Extensivity and the contracted Schrödinger equation

John M. Herbert and John E. Harriman[a)]
Theoretical Chemistry Institute and Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706

(Received 10 June 2002; accepted 30 July 2002)

We provide an extensive formulation of the contracted Schrödinger equation and other reduced eigenvalue equations. Nonextensive (unconnected) terms in these equations cancel exactly, leading to completely connected one- and two-electron equations that together are equivalent to the Schrödinger equation. We discuss how these equations can be solved for the one- and two-electron cumulants. These cumulants yield a two-electron reduced density matrix that is necessarily size consistent, even for an approximate solution. A diagram technique, introduced to aid the formal manipulations, clarifies the connection between density matrix reconstruction and solution of the CSE. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1508369]

I. INTRODUCTION

Recently several groups[1–5] have pursued a ''wave function-less'' approach to electronic structure via direct solution of the contracted Schrödinger equation (CSE). Within the set of *N*-representable[6] reduced density matrices, the CSE (Refs. 7–9) is an equivalent formulation of the electronic Schrödinger equation that couples the two-, three-, and four-electron reduced density matrices (2-, 3-, and 4-RDMs) but does not involve the electronic wave function explicitly. Direct calculation of the 2-RDM is accomplished using approximate *reconstruction functionals*,[2,4,9–13] by means of which the 3- and 4-RDMs are expressed in terms of the 2-RDM, leading to a closed equation for the latter.

Much of the recent literature on RDMs has focused on their cumulant decompositions.[14–22] Each RDM can be decomposed into a connected and an unconnected part, with the latter obtained in a known way from the lower-order RDMs. The connected part (cumulant) is an extensive (additively separable) quantity, in contrast to the RDMs themselves.[15,18] It is thus desirable that any direct method for calculating RDMs be couched in terms of the cumulants, without any unconnected (nonextensive) terms.

As formulated originally,[2,3,7–9] the CSE does not satisfy this criterion, although formally the unconnected terms present in this equation must cancel exactly, since the CSE is equivalent to the ordinary Schrödinger equation. As with other electronic structure methods, however, an approximate solution will not be extensive so long as the basic working equations contain unconnected terms. Consequently the CSE, in practical applications, need not yield a size-consistent 2-RDM. The primary purpose of this paper is to reformulate the CSE by explicitly removing all unconnected terms. This leads to a pair of coupled equations for the one- and two-electron cumulants. Any solution of these connected equations—even an approximate one—is necessarily size consistent.

The connected equations that we obtain are formally equivalent to the ''irreducible'' CSEs introduced, in a second-quantized setting, by Kutzelnigg and Mukherjee.[18–20] In that work, extensive equations were derived in a manner that does not require the original form of the CSE at all. In contrast, we proceed starting from the original (''first-quantized'') form of the CSE in terms of position-space kernels and Hilbert-space operators, explicitly demonstrating the cancellation of all unconnected terms. This derivation does not require the introduction of a basis set, and demonstrates that the final, extensive equations are equivalent to the CSE and the Schrödinger equation. Moreover, our derivation clarifies several important differences between the unconnected and the connected equations: The latter are *implicit* rather than explicit equations for the cumulants, and furthermore the electronic energy, which is an explicit parameter in the original CSE, is absent from its extensive analogues.

After introducing the CSE in Sec. II as a special case of a more general class of *reduced eigenvalue equations*,[23] in Sec. III we develop a diagram technique to facilitate formal manipulation of the terms in this equation. The diagrams also clarify the relationship between the CSE and older, Green's function-based methods in many-body theory. This is discussed in Sec. IV, where we present the extensive form of the CSE and discuss solution of the connected equations. Section V constitutes a summary.

II. CUMULANTS AND REDUCED EIGENVALUE EQUATIONS

Employing the abbreviated notation $''1'' \equiv \mathbf{x}_1$ for the composite space/spin coordinate of electron 1, let

$$\hat{W}(1,\ldots,N) = \sum_{j=1}^{N} \hat{f}(j) + \sum_{j<k}^{N} \hat{g}(j,k), \tag{1}$$

be a symmetric operator on the *N*-electron Hilbert space. Note that this implies $\hat{g}(j,k) = \hat{g}(k,j)$. We wish to consider RDM analogues of the (*N*-particle) eigenvalue equation

$$\hat{W}\Psi = w\Psi. \tag{2}$$

a)Electronic mail: harriman@chem.wisc.edu

Let the eigenvalue w be fixed and assume that Ψ is nondegenerate and unit normalized. The restriction to nondegenerate eigenstates will be relaxed in Sec. IV, but at present it is appropriate to consider only pure-state density matrices. The N-particle density matrix for the pure state Ψ is

$$\Gamma_N(1,\ldots,N;1',\ldots,N')=\Psi(1,\ldots,N)\Psi^*(1',\ldots,N'). \tag{3}$$

In terms of the reduction operator (partial trace)

$$\mathrm{tr}_{p+1,\ldots,q}=\int d\mathbf{x}_{p+1}\cdots d\mathbf{x}_q\int d\mathbf{x}'_{p+1}\cdots d\mathbf{x}'_q \times\delta(\mathbf{x}_{p+1}-\mathbf{x}'_{p+1})\cdots\delta(\mathbf{x}_q-\mathbf{x}'_q), \tag{4}$$

where $p<q\leqslant N$, RDMs for the state Ψ are obtained from the relationship

$$\Gamma_p=\left(\frac{q!(N-q)!}{p!(N-p)!}\right)\mathrm{tr}_{p+1,\ldots,q}\Gamma_q. \tag{5}$$

This definition establishes the normalization

$$\mathrm{tr}\Gamma_p=\binom{N}{p}. \tag{6}$$

From the N-particle Hilbert-space eigenvalue equation follows a hierarchy of p-particle *reduced* eigenvalue equations,[7–9,23] for $1\leqslant p\leqslant N-2$. The pth equation of this hierarchy, which couples Γ_p, Γ_{p+1}, and Γ_{p+2}, can be expressed as

$$\Omega^{(p)}(1,\ldots,p;1',\ldots,p')\equiv 0, \tag{7}$$

in which $\Omega^{(p)}$ is the p-particle kernel[3]

$$\begin{aligned}\Omega^{(p)}&(1,\ldots,p;1',\ldots,p')\\ &=\left[\sum_{j=1}^{p}\hat{f}(j)+(1-\delta_{p,1})\sum_{j<k}^{p}\hat{g}(j,k)-w\right]\Gamma_p\\ &\quad+(p+1)\mathrm{tr}_{p+1}\left\{\left[\hat{f}(p+1)+\sum_{j=1}^{p}\hat{g}(j,p+1)\right]\Gamma_{p+1}\right\}\\ &\quad+\binom{p+2}{2}\mathrm{tr}_{p+1,p+2}\{\hat{g}(p+1,p+2)\Gamma_{p+2}\}.\end{aligned} \tag{8}$$

Here, $\Gamma_n=\Gamma_n(1,\ldots,n;1',\ldots,n')$.

Following Kutzelnigg and Mukherjee,[18–20] we refer to Eq. (7) as the pth-order CSE, CSE(p). [Sometimes CSE(p) is called the $(p,p+2)$-CSE.] Strictly speaking, this terminology implies that $\hat{W}$ is an electronic Hamiltonian, which is clearly the most important case. We emphasize, however, that the formal structure of the equations is the same for any $\hat{W}$ of the form in Eq. (1). In particular, the reduced equations for $\hat{W}=\hat{S}^2$ may be useful—for spin eigenstates—as boundary conditions to enforce while solving CSE(p).[24]

The remarkable fact, first demonstrated by Nakatsuji,[7] is that for each $p\geqslant 2$, CSE(p) is equivalent (in a necessary and sufficient sense) to the original Hilbert-space eigenvalue equation, provided that the former is solved subject to appropriate boundary conditions (N-representability conditions[6]) for the RDMs. Stated differently, CSE(p), which represents a closed equation for the $(p+2)$-RDM, has a unique N-representable solution Γ_{p+2}. Absent N-representability constraints this equation has many spurious solutions.[25,26] Ultimately CSE(2) is the equation that we wish to solve, since it is the most tractable reduced equation that is still equivalent to the original Hilbert space equation. We will see that there is reason to consider CSE(1) as well, however, for it plays an instrumental role in removing unconnected terms from CSE(2).

Using a generating functional for the RDMs, one obtains a cumulant decomposition of these quantities in the usual way; for details, we refer the reader to Refs. 14, 15, and 17. Let Δ_p denote the pth-order RDM cumulant (the connected or additively separable part of Γ_p). Then[14]

$$\Gamma_1=\Delta_1, \tag{9a}$$

$$\Gamma_2=\Delta_2+\Delta_1\wedge\Delta_1, \tag{9b}$$

$$\Gamma_3=\Delta_3+\Delta_1^{\wedge 3}+3\Delta_2\wedge\Delta_1, \tag{9c}$$

and

$$\Gamma_4=\Delta_4+\Delta_1^{\wedge 4}+6\Delta_2\wedge\Delta_1\wedge\Delta_1+4\Delta_3\wedge\Delta_1+3\Delta_2\wedge\Delta_2. \tag{9d}$$

Here, "$\wedge$" denotes an antisymmetrized (Grassmann) product,[23,27]

$$\begin{aligned}(\Delta_p\wedge\Delta_q)&(1,\ldots,p+q;1',\ldots,(p+q)')\\ &=(p+q)!^{-2}\hat{\mathsf{P}}_{p+q}\hat{\mathsf{P}}'_{p+q}(\Delta_p(1,\ldots,p;1',\ldots,p')\\ &\quad\times\Delta_q(p+1,\ldots,p+q;(p+1)',\ldots,(p+q)')),\end{aligned} \tag{10}$$

where $\hat{\mathsf{P}}'_{p+q}$ and $\hat{\mathsf{P}}_{p+q}$ indicate sums over signed permutations of the primed and unprimed coordinates, respectively. "Wedge" exponents appearing in Eqs. (9c) and (9d) are defined by

$$\Delta_p^{\wedge n}=\underbrace{\Delta_p\wedge\Delta_p\wedge\cdots\wedge\Delta_p}_{n\text{ factors}}. \tag{11}$$

Equations (9a)–(9d) define the RDM cumulants and do not depend upon the validity of perturbative expansions of the RDMs. Insofar as perturbation theory is applicable, however, Δ_p is precisely the sum of connected diagrams in the expansion of Γ_p.

As mentioned in the Introduction, the cumulants Δ_p are extensive (additively separable[15]) quantities, whereas the RDMs themselves are not. With the normalization chosen in Eq. (6), extensive quantities have a trace proportional to N while nonextensive quantities scale as some higher power of N; for example, tr $\Gamma_p\sim N^p$. Note that size consistency, along with nonzero correlation energy per particle in the thermodynamic limit, are consequences of extensivity, not the definition of this term.[28]

As an example of the extensivity of the cumulants, let us evaluate tr Δ_2 and tr Δ_3. To this end, define

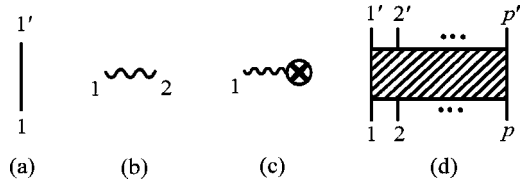


FIG. 1. Basic diagram elements used in this paper: (a) $\Gamma_1(1;1')$, (b) $\hat{g}(1,2)$, (c) $\hat{f}(1)$, and (d) $\Delta_p(1,\ldots,p;1',\ldots,p')$, for $p\geqslant 2$.

$$\eta_k = N^{-1}\ \mathrm{tr}(\Delta_1^k), \tag{12}$$

where Δ_1^k indicates an ordinary matrix product, as opposed to the wedge product $\Delta_1^{\wedge k}$. Since the eigenvalues of $\Delta_1=\Gamma_1$ lie between 0 and 1, it follows that $1\geqslant\eta_2\geqslant\eta_3\geqslant\cdots\geqslant 0$. This is true even for extended systems, where $N\to\infty$. From Eq. (9b) one obtains

$$\mathrm{tr}\ \Delta_2 = N(\eta_2-1)/2, \tag{13}$$

so clearly $\mathrm{tr}\ \Delta_2\sim N$, even as $N\to\infty$. In contrast to the 2-RDM, the trace of Δ_2 depends upon how far Γ_1 deviates from idempotency, though in any case $-N/2\leqslant\mathrm{tr}\ \Delta_2\leqslant 0$. The trace of Δ_3,

$$\mathrm{tr}\ \Delta_3 = N(1-3\eta_2+2\eta_3)/3, \tag{14}$$

also scales like N and depends upon the eigenvalues of Γ_1. Traces of Δ_2 and Δ_3 were examined for some model problems in Ref. 15.

Replacing the RDMs in $\Omega^{(2)}$ with their cumulant decompositions elucidates the unconnected terms in CSE(2). Consider as an example the following term in $\Omega^{(2)}(1,2;1',2')$:

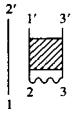
$$\begin{aligned}\hat{f}(1)&\Gamma_2(1,2;1',2') \\ &= \hat{f}(1)[\Delta_2(1,2;1',2') + \tfrac{1}{2}\Delta_1(1;1')\Delta_1(2;2') \\ &\quad - \tfrac{1}{2}\Delta_1(1;2')\Delta_1(2;1')].\end{aligned} \tag{15}$$

The first term on the right is obviously connected, and we may deduce that the second term is unconnected because its trace equals $N^2\langle\hat{f}\rangle/2$. The third term, which constitutes a *transvection*[16,22] of Δ_1 with itself, is actually connected, but differs from the second term by a coordinate permutation. If the second term is removed from CSE(2) then the third term ought to be removed as well, for otherwise we destroy the antisymmetry of $\Omega^{(2)}$.

Unconnected terms and their exchange counterparts are readily identified using a diagram technique, so before attempting to cancel all unconnected terms in CSE(2) we first introduce diagrammatic representations of $\Omega^{(1)}$ and $\Omega^{(2)}$. The diagrams are not strictly necessary, but are quite convenient and (in the authors' opinion) easier to check for mistakes than lengthy algebraic formulas. In addition, several existing reconstruction functionals for Γ_3 and Γ_4[2,3,12,13] have been derived via diagrammatic many-body perturbation theory, and a diagrammatic representation for CSE(2) clarifies the role of this equation in improving approximate reconstruction functionals. Our diagram conventions are conceived with this purpose in mind, and are basically unrelated to the CSE diagrams introduced by Mukherjee and Kutzelnigg.[19,20]

III. DIAGRAMMATIC REPRESENTATIONS

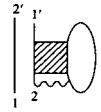
The basic diagram elements representing $\Gamma_1=\Delta_1$, $\hat{f}$, $\hat{g}$, and Δ_p are illustrated in Fig. 1. A representation of each term in $\Omega^{(p)}$ is constructed by connecting operator diagrams to cumulant diagrams at the coordinates on which the operators act. For instance



$$= \hat{g}(2,3)\,\Delta_1(1;2')\,\Delta_2(2,3;1',3') \tag{16}$$

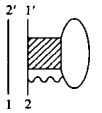
is obtained by attaching a $\hat{g}(2,3)$ diagram at the lower endpoints of a $\Delta_2(2,3;1',3')$ diagram, since according to Fig. 1(d) these endpoints stand for coordinates 2 and 3. A factor of $\Delta_1(1;2')$ is present, as indicated, but because $\hat{g}(2,3)$ operates on neither $\mathbf{x}_1$ nor $\mathbf{x}_2'$, this part of the diagram is not connected to the rest.

A trace over coordinate $\mathbf{x}_n$ is indicated by connecting the line labeled n to the line labeled n'. The labels n and n' are then deleted, since these coordinates become a single dummy integration variable. Diagrammatically, this creates a loop in the case that both $\mathbf{x}_n$ and $\mathbf{x}_n'$ are arguments of the same cumulant. Applying tr_3 to Eq. (16), one obtains



$$= \Delta_1(1;2')\ \mathrm{tr}_3\left\{\hat{g}(2,3)\,\Delta_2(2,3;1',3')\right\}. \tag{17}$$

If, on the other hand, $\mathbf{x}_n$ and $\mathbf{x}_n'$ are arguments of different cumulants, then tr_n serves to connect two cumulant diagrams:

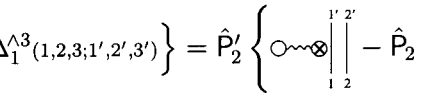


$$= \Delta_1(1;2')\ \mathrm{tr}_{3,4}\left\{\hat{g}(3,4)\,\Delta_1(2;3')\,\Delta_2(3,4;1',4')\right\}. \tag{18}$$

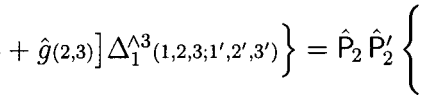
Note carefully the difference between this diagram and the previous one. These two examples illustrate that *internal* operator vertices (those not appearing at the endpoint of a cumulant line) are each associated with a coordinate integration, whereas a vertex occurring at the endpoint of a cumulant line does not imply an integration.

In close analogy to diagrammatic perturbation theory (although our diagrams are not perturbative), we have transformed the problem of generating terms in $\Omega^{(p)}$ into a problem of generating topologically distinct diagrams. Restriction to topologically inequivalent diagrams is an easy way to incorporate symmetries such as $\hat{g}(j,k)=\hat{g}(k,j)$ that reduce the number of terms in $\Omega^{(p)}$. The nontrivial terms in $\Omega^{(1)}$ and $\Omega^{(2)}$ that involve only the one-electron cumulant are

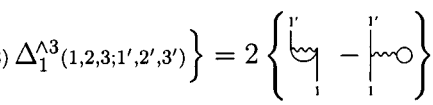
$$6\ \mathrm{tr}_3\left\{\hat{f}_{(3)}\,\Delta_1^{\wedge 3}(1,2,3;1',2',3')\right\} = \hat{\mathsf{P}}_2'\left\{ \quad - \hat{\mathsf{P}}_2 \quad \right\}, \tag{19}$$



$$6\ \mathrm{tr}_3\left\{\left[\hat{g}_{(1,3)}+\hat{g}_{(2,3)}\right]\Delta_1^{\wedge 3}(1,2,3;1',2',3')\right\} = \hat{\mathsf{P}}_2\,\hat{\mathsf{P}}_2'\left\{ \quad - \quad - \quad \right\}, \tag{20}$$

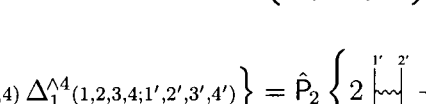


$$6\ \mathrm{tr}_{2,3}\left\{\hat{g}_{(2,3)}\,\Delta_1^{\wedge 3}(1,2,3;1',2',3')\right\} = 2\left\{ \quad - \quad \right\} + \left(\quad - \quad \right), \tag{21}$$



and

$$24\ \mathrm{tr}_{3,4}\left\{\hat{g}_{(3,4)}\,\Delta_1^{\wedge 4}(1,2,3,4;1',2',3',4')\right\} = \hat{\mathsf{P}}_2\left\{ 2 \quad + \left(\quad - \quad \right) \quad + 2\,\hat{\mathsf{P}}_2'\left(\quad - \quad \right)\right\}. \tag{22}$$



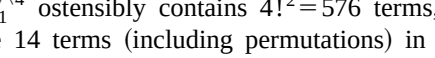
These expressions are highly compact compared to brute-force expansions of the Grassmann products $\Delta_1^{\wedge 3}$ and $\Delta_1^{\wedge 4}$. For example, $\Delta_1^{\wedge 4}$ ostensibly contains $4!^2=576$ terms, as compared to the 14 terms (including permutations) in Eq. (22).

Certain diagrams in the expressions above have no coordinate dependence. These can be related to the eigenvalue w if we decompose w into one- and two-electron contributions, $w=w_1+w_2$, with $w_1=N\langle\hat{f}\rangle$ and $w_2=\binom{N}{2}\langle\hat{g}\rangle$. Diagrammatically,

$$w_1 = \mathrm{tr}\left\{\hat{f}_{(1)}\,\Delta_1(1;1')\right\} = \quad , \tag{23}$$

and

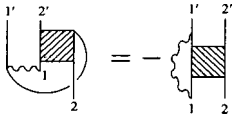
$$w_2 = \mathrm{tr}\left\{\hat{g}_{(1,2)}\,\Gamma_2(1,2;1',2')\right\} = \quad + \tfrac{1}{2}\left(\quad - \quad \right). \tag{24}$$



When w is the electronic energy, the three contributions to w_2 in the equation above are, respectively, the cumulant correlation energy, the classical electrostatic (Hartree) energy, and the exchange energy. This cumulant decomposition of w_2 provides a universal, extensive definition for these quantities, which does not depend upon any noninteracting (Hartree–Fock or Kohn–Sham) reference system.

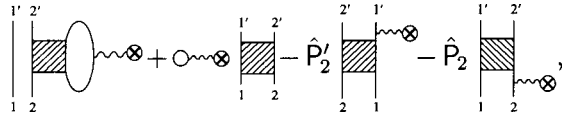
For expressions involving higher-order cumulants, one can utilize the antisymmetry of Δ_p to reduce the number of required terms. For example,

$$\quad = - \quad \tag{25}$$

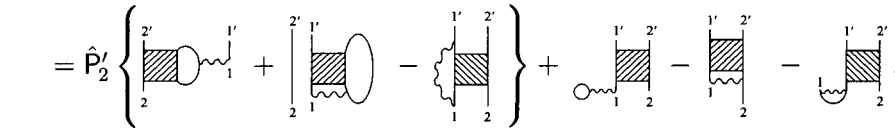


is obtained by exchanging the lines entering the top of Δ_2, corresponding to a permutation of the primed coordinates in this cumulant. Making use of this and similar relationships, we obtain

$$9\ \mathrm{tr}_3\left\{\hat{f}_{(3)}\left(\Delta_1\wedge\Delta_2\right)_{(1,2,3;1',2',3')}\right\} = \hat{\mathsf{P}}_2\,\hat{\mathsf{P}}_2' \quad + \quad - \hat{\mathsf{P}}_2' \quad - \hat{\mathsf{P}}_2 \quad , \tag{26}$$



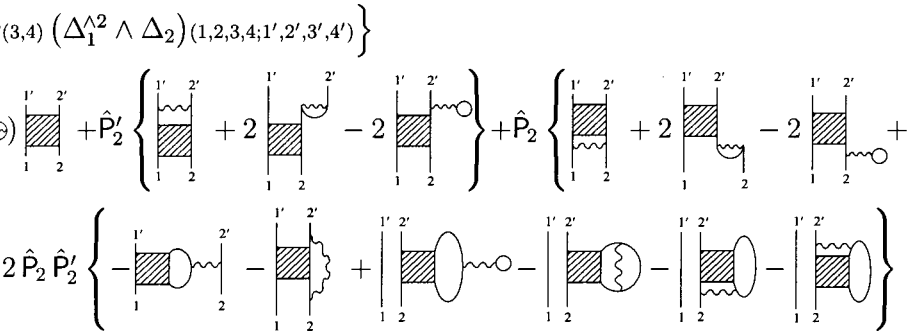
$$9\ \mathrm{tr}_3\left\{\hat{g}_{(1,3)}\left(\Delta_1\wedge\Delta_2\right)_{(1,2,3;1',2',3')}\right\} = \hat{\mathsf{P}}_2'\left\{ \quad + \quad - \quad \right\} + \quad - \quad - \quad , \tag{27}$$



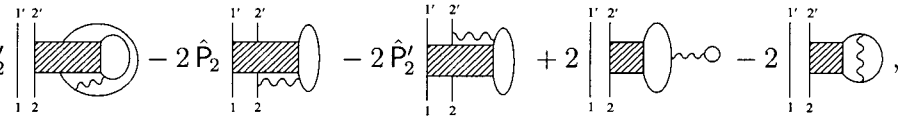
$$72\ \mathrm{tr}_{3,4}\left\{\hat{g}_{(3,4)}\left(\Delta_1^{\wedge 2}\wedge\Delta_2\right)_{(1,2,3,4;1',2',3',4')}\right\}$$

$$= \left(\quad - \quad \right) \quad + \hat{\mathsf{P}}_2'\left\{ \quad + 2 \quad - 2 \quad \right\} + \hat{\mathsf{P}}_2\left\{ \quad + 2 \quad - 2 \quad + \quad \right\}$$

$$+ 2\,\hat{\mathsf{P}}_2\,\hat{\mathsf{P}}_2'\left\{ - \quad - \quad + \quad - \quad - \quad - \quad \right\}, \tag{28}$$



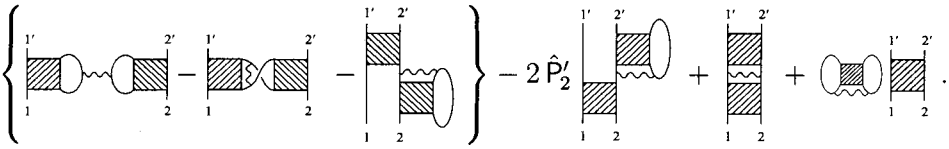
$$16\ \mathrm{tr}_{3,4}\left\{\hat{g}_{(3,4)}\left(\Delta_1 \wedge \Delta_3\right)_{(1,2,3,4;1',2',3',4')}\right\}$$



(29)

and

$$18\ \mathrm{tr}_{3,4}\left\{\hat{g}_{(3,4)}\,\Delta_2^{\wedge 2}{}_{(1,2,3,4;1',2',3',4')}\right\}$$



(30)

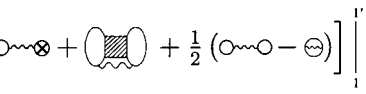
IV. THE CONNECTED EQUATIONS

A. Cancellation of unconnected terms

Using a generating functional for $\Omega^{(p)}$, Yasuda[3] has demonstrated that this kernel can be decomposed into a connected part $\Omega_c^{(p)}$ and an unconnected part $\Omega_u^{(p)}$; our diagrammatic treatment brings this to the forefront.

Clearly $\Omega^{(1)}$ as defined in Eq. (8) contains unconnected terms, for example

$$w\,\Gamma_1(1;1') =$$



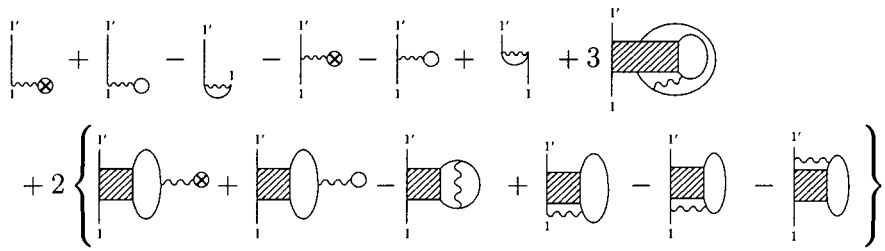
(31)

However, all unconnected terms in $\Omega^{(1)}$ are found to cancel exactly, so $\Omega_u^{(1)}=0$. Of course, an approximate solution of CSE(1) may not lead to exact cancellation, so instead of solving the equation $\Omega^{(1)}\equiv 0$, one ought to solve instead the manifestly extensive equation

$$\Omega_c^{(1)}\equiv 0. \tag{32}$$

The connected part of $\Omega^{(1)}$ is found to be

$$\Omega_c^{(1)}(1;1') =$$



(33)

Since the unconnected terms cancel exactly, the extensive equation is equivalent, in a necessary and sufficient sense, to CSE(1). Following Kutzelnigg and Mukherjee,[18–20] we refer to Eq. (32) as the first-order *irreducible* CSE, ICSE(1). To obtain an equation that is equivalent, within a finite basis set, to our ICSE(p), one must solve the Kutzelnigg–Mukherjee[19] version of ICSE(p) simultaneously with its adjoint equation. In contrast, the solution to Eq. (32) is automatically self-adjoint.

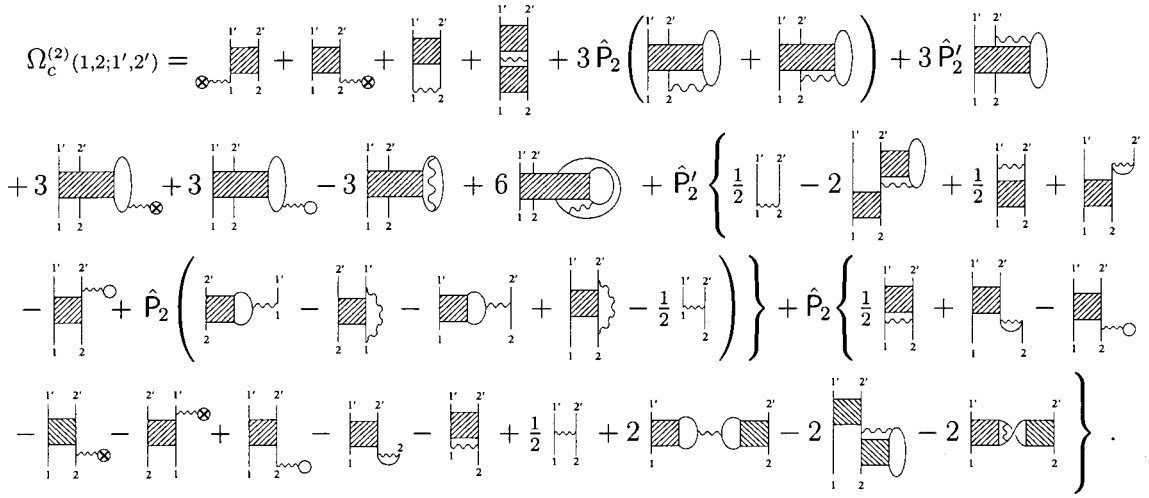
Neither CSE(1) nor ICSE(1) is equivalent to the original Hilbert-space eigenvalue equation; for that we need CSE(2). The unconnected part of $\Omega^{(2)}$ is[3]

$$\Omega_u^{(2)}=\Gamma_1\wedge\Omega^{(1)}, \tag{34}$$

which implies that $\Omega_u^{(2)}=\Gamma_1\wedge\Omega_c^{(1)}$ since $\Omega_u^{(1)}=0$. This relationship can be verified directly using the expressions in the previous section. Thus, insofar as CSE(1) is exactly satisfied—a necessary condition in order that CSE(2) be satisfied—all unconnected terms in $\Omega^{(2)}$ cancel and we obtain the extensive equation

$$\Omega_c^{(2)}\equiv 0, \tag{35}$$

which we call ICSE(2). Carrying out the cancellation is relatively easy using diagrams, and one obtains



(36)

Including permutations, this expression for $\Omega_c^{(2)}$ contains 68 terms, a significant reduction as compared to the unsimplified Grassmann products.

B. Discussion of the connected equations

Perhaps the most striking feature of ICSE(1) and ICSE(2) is the absence of the eigenvalue w in these equations. In hindsight its disappearance should not be surprising, since w appears in $\Omega^{(p)}$ as $w\Gamma_p$. The observable w scales (asymptotically) as N, as does the connected part of Γ_p, hence no part of the product $w\Gamma_p$ exhibits correct scaling in the thermodynamic limit and this entire term must cancel with some other part of CSE(p). (This is analogous to the fact that the coupled-cluster amplitude equations, which are extensive, contain the cluster amplitudes but not the electronic energy, even though these equations derive from the Schrödinger equation.) However, w is specified *implicitly* in ICSE(1) and ICSE(2), insofar as these cumulants determine Γ_2 and thus $w=\mathrm{tr}(\hat{W}\Gamma_2)$.

The absence of w in ICSE(p) has important consequences, to which we shall return later in this section. First, however, we demonstrate that ICSE(1) and ICSE(2) are independent equations that must be solved in tandem, as opposed to CSE(2), which implies CSE(1) via a trace over one electronic coordinate. The difference is that Γ_2 determines $\Gamma_1=\Delta_1$ but Δ_2 does not fully specify Δ_1 and is thus insufficient to determine Γ_2 or w.

To see this, observe that

$$\mathrm{tr}_2(\Delta_2)=\tfrac{1}{2}(\Delta_1^2-\Delta_1). \tag{37}$$

It follows that Δ_1 and $\mathrm{tr}_2(\Delta_2)$ are simultaneously diagonalizable. Let $\{\gamma_k\}$ be the eigenvalues of Δ_1, and for each γ_k let d_k be the eigenvalue of $\mathrm{tr}_2(\Delta_2)$ associated with the same eigenvector. These eigenvalues are related according to

$$d_k=\gamma_k(\gamma_k-1)/2, \tag{38a}$$

which yields

$$\gamma_k=\tfrac{1}{2}(1\pm\sqrt{1+8d_k}). \tag{38b}$$

As a function of d_k, γ_k is thus double-valued, as shown in Fig. 2, so that strictly speaking the eigenvalues of $\mathrm{tr}_2(\Delta_2)$ do not determine those of Δ_1.

For systems that do not exhibit strong multideterminant character (''nondynamical correlation''), most eigenvalues of Δ_1 should lie near 0 or 1. Given Δ_2 (and thus the d_k), it may be possible in such cases to choose, for each k, one of the two solutions γ_k in Eq. (38b), based upon the nature of the kth eigenvector (that is, based upon our intuition as to whether this eigenvector should be strongly or weakly occupied). To see this, suppose $\gamma_k=\varepsilon$ or $\gamma_k=1-\varepsilon$. Upon calculating the d_k corresponding to each, and substituting this back into Eq. (38b), one obtains in either case a choice between the solutions $\gamma_k=1-\varepsilon+O(\varepsilon^2)$ and $\gamma_k=\varepsilon+O(\varepsilon^2)$. So long as $\varepsilon^2\ll\varepsilon$, and assuming that one can ascertain whether each eigenvector of $\mathrm{tr}_2(\Delta_1)$ is strongly or weakly occupied, one may obtain Δ_1 from Δ_2, and from them Γ_2 and w. In this case one need to solve only ICSE(2).

In deriving ICSE(1) and ICSE(2) from the corresponding CSEs, we have merely identified and removed terms that cancel exactly; as such, these two connected equations, when solved simultaneously, are entirely equivalent to CSE(2) and thus equivalent to the original Hilbert-space eigenvalue equation (Schrödinger equation), provided that appropriate N-representability constraints are enforced in the solving the ICSE(p). Since necessary and sufficient N-representability

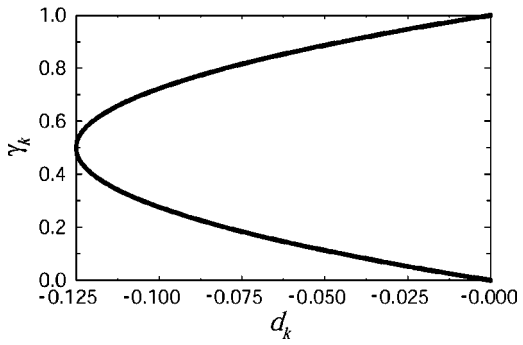


FIG. 2. An eigenvalue γ_k of Δ_1 as a (double-valued) function of the corresponding eigenvalue d_k of $\mathrm{tr}_2(\Delta_2)$.

constraints are not known, one must in practice contend with an infinite number of spurious solutions to these equations. Recent calculations[2,4,29,30] in which CSE(2) is solved starting from an *N*-representable (actually, Hartree–Fock) 2-RDM indicate that, for ground states, the solution usually converges to a 2-RDM that is nearly consistent with the necessary *P*, *Q*, and *G* conditions for *N* representability.[6] (That is, all eigenvalues of Γ_2 and also the so-called *Q* and *G* matrices are positive or else very small.)

Even in the presence of necessary and sufficient *N*-representability constraints, however, the solution of CSE(2) is only unique provided that the eigenvalue *w* is specified and fixed. Because *w* does not appear in the ICSEs, in order to obtain a unique solution ICSE(1) and ICSE(2) must be solved subject not only to *N*-representability constraints but also subject to the constraint that $w=\mathrm{tr}(\hat{W}\Gamma_2)$ remains fixed. For auxiliary constraint equations such as, for example, the $\hat{S}^2$ reduced eigenvalue equation, one would know $\langle\hat{S}^2\rangle$ in advance and could constrain $\langle\hat{S}^2\rangle=\mathrm{tr}(\hat{S}^2\Gamma_2)$. In the basic equations of our theory, however, $\hat{W}$ is an electronic Hamiltonian and such a constraint would require us to know the electronic energy in advance. Foregoing the energy constraint, ICSE(1) and ICSE(2) possess *N*-representable solutions corresponding to the ground state, the excited states, and also all superposition states that can be formed from degenerate eigenfunctions of $\hat{W}$. In practice, the same is true for CSE(2), insofar as the energy is unknown and the eigenvalue appearing in CSE(2) must be updated iteratively during the course of the solution. Viewed in this way, the fact that *w* does not appear in ICSE(1) or ICSE(2) may actually be advantageous.

We now wish to discuss how ICSE(1) and ICSE(2) can be solved. First, let us discuss the solution of CSE(2). For $w\neq 0$, CSE(2) may be written

$$\Gamma_2 = w^{-1}F_w[\Gamma_2,\Gamma_3,\Gamma_4], \tag{39}$$

where the functional $F_w=\Omega^{(2)}+w\Gamma_2$. Assuming one has approximate reconstruction functionals $\Gamma_3[\Gamma_2]$ and $\Gamma_4[\Gamma_2]$, Eq. (39) can be solved for Γ_2 by one of two means. The first option is to substitute the reconstruction functionals directly into F_w, which effectively makes F_w a functional of Γ_2 only. Upon expanding Eq. (39) in a finite basis set, this leads to a closed set of nonlinear equations for the tensor elements of Γ_2, and these can be solved, for example, by a Newton–Raphson procedure.[2–4] Alternatively, Eq. (39) can be solved by self-consistent iteration, employing the reconstruction functionals at each iteration to generate updated 3- and 4-RDMs from the current 2-RDM, and using the current 2-RDM to estimate *w*. Several algorithms for carrying out this iteration have been described.[1,5,31,32]

It does not appear that the ICSEs can be solved by self-consistent iteration, however, since these equations specify the cumulants Δ_1 and Δ_2 only implicitly. Using cumulant reconstruction functionals $\Delta_3[\Delta_1,\Delta_2]$ and $\Delta_4[\Delta_1,\Delta_2]$, however, one can certainly derive closed nonlinear equations for the elements of Δ_1 and Δ_2, which could be solved using the aforementioned Newton–Raphson procedure. Of the RDM reconstruction functionals derived to date, several[2,3,12,13] utilize the cumulant decompositions in Eqs. (9c) and (9d) to obtain the unconnected portions of Γ_3 and Γ_4, then use many-body perturbation theory to estimate the connected parts Δ_3 and Δ_4 in terms of Δ_1 and Δ_2, the latter essentially serving as a renormalized pair interaction. Reconstruction functionals of this type are equally useful in solving ICSE(1) and ICSE(2), but the reconstruction functionals introduced by Valdemoro and coworkers[10,11] cannot be used to solve the ICSEs because they contain no connected terms in Γ_3 or Γ_4 (and thus no contributions to Δ_3 or Δ_4).

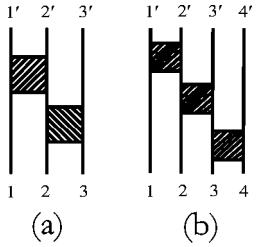


FIG. 3. Lowest-order connected corrections to (a) Δ_3 and (b) Δ_4, within a renormalized ladder-type approximation.

C. Reconstruction and solution of the reduced equations

Finally, we present some observations concerning the interconnection of the reconstruction process and the solution of either CSE(p) or ICSE(p). The perturbative reconstruction functionals mentioned above each constitute a finite-order ladder-type approximation to Δ_3 and Δ_4;[13] examples of the lowest-order corrections of this type are shown in Fig. 3. As discussed in Ref. 13, ladder-type reconstructions—even if extended to infinite order in the effective pair interaction—necessarily neglect ''true'' (that is, simultaneous, in a time-dependent context) three-electron correlations. The CSEs (or ICSEs), however, help to build these correlations back into the cumulants, which becomes clear upon examination of the diagrammatic representations of these equations, together with diagrammatic representations of the reconstruction functionals.

In Fig. 4(a) we show a typical diagram in the expansion of Δ_3 that cannot be incorporated into any ladder-type diagram because it involves simultaneous correlation between three particles.[13] As it appears in CSE(2) and ICSE(2), however, Δ_3 is always traced over coordinate $\mathbf{x}_3$, and in Fig. 4(b) we show the effect of tr_3 on the diagram in Fig. 4(a). The diagram in Fig. 4(b) *is* included in the partial trace of a *third*-order ladder-type diagram, namely, the one shown in Fig. 4(c). Thus, the presence of tr_3 in the two-particle equations allows us to incorporate three- and higher-body effects that would not otherwise be present in a ladder approximation for the three- and four-electron cumulants.

Actually, three-particle correlations such as that in Fig. 4(a) are introduced by the CSEs and ICSEs, even within a second-order ladder approximation. To understand why, consider the diagram in Fig. 4(d), which represents one of the terms in $\Omega_c^{(2)}$. Within a second-order ladder approximation to Δ_3, the diagram in Fig. 4(b) is included within Fig. 4(d). Thus, three- and higher-body effects are incorporated into the

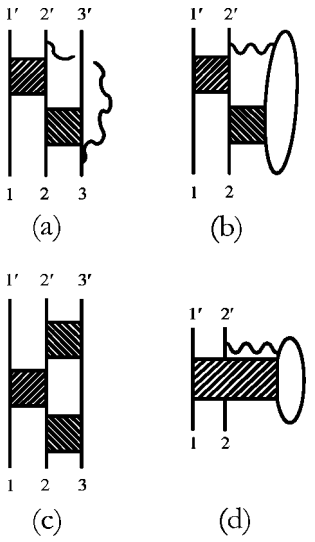


FIG. 4. Diagrams illustrating the connection between reconstruction and solution of the CSEs or ICSEs. See the text for an explanation.

cumulants Δ_3 and Δ_4 by the CSEs or ICSEs, even when these effects are absent from approximate reconstruction functionals. In effect, solution of these equations corresponds to a partial summation of the perturbation series for Γ_2 [in the case of CSE(2)] or Δ_1 and Δ_2 [in the case that ICSE(1) and ICSE(2) are solved simultaneously]. The connection between reconstruction and solution of coupled Green's function equations of motion, which are time-dependent hierarchies analogous to the CSE(p) hierarchy, has been explored in some detail.[33,34] A more thorough exploration of this connection in the present context would be welcome.

V. SUMMARY

We have introduced extensive formulations of the contracted Schrödinger equation and other reduced eigenvalue equations by explicitly canceling all unconnected terms that appear in the usual formulations of these equations. The resulting equations guarantee that size-consistent solutions are obtained for the one- and two-electron cumulants, even for an approximate solution. Moreover, the extensive formulation is the only one appropriate for extended systems. The derivation is carried out using position-space kernels and Hilbert-space operators, circumventing the introduction of a finite basis approximation.

Although the cancellation of unconnected terms is formally exact, this operation is far from trivial, resulting in simultaneous nonlinear equations for the one- and two-particle cumulants that contain these quantities implicitly rather than explicitly. The eigenvalue w, which appeared explicitly in the nonextensive form of the CSEs, is absent from the extensive equations (though it can, of course, be recovered once the cumulants are known). Since w is absent, within the extensive formulation one avoids an additional approximation (namely, approximation of the eigenvalue w) that must be introduced in order to solve the CSEs.

Finally, we have demonstrated an important connection between cumulant (or density matrix) reconstruction and solution of the CSEs or their connected counterparts. When approximate reconstruction functionals based on finite-order many-body perturbation theory are employed, simultaneous many-particle correlations must be neglected beyond a certain number of particles. The structure of the CSEs, however, plays an instrumental role in putting these correlations back into the 2-RDM.

ACKNOWLEDGMENTS

One of the authors (J.M.H.) acknowledges a graduate fellowship from the Department of Defense.

[1] F. Colmenero and C. Valdemoro, Int. J. Quantum Chem. **51**, 369 (1994).
[2] K. Yasuda and H. Nakatsuji, Phys. Rev. A **56**, 2648 (1997).
[3] K. Yasuda, Phys. Rev. A **59**, 4133 (1999).
[4] M. Nakata, M. Ehara, K. Yasuda, and H. Nakatsuji, J. Chem. Phys. **112**, 8772 (2000).
[5] D. A. Mazziotti, J. Chem. Phys. **116**, 1239 (2002).
[6] A. J. Coleman and V. I. Yukalov, "Reduced Density Matrices: Coulson's Challange" in *Lecture Notes in Chemistry* (Springer, Berlin, 2000), Vol. 72.
[7] H. Nakatsuji, Phys. Rev. A **14**, 41 (1976).
[8] L. Cohen and C. Frishberg, Phys. Rev. A **13**, 927 (1976).
[9] D. A. Mazziotti, Phys. Rev. A **57**, 4219 (1998).
[10] F. Colmenero, C. Pérez del Valle, and C. Valdemoro, Phys. Rev. A **47**, 971 (1993).
[11] F. Colmenero and C. Valdemoro, Phys. Rev. A **47**, 979 (1993).
[12] H. Nakatsuji and K. Yasuda, Phys. Rev. Lett. **76**, 1039 (1996).
[13] J. M. Herbert, J. Chem. Phys. (unpublished).
[14] D. A. Mazziotti, Chem. Phys. Lett. **289**, 419 (1998).
[15] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **110**, 2800 (1999).
[16] D. A. Mazziotti, Chem. Phys. Lett. **326**, 212 (2000).
[17] P. Ziesche, in *Many-Electron Densities and Reduced Density Matrices*, edited by J. Cioslowski (Plenum, New York, 2000), p. 33.
[18] W. Kutzelnigg and D. Mukherjee, Chem. Phys. Lett. **317**, 567 (2000).
[19] D. Mukherjee and W. Kutzelnigg, J. Chem. Phys. **114**, 2047 (2001); **114**, 8226 (2001).
[20] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **116**, 4787 (2002).
[21] J. E. Harriman, Phys. Rev. A **65**, 052507 (2002).
[22] F. Harris, Int. J. Quantum Chem. **90**, 105 (2002).
[23] J. M. Herbert and J. E. Harriman, Phys. Rev. A **65**, 022511 (2002).
[24] C. Valdemoro, L. M. Tel, and E. Pérez-Romero, Adv. Quantum Chem. **28**, 33 (1997).
[25] J. E. Harriman, Phys. Rev. A **19**, 1893 (1979).
[26] K. Yasuda, Phys. Rev. A **65**, 052121 (2002).
[27] A. J. Coleman and I. Absar, Int. J. Quantum Chem. **18**, 1279 (1980).
[28] R. J. Bartlett, in *Modern Electronic Structure Theory, Part II*, edited by D. R. Yarkony (World Scientific, Singapore, 1995), p. 1047.
[29] D. A. Mazziotti, in *Many-Electron Densities and Reduced Density Matrices*, edited by J. Cioslowski (Plenum, New York, 2000), p. 139.
[30] D. A. Mazziotti and R. M. Erdahl, Phys. Rev. A **63**, 042113 (2001).
[31] C. Valdemoro, Int. J. Quantum Chem. **60**, 131 (1996).
[32] C. Valdemoro, L. M. Tel, E. Pérez-Romero, and A. Torre, J. Mol. Struct.: THEOCHEM **537**, 1 (2001); **574**, 255 (2001).
[33] R. D. Mattuck and A. Theumann, Adv. Phys. **20**, 721 (1971).
[34] S.-J. Wang, W. Zuo, and W. Cassing, Nucl. Phys. A **573**, 245 (1994).