Phys. 40, 595 (1964).
[56] P.-O. Löwdin, Advan. Chem. Phys. 2, 207 (1959).
[57] R. McWeeny, Rev. Mod. Phys. 32, 335 (1960).
[58] D. ter Haar, Rept. Progr. Phys. 24, 304 (1961).
[59] K. F. Freed, J. Chem. Phys. 47, 3907 (1967).
[60] J. Cizek, J. Chem. Phys. 45, 4256 (1966).
[61] R. E. Stanton, J. Chem. Phys. 42, 2353 (1965).
[62] L. Szasz and J. Byrne, Phys. Rev. 158, 34 (1967).
[63] R. Yaris and J. S. Musher, J. Chem. Phys. 41, 1701 (1964).
[64] B. D. Day, Rev. Mod. Phys. 39, 719 (1967).

Generalized Perturbation Theory and its Application to the Problem of the Description of Electron Correlation in Atoms and Molecules*

Karl F. Freed†
Department of Chemistry, Harvard University, Cambridge, Massachusetts

(Received 21 August 1967)

For the case in which the zeroth-order wave function is a single Slater determinant of arbitrary orthonormal spin-orbitals, the exact wave function can be expressed in terms of a cluster-type expansion. If perturbation theory is applied in the standard (and systematic) fashion to the Schrödinger equation for the cluster functions, a generalized form of Rayleigh-Schrödinger perturbation theory (GPT) is obtained. Although GPT is quite flexible, the form presented in this work is chosen so that the terms initially calculated in GPT correspond as closely as possible to the starting points in the theories of Sinanoğlu, Kelly, Conkie, Nesbet, and others.

The first-order corrections to the wave function in GPT are written in terms of one- and two-electron functions which respectively describe the perturbative corrections for the motion of one and two electrons in the field of the remaining electrons which are in their ground-state orbitals (i.e., perturbative Bethe-Goldstone-like equations). This is a more physical starting point than the first-order wave function of ordinary Rayleigh-Schrödinger perturbation theory (OPT), and GPT should be much more rapidly convergent than OPT.

By analogy with OPT, GPT also has a $2n+1$ rule, a variational principle for the nth-order energy and the 2^nth-order wave function, and interchange theorems of double-perturbation theory. As in OPT, the equations of GPT can be solved by the use of correlated functions; however, in GPT, H_0 is not required, and the zeroth-order spin-orbitals need not all be eigenfunctions of the same one-electron Hamiltonian.

I. INTRODUCTION

The theories for the description of electron correlation in atoms and molecules of Sinanoğlu, Szász, Nesbet, and others begin with approximate Schrödinger equations (in different forms) describing the stationary states of a pair of electrons which are moving in the field of the remaining electrons which are in their ground-state orbitals.[1-10] Although the results presented in this work are valid for the case in which the zeroth-order wave function is a single Slater determinant of arbitrary orthonormal spin-orbitals, the discussion in the first three sections is limited to the simplest case of a closed-shell Hartree-Fock (HF) zeroth-order wave function.

When approaching the problem of electron correlation by ordinary Rayleigh-Schrödinger perturbation theory, the first-order corrections to the wave function are written in terms of first-order pairs. If $\{l_i \mid i=1, \cdots, N\}$ denote the occupied spin-orbitals in the zeroth-order HF wave function,[11]

$$Q_1 = \hat{1}(1) - \sum\nolimits_{l\,\text{occpd}} |(1)\rangle\langle l(1)|, \tag{1.1a}$$

$$Q_{12} = Q_1 Q_2 = Q_2 Q_1, \tag{1.1b}$$

$$|[ll']_{12}\rangle = 2^{-1/2}(1-P_{12})\,|l(1)l'(2)\rangle, \tag{1.1c}$$

where P_{12} is the operator which permutes the space and spin variables of particles one and two, then the first-order pairs $u_{ll'}{}^{(1)}(12)$ are the solutions to the uncoupled *linear* inhomogeneous integrodifferential equations,

$$Q_{12}[\epsilon_{ll'} - H_0(12)]u_{ll'}{}^{(1)}(12) \\ - Q_{12}(1/r_{12})\,|[ll']_{12}\rangle = 0, \tag{1.2}$$

where $H_0(i)$ is the HF Hamiltonian for particle i, $H_0(12) = H_0(1) + H_0(2)$, $\epsilon_{ll'} = \epsilon_l + \epsilon_{l'}$ is a sum of HF orbital energies via

$$H_0(i) l_k(i) = \epsilon_k l_k(i), \tag{1.3a}$$

and Q_{12} is a projection operator which makes any pair function strongly orthogonal to the HF orbitals which are occupied in the zeroth-order wave function, i.e.,

$$\int d(1) l_k(1) Q_{12} f(12) = 0 \tag{1.3b}$$

for all $k \leqslant N$ and $f(12)$ arbitrary.

From the homogeneous part of Eq. (1.2), we see that the pairs see the full HF Hamiltonian, and the inhomogeneity (or "forcing" term) is the product of the perturbation to the HF Hamiltonian and zeroth-order occupied spin-orbitals. This Eq. is, however, a rather poor description of the physical situation, since when electrons originally in spin-orbitals l and l' are excited, they do not see a field due to electrons in l and l'. Instead, they interact via the full electron repulsion $1/r_{12}$ (in atomic units). Furthermore the zeroth-order energy of the pair of electrons, $\epsilon_{ll'}$, is inadequate, since it counts the repulsion between the electrons twice, and hence should be replaced by $\epsilon_{ll'}-\langle ll' | (1-P_{12})/r_{12} | ll' \rangle$. Therefore if we define the two-body potential

$$V_{ll'}(12)=\frac{1}{r_{12}}-\sum_{i=1,2}[K_{ll}(i)+K_{l'l'}(i)] + \left\langle ll' \left| \frac{1-P_{12}}{r_{12}} \right| ll' \right\rangle , \tag{1.4a}$$

where $K_{ll'}(i)=\langle l(\mu) | (1-P_{i\mu})/r_{i\mu} | l'(\mu) \rangle$ (1.4b)

is the difference between Coulomb and exchange operators, the pair equations

$$Q_{12}[\epsilon_{ll'}-H_0(12)-V_{ll'}(12)Q_{12}]u_{ll'}(12) - Q_{12}(1/r_{12}) | [ll']_{12} \rangle = 0 \tag{1.5}$$

would be a much more physical starting point for a perturbative approach to the problem of the description of electron correlation in atoms and molecules. The pairs $u_{ll'}(12)$ in (1.5) are the Bethe-Goldstone (BG) correlation-correction pairs [the full BG pairs are $|[ll']_{12}\rangle + u_{ll'}(12)$] which were obtained by Sinanoğlu by varying part of the energy expression from a simple cluster-type wave function[2,3] and were also shown in paper I[12] to result from a summation of a certain class of contributions to the wave function to all orders of perturbation theory. These BG correlation functions (which are different from Nesbet's BG pairs as discussed below) are the perturbative pair corrections for a pair of electrons moving in the field of the remaining electrons which are in their ground state orbitals. As shown in paper I,[12] a knowledge of these BG pairs would enable the evaluation of the vast majority of diagrams calculated by Kelly (to all orders of OPT) for the Be atoms.[13,14] In Kelly's Be calculation, he obtained better than 90% of the correlation energy. Furthermore, when these BG correlation functions are small in the sense that $\langle u_{ll'} | u_{ll'} \rangle \ll 1$, these BG pairs are essentially equivalent to Sinanoğlu's "exact pairs" which describe the *stationary states* of a pair in the field of the remaining electrons.[2,12,15]

These BG pairs can be taken to be the first approximations in a generalized form of Rayleigh-Schrödinger perturbation theory (GPT). (GPT is sufficiently flexible so that Sinanoğlu's exact pairs could also be chosen to determine the first-order corrections to the wave functions.) As discussed in paper I, these BG pairs can be solved by the use of a correlated basis set and (1.5) may be solved variationally. GPT may also be written in the form of what is commonly called a variation-perturbation theory.[11] It also has a $2n+1$ rule that knowledge of the wave function through nth order of GPT enables a calculation of the energy through $(2n+1)$st order of GPT to be made, and double GPT also has interchange theorems.[11] Furthermore, in the general case (discussed in Sec. IV) of a zeroth-order wave function which is a single Slater determinant, the spin-orbitals need not be eigenfunctions of the same one-electron Hamiltonian, and we do not require a knowledge of the Hamiltonian of which the zeroth-order wave function is an eigenfunction.[16]

As mentioned above, in paper I the BG pairs were obtained by summing certain terms in OPT to all orders, and this summation procedure was extended to provide approximate equations for the description of single and triple excitations in the wave function and correlation energy.[12] However, there is often an uneasiness associated with the summation of certain terms to all orders of OPT since it is hoped that OPT converges (at least asymptotically), even though there are a large number of terms in higher orders which are not calculated, partly because of the cancellation between different terms in the same order of perturbation theory. Thus, it may be argued, summation of only some terms to all orders may not enable this cancellation to occur, and the final results obtained from such a theory may be completely erroneous.

Physically, we expect that the perturbative BG pair functions should provide a good description of electron correlation in atoms and molecules, and therefore any systematic form of perturbation theory in which these pairs determine the first-order corrections should be expected to be more rapidly convergent than OPT. GPT can therefore be used to justify *certain* summations in OPT, even though a large number of contributions have been omitted. The results of diagram summation to all orders of OPT in paper I can be used to express a given order of GPT (for the correlation energy and/or wave function) in terms of OPT.

The starting point in the derivation of GPT is the Schrödinger equation for the wave function when it is written as a cluster-type wave function. These equations also form the starting point for a generalized form of configuration interaction (CI) (in which a correlated basis set may be employed) which reduces to ordinary CI if orbital expansions are used. The "exact pairs" of Sinanoğlu[2,3] and the BG equations of Nesbet[8] are approximations to generalized CI.

The nice property of GPT as opposed to generalized CI or the BG equations of Nesbet is that in the latter method, (in order to be completely consistent) for a given size of basis set, successively larger secular equations must be solved, while in higher orders of GPT the sizes of the matrices do not greatly increase. The flexibility of GPT enables us to use our physical intuition to speed up convergence by tailoring GPT to the specific needs of a given problem.

II. GENERALIZED CI

The zeroth-order wave function Φ_0 is a closed-shell HF wave function[17] (the case of a single Slater determinant of arbitrary orthonormal spin-orbitals is discussed in Sec. IV), where

$$\Phi_0 = \mathcal{A}_N \prod_{i=1}^{N} l_i(i) = \mathcal{A}_N \varphi_0, \tag{2.1}$$

$$\mathcal{A}_N = \frac{1}{\sqrt{N!}} \sum_{i=1}^{N} (-1)^P P, \tag{2.2}$$

and the $\{l_i \,|\, i=1,\cdots,N\}$ are the ground-state HF spin-orbitals. If we define

$$\varphi_{l_k} = \left[\prod_{i=1}^{N} l_i(i)\right] / l_k(k),$$

$$\varphi_{l_j l_k} = \left[\prod_{i=1}^{N} l_i(i)\right] / l_j(j) l_k(k), \text{ etc.,} \tag{2.3}$$

then the exact wave function ψ may be written as a cluster expansion

$$\psi = \mathcal{A}_N \Bigg(\varphi_0 + \sum_l \varphi_l w_l + \frac{1}{\sqrt{2!}} \sum_{l>l'} \varphi_{ll'} w_{ll'} + \frac{1}{\sqrt{3!}} \sum_{l>l'>l''} \varphi_{ll'l''} w_{ll'l''} + \cdots \Bigg), \tag{2.4}$$

where $\{w_l\}$, $\{w_{ll'}\}$, and $\{w_{ll'l''}\}$ are one-, two-, and three-particle "cluster" functions, respectively. These cluster functions are antisymmetrized so that

$$\mathcal{A}_m w_{l_1 \ldots l_m} = (m!)^{1/2} w_{l_1 \ldots l_m}, \tag{2.5a}$$

and they are strongly orthogonal to the ground state HF orbitals, i.e.,

$$Q_1 Q_2 \cdots Q_m w_{l_1 \ldots l_m} = w_{l_1 \ldots l_m}. \tag{2.5b}$$

We should note that Eq. (2.4) is a perfectly general way of writing the *exact* wave function when the initial approximation is in the form of a single determinant (with the proper symmetry, of course). In practice we are usually forced to truncate (2.4) either by keeping only some of the cluster functions and/or by expanding these functions in a limited basis set. For instance, if only the one- and two-particle cluster functions are kept, and the energy from such a wave function is varied with respect to these cluster functions as well as the ground state orbitals (in order to maintain self-consistency between the $\{l_i\}$ and the $\{w_l, w_{ll'}\}$), the nonlinear, coupled eigenvalue equations of Szász are obtained[5,6] (provided the $\{l_i\}$ are kept orthonormal and strongly orthogonal to the $\{w_l, w_{ll'}\}$).[7] Such nonlinear equations could serve as the basis for a perturbative expansion; however, a linear set of equations is much more suitable. Besides, we usually can assume that we initially have a good approximate wave function Φ_0 which would not be appreciably altered by the inclusion of the single and double excitations (see Sec. V of paper I).

The cluster functions in (2.4) could be decomposed quite generally, following Sinanoğlu, into what he calls linked and unlinked parts.[1,2] However, if we vary the energy from such a wave function, we would again get nonlinear equations. Therefore, the correlation energy from the wave function in (2.4) is varied by making arbitrary changes in all of the cluster functions to obtain the Schrödinger equation for the cluster functions as a set of coupled *linear* inhomogeneous, integrodifferential, eigenvalue equations. The eigenfunctions are the cluster functions, and the eigenvalue is the correlation energy $\Delta E = E - E_{\text{HF}}$.

Before explicitly writing down the resulting equations, a few interesting points should be noted. We can expand the cluster functions in terms of a correlated basis set if desired, and if a linear expansion is made, the resulting equations for the expansion coefficients are just the secular equations. Thus, the Schrödinger equation for the cluster functions can serve as a basis for a generalized CI as well as providing the starting point for generalized perturbation theory.

The equations for Sinanoğlu's "exact pairs" result from keeping only one of the pair functions $w_{ll'}$ in (2.4) nonzero and varying the correlation energy. This gives us a set of uncoupled linear inhomogeneous eigenvalue equations for each "exact pair" and pair correlation energy.[1,2,9] (Sinanoğlu's derivation differs from this procedure.) Nesbet's BG equations are also special cases of generalized CI. His BG pairs result if only w_l, $w_{l'}$, $w_{ll'}$ are kept nonzero and are expanded in an orbital basis set, while his three-particle BG equations are obtained by keeping w_l, $w_{l'}$, $w_{l''}$, $w_{ll'}$, $w_{ll''}$, $w_{l'l''}$, $w_{ll'l''}$.[8] Although Nesbet's N-body BG equation (i.e., the complete CI equation) gives the exact wave function, the necessary computational work increases enormously when going, e.g., from Nesbet's two- to his three-particle BG equations when the same size basis set is chosen. This points to the need for perturbative solutions in which at each stage smaller matrices must be handled.

Since the equations of generalized CI are *linear* and coupled, they can be written most compactly in matrix notation. Let $\vec{\mathrm{w}}$ be a column vector of the 2^N-1 different cluster functions whose indices are ordered (superscript T denotes transpose of a matrix),

$$\vec{\mathrm{w}}^T = (w_{l_1}, w_{l_2}, \ldots, w_{l_N}, w_{l_1 l_2}, \ldots, w_{l_{N-1} l_N}, \ldots, w_{l_1 \ldots l_N}). \tag{2.6}$$

The *operators* which act on these cluster functions can be collected into a $(2^N-1)\times(2^N-1)$ matrix $\Delta\overleftrightarrow{\mathrm{H}}$ whose elements are labeled by the indices which label the cluster functions. The components of $\Delta\overleftrightarrow{\mathrm{H}}$ inside the one-, two-, and three-electron blocks of this matrix are summarized in Table I. The remaining components of $\Delta\overleftrightarrow{\mathrm{H}}$ may be obtained by induction. We should note that the diagonal components are like $E^0 - H$ operators for n particles moving in the field of the remaining $N-n$ electrons, which are in their ground-state orbitals. Define the (2^N-1)-dimensional column vector $\vec{\mathrm{I}}$ whose components are

$$\vec{\mathrm{I}}_{l_i l_j} = Q_{12}(1/r_{12})\,|[l_i l_j]_{12}\rangle \tag{2.7}$$

$$= 0 \text{ otherwise.}$$

The Schrödinger equation for the eigenvalue ΔE

TABLE I. Components of the one-, two-, and three-particle blocks of ΔH for the case in which the zeroth-order approximation is a closed-shell Hartree-Fock wave function. Other components may be found by induction.

$\Delta\overleftrightarrow{H}_{l,l}$	$=Q_1[H_0(1)-K_{ll}(1)Q_1-\epsilon_l]$.
$\Delta\overleftrightarrow{H}_{l,l'}$	$=-Q_1K_{l'l}(1)Q_1,\quad l\neq l'$.
$\Delta\overleftrightarrow{H}_{ll',ll'}$	$=Q_{12}[H_0(12)+V_{ll'}(12)-\epsilon_{ll'}]Q_{12}$.
$\Delta\overleftrightarrow{H}_{ll',ll''}$	$=Q_{12}\left\langle [ll'']_{12}\middle\vert\frac{1}{r_{12}}\middle\vert [ll']_{12}\right\rangle-Q_{12}\sum_{i=1}^{2}K_{l''l'}(i)Q_{12},\quad l''\neq l'$.
$\Delta\overleftrightarrow{H}_{ll',l''l'''}$	$=Q_{12}\left\langle [l''l''']\middle\vert\frac{1}{r_{12}}\middle\vert [ll']_{12}\right\rangle,\quad l$ or $l'\neq l''$ or l'''.
$\Delta\overleftrightarrow{H}_{ll'l'',ll'l''}$	$=Q_{123}[H_0(123)+\vec{V}_{ll'l''}(123)-\epsilon_{ll'l''}]$, where $\epsilon_{ll'l''}=\epsilon_l+\epsilon_{l'}+\epsilon_{l''},\quad Q_{123}=Q_1Q_2Q_3$,
and $\vec{V}_{ll'l''}(123)$	$=Q_{123}\left(\sum_{i>j=1,2,3}\frac{1}{r_{ij}}-\sum_{a=l,l',l''}\sum_{i=1}^{3}K_{aa}(i)+\sum_{a>b=l,l'l''}\left\langle [ab]_{12}\middle\vert\frac{1}{r_{12}}\middle\vert [ab]_{12}\right\rangle\right)Q_{123}$.
$\Delta\overleftrightarrow{H}_{ll'l'',ll'l'''}$	$=Q_{123}\left(\sum_{a=l,l'}\left\langle [al''']_{12}\middle\vert\frac{1}{r_{12}}\middle\vert [al'']_{12}\right\rangle-\sum_{i=1}^{3}K_{l'''l''}(i)\right)Q_{123},\quad l''\neq l'''$.
$\Delta\overleftrightarrow{H}_{ll'l'',ll'''l''''}$	$=Q_{123}\left\langle [l'''l'''']\middle\vert\frac{1}{r_{12}}\middle\vert [l'l'']_{12}\right\rangle,\quad l'$ or $l''\neq l'''$ or l''''.
$\Delta\overleftrightarrow{H}_{ll'l'',l'''l''''l'''''}$	$=0,\quad l,\ l',$ or $l''\neq l''',\ l'''',$ or l'''''.
$\Delta\overleftrightarrow{H}_{ll',l}$	$=Q_{12}\left(\frac{1}{r_{12}}-\sum_{i=1}^{2}K_{l'l'}(i)\right)\frac{(1-P_{12})}{\sqrt{2}}\vert l(2)\rangle Q_1$.
$\Delta\overleftrightarrow{H}_{ll',l''}$	$=-Q_{12}\sum_{i=1}^{2}K_{l''l'}(i)\frac{(1-P_{12})}{\sqrt{2}}\vert l(2)\rangle Q_1,\quad l''\neq l$ or l'.
$\Delta\overleftrightarrow{H}_{l'',ll'}$	$=\Delta\overleftrightarrow{H}_{ll',l''}{}^{\dagger}$ (all l'')
$\Delta\overleftrightarrow{H}_{ll'l'',ll'}$	$=Q_{123}\left(\sum_{i>j}^{3}\frac{1}{r_{ij}}-\sum_{i=1}^{3}\sum_{a=ll'}K_{aa}(i)\right)\frac{(1-P_{13}-P_{23})}{\sqrt{3}}\vert l''(3)\rangle Q_{12}$.
$\Delta\overleftrightarrow{H}_{ll'l'',ll'''}$	$=Q_{123}\left(-\sum_{i=1}^{3}K_{l'''l'}(i)\right)\frac{(1-P_{13}-P_{23})}{\sqrt{3}}\vert l''(3)\rangle Q_{12},\quad l'''\neq l',l''$.
$\Delta\overleftrightarrow{H}_{ll'l'',l'''l''''}$	$=0,\quad l''',l''''\neq l,l'$ or l''.
$\Delta\overleftrightarrow{H}_{l'''l'''',ll'l''}$	$=\Delta\overleftrightarrow{H}_{ll'l'',l'''l''''}{}^{\dagger}$, (all l''',l'''').
$\Delta\overleftrightarrow{H}_{ll'l'',l}$	$=Q_{123}\sum_{i>j=1}^{3}\frac{1}{r_{ij}}\frac{(1-P_{13}-P_{12})}{\sqrt{3}}\vert [l'l'']_{23}\rangle Q_1$.
$\Delta\overleftrightarrow{H}_{ll'l'',l'''}$	$=0,\quad l'''\neq l,l',l''$
$\Delta\overleftrightarrow{H}_{l,ll'l''}$	$=\Delta\overleftrightarrow{H}_{ll'l'',l}{}^{\dagger}$.

$=E-E_{\text{HF}}$ and eigenfunction $\vec{w}$ is [see (3.14) and (3.15)]

$$\Delta\overleftrightarrow{H}\vec{w}+\vec{I}=\Delta E\vec{w}. \tag{2.8}$$

The elements of $\Delta\overleftrightarrow{H}$ are defined only when the row (and column) indices are distinct, i.e., $\Delta H_{l_il_j\ldots,l_ml_n\ldots}$ is defined for $l_i\neq l_j$, $l_m\neq l_n$, etc. The matrix multiplication in (2.8) is, of course, only defined when the column indices of $\Delta\overleftrightarrow{H}$ match (when in the same order as) the indices of $\vec{w}$. Because the functions in $\vec{w}$ are antisymmetric, they are invariant if multiplied by $(-1)^PP$ where P permutes the spin-orbital indices. The elements of $\Delta\overleftrightarrow{H}$ as presented in Table I do not necessarily have their row indices ordered. Hence, if P_r and P_c are the permutations of the row and column indices of an element of $\Delta\overleftrightarrow{H}$ necessary to put them in the same order as those of the elements of $\vec{w}$ which this element of $\Delta\overleftrightarrow{H}$ multiplies, a factor of $(-1)^{P_r+P_c}$ must be introduced. (In Sec. III we shall encounter terms of the form $\langle \vec{\mathrm{w}}^T | \Delta\overleftrightarrow{\mathrm{H}} | \vec{\mathrm{w}} \rangle$ for which both row and column indices of $\Delta\overleftrightarrow{\mathrm{H}}$ must be ordered properly.)

If Φ_0 were chosen as a single Slater determinant of arbitrary orthonormal spin-orbitals (see Sec. IV), the components of $\Delta\overleftrightarrow{\mathrm{H}}$ would be slightly modified, and the inhomogeneity would also contain a nonzero one-electron part.

In order to obtain a perturbative expansion of $\vec{\mathrm{w}}$ and ΔE, $\Delta\overleftrightarrow{\mathrm{H}}$ is separated as

$$\Delta\overleftrightarrow{\mathrm{H}} = \Delta\overleftrightarrow{\mathrm{H}}^{(0)} + \lambda\Delta\overleftrightarrow{\mathrm{H}}^{(1)} + \lambda^2\Delta\overleftrightarrow{\mathrm{H}}^{(2)} + \cdots, \tag{2.9a}$$

then $\vec{\mathrm{w}}$ and ΔE may be expanded as

$$\vec{\mathrm{w}} = \vec{\mathrm{w}}^{(0)} + \lambda\vec{\mathrm{w}}^{(1)} + \lambda^2\vec{\mathrm{w}}^{(2)} + \cdots, \tag{2.9b}$$

$$\Delta E = \Delta E^{(0)} + \lambda\,\Delta E^{(1)} + \lambda^2\Delta E^{(2)} + \cdots, \tag{2.9c}$$

and if $\vec{\mathrm{I}}$ is also expanded in powers of the perturbation parameter λ, terms of the same order in λ are equated to give the equations of GPT. It is clear that $\Delta\overleftrightarrow{\mathrm{H}}$ may be decomposed according to (2.9a) in a large number of ways. Thus we could generate a large number of forms for GPT. The order parameter λ, could in principle, be assigned to the elements of $\Delta\overleftrightarrow{\mathrm{H}}$ in conjunction with the size of these operators in a given problem. The particular choice for the decomposition of $\Delta\overleftrightarrow{\mathrm{H}}$ has been made in order that $\vec{\mathrm{w}}^{(1)}$ be composed of the perturbative BG pairs of (1.5).

III. GENERALIZED PERTURBATION THEORY

GPT is very similar in structure to OPT.[11] In fact the two different types of perturbation theories are formally identical in the sense that given a general property of OPT (i.e., a property which we shall call independent of the specific nature of the Hamiltonian), a similar property exists in GPT. In order to demonstrate this formal similarity, OPT is written in a special form.

Let

$$H^{(0)}\varphi^{(0)} = E^{(0)}\varphi^{(0)} \tag{3.1}$$

be the unperturbed problem. We seek a solution to the Schrödinger equation

$$H\psi = E\psi \tag{3.2}$$

where

$$H = H^{(0)} + \lambda H^{(1)}, \tag{3.3a}$$

$$\psi = \varphi^{(0)} + \chi, \tag{3.4a}$$

$$\langle \varphi^{(0)} | \psi \rangle = \langle \varphi^{(0)} | \varphi^{(0)} \rangle = 1, \tag{3.4b}$$

$$E^{(1)} = \langle \varphi^{(0)} | H^{(1)} | \varphi^{(0)} \rangle.$$

OPT is obtained by substituting the expansions

$$\chi = \lambda\chi^{(1)} + \lambda^2\chi^{(2)} + \cdots \tag{3.3b}$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \cdots \tag{3.3c}$$

into the inhomogeneous Schrödinger equation for χ

$$[(H^{(0)} - E^{(0)}) + \lambda(H^{(1)} - E^{(1)})]\chi + \lambda(H^{(1)} - E^{(1)})\varphi^{(0)} = \Delta E(\varphi^{(0)} + \chi), \tag{3.5}$$

where $\Delta E = E - E^{(0)} - E^{(1)} = \langle \varphi^{(0)} | H^{(1)} | \chi \rangle$ (3.6)

and then equating the coefficients of λ^n for all $n = 0, 1, \cdots \infty$.[11]

In (3.5) the first term is of the form $\Delta H\chi$ and the second represents the inhomogeneity, but the right-hand side contains both $\varphi^{(0)}$ and χ. The left-hand side of (2.8) is of the same form as the left-hand side of (3.5), but the right-hand side of (2.8) does not contain the unperturbed function explicitly. However, if we let

$$P^{(0)} = 1 - |\varphi^{(0)}\rangle\langle\varphi^{(0)}|. \tag{3.7a}$$

then $P^{(0)}\chi \equiv \chi$. (3.7b)

and (3.5) becomes using (3.6) and (3.7)

$$P^{(0)}[(H^{(0)} - E^{(0)}) + \lambda(H^{(1)} - E^{(1)})]P^{(0)}\chi + \lambda P^{(0)}H^{(1)}\varphi^{(0)} = \Delta E\chi. \tag{3.8}$$

In (3.8) $\varphi^{(0)}$ no longer appears multiplied by ΔE, and hence (3.8) is formally identical to (2.8) if we take

$$\Delta\overleftrightarrow{\mathrm{H}} = \Delta\overleftrightarrow{\mathrm{H}}^{(0)} + \lambda\Delta\overleftrightarrow{\mathrm{H}}^{(1)}, \quad \vec{\mathrm{I}} = \lambda\vec{\mathrm{I}}^{(1)}, \tag{3.9}$$

and make the substitutions

$$P^{(0)}[(H^{(0)} - E^{(0)}) + \lambda(H^{(1)} - E^{(1)})]P^{(0)} \to \Delta\overleftrightarrow{\mathrm{H}}, \tag{3.10a}$$

$$\chi \to \vec{\mathrm{w}}, \tag{3.10b}$$

$$\lambda P^{(0)}H^{(1)}\varphi^{(0)} \to \vec{\mathrm{I}} \tag{3.10c}$$

and $\Delta E \to \Delta E$ (3.10d)

and replace ordinary multiplication by matrix multiplication. Thus, any property of OPT which is independent of the structure of $H^{(0)}$, $H^{(1)}$, $\varphi^{(0)}$ is also a property of GPT because of the analogy (3.10) between Eqs. (2.8) and (3.8).

If in (2.9a), we endowed $\Delta\overleftrightarrow{\mathrm{H}}$ with a second-order part also, some relationships between OPT and GPT could also be established provided $\vec{\mathrm{I}}$ also has a second-order component.

Some properties of GPT which can be inferred from Eqs. (3.10) are as follows:[11]

(1) The "$2n+1$ rule of perturbation theory" which states that knowledge of the wave function (i.e., $\vec{\mathrm{w}}$) through nth order of GPT enables the calculation of ΔE through $(2n+1)$st order of GPT,

(2) the existence of a variational principle for $\vec{\mathrm{w}}^{(n)}$ and $\Delta E^{(2n)}$, and

(3) the interchange theorem of double-perturbation theory which could be demonstrated by including a second perturbation μW in (3.5) and (3.8) which would give rise to extra terms in these equations which would be analogous to a component of $\Delta\overleftrightarrow{\mathrm{H}}$ and $\vec{\mathrm{I}}$ of order μ.

In the remainder of this section, the lowest-order equations of GPT are discussed along with a brief presentation of examples of properties (1) and (2) above. Generalized double-perturbation theory is briefly discussed in Sec. V.

From (3.6), using (3.10), we have

$$\Delta E = \vec{\mathrm{I}}^{\dagger}\vec{\mathrm{w}} \equiv \langle \vec{\mathrm{I}}^{T} | \vec{\mathrm{w}} \rangle \equiv \sum_{l_1 > l_2 > \ldots}^{2^N - 1} \langle I_{l_1 l_2 \ldots} | w_{l_1 l_2 \ldots} \rangle, \tag{3.11}$$

i.e., matrix multiplication is implied. Equation (3.11) thus presents the correlation energy of a closed-shell HF system as the sum of pair-correlation energies, since for this case $\vec{\mathrm{I}}$ only has two-particle components [see (2.7)]. In order to fully utilize the analogy (3.10), the decompositions of

$\Delta\overleftrightarrow{H}$ and $\vec{I}$ in (3.9) have implicitly been assumed. Thus the lowest order of $\vec{w}$ is first order,[18] and from (3.11), the lowest order of ΔE is second order.

The equation for the first-order "wave function" $\vec{w}^{(1)}$ is therefore

$$\Delta\overleftrightarrow{H}^{(0)}\vec{w}^{(1)}+\vec{I}=\vec{0}. \tag{3.12}$$

If $\Delta\overleftrightarrow{H}^{(0)}$ is taken to be the diagonal part of $\Delta\overleftrightarrow{H}$, and $\Delta\overleftrightarrow{H}^{(1)}$ is the off-diagonal part; then comparing (3.12), (2.7), and Table I with (1.5) shows that $\vec{w}^{(1)}$ is composed solely of the BG perturbative pairs describing the correlation corrections for a pair of electrons moving in the field of the remaining electrons, which are in their ground-state orbitals.[19,20] From the numerical calculations of Kelly, Nesbet, Sinanoğlu, and Geller *et al.*[4,8,10,11,12,21] in which they evaluate a correlation energy which should be roughly similar to that obtainable from

$$\Delta E^{(2)}=\vec{I}^{\dagger}\vec{w}^{(1)}, \tag{3.13}$$

we can see that the lowest order of generalized perturbation theory when starting from closed-shell HF should be expected to provide quite good correlation energies and probably good wave functions.

Functions obtained from GPT could be used to calculate rigorous upper bounds to the correlation energy by using the cluster-type wave function (2.4). Equivalently, $(\Delta E)_t$ defined by

$$(\Delta E)_t=\frac{\vec{w}_t{}^{\dagger}\Delta\overleftrightarrow{H}\vec{w}_t+\vec{w}_t{}^{\dagger}\vec{I}+\vec{I}^{\dagger}\vec{w}_t}{1+\vec{w}_t{}^{\dagger}\vec{w}_t}, \tag{3.14}$$

where $\vec{w}_t$ denotes a set of antisymmetrized trial cluster functions, is always a rigorous upper bound to the correlation energy. In fact, the variational principle

$$\delta(\Delta E)_t/\delta\vec{w}_t=0, \tag{3.15}$$

is equivalent to the Schrödinger equation (2.8).[22]

$\Delta E^{(2)}$ and $\vec{w}^{(1)}$ can be obtained variationally since

$$F^{(2)}[\vec{w}_t]=\vec{I}^{\dagger}\vec{w}_t+\vec{w}_t{}^{\dagger}\vec{I}+\vec{w}_t{}^{\dagger}\Delta\overleftrightarrow{H}^{(0)}\vec{w}_t \tag{3.15a}$$

and $\delta F^{(2)}/\delta\vec{w}_t=0$ (3.16b)

are equivalent to (3.12), and by (3.13), the stationary value of $F^{(2)}$ is $\Delta E^{(2)}$. Equations (3.16) actually correspond to a variational solution for each individual BG pair and pair-correlation energy.[2,3,12]

The second-order functions are determined by the equations

$$\Delta\overleftrightarrow{H}^{(0)}\vec{w}^{(2)}+\Delta\overleftrightarrow{H}^{(1)}\vec{w}^{(1)}=0. \tag{3.17}$$

Since $\vec{w}^{(1)}$ consists only of pair functions (for closed-shell HF), $\vec{w}^{(2)}$ consists of one-, two-, three-, and four-electron functions just as in OPT. Equations (3.17) are *uncoupled linear* inhomogeneous equations for these one-, ..., four-electron functions. Thus, the solution of the second-order functions of GPT (which would give the fifth-order energies!) involves a set of smaller matrices than obtained in Nesbet's three-body BG equations (i.e., "second order" in his theory).

OPT has the structure dependent property that the second-order four-particle functions are completely expressible in terms of the first-order pairs.[12,23,24] This property is lost when we use GPT, but such a factorization might be expected to give a reasonable approximation.[1,2,12] As discussed in paper I, we may wish to try to optimize this approximation by introducing a set of variational parameters in this factorization and by using the usual variational principle of (3.14) and (3.15) to determine these parameters. The results of paper I can be used to express $\vec{w}^{(1)}$, $\vec{w}^{(2)}$, ... as well as the cluster functions obtained from the decomposition of $\Delta\overleftrightarrow{H}$ as discussed in Footnote 20, in terms of contributions from OPT.[25]

Since (3.11) implies that

$$\Delta E^{(3)}=\vec{I}^{\dagger}\vec{w}^{(2)}, \tag{3.18}$$

multiplying (3.12) by $\vec{w}^{(2)\dagger}$, (3.17) by $\vec{w}^{(1)\dagger}$, taking the complex conjugate of one of the resulting equations and subtracting converts (3.18) to

$$\Delta E^{(3)}=\vec{w}^{(1)\dagger}\Delta\overleftrightarrow{H}^{(1)}\vec{w}^{(1)}, \tag{3.18a}$$

which is the simplest example of the $2n+1$ rule of GPT. By similar, *and standard*,[11] methods we could demonstrate properties (1) and (2) in higher orders of generalized perturbation theory.

Now that we have established GPT for a closed-shell HF system, it is of interest to generalize this method to cover the case in which the zeroth-order wave function is a single Slater determinant of arbitrary orthonormal spin-orbitals. This lifts the restriction that HF orbitals be used for closed shells and also enables one to treat some open-shell systems where, in practice, restricted HF orbitals may be available. This treatment of an "arbitrary" single-determinantal Φ_0 is given in the next section.

IV. "ARBITRARY" SINGLE SLATER DETERMINANT

If GPT were only applicable to cases in which the zeroth-order wave function is a closed-shell HF wave function, it would be extremely restrictive. The extension of the results to cover an "arbitrary" single Slater determinant as the zeroth-order wave function is not completely general, but it is obvious that generalized perturbation theory could be obtained when Φ_0 is a symmetry-restricted sum of Slater determinants. The algebra would, of course, be tedious.[26]

Often we start with a zeroth-order wave function of the form

$$\Phi_0=\mathcal{A}_N\prod_{i=1}^{N}p_i(i), \tag{4.1}$$

where the $\{p_i|i=1,\cdots,N\}$ are called the ground-state, or occupied, spin-orbitals. These ground-state orbitals may all be eigenfunctions of the same one-electron Hamiltonian $h_0(i)=t(i)+u(i)$, where $t(i)$ contains the kinetic-energy operator and the attraction between the electrons and the nuclei and $u(i)$ is some one-body potential which describes some of the effects of the electron repulsion. For the purposes of GPT, it is not necessary to require that such a $h_0(i)$ [or $u(i)$] exist; and furthermore, if Φ_0 is the eigenfunction of some Hamiltonian, we do not require a knowledge of it.[16] It is convenient, however, to require that the spin-orbitals $\{p_i\}$ form an orthonormal set.

If we define the single-particle energies

$$\epsilon_p = \langle p \,|\, t + K_0 \,|\, p \rangle, \tag{4.2}$$

where

$$K_0(i) = \sum_{p\ \text{occpd}} \left\langle p(\mu) \left| \frac{1-P_{i\mu}}{r_{i\mu}} \right| p(\mu) \right\rangle = \sum_{p\ \text{occpd}} K_{pp}(i), \tag{4.3}$$

then

$$E^{(0)} + E^{(1)} = \langle \Phi^{(0)} \,|\, H \,|\, \Phi^{(0)} \rangle \tag{4.4}$$

$$= \sum_p \epsilon_p - \sum_{p > p'} \left\langle [pp']_{12} \left| \frac{1}{r_{12}} \right| [pp']_{12} \right\rangle. \tag{4.5}$$

in analogy with the form of the energy obtained when $\Phi^{(0)}$ is a HF wave function.

Define

$$H_0(i) \equiv t(i) + K_0(i) \tag{4.6}$$

as the one-electron Hamiltonian for an electron moving in the field of the N electrons which are in their ground-state spin-orbitals. *The* $\{p_i\}$ *are not eigenfunctions of this* H_0. However, if we let $\Delta\overleftrightarrow{\mathrm{H}}(l)$ denote the Hamiltonian of GPT for a closed-shell HF system as given in Table I, and if $\Delta\overleftrightarrow{\mathrm{H}}(p)$ denotes the Hamiltonian appropriate to the zeroth-order wave function (4.1), then

$$\Delta\overleftrightarrow{\mathrm{H}}(p) = \Delta\overleftrightarrow{\mathrm{H}}(l)\big|_{l \to p} + \delta\Delta\overleftrightarrow{\mathrm{H}}(p) \tag{4.7}$$

where $\Delta\overleftrightarrow{\mathrm{H}}(l)\,|_{l \to p}$ implies that in Table I, e.g.,

$\epsilon_l \to$ the ϵ_p of (4.2),

$H_0 \to$ the H_0 of (4.3),

$Q_1 \to 1 - \sum_{p\ \text{occpd}} |p^{(1)}\rangle\langle p^{(1)}|,$

i.e., $l \to p$, $l' \to p'$, etc. The extra term $\delta\Delta\overleftrightarrow{\mathrm{H}}(p)$ only has a few non-zero components, and these are summarized in Table II for the one-, two-, and three-electron blocks of $\delta\Delta\overleftrightarrow{\mathrm{H}}(p)$. The same sign convention holds for $\delta\Delta\overleftrightarrow{\mathrm{H}}(p)$ as does for $\Delta\overleftrightarrow{\mathrm{H}}(l)$ and $\Delta\overleftrightarrow{\mathrm{H}}(l)\,|_{l \to p}$. The inhomogeneity now becomes

$$\vec{\mathrm{I}}(p) = \vec{\mathrm{I}}(l)\Big|_{l \to p} + \delta\vec{\mathrm{I}}(p), \tag{4.8}$$

$$\text{where}\quad \mathrm{I}(l)_{l_i l_j}\Big|_{l \to p} \equiv Q_{12}(p) \frac{1}{r_{12}} |[p_i p_j]_{12}\rangle$$
$$= 0 \text{ otherwise}, \tag{4.9}$$

$$\text{and}\quad \delta\vec{\mathrm{I}}_{p_i}(p) = Q_1(p) H_0(1) \,|\, p_i(1)\rangle$$
$$= 0 \text{ otherwise}, \tag{4.10}$$

where, obviously, the $H_0(1)$ in (4.10) refers to the quantity defined in (4.6).

From Table II, we see that $\delta\Delta\overleftrightarrow{\mathrm{H}}(p)$ has nonzero components when the row and column indices have the same number of spin-orbital indices and differ only in one of these indices. Furthermore, these terms are proportional to the off-diagonal matrix elements of the H_0 of (4.6). Thus, in the event that the $\{p_i \,|\, i = 1, \cdots, N\}$ are linear combinations of HF orbitals, $\delta\vec{\mathrm{I}}(p)$ in (4.10) vanishes indentically, and $\delta\Delta\overleftrightarrow{\mathrm{H}}(p)$ is written in terms of the HF off-diagonal Lagrange multipliers, since the last two terms in Table II also vanish identically in this case. Thus the Schrödinger equation

$$\Delta\overleftrightarrow{\mathrm{H}}(p)\vec{\Psi} + \vec{\mathrm{I}}(p) = \Delta E \vec{\Psi} \tag{4.11}$$

reduces to the HF result (2.8), but is also easily written in terms of localized HF spin-orbitals.[27,28] The fact that the spin-orbitals $\{p_i \,|\, i = 1, \cdots, N\}$ need only be orthonormal, enables us to use different kinds of orbital functions for different purposes. For example, in molecules Slater-type orbitals could be used for the core electrons and elliptical orbitals for the valence electrons if desired.[29]

TABLE II. Non-zero components of the one, two, and three particle blocks of $\delta\Delta\overleftrightarrow{\mathrm{H}}(p)$ for the case of a zeroth order wave function which is a single Slater determinant of arbitrary orthonormal spin orbitals. Other components may be found by induction. H_0 is defined in (4.6).

$\delta\Delta\overleftrightarrow{\mathrm{H}}_{p',p''}$	$= -Q_1 \langle p'' \,|\, H_0 \,|\, p' \rangle, \quad p'' \neq p'.$
$\delta\Delta\overleftrightarrow{\mathrm{H}}_{pp',pp''}$	$= -Q'_{12} \langle p'' \,|\, H_0 \,|\, p' \rangle, \quad p'' \neq p'.$
$\delta\Delta\overleftrightarrow{\mathrm{H}}_{pp'p'',pp'p'''}$	$= -Q_{123} \langle p''' \,|\, H_0 \,|\, p'' \rangle, \quad p''' \neq p''.$
$\delta\Delta\overleftrightarrow{\mathrm{H}}_{pp',p}$	$= Q_{12} H_0(12) \frac{(1-P_{12})}{\sqrt{2}} \,|\, p'(2)\rangle Q_1$ $= (\delta\Delta\overleftrightarrow{\mathrm{H}}_{p,pp'})^\dagger.$
$\delta\Delta\overleftrightarrow{\mathrm{H}}_{pp'p'',pp'}$	$= Q_{123} H_0(123) \frac{(1-P_{13}-P_{23})}{\sqrt{3}} \,|\, p''(3)\rangle Q_{12}$ $= (\delta\Delta\overleftrightarrow{\mathrm{H}}_{pp',pp'p''})^\dagger.$

When non-Hartree-Fock $\{p_i\}$ are used and $\Delta\overleftrightarrow{\mathrm{H}}^{(0)}(p)$ is the diagonal part of $\Delta\overleftrightarrow{\mathrm{H}}(p)$ while the off-diagonal part of $\Delta\overleftrightarrow{\mathrm{H}}(p)$ and $\vec{\mathrm{I}}(p)$ are taken to be first-order quantities, $\vec{\Psi}^{(1)}$ contains both orbital and pair functions which are the perturbative corrections for one and two electrons, respectively, moving in the field of the remaining electrons, which in turn are in their ground state orbitals. The $2n+1$ rule and variation-perturbation theory for $\Delta E^{(2n)}$ and $\vec{\Psi}^{(n)}$ still hold for these "arbitrary" $\{p_i\}$. If the $\{p_i\}$ are "good orbitals," then $\vec{\Psi}^{(1)}(p_i)$ should give results comparable to those achievable with the $\vec{\Psi}^{(1)}$ obtained from HF orbitals.

Generalized double-perturbation theory is only discussed briefly in the next section, since the results should be obvious from the analogy (3.10) of Eqs. (2.8) and (4.11) with (3.8).

V. GENERALIZED DOUBLE PERTURBATION THEORY

The discussion of generalized double perturbation is given for the case of a zeroth-order wave function which is a single Slater determinant of arbitrary orthonormal orbitals.[11] (The $\{p_i\}$ of Sec. IV.) These results reduce to the HF case. If we add the one-electron perturbation $\mu W = \mu \sum_{i=1}^{N} W(i)$ to the Hamiltonian H,[30] then the Schrödinger equation for the cluster functions is of the form [subtracting $\langle \Phi_0 \,|\, W \,|\, \Phi_0 \rangle$ in the same way that $E^{(1)}$ is subtracted in (3.5)]

$$[\Delta\overleftrightarrow{H}^{(0,0)}(p)+\lambda\Delta\overleftrightarrow{H}^{(1,0)}(p)+\mu\Delta\overleftrightarrow{W}^{(0,1)}]\vec{w} + \lambda\vec{I}^{(1,0)}(p)+\mu\vec{I}^{(0,1)}(p)=\Delta E\vec{w}, \tag{5.1}$$

where

$$\Delta\overleftrightarrow{H}^{(0,0)}(p)=\Delta\overleftrightarrow{H}^{(0)}(p), \tag{5.2a}$$

$$\Delta\overleftrightarrow{H}^{(1,0)}(p)=\Delta\overleftrightarrow{H}^{(1)}(p), \tag{5.2b}$$

$$\vec{I}^{(1,0)}(p)=\vec{I}(p), \tag{5.2c}$$

and the quantities on the right-hand side of Eqs. (5.2) are defined in Sec. IV. The non-zero components of $\vec{I}^{(0,1)}$ are

$$\vec{I}^{(0,1)}(p)=Q_1[W(1)-\langle p|W|p\rangle]|p(1)\rangle. \tag{5.3}$$

and $\Delta\overleftrightarrow{W}^{(0,1)}$ is summarized in Table III, where the sign conventions as to the interchange of row or column indices is the same for $\Delta\overleftrightarrow{W}^{(0,1)}$ as discussed for $\Delta\overleftrightarrow{H}$ as discussed for $\Delta\overleftrightarrow{H}$ in Sec. II.

If the cluster functions and ΔE are both expanded in a double power series in λ and μ[11]

$$\vec{w}=\sum_{\substack{n=0\\(m+n\neq 0)}}^{\infty}\sum_{m=0}^{\infty}\lambda^n\mu^m\vec{w}^{(n,m)} \tag{5.4a}$$

$$\Delta E=\sum_{\substack{n\neq 0\\(m+n\neq 0,1)}}^{\infty}\sum_{m=0}^{\infty}\lambda^n\mu^m\Delta E^{(n,m)}, \tag{5.4b}$$

and if (5.4a) and (5.4b) are both substituted into (5.1), then, by equating coefficients of $\lambda^n\mu^m$, the equations of double-perturbation theory are obtained. The important point is that, because of the analogy of (5.1) with the equation resulting from taking $\lambda(H^{(1)}-E^{(1)})\rightarrow\lambda(H^{(1)}-E^{(1)})+\mu(W-\langle\Phi_0|W|\Phi_0\rangle)$ in (3.8), Dalgarno's interchange theorems are also valid in generalized double-perturbation theory.[11]

As the simplest example,

$$\Delta E^{(1,1)}=\vec{I}^{(1,0)\dagger}\vec{w}^{(0,1)}+\vec{I}^{(0,1)\dagger}\vec{w}^{(1,0)}. \tag{5.5}$$

From (5.1) and Eqs. (5.4), the first-order cluster functions are determined by the equations

$$\Delta\overleftrightarrow{H}^{(0,0)}(p)\vec{w}^{(1,0)}+\vec{I}^{(1,0)}(p)=0 \tag{5.6a}$$

$$\Delta\overleftrightarrow{H}^{(0,0)}(p)\vec{w}^{(0,1)}+I^{(0,1)}(p)=0 \tag{5.6b}$$

We should note that the orbital corrections of $\vec{w}^{(1,0)}$ and $\vec{w}^{(0,1)}$ represent perturbative corrections for an electron *moving in the field of the remaining electrons*, which are in their ground state orbitals. Thus, these orbital corrections contain orbital corrections to all orders in OPT. Multiplying (5.6a) by $\vec{w}^{(0,1)\dagger}$, (5. 5b) by $\vec{w}^{(1,0)\dagger}$, and taking the complex conjugate of the latter equation and subtracting it from the former gives

$$\Delta E^{(1,1)}=\vec{I}^{(1,0)\dagger}\vec{w}^{(0,1)}+\vec{w}^{(0,1)\dagger}\vec{I}^{(1,0)} \tag{5.7}$$

when the above result is substituted into (5.5). The interpretation of (5.7) is already well understood,[11] and need not be discussed again here. Higher-order interchange theorems can likewise be established by analogy between OPT and GPT.

There is one further point worth noting. When second-order properties, i.e., $\Delta E^{(0,2)}$, are calculated (for one-electron W) starting from a closed shell HF wave function (or an unrestricted HF function for open shells), Eqs. (5.6b) correspond to an uncoupled perturbation theory (containing all orders in OPT). If $\Delta\overleftrightarrow{H}^{(0,0)}$ were taken to contain

TABLE III. Non-zero components of the one-, two-, and three-particle blocks of $\Delta\overleftrightarrow{W}$ for the case of a zeroth order wave function which is a single Slater determinant of arbitrary orthonormal spin-orbitals. Other components may be found by induction.

$$\Delta\overleftrightarrow{W}_{p,p}=Q_1W(1)Q_1.$$

$$\Delta\overleftrightarrow{W}_{p,p'}=-Q_1\langle p'|W|p\rangle,\quad p\neq p'.$$

$$\Delta\overleftrightarrow{W}_{pp',pp'}=Q_{12}\Big(\sum_{i=1}^{2}W(i)-\sum_{a=p,p'}\langle a|W|a\rangle\Big)Q_{12}.$$

$$\Delta\overleftrightarrow{W}_{pp',p''p'}=-Q_{12}\langle p''|W|p\rangle,\quad p''\neq p.$$

$$\Delta\overleftrightarrow{W}_{pp'p'',pp'p''}=Q_{123}\Big(\sum_{i=1}^{3}W(i)-\sum_{a=p,p',p''}\langle a|W|a\rangle\Big)Q_{123}$$

$$\Delta\overleftrightarrow{W}_{pp'p'',pp'p'''}=-Q_{123}\langle p'''|W|p''\rangle,\quad p''\neq p'''$$

$$\Delta\overleftrightarrow{W}_{pp',p}=Q_{12}\sum_{i=1}^{2}W(i)\frac{(1-P_{12})}{\sqrt{2}}|p'(2)\rangle Q_1.$$

$$\Delta\overleftrightarrow{W}_{p,pp'}=\Delta\overleftrightarrow{W}_{pp',p}^{\dagger}.$$

$$\Delta\overleftrightarrow{W}_{pp'p'',pp'}=Q_{123}\sum_{i=1}^{3}W(i)\frac{(1-P_{13}-P_{23})}{\sqrt{3}}|p''(3)\rangle Q_{12}.$$

$$\Delta\overleftrightarrow{W}_{pp',pp'p''}=\Delta\overleftrightarrow{W}_{pp'p'',pp'}^{\dagger}.$$

the whole one-electron block of $\Delta\overleftrightarrow{H}$, Eq. (5.6b) would correspond to a coupled HF perturbation theory. However, this coupled perturbation theory is not self-consistent since the excited orbitals do not modify the HF field.[31]

V. CONCLUSION

If the exact wave function is written formally in terms of cluster functions, application of the techniques of OPT to the Schrödinger equation for these cluster functions leads to GPT. The simple case of a zeroth-order wave function which is a single Slater determinant of arbitrary orthonormal spin-orbitals is discussed, but it is obvious that more general cases can also be treated in a similar fashion. Furthermore, if the wave function were written formally in another manner (either exact or approximate), e.g., a separated-pairs model, any zeroth-order wave function could be improved by the use of an analogous GPT which would result from a perturbative expansion of the Schrödinger equation peculiar to the chosen form of the wave function.

The form of GPT presented here is quite flexible, and different decompositions of the "Hamiltonian" $\Delta\overleftrightarrow{H}$ can lead to slightly different models for the lowest-order perturbative corrections. This flexibility of GPT may also enable decompositions of $\Delta\overleftrightarrow{H}$ into orders which are in line with the numerical orders of magnitude of these terms.

For the cluster-type wave functions, the simplest and most physical form of GPT leads to first-order corrections to the wave function which are written in terms of orbital and pair functions describing respectively the motion of one and two electrons in the field of the remaining electrons, which are in their ground state orbitals. These functions are expected to give accuracy comparable to that obtainable from a calculation of Sinanoğlu's "exact pairs," Nesbet's BG pairs, or a summation of the diagrams of many-body perturbation theory as is done by Kelly.

GPT is a systematic form of perturbation theory and does not require the use of HF orbitals. As Conkie discusses,[32] the time spent in obtaining accurate HF orbitals could probably be more efficiently spent in determining corrections to fairly good orbitals which are easily obtained. However, if HF orbitals are available, GPT can be used with some sort of localized orbitals[12] which would "minimize" higher-order corrections.

GPT should be much more rapidly convergent than OPT. From the numerical calculations of Kelly, Sinanoğlu, Geller *et al.*, and Nesbet, first-order GPT should give very good correlation energies. Furthermore, the lowest-order equations of GPT are "more physical" than those of OPT and contain more information about correlation (e.g., the fact that certain orbitals are no longer occupied when electrons are excited, the full electron repulsion between the excited electrons, etc.).

Because of the formal analogy between GPT and OPT, many of the theorems of OPT are also valid for GPT,[33] namely, the "$2n+1$ rule," variational principles of $\Delta E^{(2n)}$ and $\vec{w}^{(n)}$, and the interchange theorems of double-perturbation theory. The spin-orbitals of the zeroth-order wave function need not be eigenfunctions of a one-electron Hamiltonian, and we do not require a knowledge of the approximate Hamiltonian of which the zeroth-order wave function is an eigenfunction.[34]

---

*This research is supported by a grant from the National Science Foundation.

†National Science Foundation predoctoral fellow. Present address: James Franck Institute and Department of Chemistry, University of Chicago, Chicago, Illinois

[1] O. Sinanoğlu, J. Chem. Phys. 36, 706 (1962).

[2] O. Sinanoğlu, J. Chem. Phys. 36, 3198 (1962).

[3] D. F. Tuan and O. Sinanoğlu, J. Chem. Phys. 41, 2677 (1964).

[4] V. McKoy and O. Sinanoğlu, J. Chem. Phys. 41, 2689 (1964).

[5] L. Szász, Phys. Rev. 126, 169 (1962).

[6] L. Szász, Phys. Rev. 132, 936 (1963).

[7] J. C. Porter, Phys. Rev. 140, A732 (1965).

[8] R. K. Nesbet, Phys. Rev. 155, 51, 56 (1967).

[9] W. R. Conkie, Can. J. Phys. 43, 102 (1965).

[10] M. Geller, H. S. Taylor, and H. B. Levine, J. Chem. Phys. 43, 1727 (1965).

[11] For a review of OPT see J. O. Hirshfelder, W. B. Brown, and S. T. Epstein, Adv. Quantum Chem. 1, 256 (1964).

[12] Karl F. Freed, preceding paper [Phys. Rev. 173, 1 (1968)]; herein referred to as Paper I.

[13] H. P. Kelly, Phys. Rev. 131, 684 (1963).

[14] H. P. Kelly, Phys. Rev. 131, B896 (1964).

[15] In the light of this relationship between the BG pairs and the "exact pairs," the similarity of the results of Kelly (Refs. 13 and 14) with those obtained by Geller *et al.* (Ref. 10) [a calculation of the $(1s)^2$ and $(2s)^2$ "exact pairs" for Be] is not too surprising.

[16] S. T. Epstein, in Perturbation Theory and Its Application in Quantum Mechanics, edited by C. H. Wilcox (John Wiley & Sons, Inc., 1966), p. 49.

[17] C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951).

[18] The zeroth-order functions would be eigenfunctions of $\Delta\overleftrightarrow{H}^{(0)}$, i.e., $\Delta\overleftrightarrow{H}^{(0)}\vec{w}^{(0)}=0$. If $\vec{w}^{(0)}\neq 0$, there would be excited states of the form $\mathfrak{a}_N\varphi_{ll'}\cdots W_{ll'}\cdots=\Phi_{ex}$ such that $H\Phi_{ex}=E_{HF}\Phi_{ex}$.

[19] if $\Delta\overleftrightarrow{H}^{(0)}$ contains the diagonal part of the two-electron block of $\Delta\overleftrightarrow{H}$ minus $\epsilon^v{}_{ll'}$ along the diagonal in this block, where the $\epsilon^v{}_{ll'}$ are the lowest eigenvalues of the resulting *uncoupled* equations of the form (3.12), then the $\vec{w}^{(1)}$ are just Sinanoğlu's "exact pairs." (See Refs. 2 and 3.)

[20] If $\Delta\overleftrightarrow{H}^{(0)}$ contains the whole two-electron block of $\Delta\overleftrightarrow{H}$, but not any part of $\Delta\overleftrightarrow{H}$ connecting the two- and one-, three-, or four-electron blocks, then the $\vec{w}^{(1)}$ are the "complete pairs" of Paper I. The higher approximations in Paper I can also be generated from GPT by a suitable choice of $\Delta\overleftrightarrow{H}^{(0)}$.

[21] H. P. Kelly, Phys. Rev. 144, 39 (1966).

[22] As usual, if the state of interest is not the lowest state of a given symmetry, we must require that the wave function of the form (2.4) be orthogonal to the lower states of the same symmetry.

[23] R. E. Stanton, J. Chem. Phys. 42, 2353 (1965).

[24] J. Čížek, J. Chem. Phys. 45, 4256 (1966).

[25] In the notation of Paper I, $\{ll'|G_0^{(2)-1}|ll'\} \equiv -\Delta\overleftrightarrow{\mathrm{H}}_{ll',ll'}$, and hence $w^{(1)}_{ll'} = \{ll'|G_0^{(2)}|ll'\}Q_{12}(1/r_{12})|[ll']_{12}\rangle$, where the diagrammatic expansion of this expression is implicitly given in Paper I. The $\{\alpha|G^{(2)-1}|\beta\}$ of Paper I is identical to the negative of the whole two-electron block of $\Delta\overleftrightarrow{\mathrm{H}}$ and substituting $G^{(2)}$ for $G_0^{(2)}$ above, $\sum_{l''>l'''} \times\{ll'|G^{(2)}|l''l'''\}Q_{12}(1/r_{12})|[l''l''']_{12}\rangle$ gives the "complete pairs" discussed in Ref. 20, whose diagrammatic expansion in powers of OPT is also given in Paper I. The off-diagonal components of $\Delta\overleftrightarrow{\mathrm{H}}$ connecting cluster functions of say one (or three) and two electrons are just the $V^{(1,2)}$ (or $V^{(3,2)}$) operators of Paper I. The process of obtaining expansions of $\vec{\mathrm{w}}^{(1)}$, $\vec{\mathrm{w}}^{(2)}, \cdots$, in orders of OPT, therefore, need not be presented here in detail.

[26] H. J. Silverstone and O. Sinanoğlu, J. Chem. Phys. 44, 1899, 3608 (1966).

[27] We could look for the set of localized HF orbitals that reduce the size of higher-order corrections, i.e., that make the off-diagonal elements of $\Delta\overleftrightarrow{\mathrm{H}}$ as small as possible.

[28] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. 34, 457 (1963); J. Chem. Phys. 43, S97 (1965).

[29] J. C. Browne and F. A. Matsen, Phys. Rev. 135A, 1227 (1964).

[30] We can also obtain the results for the case in which W is a sum of two electron operators since $\Delta\overleftrightarrow{\mathrm{H}}(p) = \Delta\overleftrightarrow{t}(p) + \Delta\overleftrightarrow{\mathrm{v}}(p)$, where $\Delta\overleftrightarrow{\mathrm{t}}(p)$ is the part of $\Delta\overleftrightarrow{\mathrm{H}}(p)$ which contains only one-electron operators (and hence is of the same structure as $\Delta\overleftrightarrow{\mathrm{W}}$, and $\Delta\overleftrightarrow{\mathrm{v}}(p)$ is the remainder of $\Delta\overleftrightarrow{\mathrm{H}}(p)$ is which $1/r_{ij}$ occurs explicitly. [Note $\Delta\overleftrightarrow{\mathrm{v}}(p)$ occurs explicitly in ϵ_p through the operator K_0.]

[31] By extending the discussion presented in Paper I to include external perturbations, we could also obtain the diagrammatic expansion of these three forms of HF perturbation theory in powers of OPT.

[32] W. R. Conkie, J. Chem. Phys. 43, 3408 (1965).

[33] We could also form a first-order iterative GPT (Ref. 11) by taking the first-order wave function of GPT as a new zeroth-order wave function and finding the $\Delta\overleftrightarrow{\mathrm{H}}$ and $\vec{\mathrm{I}}$ appropriate to this wave function. The first-order corrections to this improved wave function describe the motion of electrons in the field of the remaining electrons which "spend their time" in the ground-state spin-orbitals and the original first-order spin-orbitals and pair functions.

[34] R. Yaris, J. Chem. Phys. 44, 3894 (1966).

Lifetimes and Alignment Depolarization Cross Sections of the $(3p^5\,4p)$ $^3D_3$ and $^1D_2'$ Levels in Argon and the $(1s2p)$ ^3P Term in Helium*

Donald A. Landman
New York University, University Heights, New York, New York

(Received 1 March 1968)

Hanle-effect measurements have been made of the radiative lifetimes of the $(3p^54p)^3D_3$ and $^1D_2'$ levels in argon ($2p_9$ and $2p_6$, respectively, in Paschen's notation) and the $(1s2p)^3P$ term in helium. The resulting free-atom lifetimes are $\tau(^3D_3) = 3.16(16) \times 10^{-8}$ sec, $\tau(^1D_2') = 2.90(17) \times 10^{-8}$ sec, and $\tau(^3P) = 9.90(84) \times 10^{-8}$ sec. In addition, the following values for the alignment depolarization cross sections due to collisions between ground-state and excited-state atoms have been obtained: $\bar{\sigma}_2(^3D_3) = 2.14(19) \times 10^{-14}$ cm^2 at $T = 292(5)$°K, $\bar{\sigma}_2(^1D_2') = 1.14(22) \times 10^{-14}$ cm^2 at $T = 292(5)$°K, $\bar{\sigma}_2(^3P) = 5.63(21) \times 10^{-15}$ cm^2 at $T = 303(5)$°K, and $\bar{\sigma}_2(^3P) = 5.84(33) \times 10^{-15}$ cm^2 at $T =$ liquid-nitrogen temperature.

I. INTRODUCTION

In this paper, I report on measurements of the radiative lifetimes of the $(3p^54p)\,^3D_3$ and $^1D_2'$ levels in Ar ($2p_9$ and $2p_6$, respectively, in Paschen's notation) and the $(1s2p)\,^3P$ term in He.[1] The measurements were made by means of the zero-field level-crossing method (Hanle effect).[2] In addition to the free-atom lifetimes, values of the alignment depolarization cross sections due to collisions between ground-state and excited-state atoms[3] are presented. Measurements such as these are of particular value in the interpretation of data in such fields as gaseous discharges, astrophysics, and atomic scattering.

With the zero-field level-crossing method, one observes the scattering of resonance radiation by an assembly of atoms as a function of an applied magnetic field. This field is swept through zero, resulting thereby in the crossing of the Zeeman substates of each atomic level. The form of the observed signal depends on the relative directions and polarizations of the incident and reemitted light, the direction of the applied magnetic field, and the coherence time (apparent lifetime) and g factor of the excited level,[4,5] as described in Sec. III.

In this work, the metastable levels $(3p^54s)\,^3P_2$ in Ar ($1s_5$ in Paschen's notation) and $(1s2s)\,^3S_1$ in He served as effective ground states from which the levels of interest (see Figs. 1 and 2) were excited optically. The feasibility of using the $^3S_1$ He level in this manner was demonstrated by Colegrove and Franken[6] and, as in their experiments, the metastable levels were populated by running a discharge in the scattering cell. Since radiative transitions from these levels to the ground states are forbidden, a substantial density of metastable atoms can be obtained.