Theor Chem Acc (2011) 128:115–125
DOI 10.1007/s00214-010-0771-1

REGULAR ARTICLE

Direct calculation of the Coulomb matrix: Slater-type orbitals

Ignacio Ema · Rafael López · Guillermo Ramírez · Jaime Fernández Rico

Received: 30 October 2009 / Accepted: 24 May 2010 / Published online: 15 June 2010

Abstract It is proved that the evaluation of the Coulomb potential and the calculation of its matrix elements can be carried out in separate steps whose costs formally increase as the square of the number of basis functions. The resulting method for computing the Coulomb matrix is reported, and its main features are tested with a trial program for Slater functions. A comparison of the Coulomb matrices obtained with this method and those computed from the repulsion integrals shows that the current procedure is potentially exact, highly accurate in practice, and much less expensive. The effects of basis product cutoff and long-range separation have been analyzed finding that the method tends to linear scaling with the size of the system. Moreover, the storage requirements are very low since two-electron integrals are completely absent, and it is well suited to be used in the density functional context.

Keywords Coulomb operator · Electron density · Density functional

Dirección General de Investigación Científica y Técnica (CTQ2007-63332).

I. Ema · R. López (✉) · G. Ramírez · J. F. Rico
Facultad de Ciencias, Departamento de Química Física Aplicada. 01.14, Universidad Autónoma de Madrid, Madrid, Spain
e-mail: rafael.lopez@uam.es

I. Ema
e-mail: nacho.ema@uam.es

G. Ramírez
e-mail: guillermo.ramirez@uam.es

J. F. Rico
e-mail: jaime.fernandez@uam.es

1 Introduction

In previous works [1, 2], a partition of the molecular density, $\rho(\mathbf{r})$, into atomic contributions, $\rho^A(\mathbf{r})$, accompanied by an expansion of the atomic contributions in real harmonics times radial factors, was reported.

This approach has two appealing features for the calculation of several functionals of the density needed in practical applications: (1) both partition and expansion are exact in the sense that, for every $\mathbf{r}$, the sum of atomic fragments is just the density, $\rho(\mathbf{r})$, and the expansion of every atomic fragment exactly converges to its density, $\rho^A(\mathbf{r})$; (2) the atomic fragments, $\rho^A(\mathbf{r}_A)$, are minimally deformed in the sense that they yield the quickest convergent expansion of the long-range Coulomb potential compatible with the LCAO expression of the density in the working basis set.

Although this partition was initially aimed at facilitating the calculation of several quantities (density gradient, molecular electrostatic potential, force fields, and alike), it became soon evident that it could be an aid for a better understanding of some basic chemical concepts (molecular structure [3], chemical forces [4], and so forth), and it was successfully applied in this context.

In this paper, we take up the original aim by developing an efficient procedure for the calculation of the Coulomb matrix directly from the Coulomb potential which does not require the calculation of any two-electron integral.

The underlying idea, advanced several years ago [5, 6], is that the calculation of the Coulomb potential and the evaluation of its matrix elements are different processes which can be carried out into separate steps whose computational costs formally increase as the square of the number of basis functions, m, instead of the m^4 process required when the repulsion integrals are used. The exploration of this

possibility with the aid of the density partition and expansion method is the purpose of the present work.

In Sect. 2, the partition of the density is briefly revisited. Although it has been discussed in detail elsewhere [1, 2], the key points are summarized herein for completeness and to remark the fact that the resulting expansion is exact. Sections 3 and 4 detail the evaluation of the Coulomb potential and the calculation of its matrix elements and the Coulomb energy, stressing that both processes are potentially exact and formally m^2 dependent. In Sect. 5, the formalism presented in the previous sections is applied to the development of a trial program for the calculation of the Coulomb matrix in case of Slater functions. It is proved that the matrix elements computed with this procedure may coincide with those attained from two-electron integrals with any prescribed accuracy, provided that sufficient terms are included in the expansion of the atomic fragments. It is also proved that, as a consequence of the effect of basis product cutoff and the long-range separation, the current method scales linearly (instead of quadratically) with both the number of basis functions and the number of centers for large systems.

As it is well known, the formal scaling of the repulsion integrals with m^4 is not realized in practice because basis product cutoff [7, 8] can reduce this scaling to m^2 for size increasing systems, and further savings may come from neglecting small short-range terms. Another usual way to circumvent the integrals problem is based on the substitution of the LCAO density by an approximate density represented in terms of an auxiliary basis set [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]. The relationship between these approaches and our proposal is commented in Sect. 6.

Finally, in Sect. 7, the main features and capabilities of this procedure are pointed out.

2 Expansion of the molecular density in spherical harmonics centered at the nuclei

In the method of deformed atoms in molecules (DAM) [21], the electron density is partitioned into atomic fragments and expanded as:

$$\rho(\mathbf{r}) = \sum_A \rho^A(\mathbf{r}_A) = \sum_A \sum_{l=0}^{\infty} \sum_{m=-l}^{l} z_l^m(\mathbf{r}_A) f_{lm}^A(r_A) \tag{1}$$

with $\mathbf{r}_A = \mathbf{r} - \mathbf{R}_A$, $\mathbf{R}_A$ being the position of nucleus A, $r_A = |\mathbf{r}_A|$, $z_l^m(\mathbf{r}_A)$, the real spherical harmonics, and $f_{lm}^A(r_A)$, the radial factors.

The calculation of the radial factors, $f_{lm}^A(r_A)$, will be discussed below. The regular real spherical harmonics are defined as:

$$z_l^m(\mathbf{r}) = r^l z_l^m(\mathbf{r}/r) = (-1)^m r^l P_l^{|m|}(\cos\theta)\, \Phi_m(\phi) \tag{2}$$

where $P_l^{|m|}(\cos\theta)$ are the associated Legendre functions—see Ref. [22] eq 8.751.1–, $\Phi_m(\phi) = \cos m\phi$ for $m \geq 0$, and $\Phi_m(\phi) = \sin |m|\phi$ for $m < 0$.

A detailed discussion on the practical implementation of this partition has been previously reported [1, 2]. A revision of the key points is given herein for completeness.

The starting point is the LCAO expression of the density which can be written as:

$$\rho(\mathbf{r}) = \sum_{A=1}^{N} \left\{ \sum_{a\in A} \sum_{a'\in A} \rho_{aa'}^{AA}\, \chi_a(\mathbf{r}_A)\, \chi_{a'}(\mathbf{r}_A) + \sum_{B\neq A} \sum_{a\in A} \sum_{b\in B} \rho_{ab}^{AB}\, \chi_a(\mathbf{r}_A)\, \chi_b(\mathbf{r}_B) \right\} \tag{3}$$

where A and B label the N nuclei; a and b, the basis functions, $\chi_a(\mathbf{r})$, $\chi_b(\mathbf{r})$, centered respectively at $\mathbf{R}_A$ and $\mathbf{R}_B$, and ρ_{ab}^{AB} denotes the corresponding AB block of the density matrix.

The one-center term of the r.h.s. of Eq. 3 can be trivially expanded in products of spherical harmonics times radial factors. The two-center terms, which contain charge distributions $\chi_a(\mathbf{r}_A)\ \chi_b(\mathbf{r}_B)$ localized in the region between the A and B centers, are decomposed into one part $d_{ab}^A(\mathbf{r}_A)$ assigned to A and the rest, $d_{ab}^B(\mathbf{r}_B)$, assigned to B, so that collecting the terms attributed to A, one can define:

$$\rho^A(\mathbf{r}) = \sum_a \sum_{a'} \rho_{aa'}^{AA}\, \chi_a(\mathbf{r}_A)\, \chi_{a'}(\mathbf{r}_A) + 2 \sum_{B\neq A} \sum_a \sum_b \rho_{ab}^{AB}\, d_{ab}^A(\mathbf{r}_A) \tag{4}$$

and the expansion of $\rho^A(\mathbf{r})$ in spherical harmonics results:

$$\rho^A(\mathbf{r}) = \sum_{l=0}^{2L} \sum_{m=-l}^{l} z_l^m(\mathbf{r}_A) \sum_a \sum_{a'} \rho_{aa'}^{AA} f_{lm}^{aa'}(r_A) + 2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} z_l^m(\mathbf{r}_A) \sum_{B\neq A} \sum_a \sum_b \rho_{ab}^{AB} f_{lm}^{ab}(r_A) \tag{5}$$

where L is the highest value of l in the basis set functions centered at nucleus A.

Intuition says us that, in order to get quickly convergent expansions centered at $\mathbf{R}_A$ and $\mathbf{R}_B$, the fragment $d_{ab}^A(\mathbf{r}_A)$ should contain the part of the charge distribution closer to A and $d_{ab}^B(\mathbf{r}_B)$, that in the neighborhood of B. The correctness of this intuitive envision and the way of implementing the method have been thoroughly discussed in previous works for both Gaussian [1] and Slater [2] basis sets. Here, we will illustrate the method by considering the case of a $1s(\mathbf{r}_A)\ 1s(\mathbf{r}_B)$ STO charge distribution which has the transform [2]:

$$1s(\mathbf{r}_A)1s(\mathbf{r}_B) \equiv N_{1s}N'_{1s}\, e^{-\zeta r_A}\, e^{-\zeta' r_B} = N_{1s}N'_{1s}\frac{\zeta\zeta'}{\pi}\int_0^1 \mathrm{d}u\,[u(1-u)]^{-1/2} \times \widehat{k}_1\left(\zeta_u\sqrt{R_{AB}^2+\frac{r_u^2}{u(1-u)}}\right) \quad (6)$$

with $\zeta_u^2 = \zeta^2 u + \zeta'^2(1-u)$, $\mathbf{r}_u = \mathbf{r} - \mathbf{R}_u$ and $\mathbf{R}_u = \mathbf{R}_A + u\,\mathbf{R}_{AB}$, and $\widehat{k}_\nu(z) = z^\nu K_\nu(z)$, where $K_\nu(z)$ is the Macdonald function—see Ref. [22] eqs 8.446, 8.468 and 8.532.1.

According to Eq. 6, the $1s(\mathbf{r}_A)\ 1s(\mathbf{r}_B)$ distribution is decomposed in a continuous set of infinitesimal spherical charge distributions centered along the points $\mathbf{R}_u$ of the AB segment. Since the partition criteria assign to A the charge distributions placed in the half segment closest to A (corresponding to $0 \le u \le 1/2$), it follows:

$$d^A_{1s1s}(\mathbf{r}_A) = N_{1s}N'_{1s}\frac{\zeta\zeta'}{\pi}\int_0^{1/2}\mathrm{d}u \times [u(1-u)]^{-1/2}\,\widehat{k}_1\left(\zeta_u\sqrt{R_{AB}^2+\frac{r_u^2}{u(1-u)}}\right) \quad (7)$$

and, the rest is $d^B_{1s1s}(\mathbf{r}_B)$. The partition of the $1s(\mathbf{r}_A)\ 1s(\mathbf{r}_B)$ distribution in the fragments $d^A_{1s1s}(\mathbf{r}_A)$ and $d^B_{1s1s}(\mathbf{r}_B)$ is exact, and the only pending steps are the expansion of the Macdonald functions in spherical harmonics centered at $\mathbf{R}_A$ and the solution of the integrals.

In ref [2], a solution to the first problem is given. An analytical solution for the integrals and the generalization to functions with high-order spherical harmonics were reported in reference [21]. The algorithm developed for STO in this way enables to attain the expansion appearing in Eq. 1 in an accurate and efficient manner.

Every radial factor, initially obtained as a numerical array, is further fitted to a piecewise function consisting of exponential times polynomials:

$$f^A_{\mathrm{lm}}(r_A) = e^{-\xi_k r_A}\sum_{p=0}^{n_k} t_k^p\, c_p^k(A,l,m), \quad \lambda_{k-1} \le r_A < \lambda_k \quad \text{for } k = 1 \text{ to } N, \quad \lambda_N \le r_A < \infty \quad (8)$$

with

$$t = \frac{2r - \lambda_k - \lambda_{k-1}}{\lambda_k - \lambda_{k-1}} \quad (9)$$

for the first N intervals, and for the last one:

$$t = \xi_{N+1}\,(r - \lambda_N) \quad (10)$$

The number of intervals $(N+1)$, their boundaries (λ_k) and the exponents (ξ_k) are characteristic of each atom and have been object of a thorough study. The polynomial coefficients are obtained by Lagrange interpolation to the previously tabulated values corresponding to the roots of tenth-degree Chebyshev polynomials on t, for the first N intervals, and to the abscissae of a 10-points Laguerre quadrature rule for the last interval. This fit gives the radial factors with an accuracy about ten decimal places thus ensuring that this accuracy can be attained in the expansion by simply increasing the value of $l_{\max}$.

3 The Coulomb operator

The partition of Eq. 1 allows us to write the Coulomb operator, $J(\mathbf{r})$, as a sum of atomic contributions:

$$J(\mathbf{r}) = \int \mathrm{d}\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = \sum_{A=1}^{N}\int \mathrm{d}\mathbf{r}'\,\frac{\rho^A(\mathbf{r}'_A)}{|\mathbf{r}-\mathbf{r}'|} = \sum_{A=1}^{N} J^A(\mathbf{r}_A) \quad (11)$$

which, taking into account the expansion of Eq. 1 and the piecewise fit of Eq. 8 are given by:

$$J^A(\mathbf{r}_A) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} z_l^m(\mathbf{r}_A)\left[\frac{Q_{\mathrm{lm}}(r_A)}{r_A^{2l+1}} + q^A_{lm}(r_A)\right] \equiv \sum_{l=0}^{\infty}\sum_{m=-l}^{l} z_l^m(\mathbf{r}_A)\, V^A_{\mathrm{lm}}(r_A) \quad (12)$$

where

$$Q_{lm}(r_A) = \frac{4\pi}{2l+1}\int_0^{r_A}\mathrm{d}r'\,r'^{2l+2} f^A_{lm}(r') = \frac{4\pi}{2l+1}\left[\int_0^{\lambda_{k-1}}\mathrm{d}r'\,r'^{2l+2} f^A_{lm}(r') + \int_{\lambda_{k-1}}^{r_A} dr'\,r'^{2l+2}\,e^{-\xi_k r'}\sum_{p=0}^{n_k} c_p^k\, t_k^p\right] \quad (13)$$

and

$$q_{lm}(r_A) = \frac{4\pi}{2l+1}\int_{r_A}^{\infty}\mathrm{d}r'\,r'\,f^A_{lm}(r') = \frac{4\pi}{2l+1}\left[\int_{r_A}^{\lambda_k}\mathrm{d}r'\,r'\,e^{-\xi_k r'}\sum_{p=0}^{n_k} c_p^k(A,l,m)\,t_k^p + \int_{\lambda_k}^{\infty}\mathrm{d}r'\,r'\,f^A_{lm}(r')\right] \quad (14)$$

with $\lambda_{k-1} \le r_A < \lambda_k$.

Notice that, for sufficiently large r_A, Eq. 12 is reduced to the long-range limit:

$$J^A_{\text{long}}(\mathbf{r}_A) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \frac{z_l^m(\mathbf{r}_A)}{r_A^{2l+1}} \frac{4\pi}{2l+1} \int_0^{\infty} \mathrm{d}r'\, r'^{2l+2} f^A_{lm}(r') = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \frac{z_l^m(\mathbf{r}_A)}{r_A^{2l+1}} Q^A_{\text{lm}} \tag{15}$$

where the constants Q^A_{lm} are the atomic charges and multipoles.

The first term in the r.h.s. of Eq. 13, the second of Eq. 14, and the multipolar moments of Eq. 15 are constants that can be computed once and stored. The pending integrals are easily obtained from the piecewise representation so that the Coulomb potential for arbitrary $\mathbf{r}_A$ can be computed in an accurate and economical way [23].

4 The Coulomb matrix and the Coulomb energy

We will consider now the matrix elements of $J(\mathbf{r})$. These are either of one-center, $J_{aa'}$, or two-center, J_{ab}, types:

$$J_{aa'} = \int \mathrm{d}\mathbf{r}\, J(\mathbf{r})\, \chi_a(\mathbf{r}_A)\, \chi_{a'}(\mathbf{r}_A) \tag{16}$$

$$J_{ab} = \int \mathrm{d}\mathbf{r}\, J(\mathbf{r})\, \chi_a(\mathbf{r}_A)\, \chi_b(\mathbf{r}_B) \tag{17}$$

As noted above, the $\chi_a(\mathbf{r}_A)\ \chi_{a'}(\mathbf{r}_A)$ distributions can be trivially expanded in spherical harmonics times radial factors so that:

$$J_{aa'} = \sum_{\lambda=0}^{L_a+L_{a'}} \sum_{\mu=-\lambda}^{\lambda} \int \mathrm{d}\mathbf{r}\, J(\mathbf{r})\, z_\lambda^\mu(\mathbf{r}_A) f^{aa'}_{\lambda\mu}(r_A) \tag{18}$$

Moreover, the partition of the two-center charge distributions gives:

$$J_{ab} = \int \mathrm{d}\mathbf{r}\, J(\mathbf{r})\, d^A_{ab}(\mathbf{r}_A) + \int \mathrm{d}\mathbf{r}\, J(\mathbf{r})\, d^B_{ab}(\mathbf{r}_B) \equiv J'_{ab} + J''_{ab} \tag{19}$$

and expanding the fragment $d^A_{ab}(\mathbf{r}_A)$, one has:

$$J'_{ab} = \sum_{\lambda=0}^{\infty} \sum_{\mu=-\lambda}^{\lambda} \sum_{\mu=-\lambda}^{\lambda} \int \mathrm{d}\mathbf{r}\, J(\mathbf{r}) f^{ab}_{\lambda\mu}(r_A) \tag{20}$$

with a similar expression for J''_{ab}.

We will now expand $J(\mathbf{r})$ in spherical harmonics centered at $\mathbf{R}_A$, distinguishing the contribution of atom A, its near neighbors, and the atoms placed far away from A, with just the single test on the distances commented in the appendix, so that:

$$J(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} z_l^m(\mathbf{r}_A) \times \left[V^A_{lm}(r_A) + \sum_{B}^{\text{neighbor}\,A} V^B_{lm}(r_A) + \sum_{C}^{far\,A} V^C_{lm}(r_A) \right] \tag{21}$$

The fast and accurate calculation of $V^A_{lm}(r_A)$ was already treated in Eqs. 12–14. For the sake of simplicity, in the trial program the contribution of the near neighbors is obtained by numerical integration:

$$V'^A_{lm}(r_A) \equiv \sum_{B}^{\text{neighbor}\,A} V^B_{lm}(r_A) = \frac{(2l+1)\,(-1)^m}{2\pi\,(1+\delta_{m0})} \frac{(l-|m|)!}{(l+|m|)!} \frac{1}{r_A^l} \times \int_0^{2\pi} d\phi_A \int_0^{\pi} \sin\theta_A\, \mathrm{d}\theta_A\, z_l^m(\mathbf{r}_A/r_A) \sum_{B}^{\text{neighbor}\,A} J^B(\mathbf{r}) \tag{22}$$

with $J^B(\mathbf{r})$ as in Eq. 12.

Notice, however, that this numerical integration can be replaced in refined versions by direct expansions. These expansions are specially simple in the third term, since if $R_{AC} = |\,\mathbf{R}_C - \mathbf{R}_A|$ is sufficiently large, Eq. 15 says that this term is a simple linear combination of irregular harmonics centered at atoms placed far away from A. Thus, the expansion only requires the translation of these harmonics to center A [24] giving:

$$\sum_{C}^{\text{far}\,A} V^C_{lm}(r_A) = (2-\delta_{m0}) \frac{(l-|m|)!}{(l+|m|)!} B^A_{lm} \tag{23}$$

where the general expressions of the constants B^A_{lm} are reported in the "Appendix".

In this way, we rewrite Eq. 21 as:

$$J(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} z_l^m(\mathbf{r}_A) \left[V_{lm}(r_A) + (2-\delta_{m0}) \frac{(l-|m|)!}{(l+|m|)!} B^A_{lm} \right] \tag{24}$$

with $V_{lm}(r_A) = V^A_{lm}(r_A) + V'^A_{lm}(r_A)$.

Once the expansion of Eq. 21 is known, the one-center matrix elements $J_{aa'}$ of Eq. 16 can be calculated by:

$$J_{aa'} = N_a N_{a'} \sum_{l=0}^{L_a+L_{a'}} \sum_{m=-l}^{l} \left[\frac{2\pi\,(1+\delta_{m0})}{2l+1} \frac{(l+|m|)!}{(l-|m|)!} \times \int_0^{\infty} \mathrm{d}r\, r^{2l+2}\, V_{lm}(r) f^{aa'}_{lm}(r) + Q^{aa'}_{lm}\, B^A_{lm} \right] \tag{25}$$

where $Q^{aa'}_{lm}$ are the multipolar moments of the unnormalized distribution $\chi_a(\mathbf{r}_A)\ \chi_{a'}(\mathbf{r}_A)$, and N_a, N'_a are the normalization factors.

Similarly, the contributions J_{ab}' of $d^A_{ab}(\mathbf{r}_A)$ to the two-center elements are given by:

$$J'_{ab} = N_a N_b \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left[\frac{2\pi\,(1+\delta_{m0})}{2l+1} \frac{(l+|m|)!}{(l-|m|)!} \times \int_0^{\infty} dr\, r^{2l+2}\, V_{lm}(r) f'^{ab}_{lm}(r) + Q^{A,ab}_{lm}\, B^A_{lm} \right] \tag{26}$$

$Q_{lm}^{A,ab}$ being the multipolar moments of the fragments $d_{ab}^{A}(\mathbf{r}_A)$.

In the current approach, the Coulomb energy:

$$E_c = \frac{1}{2}\int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\,\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{27}$$

can be computed either from the density, $\boldsymbol{\rho}$, and Coulomb, $\mathbf{J}$, matrices:

$$E_c = \frac{1}{2}\mathrm{Tr}(\boldsymbol{\rho}\,\mathbf{J}) \tag{28}$$

or from the atomic densities, $\rho^A(\mathbf{r}_A)$, and the Coulomb potential, $J(\mathbf{r}_A)$:

$$E_c = \frac{1}{2}\sum_{A=1}^{N}\int \mathrm{d}\mathbf{r}\,\rho^A(\mathbf{r}_A)\,J(\mathbf{r}_A) \tag{29}$$

Equation 28 is the standard procedure in methods based on integrals but it is not recommended in the present context because Eq. 29 is more efficient. This equation gives the Coulomb energy as a sum of atomic contributions which, bearing in mind Eqs. 1 and 24, are:

$$E_c = \sum_{A=1}^{N} E_c^A = \sum_{A=1}^{N}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\frac{2\pi\,(1+\delta_{m0})}{2l+1}\,\frac{(l+|m|)!}{(l-|m|)!} \times \int_0^{\infty} \mathrm{d}r_A\, r_A^{2l+2}\, V_{lm}(r_A) f_{lm}^A(r_A) + Q_{lm}^A\, B_{lm}^A\right] \tag{30}$$

Since $V_{lm}(r_A)$ and B_{lm}^A are those of Eqs. 25 or 26, and $f_{lm}^A(r_A), Q_{lm}^A$ (the atomic radial factor and multipoles) are known, Eq. 30 can be solved together with Eqs. 25 and 26 making the calculations of E_c very inexpensive.

5 Algorithm performance

In order to test the reliability and efficiency of this procedure, we have elaborated a trial program which proceeds in three steps. In the first one, both the partition of the density and the expansion of the fragments—see Eq. 1—are carried out following the scheme outlined in Sect. 2. Most algorithms required in this step have been taken from those previously developed and implemented in the DAM program [21].

In the second step, $J(\mathbf{r})$ is expanded at each nucleus in spherical harmonics times radial factors, according to Eq. 21 with the aif of Eqs. 12, 22 and 23. The values of $V_{lm}^A(r_A)$ as well as those of $J^B(\mathbf{r})$, at the abscissae of the Lebedev-Markov [25, 26, 27, 28] quadrature rules, are obtained by an efficient algorithm based on Eqs. 11–17, which was previously elaborated for the DAMPOT program [23]. The calculation of the B_{lm}^A constants, Eq. 45, is very inexpensive because the Q_{lm} multipolar moments are previously evaluated and stored in step 1, and the irregular harmonics are obtained by fast and stable recurrence relations.

Finally, once the radial factors of Eq. 24 are known for a given center, the one-center matrix elements, $J_{aa'}$, the contributions J'_{ab} to the two-center elements and to the Coulomb energy are evaluated in the third step by Eqs. 25 and 26.

Although, as stressed above, this procedure is potentially exact, the accuracy achieved in its practical implementation is determined by (1) the numerical integration of Eqs. 22 and 25 or 26, and (2) the truncation of the infinite expansions of Eqs. 1, 21 and 26.

The Lebedev-Markov angular quadrature rule of Eq. 22 was carried out following the prescription of [29], and its accuracy was determined from the convergence at several values of r_A for an increasing number of quadrature points. We observe that, for a fixed order in the quadrature rule, the radial factors corresponding to lower values of l are more accurate than those of higher l. Furthermore, for l fixed, the radial factors have a poor accuracy at values of r corresponding to the distances to other nuclei, even for very high-order quadratures. Because of this, the angular integration scheme introduces an error that limits the accuracy attainable, but we find that in the critical zones a quadrature with 1,454 points gives an error lower than that of the truncation for $l_{\max} \leq 10$, and this is sufficient for the purpose of the present study.

Moreover, in all the cases treated, the one-dimension integrals of Eqs. 25, 26, and 30 were computed with a similar accuracy using a 64 points Moebius-Legendre quadrature rule [30, 31] with the Moebius ε parameter fixed in a value of -0.9.

Once the error of the numerical integration is sufficiently low, the accuracy in the final result is determined by the truncation of the infinite expansions of Eqs. 1, 21, and 26. For the sake of simplicity, all these expansions are truncated to the same value of l, (i.e. $l_{\max} \leq 10$).

To test the dependence of the accuracy on $l_{\max}$, conventional Hartree-Fock calculations have been made for series of molecules with increasing sizes. For each molecule, the SCF process was stopped at a given cycle, and the density matrix $\boldsymbol{\rho}$, the Coulomb matrix $\mathbf{J}^{\mathrm{int}}$, and the Coulomb energy E_c^{int} were written to a file.

Next, the $\boldsymbol{\rho}$ matrix was read by the trial program which carried out the direct computation of the Coulomb matrix, $\mathbf{J}^{\mathrm{pot}}$, and Coulomb energy, E_c^{pot}, with the algorithm reported herein.

The accuracy of the results was characterized by means of two quantities: the highest difference between the elements of the two Coulomb matrices:

$$\Delta(l_{\max}) = \max_{(\mathrm{r,s})} \left|J_{rs}^{\mathrm{pot}}(l_{\max}) - J_{rs}^{\mathrm{int}}\right| / E_H \tag{31}$$

and the difference between the Coulomb energies:

$$\epsilon(l_{\max}) = |E_c^{\text{pot}}(l_{\max}) - E_c^{\text{int}}| / E_H \quad (32)$$

The observed trends are illustrated in Fig. 1 where $\log_{10} \Delta(l_{\max})$ corresponding to a sample of 13 molecules of increasing size (C_nH_{2n+2}, n from 4 to 16) is plotted against m for l_{max} = 6, 7, 8, 9 and 10. For the highest value of $l_{\max}$ available ($l_{\max}$ = 10) a maximum error of 0.1mE_H is attained in the elements of the Coulomb matrix. The number of accurate places seems to increase in roughly one further correct figure every three additional terms in the expansion.

The errors in the Coulomb energies, $\varepsilon(l_{\max})$, are reported in Fig. 2a. In order to facilitate comparison, it is a common practice—see Sect. 6—to normalize this error in terms of error per atom. Fig 2b shows this normalized error as a function of the number of atoms in this series of molecules. As it can be seen, the error ranges from roughly 50 μE_H per atom for $l_{\max}$ = 6 to 0.5 μE_H per atom for $l_{\max}$ = 10.

Table 1 shows that the computational cost increases with the length of the expansion ($l_{\max}$) in an approximately linear way. Noticeably, these conclusions seem to be independent of the size of the system.

Since efficiency is related with the ratio accuracy/cost, we have examined also the dependence of the cost of the three steps with the size of the system. In the first step, the charge distributions $D_{ab}(\mathbf{r}) \equiv \chi_a(\mathbf{r}_A)\ \chi_b(\mathbf{r}_B)$ are partitioned, expanded and accumulated, weighted with the density, at each of their centers. The computational cost should thus increase as the number of nonnegligible charge distributions, i.e. it should have a formal m^2 dependence but, due to the basis product cutoff—see Sect. 6—it must tend to linear scaling in large systems.

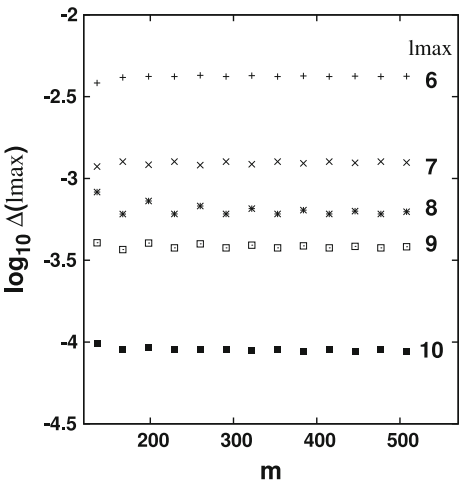

Fig. 1 Maximum error ($\Delta(l\text{max})$) in the Coulomb matrix for the n–alkane series (n: 4,…,16) as a function of the number of basis functions, m, for several values of $l_{\max}$

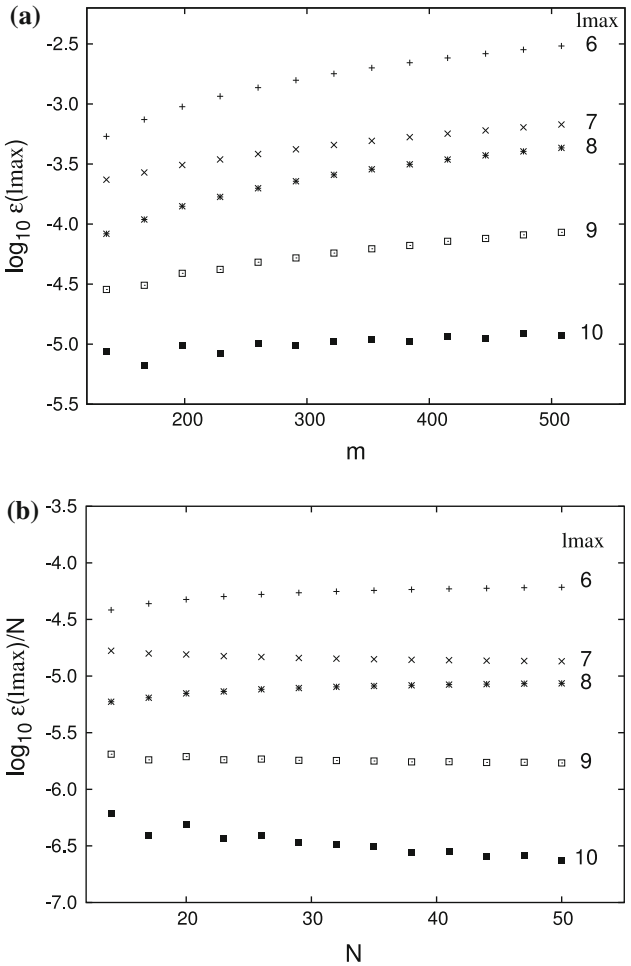

Fig. 2 Error in the Coulomb energy for the n–alkane series (n: 4,…,16): **a** total error in terms of the number of basis functions, m; **b** error per atom as a function of the number of atoms, N

Table 1 Computational time in seconds for the Coulomb matrix in n-hexane as a function of the expansion length

Basis set (functions)	VB1(198)				
$l_{\max}$	6	7	8	9	10
Step 1	16.8	17.7	18.7	20.1	21.3
Step 2	48.4	60.7	76.1	92.9	103.5
Step 3	9.6	10.4	11.4	12.5	14.0
Basis set (functions)	VB2(420)				
$l_{\max}$	6	7	8	9	10
Step 1	56.4	61.0	65.0	69.5	74.9
Step 2	50.4	64.4	82.4	101.7	121.4
Step 3	31.4	34.5	37.8	42.7	44.8

In the third step, the charge distributions are partitioned, expanded, and integrated with the Coulomb potential, so that the computational cost should follow the same trends as in step one.

However, in the second step, the Coulomb potential is expanded at each center by using the atomic potentials of all the centers and, therefore, it scales formally as the square of the number of atoms (N^2). Nevertheless, as stressed in Eq. 21, we separate the contributions of the atom placed at the expansion center and its neighbors from the contributions of the atoms placed far away. Since for large systems the number of neighbors is independent of the size, the first contribution scales linearly with the number of atoms, and the quadratic behavior must be entirely due to the second contribution, which can be computed from the long-range formula in an extremely fast process. In summary, for large systems, step 2 linearly scales with the number of atoms if the cost of the long-range contribution is smaller than that of the first contribution and quadratically otherwise.

All these points have been carefully tested, and some representative results are shown in Tables 2 and 3. In the first rows of Table 2, the computational cost of the three steps is compared with the cost of the STO integrals for the series of *n*-alkanes with $n = 1$–16 (from 5 to 50 atoms). Calculations were carried out with the VB1 (C[5:3:1], H[3:1]) STO basis set [32] at standard geometries. In the bottom of the table, the computational cost with the VB2 basis set (C[6:4:2:1], H[4:2:1]) is reported for $n = 1$–7. Integrals were computed with the SMILES package [33] which can handle up to about 500 STO at an affordable computational cost in these systems with high local symmetry.

Notice that the computational cost of the current procedure is several orders of magnitude lower than that of the integrals and savings are greater for high-quality basis sets.

Table 3 contains the computational time of the three steps for the *n*-alkanes with $n = 17$–33 (53–101 atoms). The cost of the long-range contributions included in step 2 is given between parenthesis. As it can be seen, the cost of this quadratically dependent part is extremely small so that the whole cost in these systems tends to increase linearly with the number of basis functions or the number of atoms. Extrapolation on these results shows that the cost of the quadratic contribution will be comparable to the linear one for several thousands of atoms. Therefore, one can be confident that the scaling will remain linear for the larger affordable systems (hundreds of atoms).

As a practical example, we plot in Fig. 3 the whole computational cost ($step1 + step2 + step3$) for the alkanes in Table 3. The fit of these data to a second order polynomial gives a coefficient for the linear term of 15.3 ± 1.2, whereas that of the quadratic term is -0.005 ± 0.008. The linear behavior is evident.

We have also analyzed the behavior of the method for compact systems like graphenes, helicenes, and other systems alike, finding that it is satisfactory also in these cases.

Table 2 Cost in seconds of STO integrals and the three steps involved in the direct computation of the Coulomb matrix ($l_{max} = 10$)

Molecule	Size[a]	Integrals[b]	Step 1	Step 2	Step 3
VB1 basis set[c]					
CH_4	43	39	0.9	5.4	0.5
C_2H_6	74	279	2.9	15.9	1.6
C_3H_8	105	1,534	6.0	31.4	3.2
C_4H_{10}	136	3,150	10.2	54.9	5.6
C_5H_{12}	167	5,233	15.2	77.6	9.8
C_6H_{14}	198	7,469	21.3	103.5	14.0
C_7H_{16}	229	10,117	27.9	135.6	18.1
C_8H_{18}	260	12,298	34.9	173.2	23.0
C_9H_{20}	291	14,638	42.3	191.3	27.7
$C_{10}H_{22}$	322	16,968	50.2	233.4	32.9
$C_{11}H_{24}$	353	19,400	57.8	255.2	37.7
$C_{12}H_{26}$	384	21,980	65.4	292.6	42.7
$C_{13}H_{28}$	415	24,854	72.6	315.5	47.4
$C_{14}H_{30}$	446	28,056	80.0	353.0	52.5
$C_{15}H_{32}$	477	31,653	88.1	375.4	57.2
$C_{16}H_{34}$	508	36,003	95.1	412.7	62.5
VB2 basis set[c]					
CH_4	95	392	3.3	5.6	2.0
C_2H_6	160	2,952	10.4	16.2	6.1
C_3H_8	225	16,502	21.2	32.4	12.6
C_4H_{10}	290	35,742	35.2	56.9	20.9
C_5H_{12}	355	59,505	53.3	83.6	31.9
C_6H_{14}	420	87,478	74.9	121.4	44.8
C_7H_{16}	485	118,483	99.3	156.5	71.8

[a] Basis set size

[b] Time for computing integrals

[c] See Ref. [32] for basis sets

As an example, we report in the last rows of Table 3 the results for three graphene-like molecules ($C_{6n^2}H_{6n}$): coronene ($n = 2$), circumcoronene ($n = 3$), and circumcircumcoronene ($n = 4$).

All the times reported in the tables have been measured on a computer with a Pentium IV processor at 3GHz.

6 Comparison with other alternatives

We will compare first the practical possibilities of the current method with those based on integrals. As it was stressed in the introduction, the formal scaling of the conventional approach based on integrals (m^4) is not realized in practice because product cutoff techniques reduce this scaling to m^2 in the limit of very large systems.

To our knowledge, the importance of the cutoff was first recognized by Monkhorst and Harris [7] in a study of the expensive integrals with STO, but it was soon incorporated

Table 3 Time in seconds for the three steps involved in the Coulomb matrix computation ($l_{max} = 10$) VB1 basis set

Molecule	Size[a]	Step 1	Step 2	Step 3
$C_{17}H_{36}$	539	102.3	459.9 (0.39)	57.9
$C_{18}H_{38}$	570	109.4	485.6 (0.46)	61.9
$C_{19}H_{40}$	601	116.4	525.8 (0.53)	67.6
$C_{20}H_{42}$	632	124.1	558.0 (0.60)	70.1
$C_{21}H_{44}$	663	132.3	577.9 (0.68)	75.4
$C_{22}H_{46}$	694	142.0	613.5 (0.77)	79.1
$C_{23}H_{48}$	725	147.5	647.3 (0.86)	83.1
$C_{24}H_{50}$	756	155.6	674.3 (0.96)	87.4
$C_{25}H_{52}$	787	161.8	710.1 (1.05)	91.6
$C_{26}H_{54}$	818	169.5	746.3 (1.16)	96.4
$C_{27}H_{56}$	849	177.9	764.0 (1.27)	100.4
$C_{28}H_{58}$	880	186.8	807.2 (1.39)	104.9
$C_{29}H_{60}$	911	192.9	827.4 (1.51)	109.7
$C_{30}H_{62}$	942	204.7	860.8 (1.64)	113.2
$C_{31}H_{64}$	973	208.3	880.5 (1.77)	117.4
$C_{32}H_{66}$	1,004	215.4	930.0 (1.91)	124.6
$C_{33}H_{68}$	1035	223.7	945.9 (2.04)	126.6
$C_{24}H_{12}$	528	123.7	343.2 (0.06)	72.5
$C_{54}H_{18}$	1,134	432.4	1042.4 (0.60)	260.5
$C_{96}H_{24}$	1,968	958.8	2156.2 (2.36)	566.1

[a] Basis set size

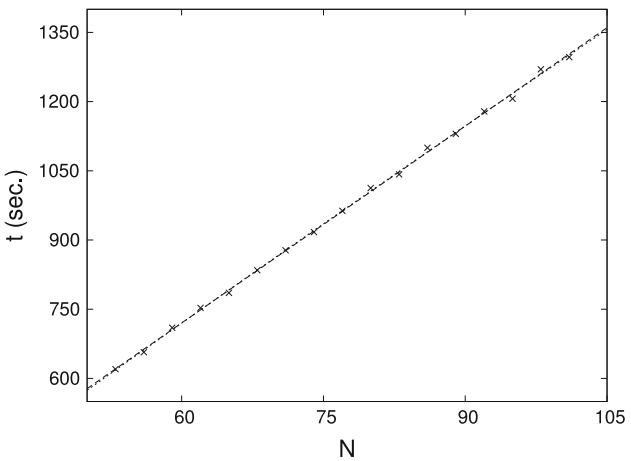

Fig. 3 Computational cost for the *n*–alkane series (*n*: 17,…,33) as a function of the number of atoms, *N*

to the first programs with Gaussian functions by Clementi [8], and improved versions have been implemented since then in every integral package, with Gaussians and STO. The foundation of this technique is simple: if a charge distribution, $D_{ab}(\mathbf{r}) = \chi_a(\mathbf{r}_A)\ \chi_B(\mathbf{r}_B)$, is sufficiently small all over space, the set of m^2 integrals containing this distribution can be neglected. Moreover, since $D_{ab}(\mathbf{r})$ tends to zero as the distance $R_{AB} = |\mathbf{R}_B - \mathbf{R}_A|$ goes to infinity, in a very extended system only charge distributions with R_{AB} smaller than a certain distance must be considered. This is the reason for the m^2 scaling of the integrals. Furthermore, in large systems, many pairs of distributions $D_{ab}(\mathbf{r})$ and $D_{cd}(\mathbf{r})$ are separated by distances sufficiently large to enable the use of long-range methods, thus rendering additional savings [34, 35, 36] in computational cost.

Nonetheless, it must be noticed that the effect of both cutoff and long-range separation depends not only on the size of the system but also on the type and quality of the basis set. In fact, these effects must be more important in GTO than in STO basis sets (due to the faster decay of the former) and, for both types of functions, they are more significant for low-quality than for high-quality basis sets, because the latter include more functions with high quantum numbers, which have nonnegligible values over larger distances from nucleus. In summary, one can expect important savings (and a quick reach of the scaling limits) in the calculation of extended systems with poor GTO basis sets, but for good STO basis sets these theoretical limits are well beyond the size of the systems that currently can be studied with integrals. Nevertheless, there are other possibilities for saving.

We notice first that good atomic basis sets give, in compact systems, molecular basis sets that tend to redundancy, and this near redundancy is even clearer in the much larger set of charge distributions, $D_{ab}(\mathbf{r})$. These distributions span a subspace $\mathscr{V}_D$ of $\mathscr{L}^2(\mathbb{R}^3)$ and the introduction of a smaller nonredundant basis of this subspace can lead to an important saving. Notice also that, in large systems, most charge distributions are two-center and this makes it difficult to represent the repulsion operator, being particularly expensive for STO. Clearly, further savings can be attained if this small-size basis set consists of one-center functions.

An almost optimal method for detection and elimination of near redundancy was reported by Beebe and Linderberg [37] in the seventies. In this method, it is implicit the use of a basis set of $\mathscr{V}_D$ obtained from Gram-Schmidt orthogonalization of the charge distributions set, with a metric determined by the repulsion operator. Since the coefficients matrix of this orthogonal basis is the inverse of the triangular matrix resulting from Cholesky decomposition of the repulsion integrals matrix, it is called [19] Cholesky basis. It gives, for a given accuracy threshold, the optimal approximation to every repulsion integral, being unbiased toward any particular combination of integrals required in different methods. Although this basis has a limited computational efficiency, because it contains two-center distributions, Aquilante, Lindh, and Pedersen have recently demonstrated [19, 20] that one can take only one-center terms without serious downgrading of accuracy.

In this regard, it is important to stress that electron density–see Eq. 1—is nothing but a particular vector of $\mathscr{V}_D$ which can be, in principle, accurately represented with small basis sets biased toward this objective though, as demonstrated in the above mentioned work[19], not necessarily good for the representation of other vectors of $\mathscr{V}_D$. This representation is the underlying idea of the so-called density fitting methods, which can be traced back to the seventies, its first practical application being, to our knowledge, that of Baerends, Ellis and Roos [9] in the context of the Hartree-Fock-Slater (or X_α) method, and in the ADF code, both using STO basis sets. This technique, which has undergone a continued development since then [10, 11, 12, 13, 14, 15, 16, 17, 18] and is currently implemented in many popular computational programs, consists of the fitting of the LCAO expression of $\rho(\mathbf{r})$—see Eq. 3—, to one-center functions $f_k(\mathbf{r})$ of an auxiliary basis set, to give the approximate density $\widetilde{\rho}(\mathbf{r})$:

$$\widetilde{\rho}(\mathbf{r}) = \sum_{k=1}^{M} \mathrm{d}_k f_k(\mathbf{r}) \tag{33}$$

where M stands for the size of the whole auxiliary basis set. The coefficients d_k are obtained by minimizing some suitable functional of $\rho(\mathbf{r}) - \widetilde{\rho}(\mathbf{r})$, such as the square square difference [9]:

$$\Delta^2 = \int \mathrm{d}\mathbf{r}\,[\rho(\mathbf{r}) - \widetilde{\rho}(\mathbf{r})]\,[\rho(\mathbf{r}) - \widetilde{\rho}(\mathbf{r})] \tag{34}$$

or the Coulomb energy of the density difference [15, 16, 17, 18]:

$$\Delta'^2 = \int \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,[\rho(\mathbf{r}) - \widetilde{\rho}(\mathbf{r})]\frac{1}{|\mathbf{r}-\mathbf{r}'|}[\rho(\mathbf{r}') - \widetilde{\rho}(\mathbf{r}')] \tag{35}$$

This latter procedure, which is equivalent to taking a Coulomb metric of $\mathscr{V}_D$, is currently preferred because it gives higher accuracy [13].

As an original and efficient alternative, Manby and Knowles [38, 39, 40] introduce two auxiliary basis sets related by the Poisson equation. The Coulomb potential is represented in the first basis set, but the expansion coefficients are obtained by fitting the density with its associated Poisson basis set.

In practice, the elements of the approximate Coulomb matrix:

$$\widetilde{J}_{ab} = \int \mathrm{d}\mathbf{r}\,\chi_a(\mathbf{r}_A)\,\chi_b(\mathbf{r}_B)\int \mathrm{d}\mathbf{r}'\sum_{k=1}^{M}\mathrm{d}_k\frac{f_k(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{36}$$

can be obtained in three steps. First, one defines:

$$a_l = \sum_{c=1}^{m}\sum_{d=1}^{m}[f_l|\chi_c\,\chi_d]\,\rho_{cd}, \quad l = 1,\ldots,M \tag{37}$$

next:

$$\mathrm{d}_k = \sum_{l=1}^{M}(\mathbf{M}^{-1})_{kl}\,a_l \quad k = 1,\ldots,M \tag{38}$$

and finally

$$\widetilde{J}_{ab} = \sum_{k=1}^{M}[\chi_a\,\chi_b|f_k]\,d_k \quad a = 1,\ldots,m;\ b = 1,\ldots,m \tag{39}$$

where $M_{kl} = [\ f_k \mid f_l\ ]$, and the bracket denotes the scalar product with the Coulomb metric, i.e the corresponding two-electron repulsion integral.

Clearly, steps 1—Eq. 37—and 3—Eq. 39—are $M \times m^2$ (or m^3) dependent and, since M is proportional to m—typically around $3m$—this representation implies a remarkable saving of computational cost with respect to the methods based on integrals (m^4 dependent), a saving still greater if the two-center charge distributions, $\chi_a(\mathbf{r}_A)$ $\chi_b(\mathbf{r}_B)$, are approximated with only auxiliary functions centered at $\mathbf{R}_A$ and $\mathbf{R}_B$:

$$\chi_a(\mathbf{r}_A)\,\chi_b(\mathbf{r}_B) \approx \sum_{p=1}^{M_A} f_k^A(\mathbf{r}_A)\,a_k^{ab} + \sum_{q=1}^{M_B} f_k^B(\mathbf{r}_B)\,b_k^{ab} \tag{40}$$

Since the auxiliary functions are products of spherical harmonics times radial factors centered at the nuclei, the terms in Eq. 33 associated with each nucleus can be separated, and this equation rewritten as:

$$\begin{aligned}\widetilde{\rho}(\mathbf{r}) &= \sum_A\sum_{k=1}^{M_A}\mathrm{d}_k f_k^A(\mathbf{r}_A) = \sum_A \widetilde{\rho}^A(\mathbf{r}_A)\\ &= \sum_A\sum_{l=0}^{L_A}\sum_{m=-l}^{l} z_l^m(\mathbf{r}_A)\,\widetilde{f}_{lm}^A(r_A)\end{aligned} \tag{41}$$

where M_A is the size of the auxiliary basis subset centered at $\mathbf{R}_A$, and L_A is the greatest l in these functions.

The formal similarity between Eqs. 41 and 1 as well as that between Eq. 6 and 4 is clear, but this similarity hides deep differences. Equations 1 and 16 or 17 imply a formally m^2 process and tend, in practice, to linear scaling (m^1) for increasingly large systems, whereas Eqs. 37 and 39 imply a formally $M \times m^2$ process which tends to slightly less than quadratic scaling ($m^{1.7}$ according to refs [41, 42]).

The origin of this difference is clear. In the present method, the cost in the partition/expansion of every charge distribution, $k(l_{\max})$, is independent of the system size, so that the formal dependence is $k(l_{\max}) \times m^2$ with $k(l_{\max})$ constant. However, in DF—or RI—methods the formal dependence is $K\ M\ m^2$ and, when the size increases, not only m^2 but also M increases, since new functions are added to the auxiliary basis set.

Moreover, Eqs. 1 and 16 or 17 are exact and allow us the systematic improvement of accuracy by simply adding more terms to the expansions. On the contrary, Eqs. 37 and 39 are approximate, the development of good auxiliary basis sets is not so simple and their systematic improvement seems highly difficult.

Auxiliary basis sets optimized for specific RI calculations have been developed by the Karlsruhe group [16, 43, 42], and we take as an example the RI–J1 and RI–J2 optimized for reproducing the Coulomb energy. The RI–J1 was reported in 1997 and it gives typical errors around 100 μE_H per atom when used together with some particular orbital basis sets [16]. The RI–J2 was reported nine years later yielding typical errors around 30 μE_H per atom and being more general. Higher accuracy would require a third generation of auxiliary basis sets.

According to Fig. 2b, similar accuracies are obtained with the current method by taking $l_{\max}$ around 6 or 7, and higher accuracy only requires to increase the value of $l_{\max}$. Notice that the errors per atom for $l_{\max} = 10$ are orders of magnitude lower than the typical errors of RI–J2.

Finally, we stress that DF—or RI—method still requires the calculation and storage of $M \times m^2$ repulsion integrals $[f_k \mid \chi_a\chi_b]$ as well as the calculation of the $M \times M$ matrix with elements $M_{kl} = [f_l \mid f_k]$ and its inverse. This is avoided in our approach.

7 Conclusions

A new method for computing the Coulomb matrix has been reported. Its reliability has been proved, and the main features have been tested with a trial program for Slater functions.

The accuracy of both the Coulomb energy and the Coulomb matrix has been tested by comparing the results of the current procedure with those obtained from integrals.

In practice, the differences depend on the number of terms in the expansion of the density, potential, and charge distributions in spherical harmonics. In this respect, we conclude that: (1) accurate results are attained with moderate expansions; (2) increasing the accuracy only requires to take more terms in these expansions, and this is not very time-consuming; and (3) it seems that the normalized accuracy (error per atom) does not increase with the size of the system.

We have also tested the computational cost of the method. In this regard, we confirm that it is much less expensive than methods based on integrals for systems of moderate size. Furthermore, for large systems the m^2 formal dependence is overcome in practice, because of the basis set product cutoff, and the separation of the long-range contributions makes it quickly tend to linear scaling with both the number of basis functions and the number of atoms.

Two further practical advantages of the current approach must be stressed. Its storage requirements are extremely reduced since two-electron integrals are completely suppressed, and it is well suited for DFT because the density, $\rho(\mathbf{r})$, its gradient, $\nabla\rho(\mathbf{r})$, and its laplacian, $\nabla^2\,\rho(\mathbf{r})$, can be efficiently computed from the expansion of Eq. 1. The conceptual advantages are also remarkable, since in this approach the physical approximations and mathematical simplifications are smoothly related.

We finally notice that the computational times with the trial program must be regarded as merely indicative because the development of a fully optimized code has not been a goal at this stage. We are currently working on an enhanced version of this procedure.

Appendix

Let $\rho^A(\mathbf{r}_A)$ and $\rho^C(\mathbf{r}_C)$ be two charge distributions fully enclosed by two spheres centered at $\mathbf{R}_A$ and $\mathbf{R}_C$ and with radii R^A and R^B, and let $R_{AC} \equiv \mid \mathbf{R}_C - \mathbf{R}_A \mid \geq R^A + R^C$. Its electrostatic interaction energy, E^{AC}, can be obtained by combining the bipolar expansion of $|\mathbf{r} - \mathbf{r}'|^{-1}$:

$$\begin{aligned}\frac{1}{|\mathbf{r}-\mathbf{r}'|} = &\sum_{l=0}^{\infty}\sum_{m=-l}^{l}(2-\delta_{m0})\frac{(l-|m|)!}{(l+|m|)!}z_l^m(\mathbf{r}_A)\\ &\times\sum_{L=0}^{\infty}\sum_{M=-L}^{L}(2-\delta_{M0})\frac{(L-|M|)!}{(L+|M|)!}\\ &\times z_L^M(\mathbf{r}'_C)\frac{\sqrt{\pi}(l+L-1/2)!}{(l-1/2)!(L-1/2)!}\sum_{m'}\alpha_{l+L,m'}^{lmLM}\frac{z_{l+L}^{m'}(\mathbf{R}_{AC})}{R_{AC}^{2l+2L+1}}\end{aligned} \tag{42}$$

with the expansions in spherical harmonics of $\rho^A(\mathbf{r}_A)$ and $\rho^C(\mathbf{r}_C)$. This gives:

$$\begin{aligned}E^{AC} = &\sum_{l=0}^{\infty}\sum_{m=-l}^{l}Q_{lm}^A\sum_{L=0}^{\infty}\sum_{M=-L}^{L}Q_{LM}^C\\ &\times\frac{\sqrt{\pi}\,(l+L-1/2)!}{(l-1/2)!\,(L-1/2)!}\sum_{m'}\alpha_{l+L,m'}^{lmLM}\frac{z_{l+L}^{m'}(\mathbf{R}_{AC})}{R_{AC}^{2l+2L+1}}\end{aligned} \tag{43}$$

where $\alpha_{l+L,\,m'}^{lm\;LM}$ are the coefficients for the decomposition of products of spherical harmonics into spherical harmonics. The sum on m', determined by $\alpha_{l+L,m'}^{lmLM}$, contains two terms at most, and Q_{lm}^A and Q_{LM}^C are the multipolar moments related to the radial factors by:

$$Q_{lm}^A = \frac{4\pi}{2l+1}\int_0^{R^A} \mathrm{d}r\, r^{2l+2} f_{lm}^A(r) \tag{44}$$

with a likewise expression for Q_{LM}^C.

In practice, the $\rho^A(\mathbf{r}_A)$ fragments extend to infinite but, since the radial factors have a quick decay (at least, exponential) one can estimate—for a fixed threshold—the size of the sphere that encloses the nonnegligible part of $\rho^A(\mathbf{r}_A)$. In a previous work [3], we showed that the $\rho^A(\mathbf{r}_A)$ given by DAM looks very much like the densities of the corresponding isolated atom, so that a conservative criterion would be to take for atoms in molecules the size of the spheres of the isolated atoms. This point has been tested and verified finding that a radius between 5 and 6 bohr should give sufficient accuracy but, for prudential reasons, we take 6.5 bohr in this trial program.

Once this radius is fixed, C is considered as a neighbor of A if their spheres intersect each other ($R_{AC} < 13$ bohr), and nonneighbor (or far away from A) otherwise. Notice that the set of atoms placed far away from A *generates a constant potential region where* $\rho^A(\mathbf{r}_A)$ *is nonnegligible*. This is the B^A_{lm} constant appearing in Eqs. 23–26 which, according to Eq. 43 reads:

$$B^A_{lm} = \sqrt{\pi} \sum_{L=0}^{L_{\max}} \frac{(-1)^L (L+l-1/2)!}{(L-1/2)!\,(l-1/2)!} \times \sum_{M=-L}^{L} \sum_{m'} c^{LM\,lm}_{L+lm'} \sum_{C} Q^C_{LM} \frac{z^{m'}_{L+l}(\mathbf{R}_{AC})}{R^{2L+2l+1}_{AC}} \tag{45}$$

where $L_{\max}$ is the order of the highest atomic multipolar moment considered.

References

1. Rico JF, López R, Ramírez G (1999) J Chem Phys 110(9):4213
2. Rico JF, López R, Ema I, Ramírez G (2002) J Chem Phys 117(2):533
3. Rico JF, López R, Ema I, Ramírez G (2002) J Chem Theory Comput 1:1083
4. Rico JF, López R, Ema I, Ramírez G (2007) Theor Chem Account 118:709
5. Rico JF, López R, Alonso JIF (1984) Phys Rev A 29:6
6. Rico JF, López R, Paniagua M, Alonso JIF (1984) Phys Rev A 29:11
7. Monkhorst HJ, Harris FE (1969) Chem Phys Lett 3(7):537
8. Clementi E (1972) Proc Nat Acad Sci 69(10):2942
9. Baerends EJ, Ellis DE, Ros P (1973) Chem Phys 2(1):41
10. Whitten JL (1973) J Chem Phys 58:4496
11. Dunlap BI, Connolly JWD, Sabin JR (1979) J Chem Phys 71:3396
12. Andzelm J, Wimmer E (1992) J Chem Phys 96(2):1280
13. Vahtras O, Almlöf J, Feyereisen MW (1993) Chem Phys Lett 213(5–6):514
14. Treutler O, Ahlrichs R (1995) J Chem Phys 102(1):346
15. Eichkorn K, Treutler O, Öhm H, Haser M, Ahlrichs R (1995) Chem Phys Lett 240(4):283
16. Eichkorn K, Weigend F, Treutler O, Ahlrichs R (1997) Theor Chem Account 97:119
17. Cohen AJ, Handy NC (2002) J Chem Phys 117(4):1470
18. Watson MA, Handy NC, Cohen AJ (2003) J Chem Phys 119(13):6475
19. Aquilante F, Lindh R, Pedersen TB (2007) J Chem Phys 127:114107
20. Pedersen TB, Aquilante F, Lindh R (2009) Theor Chem Acc 124:1
21. Rico JF, López R, Ema I, Ramírez G, Ludeña E (2004) J Comput Chem 25(11):1355
22. Gradshteyn IS, Ryzhik IM (1980) Table of integrals, series and products, 4th edn. Academic Press, New York
23. Rico JF, López R, Ema I, Ramírez G (2004) J Comput Chem 25(11):1347
24. Steinborn EO, Ruedenberg K (1973) Rotation and translation of regular and irregular solid spherical harmonics (Elsevier, 1973). Advances in Quantum Chemistry, vol 7, pp 1–81
25. Lebedev VI (1976) Dokl Akad Nauk 231(1):32
26. Lebedev VI, Skorokhodov AL (1992) Dokl Akad Nauk 324(3):519
27. Lebedev VI (1994) Dokl Akad Nauk 338(4):454
28. Lebedev VI, Laikov DN, (1999) Dokl Akad Nauk 366(6):741
29. Kakhiani K, Tsereteli K, Tsereteli P (2009) Comput Phys Commun 1806:256
30. Homeier HHH, Steinborn EO (1991) Int J Quantum Chem 39(4):625
31. Homeier HHH, Steinborn EO (1992) Int J Quantum Chem 41(4):399
32. Ema I, García de la Vega JM, Ramírez G, López R, Rico JF, Meissner J, J Paldus (2003) J Comput Chem 24(7):859
33. Rico JF, López R, Ema I, Ramírez G (2004) J Comput Chem 25(16):1987
34. Strain MC, Scuseria GE, Frisch MJ (1996) Science 271:51
35. White CA, Johnson BG, Gill PM, Head-Gordon M (1994) Chem Phys Lett 230:8
36. Challacombe M, Schwegler E, Almlof J (1996) J Chem Phys 104:4685
37. Beebe N, Linderberg J (1977) Int J Quantum Chem 12:683
38. Mamby FR, Knowles P (2001) Phys Rev Lett 87:163001
39. Mamby F, Knowles P, Lloyd A (2001) J Chem Phys 115:9144
40. Polly R, Werner H, Mamby F, Knowles P (2004) Mol Phys 102(21–22):2311
41. Sierka M, Hogekamp A, Ahlrichs R (2003) J Chem Phys 118(20):9136
42. Weigend F, Kattannek M, Ahlrichs R (2009) J Chem Phys 130:164106
43. Weigend F (2006) Phys Chem Chem Phys 8:1057