

BASIS SET CONSIDERATIONS FOR THE CALCULATION OF GRADIENTS IN THE LCAO FORMALISM

Jan ALMLÖF and Trygve HELGAKER

Department of Chemistry, University of Oslo, Blindern, Oslo 3, Norway

Received 27 April 1981; in final form 16 June 1981

Compact expressions are derived for the analytical gradient with a general MC SCF LCAO wavefunction. Cancellations among the terms involved are shown to have important computational aspects, and substantial reductions of computer time may be achieved with a properly chosen basis set.

1. Introduction

During the last few years analytically calculated derivatives of total energy expressions have become a valuable tool for the study of molecular energy surfaces [1–4]. However, although the gradient method is vastly superior to the use of pointwise calculated energies, computational economy still sets severe limitations on the size of systems that can be conveniently handled. It therefore appears important to continue the analysis of the expressions for these quantities in order to find optimal computational procedures.

In two recent papers Nakatsuji et al. [5,6] investigated how calculation of the analytical gradient for general SCF functions may be avoided by systematically enlarging the basis set with derivatives of the orbitals so that the Hellmann–Feynman theorem [7,8] is approximately satisfied. This seems to be a reasonable procedure for small molecules. But for most systems of chemical interest – for which the calculation of the gradient is of special value – such a basis set becomes prohibitively large. Nevertheless, by exploiting the same relationships it appears possible to achieve significant savings in computer time when calculating the analytical gradient using an ordinary basis set.

The present paper provides a derivation of a compact expression for the MC SCF LCAO energy gradient, which is a convenient reformulation of those given previously [1–6]. From this expression we derive

upper bounds for the contribution to the gradient from individual atomic orbitals, and suggest a simple estimate for the relative importance of these orbitals when calculating the gradient. It is finally shown how the choice of basis functions may lead to cancellations of terms in the gradient expression that are normally computed and thus to savings in computer time.

2. The gradient of an MC SCF function

For an MC SCF function we have, following the notation of Hinze [9]:

$$\sum_j F_{ij} |\varphi_j\rangle = \sum_j \epsilon_{ji} |\varphi_j\rangle, \quad (1)$$

where the summation is over orthonormal molecular orbitals and

$$F_{ij} = h\gamma_{ij} + \sum_{kl} V_{kl} \Gamma_{ij,kl}, \quad (2a)$$

$$V_{kl}(1) = \int \varphi_k^*(2) g(1,2) \varphi_l(2) d(2), \quad (2b)$$

$$\epsilon_{ij} = \langle \varphi_i | \sum_k F_{jk} | \varphi_k \rangle. \quad (2c)$$

In these expressions, h and $g(1,2)$ are the one- and two-electron parts of the hamiltonian, γ_{ij} and $\Gamma_{ij,kl}$ are the first- and second-order density matrix elements. The total energy is given by

$$E = \sum_i \epsilon_{ii} - \frac{1}{2} \sum_{ijkl} \langle \varphi_i | V_{kl} | \varphi_j \rangle \Gamma_{ij,kl} \quad (3)$$

For a function optimized on a truncated basis, the molecular orbitals do not satisfy (1), but instead

$$\sum_{ij} \langle \varphi_i | F_{ij} - \epsilon_{ji} | \varphi_j \rangle = 0. \quad (4)$$

Differentiation of this equation with respect to a parameter contained in the molecular hamiltonian (indicated by a prime) leads immediately to

$$\sum_i \epsilon'_{ii} = \sum_{ij} \langle \varphi_i | F'_{ij} | \varphi_j \rangle + 2 \text{Re} \sum_{ij} \langle \varphi'_i | F_{ij} - \epsilon_{ji} | \varphi_j \rangle. \quad (5)$$

The first, Hellmann–Feynman-like term may be written

$$\begin{aligned} \sum_{ij} \langle \varphi_i | F'_{ij} | \varphi_j \rangle &= \sum_{ij} \langle \varphi_i | h' | \varphi_j \rangle \gamma_{ij} \\ &+ \sum_{ijkl} \langle \varphi_i | V'_{kl} | \varphi_j \rangle \Gamma_{ij,kl} \\ &+ \frac{1}{2} \sum_{ijkl} \langle \varphi_i | V_{kl} | \varphi_j \rangle \Gamma'_{ij,kl} + A, \end{aligned} \quad (6)$$

where A is a sum of terms containing the derivatives of the density matrix elements as factors. As these parameters have been optimized, their contribution to the gradient will vanish, i.e. $A \equiv 0$ [1,10]. Substituting (6) in (5) and introducing (3) we obtain

$$\begin{aligned} E' &= \sum_{ij} \langle \varphi_i | h' | \varphi_j \rangle \gamma_{ij} + 2 \text{Re} \sum_{ij} \langle \varphi'_i | F_{ij} - \epsilon_{ji} | \varphi_j \rangle \\ &= -f^{\text{HF}} - f^{\text{WF}}, \end{aligned} \quad (7)$$

where we recognize the Hellmann–Feynman force $f^{\text{HF}} = -\sum_{ij} \langle \varphi_i | h' | \varphi_j \rangle \gamma_{ij} = -(H')$ and the wavefunction force f^{WF} , the latter being the sum of the integral and density forces in the terminology of Pulay [1]. The expression for the wavefunction force may be slightly rewritten by noticing that

$$\sum_j \langle F_{ij} - \epsilon_{ji} | \varphi_j \rangle \equiv Q \sum_j F_{ij} | \varphi_j \rangle, \quad (8)$$

where we have used the relation (2c) and introduced the projection operator

$$Q = 1 - \sum_i | \varphi_i \rangle \langle \varphi_i |. \quad (9)$$

Using the idempotency of the projector, we obtain

$$f^{\text{WF}} = -2 \text{Re} \sum_{ij} \langle Q \varphi'_i | Q F_{ij} \varphi_j \rangle. \quad (10)$$

This compact expression is seen to be equivalent to those obtained previously [1–6]. However, we immediately note that $Q | \varphi'_i \rangle$ lies entirely outside the space spanned by the basis set orbitals. Inclusion of the derivative of one of the orbitals in the basis set will thus eliminate its contribution to the gradient [5,6]. If *all* (first, second, ...) derivatives of all orbitals are included the wavefunction force will vanish identically; this is actually an obvious special case of the generalized Hellmann–Feynman theorem due to Hurley [10,11]: $dE/dX = \langle dH/dX \rangle$ if the trial functions are invariant to changes in the parameter X .

The wavefunction force may consequently be calculated in two equivalent ways: The normal procedure is to calculate the full integral force after which the density force is subtracted according to (7). Alternatively one may calculate projected derivatives of the orbitals and then the “projected” integral force according to (10). But, since the gradient is usually calculated as a sum over atomic orbitals, little is gained from the latter procedure unless one or more of the differentiated orbitals is completely removed when the projector is applied. In certain cases, therefore, a partial projection – as described below – appears advantageous.

3. Upper bounds to the contribution of the individual atomic orbitals to the MC SCF gradient

Using eq. (10) it is easy to set up upper bounds to the contribution of the individual atomic orbitals to the MC SCF gradient. The molecular orbitals are usually expanded over atomic orbitals

$$| \varphi_i \rangle = \sum_p C_{ip} | \chi_p \rangle, \quad (11)$$

which upon differentiation and projection gives

$$Q | \varphi'_i \rangle = Q \sum_p C'_{ip} | \chi'_p \rangle, \quad (12)$$

since $Q \sum_p C'_{ip} | \chi_p \rangle = 0$. The contribution of the differentiated orbital $| \chi'_q \rangle$ is then given by

$$f_a^{\text{WF}} = -2\text{Re}\langle Q\chi'_a | Q \sum_{ij} C_{ia}^* F_{ij} \phi_j \rangle. \quad (13)$$

Applying the Cauchy–Schwarz inequality we obtain upper bounds to the individual atomic contributions:

$$|f_a^{\text{WF}}| \leq 2 \|Q\chi'_a\| \|Q \sum_{ij} C_{ia}^* F_{ij} \phi_j\|. \quad (14)$$

The first factor on the right-hand side may be seen as a measure of the inadequacy of the basis set to describe the differentiated basis function. In a similar fashion the second factor is a weighted measure of the inability of the basis set to describe the exact molecular orbitals satisfying eq. (1). This last factor is difficult to estimate because of the intervening Fock matrix, the first is however easily calculated. Using once more the idempotency of the projector we obtain

$$\begin{aligned} \|Q\chi'_a\|^2 &= \langle \chi'_a | Q | \chi'_a \rangle \\ &= \langle \chi'_a | \chi'_a \rangle - \sum_i \langle \chi'_a | \phi_i \rangle \langle \phi_i | \chi'_a \rangle. \end{aligned} \quad (15)$$

These quantities may be computed for all orbitals prior to the calculation of the gradient and will provide a rough guide to the relative importance of the differentiated basis functions (the magnitude of the expansion coefficients may also be taken into consideration).

4. Applications

Some applications of these results will now be given. In routine MC SCF calculations the only unstable (i.e. neither constant nor optimized) parameters are the basis function centers, which are usually fixed on the atomic nuclei. Thus only partial derivatives with respect to these parameters will contribute to the gradient. The derivative of a primitive gaussian function

$$\begin{aligned} G(\alpha, A, l, m, n) &= (x - A_x)^l (y - A_y)^m (z - A_z)^n \\ &\times \exp[-\alpha(r - A)^2], \end{aligned} \quad (16)$$

with respect to the center of expansion A is a linear combination of two new gaussians with the original exponents retained, e.g.

$$\begin{aligned} dG(\alpha, A, l, m, n)/dA_x \\ = 2\alpha G(\alpha, A, l + 1, m, n) - lG(\alpha, A, l - 1, m, n). \end{aligned} \quad (17)$$

If the basis set is chosen such that the orbital exponents for s, p, d, ... functions are equal, the derivatives of this set will contain several terms lying entirely within the original set. A partial projection is then easily accomplished by simply omitting these terms when calculating the integral and density forces. This may lead to considerable savings in computer time; for a first-row atom with a minimal or DZ basis and uncontracted valence orbitals the gain is $\approx 30\%$ and increases to $\approx 50\%$ when polarization functions are added. With a 11s9p5d1f set, appropriate for near Hartree–Fock calculations on first-row transition metals, the gain exceeds 50%.

It should be noted that basis sets of the type mentioned above are already widely used in routine ab initio calculations for other reasons [12], and that the systematic construction of basis sets by repeated differentiation of s-orbitals was proposed several years ago as they give fast evaluation of one- and two-electron integrals [13,14]. With such basis sets the above procedure gives significant savings in computer time with no loss of accuracy.

The scheme discussed above obviously becomes less efficient if the basis functions are contracted. Since contraction is of less importance for valence shell orbitals – where the most important savings may be achieved – it appears reasonable to leave the valence orbitals uncontracted in gradient calculations where the above procedure is employed.

In the general case the norms of the projected differentiated orbitals (15) will not be zero. In most cases we may nevertheless expect some of them to be small, for example when the orbital exponents are not equal but nearly so or when the valence orbitals with equal exponents have been contracted. If high accuracy is not required, these contributions to the gradient may be neglected. This will be the case in the early stages of geometry optimizations as the knowledge of the force constant matrix is usually very uncertain and makes an accurate calculation of the gradient unnecessary.

References

- [1] P. Pulay, in: *Modern theoretical chemistry*, Vol. 4, ed. H.F. Schaefer III (Plenum Press, New York, 1977).
- [2] S. Kato and K. Morokuma, *Chem. Phys. Letters* 65 (1979) 19.

- [3] J.D. Goddard, N.C. Handy and H.F. Schaefer III, *J. Chem. Phys.* 71 (1979) 1525.
- [4] J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, *Intern. J. Quantum Chem.* S13 (1979) 225.
- [5] H. Nakatsuji, K. Kanda and T. Yonezawa, *Chem. Phys. Letters* 75 (1980) 340.
- [6] H. Nakatsuji, T. Hayakawa and M. Hada, *Chem. Phys. Letters* 80 (1981) 94.
- [7] J. Hellmann, *Einführung in die Quantenchemie* (Deuticke, Leipzig, 1937).
- [8] R.P. Feynman, *Phys. Rev.* 56 (1939) 340.
- [9] J. Hinze, *J. Chem. Phys.* 59 (1973) 6424.
- [10] S.T. Epstein, *Theoret. Chim. Acta* 55 (1980) 251.
- [11] A.C. Hurley, *Proc. Roy. Soc. A* 226 (1954) 170.
- [12] W.J. Hehre, R.F. Stewart and J.A. Pople, *J. Chem. Phys.* 51 (1969) 2657.
- [13] A. Gołębiewski and J. Mrozek, *Intern. J. Quantum Chem.* 7 (1973) 623.
- [14] A. Gołębiewski and J. Mrozek, *Intern. J. Quantum Chem.* 7 (1973) 1021.