## Chapter 4|The Coupled-Cluster

## A. INTRODUCTION

The CI and MCSCF methods described earlier suffer from one significant weakness. The slow convergence of the wavefunction as the configuration size is increased is a problem that becomes more severe as the number of electrons in the system grows. In fact, for extended systems the finite CI or MCSCF wavefunctions (because they contain only a finite number of electron pair interactions) become infinitesimal portions of the exact wavefunction. Perturbation theory methods, whose wavefunction usually also contains only finite numbers of interactions, sometimes provide some relief because the total energy is not calculated as an expectation value. However, it is often not appropriate to assume that the usual fluctuation potential (true electron-electron interaction minus the HF potential) is small, i.e., to assume convergence of the perturbation series. Moreover, it is quite often important to be able to properly treat systems that are not adequately described by a single-configuration zeroth-order reference wavefunction (such as is assumed in most perturbation theories).

## B. FORM OF THE WAVEFUNCTION

The coupled-cluster (CC) method (Cizek and Paldus, 1971; Harris, 1977a,b; Bartlett and Purvis, 1978) is an attempt to introduce interactions among electrons within clusters (predominantly pairs) as well as coupling among these clusters of electrons and to permit the wavefunction to contain all possible disjoint clusters. For example, we know, from the early work of Sinanoglu (1962) and others, that electron pair interactions are of utmost importance and that contributions of quadruply excited configurations to
$|0\rangle$ arise predominantly as products of doubly excited configurations. CC wavefunctions in which such electron pair interactions (clusters) are assumed to dominate still contain terms that describe disjoint products of electron pair clusters just as Sinanoglu's observations would suggest. In fact, for a system containing an even (odd) number of electrons $N(N+1)$, one has products of $2,3, \ldots, N / 2$ disjoint pair clusters in the CC wavefunction. The mechanism for introducing these cluster interactions is to write the wavefunction $|0\rangle$ in terms of a so-called cluster operator $T$ acting on a reference function describing noninteracting or noncoupled electrons $\left|0^{\circ}\right\rangle$ :

$$
\begin{equation*}
|0\rangle=\exp (T)\left|0^{0}\right\rangle \tag{4.1}
\end{equation*}
$$

The reference function $\left|0^{\circ}\right\rangle$ has, in nearly all CC developments to date, been limited to a ket corresponding to a single Slater determinant. In the treatment given in this chapter, we therefore restrict our attention to this single determinantal case. The cluster operator $T$ generates one-, two-electron, etc., clusters

$$
\begin{equation*}
T=T_{1}+T_{2}+\cdots+T_{N} \tag{4.2}
\end{equation*}
$$

with

$$
\begin{align*}
& T_{1}=\sum_{\alpha, r} t_{\alpha}^{r} r^{+} \alpha  \tag{4.3}\\
& T_{2}=\frac{1}{4} \sum_{\substack{\alpha \beta \\
r s}} t_{\alpha \beta}^{r s} r^{+} s^{+} \beta \alpha \tag{4.4}
\end{align*}
$$

etc. (the greek indices $\alpha, \beta, \gamma, \ldots$ denote spin-orbitals occupied in $\left|0^{\circ}\right\rangle$; $r, s, t, u, \ldots$ denote unoccupied spin-orbitals). To make some connection between the CC wavefunction of Eq. (4.1) and the more conventional CI and MBPT expressions for $|0\rangle$, we expand the $\exp (T)\left|0^{\circ}\right\rangle$ and collect terms of common excitation level:

$$
\begin{align*}
\exp (T)\left|0^{0}\right\rangle= & \left(1+T_{1}+T_{2}+\frac{1}{2!} T_{1}^{2}+T_{3}+\frac{1}{3!} T_{1}^{3}+T_{1} T_{2}+T_{4}\right. \\
& \left.+\frac{1}{4!} T_{1}^{4}+\frac{1}{2!} T_{2}^{2}+T_{3} T_{1}+\frac{1}{2!} T_{1}^{2} T_{2}+\cdots\right)\left|0^{0}\right\rangle \tag{4.5}
\end{align*}
$$

By grouping the terms of a given excitation level together, we see that the CC wavefunction can be rewritten as

$$
\begin{equation*}
\exp (T)\left|0^{0}\right\rangle=\left(1+C_{1}+C_{2}+C_{3}+\cdots\right)\left|0^{0}\right\rangle \tag{4.6}
\end{equation*}
$$

where the configuration excitation operators $C_{1}, C_{2}, \ldots$ are

$$
\begin{align*}
& C_{1}=T_{1}  \tag{4.7}\\
& C_{2}=T_{2}+\frac{1}{2!} T_{1}^{2}  \tag{4.8}\\
& C_{3}=T_{3}+\frac{1}{3!} T_{1}^{3}+T_{1} T_{2}  \tag{4.9}\\
& C_{4}=T_{4}+\frac{1}{4!} T_{1}^{4}+\frac{1}{2!} T_{2}^{2}+T_{3} T_{1}+\frac{1}{2!} T_{1}^{2} T_{2} \tag{4.10}
\end{align*}
$$

etc. We thus see, for example, that the quadruple excitations that would be obtained in a CI or MCSCF treatment can be viewed within the cluster framework as consisting of five separate parts. The $T_{2}^{2}$ component is thought to be the dominant term because it represents the simultaneous interactions of two distinct pairs of electrons (e.g., electron pairs that occupy spatially different molecular orbitals). The $T_{4}$ term is usually expected to be quite small since it describes the simultaneous interaction of four electrons. The single-cluster contributions to $C_{4}, C_{3}$, and $C_{2}$ can be made small by using - MCSCF orbitals.

If our normal description of chemical bonding in terms of electron pair bonds is correct, it is likely that a description of molecular structure in which $T_{2}$ is treated to high order [e.g., through $\exp \left(T_{2}\right)$ ] while $T_{1}, T_{3}, T_{4}$, etc. are either neglected or treated less rigorously, is quite accurate. For this reason we consider developing systematic procedures for truncating the expansion of $T$ given in Eq. (4.2). By truncating [approximating the cluster operator $T$ to some low-order (say pair clusters $T_{2}$ )], the resultant wavefunction contains not only these low-order clusters $T_{2}\left|0^{0}\right\rangle$ but also disjoint clusters [e.g., $(1 / 2!) T_{2} T_{2}\left|0^{0}\right\rangle,(1 / 3!) T_{2} T_{2} T_{2}\left|0^{0}\right\rangle$, etc.] that involve more highly excited configurations than are present in $T_{2}\left|0^{\circ}\right\rangle$. Of course, these higher-order excitations [e.g., quadruply excited for $(1 / 2!) T_{2} T_{2}\left|0^{\circ}\right\rangle$ ] are present in $|0\rangle$ only to the extent that their amplitudes can be described in terms of products of the amplitudes belonging to the smaller clusters (e.g., $\left.t_{\alpha \beta}^{r s} t_{y \delta}^{t u}\right)$. The fact that the product factors $T_{2} T_{2}\left|0^{0}\right\rangle$ contain only disjoint clusters arises because the operator product $\left(r^{+} s^{+} \cdots \alpha \beta \cdots\right)\left(t^{+} u^{+} \cdots \gamma \delta \cdots\right)$ vanishes if any of the hole ( $\alpha, \beta, \gamma, \ldots$ ) or particle ( $r^{+}, s^{+}, t^{+}, \ldots$ ) indices are equal. An essential point of the CC approach is that even low-order truncations of $T$ (which are usually based upon the physical assumption that electron pair interactions dominate) lead to a wavefunction that contains all of the disjoint higher excitations needed to make the resultant energy (and other physical properties) size consistent.

## C. EQUATIONS FOR THE CLUSTER AMPLITUDES

The cluster amplitudes $t_{\alpha \beta \ldots}^{r s} \ldots$ are determined by insisting that $\exp (T)\left|0^{0}\right\rangle$ satisfy the usual Schrödinger equation (or at least certain projections of this equation)

$$
\begin{equation*}
H \exp (T)\left|0^{0}\right\rangle=E \exp (T)\left|0^{0}\right\rangle \tag{4.11}
\end{equation*}
$$

which upon premultiplying by $\exp (-T)$ gives

$$
\begin{equation*}
\exp (-T) H \exp (T)\left|0^{0}\right\rangle=E\left|0^{\circ}\right\rangle \tag{4.12}
\end{equation*}
$$

The above exponential series gives, when expanded and collected together as commutators,

$$
\begin{align*}
(H & +[H, T]+\frac{1}{2!}[[H, T], T]+\frac{1}{3!}[[[H, T], T], T] \\
& \left.+\frac{1}{4!}[[[[H, T], T], T], T]\right)\left|0^{\circ}\right\rangle=E\left|0^{\circ}\right\rangle \tag{4.13}
\end{align*}
$$

The series truncates (exactly) after four commutators regardless of the level at which $\left(T_{n}\right) T$ is truncated (if at all). This exact truncation is a result of the fact that $H$ contains at most two-electron operators, which involve four general (particle or hole) operators $i^{+} j^{+} l k$. Therefore $[H, T]$ contains at most three general operators, $[[H, T], T]$ contains two, and $[[[[H, T], T], T], T]$ thus contains only (excitation) operators of the form $r^{+} s^{+} \cdots \alpha \beta \cdots$. These excitation operators clearly commute with $T$; thus the next (fifth) commutator in the series vanishes. The CC expression of the Schrödinger equation hence yields a quartic equation for the cluster amplitudes $\left(t_{\alpha \beta}^{r s} \ldots\right)$ appearing in $T$.

A closed set of equations for the desired amplitudes is obtained by insisting that the final Schrödinger equation [Eq. (4.13)], when projected against a set of low-order excitations out of $\left|0^{0}\right\rangle$, yield zero. The particular excitations are usually chosen to include up through $n$-fold excitations from $\left|0^{\circ}\right\rangle$ in the case where $T$ has been truncated at $T_{n}$. The resultant set of algebraic equations will then be equal in number to the number of amplitudes $t_{\alpha \beta}^{r s} \ldots$ in $T$. Once these amplitudes are obtained by solving the resultant nonlinear equations, the total electronic energy is computed by projecting Eq. (4.13) onto $\left|0^{\circ}\right\rangle$. We should stress that the energy expression thereby obtained is not variational in the sense that it is not given as an expectation value of the Hamiltonian. The quantity

$$
\left\langle 0^{0}\right| \exp \left(T^{+}\right) H \exp (T)\left|0^{0}\right\rangle /\left\langle 0^{0}\right| \exp \left(T^{+}\right) \exp (T)\left|0^{0}\right\rangle
$$

would give rise to a variational energy expression but the resultant commutator expansion of the exponential operators would not truncate because $T^{+}$ contains operators of the form $\alpha^{+} \beta^{+} \cdots r s \cdots$, which do not commute with $T$. It is the presence of $\exp (-T)$, rather than $\exp \left(T^{+}\right)$, in Eq. (4.12) that gives rise to the exactly closed quartic equation for $T$. Moreover, the presence of the commutators in the expression for $E$ and the fact that $T$ contains only particle creation and hole annihilation operators makes the CC-calculated energy contain only linked terms (in the sense discussed in Chapter 3). This then makes $E$ contain only size-consistent terms.

## D. HARTREE-FOCK ORBITALS AND $\boldsymbol{T} \simeq \boldsymbol{T}_{\mathbf{2}}$

Most CC calculations carried out so far have used the approximation $T \simeq T_{2}$. In this section, we treat this model in some detail since doing so will give us more insight into the structure of the CC equations. The physical motivation for approximating $T \simeq T_{2}$ relies on the fact that if the set of HF orbitals are used, the BT suggests that single excitation $T_{1}$ operators, which largely serve to optimize the spin-orbitals, should be less important than $T_{2}$. It is, however, now commonly felt that one should include both $T_{1}$ and $T_{2}$ so as to obtain a balanced or coupled description of the orbital and electron pair cluster optimization. Let us, however, continue our analysis of the $T \simeq T_{2}$ case.

To see what the solution of the above discussed nonlinear equations actually involves, let us examine these expressions in more detail for a case in which the spin-orbitals $\left\{\phi_{a} \phi_{\beta}, \ldots, \phi_{r}, \ldots, \phi_{s}\right\}$ are eigenfunctions of a HF operator having orbital energies $\left\{\varepsilon_{\alpha} \cdots \varepsilon_{r} \cdots\right\}$. The decomposition of the Hamiltonian $H$ into $H^{0}+U$ is then given as in Eq. (2.84) by

$$
\begin{equation*}
H=H^{0}+W-V_{\mathbf{H F}} \tag{4.14}
\end{equation*}
$$

where $H^{0}$ is the HF Hamiltonian

$$
\begin{equation*}
H^{0}=\sum_{i} \varepsilon_{i} i^{+} i \tag{4.15}
\end{equation*}
$$

$W$ is the full electron interaction term in Eq. (2.84) and $V_{\mathrm{HF}}$ in the HF potential [Eq. (2.91)].

The commutator expansion of $\exp (-T) H \exp (T)$ in Eq. (4.12) given in Eq. (4.13) demonstrates in an elegant manner that when Eq. (4.13) is projected against low-order excitations $\left\langle\begin{array}{c}\left\langle_{\alpha \beta}^{r s} \cdots n\right. \\ s_{\gamma}\end{array}\right|=\left\langle 0^{0}\right| \gamma \cdots \beta \alpha r^{+} s^{+} \cdots n^{+}$, it gives equations that are at most quartic in the cluster amplitudes $t_{\alpha \beta}^{r s \cdots u}$. However, it turns out that for finding equations for $t_{\alpha \beta}^{r s}$ it is equally simple to expand the exponential operators directly. To determine the total energy $E$, we project

Eq. (4.12) against $\left|0^{0}\right\rangle$ to obtain

$$
\begin{align*}
E & =\left\langle 0^{0}\right| \exp \left(-T_{2}\right) H \exp \left(T_{2}\right)\left|0^{0}\right\rangle \\
& =\left\langle 0^{\circ}\right|\left(1-T_{2}+\frac{1}{2}\left(T_{2}\right)^{2}\right) H\left(1+T_{2}+\frac{1}{2}\left(T_{2}\right)^{2} \cdots\right)\left|0^{0}\right\rangle \\
& =\left\langle 0^{0}\right| H\left|0^{0}\right\rangle+\left\langle 0^{0}\right| H T_{2}\left|0^{0}\right\rangle \\
& \left.=E_{\mathrm{HF}}+\sum_{\substack{m>n \\
\alpha>\beta}}\langle\beta \alpha||m n\rangle\right\rangle_{\alpha \beta}^{m n} \tag{4.16}
\end{align*}
$$

where we have used the fact that $\left\langle 0^{0}\right| T_{2}=0$ because of the appearance of $r^{+} s^{+} \beta \alpha$ in $T_{2}$. We have also used the fact that $\left\langle 0^{0}\right| H T_{2} T_{2}\left|0^{0}\right\rangle$ vanishes because $T_{2}^{2}\left|0^{0}\right\rangle$ is quadruply excited and hence cannot couple through $H$ to $\left\langle 0^{\circ}\right|$. The $t_{x \beta}^{m n}$ amplitudes are determined by projecting Eq. (4.12) against doubly excited kets $\left\langle{ }_{\alpha \beta}^{m n}\right|$ to obtain

$$
\begin{equation*}
0=\left\langle{ }_{\alpha \beta}^{m}\right| \exp \left(-T_{2}\right) H \exp \left(T_{2}\right)\left|0^{0}\right\rangle \tag{4.17}
\end{equation*}
$$

Expanding the exponential then allows one to see that the only nonvanishing contributions are contained in

$$
\begin{equation*}
0=\left\langle{ }_{\alpha \beta}^{m n}\right| H\left(1+T_{2}+\frac{1}{2} T_{2}^{2}\right)\left|0^{0}\right\rangle+\left\langle{ }_{\alpha \beta}^{m n}\right|\left(-T_{2}\right) H\left(1+T_{2}\right)\left|0^{0}\right\rangle \tag{4.18}
\end{equation*}
$$

which shows that we obtain only a quadratic equation for the cluster amplitudes when $T \simeq T_{2}$. Explicitly evaluating the matrix element appearing in Eq. (4.18) then leads to the following nonlinear equation for the cluster amplitudes:

$$
\begin{align*}
\left(\varepsilon_{m}+\right. & \left.\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right) t_{\alpha \beta}^{m m} \\
= & \langle m n \|||\alpha\rangle\rangle-\sum_{p>q}\langle m\| \| p q\rangle t_{\alpha \beta}^{p q}-\sum_{\gamma>\delta}\langle\gamma \delta \| \mid \alpha \beta\rangle t_{\gamma \delta}^{m n} \\
& +\sum_{\gamma p}\left(\langle\gamma n \| \mid \beta p\rangle t_{\alpha \gamma}^{m p}-\langle\gamma m \|||\beta p\rangle t_{\alpha \gamma}^{n p}-\langle\gamma n \| \alpha p\rangle t_{\beta \gamma}^{m p}+\langle\gamma m \|||\alpha p\rangle t_{\beta \gamma}^{n p}\right) \\
& +\sum_{\substack{\gamma>\delta \\
p>q}}\langle\gamma \delta \| p q\rangle\left[t_{\alpha \beta}^{p q} t_{\gamma \delta}^{m n}-2\left(t_{\alpha \beta}^{m p} t_{\gamma}^{n q}+t_{\alpha \beta}^{n q} t_{\gamma \delta}^{m p}\right)\right. \\
& \left.-2\left(t_{\alpha \gamma}^{m n} t_{\beta \delta}^{p q}+t_{\alpha \gamma}^{p q} t_{\beta \delta}^{m n}\right)+4\left(t_{\alpha \gamma}^{m p} t_{\beta \delta}^{n q}+t_{\alpha \gamma}^{m q} t_{\beta \delta}^{m p}\right)\right] \tag{4.19}
\end{align*}
$$

In the next sections we describe how solutions may be obtained to Eq. (4.19) and we discuss the relationship of the solution thereby obtained to results of MBPT.

## E. PERTURBATIVE SOLUTION TO THE COUPLED-CLUSTER EQUATIONS

We describe here how Eq. (4.19) may be solved in a manner that shows the connection between the CC and the MBPT approaches. We solve

Eq. (4.19) in an iterative manner by performing an initial guess of $\left\{t_{\alpha \beta}^{r s}\right\}$ and then inserting this value on the right-hand side into Eq. (4.19) to obtain an improved set of values of $\left\{r_{\alpha \beta}^{r s}\right\}$. These are then inserted back into the righthand side of Eq. (4.19) to again give us improved values of $\left\{t_{\alpha \beta}^{r s}\right\}$, etc. As an initial guess of the cluster amplitudes we set those amplitudes that appear on the right-hand side of Eq. (4.19) equal to zero. The motivation for this choice is that the terms containing $t$ on the right-hand side of Eq. (4.19) are assumed to be smaller than those on the left-hand side of this equation. We then obtain the following expression for the amplitudes:

$$
\begin{equation*}
t_{\alpha \beta}^{m n}=\langle m n||\alpha \beta\rangle\left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)^{-1} \tag{4.20}
\end{equation*}
$$

Inserting this value of $t_{\alpha \beta}^{m n}$ into the CC expression for the total energy as given in Eq. (4.16) yields

$$
\begin{equation*}
E=E_{\mathrm{HF}}+\sum_{\substack{m>n \\ \alpha>\beta}}\langle\beta \alpha||m n\rangle\langle m n||\alpha \beta\rangle\left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)^{-1} \tag{4.21}
\end{equation*}
$$

This expression is nothing but the result obtained in second-order perturbation theory, which is written explicitly in Eq. (3.38).

A second iteration may be carried out by inserting into the right-hand side of Eq. (4.19) the cluster amplitudes obtained above. If we then neglect the terms that are quadratic in the $t_{\alpha \beta}^{m n}$ amplitudes [the eighth through fourteenth terms on the right-hand side of Eq. (4.19)], we obtain cluster amplitudes that, when used to compute the energy $E$ via Eq. (4.16) give the same algebraic expression as is obtained in third-order MBPT (see Problem 4.1, question 1). If these cluster amplitudes are then inserted into the right-hand side of Eq. (4.19) (keeping the quadratic terms this time), we obtain new amplitudes that, when used to compute $E$, give all contributions to the fourth-order MBPT energy that arise from quadruple excitations [ $C_{4}$ in Eq. (4.10)]. From the form of our working equation, Eq. (4.19), it is further clear that the quadruple excitations obtained in this way can only arise from the $\frac{1}{2}\left(T_{2}\right)^{2}$ and $-T_{2} T_{2}$ terms. These terms, in a sense, correspond to two simultaneous interactions of two electrons (electron pair interaction). The $T_{4}$ term, which corresponds to a true four-body interaction, first enters at fifth order in perturbation theory, thus indicating that electron pair interactions are much more important than true four-body interaction (Sinanoğlu, 1962). All fourth-order energy diagrams can, of course, not be obtained by approximating $T$ with $T_{2}$, since both single and triple excitations contribute in fourth order. To obtain all fourth-order diagrams in a CC calculation would require both $T_{1}$ and $T_{3}$ to be included in the cluster expansion.

The iterative process carried out when determining the cluster amplitudes from Eq. (4.19) may be continued by inserting the cluster amplitudes from one iteration into the right-hand side of Eq. (4.19) to obtain the new ampli-
tudes. The converged result would then correspond to summing all diagrams that contain double and products of double excitations through infinite order.

## F. NEWTON-RAPHSON METHOD

Clearly, either the equations obtained by taking $T \simeq T_{2}$ [Eq. (4.19)] or the general quartic equation obtained from Eq. (4.13) are nonlinear and multivariable. Such equations can be represented in matrix form (by defining $t_{\alpha \beta}^{r s}$ as the $r s, \alpha \beta$ element of the $t$ column vector) as

$$
\begin{equation*}
\mathbf{0}=\mathbf{a}+\mathbf{b} \mathbf{t}+c \mathbf{t} \tag{4.22}
\end{equation*}
$$

where, for example $a_{r s, \alpha \beta}=\langle r s||\alpha \beta\rangle$ [see Eq. (4.19)]. The solution of these nonlinear algebraic equations represents a substantial practical difficulty in implementing the CC method. To solve these equations one can employ the perturbative analysis described above. This technique has the advantages that it is straightforward to program on a computer and that it has a close connection with MBPT.

An alternative to the above described perturbative procedure is the multivariable Newton-Raphson method. Such methods were used in the first molecular CC calculations (Paldus et al., 1972). Here, one attempts to choose $t$ such that the vector $f(t)$ defined as

$$
\begin{equation*}
f(t) \equiv \mathbf{a}+b \mathbf{t}+c \mathbf{t t} \tag{4.23}
\end{equation*}
$$

becomes equal to zero. This is done by expanding $f(t)$ about the "point" $\mathbf{t}_{0}$. Keeping only linear terms in this Taylor expansion and setting $f(t)$ equal to zero, one obtains equations for the changes $\Delta t$ in the $t$ amplitudes, which can be expressed as

$$
\begin{equation*}
f_{\alpha \beta}^{r s}(\mathbf{t})=0=f_{\alpha \beta}^{r s}\left(\mathbf{t}_{0}\right)+\sum_{\substack{u w \\ \gamma \delta}}\left(\frac{\partial f_{\alpha \beta}^{r s}}{\partial t_{y \delta}^{u s}}\right)_{\mathbf{t}_{0}} \Delta t_{\gamma \delta}^{u w} \tag{4.24}
\end{equation*}
$$

The step lengths (corrections to $\mathbf{t}_{0}$ ) can be obtained by solving the above set of linear equations and then used to update the $t$ amplitudes

$$
\begin{equation*}
\mathbf{t}=\mathbf{t}_{0}+\Delta \mathbf{t} . \tag{4.25}
\end{equation*}
$$

These values of $\mathbf{t}$ can then be used as a new $\mathbf{t}_{0}$ vector for the next application of Eq. (4.24). This multidimensional Newton-Raphson procedure, which involves the solution of a large number of coupled linear equations, is then repeated until the $\Delta t$ values are sufficiently small (convergence). Given the set of $t_{\alpha \beta}^{r s}$ amplitudes, Eq. (4.16) can then be used to compute E. Although the first applications of the coupled cluster method to quantum chemistry did employ this Newton-Raphson scheme, the numerical problems involved
in solving the large multivariable inhomogenous equations [Eq. (4.24)] has led more recent workers to use the perturbative techniques discussed in Section E. Within the perturbative framework, more sophisticated methods have been developed to solve the large number of quadratic (for $T \simeq T_{2}$ ) equations that arise. One such device is based upon the so-called reduced linear equations technique, which has also been widely used to find selected eigenvalues and eigenvectors of large CI matrices (Davidson, 1975).

## G. SUMMARY

Although the CC method possesses several advantages over CI and MCSCF approaches, the fact that the resultant set of CC equations that determine the $t_{\alpha \beta}^{r s} \cdots$ amplitudes are nonlinear and of very large dimensions even for modest-sized systems, has made the practical applications of this theory rather limited. An analysis of the relationships between the solutions of the nonlinear CC equations and the solutions of corresponding CI secular problems has recently been provided (Monkhorst and Zivkovic, 1978). This analysis thus provides some reason for optimism concerning the possibility of finding efficient mechanisms for solving the CC equations. However, at present, the nonlinear nature of the equations to be solved still makes the practical utilization of the CC method something toward which we are still working. Research aimed at achieving efficient solutions of the quadratic (or even quartic) coupled equations and at extending the CC development to open-shell and multiconfigurational reference states is necessary if the CC method is to become widely used in quantum chemistry.

## PROBLEMS

4.1 Perform a CC calculation where $T$ is approximated with $T_{2}$.

1. Show that the CC equations may be iterated to yield cluster amplitudes that, when used in the energy expression, give the MBPT third-order energy expression [see discussions following Eqs. (4.20) and (4.21)]. The third-order MBPT energy expression is given in Problem 3.2.

Carry out a CC calculation on $\mathrm{HeH}^{+}$using the minimum basis HF results found in Problem 2.1. In performing this calculation follow the steps given below.
2. First use a linear (truncated) form of the CC equation to determine the numerical values of the $t_{\alpha \beta}^{m n}$ parameters and then use these parameters to compute the corresponding correlation energy.
3. Argue that although the above linear form (question 2) of the CC equation and the perturbative solution (question 1) yield amplitudes that
have the same formal structure, the correlation energy of $\mathrm{HeH}^{+}$determined in question 2 and in third-order MBPT (question 1 and Problem 3.2) differ.
4. Use the quadratic form of the CC equations (which clearly has two solutions) to determine the values of the two sets of $t_{\alpha \beta}^{m n}$ parameters.
5. Evaluate the total energy and the correlation energy contribution for both of these two sets of solutions.
6. Show that the CC equations in Eqs. (4.16) and (4.18) and the CI eigenvalue equation that contains doubly excited states become identical for a two-basis-function two-electron problem. Why do the two configuration CI total energies of Problem 2.4 and the CC total energies of question 5 differ?

Consider now $n \mathrm{HeH}^{+}$molecules that are separated at infinite distance with each molecule described by the localized SCF orbitals of Problem 2.1.
7. Show by carrying out a perturbative solution to the CC equations as described in Section E that the correlation energy for the $n \mathrm{HeH}^{+}$molecules becomes identical to $n$ times the correlation energy of a single $\mathrm{HeH}^{+}$ molecule and that the CC model thus is size consistent.

## SOLUTIONS

## 4.1

1. When the CC amplitudes on the right-hand side of Eq. (4.19) are set equal to zero, we get

$$
t_{\alpha \beta}^{m n}=\langle m n \| \alpha \beta\rangle\left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)^{-1}
$$

Inserting this value of $t_{\alpha \beta}^{m n}$ on the right-hand side of Eq. (4.19) gives the next approximation to $t_{\alpha \beta}^{m n}$ :

$$
\begin{aligned}
t_{\alpha \beta}^{m n}= & \left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)^{-1}\left[\langle m n \| \alpha \beta\rangle-\sum_{p>q}\langle m n \| \mid p q\rangle t_{\alpha \beta}^{p q}-\sum_{\gamma>\delta}\langle\gamma \delta \| \alpha \beta\rangle t_{\gamma \delta}^{m n}\right. \\
& \left.\left.\left.\left.+\sum_{\gamma p}\left(\langle\gamma n \| \mid \beta p\rangle t_{\alpha \gamma}^{m p}-\langle\gamma m \| \mid \beta p\rangle\right\rangle_{\alpha \gamma}^{n p}-\langle\gamma \eta \| \alpha p\rangle\right\rangle_{\beta \gamma}^{m p}+\langle\gamma m\| \| \alpha p\rangle\right\rangle_{\beta \gamma}^{n p}\right)\right]
\end{aligned}
$$

The first term in the square brackets results in the second-order energy expression [Eq. (4.21)] when used in Eq. (4.16).

The second term in the square brackets gives, when inserted into Eq. (4.16), the correlation energy contribution

$$
-\sum_{\substack{m>n \\ \beta>\alpha \\ p>q}} \frac{\langle\beta \alpha||m n\rangle\langle m n \| \mid p q\rangle\langle p q \| \alpha \beta\rangle}{\left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)\left(\varepsilon_{p}+\varepsilon_{q}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)}
$$

which is identical to diagram A in Fig. 3.7.

The third term gives

$$
-\sum_{\substack{m>n \\ \alpha \gg \\ \gamma>\delta}} \frac{\langle\beta \alpha \| m n\rangle\langle\gamma \delta||\alpha \beta\rangle\langle m n \| \gamma \delta\rangle}{\left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)\left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\gamma}-\varepsilon_{\delta}\right)}
$$

which is identical to diagram B of Fig. 3.7.
The last four terms give

$$
\begin{aligned}
\sum_{\alpha>n}^{m>\beta} & \frac{\langle\beta \alpha||m n\rangle}{\left(\varepsilon_{m}+\varepsilon_{n}-\varepsilon_{\alpha}-\varepsilon_{\beta}\right)} \sum_{\gamma p}\left(\langle\gamma n \| \mid \beta p\rangle t_{\alpha \gamma}^{m p}-\langle\gamma m \| \beta p\rangle t_{\alpha \gamma}^{n p}\right. \\
& \left.-\langle\gamma n \| \alpha p\rangle t_{\beta \gamma}^{m p}+\langle\gamma m \| \alpha p\rangle t_{\beta \gamma}^{n p}\right)
\end{aligned}
$$

Substitution of variables [e.g., in the first term we substitute ( $m \rightarrow p, \alpha \rightarrow \gamma$, $\gamma \rightarrow \beta, n \rightarrow m, \beta \rightarrow \alpha, p \rightarrow n)$ ] allows these four terms to be rewritten as

$$
\begin{aligned}
& +\sum_{\substack{p>m \\
\gamma>\alpha \\
\beta n}} \frac{\langle\alpha \gamma||p m\rangle\langle\beta m \| \alpha n\rangle\langle p n \| \gamma \beta\rangle}{\left(\varepsilon_{p}+\varepsilon_{m}-\varepsilon_{\alpha}-\varepsilon_{\gamma}\right)\left(\varepsilon_{p}+\varepsilon_{n}-\varepsilon_{\gamma}-\varepsilon_{\beta}\right)} \\
& -\sum_{\substack{m>p \\
\gamma>\alpha \\
n \beta}} \frac{\langle\alpha \gamma||m p\rangle\langle\beta m \| \alpha n\rangle\langle p n \| \gamma \beta\rangle}{\left(\varepsilon_{m}+\varepsilon_{p}-\varepsilon_{\alpha}-\varepsilon_{\gamma}\right)\left(\varepsilon_{p}+\varepsilon_{n}-\varepsilon_{\gamma}-\varepsilon_{\beta}\right)} \\
& -\sum_{\substack{p>m \\
\alpha>\gamma \\
n \beta}} \frac{\langle\gamma \alpha \| p m\rangle\langle\beta m \| \mid \alpha n\rangle\langle p n \| \gamma \beta\rangle}{\left(\varepsilon_{p}+\varepsilon_{m}-\varepsilon_{\gamma}-\varepsilon_{\alpha}\right)\left(\varepsilon_{p}+\varepsilon_{n}-\varepsilon_{\gamma}-\varepsilon_{\beta}\right)} \\
& +\sum_{\substack{m>p \\
\alpha>\gamma \\
n \beta}} \frac{\langle\gamma \alpha||m p\rangle\langle\beta m \| \alpha n\rangle\langle p n \| \gamma \beta\rangle}{\left(\varepsilon_{m}+\varepsilon_{p}-\varepsilon_{\gamma}-\varepsilon_{\alpha}\right)\left(\varepsilon_{p}+\varepsilon_{n}-\varepsilon_{\gamma}-\varepsilon_{\beta}\right)}
\end{aligned}
$$

The above four terms when collected together give diagram C of Fig. 3.7. Hence, all second- and third-order diagrams have been accountered for.
2. The only nonvanishing cluster amplitude is $t_{1 \alpha \beta \beta}^{2 \alpha 2 \beta}$. Equation (4.19) gives

$$
0=\langle 22 \mid 11\rangle+t_{1 \alpha 1 \beta}^{2 \alpha 2 \beta}\left(2 \varepsilon_{1}-2 \varepsilon_{2}-\langle 22 \mid 22\rangle-\langle 11 \mid 11\rangle+4\langle 12 \mid 12\rangle-2\langle 12 \mid 21\rangle\right)
$$

which gives $t_{1 \alpha 1 \beta}^{2 \alpha 2 \beta}=0.0559$. Inserting this value in Eq. (4.16) gives the correlation energy contribution

$$
\Delta E_{\text {corr }}=-0.0070 \quad \text { a.u. }
$$

Notice that although $t_{1 \alpha 1 \beta}^{2 \alpha 2 \beta}$ is positive, the correlation energy of Eq. (4.16) is negative, because $\langle\beta \alpha||m n\rangle=-\langle 11 \mid 22\rangle$.
3. Both the third-order MBPT and the approximation in question 2 use a linear form of the CC equations. However in the MBPT solution (question 1) the cluster amplitudes used in Eq. (4.19) are determined from perturbation theory, which results in a second- and third-order correlation energy of -0.0066 (Solution 3.2, question 5). The nonperturbation solution of question 2 is different from the one obtained in question 1 and gives a correlation energy -0.0070 .
4. The quadratic CC equation reads

$$
\begin{aligned}
0= & \langle 22 \mid 11\rangle+t_{1 \alpha \beta}^{2 \alpha 2 \beta}\left(2 \varepsilon_{1}-2 \varepsilon_{2}-\langle 22 \mid 22\rangle-\langle 11 \mid 11\rangle\right. \\
& \left.\left.+4\langle 12 \mid 12\rangle-2\langle 12 \mid 21\rangle)+\langle 11 \mid 22\rangle t_{1 \alpha 1 \beta}^{2 \alpha 2}\right)_{1 \alpha 1}^{2021}\right)
\end{aligned}
$$

which gives

$$
t_{1 \alpha 1 \beta}^{2 \alpha 2 \beta}=0.0560, \quad t_{1 \alpha 1 \beta}^{2 \alpha 2 \beta}=-17.8432
$$

5. $t_{12 \alpha \beta}^{2 \alpha 2 \beta}=0.0560, \quad E=-4.2791, \quad E_{\text {corr }}=-0.0071$

$$
t_{1 \alpha 1 \beta}^{2 \alpha 2 \beta}=-17.8432, \quad E=-2.0220, \quad E_{\text {corr }}=2.2500
$$

6. The CC (Schrödinger) Equation (4.12) contains only linear terms in $T_{2}$ when applied to a two-electron system:

$$
\left(1-T_{2}\right) H\left(1+T_{2}\right)\left|0^{0}\right\rangle=E\left|0^{0}\right\rangle
$$

When this equation is projected against $\left\langle 0^{\circ}\right|$ and $\left\langle\begin{array}{l}2 \alpha 2 \beta \\ 1 \alpha 1 \beta\end{array}\right|=\left\langle{ }_{11}^{22}\right|$ one obtains

$$
\begin{array}{r}
\left.\left\langle 0^{0}\right| H\left|0^{0}\right\rangle+\left.\langle 0| H\right|_{11} ^{22}\right\rangle t=E \\
\left.\left.\left\langle{ }_{11}^{22}\right| H\left|0^{0}\right\rangle+\left.\left\langle{ }_{11}^{22}\right| H\right|_{11} ^{22}\right\rangle t-t\left\langle 0^{0}\right| H\left|0^{0}\right\rangle-\left.t^{2}\langle 0| H\right|_{11} ^{22}\right\rangle=0 \tag{B}
\end{array}
$$

where $t_{1 a 1 \beta}^{2 \alpha 2 \beta}$ is denoted $t$. Substituting Eq. (A) into Eq. (B) gives

$$
\begin{equation*}
\left.\left\langle\left\langle_{11}^{22}\right| H \mid 0^{0}\right\rangle+\left.\left\langle_{11}^{22}\right| H\right|_{11} ^{22}\right\rangle t=E t \tag{C}
\end{equation*}
$$

Equations $(\mathrm{A})$ and $(\mathrm{C})$ are nothing but the CI eigenvalue problem written out in component form for an intermediate normalized eigenvector with components $(1, t)$. The CC total energies $(-4.2791,-2.0220)$ and the CI total energies $(-4.2790,-2.0079)$ differ only because of numerical errors caused by using four significant digits in the integrals.
7. When the cluster amplitudes in the nonlinear part of Eq. (4.19) are set equal to zero (as in the first step of the perturbative solution), the only nonvanishing cluster amplitudes that remain (see the solution to question 1) are those involving all four orbitals in $t_{\alpha \beta}^{m n}$ located on the same $\mathrm{HeH}^{+}$molecule. This result is due to the fact that integrals involving orbitals on different $\mathrm{HeH}^{+}$molecules are zero. Continuing this iterative process does not introduce cluster amplitudes that couple different $\mathrm{HeH}^{+}$molecules again because integrals involving orbitals on two or more different molecules vanish. Hence,
the CC equations separate into equations for each $\mathrm{HeH}^{+}$molecule. Consequently, the correlation energy as computed via Eq. (4.16) for $n \mathrm{HeH}^{+}$molecules will be $n$ times the contribution from a single $\mathrm{HeH}^{+}$molecule.

## References

Bartlett, R. J., and Purvis, G. D. (1978). Int. J. Quantum Chem. 16, 561.
Cizek, J., and Paldus, J. (1971). Int. J. Quantum Chem. 5, 359.
Davidson, E. R. (1975). J. Comput. Phys. 17, 87.
Harris, F. E. (1977a). Int. J. Quantum Chem. S11, 403.
Harris, F. E. (1977b). "Electrons in Finite and Infinite Structures" (P. Phariseau and L. Scheive, eds.). Plenum, New York.
Monkhorst, H., and Zivković, T. (1978). J. Math. Phys. 19, 1007.
Paldus, J., Cizek, J., and Shavitt, I. (1972). Phys. Rev. A 5, 50.
Sinanoğlu, O. (1962). J. Chem. Phys. 36, 706.

