Theor Chim Acta (1989) 76: 227–245

Theoretica
Chimica Acta

A numerically stable procedure for calculating Møller–Plesset energy derivatives, derived using the theory of Lagrangians

Trygve Helgaker[1,*], Poul Jørgensen[1], and Nicholas C. Handy[2]
[1] Department of Chemistry, Århus University, DK-8000 Aarhus C, Denmark
[2] University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW, UK

(Received April 26, revised July 12/Accepted July 14, 1989)

Summary. When Møller–Plesset energy derivatives are determined in the canonical Hartree–Fock basis, singularities or instabilities may arise due to degeneracies among the occupied or unoccupied orbitals. If a non-canonical basis is used these singularities disappear. Numerically stable expressions are presented for the molecular gradient and Hessian of the second-order Møller–Plesset energy, obtained by differentiating a fully variational Lagrangian of the energy constructed in a non-canonical representation. By using a non-canonical representation, singularities and instabilities are avoided, and the variational property of the Lagrangian ensures that Wigner's $2n+1$ rule is satisfied for the orbital derivatives and that the multipliers satisfy the stronger $2n+2$ rule. It is shown that the most expensive step in the calculation of the Hessian scales as Mn^4o, where M is the number of independent Cartesian distortions, n the total number of orbitals, and o the number of occupied orbitals.

Key words: Energy derivatives — Møller–Plesset — Non-canonical basis — Lagrangian formulation

1. Introduction

Møller–Plesset perturbation theory constitutes a simple and powerful technique for calculating the correlation energy of closed-shell molecular systems. The method has recently received new interest with the realization that molecular gradients and Hessians of the second-order Møller–Plesset (MP2) energy can be calculated at a cost comparable to that of uncorrelated wave functions [1–4],

---

[*] *Present address:* Department of Chemistry, University of Oslo, Blindern, N-0315 Oslo 3, Norway

yielding results in significantly better agreement with experiment [5]. For many purposes Møller–Plesset theory provides the most cost effective tool in quantum chemistry.

Møller–Plesset energies are conveniently calculated in the canonical Hartree–Fock basis, and this representation has traditionally also been used for deriving analytical expressions for the geometrical derivatives. A problem with this approach is that derivatives of degenerate orbitals are undefined. This leads to singularities in the expressions for the molecular derivatives and makes their calculation unstable. In practice, singularities appear in calculations on highly symmetric molecules when the molecule is treated in a lower symmetry group. Instabilities, on the other hand, may arise in calculations on molecules slightly distorted from degenerate symmetry, or in calculations involving several diffuse orbitals. Accidental degeneracies or near degeneracies may also occur when tracing potential energy curves. The singularities or instabilities are spurious in the sense that they do not appear when derivatives of the energy are calculated by numerical differentiation. They occur only because we insist on expressing the derivatives in the canonical representation, in which accidentally degenerate orbitals are not uniquely defined. In this communication we demonstrate that numerically stable expressions for the derivatives of the Møller–Plesset energies may be derived using a non-canonical representation.

The first analytical expression for the second-order Møller–Plesset molecular gradient was derived and implemented ten years ago by Pople et al. [1], and in 1983 an expression for the molecular Hessian was presented by Jørgensen and Simons [6]. A significant advance was made in 1986 by Handy et al. [2] who simplified the calculation of gradients and reported the first analytical calculation of Møller–Plesset Hessians. Using the Handy–Schaefer technique [7], these authors eliminated the highest-order coupled perturbed Hartree–Fock (CPHF) equations from the expressions and made the calculation of Møller–Plesset derivatives a practical proposition. In 1988 Jørgensen and Helgaker [8] simplified the structure of the derivative expressions using a variational technique based on Lagrange's method of undetermined multipliers, and presented expressions for derivatives up to third order. The variational property of the Lagrangian ensures that the variational parameters satisfy Wigner's $2n + 1$ rule [9, 10]. In particular, the Lagrange multipliers satisfy the stronger $2n + 2$ rule [8, 11]. The above papers all use the canonical representation in deriving the analytical expressions for the derivatives and are therefore unstable. The instabilities have been considered by Handy et al. [3], who eliminated the singularities from the molecular gradient by reformulating the original equations and also derived an expression for the Hessian in a non-canonical representation, using an *ad hoc* procedure. A non-canonical formulation of the second-order Møller–Plesset gradient has also been given by Pulay and Sæbø [12], based on the Hylleraas variation method.

In this paper we use the Lagrangian technique to derive numerically stable expressions for the MP2 gradient and Hessian. We construct a variational Lagrangian of the energy in a non-canonical representation which reduces to the canonical representation at the reference geometry. Only variables which are not redundant with respect to the optimization of the total energy and their multipliers enter the Lagrangian. In the calculation of the gradient we must solve the zeroth-order equations for the Lagrange multipliers, and for the Hessian we must also solve the first-order equations for the orbital responses. As usual, by constructing a set of effective densities we avoid transforming the highest-order derivative integrals to the molecular orbital (MO) basis [2, 13]. By introducing a set of one-electron MP2 Fock matrices we have broken the time-consuming contributions to the Hessian into parts which depend on one nuclear distortion only. As a consequence, the most expensive step scales as Mn^4o. Here M is the number of independent geometrical distortions, n the total number of orbitals, and o and v the number of occupied and virtual orbitals, respectively. The additional terms introduced by the non-canonical representation scale as Mv^3o^2. The non-canonical expression for the Hessian is therefore computationally not much more demanding than the corresponding expression obtained in the unstable canonical representation.

In Sect. 2 we summarize Møller–Plesset perturbation theory, providing background and notation for the subsequent development, and we construct the Lagrangian of the second-order Møller–Plesset energy. Expressions for the MP2 molecular gradient and Hessian are presented in Sect. 3 together with the necessary response equations. Section 4 discusses the calculation of these expressions, and in Sect. 5 we compare our derivative expressions with previously derived numerically unstable expressions. The last section contains some concluding remarks.

2. Møller–Plesset perturbation theory

In this section we discuss the calculation of Møller–Plesset energies in canonical and non-canonical representations. We point out that although the canonical representation is convenient for calculating energies, it may introduce singularities and numerical instabilities when used in derivative calculations. We then construct a variational Lagrangian for the second-order Møller–Plesset energy in a non-canonical representation.

2.1. The Møller–Plesset energy at the reference geometry

We write the Hamiltonian at the reference geometry as

$$H = \sum_{pq} h_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs}(E_{pq}E_{rs} - \delta_{rq}E_{ps}), \tag{1}$$

where h_{pq} and g_{pqrs} are the usual one- and two-electron integrals in Mulliken notation and the summations are over the full set of orbitals. The operators E_{pq} are defined in terms of creation and annihilation operators as

$$E_{pq} = \sum_{\sigma} a^{+}_{p\sigma} a_{q\sigma} \tag{2}$$

summing over spin. We use p, q, r, s to denote general (unspecified) orbitals.

In Møller–Plesset perturbation theory the unperturbed Hamiltonian is the Fock operator [14]

$$F = \sum_{pq} F_{pq} E_{pq}, \tag{3}$$

$$F_{pq} = \frac{1}{2} \sum_{\sigma} \langle HF|[a^{+}_{p\sigma}, [a_{q\sigma}, H]]_{+}|HF\rangle, \tag{4}$$

where the Hartree–Fock state $|HF\rangle$ fulfills the Brillouin condition

$$\langle HF|[E_{ia}, H]|HF\rangle = 0. \tag{5}$$

The indices *ijkl* and *abcd* denote occupied and unoccupied orbitals, respectively. No matrix elements connect the Hartree–Fock state and singly excited states $E_{ai}|HF\rangle$, and the Hartree–Fock state interacts directly with doubly excited states only. Inserting the Hamiltonian of Eq. (1) into Eq. (4) we obtain an explicit expression for the Fock matrix

$$F_{pq} = h_{pq} + \sum_{i} k_{pqii}, \tag{6}$$

where

$$k_{pqrs} = 2g_{pqrs} - g_{psrq} \tag{7}$$

is introduced for convenience.

The second-order Møller–Plesset energy is calculated from the expression

$$E = \langle HF|H|HF\rangle + \sum_{uv} \langle HF|H|u\rangle[\boldsymbol{G}^{-1}]_{uv}\langle v|H|HF\rangle, \tag{8}$$

where

$$G_{uv} = \delta_{uv}\langle HF|F|HF\rangle - \langle u|F|v\rangle. \tag{9}$$

Only doubly excited states enter the second-order perturbation expansion. The Brillouin condition, Eq. (5), implies that the Fock matrix is block diagonal since for occupied-virtual rotations [15]

$$\sum_{\sigma} \langle HF|[a^{+}_{i\sigma}, [a_{a\sigma}, H]]_{+}|HF\rangle = \langle HF|[E_{ia}, H]|HF\rangle. \tag{10}$$

The Møller–Plesset energy, Eq. (8), is invariant to unitary rotations among occupied and among virtual orbitals. The energy is therefore best calculated in the canonical basis, in which the Fock matrix and therefore $\boldsymbol{G}$ are diagonal:

$$\sum_{\sigma} \langle HF|[a^{+}_{p\sigma}, [a_{q\sigma}, H]]_{+}|HF\rangle = 2\varepsilon_p \delta_{pq}. \tag{11}$$

In terms of integrals the orbital energies are given as

$$\varepsilon_p = h_{pp} + \sum_{i} k_{ppii}. \tag{12}$$

Note that the canonical condition, Eq. (11), contains the Brillouin condition, Eq. (5), as a special case, and that it determines the orbitals uniquely except when the orbital energies ε_p are degenerate.

Before considering the Møller–Plesset energy at displaced geometries let us write the second-order energy explicitly in the canonical basis. For this purpose we use the biorthonormal basis [16]

$$|{}^{ab}_{ij}\rangle = \tfrac{1}{3}(1+\delta_{ab}\delta_{ij})^{-1}(2E_{ai}E_{bj}+E_{aj}E_{bi})|HF\rangle, \tag{13}$$

$$\langle{}^{ab}_{ij}| = \tfrac{1}{2}\langle HF|E_{ia}E_{jb}, \tag{14}$$

where $ai \geqslant bj$. In this representation the Hamiltonian matrix elements between the Hartree–Fock state and the excited states are

$$\langle HF|H|{}^{ab}_{ij}\rangle = \frac{2}{1+\delta_{ab}\delta_{ij}} g_{iajb}, \tag{15}$$

$$\langle{}^{ab}_{ij}|H|HF\rangle = k_{iajb}, \tag{16}$$

and the only nonvanishing elements of $\boldsymbol{G}$ are

$$\langle{}^{ab}_{ij}|G|{}^{ab}_{ij}\rangle = \varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b. \tag{17}$$

Inserting the above elements, Eqs. (15–17), into the energy expression, Eq. (8), we obtain

$$E = \langle HF|H|HF\rangle + \sum_{iajb} \frac{g_{iajb}k_{iajb}}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b} \tag{18}$$

without restrictions on the summation range.

For future reference we write the first-order Møller–Plesset correction to the wave function

$$|MP1\rangle = \sum_u |u\rangle G^{-1}_{uu}\langle u|H|HF\rangle \tag{19}$$

in the canonical representation. Inserting the matrix elements above we find

$$|MP1\rangle = \sum_{ai \geqslant bj} |{}^{ab}_{ij}\rangle t_{iajb}, \tag{20}$$

$$\langle MP1| = \sum_{ai \geqslant bj} \frac{2}{1+\delta_{ab}\delta_{ij}} u_{iajb}\langle{}^{ab}_{ij}|, \tag{21}$$

where the amplitudes are given by

$$t_{iajb} = \frac{k_{iajb}}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b}, \tag{22}$$

$$u_{iajb} = \frac{g_{iajb}}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b}. \tag{23}$$

The second-order Møller–Plesset energy may be written in the following equivalent forms

$$E = \langle HF|H|HF\rangle + \sum_{iajb} g_{iajb}t_{iajb} = \langle HF|H|HF\rangle + \sum_{iajb} u_{iajb}k_{iajb} \tag{24}$$

in terms of integrals and amplitudes.

2.2. The Møller–Plesset energy at displaced geometries

In order to consider the variation of the Møller–Plesset energy with geometry we must relate the Hamiltonian $H(x)$ and the unperturbed states $|HF(x)\rangle$ and $|u(x)\rangle$ at displaced geometries to the Hamiltonian and states at the reference geometry. We therefore express the Hamiltonian and states at displaced geometries in terms of a set of symmetrically orthonormalized molecular orbitals (OMO's) which reduces to the canonical basis at the reference geometry. The orthonormality of the basis ensures that we may neglect the geometry dependence of the creation and annihilation operators and the states constructed from such operators when taking derivatives of the energy [6, 17]. For ease of notation we therefore omit any reference to the geometry in the OMO creation and annihilation operators. The OMO basis is chosen because of the simple analytical expressions for its derivatives, but any other orthonormal basis may be used provided its geometry dependence is well defined. As shown elsewhere [17, 18] the Hamiltonian may be written as

$$H(\boldsymbol{x}) = \sum_{pq} \mathscr{h}_{pq}(\boldsymbol{x}) E_{pq} + \frac{1}{2} \sum_{pqrs} \mathscr{g}_{pqrs}(\boldsymbol{x})[E_{pq}E_{rs} - \delta_{rq}E_{ps}], \tag{25}$$

where the integrals are calculated in terms of the OMO's as

$$\mathscr{h}_{pq}(\boldsymbol{x}) = \sum_{tu} h_{tu}(\boldsymbol{x})[\boldsymbol{S}^{-1/2}(\boldsymbol{x})]_{tp}[\boldsymbol{S}^{-1/2}(\boldsymbol{x})]_{uq}, \tag{26}$$

$$\mathscr{g}_{pqrs}(\boldsymbol{x}) = \sum_{tuvw} g_{tuvw}(\boldsymbol{x})[\boldsymbol{S}^{-1/2}(\boldsymbol{x})]_{tp}[\boldsymbol{S}^{-1/2}(\boldsymbol{x})]_{uq}[\boldsymbol{S}^{-1/2}(\boldsymbol{x})]_{vr}[\boldsymbol{S}^{-1/2}(\boldsymbol{x})]_{ws}. \tag{27}$$

Indices *tuvw* denote general orbitals. The Hartree–Fock state, the excited states, and the Hartree–Fock orbitals at displaced geometries may be written in the OMO basis as

$$|HF(\boldsymbol{x})\rangle = \exp(-\kappa)|HF\rangle, \tag{28}$$

$$|u(\boldsymbol{x})\rangle = \exp(-\kappa)|u\rangle, \tag{29}$$

$$a_{p\sigma}^{+}(\boldsymbol{x}) = \exp(-\kappa)a_{p\sigma}^{+}\exp(\kappa), \tag{30}$$

where $|HF\rangle$ and $|u\rangle$ are constructed from OMO creation operators. The antisymmetric orbital rotation operator κ is given by

$$\kappa = \sum_{p>q} \kappa_{pq}(\boldsymbol{x})E_{pq}^{-} = \sum_{p>q} \kappa_{pq}(\boldsymbol{x})[E_{pq} - E_{pq}], \tag{31}$$

where for ease of notation we omit the geometry argument in $\kappa(\boldsymbol{x})$. The exact summation range and the determination of the parameters κ_{pq} are discussed later.

Replacing the geometry-independent states and operators of the previous section by their geometry-dependent counterparts, Eqs. (28–30), we obtain the following expression for the second-order Møller–Plesset energy at displaced geometries:

$$E(\boldsymbol{x}) = \langle HF|\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)|HF\rangle + \sum_{uv}\langle HF|\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)|u\rangle \times [\boldsymbol{G}(\boldsymbol{x})^{-1}]_{uv}\langle v|\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)|HF\rangle, \tag{32}$$

where

$$G_{uv}(\boldsymbol{x}) = \delta_{uv}\langle HF|\exp(\kappa)F(\boldsymbol{x})\exp(-\kappa)|HF\rangle - \langle u|\exp(\kappa)F(\boldsymbol{x})\exp(-\kappa)|v\rangle \quad (33)$$

and

$$\exp(\kappa)F(\boldsymbol{x})\exp(-\kappa) = \frac{1}{2}\sum_{pq\sigma}\langle HF|[a^{+}_{p\sigma},[a_{q\sigma},\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)]]_{+}|HF\rangle E_{pq}. \quad (34)$$

To specify the Møller–Plesset energy completely at displaced energies, we must determine the geometry dependence of the operator κ.

As noted above it is convenient to calculate the Møller–Plesset energy at the reference geometry in the canonical representation since the unperturbed Hamiltonian then becomes diagonal. It is therefore tempting to use the canonical representation at displaced geometries as well and require that the canonical condition, Eq. (11), is fulfilled at all geometries:

$$\sum_{\sigma}\langle HF|[a^{+}_{p\sigma},[a_{q\sigma},\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)]]_{+}|HF\rangle = 2\varepsilon_{p}(\boldsymbol{x})\delta_{pq}. \quad (35)$$

Expanding this equation around the reference geometry and setting each term equal to zero we arrive at the CPHF equations [19, 20] which determine the geometry dependence of ε_p and κ_{pq} for *all* $p > q$. Unfortunately, this procedure breaks down when the Hartree–Fock orbitals at the reference geometry contains degenerate or nearly degenerate orbitals, since the CPHF equations then become singular or numerically unstable. These problems are avoided if, instead of insisting on the canonical condition of Eq. (11), we use the weaker Brillouin condition of Eq. (5) at displaced geometries:

$$\langle HF|[E_{ia},\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)]|HF\rangle = 0. \quad (36)$$

Expanding this equation around the reference geometry and setting each term equal to zero we obtain the Hartree–Fock response equations [15, 20] which determine the geometry dependence of κ_{ai}. Only occupied–unoccupied rotations now enter the orbital rotation operator [15]

$$\kappa = \sum_{ai}\kappa_{ai}(\boldsymbol{x})E^{-}_{ai}. \quad (37)$$

Note that although we do not require the canonical condition to be satisfied at displaced geometries, it still holds at the reference geometry. Therefore, when expanding the Fock operator, Eq. (34), in geometrical displacements, the undifferentiated term is diagonal and the remaining terms block diagonal only. In the derivative expressions only terms containing the undifferentiated Fock operator are inverted. The non-diagonal derivative Fock matrices are easily constructed and enter the calculation of the Hessian through one-index transformations scaling as Mn^4o.

2.3. The Møller–Plesset energy Lagrangian

In the previous section we gave an expression for the Møller–Plesset energy which is valid at all geometries and reduces to the canonical expression at the reference geometry. Together with the Brillouin constraints of Eq. (36) it completely determines the energy at displaced geometries. To avoid constrained variables when taking derivatives we use Lagrange's method of undetermined multipliers and construct the Lagrangian

$$\begin{aligned} W(\boldsymbol{x}, \boldsymbol{\kappa}, \boldsymbol{\zeta}) = {} & \langle HF|\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)|HF\rangle \\ & + \sum_{uv} \langle HF|\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)|u\rangle[G(\boldsymbol{x})^{-1}]_{uv}\langle v|\exp(\kappa)H(\boldsymbol{x}) \\ & \times \exp(-\kappa)|HF\rangle + \sum_{ia}\zeta_{ia}\langle HF|[E_{ia}, \exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)]|HF\rangle \end{aligned} \tag{38}$$

introducing one multiplier ζ_{ia} for each Brillouin constraint [8, 11]. Note that the geometry dependence of the Lagrangian is completely isolated in the operator $\exp(\kappa)H(\boldsymbol{x})\exp(-\kappa)$. The Lagrangian fulfills the variational conditions

$$\frac{\partial W(\boldsymbol{x}, \boldsymbol{\kappa}, \boldsymbol{\zeta})}{\partial \boldsymbol{\kappa}} = \mathbf{0}, \tag{39}$$

$$\frac{\partial W(\boldsymbol{x}, \boldsymbol{\kappa}, \boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}} = \mathbf{0}, \tag{40}$$

at all geometries. Expanding these conditions in geometrical displacements and setting each term equal to zero, we obtain the Møller–Plesset response equations which determine $\boldsymbol{\kappa}$ and $\boldsymbol{\zeta}$.

The Lagrangian is of no use for calculating energies since the constraints are trivially fulfilled by first carrying out a separate Hartree–Fock calculation and then using the obtained orbitals and orbital energies in the subsequent Møller–Plesset calculation. In contrast, when calculating energy derivatives the Lagrangian is very useful since its variational character ensures that we may take advantage of the simplifications which occur for fully variational wave functions. In particular, Wigner's $2n+1$ rule holds for the non-linear parameters κ_{ia} while the linear parameters ζ_{ia} obey the stronger $2n+2$ rule. If we had used the canonical condition at displaced geometries, occupied–occupied κ_{ij} rotations, unoccupied–unoccupied κ_{ab} rotations, orbital energies ε_p, and their multipliers would also enter the Lagrangian [8].

3. Møller–Plesset energy derivatives

In the following we first present the general expressions for derivatives of variational Lagrangians. Inserting the Lagrangian constructed above we then obtain the Møller–Plesset derivatives and response equations.

3.1. *Derivatives of variational Lagrangians*

General expressions for derivatives of variational Lagrangians have been reported elsewhere [8, 11] and are reproduced here for easy reference. We denote the total derivatives of a function $f(\boldsymbol{x}, \boldsymbol{\kappa})$ with respect to $\boldsymbol{x}$ by $f^{(m)}$, and the partial derivatives with respect to $\boldsymbol{x}$ and $\boldsymbol{\kappa}$ by $f^{(mn)}$. The gradient and Hessian of the Lagrangian

$$W(\boldsymbol{x}, \boldsymbol{\kappa}, \zeta) = E(\boldsymbol{x}, \boldsymbol{\kappa}) + \zeta e(\boldsymbol{x}, \boldsymbol{\kappa}) \tag{41}$$

are then given by

$$W^{(1)} = E^{(10)} + \zeta^{(0)} e^{(10)}, \tag{42}$$

$$W^{(2)} = E^{(20)} + 2E^{(11)}\boldsymbol{\kappa}^{(1)} + E^{(02)}\boldsymbol{\kappa}^{(1)}\boldsymbol{\kappa}^{(1)} + \zeta^{(0)}(e^{(20)} + 2e^{(11)}\boldsymbol{\kappa}^{(1)} + e^{(02)}\boldsymbol{\kappa}^{(1)}\boldsymbol{\kappa}^{(1)}), \tag{43}$$

where $\zeta^{(0)}$ and $\boldsymbol{\kappa}^{(1)}$ are obtained from the response equations

$$\zeta^{(0)} e^{(01)} = -E^{(01)}, \tag{44}$$

$$e^{(01)}\boldsymbol{\kappa}^{(1)} = -e^{(10)}. \tag{45}$$

We emphasize that these equations are completely general and can be applied to any wave function containing constrained variables. We note that the gradient requires $\zeta^{(0)}$ only, whereas the Hessian requires $\zeta^{(0)}$ and $\boldsymbol{\kappa}^{(1)}$. In the next section the above expressions are used to determine the molecular gradient and Hessian of the second-order Møller–Plesset energy.

3.2. *The Møller–Plesset molecular gradient and Hessian*

Inserting the Lagrangian, Eq. (38), into the derivative expressions, Eqs. (42) and (43), we obtain the Møller–Plesset molecular gradient

$$\begin{aligned} W^{(1)} = \langle HF|H^{(1)}|HF\rangle &+ 2\sum_u \frac{\langle HF|H^{(1)}|u\rangle\langle u|H^{(0)}|HF\rangle}{G_{uu}^{(0)}} \\ &- \sum_{uv} \frac{\langle HF|H^{(0)}|u\rangle G[H^{(1)}]_{uv}\langle v|H^{(0)}|HF\rangle}{G_{uu}^{(0)}G_{vv}^{(0)}} + \sum_{ia} \zeta_{ia}^{(0)}\langle HF|[E_{ia}, H^{(1)}]|HF\rangle \end{aligned} \tag{46}$$

and the Møller–Plesset molecular Hessian

$$\begin{aligned} W^{(2)} = \langle HF|J^{(2)}|HF\rangle &+ 2\sum_u \frac{\langle HF|J^{(2)}|u\rangle\langle u|H^{(0)}|HF\rangle}{G_{uu}^{(0)}} \\ &- \sum_{uv} \frac{\langle HF|H^{(0)}|u\rangle G[J^{(2)}]_{uv}\langle v|H^{(0)}|HF\rangle}{G_{uu}^{(0)}G_{vv}^{(0)}} \\ &+ \sum_{ia} \zeta_{ia}^{(0)}\langle HF|[E_{ia}, J^{(2)}]|HF\rangle + 2\sum_u \frac{V_u^{(1)}\tilde{V}_u^{(1)}}{G_{uu}^{(0)}}. \end{aligned} \tag{47}$$

In these equations the tilde indicates conjugation and $G[A]$ means that the Fock operator entering the matrix is calculated from the operator A rather than the Hamiltonian H as in Eq. (4). We have also used the following notation:

$$J^{(1)} = H^{(1)} + [\kappa^{(1)}, H^{(0)}], \tag{48}$$

$$J^{(2)} = H^{(2)} + 2[\kappa^{(1)}, H^{(1)}] + [\kappa^{(1)}, \kappa^{(1)}, H^{(0)}], \tag{49}$$

$$V_u^{(1)} = \langle HF|J^{(1)}|u\rangle - \sum_v \frac{\langle HF|H^{(0)}|v\rangle G[J^{(1)}]_{vu}}{G_{vv}^{(0)}}. \tag{50}$$

To display the structure of the molecular derivatives more clearly we introduce the matrix

$$\varrho_{pq} = \begin{bmatrix} \mathscr{D}_{ij} & \zeta_{ia} \\ \zeta_{ia} & \mathscr{D}_{ab} \end{bmatrix}, \tag{51}$$

where the diagonal blocks contain the one-electron density

$$\mathscr{D}_{pq} = \langle MP1|E_{pq}|MP1\rangle - \langle MP1|MP1\rangle\langle HF|E_{pq}|HF\rangle. \tag{52}$$

Inserting the biorthonormal expressions for the first-order Møller–Plesset correction, Eqs. (13) and (14), we find that the only non-vanishing density elements are

$$\mathscr{D}_{ij} = -2\sum_{akb} u_{iakb} t_{jakb}, \tag{53}$$

$$\mathscr{D}_{ab} = 2\sum_{ijc} u_{iajc} t_{ibjc}. \tag{54}$$

Collecting terms, we may now write the derivatives more succinctly as

$$W^{(1)} = \langle HF|H^{(1)}|HF\rangle + 2\langle HF|H^{(1)}|MP1\rangle + \operatorname{Tr} \varrho F[H^{(1)}] \tag{55}$$

and

$$W^{(2)} = \langle HF|J^{(2)}|HF\rangle + 2\langle HF|J^{(2)}|MP1\rangle + \operatorname{Tr} \varrho F[J^{(2)}] + 2\sum_u \frac{V_u^{(1)}\tilde{V}_u^{(1)}}{G_{uu}^{(0)}}. \tag{56}$$

We note the similarity between the expressions for the gradient and the Hessian. In fact, the first three contributions to the Hessian are identical to the terms in the gradient except that the Hamiltonian $H^{(1)}$ is replaced by $J^{(2)}$. We shall see that these terms are evaluated in much the same way for the gradient and the Hessian, and that in either case there is no need to transform the highest-order derivative integrals to the MO (molecular orbital) basis. The Hessian requires, however, the transformation of first-derivative integrals with two occupied and two virtual indices.

The gradient and Hessian expressions presented above may be given a simple physical interpretation. The gradient, Eq. (55), has no terms describing the response of the orbitals to nuclear displacements, and contains in this sense static terms only. On the other hand, to calculate the Hessian we need the first-order response of the orbitals which enters each of the four terms in Eq. (56). This is in accordance with Wigner's $2n+1$ rule. If we regard the second-order Møller–Plesset energy as obtained by combining matrix elements g_{iajb} with amplitudes

t_{iajb} (as displayed in Eq. (24)), the gradient contains the first derivatives of the matrix elements and the undifferentiated amplitudes. In contrast, the Hessian also contains the first derivatives of the amplitudes

$$\tilde{t}_u^{(1)} = \frac{\tilde{V}_u^{(1)}}{G_{uu}^{(0)}} \tag{57}$$

in the last term in Eq. (56).

3.3. *The Møller–Plesset response equations*

To determine the response equations we proceed in exactly the same way as for the derivatives above. We obtain for the Lagrange multipliers $\boldsymbol{\zeta}^{(0)}$

$$\begin{aligned}\sum_{jb} \zeta_{jb}^{(0)} \langle HF|[E_{jb}, [E_{ai}^{-}, H^{(0)}]]|HF\rangle &= -2\sum_{u} \frac{\langle HF|[E_{ai}^{-}, H^{(0)}]|u\rangle\langle u|H^{(0)}|HF\rangle}{G_{uu}^{(0)}} \\ &\quad + \sum_{uv} \frac{\langle HF|H^{(0)}|u\rangle G\{[E_{ai}^{-}, H^{(0)}]\}_{uv}\langle v|H^{(0)}|HF\rangle}{G_{uu}^{(0)}G_{vv}^{(0)}},\end{aligned} \tag{58}$$

and for the orbital responses $\boldsymbol{\kappa}^{(1)}$

$$\sum_{bj} \langle HF|[E_{ia}, [E_{bj}^{-}, H^{(0)}]]|HF\rangle\kappa_{bj}^{(1)} = -\langle HF|[E_{ia}, H^{(1)}]|HF\rangle. \tag{59}$$

The linear equations (58) and (59) have the same coefficient matrix

$$A_{aibj} = \langle HF|[E_{ia}, [E_{bj}^{-}, H^{(0)}]]|HF\rangle, \tag{60}$$

which may be shown to be symmetric in *ai* and *bj*. Using Eq. (52) and collecting terms we may write the linear equations as

$$\sum_{bj} A_{aibj}\zeta_{jb}^{(0)} = -2\langle HF|[E_{ai}^{-}, H^{(0)}]|MP1\rangle - \sum_{pq} \mathscr{D}_{pq}F\{[E_{ai}^{-}, H^{(0)}]\}_{pq} \tag{61}$$

and

$$\sum_{bj} A_{aibj}\kappa_{bj}^{(1)} = -\langle HF|[E_{ia}, H^{(1)}]|HF\rangle. \tag{62}$$

Equation (61) corresponds to the Z-vector equation of Handy et al. [2], and Eq. (62) is the usual equation for the first-order response of the Hartree–Fock orbitals, independent of Møller–Plesset theory. Equations (61) and (62) differ in the right-hand sides only.

4. Computational considerations

In this section we analyze the expressions for the molecular gradient and Hessian derived in the previous section with emphasis on computational aspects. In particular, we identify all terms in the derivatives whose calculation is proportional to the fifth power in the number of orbitals (occupied o, virtual v or general n) in order to pinpoint the computationally most demanding steps. The first contributions to the gradient (55) and the Hessian (56) are simply the derivatives of the Hartree–Fock energy. Their calculation is well known and involves only terms to the fourth power (or lower) in the number of orbitals. In the following we concentrate on the remaining contributions to the derivatives.

4.1. The calculation of $\langle HF|H^{(1)}|MP1\rangle$ and $\langle HF|J^{(2)}|MP1\rangle$

The second contributions to the gradient, Eq. (55), and Hessian, Eq. (56), are

$$2\langle HF|H^{(1)}|MP1\rangle = 2\sum_{iajb} t_{iajb}\, \mathcal{g}^{(1)}_{iajb}, \tag{63}$$

$$2\langle HF|J^{(2)}|MP1\rangle = 2\sum_{iajb} t_{iajb}\, \mathcal{j}^{(2)}_{iajb}, \tag{64}$$

where t_{iajb} are the only non-vanishing elements of the transition density between the Hartree–Fock state and the first-order Møller–Plesset correction. The integrals $\mathcal{g}^{(1)}$ and $\mathcal{j}^{(2)}$ are given by [use Eqs. (26, 27) and Eq. (49)]

$$\mathcal{g}^{(1)} = \boldsymbol{g}^{(1)} - \tfrac{1}{2}\{\boldsymbol{S}^{(1)}, \boldsymbol{g}^{(0)}\}, \tag{65}$$

$$\begin{aligned}\mathcal{j}^{(2)} = \boldsymbol{g}^{(2)} &- \tfrac{1}{2}\{\boldsymbol{S}^{(2)} - \boldsymbol{S}^{(1)}\boldsymbol{S}^{(1)}, \boldsymbol{g}^{(0)}\} + \{2\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)}, \boldsymbol{g}^{(1)}\} \\ &+ \{\boldsymbol{\kappa}^{(1)}, \{\boldsymbol{\kappa}^{(1)}, \boldsymbol{g}^{(0)}\}\} + \tfrac{1}{4}\{\boldsymbol{S}^{(1)}, \{\boldsymbol{S}^{(1)}, \boldsymbol{g}^{(0)}\}\} - \{\boldsymbol{\kappa}^{(1)}, \{\boldsymbol{S}^{(1)}, \boldsymbol{g}^{(0)}\}\}.\end{aligned} \tag{66}$$

Braces denote one-index transformations, for example in Eq. (65):

$$\{\boldsymbol{S}^{(1)}, \boldsymbol{g}^{(0)}\}_{pqrs} = \sum_t (S^{(1)}_{pt} g^{(0)}_{tqrs} + S^{(1)}_{qt} g^{(0)}_{ptrs} + S^{(1)}_{rt} g^{(0)}_{pqts} + S^{(1)}_{st} g^{(0)}_{pqrt}), \tag{67}$$

where the summation runs over the full set of orbitals. General expressions for differentiated integrals in terms of one-index transformations are given in [18, 20]. We also note that commutators of the Hamiltonian give rise to one-index transformed integrals. For example, in Eq. (66) the one-index transformations with $\boldsymbol{\kappa}^{(1)}$ arise from the commutators $[\kappa^{(1)}, H^{(1)}]$ and $[\kappa^{(1)}, \kappa^{(1)}, H^{(0)}]$ of Eq. (49).

To calculate the terms Eqs. (63) and (64) containing the integrals Eqs. (65) and (66) it is helpful to introduce the intermediates

$$\Phi^{(n)}_{ip} = 2\sum_{ajb} t_{iajb} g^{(n)}_{pajb}, \tag{68}$$

$$\Phi^{(n)}_{ap} = 2\sum_{ijb} t_{iajb} g^{(n)}_{ipjb}. \tag{69}$$

The operation count for the calculation of $\boldsymbol{\Phi}^{(n)}$ for one geometrical distortion is proportional to nv^2o^2. The matrix $\boldsymbol{\Phi}^{(n)}$ shares many of the properties of the generalized Fock matrix of Siegbahn et al. [21] and will therefore be referred to as the MP2 Fock matrix. In particular, it has the property that

$$\sum_{iajb} t_{iajb}\{\boldsymbol{A}, \boldsymbol{g}^{(n)}\}_{iajb} = \mathrm{Tr}\, \boldsymbol{A}\boldsymbol{\Phi}^{(n)}. \tag{70}$$

Equations (63) and (64) may now be written

$$2\langle HF|H^{(1)}|MP1\rangle = 2\sum_{iajb} t_{iajb} g^{(1)}_{iajb} - \mathrm{Tr}\, \boldsymbol{S}^{(1)}\boldsymbol{\Phi}^{(0)}, \tag{71}$$

$$\begin{aligned} 2\langle HF|J^{(2)}|MP1\rangle = {} & 2\sum_{iajb} t_{iajb}\, g^{(2)}_{iajb} - \mathrm{Tr}(\boldsymbol{S}^{(2)} - \boldsymbol{S}^{(1)}\boldsymbol{S}^{(1)})\boldsymbol{\Phi}^{(0)} \\ & + 2\,\mathrm{Tr}(2\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)})\boldsymbol{\Phi}^{(1)} + 2\,\mathrm{Tr}\,\boldsymbol{\kappa}^{(1)}\boldsymbol{\Phi}^{\{\kappa^{(1)}, g^{(0)}\}} \\ & - \mathrm{Tr}(2\boldsymbol{\kappa}^{(1)} - \tfrac{1}{2}\boldsymbol{S}^{(1)})\boldsymbol{\Phi}^{\{S^{(1)}, g^{(0)}\}} \end{aligned} \tag{72}$$

where we use the notation $\boldsymbol{\Phi}^{\{S^{(1)}, g^{(0)}\}}$ to indicate that the MP2 Fock matrix has been calculated from the one-index transformed integrals $\{\boldsymbol{S}^{(1)}, \boldsymbol{g}^{(0)}\}$.

The first terms in Eqs. (71) and (72) consist of amplitudes multiplied by differentiated integrals, and are best calculated by transforming the amplitudes to the contravariant AO (atomic orbital) basis in a process scaling as n^4o. Adding these densities to the two-electron Hartree–Fock densities, the first terms in Eqs. (71) and (72) may be calculated simultaneously with the Hartree–Fock derivatives at little extra cost.

The remaining terms in Eqs. (71) and (72) are easily calculated once the various MP2 Fock matrices are available. The undifferentiated Fock matrix $\boldsymbol{\Phi}^{(0)}$ is conveniently calculated in the MO basis in a process proportional to nv^2o^2. The necessary integrals (one occupied, one virtual, and two general indices) are already available in the MO basis from the calculation of the energy. In fact, $\boldsymbol{\Phi}^{(0)}$ also enters the right-hand side of the zeroth-order Lagrange multiplier equations and should therefore already be available at this stage.

The differentiated matrix $\boldsymbol{\Phi}^{(1)}$ can be calculated in the MO basis in the same way as $\boldsymbol{\Phi}^{(0)}$. This requires the transformation of derivative integrals to the MO basis (two general, one virtual, and one occupied index) scaling as Mn^4o, and the construction of the Fock matrices scaling as Mnv^2o^2. Alternatively, the $\boldsymbol{\Phi}^{(1)}$ may be constructed directly from the AO integrals according to (denoting AO's by Greek letters)

$$\Phi^{(1)}_{i\alpha} = 2\sum_{\beta\gamma\delta} t_{i\beta\gamma\delta} g^{(1)}_{\alpha\beta\gamma\delta}, \tag{73}$$

$$\Phi^{(1)}_{a\alpha} = 2\sum_{\beta\gamma\delta} t_{\beta a\gamma\delta} g^{(1)}_{\alpha\beta\gamma\delta}, \tag{74}$$

which requires storing and sorting amplitudes with three contravariant AO indices and one MO index (occupied or virtual). When $\boldsymbol{\Phi}^{(1)}$ is constructed according to Eqs. (73) and (74) we avoid calculating first-derivative integrals with two general, one virtual, and one occupied index. The construction of $\boldsymbol{\Phi}^{(1)}$ in the AO basis scales as $\min(Mn^5, 12n^5)$ since each integral contributes to no more than 12 Cartesian directions. Using translational and possibly rotational invariance, the operation count may be further reduced to $9n^5$ or $6n^5$. This suggests that for most molecules it is more efficient to construct $\boldsymbol{\Phi}^{(1)}$ in the AO basis.

Finally, two of the terms in Eq. (72) contain MP2 Fock matrices constructed from one-index transformed integrals. These matrices are best calculated in the MO basis since the undifferentiated MO integrals are already available from the calculation of the energy. The one-index transformations scale as Mn^3vo and the construction of the Fock matrices scales as Mnv^2o^2. The trace operations are inexpensive, requiring only n^2 operations each. Note that all expensive steps depend on M rather than M^2. The reason for this simplification is the use of Fock matrices as intermediates. For example, without Fock matrices the calculation of the last two terms in Eq. (72) scales as $M^2v^4o^2$.

4.2. The calculation of Tr $\boldsymbol{\varrho}\boldsymbol{F}[H^{(1)}]$ *and* Tr $\boldsymbol{\varrho}\boldsymbol{F}[J^{(2)}]$

The third contributions to the molecular gradient Eq. (55) and Hessian Eq. (56) are obtained by taking the traces of $\boldsymbol{\varrho}$ multiplied by $\boldsymbol{F}$ calculated from $H^{(1)}$ and $J^{(2)}$, both of which contain one-index transformed integrals. We have elsewhere given explicit expressions for such Fock matrices [20]. In particular, the Fock matrix constructed from the differentiated integrals $\boldsymbol{I}^{(n)}$ (where $\boldsymbol{I}$ denotes one- and two-electron integrals collectively) one-index transformed by $\boldsymbol{A}$ may be calculated as

$$F_{pq}^{\{\boldsymbol{A},\boldsymbol{I}^{(n)}\}} = \{\boldsymbol{A}, \boldsymbol{F}^{(n)}\}_{pq} + 4\sum_{is} \mathscr{P}^{(n)}_{pqis} A_{is}, \tag{75}$$

where $\boldsymbol{F}^{(n)}$ is the Fock matrix calculated from $\boldsymbol{I}^{(n)}$

$$F^{(n)}_{pq} = h^{(n)}_{pq} + 2\sum_{i} \mathscr{P}^{(n)}_{pqii} \tag{76}$$

and

$$\mathscr{P}^{(n)}_{pqrs} = g^{(n)}_{pqrs} - \tfrac{1}{4} g^{(0)}_{prqs} - \tfrac{1}{4} g^{(n)}_{psrq}. \tag{77}$$

We note that $\boldsymbol{F}^{(n)}$ may be calculated in the AO basis. Multiplying Eq. (75) by $\boldsymbol{\varrho}$ and taking the trace we obtain

$$\mathrm{Tr}\, \boldsymbol{\varrho}\boldsymbol{F}^{\{\boldsymbol{A},\boldsymbol{I}^{(n)}\}} = 2\,\mathrm{Tr}\, \boldsymbol{\varrho}\boldsymbol{A}\boldsymbol{F}^{(n)} + 2\,\mathrm{Tr}\, \boldsymbol{D}\boldsymbol{A}\boldsymbol{Q}^{(n)}, \tag{78}$$

where $\boldsymbol{D}$ is the Hartree–Fock density and

$$Q^{(n)}_{pq} = \sum_{rs} \mathscr{P}^{(n)}_{pqrs} \varrho_{rs}. \tag{79}$$

Using Eq. (78) and the expression for the differentiated integrals $\mathscr{g}^{(1)}$ [Eq. (65) and likewise for the one-electron integrals $\mathscr{h}^{(1)}$], it is straightforward to show that the third contribution to the molecular gradient is given by

$$\mathrm{Tr}\, \boldsymbol{\varrho}\boldsymbol{F}[H^{(1)}] = \mathrm{Tr}\, \boldsymbol{\varrho}\boldsymbol{F}^{(1)} - \mathrm{Tr}\, \boldsymbol{S}^{(1)}(\boldsymbol{\varepsilon}^{(0)}\boldsymbol{\varrho} + \boldsymbol{Q}^{(0)}\boldsymbol{D}), \tag{80}$$

where we have written $\boldsymbol{\varepsilon}^{(0)}$ instead of $\boldsymbol{F}^{(0)}$ to remind ourselves that the undifferentiated Fock matrix is the diagonal matrix of orbital energies. The first term in Eq. (80)

$$\mathrm{Tr}\,\boldsymbol{\varrho}\boldsymbol{F}^{(1)} = \sum_{pq} \varrho_{pq} h^{(1)}_{pq} + 2 \sum_{pqi} \varrho_{pq} \mathscr{P}^{(1)}_{pqii} \tag{81}$$

is best calculated in the AO basis, avoiding the transformation of differentiated two-electron integrals. Again, we transform the densities to the AO basis and add these to the Hartree–Fock densities, obtaining this contribution at little cost.

The Hessian contribution is more involved. To work out the final expression we need the Fock matrix constructed from doubly one-index transformed integrals, an expression for which is provided by Eq. (E.17) of [20]. Multiplying this expression by $\boldsymbol{\varrho}$ and taking the trace, we obtain

$$\begin{aligned}\mathrm{Tr}\,\boldsymbol{\varrho}\boldsymbol{F}^{\{\boldsymbol{A},\{\boldsymbol{B},\boldsymbol{I}\}\}} &= 2\,\mathrm{Tr}\,\boldsymbol{\varrho}\boldsymbol{A}(\boldsymbol{B}\boldsymbol{F} + \widetilde{\boldsymbol{B}\boldsymbol{F}}) + 2\,\mathrm{Tr}\,\boldsymbol{D}\boldsymbol{B}(\boldsymbol{A}\boldsymbol{Q} + \widetilde{\boldsymbol{A}\boldsymbol{Q}}) \\ &\quad + 8\,\mathrm{Tr}\,\boldsymbol{\varrho}\boldsymbol{A}\boldsymbol{P}^{[\boldsymbol{B}]} + 8\,\mathrm{Tr}\,\boldsymbol{\varrho}\,\boldsymbol{B}\boldsymbol{P}^{[\boldsymbol{A}]},\end{aligned} \tag{82}$$

where

$$P^{[A]}_{pq} = \sum_{is} \mathscr{P}_{pqis} A_{is}. \tag{83}$$

It is now straightforward to work out the final expression:

$$\begin{aligned}\mathrm{Tr}\,\boldsymbol{\varrho}\boldsymbol{F}[J^{(2)}] &= \mathrm{Tr}\,\boldsymbol{\varrho}\boldsymbol{F}^{(2)} - \mathrm{Tr}(\boldsymbol{S}^{(2)} - \boldsymbol{S}^{(1)}\boldsymbol{S}^{(1)})(\boldsymbol{\varepsilon}^{(0)}\boldsymbol{\varrho} + \boldsymbol{Q}^{(0)}\boldsymbol{D}) \\ &\quad + 2\,\mathrm{Tr}(2\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)})(\boldsymbol{F}^{(1)}\boldsymbol{\varrho} + \boldsymbol{Q}^{(1)}\boldsymbol{D}) + 2\,\mathrm{Tr}\,\boldsymbol{\varrho}\boldsymbol{\kappa}^{(1)}(\boldsymbol{\kappa}^{(1)}\boldsymbol{\varepsilon}^{(0)} - \boldsymbol{\varepsilon}^{(0)}\boldsymbol{\kappa}^{(1)}) \\ &\quad - \mathrm{Tr}\,\boldsymbol{\varrho}(2\boldsymbol{\kappa}^{(1)} - \tfrac{1}{2}\boldsymbol{S}^{(1)})(\boldsymbol{S}^{(1)}\boldsymbol{\varepsilon}^{(0)} + \boldsymbol{\varepsilon}^{(0)}\boldsymbol{S}^{(1)}) \\ &\quad + 2\,\mathrm{Tr}\,\boldsymbol{D}(\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)})(\boldsymbol{\kappa}^{(1)}\boldsymbol{Q}^{(0)} - \boldsymbol{Q}^{(0)}\boldsymbol{\kappa}^{(1)}) \\ &\quad + \tfrac{1}{2}\,\mathrm{Tr}\,\boldsymbol{D}\boldsymbol{S}^{(1)}(\boldsymbol{S}^{(1)}\boldsymbol{Q}^{(0)} + \boldsymbol{Q}^{(0)}\boldsymbol{S}^{(1)}) \\ &\quad + 4\,\mathrm{Tr}\,\boldsymbol{\varrho}(2\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)})\boldsymbol{P}^{[2\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)}]}.\end{aligned} \tag{84}$$

Again the first contribution is calculated in the AO basis as in Eq. (80) for the gradient. The second term is easily calculated, while the third term requires $\boldsymbol{F}^{(1)}$ and $\boldsymbol{Q}^{(1)}$. These may both be calculated in the AO basis. The differentiated Fock matrix $\boldsymbol{F}^{(1)}$ is also needed for example for the Hartree–Fock Hessian and should already be available, and $\boldsymbol{Q}^{(1)}$ may be calculated in exactly the same way. The remaining terms in Eq. (84) are all trivial once the matrix $\boldsymbol{P}^{[2\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)}]}$ has been calculated.

It should be noted that none of the terms in Eq. (84) require more than Mn^4 operations. The calculation of these terms therefore represents a minor part of the MP2 Hessian calculation. Again, it is the use of Fock matrices which makes these simplifications possible.

4.3. *The calculation of* $V^{(1)}_u$

The last contribution to the Hessian Eq. (56) contains the elements $V^{(1)}_u$

$$\begin{aligned}\frac{1 + \delta_{ab}\delta_{ij}}{2} V^{(1)}_{iajb} &= \mathscr{j}^{(1)}_{iajb} - \sum_k (F[J^{(1)}]_{ik} u_{kajb} + F[J^{(1)}]_{jk} u_{iakb}) \\ &\quad + \sum_c (F[J^{(1)}]_{ac} u_{icjb} + F[J^{(1)}]_{bc} u_{iajc}).\end{aligned} \tag{85}$$

The integrals $\mathscr{J}^{(1)}$ are given by

$$\mathscr{J}^{(1)}_{iajb} = g^{(1)}_{iajb} + \{\boldsymbol{\kappa}^{(1)} - \tfrac{1}{2}\boldsymbol{S}^{(1)}, \boldsymbol{g}^{(0)}\}_{iajb}, \tag{86}$$

which requires the transformation to the MO basis of first-derivative integrals with two occupied and two virtual indices, scaling as Mn^4o. The one-index transformations are proportional to Mnv^2o^2. Using Eq. (75) the Fock matrix calculated from $J^{(1)}$ is given by

$$\boldsymbol{F}[J^{(1)}] = \boldsymbol{F}^{(1)} + \{\boldsymbol{\kappa}^{(1)} - \tfrac{1}{2}\boldsymbol{S}^{(1)}, \boldsymbol{\varepsilon}^{(0)}\} + 2\boldsymbol{P}^{[2\boldsymbol{\kappa}^{(1)} - \boldsymbol{S}^{(1)}]}. \tag{87}$$

All components of this matrix are available from the calculation of Tr $\boldsymbol{\varrho}\boldsymbol{F}[J^{(2)}]$ of Eq. (84). The calculation of the last two contributions to Eq. (85) scales as Mv^3o^2. Finally, the elements of $\tilde{V}^{(1)}$ are related to those of $V^{(1)}$ as

$$\tilde{V}^{(1)}_{iajb} = \frac{1 + \delta_{ab}\delta_{ij}}{2}(2V^{(1)}_{iajb} - V^{(1)}_{ibja}). \tag{88}$$

Once $V^{(1)}$ and $\tilde{V}^{(1)}$ are available, the calculation of the last contribution to the Hessian, Eq. (56), is straightforward. To avoid excessive recalculation of integrals all components of $V^{(1)}$ must be kept on disk simultaneously. For large basis sets this term therefore becomes the bottleneck of the Hessian calculation. It is also found to be time-consuming because of the cost of the transformation of the first-derivative integrals.

4.4. *The response equations*

Using Eqs. (73) and (74) the zeroth-order response equations for the Lagrange multipliers, Eq. (61), may be written

$$\sum_{bj} A_{aibj}\zeta^{(0)}_{jb} = -2(\Phi^{(0)}_{ai} - \Phi^{(0)}_{ia}) - \frac{1}{2}\sum_{pq} \mathscr{D}_{pq}A_{pqai}, \tag{89}$$

where we have extended the definition of $\boldsymbol{A}$ to have the first two indices general

$$A_{pqbj} = \sum_{\sigma} \langle HF|[a^{+}_{q\sigma}, [a_{p\sigma}, [E^{-}_{bj}, H^{(0)}]]]_{+}|HF\rangle. \tag{90}$$

It is easy to verify that when the first two indices of $\boldsymbol{A}$ are virtual and occupied, respectively, the above definition reduces to the one given previously in Eq. (60). The first-order orbital-response equations, Eq. (62), become

$$\sum_{bj} A_{aibj}\kappa^{(1)}_{jb} = 2F[H^{(1)}]_{ai}, \tag{91}$$

where the right-hand side is calculated as [use Eq. (75)]

$$\boldsymbol{F}[H^{(1)}] = \boldsymbol{F}^{(1)} - \tfrac{1}{2}\{\boldsymbol{S}^{(1)}, \boldsymbol{F}^{(0)}\} - 2\boldsymbol{P}^{\boldsymbol{S}^{(1)}}, \tag{92}$$

which is identical to Eq. (87) except that the contributions from $\boldsymbol{\kappa}^{(1)}$ are absent. Both $\boldsymbol{\zeta}^{(0)}$ and $\boldsymbol{\kappa}^{(1)}$ may be determined using the technique of Pople et al. [1] or any other method based on the conjugate-gradient algorithm. Equations (89) and (91) are essentially those solved by Handy et al. [2, 3] in their implementation.

5. A comparison with numerically unstable procedures based on the canonical condition

The expression for the MP2 molecular gradient obtained by insisting on the canonical condition (indicated by left superscript c) reads [8]

$$ {}^{c}W^{(1)} = \langle HF|H^{(1)}|HF\rangle + 2\langle HF|H^{(1)}|MP1\rangle + 2\sum_{p \geqslant q} {}^{c}\zeta_{pq}^{(0)} F[H^{(1)}]_{pq}. \tag{93}$$

It is clear that the equivalence of the expressions based on the Brillouin condition, Eq. (36), and on the canonical condition, Eq. (35), is established if we can show that

$$ {}^{c}\zeta_{pq}^{(0)} = (1 + \delta_{pq})^{-1} \rho_{pq} \tag{94}$$

for $p \geqslant q$. The parameters ${}^{c}\zeta_{pq}$ are the multipliers of the canonically-based Lagrangian. There is one multiplier for each orbital energy ε_p and one for each orbital rotation κ_{ab}, κ_{ij}, and κ_{ai}. The zeroth-order canonical multipliers are determined by the CPHF equations [8]

$$ \sum_{p \geqslant q} {}^{c}\zeta_{pq}^{(0)} \, {}^{c}A_{pqrs} = 2(\Phi_{rs}^{(0)} - \Phi_{sr}^{(0)}), \tag{95}$$

where ${}^{c}A$ is identical to A above [Eq. (90)] except that all four indices may be either occupied or virtual. The matrix ${}^{c}A$ is not symmetric in the pairs of indices pq and rs except for virtual-occupied rotations. The matrix $\mathbf{\Phi}^{(0)}$ is defined as in Eqs. (68) and (69). The structure of ${}^{c}A$ allows us to obtain explicit expressions for the occupied–occupied and virtual–virtual Lagrange multipliers, and we find:

$$ {}^{c}\zeta_{ii}^{(0)} = -\sum_{ajb} t_{iajb} u_{iajb}, \tag{96}$$

$$ {}^{c}\zeta_{aa}^{(0)} = \sum_{ijb} t_{iajb} u_{iajb}, \tag{97}$$

$$ {}^{c}\zeta_{ij}^{(0)} = \frac{\Phi_{ij}^{(0)} - \Phi_{ji}^{(0)}}{\varepsilon_i - \varepsilon_j}, \tag{98}$$

$$ {}^{c}\zeta_{ab}^{(0)} = \frac{\Phi_{ab}^{(0)} - \Phi_{ba}^{(0)}}{\varepsilon_a - \varepsilon_b}. \tag{99}$$

This reveals the problem with the canonical condition: The CPHF equations are singular for degenerate orbitals. However, if we assume that the orbitals are not degenerate and substitute the expressions for $\mathbf{\Phi}^{(0)}$ given above in Eqs. (68) and (69), we find after some rearrangement that

$$ \Phi_{ij}^{(0)} - \Phi_{ji}^{(0)} = (\varepsilon_i - \varepsilon_j)\mathscr{D}_{ij}, \tag{100}$$

$$ \Phi_{ab}^{(0)} - \Phi_{ba}^{(0)} = (\varepsilon_a - \varepsilon_b)\mathscr{D}_{ab}, \tag{101}$$

and the denominator therefore cancels. This removes the instabilities in the CPHF equations but does not apply when the energies are exactly degenerate since in this case the multipliers are undetermined. One may then invoke a suitable limiting procedure, defining the multipliers by their limiting values as the orbitals become degenerate. In this way we may identify the off-diagonal multipliers ${}^c\zeta_{ij}$ and ${}^c\zeta_{ab}$ with the densities $\mathscr{D}_{ij}$ and $\mathscr{D}_{ab}$ in all cases. Likewise, the diagonal multipliers ${}^c\zeta_{ii}$ and ${}^c\zeta_{aa}$ are easily seen to be equal to $\mathscr{D}_{ii}$ and $\mathscr{D}_{aa}$ divided by two. We have now established Eq. (94) for the occupied–occupied and virtual–virtual blocks, and it remains to demonstrate that the canonical and non-canonical occupied–virtual multipliers are identical. Replacing the occupied–occupied and virtual–virtual multipliers by the densities according to Eq. (94), the CPHF equations, Eq. (95), become

$$\sum_{jb} A_{aibj}\,{}^c\zeta^{(0)}_{jb} + \frac{1}{2}\sum_{kl} \mathscr{D}_{kl}A_{klai} + \frac{1}{2}\sum_{cd} \mathscr{D}_{cd}A_{cdai} = -2(\Phi^{(0)}_{ai} - \Phi^{(0)}_{ia}). \qquad (102)$$

This expression is identical to the response equations, Eq. (89), which determine $\zeta^{(0)}_{ia}$. We have now verified that the expressions for the gradient based on the canonical and the Brillouin conditions are equivalent.

We have seen that for the gradient the singularities may be removed by a simple rearrangement of the equations which determine the zeroth-order Lagrangian. For the Hessian the elimination of singularities becomes much more difficult since they cannot be removed directly from the first-order orbital responses $\boldsymbol{\kappa}^{(1)}$. The elimination can first be carried out after the orbital responses have been inserted into the expression for the Hessian, requiring a tedious rearrangement of each term in the Hessian.

We note that the most expensive terms introduced by the non-canonical basis are the one-index transformations in Eq. (85). These transformations represent a rather small fraction of the total calculation. No significant computational advantages are therefore gained by using the canonical representation.

6. Conclusion

We have derived expressions for the molecular gradient and Hessian of the second-order Møller–Plesset energy using the variational Lagrangian technique. By using a non-canonical basis, only non-redundant parameters (orbital rotations and their multipliers) enter the calculation. For the gradient the zeroth-order equations for the Lagrange multipliers must be solved, and for the Hessian we also need the first-order orbital responses. To avoid transforming to the MO basis first-derivative integrals for gradients and second-derivative integrals for Hessians, a set of effective densities (the same set for gradients and Hessians), is constructed. By introducing Fock-like matrices we have separated the time-consuming steps into terms which are linear in the distortions, reducing the cost of the most expensive step to Mn^4o. The derived expressions are structurally as simple as those derived in the unstable canonical representation and computationally not significantly more demanding, requiring additional terms scaling as Mv^3o^2.

We believe that this formulation for the Hessian of the second-order Møller–Plesset energy is the most elegant available, being presented in terms of Fock-like derivative matrices. The scaling cost of the calculation is clearly Mn^4o. As stated earlier, Handy et al. [3] have programmed their first formulation for the Hessian in the non-canonical representation. It is available as part of the CADPAC quantum chemistry code [22]. The equivalent of the terms in Sects. 4.1 and 4.2 above are at present programmed neither as suggested in [3], where the cost scaled as $M^2v^4o^2$, nor as suggested in this paper in terms of Fock-like matrices, but they have all been written *ad hoc* to scale no more than Mn^4o.

Handy and coworkers, in their many applications with their code, have demonstrated that this MP2 Hessian availability is an extremely important tool for theoretical spectroscopy. It is therefore important that the relatively complicated formalism be presented as elegantly and efficiently as possible, and this has been the purpose of this paper.

Acknowledgments. We would like to thank Henrik Koch for discussions. This work was supported by the Norwegian Research Council for Science and the Humanities (Grant No. 432.87/006) and the Danish Natural Science Research Council (Grant No. 11-6844).

References

1. Pople JA, Krishnan R, Schlegel HB, Binkley JS (1979) Int J Quantum Chem Symp 13:225
2. Handy NC, Amos RD, Gaw JF, Rice JE, Simandiras ED, Lee TJ, Harrison RJ, Laidig WD, Fitzgerald GB, Bartlett RJ (1986) In: Jørgensen P, Simons J. Geometrical derivatives of energy surfaces and molecular properties. Reidel, Dordrecht, p 179
3. Handy NC, Amos RD, Gaw JF, Rice JE, Simandiras ED (1985) Chem Phys Lett 120:151
4. Harrison RJ, Fitzgerald GB, Laidig WD, Bartlett RJ (1986) Chem Phys Lett 124:291
5. Simandiras ED, Handy NC, Amos RD (1987) Chem Phys Lett 133:324
6. Jørgensen P, Simons J (1983) J Chem Phys 79:334
7. Handy NC, Schaefer III HF, (1984) J Chem Phys 81:5031
8. Jørgensen P, Helgaker T (1988) J Chem Phys 89:1560
9. Jørgensen P, Swanstrøm P, Yeager DL, Olsen J (1983) Int J Quantum Chem 23:959
10. Pulay P (1983) J Chem Phys 78:5043
11. Helgaker T, Jørgensen P (1989) Theor Chim Acta 75:111
12. Pulay P, Sæbø S (1986) Theor Chim Acta 69:357
13. Rice JE, Amos RD (1985) Chem Phys Lett 122:585
14. Linderberg J, Öhrn Y (1973) Propagators in quantum chemistry. Academic Press, London
15. Jørgensen P, Simons J (1981) Second quantization-based methods in quantum chemistry, Academic Press, New York
16. Pulay P, Sæbø S, Meyer W (1984) J Chem Phys 81:1901
17. Helgaker TU, Almlöf J (1984) Int J Quantum Chem 26:275
18. Simons J, Jørgensen P, Helgaker TU (1984) Chem Phys 86:413
19. Gerratt J, Mills IM (1968) J Chem Phys 49:1719
20. Helgaker T, Jørgensen P (1988) Adv Quantum Chem 19:183
21. Siegbahn PEM, Almlöf J, Heiberg A, Roos BO (1981) J Chem Phys 74:2384
22. Amos RD, Rice JE (1988) The Cambridge analytic derivatives package issue 4.0. Cambridge