A Second-Quantization Approach to the Analytical Evaluation of Response Properties for Perturbation-Dependent Basis Sets

TRYGVE U. HELGAKER AND JAN ALMLÖF
Department of Chemistry, University of Oslo, Blindern, Oslo 3, Norway

Abstract
A general theory for response properties is presented which is applicable to perturbations affecting the molecular basis set, notably nuclear derivatives. A perturbation-independent Fock space is introduced, and the necessary reorthonormalization of a truncated basis set after a perturbation is explicitly incorporated in the Hamiltonian. Explicit formulas for MCSCF first- and second-order properties are presented, and some computational aspects are briefly discussed. A brief comparison with previous results is given.

1. Introduction
Recent years have seen a rapid development in techniques for the direct analytical evaluation of molecular energy derivatives using ab initio wave functions [1]. Techniques for the calculation of first derivatives (gradients) with respect to nuclear coordinates have been efficiently implemented at both the one-configurational [2, 3] and the multiconfigurational [4-6] SCF levels, and have proved an invaluable tool for the quantum-chemical study of molecular structure [7]. Very recently the gradient technique has been extended to configuration interaction wave functions [8-10]. Also in the case of second derivatives (force constants) significant progress has been made, and implementation has been reported for one-configurational [11-13] and two-configurational [14] SCF wave functions.

There already exists an extensive literature on the evaluation of molecular derivatives [1-28], including general formulas up to third order at the MCSCF level [28]. However, nearly all of this work has been done using a generalized Fock-operator formalism [29]. This is rather unfortunate in view of the rapid and important development of second-order MCSCF techniques [30-34], which are conveniently expressed in the language of second quantization [35]. Because these techniques rely on an order-by-order expansion of the energy in independent parameters, they are especially well suited for the calculation of response properties (e.g., nuclear derivatives), as pointed out by Dalgaard and Jørgensen [30]. However, the existing theory for response properties using second-quantization techniques [30, 35, 36] has up to very recently been restricted to perturbations which leave the metric unchanged, i.e., perturbations under which the basis set remains fixed in space. Unfortunately, this requirement is generally not met by the very important class of properties for which at least one of the perturbations involved constitutes a nuclear displacement. The purpose of the present paper...
is to provide a general theory, applicable also to perturbations that affect the molecular basis set.

An important feature of the present theory is that a close analogy is maintained with the theory for perturbations which do not affect the molecular basis set. This makes the theory more transparent. In fact, formulas for higher-order properties not given here, notably third derivatives, can easily be derived from the corresponding formulas previously given for perturbation-independent basis sets [36].

In Sec. 2 we give a brief discussion of basis set truncation and its important implications for the determination of molecular energy surfaces. In Sec. 3 we introduce the Fock space and the Hamiltonian in a form convenient for the present theory. Section 4 gives a brief summary of second-order MCSCF theory with special emphasis on the calculation of response properties, while detailed formulas needed for the evaluation of first- and second-order perturbation energies are presented in Sec. 5 and 6. Section 6 also provides a comparison with the recent results of Osamura, Yamaguchi, and Schaefer [10] and Pulay [28]. Section 7 contains some concluding comments and also a comparison with the very recent work of Jørgensen and Simons [37], whose paper puts forward a theory similar to the present one.

2. Truncation of Basis Sets

In virtually all quantum-chemical calculations one must in some way determine a set of molecular orbitals in terms of which the wave function can be constructed. Ideally, these orbitals should be expanded in a complete basis. However, to be able to handle the equations one must be content to work with a truncated set. To make the most out of the situation, these finite sets must be tailored to each molecular system in question.

To a large extent such procedures must necessarily be based on physical intuition, which unfortunately also implies a certain mathematical arbitrariness in the determination of the basis set parameters. One annoying problem is the determination of the coordinates of the atomic orbitals, which in practically all cases are clamped to the atomic nuclei. However, this computationally convenient procedure creates both conceptual and practical difficulties. For example, if we want to study the energy surface of a molecular system, we must explicitly or implicitly determine the energy for different values of the nuclear coordinates. At each point of the surface the energy is thus determined using a different basis set, an especially troublesome situation since we are nearly always interested in energy differences.

Closely related problems occur in the analytical evaluation of nuclear derivatives, where the complications are most clearly revealed in the fact that the Hellmann–Feynman relation does not hold for wave functions constructed according to the above procedure. From the practical side this means that the analytical calculation of derivatives becomes much more complicated than one would be inclined to believe considering the simplicity of the Hellmann–Feynman relation.
The most obvious way to avoid these difficulties would be to fix the atomic orbitals rigidly in space during the calculations. Unfortunately, this would have a disastrous effect on the calculated properties, which is precisely the reason why we chose to fix the orbitals on the nuclei in the first place.

A more serious alternative is to optimize the coordinates of the orbitals in each calculation [38]. Notwithstanding the computational burden involved, this possibility should be considered for the following reasons: First, the wave function will be more well defined in the sense that the Hellmann–Feynman relation will be satisfied. Second, the need for polarization functions might in some cases be eliminated, thus reducing the size of the basis set. Third, there is evidence that such functions will give considerably improved one-electron properties—more than one would expect on the basis of the rather modest improvements in energy [39]. The computational cost is comparable to that of a geometry optimization using the gradient technique, provided that the orbital and nuclear coordinates are optimized concurrently. However, this optimization of the orbital centers is no remedy for the problems discussed above, as the basis set is still a function of the nuclear coordinates.

In order to avoid some of the conceptual problems discussed in this section, we suggest a second-quantization formalism using a Fock space which is invariant to perturbations that affect the atomic orbitals and thus the metric. Having set up the corresponding Hamiltonian, the determination of the analytical formulas for nuclear derivatives becomes a straightforward task. Furthermore, in this formalism the Hellmann–Feynman relation is fulfilled for first-order properties, whereas second-order properties have two separate contributions as in the constant metric case: the expectation value of the second-order perturbation of the Hamiltonian and the relaxation energy.

3. The Structure of the Fock Space and the Hamiltonian

A. The Fock Space

When studying a system subject to a perturbation, one would prefer to work with a Fock space in terms of which the system can be described for all values of the perturbation. However, we have previously seen that when restricting ourselves to the use of finite basis sets, we can generally not assume that one particular set can describe the system throughout the perturbation. This implies that we in general will have to work with several Fock spaces (or rather, one perturbation-dependent Fock space) when studying one particular molecular process. As will be demonstrated below, this perturbation dependence of the Fock space may be disguised at the cost of a little complication in the Hamiltonian.

Let \( \mathbf{x} \) be a set of parameters describing the perturbation. For the unperturbed system (i.e., \( \mathbf{x} = \mathbf{x}_0 \)) we determine an orthonormal set of molecular orbitals (MOS) as a linear combination of atomic orbitals (AOS). In the coordinate representation we have

\[
\chi_i(\mathbf{r}; \mathbf{x}_0) = \sum_{\mu} c^0_{\mu i} \phi_{\mu}(\mathbf{r}; \mathbf{x}_0),
\]

(1)
where \( \mathbf{r} \) denotes the electron coordinates. Note that the atomic orbital coordinates are not equal to the parameters \( \mathbf{\bar{x}} \), but rather a function of them as indicated in Eq. (1). Orthonormality implies (assuming real quantities)

\[
S_{ij}(\mathbf{x}_0) = \int \chi_i(\mathbf{r}; \mathbf{x}_0) \chi_j(\mathbf{r}; \mathbf{x}_0) \, d\mathbf{r} = \sum_{\mu, \nu} c_{i\mu}^0 c_{j\nu}^0 \int \phi_{\mu}(\mathbf{r}; \mathbf{x}_0) \phi_{\nu}(\mathbf{r}; \mathbf{x}_0) \, d\mathbf{r} = \sum_{\mu, \nu} c_{i\mu}^0 c_{j\nu}^0 S_{\mu\nu}^{AO}(\mathbf{x}_0) = \delta_{ij}.
\]

(2)

After the perturbation has been applied, the metric of the atomic orbitals has changed and the orbitals expressed by the old MO coefficients \( c^0 \) are no longer orthonormal:

\[
S_{ij}(\mathbf{x}) = \sum_{\mu, \nu} c_{i\mu}^0 c_{j\nu}^0 S_{\mu\nu}^{AO}(\mathbf{x}) \neq \delta_{ij}.
\]

(3)

To avoid complications connected with the use of nonorthogonal orbitals, this set is now orthonormalized. This may be done in several ways \([40, 41]\), but to preserve a one-to-one correspondence between the orbitals before and after the perturbation we choose the symmetric orthonormalization

\[
\psi_{ij}(\mathbf{r}; \mathbf{x}) = \sum_j S_{ij}^{-1/2}(\mathbf{x}) \chi_j(\mathbf{r}; \mathbf{x}),
\]

(4)

where \( S_{ij}^{-1/2} \) is shorthand for \( (S^{-1/2})_{ij} \). A Fock space can now be constructed in the usual way using this orthonormalized set of orbitals. Consider two kets in this Fock space that correspond to two sets of occupied molecular orbitals (Slater determinants). It is obvious that the inner product between these kets is preserved throughout the perturbation. It is further clear that the inner product between any two linear combinations of such kets is unaffected by the perturbation. The effect of the perturbation is thus to create a unitary transformation of Fock spaces.

We may now introduce electron creation and annihilation operators in the usual way. These operators will satisfy the usual anticommutation relations for all values of the perturbation:

\[
[\hat{a}_i^\dagger(\mathbf{x}), \hat{a}_j^\dagger(\mathbf{x})]_\pm = 0,
\]

\[
[\hat{a}_i(\mathbf{x}), \hat{a}_j(\mathbf{x})]_\pm = 0,
\]

\[
[\hat{a}_i^\dagger(\mathbf{x}), \hat{a}_j(\mathbf{x})]_\pm = \delta_{ij}.
\]

(5)

Finally, it is easily seen that the perturbation does not affect the expectation value of these operators or any string of such operators.

In other words, we may consider these Fock spaces as one Fock space, whose kets and corresponding creation and annihilation operators are constant entities, not affected by the perturbation. However, this is no longer true if we combine entities corresponding to different values of the perturbation. This will not cause
any difficulties as we will restrict ourselves to variationally based methods, i.e.,
methods in which all properties are calculated as expectation values.

B. The Hamiltonian

Having set up our Fock space, we shall now proceed to determine the
appropriate Hamiltonian. We shall see that the Fock space as defined above
introduces certain complications in the Hamiltonian, the most conspicuous being
the dependence of the two-electron operator on the perturbation.

In terms of the orthonormalized set of orbitals \{\psi_i(\vec{r}; \vec{x})\} the Hamiltonian has the usual form

\[ \hat{H}(\vec{x}) = h_{ij}(\vec{x}) \hat{a}_i^\dagger \hat{a}_j + \frac{1}{2} g_{ijkl}(\vec{x}) \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k \hat{a}_l, \] (6)

where we have used the Einstein summation convention, and the summation is
over the finite set of molecular orbitals. Note that the dependence on the perturba-
tion lies only in the (super)matrices \( h'(\vec{x}) \) and \( g'(\vec{x}) \). We shall now determine these matrices.

In the one-electron case we have

\[ h_{ij}(\vec{x}) = \int \psi_i(\vec{r}; \vec{x}) \hat{h}(\vec{r}; \vec{x}) \psi_j(\vec{r}; \vec{x}) \ d\vec{r} \]

\[ = \sum_{\mu, \nu} \int \chi_\mu(\vec{r}; \vec{x}) \hat{h}(\vec{r}; \vec{x}) \chi_\nu(\vec{r}; \vec{x}) \ d\vec{r} \ S_{\mu i}^{1/2}(\vec{x}) S_{\nu j}^{-1/2}(\vec{x}) \] (7)

\[ = \sum_{\mu, \nu} h_{\mu \nu}(\vec{x}) S_{\mu i}^{1/2}(\vec{x}) S_{\nu j}^{-1/2}(\vec{x}), \]

where \( \hat{h}(\vec{r}; \vec{x}) \) is the usual one-electron Hamiltonian (excluding the nuclear repul-
sion term) in the coordinate representation. We have now expressed \( h'(\vec{x}) \) solely
in terms of matrix elements of the original unperturbed orbitals \( \{\chi_i(\vec{r}; \vec{x})\} \). The dependence of these elements on the perturbation arises only because the matrices
\( S^{AO}(\vec{x}) \) and \( h^{AO}(\vec{x}) \) depend on the perturbation. See Eq. (3) for the overlap matrix.

Similarly, the two-electron supermatrix elements may be written

\[ g_{ijkl}(\vec{x}) = \sum_{\mu, \nu, r, s} g_{\mu \nu rs}(\vec{x}) S_{\mu i}^{-1/2}(\vec{x}) S_{\nu j}^{-1/2}(\vec{x}) S_{r k}^{1/2}(\vec{x}) S_{s l}^{-1/2}(\vec{x}), \] (8)

where \( g_{\mu \nu rs}(\vec{x}) \) is the two-electron integral

\[ g_{\mu \nu rs}(\vec{x}) = \int \text{tr} \int \chi_\mu(\vec{r}; \vec{x}) \chi_\nu(\vec{r}; \vec{x}) \chi_r(\vec{r}_1; \vec{x}) \chi_s(\vec{r}_2; \vec{x}) \ d\vec{r}_1 \ d\vec{r}_2. \] (9)

The molecular Hamiltonian is thus given as

\[ \hat{H}(\vec{x}) = h_{\mu \nu}(\vec{x}) S_{\mu i}^{-1/2}(\vec{x}) S_{\nu j}^{-1/2}(\vec{x}) \hat{a}_i \hat{a}_j + \frac{1}{2} g_{\mu \nu rs}(\vec{x}) S_{\mu i}^{-1/2}(\vec{x}) S_{\nu j}^{-1/2}(\vec{x}) S_{r k}^{1/2}(\vec{x}) S_{s l}^{-1/2}(\vec{x}) \hat{a}_i \hat{a}_j \hat{a}_k \hat{a}_l, \] (10)

where we have simplified the notation for the electron creation and annihilation
operators.
C. Notation for Derivatives

At this point it is convenient to summarize the notation which will be used for derivatives. Total derivatives are denoted by simple superscripts:

\[ A^{\alpha}(\vec{x}) = \frac{dA(\vec{x})}{dx_\alpha}. \]  

(11)

Subscripts \( A_{\alpha}(\vec{x}) \) indicate that in the expression \( A(\vec{x}) \) the (super)matrices \( h(\vec{x}) \) and \( g(\vec{x}) \) are to be replaced by their total derivatives \( h^{\alpha}(\vec{x}) \) and \( g^{\alpha}(\vec{x}) \), for example,

\[ \hat{H}_{\alpha}(\vec{x}) = \hat{h}^{\alpha}_{\rho q}(\vec{x})S^{-1/2}_{\rho /}(\vec{x})S^{-1/2}_{\mu /}(\vec{x})i^{\mu}j + \frac{1}{2}g^{\alpha}_{\rho qr}(\vec{x})S^{-1/2}_{\rho /}(\vec{x})S^{-1/2}_{\mu /}(\vec{x})S^{-1/2}_{\nu /}(\vec{x})i^{\mu}j\]  

(12)

Finally, superscripts in parentheses \( A^{(\alpha)} \) will be used to indicate that the expression \( A^{(\alpha)} \) in some (obvious) way is related to differentiation with respect to \( x_\alpha \), being, for example, a function of total derivatives with respect to \( x_\alpha \).

It is important to note that the total derivatives of \( S(\vec{x}) \) [and likewise for \( h(\vec{x}) \) and \( g(\vec{x}) \)] are given by

\[ S^{\alpha}_{\mu}(\vec{x}) = \sum_{\mu'} c^{\alpha}_{\mu} c^{\mu'}_{\mu'} [S^{\mu\alpha}_{\mu'}(\vec{x})]^{\mu'} \]  

(13)

since the overlap matrix \( S(\vec{x}) \) is defined in terms of the unperturbed orbitals \( \epsilon_0 \).

The derivatives of the AO matrix elements are in most cases easily calculated.

We finally determine the first and second derivatives of \( S^{-1/2}(\vec{x}) \) for \( \vec{x} = \vec{x}_0 \). By writing

\[ S(\vec{x}) = 1 + \Delta(\vec{x}) \]  

(14)

we can expand \( S^{-1/2}(\vec{x}) \) in a Taylor series about \( \vec{x}_0 \):

\[ S^{-1/2}(\vec{x}) = 1 - \frac{1}{2} \Delta(\vec{x}) + \frac{3}{4} \Delta^2(\vec{x}) + O[\Delta^3(\vec{x})]. \]  

(15)

Differentiation with respect to \( x_\alpha \) and \( x_\beta \) gives

\[ [S^{-1/2}(\vec{x})]_{\alpha}^{\alpha} = -\frac{1}{2} [S^{\alpha}_{\alpha} + \Delta(\vec{x})]_{\alpha}, \]  

(16a)

\[ [S^{-1/2}(\vec{x})]^{\alpha\beta}_{\alpha\beta} = -\frac{1}{2} [S^{\alpha\beta}_{\alpha\beta} + \Delta(\vec{x})]_{\alpha\beta} + \frac{3}{4} [S^{\alpha\beta}_{\alpha\beta} + \Delta^2(\vec{x})]_{\alpha\beta}, \]  

(16b)

where we have used the fact that \( \Delta(\vec{x}_0) = 0 \) and \( \Delta^2(\vec{x})_{\alpha\beta} = \Delta^2(\vec{x})_{\beta\alpha} = S^{\alpha\beta}(\vec{x})_{\alpha\beta}. \)

4. The MCSCF Wave Function, Energy, and Response Functions

In this section we give a brief summary of the MCSCF theory with special emphasis on the calculation of response properties within the coupled Hartree-Fock scheme [30–36].

A. Wave Function and Energy

The multiconfigurational reference state \( |0\rangle \) is a normalized linear combination of orthonormal configuration state functions (CSFS, denoted by Greek letters):

\[ |0\rangle = \sum_{\mu} C^{(n)}_{\mu} |\mu\rangle, \langle 0|0\rangle = 1. \]  

(17)
Each CSF is a space- and spin-symmetrized electronic configuration constructed from a set of molecular orbitals. In the simplest case these configurations may be written as an ordered product of creation operators working on the vacuum ket:

$$|\mu\rangle = \prod_i |\mu_i\rangle_{\text{vac}}.$$  \hspace{1cm} (18)

The orthogonal complement of $|0\rangle$ is given as

$$|K\rangle = \sum_{\mu=1}^{N+1} C_\mu^{(K)}|\mu\rangle, \langle K|L\rangle = \delta_{KL}; K = 1, \ldots, N.$$  \hspace{1cm} (19)

We further need operators that perform arbitrary unitary transformations of the state functions and the orbitals when working on $|0\rangle$. These are given as

$$\exp (\hat{P}), \hat{P} = \sum_{K=0} P_{K0} (|K\rangle\langle 0| - |0\rangle\langle K|),$$

$$\exp (\hat{Q}), \hat{Q} = \sum_{i>j} Q_{ij} (i^+j - j^+i),$$

where redundant parameters have been eliminated from Eq. (20b). The parameters \{P_{K0}: K \neq 0\} and \{Q_{ij}: i > j\} may be collected in column vectors $\hat{P}$ and $\hat{Q}$, which may be further combined to one vector $\hat{\lambda}$:

$$\hat{\lambda} = \begin{pmatrix} \hat{P} \\ \hat{Q} \end{pmatrix}.$$  \hspace{1cm} (21)

Note that \{P_{KL}\} and \{Q_{ij}\} can also be considered elements of real antisymmetric matrices. A unitary transformation of the reference state is then given by

$$|\tilde{0}\rangle = \exp (\hat{Q}) \exp (\hat{P})|0\rangle.$$  \hspace{1cm} (22)

We emphasize that the reference state $|0\rangle$ is independent of the perturbation—the variation of the wave function during the perturbation is described by the operators $\exp (\hat{P})$ and $\exp (\hat{Q})$.

The energy functional $\epsilon(\tilde{x}, \tilde{\lambda})$ may now be expanded in a Taylor series to second order in the independent parameters $\tilde{\lambda}$:

$$\epsilon(\tilde{x}, \tilde{\lambda}) = \langle 0 | \exp (-\hat{P}) \exp (-\hat{Q}) \hat{H}(\tilde{x}) \exp (\hat{Q}) \exp (\hat{P}) | 0 \rangle$$

$$= \langle 0 | \hat{H}(\tilde{x}) | 0 \rangle + \langle 0 | [\hat{H}(\tilde{x}), \hat{P} + \hat{Q}] | 0 \rangle$$

$$+ \frac{1}{2} \langle 0 | [\hat{P}, \hat{P}, \hat{H}(\tilde{x})] | 0 \rangle + \frac{1}{2} \langle 0 | [\hat{Q}, \hat{Q}, \hat{H}(\tilde{x})] | 0 \rangle$$

$$+ \langle 0 | [\hat{P}, [\hat{Q}, \hat{H}(\tilde{x})]] | 0 \rangle + O(\tilde{\lambda}^3).$$

\hspace{1cm} (23)

We have here introduced the symmetric double commutator

$$[\hat{A}, \hat{B}, \hat{C}] = \frac{1}{2} [\hat{A}, [\hat{B}, \hat{C}]] + \frac{1}{2} [\hat{B}, [\hat{A}, \hat{C}]]$$

\hspace{1cm} (24)

to ensure that the Hessian matrix given below is Hermitian. If we introduce the gradient $\tilde{\tilde{f}}(\tilde{x})$ and the Hessian matrix $G(\tilde{x})$, Eq. (23) may be written as

$$\epsilon(\tilde{x}, \tilde{\lambda}) = E_0(\tilde{x}) + \tilde{\tilde{f}}(\tilde{x})\tilde{\lambda} + \frac{1}{2} \tilde{\lambda} G(\tilde{x})\tilde{\lambda} + O(\tilde{\lambda}^3).$$

\hspace{1cm} (25)
Explicit formulas for the gradient and the Hessian are obtained by substituting the Hamiltonian (6) and the operators (20) in Eq. (23). For the two parts of the gradient we find

$$V_K(\tilde{x}) = 2\langle K | \hat{H}(\tilde{x}) | 0 \rangle,$$

$$W_{ij}(\tilde{x}) = 2[ F_{ji}(\tilde{x}) - F_{ij}(\tilde{x}) ],$$

where $F(\tilde{x})$ is the generalized Fock matrix [32]

$$F_{ij}(\tilde{x}) = D_{ij} h'_{ji}(\tilde{x}) + 2 P_{ijk\gamma} g'_{ij\gamma}(\tilde{x}),$$

$$D_{ij} = \langle 0 | i^* j | 0 \rangle,$$

$$P_{ijkl} = \frac{1}{2} \langle 0 | i^* j^* k l | 0 \rangle.$$ 

In a similar way the Hessian is divided into four blocks:

$$G_{KL}^{(CC)}(\tilde{x}) = 2[ \langle K | \hat{H}(\tilde{x}) | L \rangle - \delta_{KL} \langle 0 | \hat{H}(\tilde{x}) | 0 \rangle ], \quad K, L = 1, \ldots, N,$$

$$G_{ijkl}^{(CO)}(\tilde{x}) = w_{ijkl}(\tilde{x}) - w_{ijlk}(\tilde{x}) - w_{jikl}(\tilde{x}) + w_{ijk}(\tilde{x}),$$

$$G_{ijkl}^{(OC)}(\tilde{x}) = 2[ F_{ji}^{0K}(\tilde{x}) + F_{ji}^{K0}(\tilde{x}) - F_{ji}^{0K}(\tilde{x}) - F_{ij}^{0K}(\tilde{x}) ], \quad K = 1, \ldots, N.$$ 

In the above equations we have introduced the supermatrix [32]

$$w_{ijkl}(\tilde{x}) = \delta_{ik} [ F_{kl}(\tilde{x}) + F_{kl}(\tilde{x}) ] + 2 D_{ik} h'_{kl}(\tilde{x})$$

$$+ 4[ P_{ik\alpha\beta} g'_{j\alpha\beta}(\tilde{x}) + P_{ik\alpha\beta} g'_{j\alpha\beta}(\tilde{x}) + P_{i\alpha\beta k} g'_{j\alpha\beta}(\tilde{x}) ]$$

and $F^{K0}(\tilde{x})$, defined as $F(\tilde{x})$ but with transition density matrices replacing ordinary density matrices

$$D_{ij}^{K0} = \langle K | i^* j | 0 \rangle,$$

$$P_{ijkl}^{K0} = \frac{1}{2} \langle K | i^* j^* k l | 0 \rangle.$$ 

The energy $E(\tilde{x})$ is obtained by minimizing $\epsilon(\tilde{x}, \tilde{\lambda})$ with respect to $\tilde{\lambda}$. The first derivative of the energy functional is given as

$$de(\tilde{x}, \tilde{\lambda})/d\tilde{\lambda} = \tilde{f}(\tilde{x}) + G(\tilde{x}) \tilde{\lambda} + O(\tilde{\lambda}^2).$$

When this is set to zero, $\tilde{\lambda}$ is determined as a function of $\tilde{x}$. The energy $E(\tilde{x})$ is thus defined by the two sets of equations:

$$E(\tilde{x}) = E_0(\tilde{x}) + \tilde{f}(\tilde{x}) \tilde{\lambda}(\tilde{x}) + \frac{3}{2} \lambda(\tilde{x}) G(\tilde{x}) \tilde{\lambda}(\tilde{x}) + O(\tilde{\lambda}^3),$$

$$G(\tilde{x}) \tilde{\lambda}(\tilde{x}) = -\tilde{f}(\tilde{x}) + O(\tilde{\lambda}^2).$$ 

These equations must be solved iteratively, usually neglecting higher-order terms (Newton’s method). In each iteration the quantities $E_0(\tilde{x})$, $\tilde{f}(\tilde{x})$, and $G(\tilde{x})$ are recalculated using the wave function obtained in the previous iteration. On convergence

$$\tilde{\lambda}(\tilde{x}) = \tilde{f}(\tilde{x}) = 0.$$ 

This disappearance of the gradients $\tilde{V}(\tilde{x})$ and $\tilde{W}(\tilde{x})$ given by Eq. (26) is referred to as the generalized Brillouin theorem (GBT) [35].
B. Response Properties

We now turn to the MCSCF response properties. Henceforth we tacitly assume that $\tilde{x} = \tilde{x}_0$ and that the reference state $|0\rangle$ has been optimized for the unperturbed system.

Differentiation of Eq. (32a) with respect to $x_a$ gives

$$E^a = E_0^a + \tilde{f} \tilde{x}^a + O(\tilde{x}).$$

(34)

For the optimized wave function Eq. (33) holds, and Eq. (34) reduces to the simple relation

$$E^a = E_0^a.$$

(35)

For the second-order properties we obtain for the optimized wave function

$$E^{ab} = E_0^{ab} + \tilde{f}^a \tilde{x}^b + \tilde{f}^b \tilde{x}^a + \frac{1}{2} \tilde{x}^a G \tilde{x}^b + \frac{1}{2} \tilde{x}^b G \tilde{x}^a.$$

(36)

The vectors $\tilde{x}^a$ and $\tilde{x}^b$ are obtained by differentiating Eq. (32b) and using Eq. (33):

$$\tilde{f}^a + G \tilde{x}^a = 0$$

(37)

and the same for $\tilde{x}^b$. Using Eq. (37) it is possible to eliminate any of the quantities $\tilde{f}^a$, $G$, or $\tilde{x}^a$ from Eq. (36). Eliminating $\tilde{x}^a$ and $\tilde{x}^b$ we find

$$E^{ab} = E_0^{ab} - \tilde{f}^a G^{-1} \tilde{f}^b,$$

(38)

where we have used the fact that $G$ is Hermitian. Note that the same Hessian is used for all perturbations. These are the coupled multiconfigurational Hartree-Fock (CMCHF) equations given elsewhere [30, 35, 36].

Considering the potentially huge dimensions of the Hessian matrix, the working CMCHF equations are better represented by

$$E^{ab} = E_0^{ab} + \tilde{f}^a \tilde{x}^b,$$

(39a)

$$G \tilde{x}^b = -\tilde{f}^b.$$

(39b)

These equations are still not suitable for large-scale calculations, as the configuration parts of the gradient and the Hessian are written in terms of the (unknown) eigenvectors of the $N$-dimensional orthogonal complement of the reference state. However, they may easily be transformed to a CSF basis using a method developed by Lengsfield [33] for solving Eqs. (32).

For this purpose we introduce the augmented Hessian matrix $G_{aug}$ which has exactly the same structure as $G$ [see Eq. (28)], but is defined for all the $(N + 1)$ eigenvectors $|0\rangle, |1\rangle, \ldots, |N\rangle$. Due to the generalized Brillouin theorem

$$G_{0k}^{(CC)} = V_k = 0,$$

(40a)

$$G_{0ij}^{(CO)} = 2 W_{ij} = 0,$$

(40b)

for an optimized wave function. Noting that $G_{00}^{(CC)} = 0$, we obtain

$$G_{aug} = \begin{pmatrix} 0 & \tilde{0} \\ \tilde{0} & G \end{pmatrix}.$$  

(41)
Equation (39b) can therefore be written in the equivalent form
\[
\left[ \begin{array}{cc}
\tilde{z} & \tilde{\Theta} \\
0 & 0 \\
\end{array} \right] + G^{\text{aug}} \left[ \begin{array}{c}
\chi \\
r_b \\
\end{array} \right] = \left[ \begin{array}{c}
0 \\
r_b \\
\end{array} \right],
\]
(42)
where \( z \neq 0 \) to ensure a nonsingular matrix. The \((N+1)\text{C}_1\) eigenvectors define a unitary matrix
\[
U_{\text{C}_1} = [\tilde{C}^{(0)}, C] = [\tilde{C}^{(0)}, \tilde{C}^{(1)}, \ldots, \tilde{C}^{(N)}]
\]
(43)
and we introduce the projectors
\[
R = \tilde{C}^{(0)}\tilde{C}^{(0)}, \quad R' = 1 - R.
\]
(44a)
(44b)
The transformation to the \text{CSF} basis may now be performed using the matrix
\[
U = \begin{pmatrix} U_{\text{C}_1} & 0 \\ 0 & 1 \end{pmatrix}
\]
(45)
giving the final equations
\[
\begin{pmatrix}
zR + G^{(\text{CC})}_{\text{CSF}} \\
G^{(\text{OC})}_{\text{CSF}} \\
G^{(\text{CO})}_{\text{CSF}} \end{pmatrix} \begin{pmatrix}
\tilde{C}^{b} \\
\tilde{Q}^{b} \\
\tilde{W}^{b} \end{pmatrix} = -\begin{pmatrix}
R\tilde{V}^{b}_{\text{CSF}} \\
\tilde{P}^{b} \\
\tilde{V}^{b} \end{pmatrix},
\]
(46)
where the matrices \(G_{\text{CSF}}\) are defined as in Eq. (28) with \( \tilde{x} = \tilde{x}_0 \), but in terms of the \text{CSFs}. These equations may be solved by the conjugate gradient method [42] or any similar method [11,43] for the solution of large linear equations.

Finally, the second-order perturbation energy may be calculated using Eq. (39a) which now takes the form
\[
E^{ab} = E_0^{ab} + (R\tilde{V}^{a}_{\text{CSF}})\tilde{C}^{b} + \tilde{W}^{a}\tilde{Q}^{b}.
\]
(47)
Hence we see that the second-order response properties can be calculated without any reference to the orthogonal complement of \text{C}_1\ eigenvectors.

In Sec. 5 and 6 we will give the explicit formulas needed for the calculation of \( E^a \) and \( E^{ab} \) using the Hamiltonian (10) and the methods described in Sec. 3.

5. First-Order Response Properties

In this section we will consider first-order response properties in some detail to illustrate the technique involved. According to Eq. (35) the gradient is given by
\[
E^a = \langle 0|\hat{H}|0\rangle^a.
\]
(48)
The expectation value of the Hamiltonian as given by Eq. (6) is
\[
\langle 0|\hat{H}(\tilde{x})|0\rangle = h'_0(\tilde{x})\langle 0|i'j'0\rangle + \frac{1}{2}g'_{ijkl}(\tilde{x})\langle 0|i'j'lk0\rangle
\]
(49)
and according to the discussion in Sec. 3 we have
\[
\langle 0|i'j'0\rangle^a = 0, 
\]
(50a)
\[
\langle 0|i'j'lk0\rangle^a = 0. 
\]
(50b)
This implies that the wave function in the chosen representation obeys the Hellmann–Feynman relation

\[ E^a = \langle 0 | \hat{H}^a | 0 \rangle \]

\[ = (h^a)_{ij} \langle 0 | i^j | 0 \rangle + \frac{1}{2} (g^a_{ijkl})^2 \langle 0 | i^j j^k | 0 \rangle, \]

(51)
i.e., the gradient is given as the expectation value of the differentiated Hamiltonian.

To calculate this expectation value directly we need the one- and two-electron integral derivatives. Differentiating Eqs. (7) and (8) we find

\[ (h^a)_{ij} = s^a_{ij} - \frac{1}{2} \sum_p \left( s^a_{ip} s^a_{pj} + s^a_{pj} s^a_{ip} \right) \]

(52a)

\[ (g^a_{ijkl})^2 = g^a_{ijkl} - \frac{1}{2} \sum_p \left( g^a_{ipkj} s^a_{pj} + g^a_{ipkl} s^a_{kj} + g^a_{ipjl} s^a_{kp} + g^a_{ipjl} s^a_{ik} \right), \]

(52b)

where we have used Eq. (16a) and the fact that \( S(\tilde{\alpha}) = 1 \). The one-index transformations above are very similar to those used by Jørgensen, Olsen, and Yeager [44] for calculating cubic contributions to the Taylor expansions (32), and very recently by Jensen and co-workers [45, 46] in connection with “direct MCSCF” techniques. In addition we notice that once the integrals (52) have been calculated, gradients of Hamiltonian matrix elements in CSF basis

\[ \langle \mu | \hat{H} | \nu \rangle = \langle \mu | \hat{H}^a | \nu \rangle \]

(53)

may be calculated using integral driven techniques. This will prove useful for calculating second and higher derivatives.

However, if we are only interested in the gradient of a single electronic state (e.g., the ground state), the above procedure is not optimal as it requires integral derivatives in MO basis. To be able to express the gradient in a form more suitable for single-state calculations, it is convenient to rewrite the first-order perturbation of the Hamiltonian.

According to Eq. (52a) the one-electron part of the differentiated Hamiltonian is

\[ \hat{H}_i^a = (h^a_{ij} - \frac{1}{2} s^a_{ip} s^a_{pj} - \frac{1}{2} h^a_{ip} s^a_{pj} ) i^j, \]

(54)

where once again we have suppressed the summation sign. If we assume real quantities, this expression may be rewritten in the form

\[ \hat{H}_i^a = h^a_{ij} i^j j - \frac{1}{2} s^a_{ij} h^a_{ij} (i^j j^i j + j^i i^j j). \]

(55)

By rearranging the two-electron part in the same way and adding it to Eq. (55) we obtain

\[ \hat{H}^a = \hat{H}_a - \frac{1}{2} S_{ij}^a [ h_{ij} (i^j j^i j + g_{ijkl} i^j j^l k + k^l i^j j i) ]. \]

(56)

Here we have used the notation

\[ \hat{H}_a = h^a_{ij} i^j j + \frac{1}{2} g^a_{ijkl} i^j j^l k \]

(57)
introduced in Sec. 3. The expectation value of the operator (56) is

$$
\langle 0 | \hat{H}^a | 0 \rangle = \langle 0 | \hat{H}_a | 0 \rangle - S_{ip}^a (D_{ij} h_{pj} + 2 P_{ijkl} g_{ijkl}).
$$

(58)

This gives us the desired expression for the gradient

$$
E^a = E_a - \text{Tr} (S^a F)
$$

(59)

which may further be rewritten in terms of AO integrals. This expression may be compared with previous results. Using the terminology introduced by Pulay [1] the first term includes both the Hellmann–Feynman and integral forces, while the second is the density force. In our representation it may be looked upon as a reorthonormalization term of the Hamiltonian.

Thus we see that the gradient may be calculated in two different ways. The first involves a direct calculation of the integral derivatives (implicit reorthonormalization), while the second involves a separate calculation of the reorthonormalization term (explicit reorthonormalization). For MCSCF gradients the latter is preferable as it allows the calculation to be carried out in an AO basis.

6. Second-Order Response Properties

A. Formulas for Second-Order Properties

We now turn to the evaluation of second-order properties according to Eqs. (46) and (47). For convenience we first summarize all the pertinent formulas. Notation—when not selfexplanatory—will be explained later.

$$
E^{ab} = E^{ab}_0 - 2(M^{(a)})^{(b)} - 2(A^{(a)})^{(b)},
$$

(60a)

$$
\left( \begin{array}{c}
A^{(CC)} \\
A^{(CO)} \\
A^{(OO)}
\end{array} \right) \left( \begin{array}{c}
\hat{X}^{(b)} \\
\hat{Y}^{(b)} \\
\hat{N}^{(b)}
\end{array} \right) = \left( \begin{array}{c}
M^{(b)} \\
F^{(b)} \\
N^{(b)}
\end{array} \right),
$$

(60b)

$$
E^{ab}_0 = E_{ab} - \text{Tr} [S^a F_b + S^b F_a + (S^{ab} - S^a S^b) F] + \frac{1}{4} \sum_{ijkl} S_{ij}^a w_{ijkl} S_{kl}^b
$$

(61)

$$
M^{(a)}_{(\mu)} = -\langle \mu | \hat{H}^a | 0 \rangle - C^{(0)}_{(\mu)} | 0 \rangle \hat{H}^a | 0 \rangle
$$

$$
= -E^{a0}_{(\mu)} + \frac{1}{2} \text{Tr} [S^a (F^{0\mu} + F^{a0})] + C^{(0)}_{(\mu)} E^a,
$$

(62a)

$$
N_{(a)}^{(\mu)} = (F^0_{(a)})_{ij} - (F^{a0})_{ji}
$$

$$
= (F_{(a)})_{ij} - (F^{0a})_{ji} - \sum_{k,l} (w_{ijkl} - w_{jikl}) S_{kl}^a,
$$

(62b)

$$
A^{(CC)}_{(\mu\nu)} = \langle \mu | \hat{H} | \nu \rangle - \delta_{\mu\nu} E^{a0} + z C^{(0)}_{(\mu)} C^{(0)}_{(\nu)},
$$

(63a)

$$
A^{(CO)}_{ijkl} = \frac{1}{2} (w_{ijkl} - w_{klji} - w_{ijlk} + w_{jikl}),
$$

(63b)

$$
A^{(OO)}_{(a\mu)(b\nu)} = F_{(a\mu)}^{0b} + F_{(a\nu)}^{b0} - F_{(a\mu)}^{b0} - F_{(a\nu)}^{0b}.
$$

(63c)

In the above equations the Hessian matrix and the differentiated gradients have been multiplied by a factor $\frac{1}{4}$ for convenience. Where two alternative expressions appear, the first involves implicit reorthonormalization while the second is the
BASIS SETS

equivalent expression using explicit reorthonormalization. The derivation of Eqs.
(60)–(63) is in all cases straightforward, although in some cases tedious. We will
now treat the separate cases.

\( E_{ab}^{(a)} \) may obviously be determined in a way similar to \( E_{00}^{(a)} \), i.e., as the expectation
value of the second-order perturbation of the Hamiltonian

\[
\langle 0 | \hat{H} | 0 \rangle^{ab} = \langle 0 | \hat{H}^{ab} | 0 \rangle.
\]  

By straightforward differentiation we arrive at Eq. (61), which gives the second-
order variation in the energy prior to the relaxation of the wave function. We note
that this term might be treated using implicit reorthonormalization by first
calculating the integral derivatives obtained by differentiating Eqs. (7) and (8)
twice. However, use of Eq. (61) is advantageous, as it avoids transformation of
the second integral derivatives to MO basis.

We now determine the differentiated gradients. The configuration part is given by

\[
\tilde{M} = -\frac{1}{2} R \tilde{V}^a_{\text{CSF}},
\]  

where \( (V^a_{\text{CSF}})_\mu = 2\langle \mu | \hat{H}^a | 0 \rangle \), see Eq. (26a). Inserting this expression and using
(44b) we obtain

\[
M^{(a)}_\mu = -\langle \mu | \hat{H}^a | 0 \rangle + C^{(0)}_\mu \langle 0 | \hat{H}^a | 0 \rangle.
\]  

In analogy with the expression for the gradient \( E^{(a)} \) we find

\[
\langle \mu | \hat{H}^a | 0 \rangle = E^{(0)}_\mu - \frac{1}{2} \text{Tr} [S^a (F^{(0)} + F^{(0)} - F_{\mu 0}^a)],
\]  

where as before the superscript \( \mu 0 \) indicates that transition density matrices
replace ordinary density matrices in the expressions. Combining Eqs. (66) and
(67) we arrive at Eq. (62a). It is obvious that for large CI expansions these elements
are best calculated using implicit reorthonormalization, as this allows us to employ
direct CI techniques. Moreover, in many cases the MO set may be divided into
inactive, active, and secondary orbitals. In such cases only differentiated two-
electron integrals within the relatively small active space are needed in the direct
CI scheme [32]. This fortunate fact will drastically reduce the number of integral
derivatives to be calculated. The remaining integrals either do not contribute at
all to the CI gradient (all indices within the inactive space), or enter the equations
through a one-electron potential [e.g., the inactive Fock matrix of Ref. 32]. The
latter integrals are conveniently handled using explicit reorthonormalization.

The orbital part of the differentiated gradient is [see Eq. (26b)]

\[
N^{(a)}_i = -\frac{1}{2} W^a_{ij} = (F^a_{0j})_i - (F^a_0)_ij.
\]  

The subscript \( \phi \) has been added to remind us that this Fock matrix has been
defined in terms of the unperturbed reference state \( |0\rangle \), which is independent of
the perturbation. [The total derivative \( F^{(0)}_{\phi i} \), of a Fock matrix defined in terms of
a wave function which is reoptimized for each value of the perturbation would
be Hermitian (GBT) and not equal to \( F^a_0 \)—compare the difference between \( E^{ab} \)
and \( E_0^{ab} \).] After straightforward differentiation of Eq. (27a) we obtain

\[
(F_0^a \delta)_{ij} = (F_a^a \delta)_{ij} - \frac{1}{4} \sum_{k,l} w_{ijkl} S^a_{kl}.
\]  

Combining Eqs. (68) and (69) we arrive at Eq. (62b). As the Fock matrix involves only the reference state it is advantageous to use explicit reorthonormalization for the orbital part of the differentiated gradient. Various Fock matrices have been introduced in different contexts, e.g., the inactive and active Fock matrices of Ref. 32. However, the total derivatives are always given by equations analogous to Eq. (69), obtained by differentiation of the given Fock matrix.

Returning to Eq. (60) we see that the second-order properties have a very simple interpretation: \( E^{ab} \) consists of a Hellmann–Feynman-like term (which includes reorthonormalization of the Hamiltonian) and a relaxation term. This simplicity is made possible by use of second quantization, in which the Hamiltonian is written explicitly in terms of the molecular basis set. In this way the “unphysical” need for reorthonormalization of the incomplete basis set can be incorporated in the Hamiltonian.

B. Comparison with Previous Results

Expressions for MCSCF energy derivatives up to third order have recently been presented by Pulay [28]. Also, the coupled MCSCF equations have been given by Osamura, Yamaguchi, and Schaefer [10] for applications to CI gradients. It is therefore of some interest to compare their results with the present work.

The approach of the two above-mentioned references is different from ours in that it is based on a formalism in which the wave function is determined by constrained optimization of dependent parameters. In contrast, we have explicitly derived equations for the variation of the independent parameters \( \tilde{P} \) and \( \tilde{Q} \). We shall now see how this is reflected in a small but interesting difference between the coupled equations presented above and the equations of Refs. 10 and 28.

For this purpose we rewrite Eq. (60b) slightly by subtracting from both sides the vector \( \tilde{T} \) given by

\[
\tilde{T} = \frac{1}{2} (A_{(CO)}^{(CO)}) S^a,
\]  

where \( S^a \) is the vector consisting of the elements \( \{S^a_{ij}, i > j\} \). This leads to a new set of linear equations

\[
\begin{pmatrix}
A^{(CC)} & A^{(CO)} \\
A^{(OC)} & A^{(OO)}
\end{pmatrix}
\begin{pmatrix}
\tilde{X}^{(a)} \\
\tilde{U}^{(a)}
\end{pmatrix}
= \begin{pmatrix}
\tilde{M}^{(a)}_t \\
\tilde{N}^{(a)}_t
\end{pmatrix},
\]  

where the differentiated gradient is now given by

\[
(M_t^{(a)})_\mu = -E_{\mu 0} + C^{(0)}_\mu E^a + \sum_{k \neq l} \left( 1 - \frac{1}{2} \delta_{kl} \right) (F_{k\mu}^{0} + F_{l\mu}^{0}) S^a_{kl},
\]  

\[
(N_t^{(a)})_{ij} = (F_a)_{ij} - (F_a)_{ji} - \frac{1}{2} \sum_{k \neq l} \left( 1 - \frac{1}{2} \delta_{kl} \right) (w_{ijkl} - w_{jikl}) S^a_{kl},
\]
The orbital part of the solution of Eq. (71) is related to the old solution by
\[ \mathbf{U}^{(a)} = \mathbf{Y}^{(a)} - \frac{1}{2} \mathbf{S}^a, \]  
while the configuration part is unaffected.

The set of equations (71) is identical to the coupled equations presented by Osamura et al. [10]. Except for the term \( z R \) in the (CC) block of the Hessian matrix [see Eq. (63a)], our results are also equivalent to those presented by Pulay [28]. This term has been introduced to ensure a unique solution of the coupled MCSCF equations. Absence of this term would lead to solutions containing the reference state vector \( \mathbf{C}^{(0)} \) as a component. However, this would not affect the final results as given by Eq. (60a) since \( \mathbf{C}^{(0)} \) and \( \mathbf{M}^{(a)} \) are orthogonal.

By introducing the diagonal elements \( \{+S^a_{\alpha}\} \) we can write Eq. (73) in matrix form
\[ \mathbf{U}^{(a)} = \mathbf{Q}^a - \frac{1}{2} \mathbf{S}^a, \]  
where \( \mathbf{Q}^a \) is antisymmetric and \( \mathbf{S}^a \) symmetric. The variation in the molecular orbitals may now be written
\[ c^{(a)} = c^0 U^{(a)}. \]  
In Ref. 28 \( c^{(a)}, \mathbf{X}^{(a)}, \) and the first derivatives of the Fock matrix \( F_{\alpha \beta}^a \) are all explicitly determined from the coupled MCSCF equations. The second derivative \( E_{\alpha \beta} \) is then obtained by an equation which is equivalent to Eq. (60a), but which has a very different structure corresponding to a different partitioning of the energy variation. However, it can be shown by a straightforward but tedious comparison that the two expressions are indeed equivalent.

7. Final Comments

We have presented a general analysis of response properties with emphasis on a strategy for full-scale applications. Our work will hopefully improve the understanding of MCSCF response properties. It further brings the notoriously troublesome properties related to nuclear displacement on an equal footing with constant-metric properties.

In this context we would like to draw attention to the important work by Jørgensen and Simons [37], which appeared upon the completion of the present article. Using techniques similar to ours, they have treated molecular gradients and Hessians for a variety of wave functions. However, while Jørgensen and Simons have employed symmetrically orthonormalized atomic orbitals, we have preferred to work in a symmetrically orthonormalized MO basis. In our opinion, the latter approach offers a more direct connection to previous work. We also feel that it offers certain advantages from a computational viewpoint.

Acknowledgments

We would like to thank H. J. Jensen and P. Jørgensen for placing the results of Ref. 45 at our disposal prior to publication. Many helpful discussions with P. Jørgensen and A. Heiberg are also acknowledged.
Bibliography


Received July 8, 1983
Accepted for publication November 8, 1983