

Gradient Theory

Trygve Helgaker

University of Oslo, Norway

1	Introduction	1157
2	Molecular Properties Calculated as Derivatives	1157
3	Numerical Versus Analytical Evaluation of Molecular Properties	1159
4	Molecular Properties for Fully Variational Wavefunctions	1159
5	Energy Derivatives for SCF Wavefunctions	1161
6	Second-quantization Representation of Derivatives	1163
7	The Molecular Gradient for Nonvariational Wavefunctions	1164
8	Conclusions	1166
9	Appendix: Elements of Second Quantization	1167
10	Related Articles	1168
11	References	1169

1 INTRODUCTION

The efficient evaluation of derivatives of the electronic energy of molecular systems is important in most applications of *ab initio* theory to problems of chemical interest. For example, studies of chemical structure and reactivity require the identification and characterization of critical points on the Born–Oppenheimer potential-energy surface, which for all but the simplest molecules cannot be carried out without an efficient means of extracting the molecular gradient (i.e., the forces acting on the nuclei) from the electronic wavefunction. Other examples are encountered in the study of spectroscopic constants, many of which may be identified as the derivatives of the molecular electronic energy with respect to parameters such as the nuclear positions, externally applied electric or magnetic field vectors, the nuclear magnetic dipole moments, and the electric quadrupole moments.

In this article, we consider the evaluation of derivatives of the electronic energy for molecular systems. Our emphasis throughout will be on principles rather than details, although in a few cases enough details will be provided to give the reader some understanding of the technical considerations associated with the efficient evaluation of molecular properties from *ab initio* wavefunctions. Also, apart from a short introductory section, the present article does not cover any applications of derivative theory – that is, we do not discuss in depth the use or usefulness of the derivatives, nor do we consider the relationship of the calculated derivatives to experimental measurements and the accuracy that may be expected in the evaluation of the derivatives for the standard wavefunctions and basis sets of *ab initio* theory. For such information, the reader is referred to *Benchmark Studies on Small Molecules; Molecular Magnetic Properties; NMR Chemical Shift Computation: Ab Initio*; and *Spectroscopy: Computational Methods*. For a more detailed exposition of gradient theory and for more complete bibliographies than that attempted here, the reader is referred to more complete reviews and to conference proceedings.^{1–6} Parts of this article assume some knowledge of second quantization as applied to electronic structure theory.⁷ For those

unfamiliar with this particular formalism, the essential elements of second quantization are described in the Appendix (Section 9).

2 MOLECULAR PROPERTIES CALCULATED AS DERIVATIVES

When a molecular electronic system is perturbed in some manner, its total electronic energy changes

$$E(\boldsymbol{\mu}) = E^{(0)} + \boldsymbol{\mu}^T \mathbf{E}^{(1)} + \frac{1}{2} \boldsymbol{\mu}^T \mathbf{E}^{(2)} \boldsymbol{\mu} + O(\boldsymbol{\mu}^3) \quad (1)$$

The coefficients $\mathbf{E}^{(n)}$ of this expansion describe the response of the molecular system to the external perturbation and are known as molecular properties. The molecular properties are characteristic of the molecular system and its quantum state. When the perturbation is static, the properties may be calculated by differentiation at $\boldsymbol{\mu} = 0$

$$\mathbf{E}^{(1)} = \left. \frac{dE}{d\boldsymbol{\mu}} \right|_0 \quad (2)$$

$$\mathbf{E}^{(2)} = \left. \frac{d^2E}{d\boldsymbol{\mu}^2} \right|_0 \quad (3)$$

and are then referred to as time independent or static. Static molecular properties play an important role in many areas of chemical research, as illustrated by the following list of examples.

2.1 Electric Properties

In the presence of a uniform electrostatic field, the energy may be expanded in the field strength \mathbf{F} in the following manner:

$$E(\mathbf{F}) = E_0 - \mathbf{F}^T \mathbf{d}_0 - \frac{1}{2} \mathbf{F}^T \boldsymbol{\alpha} \mathbf{F} + O(\mathbf{F}^3) \quad (4)$$

where \mathbf{d}_0 is the permanent molecular *dipole moment* and $\boldsymbol{\alpha}$ is the *dipole-polarizability* tensor at zero field:

$$\mathbf{d}_0 = - \left. \frac{dE}{d\mathbf{F}} \right|_0 \quad (5)$$

$$\boldsymbol{\alpha} = - \left. \frac{d^2E}{d\mathbf{F}^2} \right|_0 \quad (6)$$

The permanent molecular dipole moment and the polarizability tensor are the central parameters that characterize the interaction of a molecular system with an external electric field, generated by neighboring molecules or in an experimental apparatus. Other electric properties are also important, however. For example, expanding the energy to higher orders in the field strength \mathbf{F} , we obtain the molecular hyperpolarizability tensors of different orders, which are needed for an accurate description of the system in strong fields. Interactions with nonuniform electric fields arise from the presence of molecular quadrupole and higher multipole moments. The quadrupole moment, which is particularly important for nonpolar systems, may be calculated as the derivative of the energy with respect to the electric field gradient.

2.2 Molecular Forces and Force Constants

In the Born–Oppenheimer approximation, the nuclei move (classically speaking) on the electronic *potential-energy surface* $E(\mathbf{x})$, which is a function of the nuclear geometry \mathbf{x} . Expanding the electronic energy around some reference geometry \mathbf{x}_0 , we obtain

$$E(\mathbf{x}) = E^{(0)} + \Delta\mathbf{x}^T \mathbf{F} + \frac{1}{2} \Delta\mathbf{x}^T \mathbf{G} \Delta\mathbf{x} + O(\Delta\mathbf{x}^3) \quad (7)$$

where the nuclear displacements are given by

$$\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0 \quad (8)$$

The lowest-order molecular properties now correspond to the molecular gradient \mathbf{F} and the molecular Hessian \mathbf{G} :

$$\mathbf{F} = \left. \frac{dE}{d\mathbf{x}} \right|_{\mathbf{x}_0} \quad (9)$$

$$\mathbf{G} = \left. \frac{d^2E}{d\mathbf{x}^2} \right|_{\mathbf{x}_0} \quad (10)$$

calculated at $\mathbf{x} = \mathbf{x}_0$. The molecular gradient contains the negative molecular forces acting on the atomic nuclei, whereas the Hessian contains the quadratic force constants. Together with the energy at the expansion point $E^{(0)}$, the molecular gradient and the Hessian define a harmonic potential-energy surface. For a more accurate representation of the true surface, we must also take into account its anharmonicity – that is, the cubic, quartic, and possibly higher order force constants, which may be obtained by taking higher derivatives of the energy with respect to the nuclear geometry.

In molecular electronic-structure theory, there are two main uses of the geometrical derivatives: (1) the localization and characterization of critical points on the energy surface, in particular minima and saddle points (corresponding to molecular equilibrium configurations and transition states); and (2) the calculation of rovibrational spectroscopic constants and energy levels. For the efficient *optimization* of energy surfaces, the molecular gradient is a *sine qua non* – without an efficient means of calculating forces from the wavefunction, it becomes impractical to determine the molecular equilibrium structure of any but the smallest systems. Information about the molecular Hessian is usually required as well, but may, for minimizations at least, be obtained by comparing the gradients generated at different points along the optimization path. Higher than second derivatives are only very seldom required for optimizations but are needed for the accurate calculation of vibrational energy levels. Although harmonic frequencies may be adequate for qualitative investigations of molecular vibrations, cubic and quartic force constants are needed for quantitative comparison with experiment.

2.3 Magnetic Properties

As a final example of molecular properties, consider the molecular electronic system in the presence of a static external magnetic induction \mathbf{B} and nuclear magnetic moments \mathbf{M}_K , corresponding to the physical situation encountered in an NMR experiment. Expanding the energy of a closed-shell electronic system in the induction and in the nuclear magnetic moments,

we obtain

$$E(\mathbf{B}, \mathbf{M}_K) = E_0 + \frac{1}{2} \mathbf{B}^T \mathbf{E}^{(20)} \mathbf{B} + \sum_K \mathbf{B}^T \mathbf{E}_K^{(11)} \mathbf{M}_K + \frac{1}{2} \sum_{K \neq L} \mathbf{M}_K^T \mathbf{E}_{KL}^{(02)} \mathbf{M}_L \quad (11)$$

Only the second-order terms are included here – the first-order terms vanish for *closed-shell systems*, and the higher order terms are exceedingly small (because of the smallness of the magnetic perturbations) and can also be neglected. Of the three second-order terms in equation (11), the first represents the direct interaction of the molecular system with the external field and is represented by the magnetizability tensor at zero field:

$$\xi = \mathbf{E}^{(20)} = - \left. \frac{d^2E}{d\mathbf{B}^2} \right|_0 \quad (12)$$

However, the magnetizability tensor does not enter the effective NMR spin Hamiltonian, from which the NMR spectrum may be generated. The next term in the expansion (11) measures the coupling of the magnetic moments \mathbf{M}_K to the external magnetic induction \mathbf{B} :

$$\mathbf{E}_K^{(11)} = \left. \frac{d^2E}{d\mathbf{M}_K d\mathbf{B}} \right|_0 = -1 + \sigma_K \quad (13)$$

This coupling of the nuclei to the field is trivially equal to -1 in vacuum (the nuclear Zeeman interaction) but it is modified in the presence of electrons as described by the nuclear magnetic shielding tensor σ_K . The last term in equation (11) measures the coupling between the nuclear magnetic moments (and, therefore, between the nuclear spins):

$$\mathbf{E}_{KL}^{(02)} = \left. \frac{d^2E}{d\mathbf{M}_K d\mathbf{M}_L} \right|_0 = \frac{\mu_0}{4\pi} \frac{R_{KL}^2}{R_{KL}^5} (1 - 3\mathbf{R}_{KL} \mathbf{R}_{KL}^T) + \mathbf{K}_{KL} \quad (14)$$

which in vacuum is equal to the classical dipole interaction between the magnetic moments \mathbf{M}_K and \mathbf{M}_L (the first term in equation 14), but in the presence of the electrons is modified by the (reduced) indirect nuclear spin–spin coupling tensor \mathbf{K}_{KL} . The constant μ_0 in equation (14) is the permeability constant. For a freely rotating molecule, the direct dipole–dipole interaction vanishes and the interactions between the nuclear magnetic moments are completely described by the indirect nuclear spin–spin coupling tensor \mathbf{K}_{KL} .

The preceding examples, which represent only a small fraction of the large number of molecular properties studied experimentally and theoretically, should suffice to illustrate the important relationship that exists between the static molecular properties and derivatives of the molecular electronic energy. In the following, we shall discuss the evaluation of such properties with emphasis on the first- and second-order derivatives. Although our presentation will be in terms of geometrical perturbations, the methods developed are generally applicable. There are, however, two important restrictions on the applicability of the presented theory – it is inadequate for the calculation of time-dependent (dynamic) properties and non-adiabatic properties. For the calculation of such properties, generalizations of the present theory are required, as described

in *Nonadiabatic Derivative Couplings* and *Time-dependent Multiconfigurational Hartree Method*.

3 NUMERICAL VERSUS ANALYTICAL EVALUATION OF MOLECULAR PROPERTIES

There are two different approaches that may be taken to the calculation of static molecular properties: the energy derivatives may be calculated numerically or analytically. The numerical procedure involves the evaluation of derivatives by finite differences or polynomial fitting; the analytical procedure involves the calculation of derivatives directly from analytical expressions. The analytical approach requires considerable programming effort but offers greater speed, precision, and convenience than does the numerical approach, which may experience difficulties related to numerical stability and computational inefficiency. The numerical approach is simple in the sense that, at the level of electronic structure theory, it does not usually (but not always) require special programming – we may simply repeat the calculation of the energy for different values of the perturbational parameter. (As an example of a property that cannot easily be calculated by finite differences, we note that the calculation of indirect spin–spin coupling constants by numerical differentiation would involve the evaluation of the electronic energy in the presence of paramagnetic electrons and nuclei – a nontrivial task compared with the standard evaluation of the electronic energy for a spin-free, nonrelativistic electronic Hamiltonian.) Thus, for most properties of general interest and importance in quantum chemistry, the analytical approach is the preferred one; this is especially true for the molecular gradient, the analytical evaluation of which is vastly superior to the numerical approach for any but the smallest systems.

4 MOLECULAR PROPERTIES FOR FULLY VARIATIONAL WAVEFUNCTIONS

In general, we will write the electronic energy function in the form $E(\mathbf{x}, \boldsymbol{\lambda})$, where \mathbf{x} is a set of external parameters that characterize the physical system and $\boldsymbol{\lambda}$ is a set of *wavefunction* or electronic parameters that determine the electronic state. We shall think of the external parameters \mathbf{x} as representing the molecular geometry, but note that the results obtained here hold for other perturbations as well. The electronic parameters $\boldsymbol{\lambda}$ may correspond (directly or indirectly) to the MOs, to the CI coefficients, to the CC amplitudes, or to any other set of parameters in terms of which the wavefunction is expressed.

In the present section, we assume that the electronic energy is fully variational with respect to the electronic parameters $\boldsymbol{\lambda}$. Thus, we shall assume that the electronic energy may be calculated from the expression

$$E(\mathbf{x}) = E(\mathbf{x}; \boldsymbol{\lambda}^*) \quad (15)$$

where the parameters $\boldsymbol{\lambda}^*$ represent the optimal value of $\boldsymbol{\lambda}$ and where the optimized energy function $E(\mathbf{x}; \boldsymbol{\lambda}^*)$ satisfies the variational conditions for all values of the external

parameters

$$\left. \frac{\partial E(\mathbf{x}; \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \right|_* = 0 \quad \text{for all } \mathbf{x} \quad (16)$$

where the partial derivatives are calculated at $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$. To ensure that the variational conditions (16) are always fulfilled, the electronic parameters must change in a particular manner as the molecule deforms. The variational conditions therefore implicitly determine the dependence of the electronic parameters $\boldsymbol{\lambda}(\mathbf{x})$ on \mathbf{x} .

At this point, a few remarks on the meaning of the term ‘fully variational’ are in order. We take the term fully variational to imply that the conditions (16) are satisfied for all the electronic parameters that define the wavefunction. This term is introduced to avoid confusion with the term ‘variational’ in the looser sense of ‘obtained by application of the *variational principle*.’ As an example, the energy of a truncated CI expansion is variational in the sense that the CI coefficients have been obtained by the application of the variational principle in such a way that the calculated ground-state energy represents an upper bound to the true energy. Nevertheless, the truncated-CI energy is not fully variational since the conditions (16) hold only for variations in the CI coefficients and not for variations in the MOs.

Since many of the wavefunction models in quantum chemistry are not fully variational, it would seem that the theory described in the present section is of limited practical interest. We will find later, however, that the principles and techniques developed here for fully variational wavefunctions may be modified and extended to all other wavefunctions. The results obtained in the present section are therefore of general interest and should be understood before the evaluation of derivatives for energies of nonvariational wavefunctions is attempted.

4.1 Molecular Gradients

Let us consider the molecular gradient for the optimized variational electronic energy $E(\mathbf{x})$ (equation 15). Using the chain rule, we obtain

$$\frac{dE(\mathbf{x})}{d\mathbf{x}} = \left. \frac{\partial E(\mathbf{x}; \boldsymbol{\lambda})}{\partial \mathbf{x}} \right|_* + \left. \frac{\partial E(\mathbf{x}; \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \right|_* \frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{x}} \quad (17)$$

where the differentiation is carried out at $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$. The first term on the right-hand side represents the explicit dependence of the electronic-energy function on \mathbf{x} and arises, for example, from the dependence of the Hamiltonian on the geometry; the second term represents the implicit dependence of the energy function on \mathbf{x} and arises, for example, since the MO coefficients or cluster amplitudes change as the molecule deforms. The derivatives of the electronic parameters with respect to the external parameters $\partial \boldsymbol{\lambda} / \partial \mathbf{x}$ tell us how, to first order, the wavefunction changes when the perturbation is applied.

Combining equations (16) and (17) we obtain the following simple expression for the molecular gradient for a fully variational wavefunction:

$$\frac{dE(\mathbf{x})}{d\mathbf{x}} = \left. \frac{\partial E(\mathbf{x}; \boldsymbol{\lambda})}{\partial \mathbf{x}} \right|_* \quad (18)$$

In short, in order to calculate the molecular gradient for a fully variational wavefunction, we need not evaluate the response

of the wavefunction $\partial\lambda/\partial\mathbf{x}$. This is an extremely important result, which forms the basis for all computational techniques developed for the evaluation of molecular gradients (as well as higher order properties).

4.2 The Hellmann–Feynman Theorem

Let us now be more specific and consider the evaluation of the molecular gradient for a fully variational wavefunction the energy of which corresponds to an expectation value:

$$E(\mathbf{x}; \lambda) = \langle \lambda | \hat{H}(\mathbf{x}) | \lambda \rangle \quad (19)$$

Invoking equation (18), we obtain the following expression for the molecular gradient

$$\frac{dE}{d\mathbf{x}} = \left\langle \lambda^* \left| \frac{\partial \hat{H}}{\partial \mathbf{x}} \right| \lambda^* \right\rangle \quad (20)$$

In standard notation (and in *atomic units*), the nonrelativistic, spinless, and field-free electronic *Hamiltonian* operator is given by

$$\hat{H}(\mathbf{x}) = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{iK} \frac{Z_K}{r_{iK}} + \sum_{K>L} \frac{Z_K Z_L}{R_{KL}} \quad (21)$$

Inserting this Hamiltonian in equation (20), we arrive at the following simple expression for the gradient of the K th nucleus:

$$\mathbf{F}_K = -Z_K \left\langle \lambda^* \left| \sum_i \frac{\mathbf{r}_{iK}}{r_{iK}^3} \right| \lambda^* \right\rangle + Z_K \sum_{L \neq K} Z_L \frac{\mathbf{R}_{LK}}{R_{LK}^3} \quad (22)$$

where \mathbf{R}_{LK} , for example, is the position of nucleus L relative to nucleus K . This result is known as the Hellmann–Feynman theorem. Clearly, for any fully variational energy in the form of equation (19), we may calculate the forces acting on the nuclei by a simple integration over one-electron operators. The Hellmann–Feynman expression (equation 22) thus lends itself to a simple classical interpretation. Although originally stated only for molecular distortions, the Hellmann–Feynman theorem in the form of equation (20) holds for all first-order properties.

4.3 Molecular Hessians

We now proceed to a consideration of the molecular Hessian – that is, the matrix of second derivatives of the molecular electronic energy with respect to geometrical distortions. Differentiating the molecular gradient in the form of equation (18), we obtain from the chain rule

$$\begin{aligned} \frac{d^2 E(\mathbf{x})}{d\mathbf{x}^2} &= \left[\left(\frac{\partial}{\partial \mathbf{x}} + \frac{\partial \lambda}{\partial \mathbf{x}} \frac{\partial}{\partial \lambda} \right) \frac{\partial E(\mathbf{x}; \lambda)}{\partial \lambda} \right]_* \\ &= \frac{\partial^2 E(\mathbf{x}; \lambda)}{\partial \mathbf{x}^2} \Big|_* + \frac{\partial^2 E(\mathbf{x}; \lambda)}{\partial \mathbf{x} \partial \lambda} \Big|_* \frac{\partial \lambda}{\partial \mathbf{x}} \end{aligned} \quad (23)$$

We conclude that for a fully variational wavefunction only the first-order response of the wavefunction $\partial\lambda/\partial\mathbf{x}$ is required to

calculate the energy to second order. In particular, the second-order response of the wavefunction $\partial^2\lambda/\partial\mathbf{x}^2$ is not needed for the evaluation of the molecular Hessian.

Since we can no longer manage without the first-order response, let us consider its evaluation. We have already noted that the variational conditions (equation 16) determine the dependence of the wavefunction on \mathbf{x} . Differentiating these conditions with respect to \mathbf{x} and applying the chain rule, we obtain

$$\left[\frac{d}{d\mathbf{x}} \frac{\partial E(\mathbf{x}; \lambda)}{\partial \lambda} \right]_* = \frac{\partial^2 E(\mathbf{x}; \lambda)}{\partial \mathbf{x} \partial \lambda} \Big|_* + \frac{\partial^2 E(\mathbf{x}; \lambda)}{\partial \lambda^2} \Big|_* \frac{\partial \lambda}{\partial \mathbf{x}} = 0 \quad (24)$$

Introducing the following notation for the electronic gradient and the electronic Hessian of the optimized wavefunction

$$\mathcal{F}(\mathbf{x}) = \frac{\partial E(\mathbf{x}; \lambda)}{\partial \lambda} \Big|_* \quad (25)$$

$$\mathcal{G}(\mathbf{x}) = \frac{\partial^2 E(\mathbf{x}; \lambda)}{\partial \lambda^2} \Big|_* \quad (26)$$

we note that equation (24) may be written in the form of a set of linear symmetric equations:

$$\mathcal{G}(\mathbf{x}) \frac{\partial \lambda}{\partial \mathbf{x}} = -\frac{\partial \mathcal{F}(\mathbf{x})}{\partial \mathbf{x}} \quad (27)$$

These equations are known as the response equations, since they determine the first derivatives (i.e., the first-order responses) of the wavefunction to the perturbation.

In the response equation (equation 27), we note that the electronic Hessian $\mathcal{G}(\mathbf{x})$ is the same for all perturbations – only the differentiated electronic gradient $\partial\mathcal{F}(\mathbf{x})/\partial\mathbf{x}$ on the right-hand side depends on the perturbation. An analogy with Hooke’s law is helpful: the electronic Hessian $\mathcal{G}(\mathbf{x})$ plays the role of the force constant, and the perturbed gradient $-\partial\mathcal{F}(\mathbf{x})/\partial\mathbf{x}$ represents the force. For the unperturbed system, the electronic gradient $\mathcal{F}(\mathbf{x})$ is zero and the wavefunction is optimal or stable. When the perturbation is turned on, the wavefunction of the original unperturbed system is no longer stable – the perturbation introduces a ‘force’ $-\partial\mathcal{F}(\mathbf{x})/\partial\mathbf{x}$ to which the wavefunction responds by restabilizing itself or ‘relaxing’ by the amount $\partial\lambda/\partial\mathbf{x}$. The ‘relaxation’ $\partial\lambda/\partial\mathbf{x}$ is proportional to the ‘force’ $-\partial\mathcal{F}(\mathbf{x})/\partial\mathbf{x}$ and inversely proportional to the ‘force constant’ $\mathcal{G}(\mathbf{x})$.

4.4 The $2n + 1$ Rule

We have established that for a fully variational wavefunction we may calculate the molecular gradient from the zero-order response of the wavefunction (i.e., from the unperturbed wavefunction) and the molecular Hessian from the first-order response of the wavefunction. In general, the $2n + 1$ rule is obeyed: for fully variational wavefunctions, the derivatives (responses) of the wavefunction to order n determine the derivatives of the energy to order $2n + 1$. This means, for instance, that we may calculate the energy to third order with a knowledge of the wavefunction to first order, but that the calculation of the energy to fourth order requires a knowledge of the wavefunction response to second order. These relationships are illustrated in Table 1.

Table 1 The Orders of the Responses Needed to Calculate the Energy to a Given Order in the Perturbation According to the $2n + 1$ Rule for the Wavefunction Parameters and the $2n + 2$ rule for the Lagrange Multipliers

Energy	Wavefunction responses	Multiplier responses
$E^{(0)}$	$\lambda^{(0)}$	
$E^{(1)}$	$\lambda^{(0)}$	$\bar{\lambda}^{(0)}$
$E^{(2)}$	$\lambda^{(0)}, \lambda^{(1)}$	$\bar{\lambda}^{(0)}$
$E^{(3)}$	$\lambda^{(0)}, \lambda^{(1)}$	$\bar{\lambda}^{(0)}, \bar{\lambda}^{(1)}$
$E^{(4)}$	$\lambda^{(0)}, \lambda^{(1)}, \lambda^{(2)}$	$\bar{\lambda}^{(0)}, \bar{\lambda}^{(1)}$
\vdots	\vdots	\vdots
$E^{(2n)}$	$\lambda^{(0)}, \lambda^{(1)}, \dots, \lambda^{(n)}$	$\bar{\lambda}^{(0)}, \bar{\lambda}^{(1)}, \dots, \bar{\lambda}^{(n-1)}$
$E^{(2n+1)}$	$\lambda^{(0)}, \lambda^{(1)}, \dots, \lambda^{(n)}$	$\bar{\lambda}^{(0)}, \bar{\lambda}^{(1)}, \dots, \bar{\lambda}^{(n-1)}, \bar{\lambda}^{(n)}$
$E^{(2n+2)}$	$\lambda^{(0)}, \lambda^{(1)}, \dots, \lambda^{(n)}, \lambda^{(n+1)}$	$\bar{\lambda}^{(0)}, \bar{\lambda}^{(1)}, \dots, \bar{\lambda}^{(n-1)}, \bar{\lambda}^{(n)}$
\vdots	\vdots	\vdots

5 ENERGY DERIVATIVES FOR SCF WAVEFUNCTIONS

Having seen how the evaluation of energy derivatives for fully variational wavefunctions is simplified by the $2n + 1$ rule, let us now consider the molecular gradient for the simplest model of *ab initio* theory – the *self-consistent field (SCF) model*. We will see that, although the result for fully variational wavefunctions (equation 18) may be applied also to SCF energies, we cannot do this without first considering carefully the functional form of the energy expression. With minor modifications, the formalism presented here may be applied also to density-functional theory (DFT).

5.1 The SCF Energy

The SCF wavefunction is constructed as an antisymmetrized (and possibly spin and space symmetrized) product of orthonormal MOs expanded in a finite set of nonorthogonal AOs fixed on the atomic nuclei:

$$\phi_p(\mathbf{r}; \mathbf{x}) = \sum_{\mu} C_{\mu p} \chi_{\mu}(\mathbf{r}; \mathbf{x}) \quad (28)$$

The SCF energy may be written in the general form

$$E_{\text{SCF}} = \sum_{pq} D_{pq} h_{pq} + \frac{1}{2} \sum_{pqrs} d_{pqrs} g_{pqrs} + \sum_{K>L} \frac{Z_K Z_L}{R_{KL}} \quad (29)$$

where the three terms represent the one-electron contributions to the energy, the two-electron contributions, and the nuclear-repulsion interactions. The quantities h_{pq} and g_{pqrs} are the one- and two-electron integrals over the MOs (assuming real orbitals)

$$h_{pq}(\mathbf{x}) = \int \phi_p(\mathbf{r}; \mathbf{x}) \left[-\frac{1}{2} \nabla^2 - \sum_K \frac{Z_K}{r_K} \right] \phi_q(\mathbf{r}; \mathbf{x}) \, d\mathbf{r} \quad (30)$$

$$g_{pqrs}(\mathbf{x}) = \int \frac{\phi_p(\mathbf{r}_1; \mathbf{x}) \phi_q(\mathbf{r}_1; \mathbf{x}) \phi_r(\mathbf{r}_2; \mathbf{x}) \phi_s(\mathbf{r}_2; \mathbf{x})}{r_{12}} \, d\mathbf{r}_1 d\mathbf{r}_2 \quad (31)$$

The integrals depend explicitly on the molecular geometry \mathbf{x} because of the presence of the nuclear-attraction operator in equation (21) and because the MOs are expanded in AOs

fixed on the atomic nuclei. The detailed form of the *density matrix* elements D_{pq} and d_{pqrs} in equation (29) depends on the particular wavefunction considered. For closed-shell, restricted *Hartree-Fock* (RHF) wavefunctions, the energy may be written in the simple form

$$E_{\text{RHF}} = 2 \sum_i h_{ii} + \sum_{ij} (2g_{iijj} - g_{ijji}) + \sum_{K>L} \frac{Z_K Z_L}{R_{KL}} \quad (32)$$

where the summations are over the full set of occupied MOs. (In the following, we will use the indices i, j, k, l for occupied MOs, a, b, c, d for virtual MOs, p, q, r, s for unspecified (occupied or virtual) MOs, and μ, ν, ρ, σ for AOs.)

5.2 The Hartree-Fock Equations

The SCF energy is obtained by minimizing the expectation value of the energy (equation 29) with respect to the coefficients $C_{\mu i}$ of the occupied MOs subject to the constraints that the MOs remain orthonormal:

$$S_{ij} = \langle \phi_i | \phi_j \rangle = \delta_{ij} \quad (33)$$

Such constrained optimizations are conveniently carried out by Lagrange's method of undetermined multipliers. Introducing one Lagrange multiplier for each constraint in equation (33), we arrive at the SCF Lagrangian:

$$L_{\text{SCF}} = E_{\text{SCF}} - \sum_{ij} \varepsilon_{ij} (S_{ij} - \delta_{ij}) \quad (34)$$

where, for real orbitals, the multipliers are symmetric in the two indices.⁸ We note that, for the optimized wavefunction, the Lagrangian is equal to the SCF energy, since the conditions (33) then apply.

The SCF energy is now determined by an unconstrained minimization of the Lagrangian (34) with respect to the MO coefficients and the Lagrange multipliers. The variational conditions on the coefficients and the multipliers are then

$$\frac{\partial L_{\text{SCF}}}{\partial C_{\mu i}} = \frac{\partial E_{\text{SCF}}}{\partial C_{\mu i}} - \sum_{kl} \varepsilon_{kl} \frac{\partial S_{kl}}{\partial C_{\mu i}} = 0 \quad (35)$$

$$\frac{\partial L_{\text{SCF}}}{\partial \varepsilon_{ij}} = S_{ij} - \delta_{ij} = 0 \quad (36)$$

The conditions on the multipliers are clearly equivalent to the MO orthonormality conditions (equation 33), whereas the conditions on the MOs may be written in the form

$$\frac{\partial E_{\text{SCF}}}{\partial C_{\mu i}} = \sum_{kl} \varepsilon_{kl} \frac{\partial S_{kl}}{\partial C_{\mu i}} \quad (37)$$

After some algebra, these conditions may be written in the standard form (the HF equations):

$$\mathbf{F}^{\text{AO}} \mathbf{C} = \mathbf{S}^{\text{AO}} \mathbf{C} \boldsymbol{\varepsilon} \quad (38)$$

where \mathbf{F}^{AO} and \mathbf{S}^{AO} are symmetric square matrices of dimension equal to the number of AOs, $\boldsymbol{\varepsilon}$ is a symmetric square matrix of dimension equal to the number of occupied MOs, and \mathbf{C} is a rectangular matrix with one column for each occupied MO and one row for each AO. For a closed-shell RHF

state, the Fock matrix F^{AO} becomes

$$F_{\mu\nu}^{\text{AO}} = h_{\mu\nu}^{\text{AO}} + \sum_{\rho\sigma} P_{\rho\sigma} \left(g_{\mu\nu\rho\sigma}^{\text{AO}} - \frac{1}{2} g_{\mu\rho\sigma\nu}^{\text{AO}} \right) \quad (\text{closed-shell RHF}) \quad (39)$$

$$P_{\mu\nu} = 2 \sum_i C_{\mu i} C_{\nu i} \quad (\text{closed-shell RHF}) \quad (40)$$

The AO integrals in equations (38) and (39) are related to the MO integrals as

$$S_{pq} = \sum_{\mu\nu} C_{\mu p} C_{\nu q} S_{\mu\nu}^{\text{AO}} \quad (41)$$

$$h_{pq} = \sum_{\mu\nu} C_{\mu p} C_{\nu q} h_{\mu\nu}^{\text{AO}} \quad (42)$$

$$g_{pqrs} = \sum_{\mu\nu\rho\sigma} C_{\mu p} C_{\nu q} C_{\rho r} C_{\sigma s} g_{\mu\nu\rho\sigma}^{\text{AO}} \quad (43)$$

and may be calculated as in equations (30) and (31), but with the MOs replaced by AOs. Multiplying equation (38) from the left by the MO coefficients and invoking the orthonormality conditions on the MOs, we obtain

$$\varepsilon = C^T F^{\text{AO}} C \quad (44)$$

For the optimized wavefunction we may thus identify the Lagrange multipliers in equation (34) with the Fock matrix in the MO basis.

5.3 The SCF Molecular Gradient

Let us now consider the evaluation of molecular gradients for SCF energies. Since we have succeeded in recasting the SCF energy functional in a fully variational (unconstrained) form (equation 34), we can apply the Hellmann–Feynman theorem and obtain the following expression for the SCF molecular gradient:

$$\frac{dE_{\text{SCF}}}{d\mathbf{x}} = \frac{dL_{\text{SCF}}}{d\mathbf{x}} = \frac{\partial L_{\text{SCF}}}{\partial \mathbf{x}} = \frac{\partial E_{\text{SCF}}}{\partial \mathbf{x}} - \sum_{ij} \varepsilon_{ij} \frac{\partial S_{ij}}{\partial \mathbf{x}} \quad (45)$$

We have here first used the fact that the total derivatives of the SCF energy and of the SCF Lagrangian are identical for the optimized wavefunction. Next, we have invoked the Hellmann–Feynman theorem for the fully variational Lagrangian, thereby reexpressing the SCF molecular gradient as the partial derivatives of the Lagrangian. Finally, we have inserted the expression for the Lagrangian (equation 34).

A more explicit expression for the SCF molecular gradient is obtained by substituting into equation (45) the expression for the SCF energy in terms of density-matrix elements and MO integrals:

$$\begin{aligned} \frac{dE_{\text{SCF}}}{d\mathbf{x}} &= \sum_{ij} D_{ij} \frac{\partial h_{ij}}{\partial \mathbf{x}} + \frac{1}{2} \sum_{ijkl} d_{ijkl} \frac{\partial g_{ijkl}}{\partial \mathbf{x}} - \sum_{ij} \varepsilon_{ij} \frac{\partial S_{ij}}{\partial \mathbf{x}} + \mathbf{F}_{\text{nuc}} \\ &= \sum_{\mu\nu} D_{\mu\nu}^{\text{AO}} \frac{\partial h_{\mu\nu}^{\text{AO}}}{\partial \mathbf{x}} + \frac{1}{2} \sum_{\mu\nu\rho\sigma} d_{\mu\nu\rho\sigma}^{\text{AO}} \frac{\partial g_{\mu\nu\rho\sigma}^{\text{AO}}}{\partial \mathbf{x}} \\ &\quad - \sum_{\mu\nu} \varepsilon_{\mu\nu}^{\text{AO}} \frac{\partial S_{\mu\nu}^{\text{AO}}}{\partial \mathbf{x}} + \mathbf{F}_{\text{nuc}} \end{aligned} \quad (46)$$

where \mathbf{F}_{nuc} represents the nuclear–nuclear contributions. In the second expression, we have expanded the MO integrals

in equations (41) to (43) in the AO integrals, associating the MO coefficients instead with the density-matrix elements and with the Lagrange multipliers:

$$D_{\mu\nu}^{\text{AO}} = \sum_{ij} C_{\mu i} C_{\nu j} D_{ij} \quad (47)$$

$$d_{\mu\nu\rho\sigma}^{\text{AO}} = \sum_{ijkl} C_{\mu i} C_{\nu j} C_{\rho k} C_{\sigma l} d_{ijkl} \quad (48)$$

$$\varepsilon_{\mu\nu}^{\text{AO}} = \sum_{ij} C_{\mu i} C_{\nu j} \varepsilon_{ij} \quad (49)$$

For closed-shell RHF wavefunctions, the density-matrix elements are given by

$$D_{\mu\nu}^{\text{AO}} = P_{\mu\nu} \quad (\text{closed-shell RHF}) \quad (50)$$

$$d_{\mu\nu\rho\sigma}^{\text{AO}} = P_{\mu\nu} P_{\rho\sigma} - \frac{1}{2} P_{\mu\rho} P_{\nu\sigma} \quad (\text{closed-shell RHF}) \quad (51)$$

greatly simplifying the evaluation of the molecular gradient. For more general SCF wavefunctions, the same expression (equation 46) may be used, but the density-matrix elements are different. The expressions for the molecular gradient of an SCF wavefunction was first derived by Bratoz in 1958, who also considered second derivatives.⁹ The modern development begins with the work by Pulay in 1969 for gradients^{10,11} and with the work by Pople and co-workers for Hessians.¹²

The above expression for the SCF gradient (equation 46) should be compared with the simpler Hellmann–Feynman expression (equation 22). Besides the terms identical to those in the Hellmann–Feynman expression, the SCF gradient (equation 46) contains a number of terms that arise from the dependence of the AOs on the geometry – since the AOs are fixed to the atomic nuclei, they are displaced relative to one another when the molecule is distorted. Note in particular that there are contributions from the two-electron integrals and from the overlap integrals to the SCF gradient. In the limit of a complete basis, all contributions from the AOs cancel and the SCF gradient becomes identical to the Hellmann–Feynman expression (equation 22). We also note that, although the SCF gradient is indeed more complicated than the Hellmann–Feynman expression, there is one important point of similarity – neither expression requires the first-order response of the wavefunction (i.e., the solution of linear equations) for the calculation of the gradient.

5.4 Computational Aspects

Let us comment briefly on the evaluation of molecular gradients according to equation (46). The number of two-electron integrals and densities is usually so large that they cannot be kept in computer memory or stored on disk. In contrast, the number of one-electron integrals and densities is much smaller and may (except for large systems) be kept in memory. In the evaluation of the molecular gradient according to equation (46), we must then proceed in such a manner that storage of two-electron integrals and densities can be avoided. The two-electron densities may be calculated from equation (51) and present no problems. To avoid storage of the two-electron integrals, we multiply each derivative integral $\partial g_{\mu\nu\rho\sigma}^{\text{AO}} / \partial \mathbf{x}$ by the density elements as soon as the integral has been generated. In this way, the storage of two-electron integrals is avoided altogether.

The cost of evaluating the molecular gradient depends on the number of AO basis functions rather than on the number of geometrical degrees of freedom – even in a large molecule containing several hundred atoms, each two-electron integral depends on the position of, at most, four of these atoms. Therefore, each two-electron integral may never contribute to more than 12 components of the molecular gradient. Furthermore, the simultaneous evaluation of all these components is much faster than the separate evaluation of each component. For each two-electron integral, therefore, the full set of 12 Cartesian derivatives $\partial g_{\mu\nu\rho\sigma}^{\text{AO}}/\partial\mathbf{x}$ is first evaluated, and then these derivatives are all multiplied by the same density-matrix element and added to the appropriate component of the molecular gradient.

The evaluation of derivative AO integrals is straightforward, since the Cartesian derivatives of *Gaussian orbitals* are just linear combinations of Gaussians of different quantum numbers. For example, the derivative of a Gaussian s orbital is a p orbital, the derivative of a Gaussian p orbital is a linear combination of an s orbital and a d orbital, and so on. The derivative integrals can therefore be calculated by a simple modification of the usual scheme for undifferentiated integrals. For a discussion of integral evaluation, see *Integrals of Electron Repulsion*.

6 SECOND-QUANTIZATION REPRESENTATION OF DERIVATIVES

From the discussion in the preceding sections, it appears that the evaluation of molecular gradients (and, therefore, also molecular Hessians) for approximate wavefunctions follows a path quite different from that for exact wavefunctions. The relationship between the theories for exact and approximate wavefunctions is nevertheless a close one, as becomes particularly apparent in the formalism of second quantization. Indeed, in this formalism, the SCF gradient may be calculated from an expression formally identical to that for exact wavefunctions (equation 20). Such an equivalence of approximate and exact theories is useful, since it allows for a simpler formal manipulation and a greater transparency of the derived expressions. We therefore consider here the evaluation of molecular gradients and Hessians using the formalism of second quantization. For readers unacquainted with second quantization, a short introduction is given in the Appendix. For a more thorough introduction, see the monograph by Jørgensen and Simons.⁷

6.1 SCF Molecular Gradients in Second Quantization

In second quantization, the SCF energy is written as an expectation value of the Hamiltonian operator (see equation 119) in the form

$$E(\mathbf{x}, \boldsymbol{\kappa}) = \langle \boldsymbol{\kappa} | \hat{H}(\mathbf{x}) | \boldsymbol{\kappa} \rangle = \langle 0 | \exp(\hat{\kappa}) \hat{H}(\mathbf{x}) \exp(-\hat{\kappa}) | 0 \rangle \quad (52)$$

For a closed-shell RHF state, the orbital-rotation operator $\hat{\kappa}$ in equation (116) takes the form

$$\hat{\kappa} = \sum_{ai} \kappa_{ai} E_{ai}^- \quad (53)$$

where the parameters κ_{ai} are real and the summation is over all pairs of occupied and virtual MOs (recalling that the indices

$i, j, k \dots$ are used for occupied MOs and the indices $a, b, c \dots$ for virtual MOs). Like the exact functional (equation 19), the functional (equation 52) makes a clear distinction between those parameters \mathbf{x} that determine the physical system (isolating these in the Hamiltonian) and those that determine the electronic state (isolating these in the wavefunction).

At the reference geometry \mathbf{x}_0 , the wavefunction $|0\rangle$ is variational with respect to the full set of orbital-rotation parameters in equation (53) – no constraints arise, since the rotations are unitary for all values of κ_{ai} . According to equation (122), the variational conditions may then be written in the form

$$\langle 0 | [E_{ai}^-, \hat{H}(\mathbf{x})] | 0 \rangle = 0 \quad (54)$$

In accordance with the Hellmann–Feynman theorem, the RHF molecular gradient now takes the form

$$\frac{dE(\mathbf{x})}{d\mathbf{x}} = \left\langle 0 \left| \frac{\partial \hat{H}(\mathbf{x})}{\partial \mathbf{x}} \right| 0 \right\rangle \quad (55)$$

which is identical to that for the exact wavefunction (equation 20). It remains only to evaluate the derivatives of the second-quantization Hamiltonian (equation 104):

$$\begin{aligned} \hat{H}(\mathbf{x}) = & \sum_{pq} \tilde{h}_{pq}(\mathbf{x}) E_{pq} + \frac{1}{2} \sum_{pqrs} \tilde{g}_{pqrs}(\mathbf{x}) (E_{pq} E_{rs} - \delta_{rq} E_{ps}) \\ & + \sum_{K>L} \frac{Z_K Z_L}{R_{KL}} \end{aligned} \quad (56)$$

where the summations are over the full set of MOs and where the integrals (see equations 107 and 108) are given by

$$\tilde{h}_{pq}(\mathbf{x}) = \sum_{p'q'} h_{p'q'}(\mathbf{x}) S_{p'p}^{-1/2}(\mathbf{x}) S_{q'q}^{-1/2}(\mathbf{x}) \quad (57)$$

$$\begin{aligned} \tilde{g}_{pqrs}(\mathbf{x}) = & \sum_{p'q'r's'} g_{p'q'r's'}(\mathbf{x}) S_{p'p}^{-1/2}(\mathbf{x}) \\ & \times S_{q'q}^{-1/2}(\mathbf{x}) S_{r'r}^{-1/2}(\mathbf{x}) S_{s's}^{-1/2}(\mathbf{x}) \end{aligned} \quad (58)$$

In short-hand notation, $S_{pq}^{-1/2} = [S^{-1/2}]_{pq}$. Noting that the orbitals are orthonormal for the unperturbed system

$$S^{-1/2}(\mathbf{x}_0) = \mathbf{1} \quad (59)$$

we obtain

$$\frac{\partial \tilde{h}_{pq}}{\partial \mathbf{x}} = \frac{\partial h_{pq}}{\partial \mathbf{x}} - \frac{1}{2} \left\{ \frac{\partial S}{\partial \mathbf{x}}, \mathbf{h} \right\}_{pq} \quad (60)$$

$$\frac{\partial \tilde{g}_{pqrs}}{\partial \mathbf{x}} = \frac{\partial g_{pqrs}}{\partial \mathbf{x}} - \frac{1}{2} \left\{ \frac{\partial S}{\partial \mathbf{x}}, \mathbf{g} \right\}_{pqrs} \quad (61)$$

where we have introduced the following notation for one-index transformations:

$$\{\mathbf{A}, \mathbf{h}\}_{pq} = \sum_o (A_{po} h_{oq} + A_{qo} h_{po}) \quad (62)$$

$$\{\mathbf{A}, \mathbf{g}\}_{pqrs} = \sum_o (A_{po} g_{oqrs} + A_{qo} g_{pors} + A_{ro} g_{pqos} + A_{so} g_{pqro}) \quad (63)$$

Inserting these expressions for the integrals into the Hamiltonian operator in equation (55) and carrying out some algebra, we recover the expression for the SCF molecular gradient derived in Section 5.

Although the same final expressions are obtained in the first and second quantizations, the derivations are quite different. In second quantization, all basis-set effects are isolated in the Hamiltonian (equation 56) and thus do not appear until the derivatives of the integrals are taken.^{13,14} Since the same Hamiltonian is used for all wavefunctions, we have then solved the problems associated with atom-fixed AOs once and for all, in the sense that all correction terms that arise from the AOs appear automatically upon differentiation of the Hamiltonian – for any wavefunction model. In first quantization, on the other hand, we must incorporate the orthonormality conditions in the energy functional itself (which is different for each wavefunction model) through the introduction of Lagrange multipliers.

6.2 SCF Molecular Hessians in Second Quantization

Let us briefly consider the evaluation of the SCF molecular *Hessian* in second quantization. Proceeding in the manner outlined in Section 4, we arrive at the following expression for the molecular Hessian

$$\frac{d^2E}{d\mathbf{x}^2} = \left\langle 0 \left| \frac{\partial^2 \hat{H}}{\partial \mathbf{x}^2} \right| 0 \right\rangle + \left\langle 0 \left| \left[\frac{\partial \hat{K}}{\partial \mathbf{x}}, \frac{\partial \hat{H}}{\partial \mathbf{x}} \right] \right| 0 \right\rangle \quad (64)$$

The first term is a simple expectation value of the Hamiltonian (differentiated twice with respect to the molecular geometry), whereas the second term contains the first-order response of the wavefunction to the perturbation:

$$\frac{\partial \hat{K}}{\partial \mathbf{x}} = \sum_{ai} \frac{\partial \kappa_{ai}}{\partial \mathbf{x}} E_{ai}^- \quad (65)$$

The derivatives of the orbital-rotation parameters are obtained from the response equations:

$$\sum_{bj} \langle 0 | [E_{ai}^-, [E_{bj}^-, \hat{H}]] | 0 \rangle \frac{\partial \kappa_{bj}}{\partial \mathbf{x}} = - \left\langle 0 \left| \left[E_{ai}^-, \frac{\partial \hat{H}}{\partial \mathbf{x}} \right] \right| 0 \right\rangle \quad (66)$$

where no symmetrization of the nested commutator (as in equation 123) is necessary since the gradient (equation 54) vanishes.

The time-consuming step in the evaluation of the SCF molecular Hessian is the calculation of two-electron contributions from $\partial^2 \hat{H} / \partial \mathbf{x}^2$, since each integral now makes 78 contributions to the Hessian (although some of these are related by rotational and translational symmetries and, therefore, need not be calculated explicitly). Again, these integrals may be obtained by a simple modification of the usual scheme for the calculation of undifferentiated integrals.

The solution of the linear equations (equation 66) is usually less time-consuming than the calculation of the two-electron contribution to the Hessian (equation 64), but is treated in somewhat more detail here in order to demonstrate some general principles. The number of orbital rotations is usually so large that the electronic Hessian cannot be constructed or stored explicitly. Instead, iterative techniques are used, where the key step is the evaluation of matrix–vector products such as

$$\sigma_{ai} = \sum_{bj} \langle 0 | [E_{ai}^-, [E_{bj}^-, \hat{H}]] | 0 \rangle v_{bj} \quad (67)$$

We now introduce the effective Hamiltonian

$$\hat{H}_v = \sum_{bj} [E_{bj}^-, \hat{H}] v_{bj} \quad (68)$$

which is Hermitian (being a linear combination of commutators of the anti-Hermitian operators E_{bj}^- with the Hermitian operator \hat{H}) and may be evaluated in the following way in terms of one-index transformed integrals (equations 62 and 63):

$$\hat{H}_v = \sum_{pq} \{v, \mathbf{h}\}_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} \{v, \mathbf{g}\}_{pqrs} (E_{pq} E_{rs} - \delta_{rq} E_{ps}) \quad (69)$$

Here v is an antisymmetric matrix containing zero elements in the diagonal (occupied–occupied and virtual–virtual) blocks and the elements v_{bj} in the off-diagonal blocks:

$$v = \begin{bmatrix} 0 & \vdots & -v_{bj} \\ \vdots & \ddots & \vdots \\ v_{bj} & \vdots & 0 \end{bmatrix} \quad (70)$$

The matrix–vector product (equation 67) may thus be evaluated in the same manner as an electronic gradient:

$$\sigma_{ai} = \langle 0 | [E_{ai}^-, \hat{H}_v] | 0 \rangle \quad (71)$$

thereby bypassing the construction of the electronic Hessian altogether. In practice, the evaluation of the effective gradient (equation 71) is carried out in the AO basis, thus avoiding the transformation of integrals to the MO basis.

7 THE MOLECULAR GRADIENT FOR NONVARIATIONAL WAVEFUNCTIONS

In the preceding sections we have seen how the molecular gradient and Hessian can be calculated for fully variational wavefunctions. In particular, we have applied our results for fully variational wavefunctions to the most basic of *ab initio* models – the HF model. We now consider the evaluation of molecular gradients and Hessians for models more complicated than the HF model. As our example, we choose the truncated CI model. This wavefunction is chosen for its simplicity, which allows us to keep the technical details to a minimum, thus making the discussion more transparent. However, the scheme we shall follow may be applied to any computational model (e.g., the CC and MP models) of *ab initio* theory.

7.1 The CI Molecular Gradient

In the truncated CI model, the wavefunction at \mathbf{x} is written as a linear combination of Slater determinants:

$$|\text{CI}\rangle = \sum_{\mu} C_{\mu} |\mu\rangle \quad (72)$$

The determinants in equation (72), which constitute a subset of the full set of determinants that may be constructed in a given orbital basis, are constructed from a precalculated set of MOs, often obtained from an HF calculation at \mathbf{x} . The CI energy $E_{\text{CI}}(\mathbf{x})$ is thus determined by minimizing the expectation value of the Hamiltonian operator

$$E_{\text{CI}}(\mathbf{x}; \mathbf{C}, \boldsymbol{\kappa}) = \frac{\langle \text{CI} | \hat{H}(\mathbf{x}) | \text{CI} \rangle}{\langle \text{CI} | \text{CI} \rangle} \quad (73)$$

with respect to the CI coefficients, subject to the constraints that the MOs are obtained from a preceding HF calculation and, therefore, satisfy the conditions

$$\frac{\partial E_{\text{SCF}}(\mathbf{x}; \boldsymbol{\kappa})}{\partial \kappa_{ai}} = 0 \quad (74)$$

in the independent exponential parametrization of Section 6.

Because of the HF constraints on the MOs, the CI energy is nonvariational with respect to variations in the MO coefficients. Naively, we might attempt to calculate the derivatives from an expression such as

$$\begin{aligned} \frac{dE_{\text{CI}}}{d\mathbf{x}} &= \frac{\partial E_{\text{CI}}}{\partial \mathbf{x}} + \sum_{\mu} \frac{\partial E_{\text{CI}}}{\partial C_{\mu}} \frac{\partial C_{\mu}}{\partial \mathbf{x}} + \sum_{ai} \frac{\partial E_{\text{CI}}}{\partial \kappa_{ai}} \frac{\partial \kappa_{ai}}{\partial \mathbf{x}} \\ &= \frac{\partial E_{\text{CI}}}{\partial \mathbf{x}} + \sum_{ai} \frac{\partial E_{\text{CI}}}{\partial \kappa_{ai}} \frac{\partial \kappa_{ai}}{\partial \mathbf{x}} \end{aligned} \quad (75)$$

where the simplification occurs because of the variational conditions on the CI coefficients:

$$\frac{\partial E_{\text{CI}}}{\partial C_{\mu}} = 0 \quad (76)$$

However, since the CI energy is nonvariational with respect to the orbital-rotation parameters

$$\frac{\partial E_{\text{CI}}}{\partial \kappa_{ai}} \neq 0 \quad (77)$$

we would still have to calculate the partial derivatives of the orbital-rotation parameters with respect to all Cartesian directions by solving the linear equations

$$\sum_{bj} \frac{\partial^2 E_{\text{SCF}}}{\partial \kappa_{ai} \partial \kappa_{bj}} \frac{\partial \kappa_{bj}}{\partial \mathbf{x}} = - \frac{\partial^2 E_{\text{SCF}}}{\partial \kappa_{ai} \partial \mathbf{x}} \quad (78)$$

For a large system with many degrees of freedom, such a task is an expensive undertaking. A similar situation arises for any other nonvariational computational model. In the coupled-cluster model, for example, we would need to calculate the partial derivatives of the cluster amplitudes and the orbital-rotation parameters with respect to all perturbations of interest. Clearly, to make the calculation of gradients practical for such wavefunctions, we must come up with a better scheme for the evaluation of molecular gradients.

The key to solving this problem is to calculate the CI molecular gradient and other molecular properties not from the original energy expression (equation 73) but from a different, variational energy functional, the variationally optimized energy of which coincides with the nonvariational CI energy $E_{\text{CI}}(\mathbf{x})$. If such a functional can be found, then the evaluation of the molecular gradient may be carried out in exactly the same manner as for variational wavefunctions. Obviously, for this strategy to be useful, the construction of the new energy functional must be inexpensive so that what we gain from its use is not lost in its construction. Fortunately, a systematic and inexpensive procedure exists for the construction of variational energy functionals, making this strategy worthwhile in most cases.

Recalling our discussion of SCF molecular gradients in Section 5, we set up a variational CI energy functional by using Lagrange's method of undetermined multipliers. For each

SCF constraint (equation 74), we introduce one undetermined multiplier $\bar{\kappa}_{ai}$ and arrive at the following Lagrangian:

$$L_{\text{CI}}(\mathbf{x}; \mathbf{C}, \boldsymbol{\kappa}, \bar{\boldsymbol{\kappa}}) = E_{\text{CI}}(\mathbf{x}; \mathbf{C}, \boldsymbol{\kappa}) + \sum_{ai} \bar{\kappa}_{ai} \left(\frac{\partial E_{\text{SCF}}(\mathbf{x}; \boldsymbol{\kappa})}{\partial \kappa_{ai}} - 0 \right) \quad (79)$$

No constraints are required for the orthonormality of the MOs, since here we have used the second-quantization formulation (see Section 6). We now require the Lagrangian to be fully variational with respect to all parameters:

$$\frac{\partial L_{\text{CI}}}{\partial C_{\mu}} = \frac{\partial E_{\text{CI}}}{\partial C_{\mu}} = 0 \quad (80)$$

$$\frac{\partial L_{\text{CI}}}{\partial \kappa_{ai}} = \frac{\partial E_{\text{CI}}}{\partial \kappa_{ai}} + \sum_{bj} \bar{\kappa}_{bj} \frac{\partial^2 E_{\text{SCF}}}{\partial \kappa_{bj} \partial \kappa_{ai}} = 0 \quad (81)$$

$$\frac{\partial L_{\text{CI}}}{\partial \bar{\kappa}_{ai}} = \frac{\partial E_{\text{SCF}}}{\partial \kappa_{ai}} = 0 \quad (82)$$

The first and last equations are trivially satisfied: the conditions given by equation (80) merely represent the standard variational conditions on the CI energy, whereas the conditions (82) are simply the variational SCF conditions on the orbitals. The only conditions that require special attention are the variational conditions on the orbitals (equation 81). To satisfy these conditions, we solve the following linear equations:

$$\sum_{bj} \frac{\partial^2 E_{\text{SCF}}}{\partial \kappa_{ai} \partial \kappa_{bj}} \bar{\kappa}_{bj} = - \frac{\partial E_{\text{CI}}}{\partial \kappa_{ai}} \quad (83)$$

These equations have the same structure as the equations for the first-order orbital responses (equation 66), except that the right-hand side corresponds to the electronic gradient of the CI energy:

$$\sum_{bj} \langle \text{HF} | [E_{ai}^-, [E_{bj}^-, \hat{H}]] | \text{HF} \rangle \bar{\kappa}_{bj} = - \langle \text{CI} | [E_{ai}^-, \hat{H}] | \text{CI} \rangle \quad (84)$$

Once the multipliers have been determined from these equations, then the CI Lagrangian (equation 79) is fully variational. At the stationary point of the CI Lagrangian, all terms involving the multipliers in equation (79) vanish and the Lagrangian becomes identical to the original CI energy.

Since the CI Lagrangian is fully variational, we may now calculate the molecular gradient in the same way as for any other variational energy expression:

$$\frac{dE_{\text{CI}}}{d\mathbf{x}} = \frac{dL_{\text{CI}}}{d\mathbf{x}} = \frac{\partial L_{\text{CI}}}{\partial \mathbf{x}} = \frac{\partial E_{\text{CI}}}{\partial \mathbf{x}} + \sum_{ai} \bar{\kappa}_{ai} \frac{\partial^2 E_{\text{SCF}}}{\partial \mathbf{x} \partial \kappa_{ai}} \quad (85)$$

The advantage of this expression over equation (75) is that, for any number of perturbations, there is only one set of linear equations to be solved – namely those that determine the Lagrange multipliers. In contrast, in the original expression for the CI gradient (equation 75) there is one set of linear equations for each independent perturbational direction. Note that the complexity of the first-order orbital-response equations (equation 66) and the zero-order multiplier equations (equation 84) is the same, making the Lagrangian method the preferred one whenever more than one perturbation is considered. Using a different method, equation (85) was first derived by Handy and Schaefer.¹⁵

7.2 Computational Aspects

Let us briefly consider the evaluation of the CI molecular gradient from the Lagrangian in equation (85). In second quantization, we may rewrite equation (85) as

$$\frac{dE_{\text{CI}}}{d\mathbf{x}} = \left\langle \text{CI} \left| \frac{\partial \hat{H}}{\partial \mathbf{x}} \right| \text{CI} \right\rangle + \sum_{ai} \bar{\kappa}_{ai} \left\langle \text{HF} \left| \left[E_{ai}^-, \frac{\partial \hat{H}}{\partial \mathbf{x}} \right] \right| \text{HF} \right\rangle \quad (86)$$

assuming a normalized CI state. Introducing the effective Hermitian operator

$$\begin{aligned} \hat{H}^{\bar{\kappa}}(\mathbf{x}) &= \sum_{ai} \bar{\kappa}_{ai} [E_{ai}^-, \hat{H}(\mathbf{x})] \\ &= \sum_{pq} \{\bar{\kappa}, \mathbf{h}(\mathbf{x})\}_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} \{\bar{\kappa}, \mathbf{g}(\mathbf{x})\}_{pqrs} \\ &\quad \times (E_{pq} E_{rs} - \delta_{rq} E_{ps}) \end{aligned} \quad (87)$$

where $\bar{\kappa}$ is an antisymmetric matrix constructed in the same manner as \mathbf{v} in equation (70) we find that the CI gradient may be written as the sum of two expectation values, one involving the CI state and the other the SCF state:

$$\frac{dE_{\text{CI}}}{d\mathbf{x}} = \left\langle \text{CI} \left| \frac{\partial \hat{H}}{\partial \mathbf{x}} \right| \text{CI} \right\rangle + \left\langle \text{HF} \left| \frac{\partial \hat{H}^{\bar{\kappa}}}{\partial \mathbf{x}} \right| \text{HF} \right\rangle \quad (88)$$

We shall not pursue the evaluation of this expression in any detail, but note that it may be calculated in terms of derivative integrals and effective densities¹⁶

$$\frac{dE_{\text{CI}}}{d\mathbf{x}} = \sum_{pq} D_{pq}^{\text{eff}} \frac{\partial \tilde{h}_{pq}}{\partial \mathbf{x}} + \frac{1}{2} \sum_{pqrs} d_{pqrs}^{\text{eff}} \frac{\partial \tilde{g}_{pqrs}}{\partial \mathbf{x}} + \mathbf{F}_{\text{nuc}} \quad (89)$$

where \mathbf{F}_{nuc} represents the nuclear–nuclear contribution and where the effective variational density-matrix elements are given by

$$D_{pq}^{\text{eff}} = D_{pq}^{\text{CI}} - \{\bar{\kappa}, \mathbf{D}^{\text{HF}}\}_{pq} \quad (90)$$

$$d_{pqrs}^{\text{eff}} = d_{pqrs}^{\text{CI}} - \{\bar{\kappa}, \mathbf{d}^{\text{HF}}\}_{pqrs} \quad (91)$$

The term ‘variational’ is used here as these densities correspond to a variational Lagrangian formulation of the CI energy. The one-electron CI and HF densities are given by

$$D_{pq}^{\text{CI}} = \langle \text{CI} | E_{pq} | \text{CI} \rangle \quad (92)$$

$$D_{pq}^{\text{HF}} = \langle \text{HF} | E_{pq} | \text{HF} \rangle \quad (93)$$

and similarly for the two-electron densities. In conclusion, we have found that the CI gradient may be calculated in the same way as the HF gradient, but in terms of the variational densities (equations 90 and 91) involving the Lagrange multipliers.

7.3 The Lagrangian Method and the $2n + 2$ Rule

The Lagrangian method is a general one, which may be applied to any wavefunction in order to simplify the calculations of derivatives.^{3,17} The procedure may be summarized as follows:

1. The energy functional $E(\mathbf{x}; \boldsymbol{\lambda})$ is set up in the usual manner and the energy $E(\mathbf{x})$ is determined by optimizing $E(\mathbf{x}; \boldsymbol{\lambda})$ with respect to $\boldsymbol{\lambda}$ (variationally or otherwise).

2. All nonvariational conditions

$$e_i(\mathbf{x}; \boldsymbol{\lambda}) = 0 \quad (94)$$

on the wavefunction parameters $\boldsymbol{\lambda}$ are identified. Each condition is multiplied by an undetermined Lagrange multiplier and added to the original energy expression

$$L(\mathbf{x}; \boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}}) = E(\mathbf{x}; \boldsymbol{\lambda}) + \sum_i \bar{\lambda}_i e_i(\mathbf{x}; \bar{\boldsymbol{\lambda}}) \quad (95)$$

These conditions may, for example, correspond to the MO orthonormality conditions or to the SCF conditions on the MOs in CI calculations.

3. The Lagrangian is made variational by solving the equations

$$\frac{\partial L(\mathbf{x}; \boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}})}{\partial \bar{\lambda}_i} = e_i(\mathbf{x}; \boldsymbol{\lambda}) = 0 \quad (96)$$

$$\frac{\partial L(\mathbf{x}; \boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}})}{\partial \boldsymbol{\lambda}} = \frac{\partial E(\mathbf{x}; \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} + \sum_i \bar{\lambda}_i \frac{\partial e_i(\mathbf{x}; \bar{\boldsymbol{\lambda}})}{\partial \boldsymbol{\lambda}} = \mathbf{0} \quad (97)$$

The first set of equations represents the conditions (94), whereas the second set determines the Lagrange multipliers.

4. The derivatives are calculated from the Lagrangian as for any fully variational wavefunction. In particular, the molecular gradient is obtained as

$$\frac{dE(\mathbf{x})}{d\mathbf{x}} = \frac{dL(\mathbf{x})}{d\mathbf{x}} = \frac{\partial E(\mathbf{x}; \boldsymbol{\lambda})}{\partial \mathbf{x}} + \sum_i \bar{\lambda}_i \frac{\partial e_i(\mathbf{x}; \bar{\boldsymbol{\lambda}})}{\partial \mathbf{x}} \quad (98)$$

in accordance with the Hellmann–Feynman theorem.

We have already noted that, for a fully variational energy functional, the derivatives of the wavefunction parameters to order n determine the energy to order $2n + 1$. Since the Lagrange multipliers are part of the wavefunction parameters for the Lagrangian, this result must hold also for the multipliers. It turns out, however, that an even stronger $2n + 2$ rule applies to the Lagrange multipliers: The derivatives of the Lagrange multipliers to order n determine the energy to order $2n + 2$ (see Table 1). The zero-order Lagrange multipliers, for instance, determine the energy to second order. The reason for this special behavior of the multipliers is that these parameters appear in a very special manner in the Lagrangian (equation 95) – that is, linearly.

8 CONCLUSIONS

We have discussed the evaluation of molecular gradients and Hessians for variational and nonvariational wavefunctions. In both cases, we may reduce the calculation of molecular gradients to the calculation of an expectation value, combining differentiated atomic integrals with a set of density-matrix elements – no solution of perturbed equations is needed, although for nonvariational wavefunctions a set of Lagrange multipliers must be determined from a set of linear equations. For variational wavefunctions, the standard densities are used for the evaluation of the gradient; for nonvariational wavefunctions, a set of effective, variational densities is constructed from the Lagrange multipliers. In either case, the time-consuming step in the calculation of the gradient is the evaluation of the differentiated two-electron integrals.

For variational wavefunctions, the wavefunction parameters obey the $2n + 1$ rule, which states that the derivatives of the parameters to order n determine the derivatives of the energy to order $2n + 1$. The same rule is obeyed also by the wavefunction parameters of nonvariational wavefunctions, whereas the Lagrange multipliers follow the $2n + 2$ rule.

9 APPENDIX: ELEMENTS OF SECOND QUANTIZATION

9.1 Creation and Annihilation Operators

In second quantization, Slater determinants are expressed as products or ‘strings’ of creation operators a_p^\dagger working on the vacuum state

$$|\mu\rangle = a_{\mu_1}^\dagger a_{\mu_2}^\dagger \dots a_{\mu_N}^\dagger |\text{vac}\rangle \quad (99)$$

There is one creation operator a_p^\dagger for each spin orbital in the basis and each such operator is viewed as creating an electron in the associated spin orbital. Within a given orbital basis, any wavefunction may then be expressed as a linear combination of strings of creation operators working on the vacuum state.

With each creation operator a_p^\dagger , there is an associated conjugate operator a_p – known as the annihilation operator – which removes the electron from spin orbital P , undoing the effect of the creation operator. If the annihilation operator a_p is applied to any state that does not contain the spin orbital P , the state vanishes. In particular, applying an annihilation operator to the vacuum state, we obtain

$$a_p |\text{vac}\rangle = 0 \quad (100)$$

since there are no electrons to be annihilated in this state.

To conform with the antisymmetry requirements of fermions, the creation and annihilation operators are required to satisfy the anticommutation relations:

$$[a_p^\dagger, a_q^\dagger]_+ = 0 \quad (101)$$

$$[a_p, a_q]_+ = 0 \quad (102)$$

$$[a_p^\dagger, a_q]_+ = \delta_{pq} \quad (103)$$

Thus, all pairs of creation and/or annihilation operators anticommute except for the conjugate pairs of operators such as a_p^\dagger and a_p . From these relationships, all other properties of the creation and annihilation operators – often referred to as the elementary operators of second quantization – follow. We note that equation (103) holds only for orthonormal sets of spin orbitals. For nonorthonormal spin orbitals, the Kronecker delta in equation (103) must be replaced by the overlap integral between the two spin orbitals.

9.2 The Molecular Electronic Hamiltonian

In second quantization, the electronic Hamiltonian operator is expressed as a linear combination of strings of creation and annihilation operators. The following form is appropriate for a spin-free, nonrelativistic electronic system:

$$\hat{H}(\mathbf{x}) = \sum_{pq} \tilde{h}_{pq}(\mathbf{x}) E_{pq} + \frac{1}{2} \sum_{pqrs} \tilde{g}_{pqrs}(\mathbf{x}) (E_{pq} E_{rs} - \delta_{rq} E_{ps}) + \sum_{K>L} \frac{Z_K Z_L}{R_{KL}} \quad (104)$$

The summations are here over the full set of molecular orbitals. (Lower-case letters are used for the orbitals to distinguish these from the spin orbitals, for which upper-case letters were used in Section 9.1.) The excitation operators E_{pq} in equation (104) are given by

$$E_{pq} = a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta} \quad (105)$$

and represent the excitation of an electron from orbital q to orbital p – that is, the annihilation of an electron in orbital q followed by the creation of an electron of the same spin in orbital p . From the anticommutation relations (equations 101–103), the excitation operators may be shown to satisfy the commutation relations

$$[E_{pq}, E_{rs}] = E_{ps} \delta_{rq} - E_{rq} \delta_{ps} \quad (106)$$

The numerical quantities appearing in the Hamiltonian operator (equation 104) are related to the molecular integrals (equations 30 and 31) in the following manner:

$$\tilde{h}_{pq}(\mathbf{x}) = \sum_{p'q'} h_{p'q'}(\mathbf{x}) T_{p'p}(\mathbf{x}) T_{q'q}(\mathbf{x}) \quad (107)$$

$$\tilde{g}_{pqrs}(\mathbf{x}) = \sum_{p'q'r's'} g_{p'q'r's'}(\mathbf{x}) T_{p'p}(\mathbf{x}) T_{q'q}(\mathbf{x}) T_{r'r}(\mathbf{x}) T_{s's}(\mathbf{x}) \quad (108)$$

where $T_{pq}(\mathbf{x})$ are elements of the connection matrix, which for our purposes may be chosen as

$$\mathbf{T}(\mathbf{x}) = \mathbf{S}^{-1/2}(\mathbf{x}) \quad (109)$$

although different choices should be made in some cases.

The purpose of the connection matrix is to make the Hamiltonian (equation 104) valid also for nonorthogonal molecular orbitals. We recall that, although the MOs from which the unperturbed wavefunction is constructed are orthonormal, these MOs are not necessarily orthonormal for the perturbed system, since the overlap matrix may depend on the perturbation. The presence of the connection matrix in the Hamiltonian effectively solves the nonorthogonality problem by incorporating the necessary corrections for nonorthogonality in the molecular integrals – the summations in equation (104) being over the set of orthonormalized MOs. For the unperturbed system, the connection matrix becomes unity and the numerical parameters \tilde{h}_{pq} and \tilde{g}_{pqrs} in the electronic Hamiltonian (equation 104) then reduce to the standard one- and two-electron molecular integrals h_{pq} and g_{pqrs} in equations (30) and (31).

The second-quantization Hamiltonian is a linear combination of excitation operators (the one-electron part) and products of two excitation operators (the two-electron part). Applied to any electronic state written as a linear combination of determinants, the Hamiltonian operator generates a new linear combination of determinants, related to the old one by single and double spin-orbital replacements (as well as by zero replacements). The exact form of the new state may be determined from the anticommutation relations of the elementary operators (equations 101–103) and from the numerical values of the one- and two-electron molecular integrals of the Hamiltonian. Note that, although the form of the Hamiltonian operator is independent of the number of electrons, the operator is number conserving, in the sense that, working on an N -particle electronic state, it creates a new N -particle state, having merely caused a redistribution of the N electrons among the available spin orbitals.

9.3 Unitary Orbital Transformations

The algebra of second quantization as outlined above pertains to one particular choice of orthonormal spin orbitals. For a different set of (orthonormal) spin orbitals, a new set of elementary operators – related to the old one by a unitary orbital transformation – is obtained. Indicating the new operators and states by overbars, we write the transformed determinants as

$$|\bar{\mu}\rangle = \bar{a}_{\mu_1}^\dagger \bar{a}_{\mu_2}^\dagger \dots \bar{a}_{\mu_N}^\dagger |\text{vac}\rangle \quad (110)$$

To relate the new and old elementary operators to each other, we first introduce the operator

$$\hat{\kappa} = \sum_{PQ} \kappa_{PQ} \bar{a}_P^\dagger a_Q; \quad \kappa_{PQ}^* = -\kappa_{QP} \quad (111)$$

which is anti-Hermitian, the numerical parameters κ_{PQ} being the elements of an anti-Hermitian matrix κ . Taking the exponential of this operator, we obtain the new number-conserving operator

$$\exp(\hat{\kappa}) = 1 + \hat{\kappa} + \frac{1}{2}\hat{\kappa}^2 + \frac{1}{6}\hat{\kappa}^3 + \dots \quad (112)$$

which, from the anti-Hermiticity of $\hat{\kappa}$, is easily seen to be unitary. If the matrix κ is chosen real and antisymmetric rather than complex and anti-Hermitian, then the operator $\exp(\hat{\kappa})$ becomes orthogonal rather than unitary. It may now be shown that, for any two orthonormal sets of spin orbitals, there exists an anti-Hermitian matrix κ such that

$$\bar{a}_P^\dagger = \exp(-\hat{\kappa}) a_P^\dagger \exp(\hat{\kappa}) \quad (113)$$

$$\bar{a}_P = \exp(-\hat{\kappa}) a_P \exp(\hat{\kappa}) \quad (114)$$

Inserting the equation (113) in equation (110), we obtain the following simple relationship between the determinants in the old and new spin-orbital representations:

$$|\bar{\mu}\rangle = \exp(-\hat{\kappa}) |\mu\rangle \quad (115)$$

Clearly, this relationship holds also for linear combinations of determinants. Note that this representation of unitary spin-orbital rotations and unitarily transformed states is independent, in the sense that the numerical parameters κ_{PQ} may be chosen freely in accordance with the simple requirement that κ constitutes an anti-Hermitian matrix.

The operator $\hat{\kappa}$ in the form of equation (111) is needed for a general unitary transformation. Often, however, we are interested in restricted unitary operators with special properties. Thus, the following operator, where the κ_{pq} are real-valued parameters, generates real unitary (i.e., orthogonal) transformations that conserve the spin symmetry of the transformed electronic state

$$\hat{\kappa} = \sum_{p>q} \kappa_{pq} E_{pq}^- \quad (116)$$

We here use the notation

$$E_{pq}^- = E_{pq} - E_{qp} \quad (117)$$

and note that the summation in equation (116) is over pairs of orbitals rather than over pairs of spin orbitals as in equation (111).

9.4 The Electronic Energy

Let us finally consider the evaluation of the electronic energy in second quantization. We denote our reference wavefunction by $|0\rangle$ and the unitarily transformed state by

$$|\kappa\rangle = \exp(-\hat{\kappa})|0\rangle \quad (118)$$

Taking the expectation value of the transformed state, we obtain

$$E(\mathbf{x}; \kappa) = \langle \kappa | \hat{H}(\mathbf{x}) | \kappa \rangle = \langle 0 | \exp(\hat{\kappa}) \hat{H}(\mathbf{x}) \exp(-\hat{\kappa}) | 0 \rangle \quad (119)$$

The unitarily transformed Hamiltonian may now be expanded in the following power series in the anti-Hermitian operator $\hat{\kappa}$:

$$\begin{aligned} \exp(\hat{\kappa}) \hat{H}(\mathbf{x}) \exp(-\hat{\kappa}) &= \hat{H}(\mathbf{x}) + [\hat{\kappa}, \hat{H}(\mathbf{x})] \\ &+ \frac{1}{2} [\hat{\kappa}, [\hat{\kappa}, \hat{H}(\mathbf{x})]] + \dots \end{aligned} \quad (120)$$

which is known as the Baker–Campbell–Hausdorff expansion of the Hamiltonian. Inserting this expansion in (equation 119), we obtain the following compact expression for the energy of the transformed state

$$\begin{aligned} E(\mathbf{x}; \kappa) &= \langle 0 | \hat{H}(\mathbf{x}) | 0 \rangle + \langle 0 | [\hat{\kappa}, \hat{H}(\mathbf{x})] | 0 \rangle \\ &+ \frac{1}{2} \langle 0 | [\hat{\kappa}, [\hat{\kappa}, \hat{H}(\mathbf{x})]] | 0 \rangle + \dots \end{aligned} \quad (121)$$

Restricting ourselves to orthogonal transformations and taking the derivatives of the energy (equation 121) with respect to the elements of the antisymmetric matrix κ , we obtain the following expressions for the electronic gradient and the electronic Hessian at $\kappa = 0$:

$$\left. \frac{\partial E(\mathbf{x}; \kappa)}{\partial \kappa_{pq}} \right|_0 = \langle 0 | [E_{pq}^-, \hat{H}(\mathbf{x})] | 0 \rangle \quad (122)$$

$$\begin{aligned} \left. \frac{\partial^2 E(\mathbf{x}; \kappa)}{\partial \kappa_{pq} \partial \kappa_{rs}} \right|_0 &= \frac{1}{2} \langle 0 | [E_{pq}^-, [E_{rs}^-, \hat{H}(\mathbf{x})]] \\ &+ [E_{rs}^-, [E_{pq}^-, \hat{H}(\mathbf{x})]] | 0 \rangle \end{aligned} \quad (123)$$

in terms of (nested) commutators of the Hamiltonian with antisymmetrized excitation operators.

Explicit expressions for the electronic energy, gradient, and Hessian may be derived by inserting the electronic Hamiltonian in the form of equation (104) and invoking the anticommutation relations of the elementary operators. For the energy, we obtain an expression of the form of equation (29), where the density elements are given by:

$$D_{pq} = \langle 0 | E_{pq} | 0 \rangle \quad (124)$$

$$d_{pqrs} = \langle 0 | E_{pq} E_{rs} - \delta_{rq} E_{ps} | 0 \rangle \quad (125)$$

In second quantization, therefore, the density elements are expectation values of the excitation operators, the numerical values of which may be obtained by the repeated application of the anticommutation relations (equations 101–103).

10 RELATED ARTICLES

Benchmark Studies on Small Molecules; Configuration Interaction; Coupled-cluster Theory; Density Functional Theory (DFT), Hartree–Fock (HF), and the Self-consistent Field; Geometry Optimization: I; Geometry Optimization: 2;

Integrals of Electron Repulsion; Molecular Magnetic Properties; Møller–Plesset Perturbation Theory; NMR Chemical Shift Computation: Ab Initio; Nonadiabatic Derivative Couplings; Normal Modes; Reaction Path Following; Spectroscopy: Computational Methods; Time-dependent Multi-configurational Hartree Method; Transition Structure Optimization Techniques.

11 REFERENCES

1. P. Pulay, *Adv. Chem. Phys.*, 1987, **69**, 241–286.
2. T. Helgaker and P. Jørgensen, *Adv. Quantum Chem.*, 1988, **19**, 183–245.
3. T. Helgaker and P. Jørgensen, in ‘Methods in Computational Molecular Physics’, eds. S. Wilson and G. H. F. Diercksen, Plenum Press, New York, 1992, pp. 353–421.
4. R. Shepard, in ‘Modern Electronic Structure Theory’, ed. D. R. Yarkony, World Scientific, Singapore, 1995, pp. 345–458.
5. P. Pulay, in ‘Modern Electronic Structure Theory’, ed. D. R. Yarkony, World Scientific, Singapore, 1995, pp. 1191–1240.
6. P. Jørgensen and J. Simons (eds.), ‘Geometrical Derivatives of Energy Surfaces and Molecular Properties (Proc. NATO Adv. Research Workshop)’, Reidel, Dordrecht, 1986.
7. P. Jørgensen and J. Simons, ‘Second Quantization-Based Methods in Quantum Chemistry’, Academic Press, New York, 1981.
8. P. Pulay, *J. Chem. Phys.*, 1983, **78**, 5043–5051.
9. S. Bratoz, *Colloq. Int. CNRS*, 1958, **82**, 287–301.
10. P. Pulay, *Mol. Phys.*, 1969, **17**, 197–204.
11. P. Pulay, *Mol. Phys.*, 1970, **18**, 473–480.
12. J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem.*, 1979, **S13**, 225–241.
13. T. U. Helgaker and J. Almlöf, *Int. J. Quantum Chem.*, 1984, **26**, 275–291.
14. P. Jørgensen and J. Simons, *J. Chem. Phys.*, 1983, **79**, 334–357.
15. N. C. Handy and H. F. Schaefer, III, *J. Chem. Phys.*, 1984, **81**, 5031–5033.
16. J. E. Rice and R. D. Amos, *Chem. Phys. Lett.*, 1985, **122**, 585–590.
17. T. Helgaker and P. Jørgensen, *Theor. Chim. Acta*, 1989, **75**, 111–127.