,|\,\hat{c}^{(k)}\,|\,A_1 \cdots A_k\rangle$$

$$w_C^{(k)} = \frac{1}{k!}$$

$$\hat{D}_C^{(k)}(\chi) = d_C^{(k)}(\chi)\, N\Big[\prod_{i=1}^{k} (Y_{A^i}^+ Y_{A_i}^+)\Big]$$

$$= d_C^{(k)}(\chi)\, \hat{X}_{A^1}^+ \cdots \hat{X}_{A^k}^+ \hat{X}_{A_k} \cdots \hat{X}_{A_1}$$

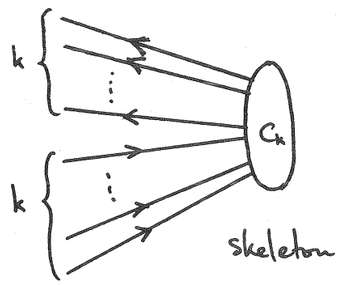


diagram

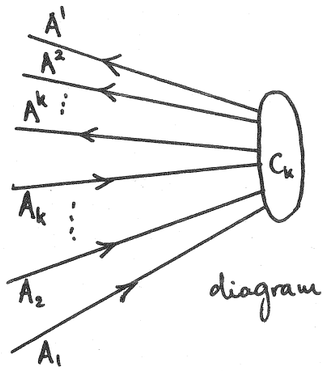
so that

$$\hat{C}_k = w_C^{(k)} \sum_{\chi} \hat{D}_C^{(k)}(\chi)$$

or the representation based on antisymmetric matrix elements using degenerate (Hugenholtz-type) diagrams:



skeleton



diagram

$$\tilde{\chi} = (A^1 A^2 \cdots A^k \,;\, A_{p_1} A_{p_2} \cdots A_{p_k})$$

$$d^{(k)}_{\tilde{c}}(\tilde{\chi}) = \langle A^1 \cdots A^k \,|\, \hat{c}^{(k)} \,|\, A_{p_1} A_{p_2} \cdots A_{p_k} \rangle$$

$$w^{(k)}_{\tilde{c}} = \frac{1}{(k!)^2}$$

$$\hat{D}^{(k)}_{\tilde{c}}(\tilde{\chi}) = d^{(k)}_{\tilde{c}}(\tilde{\chi})\, \hat{X}^+_{A^1} \cdots \hat{X}^+_{A^k} \hat{X}_{A_{p_k}} \cdots \hat{X}_{A_{p_1}}$$

$$\hat{C}_k = w^{(k)}_{\tilde{c}} \sum_{\tilde{\chi}} \hat{D}^{(k)}_{\tilde{c}}(\tilde{\chi})$$

<u>Remark:</u> Tilde on χ means that no ordering of index sets is implied; thus while

$$\chi = (A^1 A^2 \cdots A^k \,;\, A_1 A_2 \cdots A_k) = (A^{p_1} A^{p_2} \cdots A^{p_k} \,;\, A_{p_1} A_{p_2} \cdots A_{p_k})$$

we have

$$\tilde{\chi} = (A^1 A^2 \cdots A^k \,;\, A_1 A_2 \cdots A_k) = (A^1 A^2 \cdots A^k \,;\, A_{p_1} A_{p_2} \cdots A_{p_k})$$
$$= (A^{p_1} A^{p_2} \cdots A^{p_k} \,;\, A_1 A_2 \cdots A_k) \text{ etc.}$$

However, in defining $\hat{D}$ operators, same $\tilde{\chi}$ must be considered for the operator part as for the scalar factor d.

(v) Consider, finally, the projector $\hat{P}_0$,

$$\hat{P}_0 = \sum_{k=1}^{N} \hat{P}_0^{(k)}$$

$$P_0^{(k)} = \frac{1}{(k!)^2} \sum_{\substack{A^1 \cdots A^k \\ A_1 \cdots A_k}} N[(\hat{X}^+_{A^1} \hat{X}_{A_1}) \cdots (\hat{X}^+_{A^k} \hat{X}_{A_k})] \,|\Phi_0\rangle\langle\Phi_0| \times$$

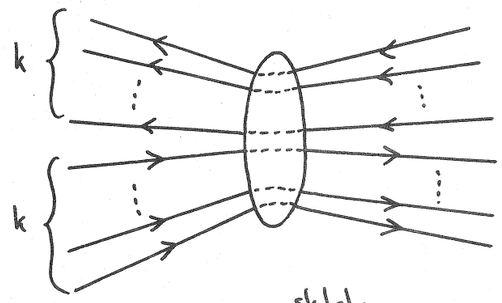
$$\times N[(\hat{X}^+_{A_k} \hat{X}_{A^k}) \cdots (\hat{X}^+_{A_1} \hat{X}_{A^1})] \cdot (-\varkappa_K^{(k)})^{-1}$$

$$= \frac{1}{(k!)^2} \sum_{\substack{A^1 \cdots A^k \\ A_1 \cdots A_k}} N[\hat{X}^+_{A^1} \hat{X}^+_{A^2} \cdots \hat{X}^+_{A^k} \hat{X}_{A_k} \cdots \hat{X}_{A_1}] \,|\Phi_0\rangle\langle\Phi_0|\, N[X^+_{A_k} \cdots X^+_{A_1} X_{A^1} \cdots X_{A^k}] \cdot (-\varkappa_K^{(k)})^{-1}$$

Notice that in $P_0^{(k)}$ there is always the same unperturbed state in each component, i.e. $|\Phi_K^{(k)}\rangle\langle\Phi_K^{(k)}|$,

or $\left|\begin{matrix} A^1 A^2 \cdots A^k \\ A_1 A_2 \cdots A_k \end{matrix}\right\rangle\left\langle\begin{matrix} A^1 A^2 \cdots A^k \\ A_1 A_2 \cdots A_k \end{matrix}\right|$.

In order that the phase (sign) of each factor is preserved, the same order of hole & particle states must be used in the corresponding ket & bra vectors. Thus, we can use C and C^+ diagrams, joined together, and indicate the corresponding states by dotted lines which must not intersect (or we can allow intersection but associate minus sign with odd number of intersections). Thus we define



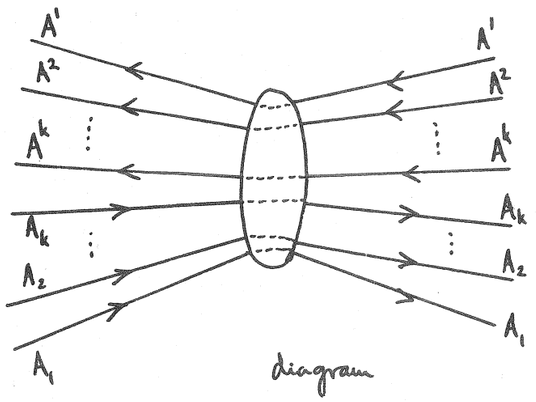
skeleton

$$\chi = (A^1 A^2 \cdots A^k \,;\, A_1 A_2 \cdots A_k)$$

or permuted on either side

$$d_P^{(k)}(\chi) = \left[\sum_{i=1}^{k} (\omega_{A_i} - \omega_{A^i})\right]^{-1}$$

$$w_P^{(k)} = \frac{1}{(k!)^2}$$



diagram

$$\hat{D}_P^{(k)}(\chi) = d_P^{(k)}(\chi) \times$$

$$\times N[X^+_{A^1} \cdots X^+_{A^k} X_{A_k} \cdots X_{A_1}] \,|\Phi_0\rangle\langle\Phi_0| \times$$

$$\times N[X^+_{A_1} \cdots X^+_{A_k} X_{A^k} \cdots X_{A^1}]$$

(notice again the convention: outgoing lines correspond to creators; ingoing to annihilators)

and $\hat{P}_0^{(k)} = w_P^{(k)} \sum_{\chi} \hat{D}_P^{(k)}(\chi)$.

It is now easy to see, that these definitions can be replaced by a simple "denominator rule", since the operators $\hat{P}_o^{(k)}$ always appear surrounded by the perturbation $\hat{W}$ operators; moreover, in $\hat{P}_o$ all configurations except the $|\Phi_o\rangle$ occur: thus, we can draw only the vertices associated with the components of the W operator, and eliminate disconnected diagrams having no oriented lines passing between any pair of neighboring vertices and simply to assign to each pair of neighboring W vertices a denominator

$$1/\left[\sum_{i=1}^{k}(\omega_{A_i}-\omega_{A^i})\right]$$

where the sum extends over the labels of all hole & particle lines passing between the given pair of neighboring vertices (i.e., all oriented lines cut by a hypothetical line separating given two neighboring vertices).

Remark: It is easy to prove that in each case the number of particle lines must equal the number of hole lines between any two vertices.

Summary of rules

(i) Draw all possible connected (linked) vacuum skeletons having n_1 Q-vertices and n_2 V-vertices, $n_1, n_2 = 0, 1, \ldots, k$ and $n_1 + n_2 = k$, and determine their weights w_R.

(ii) Label the particle and hole lines of these skeletons with particle and hole spin orbital labels, respectively, thus forming all possible kth order diagrams.

(iii) Write down the contributions from each diagram following these rules:

(a) Assign matrix elements $\langle A|\hat{q}|C\rangle$ and $\langle AB|\hat{v}|CD\rangle$ to each Q- and V-vertex, respectively, where A, B are labels of the outgoing lines and C,D are labels of the incoming lines, A and C (B and D) being associated with the same single vertex.

(b) With each pair of neighboring vertices associate a denominator factor $[\sum_i (\omega_{A_i} - \omega_{A_i'})]^{-1}$, taking the orbital energies ω_A associated with the hole (particle) lines with positive (negative) sign.

(c) Determine the numerical factor given by $w_R(-1)^{l+h}$ [cf. Eq. (76)].

(d) Multiply together all the above factors and sum over all hole and particle states labeling internal lines.