Theor Chim Acta (1992) 83: 177–183

Theoretica
Chimica Acta

On the evaluation of derivatives of Gaussian integrals*

Trygve Helgaker[1] **and Peter R. Taylor**[2]**

[1] Department of Chemistry, University of Oslo, P.O.B. 1033, Blindern, N-0315 Oslo 3, Norway

[2] ELORET Institute, Palo Alto, CA 94303, USA

Received May 29, 1991/Accepted October 24, 1991

Summary. We show that by a suitable change of variables, the derivatives of molecular integrals over Gaussian-type functions required for analytic energy derivatives can be evaluated with significantly less computational effort than current formulations. The reduction in effort increases with the order of differentiation.

Key words: Gaussian integrals – Derivative integrals – Gradients – Energy derivatives

1. Introduction

Analytic energy derivative methods have revolutionized the application of computational quantum chemistry to problems of chemical interest [1]. The location and characterization of stationary points on polyatomic molecular potential energy surfaces can be accomplished so much more efficiently using analytic derivatives than with techniques based on computing energies alone that the development and extension of analytic derivative methods has been one of the most active fields of methodological research in quantum chemistry in recent years. Given the gradient and Hessian of the energy with respect to the nuclear coordinates, a variety of strategies have been developed that are guaranteed to converge to minima on potential surfaces and that can efficiently locate other stationary points, particularly transition states. These strategies can also be used to "walk" on surfaces from one minimum to another, thereby defining a reaction coordinate, and among the most elegant and conceptually illuminating studies of this sort are the investigations of Ruedenberg and co-workers on rearrangement reactions of small hydrocarbon species (see Refs. 2–5 and references therein). It is thus a great pleasure to dedicate this contribution to Professor Ruedenberg on the occasion of his 70th birthday.

---

* Dedicated to Prof. Klaus Ruedenberg

** *Mailing address*: NASA Ames Research Center, Moffett Field, CA 94035-1000, USA

Of course, in order to perform such walks and optimizations it is imperative to evaluate the energy derivatives efficiently at the computational level of interest (Hartree–Fock or some correlated treatment). As noted above, much work has been performed in this area, and several reviews are available [1, 6, 7]. We shall concentrate here on a topic that ultimately affects the computational effort necessary to evaluate energy derivatives for any *ab initio* method that relies on a basis set expansion of Gaussian one-electron functions.

Wave functions for polyatomic molecules are invariably expanded in a basis set that is centred on the various nuclei, and so in a calculation of the energy derivative of nth order with respect to the nuclear coordinates, up to nth-order derivatives of the one- and two-electron integrals are required. These derivative integrals can involve differentiation of the operators as well as differentiation of the basis functions, but the greatest computational problems arise from the differentiation of the basis functions. Like the evaluation of integrals over Gaussians [8, 9], the calculation of integrals over *differentiated* Gaussians has been the subject of many investigations and numerous efficient computational schemes have been devised. In this work we show how the efficiency of derivative integral evaluation can be improved by some simple manipulations of variables. We shall briefly review the McMurchie–Davidson scheme [8] for computing Gaussian integrals and derivative integrals, and then show how a change of differentiation variables simplifies the formulas.

2. Derivative integral formulas

We shall expand the Gaussian charge distributions that appear in the integrals in Hermite functions, as described by McMurchie and Davidson [8] (see also Saunders [9]). Let us represent an unnormalized Cartesian Gaussian function centred at $\vec{A}$ by:

$$G_{ijk}(\vec{r}, a, \vec{A}) = x_A^i y_A^j z_A^k \exp(-ar_A^2), \tag{1}$$

where $x_A = x - A_x$, etc. We can consider one Cartesian direction, say x, represented as

$$G_i(x, a, A_x) = x_A^i \exp(-ax_A^2). \tag{2}$$

The overlap distribution of two such functions is expanded as:

$$\begin{aligned}\Omega_{ij}(x, a, b, A_x, B_x) &\equiv G_i(x, a, A_x)G_j(x, b, B_x)\\ &= \sum_{t=0}^{i+j} E_t^{ij}(a, b, A_x, B_x)\Lambda_t(x, p, P_x),\end{aligned} \tag{3}$$

where the Hermite function $\Lambda_t(x, p, P_x)$ is defined by:

$$\Lambda_t(x, p, P_x) = (\partial/\partial P_x)^t \exp(-px_P^2) \tag{4}$$

with

$$\vec{P} = \frac{a}{p}\vec{A} + \frac{b}{p}\vec{B} \tag{5}$$

and

$$p = a + b. \tag{6}$$

The expansion coefficients $E_t^{ij}(a, b, A_x, B_x)$ are obtained from:

$$E_t^{i+1,j} = \frac{1}{2p} E_{t-1}^{ij} - \frac{b}{p} R_x E_t^{ij} + (t+1) E_{t+1}^{ij}, \tag{7}$$

where

$$R_x = A_x - B_x \tag{8}$$

and

$$E_0^{00} = \exp\left(-\frac{ab}{p} R_x^2 \right). \tag{9}$$

Henceforth we shall not always list the arguments of the expansion coefficients or Hermite functions, but we wish to emphasize here that the expansion coefficients depend on a, b, and R_x only, while the Hermite functions are independent of R_x:

$$\Omega_{ij}(x, a, b, A_x, B_x) = \sum_{t=0}^{i+j} E_t^{ij}(a, b, R_x) \Lambda_t(x, p, P_x). \tag{10}$$

In terms of the Hermite functions and expansion coefficients we can express a two-electron integral:

$$\begin{aligned} \iint & x_A^i y_A^k z_A^m \exp(-a r_A^2) x_B^j y_B^l z_B^n \exp(-b r_B^2) \\ & \times r_{12}^{-1} x_C^{i'} y_C^{k'} z_C^{m'} \exp(-c r_C^2) x_D^{j'} y_D^{l'} z_D^{n'} \exp(-d r_D^2) \, dr_1 \, dr_2 \end{aligned} \tag{11}$$

as

$$\begin{aligned} & \sum_{t=0}^{i+j} E_t^{ij}(a, b, A_x, B_x) \sum_{t'=0}^{i'+j'} E_{t'}^{i'j'}(c, d, C_x, D_x) \\ & \times \sum_{u=0}^{k+l} E_u^{kl}(a, b, A_y, B_y) \sum_{u'=0}^{k'+l'} E_{u'}^{k'l'}(c, d, C_y, D_y) \\ & \times \sum_{v=0}^{m+n} E_v^{mn}(a, b, A_z, B_z) \sum_{v'=0}^{m'+n'} E_{v'}^{m'n'}(c, d, C_z, D_z) \\ & \times (tuv \,|\, r_{12}^{-1} \,|\, t'u'v'), \end{aligned} \tag{12}$$

where

$$\begin{aligned} (tuv \,|\, r_{12}^{-1} \,|\, t'u'v') = \iint & \Lambda_t(x, p, P_x) \Lambda_{t'}(x, q, Q_x) \Lambda_u(y, p, P_y) \Lambda_{u'}(y, q, Q_y) \\ & \times \Lambda_v(z, p, P_z) \Lambda_{v'}(z, q, Q_z) r_{12}^{-1} \, dr_1 \, dr_2, \end{aligned} \tag{13}$$

and q and $\vec{Q}$ are defined analogously to p and $\vec{P}$ but for the second charge distribution. Thus in practice we evaluate integrals over the Hermite function basis and combine those with the expansion coefficients to give integrals over primitive Gaussians. Some modifications to the form of Eq. (12) are desirable from the point of view of efficiency, as discussed by Saunders [9], but for schematic purposes we can use Eq. (12). The first step, evaluation of the Hermite function integrals, is fast. The second step, which we can regard as a transformation from the Hermite function basis to the Cartesian Gaussian basis, is relatively time-consuming and is certainly more expensive than calculating the Hermite function integrals. Finally, if required, we combine these integrals with basis set contraction coefficients to give the final integrals. In fact, some of the expansion steps can be taken outside the contraction step, with a consequent improvement in efficiency.

In a derivative integral we are interested in derivatives of Ω_{ij}: $\partial\Omega_{ij}/\partial A_x$ and $\partial\Omega_{ij}/\partial B_x$ for first derivatives, for example. Conventionally, we would differentiate the orbitals of Eq. (2) first and then expand the overlap distributions of the differentiated orbitals analogously to Ω_{ij} above. For example, for the derivative with respect to A_x we obtain:

$$\frac{\partial\Omega_{ij}}{\partial A_x} = \sum_{t=0}^{i+j+1} F_t^{ij}\Lambda_t. \tag{14}$$

Note that the sum here is over more terms than appear in the undifferentiated charge distribution of Eq. (3) – higher orders of differentiation would increase this summation range further. The new coefficients F_t^{ij} are defined in terms of the coefficients E_t^{ij} above by:

$$F_t^{ij} = 2aE_t^{i+1,j} - iE_t^{i-1,j}. \tag{15}$$

Analogous coefficients can be defined for higher orders of differentiation or for differentiation with respect to B_x. In this approach, then, we compute derivative integrals using the same general scheme of Eq. (12) as for undifferentiated integrals. Since the expansion of the differentiated charge distributions in Hermite functions is longer than for the undifferentiated distributions, the work required to transform from the Hermite function basis to the Cartesian Gaussian basis is greater. Further, as the order of differentiation increases this extra work becomes larger and larger. Hence this approach is not well-suited to higher derivatives.

Let us instead consider differentiation with respect to the variables P_x and R_x, for which:

$$\frac{\partial}{\partial A_x} = \frac{a}{p}\frac{\partial}{\partial P_x} + \frac{\partial}{\partial R_x} \tag{16}$$

and

$$\frac{\partial}{\partial B_x} = \frac{b}{p}\frac{\partial}{\partial P_x} - \frac{\partial}{\partial R_x}. \tag{17}$$

We recall that the Hermite functions are independent of R_x, while the expansion coefficients are independent of P_x. Hence we can expect the expressions for the differentiated charge distributions to be simpler in terms of these variables, although we must eventually transform the derivatives back to the A_x, B_x representation. We obtain for the derivatives:

$$\frac{\partial\Omega_{ij}}{\partial P_x} = \sum_{t=0}^{i+j} E_t^{ij}\frac{\partial\Lambda_t}{\partial P_x} = \sum_{t=0}^{i+j} E_t^{ij}\Lambda_{t+1} \tag{18}$$

and

$$\frac{\partial\Omega_{ij}}{\partial R_x} = \sum_{t=0}^{i+j}\frac{\partial E_t^{ij}}{\partial R_x}\Lambda_t. \tag{19}$$

Denoting $\partial E_t^{ij}/\partial R_x$ by $E_t^{ij;1}$, we obtain the expansion relation:

$$E_t^{i+1,j;1} = \frac{1}{2p}E_{t-1}^{ij;1} - \frac{b}{p}(R_x E_t^{ij;1} + E_t^{ij}) + (t+1)E_{t+1}^{ij;1} \tag{20}$$

by differentiating Eq. (7) above.

We can make several important observations about these derivative formulas. First, the combination of expansion coefficients and Hermite functions in Eq. (18) above is over exactly the same range as the summation to give undifferentiated integrals: the only difference is that the degree of the Hermite function has increased by one. Hence the code required to evaluate this term is the same as required in the undifferentiated case, and the number of operations is also the same. (It is easy to see that this holds true in any order of differentiation for this term.) As we saw above, this is not the case if we differentiate with respect to the variables A_x and B_x, because then a linear combination of different degree Hermite functions and expansion coefficients appears.

Second, calculation of the differentiated expansion coefficients $E_t^{kl;1}$ requires essentially the same code again as for the undifferentiated case, with the obvious addition of an exta term in the expansion relation, and a starting value:

$$E_0^{00;1} = -\frac{2ab}{p} R_x E_0^{00}, \tag{21}$$

obtained by differentiating Eq. (9). As noted, the index range of the coefficients that are required is the same as that for the undifferentiated case, so the actual work required to combine Hermite function integrals and expansion coefficients does not increase. (The precomputation of the expansion coefficients themselves is of course a very rapid step.)

Third, in the usual scheme the index range of the program loops over the variables t, u, v depends on the direction of differentiation (i.e., differentiation with respect to A_x, A_y, etc). Thus these loops must be executed with different ranges for each of the three directions for first derivative integrals, for example. With our transformation of variables, the loop index ranges become independent of the direction of differentiation, so the program logic is simplified and the overheads are reduced. We may also note here that this approach in no way diminishes the possibilities for vectorizing the calculation of the integral derivatives. Indeed, the simplifications to the program loop structure are likely to enhance these possibilities.

Fourth, we can obtain an additional simplification as follows. Adding Eqs. (16) and (17) we obtain:

$$\frac{\partial}{\partial B_x} = \frac{\partial}{\partial P_x} - \frac{\partial}{\partial A_x}. \tag{22}$$

Now, (in addition to saving one multiplication) this form of the expression for the derivative with respect to B_x does not depend on the orbital exponents at all. Hence we can delay the transformation to the B_x derivative until later in the calculation, for example, until after the contraction step, so that the time required for this variable transformation becomes negligible. This is most important for first derivatives, as in any order of differentiation only one term can be treated this way.

In the case of higher derivatives there is a variety of terms to be considered but the scheme remains essentially the same. For example, the nth-order differentiated expansion coefficients with respect to R_x are obtained from the recursion formula:

$$E_t^{i+1,j;n} = \frac{1}{2p} E_{t-1}^{ij;n} - \frac{b}{p}(R_x E_t^{ij;n} + nE_t^{ij;n-1}) + (t+1)E_{t+1}^{ij;n} \tag{23}$$

with starting values:

$$E_0^{00;n+1} = -\frac{2ab}{p}(R_x E_0^{00;n} + nE_0^{00;n-1}) \tag{24}$$

and the identification:

$$E_t^{ij;0} = E_t^{ij}. \tag{25}$$

Higher derivatives of the Hermite functions with respect to P_x of Eq. (19) are trivially obtained. We note further that if the two charge distributions that appear in an integral are differentiated separately, the total savings is the product of the individual reductions in work, since the two differentiations are independent. For multiple differentiation of the same charge distribution, we recall that by using our transformation of differentiation variables the summation range in the Hermite function to Cartesian Gaussian transformation is independent of the order of differentiation. Hence the savings increase as the order of differentiation increases, since in the conventional scheme the work required to accomplish this transformation increases substantially with the order of differentiation. In order to obtain an estimate of what savings are possible, we must also include an estimate of the effort required to transform back to the A_x, B_x representation. We shall now present operation counts showing that it is always preferable to use our transformation of differentiation variables.

In order to simplify the counting we consider only floating-point operations (multiplication and addition), which are weighted equally. In addition, in our count we have not taken advantage of the possibility of deferring transformation of some derivatives until after contraction: in effect, we are counting operations only for primitive Gaussians and ignoring any additional savings that might accrue from moving manipulations outside the contraction step. If anything, neglecting this possibility favours the conventional approach to derivative integrals.

We have listed operation counts for differentiation of *SS*, *PP*, and *DD* distributions in Table 1. We have not included the calculation of the Hermite function integrals, which is fast and contributes the same work to both cases, the conventional approach and our new scheme. Further, the transformation of the second charge distribution in the integral has also been excluded. We see that for the *SS* case the total operation count is not much affected by whether or not the

Table 1. Operation counts for differentiation

	SS	*PP*	*DD*
	First derivatives		
Hermite/Cartesian transformation	12	396	4 032
P_x, R_x to A_x, B_x transformation	9	81	324
Total	21	477	4 356
Conventional	24	672	6 144
	Second derivatives		
Hermite/Cartesian transformation	42	1 386	14 112
P_x, R_x to A_x, B_x transformation	93	837	3 328
Total	135	2 223	17 460
Conventional	150	3 678	30 912

transformation of variables is performed. However, for higher angular momentum functions there is a decided advantage to using the transformation of variables, and this advantage is clearly growing with the order of differentiation. As a further illustration of this, we note that for third derivatives of a *PP* distribution, for example, the conventional method would require 14 448 operations, while using the transformation of variables the work would be reduced to 8 340 operations: a savings of 42%.

Finally, some other aspects of this scheme deserve comment. We note that:

$$\frac{\partial}{\partial R_x} = \frac{b}{p}\frac{\partial}{\partial A_x} - \frac{a}{p}\frac{\partial}{\partial B_x}. \tag{26}$$

Therefore, the operation $\partial/\partial R_x$ is *not* the same as the differentiation $\partial/\partial A_x - \partial/\partial B_x$. But if A and B coincide then the differentiation with respect to R_x does not contribute to the energy derivative: only the differentiation with respect to P_x contributes. This simplification is already used in the ABACUS program [10]. We also note that the use of translational invariance to reduce the computational labour is not affected by our transformation of variables: for first derivatives, for example, we have:

$$\frac{dI}{dP_x} + \frac{dI}{dQ_x} = 0, \tag{27}$$

where I represents the two-electron integral in Eq. (11), from the use of translational invariance.

3. Conclusions

We have shown that by employing a transformation of differentiation variables, the work required to evaluate derivative integrals can be substantially reduced. The advantages of our new approach increase both with the order of differentiation and with the angular momentum of the Gaussian functions involved. Savings will be obtained in the calculation of energy derivatives for any wave function that is expanded in a Gaussian basis. In particular, the economies obtained by applying these methods to the calculation of third or higher derivative integrals will be substantial.

Acknowledgements. Helpful discussions with W. Klopper are gratefully acknowledged. PRT was supported by NASA grant NCC 2-371.

References

1. Pulay P (1987) Adv Chem Phys 69:241
2. Valtazanos P, Elbert ST, Xantheas S, Ruedenberg K (1991) Theor Chim Acta 78:287
3. Xantheas S, Valtazanos P, Ruedenberg K (1991) Theor Chim Acta 78:327
4. Xantheas S, Elbert ST, Ruedenberg K (1991) Theor Chim Acta 78:365
5. Valtazanos P, Ruedenberg K (1991) Theor Chim Acta 78:397
6. Helgaker T, Jørgensen P (1988) Adv Quantum Chem 19:183
7. Gaw JF, Handy NC (1984) Roy Soc Chem Ann Rep C, 291
8. McMurchie LE, Davidson ER (1978) J Comput Phys 26:218
9. Saunders VR (1983) in: Diercksen GHF, Wilson S (eds) Methods of computational molecular physics. Reidel, Dordrecht
10. Helgaker TU, Almlöf J, Jensen HJAa, Jørgensen P (1986) J Chem Phys 84:6266