



An efficient method for calculating effective core potential integrals which involve projection operators

Chris-Kriton Skylaris^{a,1}, Laura Gagliardi^a, Nicholas C. Handy^a,
Andrew G. Ioannou^a, Steven Spencer^a, Andrew Willetts^a, Adrian M. Simper^b

^a *Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK*

^b *BNFL, Springfield, Preston PR4 0XJ, UK*

Received 23 June 1998; in final form 18 August 1998

Abstract

An efficient approach for evaluating effective core potential integrals which involve projection operators has been implemented in the MAGIC quantum chemistry program. The methodology is presented and its performance is examined through illustrative calculations on transition metal and actinide compounds. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

The study of complexes of heavy elements by ab initio quantum chemical methods requires the incorporation of relativistic effects since it has been shown that they play a significant role in the properties of such compounds [1].

To a great extent, the incorporation of such effects can be achieved through the use of effective core potentials (ECP) (also called pseudopotentials by some authors) or by some approximate method of solving the Dirac–Schrödinger equation adjusted to a many-electron system [2,3]. The design of ECPs is a compromise between two requirements. First, they have to imitate the interactions of core electrons of a many-electron atom with the valence electrons. Se-

cond, this imitation of core electrons has to be limited only to the valence region. This means that they have to produce valence orbitals that resemble the true valence orbitals in the valence region but have no features (nodes, etc.) in the core region. This requirement, which is not so trivial, is necessary to ensure that a smaller (i.e., valence only) basis set is needed for the atom in question. The ECPs are routinely used because in most cases they afford results of comparable accuracy to all-electron methods at reduced computational cost. Even in the cases of first-, second- and third-row atoms, where relativistic effects are usually negligible, the use of ECPs is very common due to the computational savings that are achieved from the use of ‘valence only’ basis sets.

The effective core potential operators of the functional form first proposed by Kahn et al. [4,5] are widely used today. The only effort in using them in a

¹ Corresponding author. E-mail: cks22@cus.cam.ac.uk

molecular calculation rests in the evaluation of their one-electron integrals which are their matrix elements with the Cartesian Gaussian functions of the basis set.

The matrix element of the ECP operator consists of two distinct types of integrals, one of which does not involve projection operators and one of which does. A few methods and corresponding computer programs for evaluating these integrals are known [5–8]. The integral which involves the projection operators is by far the most difficult from a theoretical and practical point of view.

In this Letter we present a new method to calculate this integral which is simple in concept and implementation. The implementation was done inside the MAGIC quantum chemistry program [9]. MAGIC was written entirely from the beginning for the purpose of studying large molecules of industrial significance at the DFT level of theory. No limitations on the maximum angular momentum of basis functions and projection operators or the values of the parameters involved exist. Furthermore, the size of the molecular system to be studied is only restricted by the available computer resources. Section 2 is a brief summary of the ECP method. Sections 3 and 4 present the new methodology for evaluating the integral. Finally, some example calculations on inorganic compounds are presented in Section 5.

2. ECP integrals

An ECP represents the interaction of the core electrons with the valence electrons by a potential operator of the following semi-local form [5]:

$$U^{\text{B}} = U_L^{\text{B}}(r) + \sum_{l=0}^{L-1} \sum_{m=-l}^l |lm; \text{B}\rangle \times [U_l^{\text{B}}(r) - U_L^{\text{B}}(r)] \langle lm; \text{B}|, \quad (1)$$

where $L - 1$ is the largest angular momentum orbital belonging to the core. It is expressed in terms of spherical polar coordinates (r, θ, ϕ) with origin on the centre B of the ECP. As one would expect, U_l^{B} and U_L^{B} approach zero asymptotically at large distance from B. The functions $|lm; \text{B}\rangle$ are real spheri-

cal harmonics Z_{lm} centred on B, according to the definition

$$Z_{lm}(\theta, \phi) = \left(\frac{(2l+1)(l-|m|)!}{2\pi(l+|m|)!} \right)^{1/2} \times \begin{cases} \cos(|m|\phi) P_l^{|m|}(\cos\theta), & m > 0, \\ \frac{1}{\sqrt{2}} P_l^{0l}(\cos\theta), & m = 0, \\ \sin(|m|\phi) P_l^{|m|}(\cos\theta), & m < 0. \end{cases} \quad (2)$$

The core potential operators are expressed in analytical form by fitting the numerical potentials to linear combinations of Gaussian functions:

$$U_L^{\text{B}}(r) = \frac{N_c}{r} + \sum_k d_{kL} r^{n_{kL}} e^{-\zeta_{kL} r^2}, \quad (3)$$

$$U_l^{\text{B}}(r) - U_L^{\text{B}}(r) = \sum_k d_{kl} r^{n_{kl}} e^{-\zeta_{kl} r^2}, \quad (4)$$

where N_c is the number of core electrons. In all ECPs derived to date, the powers n_{kL} and n_{kl} are restricted to the values 0, -1 and -2 , although the method which we describe in this Letter assumes no such restriction. We will try to follow the notation of earlier works on core potential integral evaluation [5,6,8]. The basis sets we use consist of Cartesian Gaussian functions. The general form of an unnormalised Cartesian Gaussian function on centre A is

$$\phi_A = x_A^{n_x} y_A^{n_y} z_A^{n_z} \exp(-\alpha r_A^2). \quad (5)$$

Evaluating the matrix elements of the integrals between functions ϕ_A and ϕ_C results in the evaluation of two distinct types of integrals (referred to as type 1 and type 2).

$$\chi_{\text{AC}} = \int \phi_A r^{n_{kL}} \exp(-\zeta_{kL} r^2) \phi_C d\tau \quad (6)$$

and

$$\gamma_{AC} = \int_0^\infty \langle \phi_A | lm; B \rangle(r) r^{n_{kl}+2} \exp(-\zeta_{kl} r^2) \times \langle \phi_C | lm; B \rangle(r) dr \quad (7)$$

where

$$\langle \phi_A | lm; B \rangle(r) = \int \phi_A(\Omega, r) Z_{lm}(\Omega) d\Omega. \quad (8)$$

Formulae for the analytical calculation in spherical polar coordinates around the centre B of the core potential are available for the above integrals [6,8]. In this context, the calculation of the type 1 integral, which does not involve projection operators, is relatively straightforward. Calculation of the type 2 integral which involves projection operators is more involved, at least from an implementational point of view of the radial integral. The purpose of this Letter is to present a simpler method of calculating the type 2 integral.

3. Calculation of the type 2 integral

Our approach towards the evaluation of the type 2 integral consists in performing the angular integrations analytically and the radial integration by quadrature on a one-dimensional grid of points. Previous efforts [5,6,8] have focused on the analytical evaluation of the radial integrals

$$\int_0^\infty e^{-\xi r^2} i_l(br) i_m(cr) r^n dr \quad (9)$$

that emerge after the full expansion of Eq. (7) in spherical polar coordinates. We would expect the evaluation of this integral by quadrature to be inefficient compared to the analytical approach. On the other hand, quadrature makes possible the calculation of integrals which cannot be done analytically. We can therefore use it to evaluate directly the integral of Eq. (7) as a whole. This leads to an efficient method for evaluating the type 2 integral. This would not be possible without the presence of the angular projection operators that allow the isolation of the basis functions from each other, as far as the angular integration is concerned.

First, the terms $\langle \phi_A | lm; B \rangle(r)$ are evaluated analytically at every point r of the radial grid according

to the following formula which can be derived by expressing the Gaussian function ϕ_A in terms of functions centered on the ECP centre [5,6]

$$\begin{aligned} \langle \phi_A | lm; B \rangle(r) = & 4\pi \sum_{k_x=0}^{n_x} \sum_{k_y=0}^{n_y} \sum_{k_z=0}^{n_z} \binom{n_x}{k_x} \binom{n_y}{k_y} \\ & \times \binom{n_z}{k_z} \overline{BA}_x^{n_x-k_x} \overline{BA}_y^{n_y-k_y} \overline{BA}_z^{n_z-k_z} \\ & \times \sum_{\lambda=0}^{l+\nu} r^{k_x+k_y+k_z} \left[e^{-\alpha(\overline{AB}^2+r^2)} \right. \\ & \times i_\lambda(2\alpha\overline{AB}r) \left. \right] \sum_{\mu=-\lambda}^{\lambda} Z_{\lambda\mu}(\widehat{AB}) \\ & \times \int Z_{\lambda\mu}(\hat{\mathbf{r}}) Z_{lm}(\hat{\mathbf{r}}) \hat{x}^{k_x} \hat{y}^{k_y} \hat{z}^{k_z} d\Omega, \end{aligned}$$

where $\nu = n_x + n_y + n_z$ is the angular momentum of ϕ_A , \overline{BA} is the distance from the centre B of the ECP to the basis function centre A and \widehat{AB} is the unit vector from B to A. The angular integrals in the above equation are expressed as sums of triple integrals of spherical harmonics. The term $e^{-\alpha(\overline{AB}^2+r^2)} i_\lambda(2\alpha\overline{AB}r)$ has to be evaluated as the exponential of the sum of its logarithms for retaining precision. The spherical modified Bessel function of the first kind i_λ is computed by a standard implementation [10] and its asymptotic form is used for large values of its argument [11].

The computation of γ_{AC} is completed by summing the $\langle \phi_A | lm; B \rangle(r)$ and $r^{n_{kl}} \exp(-\zeta_{kl} r^2)$ terms over the radial grid

$$\gamma_{AC} = \sum_{i=1}^{n_r} \langle \phi_A | lm; B \rangle(r_i) \langle \phi_C | lm; B \rangle(r_i) r_i^{n_{kl}} \times \exp(-\zeta_{kl} r_i^2) w_i, \quad (10)$$

where n_r is the number of radial quadrature points and w_i are the corresponding weights.

A point worth noting is that the calculation of the $\langle \phi_A | lm; B \rangle(r_i)$ (which is the computationally intensive step) scales linearly with the size of the system (number of basis functions N). Of course, the radial quadrature part scales quadratically, but it has a small prefactor so its contribution is negligible for all systems we present in this Letter. If the quadratic

part starts to dominate in large systems, linear scaling could still be achieved, if necessary, by screening of integrals that can be predicted to be negligibly small.

For the calculation of first and second derivatives of γ_{AC} with our mixed analytical–numerical method, we recommend to avoid differentiating directly the ECP operator by using the principle of translational invariance as has been done by Cui et al. [12] for other ECP integral methods.

We tested our method by comparing with an implementation of a variant [13] of the method of McMurchie et al. [6], which is available in MAGIC. It was ensured that the variant code yielded accurate results; however, it was not optimised for speed. We found that our method is considerably faster than the variant which requires a prohibitively large amount of time for molecules larger than 4–5 atoms. A wide range of molecular geometries was used in the tests and they seem to have no effect on the method. For example, correct results are obtained for UO_2^{2+} (U ECP from [14], DZP basis on O) when varying the U–O distance from 0.01 to 10.0 a_0 .

4. Radial quadrature

As far as the radial quadrature is concerned, the Log3 scheme of Mura et al. [15] was found to be particularly efficient. In order to obtain the points, we start with a number n_r of equally spaced points x_i on the interval (0,1)

$$x_i = \frac{i}{n_r + 1} \quad (11)$$

then these points are mapped to points r_i on the interval $(0, +\infty)$ by

$$r_i = -\alpha \ln(1 - x_i^3), \quad (12)$$

where α is an empirically determined parameter and has the value 5 or 7, according to the type of atom [15]. The weights are given by

$$w_i = \frac{3\alpha^3 x_i^2 \ln^2(1 - x_i^3)}{(1 - x_i^3)(n_r + 1)}. \quad (13)$$

The form of Eqs. (12) and (13) allows for the use of an integration technique that converges the values of

the integrals to a certain predefined precision. We can gradually increase the number of points without having to recalculate the value of the integrand at the points we have already used in the previous step.

In particular, we can first evaluate the integrals with a number of radial points $n_r^{(1)}$ and weights given by Eq. (13). If we then choose to use $n_r^{(2)} = 2n_r^{(1)} + 1$ radial points for evaluating the integrals at improved precision, we observe that every second point

$$r_j^{(2)}, \quad j = 2, 4, \dots, (n_r^{(2)} - 1), \quad (14)$$

coincides with a point of the first set $n_r^{(1)}$ and its corresponding weight $w_j^{(2)}$ is half of the weight $w_i^{(1)}$ of $r_i^{(1)}$. Therefore calculating an integral with all the $n_r^{(2)}$ quadrature points is equivalent to calculating it using only the subset

$$r_k^{(2)}, \quad k = 1, 3, \dots, n_r^{(2)} \quad (15)$$

and adding to the result half of the value calculated in the previous step with the $n_r^{(1)}$ points.

Using this algorithm we can increase the number of quadrature points consecutively, and check the convergence of the integrals at each step. We use only the extra points added in each step, therefore avoiding carrying out the integration from the beginning with all the points. In our implementation, we start with 31 points and the results typically converge to the 10th decimal by the time we reach 2047 points or 4097 points for a few cases.

The performance of the Log3 scheme was found to be consistently better than that of the Euler–Maclaurin radial quadrature scheme [16] which was also examined. This scheme maps points to the $(0, +\infty)$ interval according to

$$r_i = \frac{x_i^2}{(1 - x_i)^2}. \quad (16)$$

The Euler–Maclaurin scheme is not as efficient for our purpose, requiring on average an order of magnitude more points in order to produce the same results as the Log3 scheme. This trend was common in all examples we tried. As an illustration, the number of points required by each type of grid for a variety of molecules, ECPs and basis sets is shown in Table 1.

Both quadrature schemes are routinely used in DFT calculations for the evaluation of the

Table 1

Comparison of the performance of the Log3 and Euler–Maclaurin radial grids. Each molecule contains only one ECP centre (the metal atom) with the accompanying valence basis set. The integrals were converged to the 10th decimal digit

Molecule	Basis	ECP	Log3 points	Euler points
RbCl	631G2DP	Wadt et al. [26]	4095	16383
AuH ₂ O ⁺	DZ	Wadt et al. [27]	4095	32767
UO ₂ ²⁺	DZP	Hay et al. [14]	2047	65535
UF ₆	DZ	Hay et al. [14]	1023	65535
AgH	PVTZ	Hay et al. [17]	255	32767
Pt(C ₂ H ₄)Cl ₃ ⁻	DZ	Wadt et al. [27]	127	65535
RbCl	631G2DP	[28]	127	255
AuH ₂ O ⁺	DZ	Andrae et al. [29]	127	511
UO ₂ ²⁺	DZP	Küchle et al. [30]	127	255
UF ₆	DZ	Küchle et al. [30]	127	255
AgH	PVTZ	Andrae et al. [29]	127	511
Pt(C ₂ H ₄)Cl ₃ ⁻	DZ	Andrae et al. [29]	127	511

Each molecule contains only one ECP centre (the metal atom) with the accompanying valence basis set. The integrals were converged to the 10th decimal digit.

exchange-correlation potential matrix elements. The reason for the better performance of the Log3 scheme could be attributed to its specific design for integration of Gaussian functions in contrast to the Euler–Maclaurin scheme which is supposed to be of more general applicability.

Having decided upon the type of quadrature and integration technique to be used, there remains the implementational question of memory usage as the number of basis functions and quadrature points increases. The values of $\langle \phi_A | lm \rangle (r_i)$ are calculated and stored in an array of size $N_{\text{BF}} \times n_r^{(k)}$ where N_{BF} is the number of basis functions. Furthermore the values of $e^{-\xi(\overline{AB^2+r^2})} i_\lambda(2\xi\overline{ABr})$ are also precalculated and stored in an array of size $N_{\text{exp}} \times (L + \nu_{\text{max}}) \times n_r^{(k)}$, where N_{exp} is the total number of primitive Cartesian Gaussian exponents and ν_{max} is the highest orbital angular momentum in the basis set. The sizes of these arrays can be substantial for large molecules when the number of basis functions approaches or exceeds 1000 and it is obviously affected by the increasing number of quadrature points in the iterations for converging the integrals. This dissipation of memory can be avoided by performing the radial integration using *batches* of radial points. In such a batch formalism, subsets of n_{batch} of the radial points are used in each step of the integration, until all the $n_r^{(k)}$ radial points are exhausted. Thus, $n_r^{(k)}$ is replaced by n_{batch} in the sizes of the relevant arrays. In the present implementation the value of n_{batch} was

set to 64. The use of batches of points makes the memory demands of the program more or less the same as those of the usual one-electron integrals.

5. Illustrative results

We are currently examining a variety of molecular systems and we present here some preliminary results as examples. We give in Table 2 some single-point energies of molecules containing various heavy elements with ECPs. Single-point energies are a means of verifying our code and provide a point of reference for others who may try this method.

The uranium and rhodium atoms were represented by ECPs with the corresponding basis sets from references [14] and [17], respectively. The geometries of the rhodium clusters were taken from crystal structures with the aid of the Quest program of the Cambridge Crystallographic Database [18], while the geometries of the 18crown6 and the uranyl were

Table 2
Single point LDA(exchange) energies

Molecule	Energy, E_h
Rh ₄ (CO) ₁₂	-1428.3449
Rh ₈ (CO) ₁₉ C	-2337.8623
UO ₂ (18crown6) ²⁺	-1104.7779
UO ₂ ²⁺ + 18crown6	-1104.2977

obtained from a molecular mechanics geometry optimisation using Cerius² [19]. The LDA exchange functional was used throughout and no correlation functional was included. Standard double-zeta basis sets on the atoms not bearing ECPs were used.

Carbonyl clusters of metals are common examples of molecules with metal–metal bonds [20]. There are numerous experimental studies on these systems due to their complexity of structure, variety of unusual reactions, resemblance to bulk metals and possible use in catalysis [21,22]. The understanding of their electronic structure is also a field of active research [23]. There are no localised orbitals and the chemical concept of a bond between two atoms due to an electron pair cannot be used. Their structures are explained often in terms of electron counting rules which have been derived from rudimentary LCAO arguments. There are no general trends and correlations with the other great class of clusters, the halide clusters whose properties are rationalised by different arguments. Our aim is to optimise the geometries of representative types of clusters at the DFT level of theory with the most important relativistic contributions accounted for through the use of ECPs. We then hope to derive conclusions concerning the stability and electronic structure of these compounds.

Ligands which bind uranium, plutonium and other actinides are of prime importance in the nuclear industry. For example, ligands such as tributyl phosphate (TBP) are used in the separation by extraction of uranium and plutonium from nuclear fission products. Also, special ligands for removing actinides from the blood in cases of contamination exist [24]. There is potential for improving the efficiency of such separative processes by designing ligands with better selectivity. Experimental investigation of such systems is quite expensive and time consuming because of the hazards involved. A code like MAGIC is expected to provide chemically useful information on such systems. As a starting model we have decided to investigate the interaction of a uranyl cation with the 18crown6 ether. Results by other workers using the molecular dynamics approach [25] claim that the complex is stable in the gaseous phase, but dissociates in water. In the gaseous phase, we find a dissociation energy of $0.4802E_h$ (301.3 kcal/mol) which indicates quite a strongly bound complex (see Table 2). A plot of a density isosurface of the

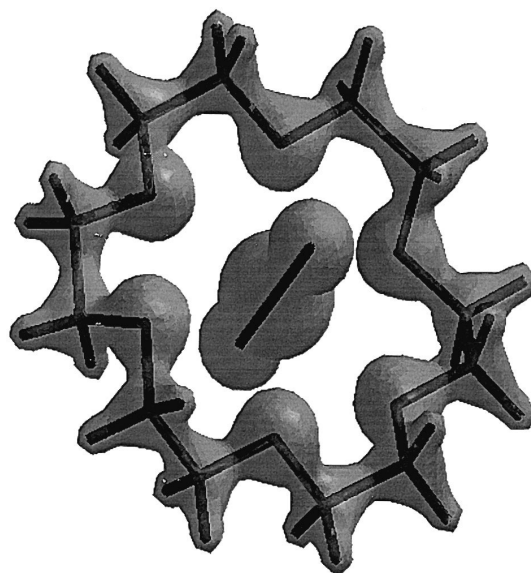


Fig. 1. Plot of a density isosurface of $UO_2(18crown6)^{2+}$. The U core is represented by an ECP and therefore its core electron density is not included in the plot.

complex is given in Fig. 1. Studies are underway to examine its stability in water using the solvent models of MAGIC. The results of these and related studies will be presented in a forthcoming publication.

6. Conclusions

An efficient approach which combines analytical angular with numerical radial integration is suggested for calculating the ECP integrals that involve projection operators. It is much simpler in concept and implementation than purely analytical integration methods. The implementation of the method has been done in the MAGIC program and is used in calculations on large molecules.

Acknowledgements

C-KS would like to thank BNFL and EPSRC for funding. LG would like to thank the European Union for support within the ‘Training Mobility Research’, Contract No. ERBFMRXCT96/0088.

References

- [1] P. Pyykkö, *Chem. Rev.* 88 (1988) 563.
- [2] C. Chang, M. Pellissier, P. Durand, *Physica Scripta* 34 (1986) 394.
- [3] M. Douglas, N.M. Kroll, *Ann. Phys.* 82 (1974) 89.
- [4] L.R. Kahn, W.A. Goddard III, *J. Chem. Phys.* 56 (1972) 2685.
- [5] L.R. Kahn, P. Baybutt, D.G. Truhlar, *J. Chem. Phys.* 65 (1976) 3826.
- [6] L.E. McMurchie, E.R. Davidson, *J. Comput. Phys.* 44 (1981) 289.
- [7] M. Pelissier, N. Komih, J.P. Daudey, *J. Comput. Chem.* 9 (1988) 298.
- [8] R.M. Pitzer, N.W. Winter, *Int. J. Quantum Chem.* 40 (1991) 773.
- [9] A. Willetts, A.M. Simper, L. Gagliardi, A.G. Ioannou, C.-K. Skylaris, S. Spencer, N.C. Handy, *MAGIC* (to be published).
- [10] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C*, 2nd ed., Cambridge University Press, Cambridge, 1994.
- [11] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series and Products*, 5th ed., Academic Press, New York, 1994.
- [12] Q. Cui, D.G. Musaev, M. Svensson, K. Morokuma, *J. Phys. Chem.* 100 (1996) 10936.
- [13] A.M. Simper, *Relativistic Calculations on Molecules*, Ph.D. Thesis, Department of Chemistry, University of Cambridge, Cambridge, 1998.
- [14] P.J. Hay, Actinide ECPs, private communication.
- [15] M.E. Mura, P.J. Knowles, *J. Chem. Phys.* 104 (1996) 9848.
- [16] C.W. Murray, N.C. Handy, G.J. Laming, *Mol. Phys.* 78 (1993) 997.
- [17] P.J. Hay, R.W. Wadt, *J. Chem. Phys.* 82 (1985) 270.
- [18] O. Kennard, F. Allen, *Chem. Design Autom. News* 8 (1993) 31.
- [19] Cerius², Version 3, Molecular Simulations Inc., San Diego, CA.
- [20] F.A. Cotton, G. Wilkinson, *Advanced Inorganic Chemistry*, 4th ed., Wiley, New York, 1980.
- [21] M.G. Richmond, *Coord. Chem. Rev.* 168 (1998) 177.
- [22] B.F.G. Johnson, Y.V. Roberts, E. Parisini, R.E. Benfield, *J. Organomet. Chem.* 478 (1994) 21.
- [23] M. McPartlin, *Polyhedron* 3 (1984) 1279.
- [24] A.J. Freeman, C. Keller (Eds.), *Handbook on the Physics and Chemistry of the Actinides*, Elsevier, Amsterdam, 1986.
- [25] P. Guilbaud, G. Wipff, *J. Phys. Chem.* 97 (1993) 5685.
- [26] R.W. Wadt, P.J. Hay, *J. Chem. Phys.* 82 (1985) 284.
- [27] P.J. Hay, R.W. Wadt, *J. Chem. Phys.* 82 (1985) 299.
- [28] Institut für Theoretische Chemie, Universität Stuttgart, ECPs and corresponding valence basis sets, <http://www.theochem.uni-stuttgart.de>.
- [29] D. Andrae, U. Haussermann, M. Dolg, H. Stoll, H. Preuss, *Theor. Chim. Acta* 77 (1990) 131.
- [30] W. Küchle, M. Dolg, H. Stoll, H. Preuss, *J. Chem. Phys.* 100 (1994) 7535.