Analytical gradient of the linear combination of Gaussian-type orbitals—local spin density energy

R. Fournier, J. Andzelm,[a] and D.R. Salahub
Département de Chimie, Université de Montréal, C.P. 6128, Succursale A, Montréal, Québec, H3C 3J7, Canada

(Received 18 November 1988; accepted 7 February 1989)

An expression has been derived for the analytical evaluation of the energy gradient within the linear combination of Gaussian-type orbitals—local spin density method. This expression is valid for any exchange-correlation energy functional which can be represented in a density gradient expansion. In practice, because the exchange-correlation terms are fitted with auxiliary functions, one has to introduce an approximation. Results are reported of tests on diatomics that show that it is possible to attain a typical accuracy of ± 0.01 a.u. on equilibrium distances, relative to the energy minimum. The formulas for molecular integral derivatives that we implemented are based on the highly efficient recurrence formulas of Obara and Saika. We report here an additional formula for angular momentum transfer which is very useful for efficient programming of the gradient. In all cases studied, the time required to compute the gradient is a fraction of the time spent to solve the self-consistent-field Kohn–Sham equations.

I. INTRODUCTION

The last 15 years has seen the emergence of many efficient methods for locating extrema of potential-energy surfaces (PES) in *ab initio* quantum chemistry.[1] These methods are based on the evaluation of the analytical gradient of the energy with respect to nuclear coordinates.[2,3] Several groups have developed working equations and/or computer programs for gradient evaluation within the framework of various *ab initio* quantum chemical methods: Hartree–Fock,[4,5] Møller–Plesset perturbation theory,[6] configuration interaction,[7–10] multiconfiguration SCF[11–13] and coupled cluster.[14,15] These methods have been of invaluable importance in the systematic search of equilibrium geometries[16] and reaction pathways,[17] and computation of vibrational frequencies[18] for molecules containing "light" atoms (roughly H to Cl). The wealth of applications of gradient techniques has been largely documented in review articles.[1–3,19]

Similar studies for "heavy" atom containing systems with *ab initio* methods are much more difficult (especially if the heavy atoms have open *d* shells). For transition metal atoms, Hartree–Fock theory is typically a bad first approximation.[20] The standard way to correct Hartree–Fock for its neglect of correlation is to expand the wave function in a combination of many determinantal functions. Such an expansion is known to have a slow convergence on the number of determinants and the cost in computer time grows rapidly with the number of electrons. This has limited the successful application of multideterminantal methods to fairly small systems when transition metal atoms are involved.[21]

Modern density functional theory (DFT) which has its roots in the work of Hohenberg, Kohn, and Sham[22] offers a practical alternative for such heavy atom systems. Most calculations have made use of the local spin density (LSD) approximation.[22,23] Accumulated experience shows that within this framework, DFT gives consistently good equilibrium geometries, even when open *d*-shell atoms are involved.[24,25(f)] In particular, the LCGTO–LSD method has provided good results for complex systems where *ab initio* methods would be difficult or practically impossible to apply.[26]

Development of analytical gradient techniques in DFT is in a rather early stage when contrasted with *ab initio* methods. Nevertheless, most interesting and encouraging progress has already been made in this area.[25] In what follows, we will give an account of the implementation of an analytical gradient technique for the linear combination of Gaussian-type orbitals–local spin density (LCGTO–LSD) method originally suggested by Sambe and Felton[27] and further developed by Dunlap *et al.*[28] and by others.

The rest of this paper is divided into four sections. In Sec. II, we give a short overview of the LCGTO–LSD method and a derivation of the equations for gradient evaluation.[29,30] Special care is taken to underline the approximations involved and the domain of applicability of these equations. In Sec. III we give a brief description of the strategy adopted to program the equations of Sec. II. In Sec. IV the results of tests for diatomics are discussed as are the implications for future applications of the method.

II. DERIVATION OF THE EQUATIONS

The total energy in the LCGTO–LSD method is given by the following expression:

$$E=\sum_{\mu\nu}P_{\mu\nu}\left[h_{\mu\nu}+\sum_{k}C_{k}\left[\mu\nu|r_{12}^{-1}|k\right]\right]-\frac{1}{2}\sum_{k}\sum_{k'}C_{k}C'_{k}\left[k|r_{12}^{-1}|k'\right]+\tilde{E}_{\mathrm{xc}}+V_{nn}\,, \quad (1)$$

where μ,ν denote atomic orbital basis functions, $P_{\mu\nu}$ denotes

[a] Present address: Cray Research, 1333 Northland Drive, Mendota Heights, MN 55120.

 0021-9606/89/116371-07$02.10

the associated density matrix element, and $h_{\mu\nu}$ is the core Hamiltonian matrix element including kinetic energy and external potential (nuclear attraction plus, possibly, an external field). Density fitting basis functions are denoted by k and k' and the C_k's (C'_k) are the associated density fit coefficients:

$$V_{nn} = \sum_{A>B} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \tag{2}$$

is the nuclear energy and

$$\tilde{E}_{xc} = \int \rho(\mathbf{r}) \tilde{\epsilon}_{xc}(\mathbf{r}) d\mathbf{r} \tag{3}$$

is the exchange-correlation (xc) energy. The tilde indicates a fitted function [e.g., $\tilde{\epsilon}_{xc}(r)$] or a quantity which depends directly on such a function (e.g., $\tilde{E}_{xc}$). The expression (3) embodies a large class of xc energy functionals, the simplest of which have an ϵ_{xc} of the form:

$$\epsilon_{xc}(\mathbf{r}) = \epsilon_{xc}(\rho^\alpha(\mathbf{r}), \rho^\beta(\mathbf{r})) . \tag{4}$$

Equation (4) means that ϵ_{xc} at a point $\mathbf{r}$ depends only on the spin up, $\rho^\alpha(\mathbf{r})$, and spin down, $\rho^\beta(\mathbf{r})$, electron densities at that point. Equations (3) and (4) define the local spin density approximation. We will also consider the case where ϵ_{xc} also depends on rotationally invariant combinations of derivatives of various orders of these spin-up and spin-down densities, that is, ϵ_{xc}'s of the form[3]

$$\epsilon_{xc}(\mathbf{r}) = \epsilon_{xc}(\rho^\alpha(\mathbf{r}), \rho^\beta(\mathbf{r}), |\nabla \rho^\alpha(\mathbf{r})|, |\nabla \rho^\beta(\mathbf{r})|, \ldots) . \tag{5}$$

The fit to the exact $\epsilon_{xc}(\mathbf{r})$ corresponding to a given density, $\tilde{\epsilon}_{xc}(\mathbf{r})$, is written as

$$\tilde{\epsilon}_{xc}(\mathbf{r}) = \sum_l C^e_l l(\mathbf{r}) . \tag{6}$$

The reader is referred to the paper by Dunlap *et al.*[28] for the description of how the fits are actually made. From here on, the equations will be written for the non-spin-polarized case, for simplicity. A generalization to the spin-polarized case is straightforward. The electron density is given in terms of the spin orbitals $\phi_i(\mathbf{r})$ by

$$\rho(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r})|^2 , \tag{7}$$

where the n_i's are occupation numbers, or alternatively, in terms of the atom-centered basis functions

$$\rho(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu} \mu^*(\mathbf{r}) \nu(\mathbf{r}) \tag{8}$$

with

$$P_{\mu\nu} = \sum_i n_i C^*_{\mu i} C_{\nu i} . \tag{9}$$

In the LCGTO–LSD method, the coefficients for expansion of the spin orbitals satisfy a set of equations analogous to the Hartree equations:

$$\sum_\nu F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_\nu S_{\mu\nu} C_{\nu i} , \tag{10}$$

where $S_{\mu\nu}$ is the usual overlap integral and $F_{\mu\nu}$ is the matrix element of our effective one electron Hamiltonian given by

$$F_{\mu\nu} \equiv \langle \mu | \hat{F} | \nu \rangle = h_{\mu\nu} + [\mu\nu | r_{12}^{-1} | \tilde{\rho}] + \langle \mu | \tilde{v}_{xc} | \nu \rangle \tag{11}$$

with $\tilde{\rho}$ and $\tilde{v}_{xc}$ given by

$$\tilde{\rho}(\mathbf{r}) = \sum_k c_k k(\mathbf{r}) , \tag{12}$$

$$\tilde{v}_{xc}(\mathbf{r}) = \sum_l C^v_l l(\mathbf{r}) . \tag{13}$$

Simply taking the derivative of Eq. (1) with respect to a parameter λ (later a nuclear coordinate) would yield an expression involving derivatives of two and three index integrals over Cartesian Gaussian functions and derivatives of the coefficients $C_{\mu i}$, C_k, and C^e_l. Evaluation of derivatives of integrals can be handled with standard techniques with minor modifications and this subject will be deferred until Sec. III. Our concern here is thus to evaluate terms involving derivatives of the various coefficients. First, we follow Pulay[4,11] and using Eqs. (10) and (11) together with the normalization constraint

$$\sum_{\mu\nu} C^*_{\mu i} S_{\mu\nu} C_{\nu i} = 1 \quad \forall i , \tag{14}$$

we obtain, after differentiation, the following:

$$\sum_{\mu\nu} P^{(\lambda)}_{\mu\nu} F_{\mu\nu} = -\sum_{\mu\nu} W_{\mu\nu} S^{(\lambda)}_{\mu\nu} . \tag{15}$$

The superscript (λ) indicates differentiation with respect to λ. The right-hand side of Eq. (15) is usually called the density force but we prefer to call it the energy-weighted overlap force (EOF). $W_{\mu\nu}$ is the so-called energy weighted density matrix:

$$W_{\mu\nu} = \sum_i n_i \epsilon_i C^*_{\mu i} C_{\nu i} . \tag{16}$$

Equations (15) and (16) take care of the derivatives of the $C_{\mu i}$'s in much the same way as in standard *ab initio* methods. Thanks to the clever choice of fitting procedure of $\rho(\mathbf{r})$ described in Ref. 28, when we make use of the fitting equation

$$\frac{\partial}{\partial C_k}\left([(\rho - \tilde{\rho}) | r_{12}^{-1} | (\rho - \tilde{\rho})] - \gamma \sum_{k'} C'_k \int k'(\mathbf{r}) d\mathbf{r} \right) = 0 , \tag{17}$$

where γ is a Lagrange multiplier, and the normalization condition:

$$\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N , \tag{18}$$

we obtain, after some algebra

$$\sum_k C^{(\lambda)}_k \left(\sum_{\mu\nu} P_{\mu\nu} [\mu\nu | r_{12}^{-1} | k] - \sum_{k'} C'_k [k | r_{12}^{-1} | k'] \right) = 0 . \tag{19}$$

Equation (19) shows exact cancellation of the only two terms involving $C^{(\lambda)}_k$. An easy way to see why all terms involving $C^{(\lambda)}_k$ add to zero is to realize that the sum of these terms can be written as

$$\sum_k \frac{\partial E}{\partial C_k} C^{(\lambda)}_k .$$

But the fitting procedure of Ref. 28 was precisely chosen because it minimizes the error on the Coulomb energy and

satisfies $\partial E/\partial C_k = 0$. Moreover, the derivative of the constraint (18) introduces no new term. Substitution of Eqs. (15) and (19) in the expression we would obtain by direct differentiation of Eq. (1) gives an intermediate expression in which all terms depend only on derivatives of molecular integrals except the following:

$$A \equiv \tilde{E}_{\rm xc}^{(\lambda)} - \sum_{\mu\nu} P_{\mu\nu}^{(\lambda)} \langle \mu | \tilde{v}_{\rm xc} | \nu \rangle \,. \tag{20}$$

In order to progress, we now have to find some useful relation to rewrite these two terms. We first assume that the functional $E_{\rm xc}$ is of the form given by Eqs. (3) and (5). This includes of course all functionals encountered in the LSD approximation. For such a functional, it can be shown that[32]

$$\frac{\partial E_{\rm xc}}{\partial \lambda} = \int \frac{\delta E_{\rm xc}[\rho]}{\delta \rho} \frac{\partial \rho}{\partial \lambda} d\mathbf{r} \,. \tag{21}$$

Moreover, in density functional theory the functional derivative of the xc energy is, by definition, the xc potential $v_{\rm xc}(\mathbf{r})$, so that Eq. (21) becomes

$$\frac{\partial E_{\rm xc}}{\partial \lambda} = \int \frac{\partial \rho(\mathbf{r})}{\partial \lambda} v_{\rm xc}(\mathbf{r}) d\mathbf{r} \,. \tag{22}$$

This last equation deserves three remarks.

(1) It is valid for any functional $E_{\rm xc}[\rho]$ of the form (3) + (5), not only local functionals of the form (3) + (4).

(2) Although the exact (unknown) functional $E_{\rm xc}[\rho]$ may not be of the form (3) + (5), this is of little practical importance for us since, in our methodology, the functionals that we use must be of this form.

(3) Given a functional, in actual calculations, we do not know the precise $E_{\rm xc}$ and $v_{\rm xc}$ corresponding to a given density. Instead, we obtain estimates of $E_{\rm xc}$ and $v_{\rm xc}$ through a fitting procedure and Eq. (22) then becomes approximate. If we now use the approximate form of Eq. (22),

$$\frac{\partial \tilde{E}_{\rm xc}}{\partial \lambda} \simeq \int \frac{\partial \rho(\mathbf{r})}{\partial \lambda} \tilde{v}_{\rm xc}(\mathbf{r}) d\mathbf{r} \,, \tag{23}$$

with Eqs. (8), (15), and (19), we obtain

$$\begin{aligned} E^{(\lambda)} = & \sum_{\mu\nu} P_{\mu\nu} \left(h_{\mu\nu}^{(\lambda)} + \sum_k C_k [\mu\nu | r_{12}^{-1} | k]^{(\lambda)} \right) \\ & - \frac{1}{2} \sum_{k,k'} C_k C'_{k'} [k | r_{12}^{-1} | k']^{(\lambda)} \\ & - \sum_{\mu\nu} W_{\mu\nu} S_{\mu\nu}^{(\lambda)} + V_{nn}^{(\lambda)} \\ & + \sum_{\mu\nu} P_{\mu\nu} [\langle \mu^{(\lambda)} | \tilde{v}_{\rm xc} | \nu \rangle + \langle \mu | \tilde{x}_{\rm xc} | \nu^{(\lambda)} \rangle] \,. \end{aligned} \tag{24}$$

This equation is amenable to efficient computation as it involves no derivative of coefficients. It is useful to regroup some terms and rewrite Eq. (24) as

$$\begin{aligned} E^{(\lambda)} = & \left(\sum_{\mu\nu} P_{\mu\nu} \langle \mu | h^{(\lambda)} | \nu \rangle + V_{nn}^{(\lambda)} \right) \\ & + \left(\sum_i \sum_{\mu\nu} n_i C_{\mu i}^* C_{\nu i} \langle \mu^{(\lambda)} | (\hat{F} - \epsilon_i) | \nu \rangle \right. \\ & \left. + \text{complex conjugate} \right) \\ & + \left(\sum_k C_k [k^{(\lambda)} | r_{12}^{-1} | (\rho - \tilde{\rho})] \right) \\ = & - (F^{\rm HF} + F^{\rm OBC} + F^{\rm DBC}) \,. \end{aligned} \tag{25}$$

In Eq. (25), $F^{\rm HF}$ is the Hellmann–Feynman force which would be the exact force if the basis sets were complete. $F^{\rm OBC}$ and $F^{\rm DBC}$ stands for "orbital basis correction" force and "density fit basis correction" force. They can be viewed as "artificial forces" correcting for the fact that these bases are incomplete. The corresponding xc basis correction force is absent as a result of approximation (23). That is, in practice, the gradient calculated with Eq. (24) will only be accurate if the xc basis is "good." The quality of commonly used bases is examined below. Instead of using this approximation, it may be interesting to write a formally exact expression. For this purpose, we define the errors on fits, $\Delta\epsilon_{\rm xc}$ and $\Delta v_{\rm xc}$ through

$$\epsilon_{\rm xc}(\mathbf{r}) = \tilde{\epsilon}_{\rm xc}(\mathbf{r}) + \Delta\epsilon_{\rm xc}(\mathbf{r}) \,, \tag{26}$$

$$v_{\rm xc}(\mathbf{r}) = \tilde{v}_{\rm xc}(\mathbf{r}) + \Delta v_{\rm xc}(\mathbf{r}) \,. \tag{27}$$

One can then derive the equivalent of Eq. (24) with an extra term Δ:

$$\Delta = \int \frac{\partial \rho}{\partial \lambda} (\Delta v_{\rm xc} - \Delta\epsilon_{\rm xc}) d\mathbf{r} - \int \rho \frac{\partial \Delta\epsilon_{\rm xc}}{\partial \lambda} d\mathbf{r} \tag{28}$$

with the LSD approximation, this becomes

$$\Delta = \int \frac{\partial \rho}{\partial \lambda} \left(\Delta v_{\rm xc} - \Delta\epsilon_{\rm xc} - \rho \frac{d\,\Delta\epsilon_{\rm xc}}{d\rho} \right) d\mathbf{r} \,. \tag{29}$$

Approximation (23) amounts to the approximation $\Delta \simeq 0$. A good fit as judged by the accuracy of the gradient expressed as Eq. (24) is one which minimizes Δ as given by Eq. (28) or (29). Equation (29) stresses the importance of having small fitting errors $\Delta v_{\rm xc}(\mathbf{r})$ and $\Delta\epsilon_{\rm xc}(\mathbf{r})$ close to nuclei where $\partial\rho(\mathbf{r})/\partial\lambda$ is large.

Andzelm *et al.* incorporated the use of model core potentials within the LCGTO–LSD method.[33] With model core potentials, the effective one-electron Hamiltonian of Eq. (11) becomes

$$F_{\mu\nu}^{\rm MP} = \langle \mu | h^{\rm MP} | \nu \rangle + [\mu\nu | r_{12}^{-1} | \tilde{\rho}] + \langle \mu | \tilde{v}_{\rm xc} | \nu \rangle - O_{\mu\nu} \,. \tag{30}$$

$h^{\rm MP}$ differs from h in that the nuclear attraction potential is replaced by an effective potential that includes the effect of frozen core orbitals. $O_{\mu\nu}$ is a matrix element of a projection operator that ensures orthogonality of valence orbitals to the core. Specifically,

$$O_{\mu\nu} = \sum_a^{\rm core} 2\epsilon_a \langle \mu | a \rangle \langle a | \nu \rangle \,, \tag{31}$$

where a denotes a core orbital. The new terms introduced in the gradient by the use of model core potentials can be evaluated in just the same way as other terms involving one-electron operators. Because the error associated with approximation (23) may contain a substantial contribution from core electrons, it is expected that the use of model core potentials could improve the accuracy of the computed gradient.

We have thus obtained in Eq. (24) the necessary working equation for evaluation of gradients. We stress that one

approximation was introduced in Eq. (23). If the xc fitting basis and the grid were complete, $\tilde{\epsilon}_{xc}$ and $\tilde{v}_{xc}$ would be exact ϵ_{xc} and v_{xc} and Eq. (23) would be exact. Section III deals with the more practical aspects of the evaluation of expression (24).

III. PROGRAMMING STRATEGY

All the coefficients appearing in Eq. (24), $P_{\mu\nu}$, $W_{\mu\nu}$, C_k, C_l^v are readily available from a self-consistent solution of the Kohn–Sham Eq. (10). The evaluation of Eq. (24) then boils down to computation of derivatives of two- and three-index molecular integrals. In our method, the basis functions of each of the three sets $\{|\mu\rangle\}$, $\{|k\rangle\}$, and $\{|l\rangle\}$ are Cartesian Gaussian functions having the form

$$g(\mathbf{r},\mathbf{R},\mathbf{l},\alpha) = N\left(\prod_{i=x,y,z}(\mathbf{r}_i - \mathbf{R}_i)^{l_i}\right)e^{-\alpha|\mathbf{r}-\mathbf{R}|^2}, \tag{32}$$

where $\mathbf{r}_i$ and $\mathbf{R}_i$ are the ith Cartesian components of the position vector of the electron and center of function, respectively, and l_i is the ith component of the angular momentum vector. N is a normalization factor equal to

$$N = \left(\frac{2\alpha}{\pi}\right)^{3/4}(4\alpha)^L\left(\prod_{i=x,y,z}(2l_i-1)!!\right)^{-1/2} \tag{33}$$

with $L = l_x + l_y + l_z$. The Cartesian Gaussians have the property that their derivatives are a combination of two Cartesian Gaussians with angular momentum $L+1$ and $L-1$:

$$\frac{\partial}{\partial R_i}g(\mathbf{r};\mathbf{R},\mathbf{l},\alpha) = 2\alpha g(\mathbf{r};\mathbf{R},\mathbf{l}+\mathbf{e}_i,\alpha) - l_i g(\mathbf{r};\mathbf{R},\mathbf{l}-\mathbf{e}_i,\alpha)\,, \tag{34}$$

where $\mathbf{e}_i$ is a unit vector. From Eq. (34), we see that computation of derivatives of molecular integrals is no more difficult than computation of integrals over higher angular momentum functions. We used the highly efficient recurrence formulas of Obara and Saika,[34] which were later generalized by the same authors,[35] for evaluation of the various integrals involved in Eq. (24). We can very easily obtain formulas for the integrals $[\mu\nu|r_{12}^{-1}|k\,]$ and $[\,k\,|r_{12}^{-1}|k'\,]$ by setting one or two exponents to zero in their four-index formulas. We found it convenient to make use of angular momentum transfer formulas. In the notation of Ref. 34, and for the case of 3-center overlap integrals, we have

$$\langle a+1_i|c|b\rangle = \langle a|c|b+1_i\rangle + (B_i - A_i)\langle a|c|b\rangle\,. \tag{35}$$

Similar formulas are easy to find for other types of integrals simply by permuting indices in the original Obara and Saika formulas and subtracting. [For example, Eq. (35) follows directly from subtraction of integral $\langle a+1_i|c|b\rangle$ and $\langle a|c|b+1_i\rangle$ as expressed in Eq. (20) of Ref. 34.] Use of these angular momentum transfer formulas reduces significantly the number of operations for computation of all necessary derivatives of molecular integrals, especially for high angular momentum functions. Moreover, we took advantage of the translational invariance of molecular integrals.[36] The present program was developed on a scalar computer CDC 855. The work of writing an integral package for a vector computer is in progress.[37] We used the so-called shell structure in which the set of all functions having the same exponent, center and angular momentum L defines a shell. Substantial savings in the time for integral evaluation are made if we calculate entire groups of integrals with functions of the same shells. In particular, we can use fitting basis sets for the electron density and exchange-correlation potential with large shells of s, p, and d symmetries. These basis sets proved to be efficient and accurate.[30,38,39] The program consists of sets of nested loops: over fitting functions and over pairs of primitive Gaussians. The inner loops have a large index limit. We avoided the use of logic statements inside the inner loops by explicitly writing the recurrence formulas for each pair of functions of a given pair of shells. These two characteristics (large index limit, absence of logic) are a prerequisite for an efficient vectorization of the code. Since we do not have to store derivatives of molecular integrals, the time consuming procedure[37] of compressing primitive integrals into contracted ones can be avoided.

IV. TESTS

Ultimately, any practical value of Eq. (24) will be judged by numerical tests. One can expect that improvements in the quality of the fitting of the exchange-correlation potential and energy would diminish the error in Eq. (24). This could be achieved by enlarging the size of the xc basis set, using a better sampling of grid points and/or using model core potentials. In what follows, we will investigate the accuracy of the equations and explore some possible improvements. We report here the results for some diatomic and triatomic molecules. These are all-electron calculations with various quality of xc basis sets and various fitting procedures. Model potential calculations and tests of the accuracy in such cases are under way and will be reported in the future.

Table I shows the equilibrium distance of N_2 computed by two methods. R_e (fit) is the equilibrium distance obtained from a third degree polynominal fit of the set of (energy, distance) points. With enough points chosen in an appropriate range, this method is precise to $\pm$ 0.001 a.u. R_e (grad) is the distance at which the gradient computed with Eq. (24) is zero. Because the gradient will eventually be used to search extrema of PES, we think that such a comparison is more meaningful than a comparison of exact (finite difference) vs approximate gradient at some given distance. All calculations reported in Table I were made with orbital and charge

TABLE I. Accuracy of the gradient computed with various choices of xc fitting procedures for the N_2 molecule.[a]

	Basis	N	Weight	R_e (fit) (a.u.)	r_e (grad) (a.u.)	ΔR_e (a.u.)
I	A	4	R	2.131	2.216	0.085
II	B	10	R^2	2.138	2.172	0.034
III	C	4	R^2	2.133	2.145	0.012
IV	B	10	R	2.134	2.124	0.010
V	C	10	R	2.134	2.124	0.010
VI	B	4	R	2.132	2.124	0.008
VII	C	4	R	2.132	2.124	0.008

[a] See text for the description of basis, fitting parameter N and weight.

density fit bases of comparable quality. The three xc bases $\{|l\rangle\}$ denoted by A,B,C increase in size (and quality) from A to C and have the following s, p, d patterns[38]:

A: 7,3,3 ,

B: 7,5,5 ,

C: 12,7,7 .

The only difference between A and B is that the latter has two additional high exponent p and d shells thus giving a better fit close to the nucleus. Basis C simply has more s, p, and d shells covering roughly the same range as in basis B. $1/N$ is the fraction of radial points kept from a Herman–Skillman-type mesh.[40] $N=4$ thus gives a grid with more points than $N=10$. Finally a weight is associated with each point of the grid used in the fitting procedure. We tried two weights: one which is proportional to R^2, the square of the distance to the nearest nucleus, or equivalently, to the volume of the grid cell, and one which is proportional to R. This last weight puts more emphasis for a good fit at points near the nuclei where the density is high and accordingly the contribution to E_{xc} is large. Dunlap has proposed a fitting procedure that minimizes the error on the xc energy.[41] We expect that this procedure would bring Eqs. (28)–(29) close to a minimum. His "best" weight, however, has to be recomputed at each iteration and this is not convenient. An approximation to this weight, which would remain unchanged through iterations, could be the best practical solution. Inspection of Table I leads us to make the following comments.

(1) R_e(fit) is not very sensitive to the choice of basis and grid whereas R_e(grad) is. This situation is reminiscent of the assessment of basis set quality from total energy vs Hellmann–Feynman force accuracy.[4,42,43]

(2) Case I shows that one has to be careful. A bad choice of XC basis can be disastrous.

(3) Pairs (II,IV) and (III,VII) seems to favor the weight proportional to R.

(4) Pairs (IV,V) and (VI, VII) shows that basis C is no better than basis B. This seems to indicate that basis B, although not very costly, is practically complete from the criterion set by Eq. (23).

(5) Pairs (IV,VI) and (V,VII) show that the finer grid is only slightly better. As in Eq. (4), this indicates that accurate gradient calculation is possible with a fitting procedure which remains economical.

Table II shows R_e(fit) and R_e(grad) for a few diatomics for two standard choices of basis and grid, "1" and "2" corresponding roughly to cases II and VI, respectively, of Table I (explicitly, for case 1: $N=10$, weight $=R^2$ and an "ordinary" xc basis was used. For 2: $N=4$, weight $=R$, and an "extended" basis was used). The same orbital and density fit bases (see footnotes to Table II) were used for both cases. In addition to R_e, we compare the vibrational frequency obtained from the fit, ω_e(fit), with what we can compute from finite difference of the gradient at R_e(grad), ω_e(grad). R_e(HF) and ω_e(HF) are the equivalent of R_e(grad) and ω_e(grad) but computed with the Hellmann–Feynman gradient instead. [It should perhaps be pointed out that the results reported for C_2, O_2, and CH in Table II are in fairly bad agreement with experimental values. This is because the calculations were done in the non-spin-polarized ("closed shell") option, sometimes with fractional occupation numbers. Therefore, the aforementioned diatomics in Table II are in states which do not correspond to pure multiplets.]

TABLE II. Accuracy of the gradient [Eq. (26)] and the Hellmann–Feynman gradient for two standard choices of basis and grid for some diatomics.[a]

	Case	R_e(fit) (a.u.)	R_e (grad)	R_e (HF)	ω_e(fit) (cm^{-1})	ω_e (grad)	ω_e (HF)
H_2	1	1.486	1.472	1.500	4160	4210	4070
	2	1.478	1.473	1.500	4250	4210	4300
C_2	1	2.578	2.581	3.30	1550	1320	1790
	2	2.566	2.569	3.20	1570	1580	1840
O_2	1	2.375	2.401	3.19	1400	1400	2550
	2	2.390	2.391	3.24	1410	1420	2470
N_2	1	2.138	2.172	3.00	2310	2330	2750
	2	2.132	2.124	2.82	2300	2200	2710
CO	1	2.194	2.208	3.14	2040	1940	2407
	2	2.195	2.183	3.06	2030	1990	2420
CH	1	2.192	2.440	3.01	2700	2390	2630
	2	2.189	2.177	2.74	2690	2710	3250

[a] The orbital basis set used have the patterns: H: (311/1), C: (5211/411), N: (5211/411), O: (5211/411). The patterns for the charge-density bases used are H: (5,2,2), C: (6,3,3), N: (9,6,6), O: (9,6,6). Cases 1 and 2 differ only in the following choices. Case 1: N = 10, weight $=R^2$, xc bases patterns: H: (6,2,2), C: (6,4,4), N: (7,5,5), O: (7,5,5). Case 2: N = 4, weight $=R$, XC bases patterns: H: (6,5,5), C: (8,7,7), N: (7,5,5), O: (10,9,9).

Table II shows that there is a systematic improvement of R_e(grad) for all diatomics when one goes from a less accurate choice of basis and grid 1 to the better choice 2. On the other hand, ω_e(grad) is always fairly good and shows no systematic improvement from 1 to 2. An interesting feature of Table II is that while the equilibrium distance computed from the Hellmann–Feynman force is very bad, the second derivative computed from it is reasonably good (except for O_2) and systematically larger than the exact value. This observation might eventually be helpful in developing a method for approximate evaluation of second derivatives.

Various terms contributing to the gradient near the equilibrium distance are shown in Table III. The quantities $E^{(\lambda)}$, F^{HF}, F^{OBC}, and F^{DBC} are defined by Eq. (25). Any error in the computed gradient can be ascribed to xc basis incompleteness and this error is thus denoted by $-F^{\mathrm{xcBC}}$ in analogy with the denominations $-F^{\mathrm{OBC}}$ and $-F^{\mathrm{DBC}}$. The values of $-F^{\mathrm{xcBC}}$ reported in Table III are estimates calculated from the force constant and the error made on equilibrium distance (see Table II, case 2). The precise values reported are not exact but the order of magnitude is right. What we call the energy-weight overlap force (EOF) and the exchange-correlation overlap force (xcOF) are just the absolute values of the third and fifth terms of Eq. (24). First, it should be noted that both F^{OBC} and F^{DBC} are large and must be computed exactly otherwise the gradient would be inaccurate. The fact that the force due to incompleteness of the xc basis, which is neglected in $E^{(\lambda)}$, is much smaller than the force due to incompleteness of the charge density fit basis reflects the relative smoothness of the xc energy function and

TABLE III. Basis correction force and overlap force contributions to the gradient for some diatomics close to their equilibrium distance. (All values in a.u.)

	$E^{(\lambda)}$	$-F^{HF}$	$-F^{OBC}$	$-F^{DBC}$	$-F^{xcBC}$	EOF	xcOF
H_2	−0.0017	−0.0220	+0.0563	−0.0361	+0.0017	0.2767	0.4520
C_2	−0.0835	−1.092	−2.249	+3.258	−0.0017	0.4596	0.6794
O_2	−0.0142	−3.233	−4.943	+8.162	−0.0006	0.7115	0.6048
N_2	+0.0151	−2.712	−3.959	+6.686	+0.0103	1.017	1.399
CO	+0.0352	−2.646	−2.249	+4.930	+0.0125	1.049	1.361
CH	+0.0016	−0.4146	−0.3283	+0.7445	+0.0032	0.2093	0.3191

the relative ease of its fitting. Various authors noted that the use of a basis that includes derivatives of "parent functions" (those of a usual basis) can improve very much the accuracy of the Hellmann–Feynman force.[26b,44] However, this method has a major drawback; the basis needed is much larger than the usual basis and this is very expensive at the SCF step because the matrix to be diagonalized is then 2–3 times larger. Second, it may be interesting to note that EOF and xcOF scale roughly with the bond order of the molecule, not with the total energy. Because xcOF is just the approximation to the two terms of Eq. (20), we can hope that, contrary to the "Hellmann–Feynman force only" approximation, our approximation will not deteriorate as we go to heavier elements.

In Table IV we show the results of geometry optimization for three molecules (HCN, H_2O, O_3). HCN was constrained to be linear and H_2O and O_3 were constrained to have C_{2v} symmetry. The basis and grid chosen for these cases were similar to choice IV of Table I, which seems to be a good compromise between accuracy and economy. The agreement between the geometry predicted from the criterion of minimizing the energy on one hand, and minimizing the gradient norm on the other, is satisfactory. The experimental values[45] given for comparison, show how accurate the geometries calculated with the LCGTO–LSD method can be. The results for ozone can be considered satisfactory as the geometry of this molecule seems to be very sensitive to the choice of basis set.[46] Indeed, ozone is an interesting case for which the Hartree–Fock method gives poor results but LSD methods give reasonably accurate geometries and excitation energy. The quality of the LSD wave function for O_3 has been previously discussed.[46]

Finally, Table V shows the time for one SCF iteration and the time needed to compute the gradient. The time required to compute the integrals is not shown as the program used is inefficient. It is currently being rewritten and we expect that evaluation of integrals will roughly be twice as fast as the evaluation of the gradient, on a scalar computer.[47] On a vector computer, the new integral package has been written and it runs at about 100 M flops per processor with 96% of the code being multitasked.[30] The essential thing to be seen in Table V is that, because we do not have to evaluate numerous and costly four-index integrals in our method, the computation of the gradient takes only a fraction of the time needed in the SCF step. The ratio (SCF/gradient) may decrease a little as we go to larger systems but almost surely never to the point where gradient evaluation would cost more than SCF.

TABLE IV. Equilibrium geometries obtained by minimizing the energy (E) and by minimizing the norm of the gradient $E^{(\lambda)}$.[a,b]

		E	$E^{(\lambda)}$	Experiment[c]
HCN	R(HC)	2.04	2.01	2.0107
	R(CN)	2.20	2.18	2.1845
H_2O	R(OH)	1.835	1.824	1.810
	θ	105.3	105.5	104.45
O_3	R(OO')	2.37	2.36	2.415
	θ	118.1	120.6	116.8

[a] The orbital basis set used have the patterns: H: (311/1), C: (521/41/1), N: (5211/411/1), O: (521/41/1). The charge density fit basis and exchange-correlation fit basis are as described for case 2 of Table II. For all three molecules, we used $N = 10$ and weight $= R$.
[b] Distances are in atomic units, angles in degrees.
[c] Reference 45.

V. CONCLUSION

An expression has been derived for the analytic evaluation of the gradient of the energy in the LCGTO–LSD method. The Hellmann–Feynman forces and those due to orbital and charge density basis set incompleteness are calculated exactly. To arrive at Eq. (24), we had to introduce an approximate relation which becomes exact in the limit of complete xc basis and fitting grid. Tests on diatomic and triatomic molecules showed that this approximation had little effect if reasonable care is taken and that a very accurate gradient can be calculated with an economical choice of xc basis and grid. We showed that Eq. (24) is valid not only

TABLE V. Comparative timings for SCF solution and gradient evaluation.[a]

	SCF (1 iter.)	Gradient	
	CPU	CPU	I/O
H_2[b]	2.4	2.4	3.7
CH[b]	3.4	6.9	9.0
CO[b]	6.2	13.1	8.6
O_3[c]	6.1	54	9.3

[a] The tests were done on a CYBER 170/855 and the time is given in seconds.
[b] We made no use of symmetry.
[c] We used C_{2v} symmetry at the SCF step but made no use of symmetry for gradient evaluation.

within the LSD approximation, but for a large class of xc energy functionals defined by Eqs. (3) and (5). This class includes many improved functionals, as compared to LSD, and all functionals that we could eventually implement within the framework of our method. The comparative timings in Table V show that our method for computing the gradient is not only accurate but also very economical when implemented as outlined in Sec. III. Finally, the possibility of routinely computing an accurate approximation to the LCGTO–LSD energy gradient opens up the possibility of efficient searches for extrema of potential-energy surfaces in complex systems with this method.

ACKNOWLEDGMENTS

Financial support from the Natural Sciences and Engineering Research Council–Canada and the Foundation Formation de Chercheurs det Aide à la Recherche–Québec are gratefully acknowledged as is the providing of computing resources by the Centre de Calcul de l'Université de Montréal. One of us (R.F.) would like to thank the FCAR and Alcan for scholarships. Part of the work of J. A. was performed at Cray Research.

[1]H. B. Schlegel, Adv. Chem. Phys. **67**, 249 (1987).
[2]P. Pulay, Adv. Chem. Phys. **69**, 241 (1987).
[3]*Geometrical Derivatives of Energy Surfaces and Molecular Properties*, edited by P. Jørgensen and J. Simons (Reidel, Dordrecht, 1986).
[4]P. Pulay, Mol. Phys. **17**, 197 (1969).
[5]G. Fitzgerald and H. F. Schaefer III, J. Chem. Phys. **83**, 1162 (1983).
[6]J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. Symp. **13**, 225 (1979).
[7]B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi, and H. F. Schaefer III, J. Chem. Phys. **72**, 4652 (1980).
[8]R. Krishnan, H. B. Schlegel, and J. A. Pople, J. Chem. Phys. **72**, 4654 (1980).
[9]M. Dupuis, J. Chem. Phys. **74**, 5758 (1981).
[10]T. J. Lee, J. Chem. Phys. **87**, 2825 (1987).
[11]P. Pulay, in *Methods of Electronic Struture Theory*, edited by H. F. Schaefer III (Plenum, New York, 1977), Vol. 4.
[12]S. Kato and K. Morokuma, Chem. Phys. Lett. **65**, 19 (1979).
[13]J. D. Goddard, N. C. Handy, and H. F. Schaefer, J. Chem. Phys. **71**, 1525 (1979).
[14]L. Adamowicz, W. D. Laidig, and R. J. Bartlett, Int. J. Quantum Chem. Symp. **18**, 245 (1984).
[15]G. Fitzgerald, R. J. Harrison, W. D. Laidig, and R. J. Bartlett, Chem. Phys. Lett. **117**, 433 (1985).
[16]L. A. Curtis and J. A. Pople, J. Chem. Phys. **88**, 7405 (1988); (b) M. W. Wong, R. H. Nobes, and L. Radom, J. Mol. Struct. **163**, 151 (1988).
[17]K. Ishida, K. Morokuma, and A. Komornicki, J. Chem. Phys. **66**, 2153 (1977); B. D. Joshi and K. Morokuma, *ibid.* **67**, 4880 (1977); T. Carrington, L. M. Hubbard, H. F. Schaefer, and W. H. Miller, *ibid.* **80**, 4347 (1980); S. K. Gray, W. H. Miller, Y. Yamaguchi, and H. F. Schaefer, J. Am. Chem. Soc. **103**, 1900 (1981); Y. Osamura, H. F. Schaefer, S. K. Gray, and W. H. Miller, *ibid.* **103**, 1904 (1981); C. Sosa and H. B. Schlegel, *ibid.* **109**, 4193 (1987); W. D. Allen and H. F. Schaefer III, J. Chem. Phys. **89**, 329 (1988).
[18]P. Pulay, Mol. Phys. **21**, 329 (1971); P. Pulay and J. Meyer, J. Mol. Spectrosc. **40**, 59 (1971); J. Meyer and P. Pulay, J. Chem. Phys. **56**, 2109 (1972); **57**, 3337 (1972); H. B. Schlegel, S. Wolfe, and F. Bernardi, *ibid.* **63**, 3632 (1975); **67**, 4181 (1977); **67**, 4194 (1977); I. L. Alberts and N. C. Handy, *ibid.* **89**, 2107 (1988); B. H. Besler, G. E. Scuseria, A. C. Scheiner, and H. F. Schaefer III, *ibid.* **89**, 367 (1988)
[19]H. F. Schaefer III and Y. Yamaguchi, J. Mol. Struct. **135**, 369 (1986).
[20]M. M. Goodgame and W. A. Goddard III, J. Phys. Chem. **85**, 215 (1981); A. D. McLean and B. Liu, Chem. Phys. Lett. **101**, 144 (1983).
[21]J. O. Noell, M. D. Newton, P. J. Hay, R. L. Martin, and F. W. Bobrowicz, J. Chem. Phys. **73**, 2360 (1980); S. P. Walch and C. W. Bauschlicher, Jr., in *Comparison of Ab initio Quantum Chemistry with Experiment for Small Molecules*, edited by R. J. Bartlett (Reidel, Dordrecht, 1985); S. R. Langhoff, C. W. Bauschlicher, Jr., and H. Partridge, J. Chem. Phys. **89**, 396 (1988); S. R. Langhoff and C. W. Bauschlicher, Jr., *ibid.* **89**, 2160 (1988).
[22]P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964); W. Kohn and L. Sham, Phys. Rev. A **140**, 1133 (1965).
[23]U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).
[24]D. R. Salahub, Adv. Chem. Phys. **69**, 447 (1987).
[25](a) C. Satoko, Chem. Phys. Lett. **83**, 111 (1981); (b) P. Bendt and A. Zunger, Phys. Rev. Lett. **50**, 1684 (1983); (c) C. Satoko, Phys. Rev. B **30**, 1754 (1984); (d) F. W. Averill and G. S. Painter, *ibid.* **32**, 2141 (1985); (e) **34**, 2088 (1986); (f) L. Verluis and T. Ziegler, J. Chem. Phys. **88**, 322 (1988).
[26]D. R. Salahub, in *Applied Quantum Chemistry Symposium*, edited by V. H. Smith, H. F. Schaefer III, K. Morokuma, Jr. (Reidel, Dordrecht, 1986); G. Chiarello, J. Andzelm, R. Fournier, N. Russo, and D. R. Salahub, Surf. Sci. **202**, L621 (1988).
[27]H. Sambe and R. H. Felton, J. Chem. Phys. **62**, 1122 (1975).
[28]B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).
[29]A preliminary outline of the equations has been given elsewhere, see Ref. 30.
[30]J. Andzelm, E. Wimmer, and D. R. Salahub, in *The Challenge of d and f electrons: Theory and Computation*, edited by D. R. Salahub and M. C. Zerner, ACS Symposium series (in press).
[31]W. Kohn and P. Vashishta, in *Theory of the Inhomogeneous Electron Gas* (Plenum, New York, 1983).
[32]See, for example, I. M. Gel'fand and S. V. Fomin, *Calculus of Variations* (Prentice Hall, Englewood Cliffs, NJ, 1963).
[33]J. Andzelm, E. Radzio, and D. R. Salahub, J. Chem. Phys. **83**, 4573 (1985).
[34]S. Obara and A. Saika, J. Chem. Phys. **84**, 3963 (1986).
[35]S. Obara and A. Saika, J. Chem. Phys. **89**, 1540 (1988).
[36]A. Banerjee, O. Jensen, and J. Simons, J. Chem. Phys. **82**, 4566 (1985); H. F. King and A. Komornicki, *ibid.* **84**, 3579 (1986); A. Banerjee, J. O. Jensen, and J. Simons, *ibid.* **84**, 3581 (1986).
[37]J. Andzelm (unpublished).
[38]J. Andzelm, E. Radzio, and D. R. Salahub, J. Comp. Chem. **6**, 520 (1985).
[39]J. Andzelm, N. Russo, and D. R. Salahub, Chem. Phys. Lett. **142**, 169 (1987).
[40]F. Herman and S. Skillman, *Atomic Structure Calculations* (Prentice Hall, Englewood Cliffs, NJ, 1963).
[41]B. I. Dunlap, J. Phys. Chem. **90**, 5524 (1986).
[42]C. W. Kern and M. Karplus, J. Chem. Phys. **40**, 1374 (1964).
[43]W. Fink and L. C. Allen, J. Chem. Phys. **46**, 3270 (1967).
[44]H. Nakatsuji, K. Kanda, M. Hada, and T. Yonezawa, J. Chem. Phys. **77**, 3109 (1982); **79**, 2493 (1983); P. Pulay, *ibid.* **79**, 2491 (1983).
[45]G. Herzberg, *Molecular Spectra and Molecular Structure* (Van Nostrand, Princeton, 1966), Vol. 3, and references therein.
[46]M. Morin, A. E. Foti, and D. R. Salahub, Can. J. Chem. **63**, 1982 (1985).
[47]A. St-Amant (unpublished).