PHYSICAL REVIEW VOLUME 126, NUMBER 3 MAY 1, 1962

Perturbation Theory for the Fock-Dirac Density Matrix*

R. McWeeny†
Quantum Chemistry Group, Uppsala University, Uppsala, Sweden

(Received October 16, 1961)

In Hartree-Fock theory and its various generalizations, it is customary to solve an eigenvalue problem involving an effective one-body Hamiltonian. The eigenvectors determine the Fock-Dirac density matrix, which also appears in the effective Hamiltonian, and solution proceeds iteratively until self-consistency is achieved.

An alternative (necessary and sufficient) condition for a solution is that the density matrix ($\boldsymbol{\rho}$) is idempotent and commutes with the Hamiltonian ($\mathbf{h}$). The change in $\boldsymbol{\rho}$, accompanying a change $\boldsymbol{\Delta}$ in $\mathbf{h}$, can then be expressed as a perturbation series. Formulas for the perturbation, to all orders, are obtained in terms of the unperturbed Hamiltonian and density matrix. It is also shown that the whole perturbation may be obtained directly, without separating the orders, and that the approach is related to earlier steepest-descent methods.

I. INTRODUCTION

It is well known that the energy of a system with Hamiltonian

$$\mathcal{H}=\sum_i \mathcal{H}_1(i)+\frac{1}{2!}\sum_{i,j}{}'\mathcal{H}_2(i,j)+\frac{1}{3!}\sum_{i,j,k}{}'\mathcal{H}_3(i,j,k)+\cdots, \tag{1.1}$$

can be written in the form

$$E=\operatorname{Tr}\mathcal{H}_1\rho_1(1;1')+\frac{1}{2!}\operatorname{Tr}\mathcal{H}_2\rho_2(1,2;1',2')+\frac{1}{3!}\operatorname{Tr}\mathcal{H}_3\rho_3(1,2,3;1',2',3')+\cdots, \tag{1.2}$$

where $\rho_n(1,2,\cdots n;1',2',\cdots n')$ is the reduced density matrix for n particles and is normalized so that $\rho_n(1,2\cdots n;1,2\cdots n)dx_1dx_2\cdots dx_n$ is the probability of finding n particles simultaneously in (space-spin) volume elements $dx_1dx_2\cdots dx_n$ at points $1, 2, \cdots n$ of configuration space. It is also known[1-3] that when the wave function has the 1-determinant form

$$\Phi(1,2,\cdots N)=\frac{1}{\sqrt{(N!)}}\begin{vmatrix}\psi_1(1) & \psi_1(2) & \cdots & \psi_1(N)\\ \psi_2(1) & \psi_2(2) & \cdots & \psi_2(N)\\ & \cdots & & \\ \psi_N(1) & \psi_N(2) & \cdots & \psi_N(N)\end{vmatrix} \tag{1.3}$$

(each ψ_i being a 1-electron function or "spin-orbital") all the reduced density matrixes are expressible in terms of $\rho_1(1;1')$ itself: this 1-body quantity is the "Fock-Dirac" density matrix,

$$\rho(1;1')=\rho_1(1;1')=\sum_{i=1}^{N}\psi_i(1)\psi_i^*(1'), \tag{1.4}$$

and may be regarded[4] as describing an ensemble of N noninteracting particles, each with unit probability of being found in one of the states $\psi_1, \psi_2, \cdots \psi_N$.

In self-consistent field theories we determine optimum orbitals by seeking the stationary values of E. The optimum orbitals $\psi_1, \psi_2, \cdots \psi_N$ then appear as solutions of a 1-body problem,

$$\mathfrak{h}\psi=\epsilon\psi, \tag{1.5}$$

where the "effective 1-body Hamiltonian" for a particle at point 1 may be written[5]

$$\mathfrak{h}(1)=\mathcal{H}_1(1)+\int\mathcal{H}_2(1,2)(1-\mathcal{P}_{12})\rho_1(2;2')|_{2'\to 2}dx_2+\cdots, \tag{1.6}$$

and ϵ is an "orbital energy." The function Φ is then the *exact* wave function for a "model" in which N interactionless particles each move in a common field described by the potentials in $\mathfrak{h}$. It must be remembered that the permutation operators in the integrands in (1.6) also work on the operand. Thus, when only 2-body Coulombic terms appear in (1.1) the 1-electron "model" Hamiltonian is

$$\mathfrak{h}(1)=\mathcal{H}_1(1)+\mathcal{J}(1)-\mathcal{K}(1),$$

where the $\mathcal{J}$ and $\mathcal{K}$ operators are defined by

$$\mathcal{J}(1)\psi(1)=\int\mathcal{H}_2(1,2)\rho_1(2;2)dx_2\,\psi(1), \quad \text{(the "Coulomb" operator)}$$

$$\mathcal{K}(1)\psi(1)=\int\mathcal{H}_2(1,2)\rho_1(1;2)\psi(2)dx_2, \quad \text{(the "exchange" operator).}$$

---

* This work first appeared as Preprint No. 61 (15th March, 1961) from the Quantum Chemistry Group, Uppsala University, Sweden.

† On leave of absence (1960–61) from Departments of Mathematics, Physics and Chemistry, University of Keele, Staffordshire, England.

[1] P. A. M. Dirac, Proc. Cambridge Phil. Soc. **27**, 240 (1931).

[2] J. E. Lennard-Jones, Proc. Cambridge Phil. Soc. **27**, 469 (1931).

[3] P.-O. Löwdin, Phys. Rev. **97**, 1490 (1955).

[4] R. McWeeny, Revs. Modern Phys. **32**, 335 (1960).

[5] P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955), Eq. (43).

Since the potential in the 1-electron model depends upon the density matrix (i.e., upon the orbitals which appear as solutions), the equation (1.5) is usually solved iteratively until *self-consistency* is achieved.

Recently, the original self-consistent field (SCF) theory has been generalized in various ways, formal n-body interactions have acquired considerable importance, and it seems opportune to re-examine the basic problem of obtaining a self-consistent Fock-Dirac density matrix. In the generalizations due to Brueckner and others (see also Löwdin[6]), the *exact* wave function for the actual system is related to those of the model by means of a projection operator $\mathcal{O}$ (which depends upon both the ground state and a complete set of excited states of the model): thus we write

$$\Psi = \mathcal{O}\Phi,$$

where Φ is a one-determinant solution for the model. The exact energy of the actual system is then

$$E=\int(\mathcal{O}\Phi)^*\mathcal{H}(\mathcal{O}\Phi)dx=\int\Phi^*\mathcal{O}^\dagger\mathcal{H}\mathcal{O}\Phi dx=\int\Phi^*\bar{\mathcal{H}}\Phi dx;$$

where $\bar{\mathcal{H}}=\mathcal{O}^\dagger\mathcal{H}\mathcal{O}$ is a *formal* Hamiltonian of type (1.1), containing n-body terms. The same device has been used extensively by Löwdin[5] in discussing systems with definite spin in terms of a one-determinant function Φ: In this case $\mathcal{O}$ is a spin operator which projects from an arbitrary Φ a given eigenfunction of $\mathcal{S}^2$ and $\mathcal{S}_z$.

The basic SCF procedure is as follows. Given an estimated effective Hamiltonian $\mathfrak{h}_0$ with eigenfunctions $\psi_1, \psi_2, \cdots$, we compute a Fock-Dirac density matrix ρ_0 and hence a revised effective Hamiltonian,

$$\mathfrak{h}=\mathfrak{h}_0+\Delta.$$

After revising the Hamiltonian, we recalculate the eigenfunctions and obtain a new density ρ. The process is then continued until revision is no longer necessary, density and Hamiltonian being self-consistent. In any one step, we start from $\mathfrak{h}_0$ and ρ_0 and must proceed to the revised quantities, $\mathfrak{h}=\mathfrak{h}_0+\Delta$ and $\rho=\rho_0+\delta\rho$. We wish to obtain $\delta\rho$ directly, without having to solve the eigenvalue problem (1.5) for the individual eigenfunctions—which in themselves have no significance until self-consistency has actually been achieved. This problem will be solved by a perturbation method, not for the eigenfunctions but for the Fock-Dirac density matrix itself.

II. PERTURBATION THEORY

We shall introduce an arbitrary complete set of 1-electron functions since the theory may then be formulated in a convenient matrix language.[4,7] The matrices are in principle infinite (giving, for example, exact Hartree-Fock functions) but truncation is equivalent merely to the application of constraints and does not affect the formulation: the resultant solutions are then "best approximations" of truncated form.

In the present context the perturbation Δ is purely formal, and the perturbation approach is merely a device for solving the nonlinear equations of SCF theory. But the same formulation applies, of course, when Δ represents an actual perturbation. Thus, $\mathfrak{h}_0$ and ρ_0 may describe a model system of noninteracting electrons (as in simplified molecular and solid-state theories) while Δ may represent the effect of applied electric or magnetic fields. Again, the advantage of a density matrix approach is that the *whole system* of electrons is dealt with and it is unnecessary to consider separately the perturbation of each one-electron energy level.

Suppose there are n occupied orbitals in $\Phi, \psi_1, \psi_2, \cdots \psi_N$ and that these are constructed from m orthonormal basis functions $\phi_1, \phi_2, \cdots \phi_m$, where m may be assumed arbitrarily large. Collecting the functions into row matrices, we may write

$$\boldsymbol{\psi}=\boldsymbol{\phi}\mathbf{T}, \tag{2.1}$$

where $\mathbf{T}$ is a rectangular $m\times n$ matrix, and the Fock-Dirac density becomes

$$\rho(1;1')=\sum_{i,j=1}^{m}\phi_i(1)\rho_{ij}\phi_j^*(1'). \tag{2.2}$$

The array of coefficients, $\boldsymbol{\rho}=[\rho_{ij}]$, is the density matrix referred to the discrete basis and is given by

$$\boldsymbol{\rho}=\mathbf{T}\mathbf{T}^\dagger. \tag{2.3}$$

A sufficient condition for stationary energy is that the columns of $\mathbf{T}$ are eigenvectors of

$$\mathbf{hc}=\epsilon\mathbf{c}, \tag{2.4}$$

where $\mathbf{h}$ is the Hamiltonian matrix with elements $h_{ij}=\int\phi_i^*\mathfrak{h}\phi_j dx$. This is the matrix counterpart of (1.5). These results may also be written

$$\mathbf{hT}=\mathbf{T}\boldsymbol{\epsilon},\quad \boldsymbol{\epsilon}=\text{diag}(\epsilon_1\epsilon_2\cdots\epsilon_n) \tag{2.5}$$

and it readily follows that

$$\mathbf{h}\boldsymbol{\rho}=\boldsymbol{\rho}\mathbf{h}. \tag{2.6}$$

Because the n orbitals are assumed orthonormal, $\boldsymbol{\rho}$ has the projection operator property of idempotency:

$$\boldsymbol{\rho}^2=\boldsymbol{\rho}. \tag{2.7}$$

Conversely, (2.6) and (2.7) are the necessary and sufficient conditions that E [given by (1.2)] is stationary against variation of the ψ's, subject to preservation of orthonormality (see, for example, reference 4). The dependence of $\mathbf{h}$ upon $\boldsymbol{\rho}$ will be expressed by writing $\mathbf{h}=\mathbf{h}(\boldsymbol{\rho})$ but the precise form of this dependence is irrelevant. We shall suppose that a first approximation

---

[6] P.-O. Löwdin, Quantum Chemistry Group, Uppsala University Technical Note No. 48, 1960 (unpublished).

[7] C. C. J. Roothaan, Revs. Modern Phys. **23**, 69 (1951).

is available and satisfies

$$\mathbf{h}_0\boldsymbol{\rho}_0=\boldsymbol{\rho}_0\mathbf{h}_0,\qquad \boldsymbol{\rho}_0^2=\boldsymbol{\rho}_0. \tag{2.8}$$

We then calculate

$$\mathbf{h}=\mathbf{h}(\boldsymbol{\rho}_0)=\mathbf{h}_0+\boldsymbol{\Delta}, \tag{2.9}$$

and wish to find a corresponding $\boldsymbol{\rho}$ satisfying (2.6) and (2.7).

First, we proceed in the manner of ordinary perturbation theory, imagining $\boldsymbol{\Delta}$ to contain a "perturbation parameter" λ ($\lambda\rightarrow 1$) and supposing

$$\boldsymbol{\rho}=\boldsymbol{\rho}_0+\boldsymbol{\rho}^{(1)}+\boldsymbol{\rho}^{(2)}+\cdots, \tag{2.10}$$

where $\boldsymbol{\rho}^{(k)}$ is the correction term containing λ^k. On inserting (2.10) in (2.6) and (2.7), and separating the orders, we obtain

$$\mathbf{h}_0\boldsymbol{\rho}_0-\boldsymbol{\rho}_0\mathbf{h}_0=0, \tag{A0}$$

$$\mathbf{h}_0\boldsymbol{\rho}^{(1)}-\boldsymbol{\rho}^{(1)}\mathbf{h}_0+\boldsymbol{\Delta}\boldsymbol{\rho}_0-\boldsymbol{\rho}_0\boldsymbol{\Delta}=0, \tag{A1}$$

$$\cdots$$

$$\mathbf{h}_0\boldsymbol{\rho}^{(k)}-\boldsymbol{\rho}^{(k)}\mathbf{h}_0+\boldsymbol{\Delta}\boldsymbol{\rho}^{(k-1)}-\boldsymbol{\rho}^{(k-1)}\boldsymbol{\Delta}=0, \tag{Ak}$$

and

$$\boldsymbol{\rho}_0^2=\boldsymbol{\rho}_0, \tag{B0}$$

$$\boldsymbol{\rho}_0\boldsymbol{\rho}^{(1)}+\boldsymbol{\rho}^{(1)}\boldsymbol{\rho}_0=\boldsymbol{\rho}^{(1)}, \tag{B1}$$

$$\cdots$$

$$\boldsymbol{\rho}_0\boldsymbol{\rho}^{(k)}+\boldsymbol{\rho}^{(1)}\boldsymbol{\rho}^{(k-1)}+\cdots+\boldsymbol{\rho}^{(k-1)}\boldsymbol{\rho}^{(1)}+\boldsymbol{\rho}^{(k)}\boldsymbol{\rho}_0=\boldsymbol{\rho}^{(k)}. \tag{Bk}$$

Equations (A0) and (B0) are satisfied according to (2.8). To obtain the kth order correction we must solve (Ak) and (Bk), but since these equations contain the preceding orders it is necessary to proceed step by step.

1st Order Correction

Since $\boldsymbol{\rho}_0$ is a projection operator it defines a certain n-dimensional subspace $\mathbf{S}_1$ in the space spanned by the m basis function: $\mathbf{1}-\boldsymbol{\rho}_0=\boldsymbol{\rho}_0'$ defines the complementary subspace $\mathbf{S}_2$ and we may use the identity,

$$\mathbf{M}=\boldsymbol{\rho}_0\mathbf{M}\boldsymbol{\rho}_0+\boldsymbol{\rho}_0\mathbf{M}\boldsymbol{\rho}_0'+\boldsymbol{\rho}_0'\mathbf{M}\boldsymbol{\rho}_0+\boldsymbol{\rho}_0'\mathbf{M}\boldsymbol{\rho}_0' =\mathbf{M}_{11}+\mathbf{M}_{12}+\mathbf{M}_{21}+\mathbf{M}_{22},$$

to resolve an operator into its projected components[8] in $\mathbf{S}_1$ and $\mathbf{S}_2$ ($\mathbf{M}_{11}$ and $\mathbf{M}_{22}$) and in the "intersections" ($\mathbf{M}_{12}$ and $\mathbf{M}_{21}$). It is readily seen that the components are linearly independent and that $\mathbf{A}=\mathbf{B}$ is equivalent to $\mathbf{A}_{ij}=\mathbf{B}_{ij}$ (i, $j=1$, 2). If we resolve $\boldsymbol{\rho}^{(1)}$ into its components, (B1) yields four equations, one for each component. Observing that $\mathbf{A}_{ij}\mathbf{B}_{kl}=0$ ($j\neq k$), these are

$$(\mathrm{B}1)_{11}:\ 2\boldsymbol{\rho}_{11}^{(1)}=\boldsymbol{\rho}_{11}^{(1)};\quad (\mathrm{B}1)_{12}:\ \boldsymbol{\rho}_{12}^{(1)}=\boldsymbol{\rho}_{12}^{(1)};$$

$$(\mathrm{B}1)_{21}:\ \boldsymbol{\rho}_{21}^{(1)}=\boldsymbol{\rho}_{21}^{(1)};\quad (\mathrm{B}1)_{22}:\ 0=\boldsymbol{\rho}_{22}^{(1)}.$$

Thus

$$\boldsymbol{\rho}_{11}^{(1)}=\boldsymbol{\rho}_{22}^{(1)}=0 \tag{2.11}$$

while $\boldsymbol{\rho}_{12}^{(1)}$ and $\boldsymbol{\rho}_{21}^{(1)}$ are arbitrary, though Hermitian symmetry requires

$$\boldsymbol{\rho}_{21}^{(1)}=\boldsymbol{\rho}_{12}^{(1)\dagger}. \tag{2.12}$$

This means that $\boldsymbol{\rho}_0+\boldsymbol{\rho}^{(1)}$ is idempotent to first order provided $\boldsymbol{\rho}^{(1)}$ is of the form

$$\boldsymbol{\rho}^{(1)}=\boldsymbol{\rho}_{12}^{(1)}+\boldsymbol{\rho}_{21}^{(1)}, \tag{2.13}$$

containing intersection terms only.

On substituting (2.13) into (A1) and resolving $\boldsymbol{\Delta}$ into its four components, we obtain

$$\mathbf{h}_0\boldsymbol{\rho}_{11}^{(1)}-\boldsymbol{\rho}_{11}^{(1)}\mathbf{h}_0=0, \tag{A1$_{11}$}$$

$$\mathbf{h}_0\boldsymbol{\rho}_{12}^{(1)}-\boldsymbol{\rho}_{12}^{(1)}\mathbf{h}_0-\boldsymbol{\Delta}_{12}=0, \tag{A1$_{12}$}$$

$$\mathbf{h}_0\boldsymbol{\rho}_{21}^{(1)}-\boldsymbol{\rho}_{21}^{(1)}\mathbf{h}_0+\boldsymbol{\Delta}_{21}=0, \tag{A1$_{21}$}$$

$$\mathbf{h}_0\boldsymbol{\rho}_{22}^{(1)}-\boldsymbol{\rho}_{22}^{(1)}\mathbf{h}_0=0. \tag{A1$_{22}$}$$

The first and last equations are automatically satisfied from (2.11) while $(\mathrm{A}1)_{12}$ and $(\mathrm{A}1)_{21}$ are Hermitian conjugate. We need determine, therefore, only $\boldsymbol{\rho}_{12}^{(1)}$. If we assume $\mathbf{h}_0$ nonsingular (which may always be ensured by suitable choice of energy zero), the inverse exists and $(\mathrm{A}1)_{12}$ may be written

$$\mathbf{x}=\mathbf{h}_0^{-1}\mathbf{x}\mathbf{h}_0+\mathbf{h}_0^{-1}\boldsymbol{\Delta}_{12}=\mathbf{f}(\mathbf{x}),\quad (\mathbf{x}=\boldsymbol{\rho}_{12}^{(1)}). \tag{2.14}$$

A formal solution may then be obtained by iteration. Thus if we take $\mathbf{x}_0=0$, $\mathbf{x}_1=\mathbf{f}(\mathbf{x}_0)=\mathbf{h}_0^{-1}\boldsymbol{\Delta}_{12}$ and in general

$$\mathbf{x}_{k+1}=\mathbf{f}(\mathbf{x}_k)=\mathbf{h}_0^{-1}\mathbf{x}_k\mathbf{h}_0+\mathbf{h}_0^{-1}\boldsymbol{\Delta}_{12} \tag{2.15}$$

then the limit of the sequence gives

$$\boldsymbol{\rho}_{12}^{(1)}=\sum_{n=0}^{\infty}\mathbf{h}_0^{-(n+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0^{n}. \tag{2.16}$$

We examine the convergence of the process in detail in Sec. 4. Here we note that for $n\rightarrow\infty$, $\mathbf{h}_0^n$ tends to a multiple of the projection operator for the eigenvector with numerically largest eigenvalue; if this is included among the *occupied* orbitals, which define $\boldsymbol{\rho}_0$, we shall have

$$\boldsymbol{\Delta}_{12}\mathbf{h}_0^n=\boldsymbol{\Delta}_{12}\boldsymbol{\rho}_0'\mathbf{h}_0^n\rightarrow 0,$$

since $\boldsymbol{\rho}_0'$ is the projection operator for the subspace spanned by the *un*occupied orbitals. Hence by choosing the energy zero so that the occupied orbitals correspond to numerically largest eigenvalues we can at least ensure that the individual terms in (2.16) tend to zero for $n\rightarrow\infty$.

2nd Order Correction

If we resolve (B2) into its four components, we obtain

$$(\mathrm{B}2)_{11}:\ 2\boldsymbol{\rho}_{11}^{(2)}+\boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\rho}_{21}^{(1)}=\boldsymbol{\rho}_{11}^{(2)};$$

$$(\mathrm{B}2)_{12}:\ \boldsymbol{\rho}_{12}^{(2)}=\boldsymbol{\rho}_{12}^{(2)};$$

$$(\mathrm{B}2)_{21}:\ \boldsymbol{\rho}_{21}^{(2)}=\boldsymbol{\rho}_{21}^{(2)};$$

$$(\mathrm{B}2)_{22}:\ \boldsymbol{\rho}_{21}^{(1)}\boldsymbol{\rho}_{12}^{(1)}=\boldsymbol{\rho}_{22}^{(2)};$$

---

[8] The notation should not be confused with that for matrix *elements*. $\mathbf{M}_{ij}$ is still an $m\times m$ matrix.

and, hence,

$$\begin{aligned}\boldsymbol{\rho}_{11}^{(2)} &= -\boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\rho}_{21}^{(1)},\\ \boldsymbol{\rho}_{22}^{(2)} &= \boldsymbol{\rho}_{21}^{(1)}\boldsymbol{\rho}_{12}^{(1)},\end{aligned} \tag{2.17}$$

while $\boldsymbol{\rho}_{12}^{(2)}$ and $\boldsymbol{\rho}_{21}^{(2)}$ are undetermined. The (11) and (22) components of the second-order correction are thus completely determined by the *first*-order correction, and it is only necessary to obtain $\boldsymbol{\rho}_{12}^{(2)}$. This follows from (A2) which gives

$$\mathbf{h}_0\boldsymbol{\rho}_{11}^{(2)} - \boldsymbol{\rho}_{11}^{(2)}\mathbf{h}_0 + \boldsymbol{\Delta}_{12}\boldsymbol{\rho}_{21}^{(1)} - \boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\Delta}_{21} = \mathbf{0}, \quad (\mathrm{A}2)_{11}$$

$$\mathbf{h}_0\boldsymbol{\rho}_{12}^{(2)} - \boldsymbol{\rho}_{12}^{(2)}\mathbf{h}_0 + \boldsymbol{\Delta}_{11}\boldsymbol{\rho}_{12}^{(1)} - \boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\Delta}_{22} = \mathbf{0}, \quad (\mathrm{A}2)_{12}$$

$$\mathbf{h}_0\boldsymbol{\rho}_{21}^{(2)} - \boldsymbol{\rho}_{21}^{(2)}\mathbf{h}_0 + \boldsymbol{\Delta}_{22}\boldsymbol{\rho}_{21}^{(1)} - \boldsymbol{\rho}_{21}^{(1)}\boldsymbol{\Delta}_{11} = \mathbf{0}, \quad (\mathrm{A}2)_{21}$$

$$\mathbf{h}_0\boldsymbol{\rho}_{22}^{(2)} - \boldsymbol{\rho}_{22}^{(2)}\mathbf{h}_0 + \boldsymbol{\Delta}_{21}\boldsymbol{\rho}_{12}^{(1)} - \boldsymbol{\rho}_{21}^{(1)}\boldsymbol{\Delta}_{12} = \mathbf{0}. \quad (\mathrm{A}2)_{22}$$

It is readily verified that the first and last equations are satisfied when $\boldsymbol{\rho}_{11}^{(2)}$ and $\boldsymbol{\rho}_{22}^{(2)}$ are chosen according to (2.17). Thus $(\mathrm{A}2)_{11}$ becomes

$$\begin{aligned}-\mathbf{h}_0\boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\rho}_{21}^{(1)} &+ \boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\rho}_{21}^{(1)}\mathbf{h}_0 + \boldsymbol{\Delta}_{12}\boldsymbol{\rho}_{21}^{(1)} - \boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\Delta}_{21}\\ &= -[\mathbf{h}_0\boldsymbol{\rho}_{12}^{(1)} - \boldsymbol{\Delta}_{12}]\boldsymbol{\rho}_{21}^{(1)} + \boldsymbol{\rho}_{12}^{(1)}[\boldsymbol{\rho}_{21}^{(1)}\mathbf{h}_0 - \boldsymbol{\Delta}_{21}]\\ &= -\boldsymbol{\rho}_{12}^{(1)}\mathbf{h}_0\boldsymbol{\rho}_{21}^{(1)} + \boldsymbol{\rho}_{12}^{(1)}\mathbf{h}_0\boldsymbol{\rho}_{21}^{(1)} = \mathbf{0},\end{aligned}$$

where we use the fact that $\boldsymbol{\rho}_{12}^{(1)}$ and $\boldsymbol{\rho}_{21}^{(1)}$ satisfy the first order equations $(\mathrm{A}1)_{12}$ and $(\mathrm{A}1)_{21}$. Equation $(\mathrm{A}2)_{12}$ then completely determines the second order correction. If we introduce

$$\mathbf{M}_{12}^{(1)} = \boldsymbol{\rho}_{12}^{(1)}\boldsymbol{\Delta}_{22} - \boldsymbol{\Delta}_{11}\boldsymbol{\rho}_{12}^{(1)}, \tag{2.18}$$

$(\mathrm{A}2)_{12}$ may be written

$$\mathbf{x} = \mathbf{h}_0^{-1}\mathbf{x}\mathbf{h}_0 + \mathbf{h}_0^{-1}\mathbf{M}_{12}^{(1)}, \quad (\mathbf{x} = \boldsymbol{\rho}_{12}^{(2)}) \tag{2.19}$$

which is of the same form as (2.14). The second-order correction is accordingly

$$\boldsymbol{\rho}_{12}^{(2)} = \sum_{n=0}^{\infty} \mathbf{h}_0^{-(n+1)}\mathbf{M}^{(1)}{}_{12}\mathbf{h}_0^{n}. \tag{2.20}$$

*k*th Order Correction

If we resolve (B*k*) into its components, it becomes clear that $\boldsymbol{\rho}_{11}^{(k)}$ and $\boldsymbol{\rho}_{22}^{(k)}$ are again determined entirely by the lower order corrections. Thus we find

$$\begin{aligned}\boldsymbol{\rho}_{11}^{(k)} &= -\sum_{i=1}^{k-1} \boldsymbol{\rho}_{11}^{(i)}\boldsymbol{\rho}_{11}^{(k-i)} - \sum_{i=1}^{k-1} \boldsymbol{\rho}_{12}^{(i)}\boldsymbol{\rho}_{21}^{(k-i)},\\ \boldsymbol{\rho}_{22}^{(k)} &= \sum_{i=1}^{k-1} \boldsymbol{\rho}_{22}^{(i)}\boldsymbol{\rho}_{22}^{(k-i)} + \sum_{i=1}^{k-1} \boldsymbol{\rho}_{21}^{(i)}\boldsymbol{\rho}_{12}^{(k-i)},\end{aligned} \tag{2.21}$$

while $\boldsymbol{\rho}_{12}^{(k)}$ and $\boldsymbol{\rho}_{21}^{(k)}$ are undetermined. It may be verified by induction that, when the lower order equations are satisfied, the equations $(\mathrm{A}k)_{11}$ and $(\mathrm{A}k)_{22}$ are also satisfied. It then remains to determine $\boldsymbol{\rho}_{12}^{(k)}$ from $(\mathrm{A}k)_{12}$. Again we introduce $\mathbf{M}_{12}^{(k)}$ as the (12)-component of the commutator $[\boldsymbol{\rho}^{(k)}\boldsymbol{\Delta} - \boldsymbol{\Delta}\boldsymbol{\rho}^{(k)}]$:

$$\mathbf{M}_{12}^{(k)} = [\boldsymbol{\rho}_{11}^{(k)}\boldsymbol{\Delta}_{11} + \boldsymbol{\rho}_{12}^{(k)}\boldsymbol{\Delta}_{22} - \boldsymbol{\Delta}_{11}\boldsymbol{\rho}_{12}^{(k)} - \boldsymbol{\Delta}_{12}\boldsymbol{\rho}_{22}^{(k)}], \tag{2.22}$$

and obtain by iteration

$$\boldsymbol{\rho}_{12}^{(k)} = \sum_{n=0}^{\infty} \mathbf{h}_0^{-(n+1)}\mathbf{M}_{12}^{(k-1)}\mathbf{h}_0^{n}. \tag{2.23}$$

From (2.21–23), the perturbation of the density matrix may be obtained to any desired order.

III. ENERGY FORMULAS

The conventional Rayleigh-Schrödinger perturbation theory applies to the individual eigenvectors and eigenvalues of (2.4). The preceding discussion reduces to this special case when there is just one occupied orbital and $\mathbf{T}$ in (2.5) is a single column $\mathbf{c}$: we then have $\epsilon = \mathbf{c}^{\dagger}\mathbf{h}\mathbf{c} = \mathrm{Tr}\,\mathbf{h}\boldsymbol{\rho}$ and may obtain the perturbed energy accordingly. In the present context, however, the orbitals in (1.3) have no individual significance, since the wave function is invariant under unitary mixing. Thus, new orbitals

$$\bar{\boldsymbol{\psi}} = \boldsymbol{\psi}\mathbf{U} = \boldsymbol{\phi}\mathbf{T}\mathbf{U}$$

give exactly the same Fock-Dirac density,

$$\bar{\boldsymbol{\rho}} = \mathbf{T}\mathbf{U}(\mathbf{T}\mathbf{U})^{\dagger} = \mathbf{T}\mathbf{T}^{\dagger} = \boldsymbol{\rho}.$$

From (2.5), the only invariant energy is then the *sum* of the *n* orbital energies,

$$\mathcal{E} = \sum_{i=1}^{n} \epsilon_i = \mathrm{Tr}\,\boldsymbol{\epsilon} = \mathrm{Tr}\,\mathbf{T}^{\dagger}\mathbf{h}\mathbf{T} = \mathrm{Tr}\,\mathbf{h}\boldsymbol{\rho}. \tag{3.1}$$

Since this quantity is of considerable importance in the independent-particle model, it is of interest to obtain the perturbation series

$$\mathcal{E} = \mathcal{E}_0 + \mathcal{E}^{(1)} + \mathcal{E}^{(2)} + \cdots. \tag{3.2}$$

From (3.1) it follows that

$$\begin{aligned}\mathcal{E} &= \mathrm{Tr}\,(\mathbf{h}_0 + \boldsymbol{\Delta})(\boldsymbol{\rho}_0 + \boldsymbol{\rho}^{(1)} + \boldsymbol{\rho}^{(2)} + \cdots)\\ &= \mathcal{E}_0 + \mathrm{Tr}\,\boldsymbol{\Delta}\boldsymbol{\rho}_0 + \mathrm{Tr}\,\boldsymbol{\Delta}\boldsymbol{\rho}^{(1)} + \mathrm{Tr}\,\mathbf{h}_0\boldsymbol{\rho}^{(2)} + \cdots,\end{aligned} \tag{3.3}$$

where we use the fact that $\mathrm{Tr}\,\mathbf{h}_0\boldsymbol{\rho}^{(1)} = 0$ since $\boldsymbol{\rho}^{(1)}$ contains only intersection terms and, for example

$$\mathrm{Tr}\,\mathbf{h}_0\boldsymbol{\rho}_0\boldsymbol{\rho}^{(1)}\boldsymbol{\rho}_0' = \mathrm{Tr}\,\boldsymbol{\rho}_0'\mathbf{h}_0\boldsymbol{\rho}_0\boldsymbol{\rho}^{(1)} = \mathrm{Tr}\,\boldsymbol{\rho}_0'\boldsymbol{\rho}_0\mathbf{h}_0\boldsymbol{\rho}^{(1)} = 0.$$

The first term in (3.3) is the unperturbed energy, the second involves the unperturbed wave function and corresponds to the usual first order energy perturbation:

$$\mathcal{E}^{(1)} = \mathrm{Tr}\,\boldsymbol{\Delta}\boldsymbol{\rho}_0. \tag{3.4}$$

The third term depends on the first-order change in wave function, but does not give the second order energy change unless the wave function is *renormalized* to second order: we shall see presently that the effect of renormalization is embodied in the fourth term.

First we consider $\mathrm{Tr}\,\boldsymbol{\Delta\rho}^{(1)}$:

$$\mathrm{Tr}\,\boldsymbol{\Delta\rho}^{(1)}=\mathrm{Tr}\,\boldsymbol{\Delta\rho}_{12}{}^{(1)}+\mathrm{Tr}\,\boldsymbol{\Delta\rho}_{21}{}^{(1)}$$

$$=\mathrm{Tr}\sum_{n=0}^{\infty}\boldsymbol{\Delta}\mathbf{h}_0{}^{-(n+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{n}$$

$$+\mathrm{Tr}\sum_{n=0}^{\infty}\boldsymbol{\Delta}\mathbf{h}_0{}^{n}\boldsymbol{\Delta}_{21}\mathbf{h}_0{}^{-(n+1)}.$$

If we make a cyclic permutation of the factors in each term, and use the projection operator properties of $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_0'$ (which both commute with $\mathbf{h}_0$), the two terms are seen to be identical and

$$\mathrm{Tr}\,\boldsymbol{\Delta\rho}^{(1)}=2\,\mathrm{Tr}\,\boldsymbol{\Delta\rho}_{12}{}^{(1)}=2\sum_{n=0}^{\infty}\mathrm{Tr}\,\boldsymbol{\Delta}_{21}\mathbf{h}_0{}^{-(n+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{n}. \quad (3.5)$$

In order to renormalize to second order we must admit those parts of $\boldsymbol{\rho}^{(2)}$ which are determined purely by the first order correction. These are given by (2.17) and the remaining terms, $\boldsymbol{\rho}_{12}{}^{(2)}$ and $\boldsymbol{\rho}_{21}{}^{(2)}$, are in any case irrelevant in second order because when taken with $\mathbf{h}_0$ each gives zero trace. The renormalization terms are thus

$$\mathrm{Tr}\,\mathbf{h}_0\boldsymbol{\rho}_{11}{}^{(2)}=-\mathrm{Tr}\sum_{n,m=0}^{\infty}\mathbf{h}_0{}^{-n}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{n}\mathbf{h}_0{}^{m}\boldsymbol{\Delta}_{21}\mathbf{h}_0{}^{-(m+1)}$$

$$=-\mathrm{Tr}\sum_{n,m=0}^{\infty}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{(n+m)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{-(n+m+1)},$$

$$\mathrm{Tr}\,\mathbf{h}_0\boldsymbol{\rho}_{22}{}^{(2)}=\mathrm{Tr}\sum_{n,m=0}^{\infty}\mathbf{h}_0{}^{m+1}\boldsymbol{\Delta}_{21}\mathbf{h}_0{}^{-(m+1)}\mathbf{h}_0{}^{-(n+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{n}$$

$$=\mathrm{Tr}\sum_{n,m=0}^{\infty}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{(n+m+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{-(n+m+2)}.$$

If we use $s_{m,n}$ to denote the general term in the first sum, it follows that

$$\mathrm{Tr}\,\mathbf{h}_0\boldsymbol{\rho}^{(2)}=-\Big[\sum_{m,n=0}^{\infty}s_{mn}-\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}s_{mn}\Big]=-\sum_{m=0}^{\infty}s_{m,0}.$$

Hence

$$\mathrm{Tr}\,\mathbf{h}_0\boldsymbol{\rho}^{(2)}=-\mathrm{Tr}\sum_{m=0}^{\infty}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{m}\boldsymbol{\Delta}_{21}\mathbf{h}_0{}^{-(m+1)}$$

$$=-\mathrm{Tr}\,\boldsymbol{\Delta}_{21}\sum_{m=0}^{\infty}\mathbf{h}_0{}^{-(m+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{m}$$

$$=-\mathrm{Tr}\,\boldsymbol{\Delta\rho}_{12}{}^{(1)}. \quad (3.6)$$

Renormalization thus exactly halves the leading contribution (3.5) to the second-order energy, and we obtain

$$\mathcal{E}^{(2)}=\mathrm{Tr}\,\boldsymbol{\Delta\rho}_{12}{}^{(1)}=\sum_{n=0}^{\infty}\mathrm{Tr}\,\boldsymbol{\Delta}_{21}\mathbf{h}_0{}^{-(n+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{n}. \quad (3.7)$$

It is also possible to obtain the third order term in simple form. The contribution from $\boldsymbol{\rho}^{(2)}$ is

$$\mathcal{E}^{(3)}=\mathrm{Tr}\,\boldsymbol{\Delta}(\boldsymbol{\rho}_{11}{}^{(2)}+\boldsymbol{\rho}_{22}{}^{(2)})$$

$$=\mathrm{Tr}\,\boldsymbol{\Delta}(\boldsymbol{\rho}_{21}{}^{(1)}\boldsymbol{\rho}_{12}{}^{(1)}-\boldsymbol{\rho}_{12}{}^{(1)}\boldsymbol{\rho}_{21}{}^{(1)}), \quad (3.8)$$

and it is easily seen that this is the *complete* third-order term. For the remaining contributions are

$$\mathrm{Tr}\,\boldsymbol{\Delta}(\boldsymbol{\rho}_{12}{}^{(2)}+\boldsymbol{\rho}_{21}{}^{(2)})+\mathrm{Tr}\,\mathbf{h}_0(\boldsymbol{\rho}_{11}{}^{(3)}+\boldsymbol{\rho}_{22}{}^{(3)}),$$

and by using (2.21) this may be reduced to

$$\mathrm{Tr}(\boldsymbol{\Delta}_{12}-\mathbf{h}_0\boldsymbol{\rho}_{12}{}^{(1)}+\boldsymbol{\rho}_{12}{}^{(1)}\mathbf{h}_0)\boldsymbol{\rho}_{21}{}^{(2)}$$

$$+\mathrm{Tr}(\boldsymbol{\Delta}_{21}-\boldsymbol{\rho}_{21}{}^{(1)}\mathbf{h}_0+\mathbf{h}_0\boldsymbol{\rho}_{21}{}^{(1)})\boldsymbol{\rho}_{12}{}^{(2)},$$

which vanishes by virtue of $(\mathrm{A}1)_{12}$ and $(\mathrm{A}1)_{21}$. Since the odd orders of perturbation theory give upper bounds to the eigenvalues, this is a particularly useful result.

IV. CONVERGENCE. CONNECTION WITH ORTHODOX PERTURBATION THEORY

We must now determine under what conditions the iterative solution converges. This may be done by writing $\mathbf{h}_0$ in "normal form" in terms of its eigenvalues and the projection operators for the eigenvectors. Thus

$$\mathbf{h}_0=\sum_{i=1}^{m}\epsilon_i\boldsymbol{\rho}_i, \quad (4.1)$$

where $\boldsymbol{\rho}_i=\mathbf{c}_i\mathbf{c}_i{}^{\dagger}$ is the projection operator for eigenvector $\mathbf{c}_i$. In general, we can then write

$$\mathbf{f}(\mathbf{h}_0)=\sum_{i=1}^{m}f(\epsilon_i)\boldsymbol{\rho}_i. \quad (4.2)$$

On using this form in (2.16), remembering that

$$\boldsymbol{\rho}_0=\sum_{i(\mathrm{occ})}\boldsymbol{\rho}_i \quad\text{and}\quad \boldsymbol{\rho}_0'=\sum_{i(\mathrm{unocc})}\boldsymbol{\rho}_i,$$

we obtain

$$\boldsymbol{\rho}_{12}{}^{(1)}=\sum_{n=0}^{\infty}\mathbf{h}_0{}^{-(n+1)}\boldsymbol{\Delta}_{12}\mathbf{h}_0{}^{n}$$

$$=\sum_{n=0}^{\infty}\sum_{j(\mathrm{occ})}\sum_{k(\mathrm{unocc})}\epsilon_j{}^{-(n+1)}\boldsymbol{\rho}_j\boldsymbol{\Delta\rho}_k\epsilon_k{}^{n}$$

$$=\sum_{n=0}^{\infty}\sum_{j(\mathrm{occ})}\sum_{k(\mathrm{unocc})}\left(\frac{\epsilon_k{}^{n}}{\epsilon_j{}^{n+1}}\right)\Delta_{jk}\mathbf{c}_j\mathbf{c}_k{}^{\dagger};$$

where

$$\Delta_{jk}=\mathbf{c}_j{}^{\dagger}\boldsymbol{\Delta}\mathbf{c}_k.$$

This expression is a geometric series with respect to n and converges absolutely so long as $\epsilon_k<\epsilon_j$ (all unoccupied k, all occupied j). This condition can always be met by suitably placing the energy zero (i.e., adding a multiple of the unit matrix to $\mathbf{h}_0$). We can then sum to infinity and obtain the closed form,

$$\boldsymbol{\rho}_{12}{}^{(1)}=\sum_{\substack{j\,(\mathrm{occ})\\k\,(\mathrm{unocc})}}\frac{1}{\epsilon_j-\epsilon_k}\boldsymbol{\Delta}_{jk}\mathbf{c}_j\mathbf{c}_k{}^{\dagger}. \tag{4.3}$$

It is easy to see how, in the case of a single eigenvector, this result gives the usual Rayleigh-Schrödinger series. For with just *one* occupied state (j) the j summation disappears and we obtain from (3.7)

$$\epsilon_j{}^{(2)}=\mathcal{E}^{(2)}=\mathrm{Tr}\,\boldsymbol{\rho}_{12}{}^{(1)}\boldsymbol{\Delta}=\sum_{k\,(\neq j)}\frac{1}{\epsilon_j-\epsilon_k}\Delta_{jk}\Delta_{kj},$$

the usual result. The third-order term (3.8) gives on introducing (4.3),

$$\mathcal{E}^{(3)}=\mathrm{Tr}\sum_{l\,(\neq j)}\sum_{k\,(\neq j)}\frac{1}{(\epsilon_j-\epsilon_l)}\frac{1}{(\epsilon_j-\epsilon_k)}\Delta_{jl}{}^{*}\mathbf{c}_l\mathbf{c}_j{}^{\dagger}\Delta_{jk}\mathbf{c}_j\mathbf{c}_k{}^{\dagger}\boldsymbol{\Delta}$$

$$-\mathrm{Tr}\sum_{l\,(\neq j)}\sum_{k\,(\neq j)}\frac{1}{(\epsilon_j-\epsilon_l)}\frac{1}{(\epsilon_j-\epsilon_k)}\Delta_{jl}\mathbf{c}_j\mathbf{c}_l{}^{\dagger}\Delta_{jk}{}^{*}\mathbf{c}_k\mathbf{c}_j{}^{\dagger}\boldsymbol{\Delta},$$

i.e.,

$$\epsilon_j{}^{(3)}=\sum_{l,k\,(\neq j)}\frac{\Delta_{jk}\Delta_{kl}\Delta_{lj}}{(\epsilon_j-\epsilon_k)(\epsilon_j-\epsilon_l)}-\sum_{l\,(\neq j)}\Delta_{jj}\frac{\Delta_{jl}\Delta_{lj}}{(\epsilon_j-\epsilon_l)^2}$$

which again agrees with the Rayleigh-Schrödinger result.

V. PERTURBATION THEORY WITHOUT SEPARATION OF THE ORDERS

For many purposes it is not necessary to determine separately each order in the perturbation series

$$\delta\boldsymbol{\rho}=\boldsymbol{\rho}^{(1)}+\boldsymbol{\rho}^{(2)}+\cdots.$$

We therefore consider the direct determination of $\delta\boldsymbol{\rho}$ by an iterative method similar to that already used in Sec. 3. From (2.6) we obtain

$$\mathbf{h}_0\delta\boldsymbol{\rho}-\delta\boldsymbol{\rho}\mathbf{h}_0+\boldsymbol{\Delta}(\boldsymbol{\rho}_0+\delta\boldsymbol{\rho})-(\boldsymbol{\rho}_0+\delta\boldsymbol{\rho})\boldsymbol{\Delta}=\mathbf{0}. \tag{5.1}$$

Now we know that in every order $\boldsymbol{\rho}_{11}{}^{(k)}$ and $\boldsymbol{\rho}_{22}{}^{(k)}$ are determined by the preceding orders, owing to the idempotency condition (2.7), and that only the mixed component $\boldsymbol{\rho}_{12}{}^{(k)}$ is free to be chosen so as to satisfy (Ak). In general then $\delta\boldsymbol{\rho}_{11}$ and $\delta\boldsymbol{\rho}_{22}$ must be determined by $\delta\boldsymbol{\rho}_{12}$ which must be chosen so as to satisfy the (12) projection of (5.1). This is found to be

$$(\mathbf{h}_0+\boldsymbol{\Delta}_{11})\delta\boldsymbol{\rho}_{12}-\delta\boldsymbol{\rho}_{12}(\mathbf{h}_0+\boldsymbol{\Delta}_{22})+\boldsymbol{\Delta}_{12}\delta\boldsymbol{\rho}_{22}-\delta\boldsymbol{\rho}_{11}\boldsymbol{\Delta}_{12}-\boldsymbol{\Delta}_{12}=\mathbf{0}, \tag{5.2}$$

or, if we put $\delta\boldsymbol{\rho}_{12}=\mathbf{x}$ and multiply by $(\mathbf{h}_0+\boldsymbol{\Delta}_{11})^{-1}$,

$$\mathbf{x}=(\mathbf{h}_0+\boldsymbol{\Delta}_{11})^{-1}\mathbf{x}(\mathbf{h}_0+\boldsymbol{\Delta}_{22})+\mathbf{M}_{12}=\mathbf{f}(\mathbf{x}), \tag{5.3}$$

where

$$\mathbf{M}_{12}=\boldsymbol{\Delta}_{12}+\delta\boldsymbol{\rho}_{11}\boldsymbol{\Delta}_{12}-\boldsymbol{\Delta}_{12}\delta\boldsymbol{\rho}_{22}=\mathbf{M}(\mathbf{x}), \tag{5.4}$$

since $\delta\boldsymbol{\rho}_{11}$ and $\delta\boldsymbol{\rho}_{22}$ are determined by the solution $\mathbf{x}$. We shall see, however, that $\mathbf{M}_{12}$ is only weakly dependent on $\mathbf{x}$: in this case, neglecting at first the variation of $\mathbf{M}_{12}$, we can use an iterative procedure based on (5.3) which is formally similar to (2.19). Thus

$$\mathbf{x}=\sum_{n=0}^{\infty}(\mathbf{h}_0+\boldsymbol{\Delta}_{11})^{-(n+1)}\mathbf{M}_{12}(\mathbf{h}_0+\boldsymbol{\Delta}_{22})^{n}. \tag{5.5}$$

Initially, we should take $\mathbf{M}_{12}=\boldsymbol{\Delta}_{12}$. After estimating $\mathbf{x}$, the correction terms in (5.4) could then be added to give an improved formula. It remains only to examine the dependence of $\delta\boldsymbol{\rho}_{11}$ and $\delta\boldsymbol{\rho}_{22}$ upon $\mathbf{x}$ (i.e., upon $\delta\boldsymbol{\rho}_{12}$), implied by the idempotency condition (2.7).

On putting $\boldsymbol{\rho}=\boldsymbol{\rho}_0+\delta\boldsymbol{\rho}$ and separating the four components of (2.7), we obtain

$$\begin{aligned}\delta\boldsymbol{\rho}_{11}{}^2+2\delta\boldsymbol{\rho}_{11}+\delta\boldsymbol{\rho}_{12}\delta\boldsymbol{\rho}_{21}&=\delta\boldsymbol{\rho}_{11},\\ \delta\boldsymbol{\rho}_{11}\delta\boldsymbol{\rho}_{12}+\delta\boldsymbol{\rho}_{12}\delta\boldsymbol{\rho}_{22}&=\mathbf{0},\\ \delta\boldsymbol{\rho}_{21}\delta\boldsymbol{\rho}_{11}+\delta\boldsymbol{\rho}_{22}\delta\boldsymbol{\rho}_{21}&=\mathbf{0},\\ \delta\boldsymbol{\rho}_{22}{}^2+\delta\boldsymbol{\rho}_{21}\delta\boldsymbol{\rho}_{12}&=\delta\boldsymbol{\rho}_{22}.\end{aligned} \tag{5.6}$$

We can solve the first and last equations formally by completing the square:

$$(\delta\boldsymbol{\rho}_{11}+\tfrac{1}{2}\boldsymbol{\rho}_0)^2=\tfrac{1}{4}\boldsymbol{\rho}_0{}^2-\delta\boldsymbol{\rho}_{12}\delta\boldsymbol{\rho}_{21},$$

$$(\delta\boldsymbol{\rho}_{22}-\tfrac{1}{2}\boldsymbol{\rho}_0{}')^2=\tfrac{1}{4}\boldsymbol{\rho}_0{}'^2-\delta\boldsymbol{\rho}_{21}\delta\boldsymbol{\rho}_{12},\quad(\boldsymbol{\rho}_0{}'=\mathbf{1}-\boldsymbol{\rho}_0).$$

Hence, putting $\delta\boldsymbol{\rho}_{12}=\mathbf{x}$ and taking appropriate signs of the square roots, we have

$$\begin{aligned}(\delta\boldsymbol{\rho}_{11}+\tfrac{1}{2}\boldsymbol{\rho}_0)&=\tfrac{1}{2}\boldsymbol{\rho}_0(\mathbf{1}-4\mathbf{x}\mathbf{x}^{\dagger})^{\frac{1}{2}}\boldsymbol{\rho}_0,\\ (\delta\boldsymbol{\rho}_{22}-\tfrac{1}{2}\boldsymbol{\rho}_0{}')&=-\tfrac{1}{2}\boldsymbol{\rho}_0{}'(\mathbf{1}-4\mathbf{x}^{\dagger}\mathbf{x})^{\frac{1}{2}}\boldsymbol{\rho}_0{}'.\end{aligned} \tag{5.7}$$

Since $\mathbf{x}$ is a small quantity, the roots may be expanded to give

$$\begin{aligned}\delta\boldsymbol{\rho}_{11}&=-[\mathbf{x}\mathbf{x}^{\dagger}+(\mathbf{x}\mathbf{x}^{\dagger})^2+2(\mathbf{x}\mathbf{x}^{\dagger})^3+\cdots],\\ \delta\boldsymbol{\rho}_{22}&=[\mathbf{x}^{\dagger}\mathbf{x}+(\mathbf{x}^{\dagger}\mathbf{x})^2+2(\mathbf{x}^{\dagger}\mathbf{x})^3+\cdots],\end{aligned} \tag{5.8}$$

and it is easily verified by substitution that these solutions also satisfy the second and third equations of (5.6). We also obtain a compact expression for $\mathbf{M}_{12}$ in (5.4):

$$\begin{aligned}\mathbf{M}_{12}&=\boldsymbol{\Delta}_{12}+\delta\boldsymbol{\rho}_{11}\boldsymbol{\Delta}_{12}-\boldsymbol{\Delta}_{12}\delta\boldsymbol{\rho}_{22}\\ &=(\delta\boldsymbol{\rho}_{11}+\tfrac{1}{2}\boldsymbol{\rho}_0)\boldsymbol{\Delta}_{12}-\boldsymbol{\Delta}_{12}(\delta\boldsymbol{\rho}_{22}-\tfrac{1}{2}\boldsymbol{\rho}_0{}');\end{aligned}$$

i.e.,

$$\mathbf{M}_{12}=\tfrac{1}{2}[\boldsymbol{\Delta}_{12}(\mathbf{1}-4\mathbf{x}^{\dagger}\mathbf{x})^{\frac{1}{2}}+(\mathbf{1}-4\mathbf{x}\mathbf{x}^{\dagger})^{\frac{1}{2}}\boldsymbol{\Delta}_{12}]. \tag{5.9}$$

If we took account of the $\mathbf{x}$ dependence of $\mathbf{M}_{12}$ in the iterative treatment of (5.3) we should obtain an exact solution. But $\mathbf{M}_{12}$ differs from $\boldsymbol{\Delta}_{12}$ only by terms of the third order in the perturbation parameter and if we put $\mathbf{M}_{12}=\boldsymbol{\Delta}_{12}$ in (5.5) we should expect to obtain the first two orders of perturbation theory exactly, together with substantial parts of all higher orders. A more refined iteration formula would result on estimating $\mathbf{x}$ in this way and then recalculating $\mathbf{M}_{12}$ by expansion: neglecting terms of fifth and higher order we have

$$\mathbf{M}_{12}\simeq\boldsymbol{\Delta}_{12}-\boldsymbol{\Delta}_{12}\mathbf{x}^{\dagger}\mathbf{x}-\mathbf{x}\mathbf{x}^{\dagger}\boldsymbol{\Delta}_{12}. \tag{5.10}$$

Finally, we note that in obtaining a *self-consistent* solution of (2.6) the dependence of $\mathbf{h}$ (and hence of $\boldsymbol{\Delta}=\mathbf{h}-\mathbf{h}_0$) upon ϱ would most conveniently be introduced during the course of iteration. Instead of obtaining an exact solution for the $\boldsymbol{\Delta}$ calculated initially, and subsequently repeating the whole process, $\boldsymbol{\Delta}$ would be recalculated periodically and the iteration formula (5.3) modified accordingly.

VI. CONNECTION WITH THE STEEPEST-DESCENT METHOD

An alternative method of solving this problem has already been given,[9] but does not permit the separation of orders and is therefore unsuitable when a perturbation development is required (e.g. in polarizability calculations where the term of second order in the field strength determines a susceptibility). In conclusion, we shall investigate the relationship between the two approaches.

Let us adopt a "mean-value" approximation in the normal form (4.2), writing

$$(\mathbf{h}_0^{-1})_{11}=\sum_{i\ \mathrm{occ}}\epsilon_i^{-1}\varrho_i\simeq\langle\epsilon_{\mathrm{occ}}\rangle^{-1}\sum_{i\ \mathrm{occ}}\varrho_i=\langle\epsilon_{\mathrm{occ}}\rangle^{-1}\varrho_0.$$

Similarly

$$(\mathbf{h}_0)_{22}=\sum_{i\ \mathrm{unocc}}\epsilon_i\varrho_i\simeq\langle\epsilon_{\mathrm{unocc}}\rangle\sum_{i\ \mathrm{unocc}}\varrho_i=\langle\epsilon_{\mathrm{unocc}}\rangle\varrho_0'.$$

Then, from (2.16),

$$\varrho_{12}{}^{(1)}\simeq\sum_{n=0}^{\infty}\frac{\langle\epsilon_{\mathrm{unocc}}\rangle^n}{\langle\epsilon_{\mathrm{occ}}\rangle^{n+1}}\boldsymbol{\Delta}_{12};$$

i.e.,

$$\varrho_{12}{}^{(1)}\simeq-\lambda\boldsymbol{\Delta}_{12},\tag{6.1}$$

where

$$\lambda=1/\Delta\epsilon\tag{6.2}$$

and $\Delta\epsilon$ is a "mean excitation energy." This gives an approximate first-order correction,

$$\varrho^{(1)}=-\lambda(\boldsymbol{\Delta}_{12}+\boldsymbol{\Delta}_{21}),\tag{6.3}$$

which is the result obtained by the method of steepest descents in the alternative approach.[9] In some applications it is possible to make a reasonable estimate of $\Delta\epsilon$, but generally it would be more satisfactory to regard λ as a variational parameter.

The solution of the basic equations (2.5) and (2.7) occurs[9] when $\mathcal{E}$, defined in (3.1), has a stationary minimum value. The third-order estimate based upon (3.4), (3.7), (3.8) is an upper bound[10] and we may therefore determine an optimum λ value by minimization. To third order we have

$$\mathcal{E}=\mathrm{Tr}\,\mathbf{h}[\varrho_0-\lambda(\boldsymbol{\Delta}_{12}+\boldsymbol{\Delta}_{21})+\lambda^2(\boldsymbol{\Delta}_{21}\boldsymbol{\Delta}_{12}-\boldsymbol{\Delta}_{12}\boldsymbol{\Delta}_{21})];$$

and hence

$$\lambda_{\mathrm{opt}}=-\frac{\mathrm{Tr}\,\mathbf{sh}}{2\,\mathrm{Tr}(\mathbf{s}^\dagger\mathbf{s}-\mathbf{ss}^\dagger)\mathbf{h}};\quad(\mathbf{s}=\boldsymbol{\Delta}_{21}).\tag{6.4}$$

This is the result obtained in an earlier paper.[9] On the other hand, it is possible to give an improved estimate of $\delta\varrho$ by including the remaining second order terms. For on using the mean energy approximation in (2.20), we obtain

$$\varrho_{12}{}^{(2)}=-\lambda\mathbf{M}_{12}=\lambda^2(\boldsymbol{\Delta}_{12}\boldsymbol{\Delta}_{22}-\boldsymbol{\Delta}_{11}\boldsymbol{\Delta}_{12}).\tag{6.5}$$

The approximate second-order correction is then

$$\varrho^{(2)}=\lambda^2(\boldsymbol{\Delta}_{21}\boldsymbol{\Delta}_{12}-\boldsymbol{\Delta}_{12}\boldsymbol{\Delta}_{21})+\lambda^2(\varrho_{12}{}^{(2)}+\varrho_{21}{}^{(2)}).\tag{6.6}$$

The second term in (6.6), calculated according to (6.5), gives an improvement on the results obtained by the steepest descent method.

ACKNOWLEDGMENTS

The perturbation method reported in this paper, which is now being applied to a variety of problems, was developed during the author's stay with the Quantum Chemistry Group, Uppsala University. It is a pleasure to express my warmest thanks to Professor P.-O. Löwdin, Director of the Group, for his generous hospitality, and to all members of the Group for the benefit of many discussions in a friendly and stimulating atmosphere.

---

[9] R. McWeeny, Proc. Roy. Soc. (London) **A235**, 496 (1956); **A237**, 355 (1956); **A241**, 239 (1957).

[10] This is irrespective of the choice of $\varrho_{12}{}^{(1)}$ because the third-order expression can be written in the form of an expectation value, $\mathcal{E}=\mathrm{Tr}\mathbf{h}\varrho$.