European **J**ournal of **Chem**istry

Check for updates

Evaluation of Coulomb integrals with various energy operators to estimate the correlation energy in electronic structure calculations for molecules

Sandor Kristyan (iD) *

HUN-REN Research Centre for Natural Sciences, H-1117 Budapest, Magyar Tudósok Körútja 2, Hungary

* Corresponding author at: HUN-REN Research Centre for Natural Sciences, H-1117 Budapest, Magyar Tudósok Körútja 2, Hungary.
 e-mail: kristyan.sandor@ttk.mta.hu (S. Kristyan).

ABSTRACT

Using energy operators $R_{C1}^{-n}R_{D1}^{-m}$, $R_{C1}^{-n}r_{12}^{-m}$, and $r_{12}^{-n}r_{13}^{-m}$ with small (n, m) values is fundamental in electronic structure calculations. Analytical integrations of the cases (n, m) = (1, 0) and (0, 1) are based on the Laplace transformation with the integrand $\exp(-a^2t^2)$, the other cases are based on the Laplace transformation with the integrand $\exp(-a^2t)$ and the two-dimensional version of the Boys function. These analytic expressions, with Gaussian function integrands, are useful for manipulation with higher moments of interelectronic distances, for example, in correlation calculations. The equations derived help to evaluate the one-, two-, and three-electron Coulomb integrals, $\int\rho(1)R_{C1}^{-n}R_{D1}^{-m}d\mathbf{r}_1$, $\int\rho(1)\rho(2)R_{C1}^{-n}r_{12}^{-m}d\mathbf{r}_1d\mathbf{r}_2$, and $\int\rho(1)\rho(2)\rho(3)r_{12}^{-n}r_{13}^{-m}d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3$, wherein $\rho(i)$ is the one-electron density describing the electron clouds in molecules, solids, or any media or ensemble of materials. Analytical solutions to integrals are more useful than numeric solutions; however, the former is not available in many cases. We evaluate these integrals numerically, even more so, the $\int f(\rho(1))d\mathbf{r}_1$ to $\int f(\rho(1),\rho(2),\rho(3))d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3$ with the analytical function f. For this task, the commonly used density functional theory numerical integration scheme has been elaborated to 6 and 9 dimensions via Descartes product. More importantly, this numerical integration scheme works not only for Gaussian type but also for Slaterian types. Analogy is commented on in terms of the powerful empirical correction between quantum potential energy correction and the empirically corrected Newton's universal law of gravity in the explanation of dark matter and energy, as well as its relation to Hartree-Fock and Kohn-Sham formalisms.

1. Introduction

The goal of this work is to discuss some extensions of Coulomb integrals (with respect to the integrand and the evaluation) for wider applicability in quantum chemical calculations. Integrations in solving the non-relativistic Schrödinger equation are a basic and inevitable step analytically, and if not applicable, numerically. Even in basis Hartree-Fock (HF) and post-HF theories with very large basis sets, sometimes the numerical integration competes with the large number of steps in analytical integration. More importantly, there are fundamental and numerical difficulties in density functional theory (DFT) to quantitatively describe the effect of corre-lations. The equations are non-linear integro-differential equations, and finding a solution requires numerical methods. Furthermore, for example, in explicitly correlated electronic structure theory, the dependence of the wave function on the interelectronic distance r_{ij} is built via the correlation factor $f(r_{ij})$. In its development, the so-called R12 theory is supposed simply that $f(r_{12}) = r_{12}$. In the last decade, the use of Slater-type geminal $f(r_{12}) = -\exp(-a\ r_{12})/a$ represents the main stream. In the products of occupied HF orbitals as $f(r_{12})\varphi_i\varphi_j$, the necessity to calculate the three- and four-electron integrals has arisen resulting from the Coulomb and exchange operators.

The Coulomb interaction between two charges in classical physics is $Q_1Q_2r_{12}^{-n}$, one of the most important fundamental interactions in nature. The power 'n' has the rigorous value 2 that describes the force, while, as a consequence, the value n = 1 describes the energy. For electron-electron interactions, the exact theory says that the Coulomb interaction energy is represented by the two-electron energy operator r_{12}^{-1}. The values (n, m) = (1, 0) and (0, 1) provide the leading potential energy terms, i.e., the nuclear-electron attraction integral, $\int\rho(1)R_{C1}^{-1}d\mathbf{r}_1$, for the exact value and the electron-electron repulsion integral, $\frac{1}{2}\int\rho(1)\rho(2)r_{12}^{-1}d\mathbf{r}_1d\mathbf{r}_2$, for approximate value in total energy (in addition to the kinetic and nuclear-nuclear repulsion energies) in DFT [1-4], and analogously, which come before, with single Slater determinant(s) in Hartree-Fock self-consistent field (HF-SCF) and Configuration interactions (CI) [5] theory. Clearly, these two are called Coulomb integrals. The other (n, m) values are useful as correction values, *e.g.*, in correlation calculations (CC).

Using the primitive Gaussian type atomic orbital (GTO) functions from

$G_{Ai}(a, nx, ny, nz) \equiv (x_i-R_{Ax})^{nx}(y_i-R_{Ay})^{ny}(z_i-R_{Az})^{nz}\exp(-a|\mathbf{r}_i-\mathbf{R}_A|^2)$ (1)
n = 2 is called GTO, n = 1 is called Slater type atomic orbital (STO)

with a > 0 and nx, ny, nz ≥ 0 benefits the important property such as $G_{Ai}G_{Bi}$ is also the (a sum of) GTO, even more, its derivatives also preserve the form as a sum of GTO. We use the shorthand notation for the energy operator (or distance operator, or weights) mostly in the discussion of numerical integration as

$$W(1) \equiv R_{C1}^{-n} \text{ or } R_{C1}^{-n}R_{D1}^{-m}, W(1,2) \equiv R_{C1}^{-n}r_{12}^{-m} \text{ or } r_{12}^{-m},$$
$$W(1,2,3) \equiv r_{12}^{-n}r_{13}^{-m} \text{ or } R_{C1}^{-n}r_{23}^{-n}, etc. \tag{2}$$

Notice that Equation 2 is (n,m) dependent but not indicated, but must always be kept in mind. The Coulomb interaction energy for molecular systems is finally expressed with the linear combination (LC) of $\int G_{A1}R_{C1}^{-1}d\mathbf{r}_1$ (an $\int G_{A1}W(1)d\mathbf{r}_1$ type) or $\int G_{A1}G_{B2}r_{12}^{-2}d\mathbf{r}_1d\mathbf{r}_2$ (an $\int G_{A1}G_{B2}W(1,2)d\mathbf{r}_1d\mathbf{r}_2$ type). Their analytical evaluation [1,2,5] is fundamental in computation chemistry. In the approximate solution of the electronic Schrödinger equation, such as HF-SCF and DFT, *etc.*, CC is necessary, as are many other devices; for example, the integrals, $\int G_{A1} G_{B2} G_{C3}r_{12}^{-n} r_{13}^{-m}d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3$ (an $\int G_{A1} G_{B2} G_{C3} W(1,2,3) d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3$ type), come up. Only approximate numerical expressions are available to evaluate the latter; see, *e.g.*, Equation 52 in reference [6] as $\langle ijm|r_{12}^{-1}r_{13}^{-1}|kml\rangle \approx \sum_p \langle ij|r_{12}^{-1}|pm\rangle\langle pm|r_{12}^{-1}|kl\rangle$, where as an alternative, the bracket notation [1,5] is used. If derivatives appear, such as $\int(\partial\rho(1)/\partial x_1)^pR_{C1}^{-n}d\mathbf{r}_1$, $\int(\partial\rho(1)/\partial x_1)^p\rho(2)^q r_{12}^{-n}d\mathbf{r}_1d\mathbf{r}_2$, *etc.*, and ρ is given as the linear combination (LC) of GTO [1,2,7], analytical evaluation is also available in these cases.

While the products and derivatives of GTO functions are of GTO type and analytical (and numerical) integration is available, the products and derivatives of STO functions are not of STO type and only numerical integration is available. In this relation, we mention the widely used devices, such as STO functions (powerful in Computational quantum chemistry (CQC)) that are expanded to LC of 3, 6, *etc.* GTO functions (less powerful than STO in CQC). Besides the analytical integration of cases in the main title, for their numerical integration, we expand the powerful numerical integration scheme for 3 dimensions [8-11] widely used in DFT for CC to 6 and 9 dimensions with the device of the Descartes product. Further-more, integrals such as $\int\rho(1)\rho(2)\rho(3)\rho(4)r_{12}^{-n}r_{34}^{-m}d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3d\mathbf{r}_4$ break up into products of simpler elements such as $\int\rho(1)\rho(2)r_{12}^{-n}d\mathbf{r}_1d\mathbf{r}_2$ and $\int\rho(3)\rho(4)r_{34}^{-m}d\mathbf{r}_3d\mathbf{r}_4$, *etc.*, and fall into the cases discussed. Last, we mention that, in addition to the ones in Equation 2, $W = a/r_{12}$ with the real value "a", as a scaling, is also a choice. The value a = 0 removes the electron-electron potential, and tuning "a" from unity can effectively account for the correlation energy in HF-SCF calculations, as detailed in references [12,13]. Furthermore, the analytical integral evaluation does not change with this "a" in this way, for example, in HF-SCF or DFT. Altogether, calculating these integrals in the main title analytically or numerically provides important energy (main and correction) values in CQC (see R12 theory also).

Important is that, if, *e.g.*, the simple $\int G_{A1}W(1)d\mathbf{r}_1$ comes up among the tasks of integrations (*i.e.*, not a general $\int f(\rho)W$ wherein f changes this situation), only n, m < 3 makes sense in the title; otherwise, this integral is infinite. It originates from (switching to spherical coordinates using $dxdydz = r^2\sin(\theta)drd\theta d\phi$ yielding) $\int_{R3} r^k \exp(-a\ r^p)d\mathbf{r} = 4\pi\int_{(0,\infty)}r^{k+2}\exp(-a r^p)dr$ = finite for p= 1 or 2 if k > -3, *etc.*

2. Analytical evaluations for values n, m = 0, 1, 2

Using Laplace transformations $R_{C1}^{-1} = \pi^{-1/2}\int_{(-\infty,\infty)} \exp(-R_{C1}^2t^2)dt$ and $R_{C1}^{-2} = \int_{(-\infty,0)} \exp(R_{C1}^2t)dt = \int_{(0,\infty)} \exp(-R_{C1}^2t)dt$, the one-electron spherical Coulomb integral, $V_{P,C}^{(n)} \equiv \int_{(R3)} \exp(-p R_{P1}^2) R_{C1}^{-n} d\mathbf{r}_1$, with $G_{P1}(p,0,0,0)$ can be evaluated [14,15]. The

case n = 1 is widely used, n = 2 with Boys function and $v \equiv p\ R_{CP}^2$ is

$$V_{P,C}^{(2)} = (2\pi^{3/2}/p^{1/2}) \int_{(0,1)} \exp(p\ R_{CP}^2\ (w^2-1))dw =$$
$$(2\pi^{3/2}/p^{1/2})e^{-v}F_0(-v) \tag{3}$$

For nonspherical $G_{P1}(p, nx, ny, nz)$ in one-electron Coulomb integral with energy operator R_{C1}^{-2}, no further trick is needed; the only formula necessary is how to shift the center of the polynomials, see Appendices 1-4 in reference [14]. With the notations $^{full}V_{P,C}^{(n)}$ and $V_{P,C}^{(n)}$, the former stands for any (spherical and non-spherical) nx+ny+nz ≥ 0 quantum number, while the latter denotes the simplest spherical (1s-like) case, nx = ny = nz = 0. In more detail, it is the integral $^{full}V_{P,C}^{(2)} \equiv \int_{(R3)}G_{P1}(p,nx1,ny1,nz1) R_{C1}^{-2} d\mathbf{r}_1$. For n = 2, see references [14,15].

For a one-electron spherical Coulomb integral with the energy operator $R_{C1}^{-n}R_{D1}^{-m}$ with n, m = 1, 2 one can also evaluate the $V_{P,CD}^{(n,m)} \equiv \int_{(R3)}\exp(-pR_{P1}^2)R_{C1}^{-n}R_{D1}^{-m}d\mathbf{r}_1$. Using the same tools as above, the algorithm is straightforward. For example, with (n,m) = (1,2) and using $g \equiv p+t^2+u$ and $f \equiv p\ t^2 R_{PC}^2 + p\ u\ R_{PD}^2 + u\ t^2 R_{CD}^2$ the

$$V_{P,CD}^{(1,2)} = \pi\int_{t=(-\infty,\infty)}\int_{u=(0,\infty)}g^{-3/2}\exp(-f/g)\ du\ dt \tag{4}$$

The two- and three-electron spherical Coulomb integrals are as follows: The two-electron spherical Coulomb integral with r_{12}^{-2}, the (n,m) = (2,0) and (0,2) cases are $V_{PQ}^{(n)} \equiv \int_{(R6)} \exp(-p R_{P1}^2) \exp(-q R_{Q2}^2) r_{12}^{-n} d\mathbf{r}_1d\mathbf{r}_2$. The case n = 1 is widely used [15] and the case n = 2 is in Equations 5 and 6, wherein $v \equiv pqR_{PQ}^2/(p+q)$:

$$V_{P,C}^{(2)} = \int_{(R3)} \exp(-p R_{P1}^2)r_{12}^{-2} d\mathbf{r}_1 =$$
$$(2\pi^{3/2}/p^{1/2}) \int_{(0,1)} \exp(p R_{P2}^2 (w^2-1))dw \tag{5}$$

$$V_{PQ}^{(2)} = 2\pi^3(pq)^{-1/2}(p+q)^{-1}\int_{(0,1)} \exp(v(w^2-1))dw =$$
$$(2\pi^3(pq)^{-1/2}(p+q)^{-1})e^{-v}F_0(-v) \tag{6}$$

The two-electron spherical Coulomb integral with mixed term $R_{C1}^{-n}r_{12}^{-m}$ and n = m = 1 is

$$\int_{(R6)}\exp(-pR_{P1}^2)\exp(-qR_{Q2}^2)R_{C1}^{-1}r_{12}^{-1}d\mathbf{r}_1d\mathbf{r}_2 =$$
$$(2\pi^2/q)\int_{u=(0,1)}\int_{t=(-\infty,\infty)} g^{-3/2}\exp(-f/g)\ dtdu \tag{7}$$

In Equation 7, $f \equiv pqR_{PQ}^2u^2+pR_{PC}^2t^2+qR_{QC}^2u^2t^2$ and $g \equiv p+qu^2+t^2$. Alternatively, the $\mathbf{R}_W = (p\mathbf{R}_P+qu^2\mathbf{R}_Q)/(p+qu^2)$ yields

$$\int_{(R6)}\exp(-pR_{P1}^2)\exp(-qR_{Q2}^2)R_{C1}^{-1}r_{12}^{-1}d\mathbf{r}_1d\mathbf{r}_2 =$$
$$(4\pi^2/q)\int_{(0,1)}F_0(gR_{WC}^2)g^{-1}\exp(-f/g)du \tag{8}$$

In Equation 8, $f \equiv pqR_{PQ}^2u^2$ and $g \equiv p+qu^2$, where R_{WC} depends on u as $gR_{WC}^2 = (p+qu^2)|\mathbf{R}_W-\mathbf{R}_C|^2 = |p\mathbf{R}_P+qu^2\mathbf{R}_Q - g\mathbf{R}_C|^2$. Unlike Equation 8, Equation 7 calls for the two-dimensional version of the Boys function; see Appendix 1. The algorithm is straight-forward for other values of (n,m).

The three-electron spherical Coulomb integral with $r_{12}^{-n}r_{13}^{-m}$ is $V_{PQS}^{(n,m)} \equiv \int_{(R9)} \exp(-p R_{P1}^2) \exp(-q R_{Q2}^2) \exp(-s R_{S3}^2) r_{12}^{-n}r_{13}^{-m} d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3$. For this, Equation 5 provides the key substitutions to integrate with $\mathbf{r}_2$ and $\mathbf{r}_3$. For example, for n = m = 1,

$$V_Q^{(n=1)} = \int_{(R3)}\exp(-qR_{Q2}^2)r_{12}^{-1}d\mathbf{r}_2 = (2\pi/q)\int_{(0,1)} \exp(-qR_{Q1}^2 u^2)du =$$
$$(2\pi/q)F_0(qR_{Q1}^2) \tag{9}$$

$$V_S^{(m=1)} = \int_{(R3)}\exp(-sR_{S3}^2)r_{13}^{-1}d\mathbf{r}_3 = (2\pi/s)\int_{(0,1)}\exp(-sR_{S1}^2 t^2)dt =$$
$$(2\pi/s)F_0(sR_{S1}^2) \tag{10}$$

Using the same tools again as above, with $f \equiv pqR_{PQ}^2u^2 + psR_{PS}^2t^2 + qsR_{QS}^2u^2t^2$ and $g \equiv p + qu^2 + st^2$, Equations 9 and 10 yield

$$V_{PQS}^{(1,1)} = (4\pi^{7/2}/(qs))\int_{(0,1)}\int_{(0,1)} g^{-3/2}\exp(-f/g)du\,dt \qquad (11)$$

Integration in Equation 11 can be done numerically, which is still more stable and reliable than equation 52 in the reference [6]. (The latter is basis-set dependent and much more difficult.) For n and/or m = 2 cases, Equation 5 must be applied analogously in the evaluation, and the algorithm is straightforward again.

Equations 9 and 10 were applied to Equation 11 that yields a two-dimensional integral on the unit square. Another way is to use Equations 8-10 which require a one-dimensional integral on a unit segment only, and with $\mathbf{R}_V \equiv (p\mathbf{R}_P + qu^2\mathbf{R}_Q)/(p + qu^2)$ the

$$V_{PQS}^{(1,1)} = (4\pi^{7/2}/(qs)) \int_{(0,1)} h(u)\,g^{-1}\exp(-f/g)\,du \qquad (12)$$

In Equation 12, $h(u) \equiv \int_{(0,c)}\exp(-g\,s\,R_{VS}^2\,w^2)dw$, $c \equiv (g+s)^{-1/2}$, $f \equiv pqR_{PQ}^2u^2$, and $g \equiv p + qu^2$. Equations 11 and 12 yield the same value for $V_{PQS}^{(1,1)}$. The $h(u)$ in Equation 12 is the prestage of the Boys function F_0. Again, Equation 12 is a two-dimensional version of the Boys function, where a one-dimensional Boys function is in the integrand.

Only these small, although useful, integer n and m values are known for analytical evaluation; for more flexible real values, numerical integration is necessary, as introduced in the next chapter.

3. Numerical evaluations for real n, m values

If beside or instead of the GTO, the STO is also used and/or non-integer or higher-value integer powers come up in electron-electron or electron-nuclear distances; analytic evaluations are not known and numerical integration is necessary. The numerical integration scheme introduced is practically the same for integer, non-integer, positive, or negative, *i.e.*, real (n, m), as well as for the use of GTO or STO. We tested the numerical integration scheme for n, m = 0, 1, 2, and GTO, where analytical integration is available for comparison, but the scheme works for all cases mentioned.

The extension of the known DFT numerical integration scheme from 3 to 6 and 9 dimensions follows. This numerical integration scheme comes from Becke's method for radial integration [9] and Lebedev's method for spherical integration [10,11], as well as the concept of Voronoi polygons (that is, geometrically atomic partition of the molecular frame). For a one-electron density with the shape of the LC of GTO or STO in Equation 1, this scheme falls into the simple sum.

$$\int f(\rho(1))d\mathbf{r}_1 \approx \sum_{k=1,...,L} c_k\,f(\rho(\mathbf{q}_k)) \qquad (13)$$

Mathematically, we are interested in smooth (differentiable) functions f and ρ, as well as ρ, which is peaking in the nucleus. Equation 13 works with special rigorous weights $\{c_k\}$ and coordinates $\{\mathbf{q}_k\}$ in 3 dimensions [9-11], estimating the integral accurately. Typical choices for $L = L_rL_s$ include $L_r = 20$, ..., 200 radial points (Chebyshev) and $L_s = 20, 86, 302$ spherical points (Lebedev). Equation 13 is widely tested in CC via DFT in CQC. We mention that the Voronoi participation of atoms in molecular frames provides another definition of partial charges [7] using numerical integration in Equation 13.

The f in Equation 13 are special functions [16] for CC in current DFT methods, typically nonlinear in ρ, so analytical integration is generally not available. The main idea is the generalization [17] of Equation 13 as

$$\int f(\rho(1),\rho(2),\rho(3))d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3 \approx \sum_{A,B,C} c_A c_B c_C\,f(\rho(\mathbf{q}_A),\rho(\mathbf{q}_B),\rho(\mathbf{q}_C))$$
with $f \equiv \rho(1)\rho(2)\rho(3)W(1,2,3)$ \qquad (14)

The f in Equation 14 is the only a particular case of CC, see Appendix 2. If $W = W(1)W(2)W(3)$, *e.g.*, $W = R_{D1}^{-3}R_{E2}^4$ or the simplest $W = 1$, *etc.*, that is, no coupling as $W = r_{12}^{-n}$ for example, Equation 14 reduces to the product of three in Equation 13. Otherwise, Equation 14 is an effective extension of Equation 13 from 3 to 9 dimensions. The extension to the 6 dimension case, $\int f(\rho(1)\rho(2))d\mathbf{r}_1d\mathbf{r}_2$, is obviously a simple algebraic reduction of Equation 14 to the analogous expression with the sum $\sum_{A,B}$. Importantly, the set $\{c_k,\mathbf{q}_k\}$ in Equation 14 is exactly the same kind/set as in Equation 13, see a simple restriction below. Simply put, A, B, and C run for the same 1, 2, ..., L points, so Equation 14 contains L^2 terms for the case $f(\rho(1)\rho(2))$ or L^3 terms for the case $f(\rho(1)\rho(2)\rho(3))$ in specific calculations. Important is to choose (slightly) different point sets $\{c_k,\mathbf{q}_k\}$ among the two (i = 1, 2) as well as among the three (i = 1, 2, 3) variables $\mathbf{r}_i$ (via L_r and L_s), as detailed in the reference [17]. The latter technically avoids singularities from reciprocal values, *e.g.*, calculating a term r_{13}^{-2} from the same upcoming coordinates, *e.g.*, from $\mathbf{q}_{11} \in \{\mathbf{q}_A\}$ and $\mathbf{q}_{11} \in \{\mathbf{q}_C\}$.

An improvement in accuracy is as follows: Equation 14 with the specific f in more detail, is

$$\int \rho(1)\rho(2)\rho(3)W(1,2,3)d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3 \approx$$
$$\sum_{A,B,C} c_A c_B c_C\,\rho(\mathbf{q}_A)\rho(\mathbf{q}_B)\rho(\mathbf{q}_C)\,W(\mathbf{q}_A,\mathbf{q}_B,\mathbf{q}_C) \qquad (15)$$

For example, if $W = r_{12}^{-1}r_{13}^{-2}$ in Equation 2, then $W = |\mathbf{q}_A-\mathbf{q}_B|^{-1}|\mathbf{q}_A-\mathbf{q}_C|^{-2}$. Normalizing to the case $W = 1$ is possible to improve the accuracy. If $W = 1$, Equation 15 reduces to $N^3 = (\int\rho(1)d\mathbf{r}_1)^3 = \int\rho(1)\rho(2(\rho(3)d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3 \approx \sum_{A,B,C} c_A c_B c_C\,\rho(\mathbf{q}_A)\rho(\mathbf{q}_B)\rho(\mathbf{q}_C) = (\sum_A c_A\,\rho(\mathbf{q}_A))^3$ if the same mesh is used. Dividing these two equations yields

$$\int \rho(1)\rho(2)\rho(3)W(1,2,3)d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3 \approx$$
$$[\int\rho(1)d\mathbf{r}_1/\sum_A c_A\rho(\mathbf{q}_A)]^3\sum_{A,B,C} c_A c_B c_C\rho(\mathbf{q}_A)\rho(\mathbf{q}_B)\rho(\mathbf{q}_C)W(\mathbf{q}_A,\mathbf{q}_B,\mathbf{q}_C) \qquad (16)$$

For the left side, the $\int\rho(1)d\mathbf{r}_1$ can be evaluated analytically in practice because it is the LC of GTO in Equation 1. One must keep in mind that Equation 16 supposes that the same mesh is used for all three variables. If three different meshes are chosen (see above) and, even more, three different densities (distinguished by indexes α, β, γ as in the test below), the factor is

$$[\int\rho_\alpha(1)d\mathbf{r}_1\int\rho_\beta(1)d\mathbf{r}_1\int\rho_\gamma(1)d\mathbf{r}_1/(\sum_A c_A\rho_\alpha(\mathbf{q}_A)\sum_B c_B\rho_\beta(\mathbf{q}_B)\sum_C c_C\rho_\gamma(\mathbf{q}_C))] \qquad (17)$$

In Equation 17, the indexes A, B, and C represent the three (generally different) meshes for variables $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$, and again, the particular normalization (*e.g.*, spin density) makes the nominator simpler. The reduction of 9 dimensions (Equations 16 and 17) to 3 dimensions (*i.e.*, to $\int\rho(1)W(1)d\mathbf{r}_1$) and 6 dimensions (*i.e.*, to $\int\rho(1)\rho(2)W(1,2)d\mathbf{r}_1d\mathbf{r}_2$) is straightforward. In the numerical integration tests below, we use the not-normalized $\rho(i) = G_{Ai}(a,0,0,0)$ mostly. If the integration device in Equation 13 is 100% accurate, the factor is the unit in Equation 17, regardless of whether the $\rho(1)$ is normalized (*e.g.*, to N) or not. Jumping from 3 (Equation 13) to 6 and 9 dimensions (Equations 14-17) increases the computation time as L *vs.* L^2 and L *vs.* L^3, respectively. Equations 14-17 are mathematically the Descartes product of the scheme for 3 dimensions in Equation 13.

We provide brief and powerful tests for the numerical integration scheme for 6 and 9 dimensions in Equations 14-17.

Table 1. Error in numerical integrations (Equations 13-17) in comparison to analytical integrations (Equations 3-12) demonstrated.

W in Equation 2	Type in Equation 1	L_r (Radial points)	Equations used	Error in numerical integration %
1	GTO and STO	Low L_r vs. high L_r	Equations 15-17	10^{-4} *vs.* 10^{-9}
R_{Q1}^{-1}	GTO	High L_r,	Equation 15 vs. Equations 16-17	10^{-2} - 10^{-8} *vs.* 10^{-2} - 10^{-10}
R_{Q1}^{-2}	GTO	Low and high L_r	Equation 15 and Equations 16-17	0.5 - 0.0001
r_{12}^{-1}	GTO	Low and high L_r	Equation 15 and Equations 16-17	4 - 10^{-9}
r_{12}^{-2}	GTO	Low and high L_r	Equation 15 and Equations 16-17	5 - 10^{-6}
$R_{S1}^{-1}r_{12}^{-1}$	GTO	Low and high L_r	Equation 15 and Equations 16-17	5 - 0.001
$r_{12}^{-1}r_{13}^{-1}$	GTO	Low L_r	Equation 15 and Equations 16-17	1 - 10^{-5}

The numerical integration scheme for 3 dimensions in Equation 13 is widely used and tested for integrations in DFT and CC; see the textbook properties in Appendix 2 in this respect. When $f(\rho(1))$ is the LC of the primitive GTO in Equation 13 and an analytical integral is available, the numerical integration follows it to many decimal digits. However, when $f(\rho(1))$ is a non-linear function of ρ and the analytical integration is not available (and the analytical integration of GTO loses its relevance in this situation), Equation 13 still works, as experienced by reasonable results in CQC for CC. Our hypothesis is the same for Equations 14-17 with a more general f, even though we have tested it for particular cases of f only in Equations 14-17. Another benefit of numerical integration in Equations 15-17 is that it works not only for GTO but also for STO (of which analytical evaluation is far more difficult if it exists). Furthermore, it works for any positive or negative real value of n and m in Equation 2 with the same computational costs. For much larger than $|n|$, $|m| \in [0, 4]$ values, the radial and spherical schemes reported in references [9-11] are not high-level enough, but their extension is straightforward.

Furthermore, with Equations 15-17, all HF-SCF or DFT integrals can be evaluated numerically for both STO and GTO, while today only GTO is used for analytical integration during HF-SCF (for all) or DFT (all, except the CC part). The advantage of STO (which has the form of exact atomic wave functions) to GTO is that much fewer basis functions are necessary because GTO is based on the approximation $\exp(-p|\mathbf{r}_1-\mathbf{R}_P|) \approx \sum_i c_i G_{Pi}(a_i,0,0,0)$; see the idea of STO-3G, *etc.* basis sets [18] in HF-SCF or DFT routines, for example. For this reason, using a higher-level GTO basis, the number of analytical operations is sometimes about the same or larger than the numerical operations via the mesh points in Equations 13-17.

In the test below, we keep in mind the values that are important in practice: We have chosen the magnitude of the GTO exponents from the STO-3G basis set [17], that is, for analytical Equations 3-12 we used the values $R_Q = (d, 0, 0)$ and $R_S = (0, d, 0)$ with d = 0, 2, 8, 20, and p, q, s = 0.3, 1, 3, 15, 70 for GTO and p, q = 0.3, 1, 3, 6 for STO, as well as G_{Ai}(p or q, 0, 0, 0) was used from Equation 1. Additionally, we have compared the numerical Equation 15 with the improved numerical Equations 16 and 17 using L_r = {20, 22, 24} (low), {50, 52, 54}, and {100, 102} (high) values for the mesh points in the integration for $d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3$ and $d\mathbf{r}_1 d\mathbf{r}_2$ together with Equation 2. The slightly different L_r values in the set (containing 3, 3, and 2 values, respectively, *e.g.*, 100 *vs.* 102) were used to avoid singularity by r_{12}^{-n}. The % error was defined as 100(1-(numerical/analytical)). The test values obtained are summarized in Table 1.

The conclusion of this test is as follows. Few radial points (L_r = 20) compete with larger ones in numerical integration, which is good for reducing the computation time. An analytical expression for a GTO (like in Equation 3, *etc.*) is one equation with a few operations in contrast to $(L_r L_s)^m$ operations (m = 1, 2, 3) in numerical integration by Equations 13-17, the latter is the same for this single GTO or for the entire ρ_{model}. However, in HF-SCF or DFT routines, a large LC of GTO has to be evaluated when expressing the ρ_{model}, so for expressions like the ones in the abstract, the number of operations can be the same for analytical and numerical integrations. The W in Equation 2 can deform the angular and radial symmetry if $W \neq 1$, and it can

decrease the numerical accuracy; however, this situation is the same in general CC formulas when using Equation 13 in 3 dimensions as well. In summary, the numerical integration scheme in Equations 13-17 is an useful, accurate, recommended, and the only option if analytical integration is not available.

4. The higher moment Coulombic energy operators in practice

In short, CI (including correlation effects) and HF-SCF (without correlation correction) methods use analytical integration [5,19], while DFT uses numerical integration (Equation 13) for correlation effects [1,2] today. Numerical integration in CC is widely used, accepted, and indispensable in approximate solutions of the Schrödinger equation for electronic structures of molecular systems, where certain non-linear $f(\rho(1), ...)$ functions with or without energy operators (Equation 2) come up.

Interestingly, two typical quantum phenomena show a simple, approximate linear relationship. First, the correlation energy [7,20], which is more a mathematical correction than a physical phenomenon, obeys the equation.

$$-0.045 (N-1) < E_{corr}[\text{hartree}] < -0.030 (N-1) \qquad (18)$$

Equation 18 is a quasi-linear relationship (regression 97.2%, standard deviation 49 kcal/mole) with the number of electrons, N, and it slightly fluctuates with the nuclear frame and molecular charge [21]; see Appendix 3. The slight fluctuation means that Equation 18 holds for any molecular system, but since the gap increases linearly with N as 0.015(N-1), it is far from being chemically accurate. However, it is excellent for predicting magnitude. To reach chemical accuracy, an adequate CC is needed. The E_{corr} is a computation handicap to correct the single-determinant approximation for the ground-state wave function. (If H_{ee} is algebraically removed from the Hamiltonian, the single-determinant form is an accurate solution [13] for ground and excited eigenfunctions; see Appendix 3.) An important restriction is that Equation 18 is based on HF-SCF/6-31G* level calculations and can be tuned accordingly for other bases. Second, the real phenomenon of zero-point energy [20,22] for neutral molecules ($N = \sum_{A=1}^{M} Z_A$ and M > 1) obeys

$$ZPE[\text{hartree}] \approx 0.012 \sum_{A=1}^{M} m_A^{-1/2} \text{ and}$$
$$0 < ZPE[\text{hartree}] < 0.0036(N-1) \qquad (19)$$

Equation 19 is a quasi-linear relationship (regression 99.5%, standard deviation 3 kcal/mole) with the sum of the reciprocal square root of mass (in a.u. in the nuclear frame), and it visibly fluctuates with N. (The latter is clear, *e.g.*, HCl with M = 2 and C_2H_6 with M = 8, both have N = 18 electrons, while ethane has larger degrees of vibrational modes via 3M−5 for linear and 3M−6 for nonlinear molecules.) For chemically accurate (1 kcal/mol) results, Equation 29 is good only for magnitudes, but for accurate Zero-point energy (ZPE), it needs an accurate potential energy surface (including CC if necessary) and an accurate frequency (ν) calculation, finishing with the correct equation, $ZPE = \frac{1}{2}\sum_{1}^{3M-6} h_{Planc}\nu_i$. For a medium-to-large molecule

consisting, *e.g.,* of 50 atoms (144 modes of vibration), the total zero-point energy is substantial. Demonstrating examples for Equations 18 and 19 are given in Appendix 4.

As indicated in the Introduction and reference [12], a very powerful correlation calculation has been reported by modifying the Coulomb energy operator as $r_{ij}^{-1} \rightarrow 0.99762\ r_{ij}^{-1}$ (lowering the increment of electron-electron repulsion), resulting from a careful fit using the sets of G2 and G3 molecules [23]. Its additional advantage is that existing HF-SCF routines need modifications only in a few program lines. Algebraically, this modification is more of a tuning of the Slater determinant wave function than of a modification of the Coulomb energy operator.

Besides many devices [1,2,5,12], a powerful empirical correction in CC is the modification in the approximation of V_{ee} and V_{ne} in Equation 21 as $r_{12}^{-1} \rightarrow r_{12}^{-n}$ and $R_{C1}^{-1} \rightarrow R_{C1}^{-n}$, respectively, with $n = 1+\varepsilon$ for electronic potential energy correction during or after, *e.g.,* an HF-SCF/basis calculation with a small fitted value ε. For example, the sets of G2 and G3 molecules [23] and the atomic energies of CI [24] can be used in this fit; they will be reported in other works. The textbook fact is that the kinetic energy operator in quantum mechanics, T, is constructed from the kinetic energy in classical physics ($T = \frac{1}{2}p^2/m$) by transforming it into the differential operator $-\frac{1}{2}\nabla^2$. However, the operator for the Coulomb potential energy of two charged particles (between macroscopic objects or between microscopic electrons, protons, or nuclei) is the multiplicative operator r_{12}^{-1} or $-R_{A1}^{-1}$, *i.e.,* the same algebraic form (reciprocal distance) holds in classical and quantum physics. Just as we were surprised that the kinetic energy operator at the birth of quantum mechanics, circa 1925, was different from classical mechanics, we should have been surprised thereafter (in the author's opinion) that the potential energy operator is, however, the same. In the calculation of the quantum scale, we also call Newton's universal law of gravity for the force between macroscopic objects $m_1\ m_2/r^2$, in which the mechanical potential energy is $m_1\ m_2/r$. On a cosmological scale, where many large bodies participate in the universe, to interpret the yet unexplained measurements that force us to suppose the existence of dark matter and energy, this law has been empirically corrected to resolve this issue as $m_1\ m_2/r^m$ with m = $2+\zeta$, where is a small fitted value. Empirically, it provides very simple and remarkable results to explain the discrepancies in cosmological measurements in relation to dark matter and energy. It originates from Milgrom's (1983) [28] claim that if the gravitational force came to vary inversely linearly with the radius (as opposed to the inverse square of the radius, as in Newton's law of gravity), the dark matter and energy could be explained. It is important to mention that there are strong doubts in the scientific community about this correction with ζ as a plausible physical explanation, even though it empirically works. However, the semiempirical CC (based on physical considerations and improved with some empirical parameters) at the quantum scale is totally accepted in the scientific community.

This modified approximation just mentioned for V_{ee} and V_{ne} in Equation 21 using r_{12}^{-n} and R_{C1}^{-n} from Equation 2, respectively, is a powerful tuning device to include correlation effects during a CQC algorithm via n around 1. The n = 1 is the theoretical power of the Coulomb energy operators (see Equations 2 and 21). Only a simple example is given next for demonstration. With a hydrogen atom accurate 1s orbital $f_i = \exp(-r_i)$, the energy integral $\langle f_1 f_2|r_{12}^{-n}f_1 f_2\rangle\ /\ \langle f_1 f_2|f_1 f_2\rangle = \langle f_1 f_2|r_{12}^{-n}f_1 f_2\rangle)/\pi^2$ with the n = 0.9, 1 and 1.1 yields 0.647791, 0.632432, and 0.622287 h, respectively. The deviation from n = 1 is 0.632432 - 0.647791 = -0.015359 h $\approx$ -9.6 kcal/mol for n = 0.9, and 0.632432 - 0.622287 = 0.010145 h $\approx$ 6.4 kcal/mol for n = 1.1. The energy integral $\langle f_1 f_2|R_{C1}^{-n}f_1 f_2\rangle\ /\ \langle f_1 f_2|f_1 f_2\rangle = \langle f_1|R_{C1}^{-n}f_1\rangle)/\pi$ with the n = 0.9, 1 and 1.1 yields 0.511865, 0.483547,

and 0.460405 h, respectively. The deviation from n = 1 is 0.483547 - 0.511865 = -0.028318 h $\approx$ -17.8 kcal/mol for n= 0.9, and 0.483547 - 0.460405 = 0.023142 h $\approx$ 14.5 kcal/mol for n = 1.1. Importantly, Equation 18 predicts about 0.03 h $\approx$ 18.8 kcal/mol correlation effect in a system per electron, a consistent value with these energy integrals as a function of n in magnitudes. For these values, the 3 and 6 dimensional cases of numerical integration in Equations 13 and 14 for n $\neq$ 1, and analytical integration for n = 1 were used.

Another device is introduced, where the square of the Hamiltonian operator [13] can be used to calculate the correlation effects, together with the use of energy operators in Equation 2. Applying the Hamiltonian twice to the ground state wave function simply yields $H^2\Psi_0 = E_{0,electr}H\Psi_0 = E_{0,electr}^2\Psi_0$, that is, $E_{0,electr}^2 = \langle\Psi_0|H^2\Psi_0\rangle/\langle\Psi_0|\Psi_0\rangle$, wherein the H^2 preserves the linearity and hermetic property of H itself. For example, if the HF-SCF/basis single determinant S_0 approximates Ψ_0, the approximation $|E_{0,electr}| \approx (\langle S_0|H^2 S_0\rangle/\langle S_0|S_0\rangle)^{1/2}$ can improve the approximation $E_{0,electr} \approx \langle S_0|H|S_0\rangle/\langle S_0|S_0\rangle$ as follows: The non-associative property must be kept in mind; see Appendix 5. If the approximate S_0 is used, the optimization of $\langle S_0|H^2 S_0\rangle/\langle S_0|S_0\rangle$ is a more complex task than that of the $\langle S_0|HS_0\rangle/\langle S_0|S_0\rangle$ by e.g. a HF-SCF. The variation principle is also not as simple: With the help of a primitive picture, because the $\langle S_0|HS_0\rangle/\langle S_0|S_0\rangle$ depends on LCAO parameters (let it be now a single c) as *e.g.,* a U-shape $(c-1)^2-1$ parabola with minimum at c = 1 under the c-axis, it follows that the $\langle S_0|H^2 S_0\rangle/\langle S_0|S_0\rangle$ depends on c as a W-shape curve $((c-1)^2-1)^2$. The latter lies on the c-axis with its two minimum points at c = 0 and 2, as well as the former global minimum at c = 1 has transformed into a local maximum at c = 1. In real systems, the global energy minimum we seek transforms into a local maximum with respect to Linear combination of atomic orbitals (LCAO) parameters or an increasing function. In this procedure, the Coulomb integrals with operators H_{ne}^2, $H_{ne}H_{ee}$, and H_{ee}^2 come up, that is, with the energy operators $R_{Ai}R_{Bj}$, $R_{Ai}r_{jk}$, and $r_{ij}r_{kl}$ with A, B = 1..., M nuclei and i, j, k, l = 1..., N electrons from Equation 2. All these can be evaluated numerically with STO basis sets or analytically with GTO basis sets, as detailed in the previous chapters. Grouping as H= $H_\nabla+(H_{ne}+H_{ee})$ and using $\langle S_0|H^2|S_0\rangle = \langle HS_0|HS_0\rangle$ to avoid $H_\nabla H_{ne}$ and $H_\nabla H_{ee}$, it yields $E_{electr,0}^2 = (T+V_{ne}+V_{ee})^2 = T^2 + 2T(V_{ne}+V_{ee}) + (V_{ne}+V_{ee})^2 \approx \langle S_0|H^2 S_0\rangle = \langle H_\nabla S_0|H_\nabla S_0\rangle + 2\langle H_\nabla S_0|(H_{ne}+H_{ee})S_0\rangle + \langle S_0|(H_{ne}+H_{ee})^2 S_0\rangle$ with normalization $\langle S_0|S_0\rangle = 1$, an *ab initio* [5] term. The identification of the three integrals is obvious; for example, the kinetic operator H_∇ comes up in the first and middle terms, as well as the first being a pure kinetic part of T^2 and the last being a pure Coulombic part among electrons and nuclei. In addition to the analytical evaluation, these integrals can be approximated with DFT using ρ as follows: The kinetic energy functional (for part of T^2) with one-electron orbitals is $\langle H_\nabla S_0|H_\nabla S_0\rangle \approx (1/4)\sum_1^N\int(\nabla^2\varphi_i)^2 d\mathbf{r}_1$, corresponding to the emblematic Equation 20 in Appendix 6 (for full T). The $\int\rho(1)\rho(2)W(1,2)d\mathbf{r}_1 d\mathbf{r}_2$ terms here with W from Equation 2 and n, m = 1, 2 correspond to the emblematic Equation 11 in Appendix 6.

An example with a hydrogen atom follows. Choose a GTO from Equation 1 as $S_0 = c\ \alpha\ G(\mathbf{r}_1)$ with G$\equiv$ $\exp(-a\ r_2)$ to approximate the exact $\Psi_0 = c\ \alpha\ \exp(-r)$ and $E_{e.electr,0} = -0.5$ h, as we know from textbooks. The spin integrals out from $\langle|\rangle$, so one can start with S_0: = G and the c will drop via the normalization in the denominator. The H= $-\frac{1}{2}\nabla^2-1/r$ yields $\langle S_0|H^k S_0\rangle = \int(-2a^2 r^2 +3a-1/r)^k G^2 d\mathbf{r}$ with k = 1 and 2. The expression $E_k(a) \equiv \langle S_0|H^k S_0\rangle/\langle S_0|HS_0\rangle$ for k = 0, 1 and 2 yields $E_0 = 1$, $E_1 = a^{1/2}(2/\pi)^{1/2}(-8+3(2a\pi)^{1/2})/4$ and $E_2 = 8a^{3/2}(2/\pi)^{1/2}(-1+(16+15a)(2\pi/a)^{1/2}/64)$, because now an analytical evalua-tion is available. The E_1 has a minimum (variation principle) at a = 0.282942 bohr^{-2} and an energy value $E_1 = -0.424414$ h with error 0.075586 h $\approx$ 47.4 kcal/mol from value -0.5. This error is

due to the weak basis set, {G}, used, i.e. a "basis set error", since there is only N = 1 electron in the system (and no correlation energy). The E_2 is a monotonically increasing function with a. The E_2 with a = 0.282942 (energy optimized value of dE_1/da= 0) yields -$(E_2)^{1/2}$ = -0.686516 h; it is under -0.5 since the variation principle does not hold for the energy integral E_2. Both approximate $E_{electr,0}$; taking the average as $(E_1 - (E_2)^{1/2})/2$ = (-0.424414 -0.686516)/2 = -0.555465 h, it deviates with 0.055465 h ≈ 34.80 kcal/mol from value -0.5, a remarkable improvement over 47.4 kcal/mol. For a non-bonding state, the average $(E_1 + (E_2)^{1/2})/2$ must be taken. With more advanced basis sets (STO-3G, 6-31G*, and up), the energy gap is narrower around the theoretical value of -0.5. Tests for many atoms and molecules will be reported in other works.

In relation to the quality of points, the accurate atomic wave function STO, and the approximate GTO for orbitals, we mention that, in 3D physical space, the cardinality of irrational numbers (which, by definition, cannot be the ratio of two integers, *e.g.*, $\sqrt{2}$, e, π, *etc.*) is larger than that of rational numbers (which, by definition, are the ratio of two integers). These two sets, constituting the real numbers, are incountable infinite and countable infinite sets, respectively. Importantly, any small neighborhood of any irrational point contains dense, *i.e.*, infinite, irrational and rational points, and vice versa. Rational values have infinite segments periodically (*e.g.*, segment 142857 in 1/7), while irrationals do not. The subdivision of real numbers into rational and irrational ones also serves an "elementary or atomic" separation in physical space, still considered totally irrelevant for natural sciences such as physics, chemistry, or biology. The elementary concepts of the physical world are matter, time, space (mathematical coordinates), and information (laws, many in mathematical forms). The rational and irrational coordinates may have importance or consequence in electromagnetic and gravitational wave propagation and in the explanation of dark matter and energy; the latter is present everywhere to our knowledge but difficult to detect and explain. For this, recall that if the genus (corresponding to powers in their algebraic expression, see, *e.g.*, Equation 2) of any curve (as a physical trajectory) greater than 1 over the field of rational numbers has only finitely many rational points. (The genus of a curve with degree n is $(n-1)(n-2)/2$; the singularities, for example, the crossing point in the shape of number 8, even decrease this value.) A linear function passes through all real points, but, for example, the only pairs of rational numbers on the graph of the simple y^2= x(x-1)(x+1) are the points (1, 0), (-1, 0), and (0, 0); otherwise, it passes through irrational points, and the proof is elementary. The general statement was conjectured in 1922 by L. Mordell [29] and proved by G. Faltings in 1983 [30]. Important related cases are, *e.g.*, the trigonometric sin function (the basic function in wave theory; see Niven's theorem stating rational values are at 0, 30, and 90° only), GTO and STO (see Equation 1 at nodes and at nuclei), the space-time distance of any two points in space representing two events, i.e., the invariant $x^2+y^2+z^2+(ict)^2$ (with imaginary unit, or with quaternion unit the alternative $(ix)^2+(jy)^2+(kz)^2+(ct)^2$), gravitational or electronic potential proportional to $(x^2+y^2+z^2)^{-1/2}$ (see Equation 2), *etc.*, which obey this property, that is, practically the irrational points are involved in their graph or trajectory if the powers (genus) are higher.

5. Conclusions

The wider domain of Coulomb integrands and their analytical and numerical integrations have been discussed and demonstrated for more effective applications in quantum chemical calculations. These are useful in correlation calculations to evaluate the two- to three-electron integrals more accurately and more rapidly. The power and flexibility of the

distance operator r_{12}^x with a real value x around unity have also been demonstrated.

Appendix

Appendix 1

The two-dimensional Boys function, its preequation, and its integration are commented on next. The one-dimensional Boys function F_0 via the term $g^{-3/2}\exp(-f/g)$ in the integrand is in the preequations Equation 3. In the middle stage of the deduction, for example, in Equation 3, the $V_{P,C}(2) = \pi^{3/2}\int_{(-\infty,0)}g^{-3/2}\exp(f/g)dt$ with f≡pR_{CP}^2t and g≡p-t comes up. In the two-dimensional case, the same term appears in the integrand, but instead of the function set {f(t), g(t)}, the {f(u, t), g(u, t)} is present; see Equations 4, 7, 11. The right-hand side of Equations 8 and 12 with the one-dimensional Boys function in its integrand is the two-dimensional Boys function. The $g^{-3/2}\exp(f/g)$ is the core part of the integrands in all these cases. A finer property is that f = f($(-u)^K$,$(-t)^L$) and g = g($(-u)^K$,$(-t)^L$) are 2nd and 1st order polynomials in the variables $(-u)^K$ and $(-t)^L$, respectively, with K, L = 1 or 2. The -u → u and -t → t transformations can also be used. The K, L = 1 generates $\exp(w^2)$, while the L = 2 generates $\exp(-w^2)$ in the integrand. Over a finite domain, the integral of $\exp(w^2)$ is finite.

Appendix 2

In relation to analytical vs. numerical integration, we mention that the Coulomb energies between nuclei (with index n) and electrons (with index e) with a real (*i.e.* exact) one-electron density, $\rho(1)\geq 0$, are the $V_{ne} = N\sum_C Z_C\int\rho(1)R_{C1}^{-1}d\mathbf{r}_1$ in HF-SCF, DFT, *etc.*, and $V_{ee}\approx \frac{1}{2}\int\rho(1)\rho(2)r_{12}^{-1}d\mathbf{r}_1d\mathbf{r}_2$ in DFT, for example. Different theories use different physically realistic $\rho_{model}(1) \geq 0$. The HF-SCF theory uses one real-valued energy optimized Slater determinant, S, for the ground state to approximate the wave function yielding $V_{ne}\approx N\sum_C Z_C\int S^2 R_{C1}^{-1}$ and $V_{ee} \approx \binom{N}{2}\int S^2 r_{12}^{-1}$, where the integral is for N spin-orbit full space $d\mathbf{x}_1...$, $d\mathbf{x}_N$. Similar expressions are used in configuration interaction (CI) methods [5,19] for more accurate ground and excited states. The ratio holds as $\binom{N}{2}\int S^*Sd\mathbf{x}_1...$, $d\mathbf{x}_N/\frac{1}{2}\int\rho(1)\rho(2)d\mathbf{r}_1d\mathbf{r}_2 = (N-1)/N \to 1$ if N → ∞. In this ratio, if the operator W = r_{12}^{-1} is applied to the integrands (for energy calculations, which are physically exact in the nominator), a 'self-interaction' phenomenon appears in the (approximate) denominator, which must be taken into account in DFT for the ground state.

The HF level ground state one-electron density is $\rho_{HF-SCF}(1)\equiv N\int S^2 ds_1 d\mathbf{x}_2...d\mathbf{r}_N$. As a consequence, the V_{ne} and V_{ee} can be approximated with the LC of GTO functions from Equation 1, and analytical integration is available. Contrary, the very first correlation energies in the history of DFT [1,2] using $\rho_{DFT}(1)$ to approximate $\rho(1)$, the local density approximation for exchange energy ($C_{LDA}\int\rho(1)^{4/3}d\mathbf{r}_1$), or the improvement of kinetic energy functional [25] by adding the von Weizsäcker (1935) [31] correction ($C_W\int|\nabla\rho(1)|^2\rho(1)^{-1}d\mathbf{r}_1$), along with the later and very advanced exchange-correlation DFT functional as $E_{xc}[\rho(1)] = \int\varepsilon_{xc}(\rho(1))d\mathbf{r}_1$ starting from GGA (including the ρ_{DFT} and $\nabla\rho_{DFT}$ in the exchange-correlation potential in ε_{xc}) and its developed meta-GGA (including the $\nabla^2\rho_{DFT}$), are not LC of GTO, and numerical integration is necessary.

Appendix 3

An elementary demonstration of the origin of the correlation effect as follows: The real eigensolution of operator $-\frac{1}{2}(\nabla_1^2+\nabla_2^2) + r_{12}^{-1}$ for the ground state (two free electrons) is $E_{electr,0}$ = -1/4 and non-Slater $\Psi_0 = s(1,2)\exp(\frac{1}{2}r_{12})$ with spin function $s(1,2)\equiv (\alpha_1\beta_2-\alpha_2\beta_1)$, while its HF approximation would be the Slater determinant $S_0= s(1,2)\exp(a(r_1^2+r_2^2))$ with energy optimized value for a. (In this case, the Ψ_0 is not well-behaved!)

The correlation effect and basis set (STO-1G) error come from the difference between the two. Recall that the advanced correlation energy per electron has been predicted using the "three-dimensional uniform electron gas".

Furthermore, the real eigensolution of operator $-\frac{1}{2}(\nabla_1^2+\nabla_2^2) -(R_{01}^{-1}+R_{02}^{-1})$ for the ground state (two electrons around a proton in the origin with a hypothetical neglect of electron-electron interaction) is $E_{electr,0} = -1$ and Slater determinant $\Psi_0 = s(1,2)f_1f_2$ with orbitals $f_i \equiv \exp(-R_{0i})$ and $i = 1, 2$. The $\langle f_1f_2|r_{12}^{-1}|f_1f_2\rangle / \langle f_1f_2|f_1f_2\rangle = 0.632432$ is a weak approximation for V_{ee}. These far-form energy optimized values yield a total ground state energy of $-1+0.632432 = -0.367568$, a variation value in comparison to the CI value of **-0.527716 h** for hydrogen anion. For H⁻, this $s(1,2)f_1f_2$ is not an accurate ground state wave function, but, importantly, Equation 2 can tune even in this primitive algorithm. Just for curiosity, the real eigensolution of operator $-\frac{1}{2}(\nabla_1^2+\nabla_2^2) + r_{12}^2$ for the ground state is $(3, \exp(-\frac{1}{2}r_{12}^2))$, as well as of operator $-\frac{1}{2}\nabla_1^2 + R_{01}^2$ is $(3(\frac{1}{2})^{1/2}, \exp(-(\frac{1}{2})^{1/2}R_{01}^2))$. In the latter, $-R_{01}$ yields a complex solution, and in both cases the eigenfunctions behave well. The polynomial multiplier of the derivatives of exponential drops luckily at $1/r$, $1/R$, r^2, and R^2 potentials in this way, yielding integer and half-integer constants. The r^x and R^x cases for other positive and negative real x values possibly do not have analytical forms. Besides the fact that STO can be expanded into the LC of GTO (the simplest widely used in practice is called STO-3G), it is remarkable from the above that STO solves when Coulomb potentials, $1/r_{12}$ or $1/R_{01}$, are involved, in contrast to GTO, which solves when potentials r_{12}^2 or R_{01}^2 are involved.

Appendix 4

The reason the CC is necessary is exampled on the potassium 1st ionization potential on HF-SCF/basis level [26] with the good p-GCHF-5Z basis ($Z_A = 19$ and to avoid ZPE, an atom is chosen): The approximated ground-state electronic energy difference for ionization K $\rightarrow$ K⁺ is -599.0171688 -(-599.1646910) = 0.1475222 h, the experimental ionization potential is 0.1595125 h $\approx$ 100.1 kcal/mol, and the error is 0.1595125 - 0.1475222 = 0.0119903 h $\approx$ 7.5 kcal/mol. It is a good example to represent the magnitude of values, *e.g.*, 599.0 vs. 0.1595 ($\approx$ 2.7%). The HF-SCF, post-HF-SCF, DFT, *etc*. methods yield more accurate results for (the most important) energy differences than for absolute energies. Equation 18 with a particular constant such as b(N-1) drops for energy differences; however, it is good for estimating magnitudes. For example, for a K atom (N = 19 electrons), the ground state energy is -599.149001 and -599.73046 h at HF-SCF/6-311G and at the accurate G3 level, respectively. The E_{corr} = -599.73046 + 599.149001 = -0.581459 h $\approx$ 364.9 kcal/mol, and from Equation 18 the -0.045*18 = -0.81 < -0.581459 < -0.030*18 = -0.54 holds. Similarly, for (MP2 level) geometry optimized ground state total energy of benzene (C_6H_6, N = 42 electrons), the G2 [23] and HF-SCF/3-21G level $E_{corr}/(N-1)$ = (-231,87667 + 229,41677)/(42-1)= -0,06 h, while the G3 [23] and HF-SCF/6-31G* level $E_{corr}/(N-1)$ = (-232.14840 + 230.70204)/(42-1) = -0.0353 h. It demonstrates the basis dependence of E_{corr} and Equation 18; more precisely, the basis set error is added to the correlation effect. In detail, $E_{corr} < 0$ is the basis set limit coming from a single determinant approximation; however, the applied (not infinite, i.e., not complete) basis in practice also causes an error < 0 by the variation principle.

The ZPE approximation in Equation 19 for uniatomic molecules (*e.g*., H_2, C_2, N_2, Cl_2, tricarbon C_3, ozone O_3, *etc*.) simplifies to ZPE[X_M, h] $\approx$ c M/$\sqrt{m_X}$ and to the ratio ZPE[X_M]/ZPE[Y_M] $\approx$ (m_Y/m_X)$^{1/2}$. For example [27], with a larger mass difference as in ZPE[Cl_2]/ZPE[H_2], the approximate (m_H/m_{Cl})$^{1/2}$ = (1/35)$^{1/2}$ $\approx$ 0,17 compares to the experimental 279.22 cm⁻¹/2179.3 cm⁻¹ $\approx$ 0.13, and for smaller mass

differences as in ZPE[N_2]/ZPE[C_2], the approximate (m_C/m_N)$^{1/2}$ = (12/14)$^{1/2}$ $\approx$ 0.9 compares to the experimental 1175.78 cm⁻¹/924.0 cm⁻¹ $\approx$ 1.3. Clearly, *e.g.,* an HF-SCF/basis level frequency calculation is more accurate than the left in Equation 19 (which, for example, cannot account for isomers), but even the HF-SCF/basis needs improved energy wells via correlation calculation for acceptable ZPE values.

Appendix 5

The nonassociative property in the application of H^2 is fundamental, because the notation H^2 itself can be misleading. In the operator algebra of CQC, the strict tacit agreement is the interpretation $H^2\Psi \equiv H(H\Psi)$ and not $(H^2)\Psi$. Simple examples demonstrate the trap. The differential operator O= d/dx yields the associative $O(Of) = (O^2)f = d^2f/dx^2$, or the matrix operators A and B yield an associative $A(Bf) = (AB)f$, but O = d/dx+x yields non-associative $O(Of) - (O^2)f = x \, df/dx \neq 0$. For the eigenvalue equation of the latter, $f = c \exp(ax-x^2/2)$ solves $Of = af$ and $O(Of) = a^2f$ in contrast to $(OO)f = (a^2-ax+x^2)f \neq a^2f$. The twin textbook property is the fundamental commutator [1,5], for example, $[x,d/dx] \equiv x(d/dx) - (d/dx)x \neq 0$, for which $f = c \, x^a$ solves $x \, df/dx = a \, f$ *vs.* only $f = 0$ solves $d(xf)/dx = a \, f$, and similarly, with general matrix operators, $[A,B]f \equiv (AB-BA)f \neq 0$ generally, *i.e.* non-commutative. The associative interpretation must also be used in the commutator, as $(d/dx)(xf) = f + x \, df/dx$ and not $((d/dx) x)f = f$ demonstrates.

Appendix 6

For any molecular system with N electrons, the magnitude [7,20] of the 'correlation energy' is estimated in Equation 18, a negative value according to the variation principle in HF-SCF. It is about 1-2 % of the ground state $E_{electr,0}$, but depends on the nuclear frame and is far above the chemical accuracy (1 kcal/mol). In the HF-SCF formalism, the 'correlation energy' is defined [5] with the energy optimized ground state Slater determinant $E_{electr,0} = \langle S_0|HS_0\rangle + E_{corr}$, *i.e.*, it comes from the error in the approximation $\Psi_0 \approx S_0$. Besides the E_{corr}, there is a basis set error as well. A typical calculation for E_{corr} was, *for example,* the emblematic Moller-Plesset method [5], but it did not bring the long-awaited accuracy and DFT was born. In the KS formalism [1,2], the approximation of kinetic energy (T) is responsible for the 'correlation' energy, notated E_c, for

$$T \equiv \langle\Psi_0|H_\nabla\Psi_0\rangle \approx T_s[\rho] \equiv (-\tfrac{1}{2})\sum_1^N\int\varphi_i\nabla^2\varphi_i d\mathbf{r}_1 \qquad (20)$$

The approximation [1,2] for electron-electron repulsion (V_{ee}) is responsible for 'exchange' energy, notated E_x, and the nuclear-electron attraction (V_{ne}) is 100% accurate if the real (accurate) ρ is used in Equation 21:

$$V_{ee} \equiv \langle\Psi_0|H_{ee}\Psi_0\rangle \approx \tfrac{1}{2}\int\rho(1)\rho(2) \, r_{12}^{-1} \, d\mathbf{r}_1 d\mathbf{r}_2 \text{ and } V_{ne}\equiv$$
$$\langle\Psi_0|H_{ne}\Psi_0\rangle = \sum_{C=1}^M\int\rho(1) \, R_{C1}^{-1} \, d\mathbf{r}_1 \qquad (21)$$

The ρ in Equations 20 and 21 is the ground-state one-electron density; it should be indexed as ρ_0. Furthermore, $E_c >0$, $E_x < 0$, and $E_{corr} \approx E_{xc} \equiv E_x+E_c < 0$. The E_{corr}, E_c, and E_x are about in the same magnitude. The KS formalism (calculating E_{xc} during) approximates $E_{electr,0}$ and energy differences better than the HF-SCF (with calculating E_{corr} after) in the routines of today. The price of convenient approximations in Equations 20 and 21 and empirical functionals for the correlation effect (E_{xc}) is that the variation principle does not hold in DFT, *i.e.*, the E_{xc} is non-variational, unless it is built into the algorithm.

Abbreviations

CC = Correlation calculation; CQC = Computational quantum chemistry; CI = Configuration interaction methods; DFT = Density Functional Theory; $E_{total,0}$ = $E_{electr,0}$ + V_{nn}; $F_L(v) \equiv \int_{(0,1)}$

exp($-vt^2$) t^{2L} dt, the Boys function, L ≥ 0 integer; GTO = Primitive Gaussian type atomic orbital, the G_{Ai}(a, nx, ny, nz) in Equation 1 with n = 2; HF-SCF/basis = Hartree-Fock self-consistent field method on a chosen basis set level; h = Hartree; KS formalism = Kohn-Sham formalism; LC = Linear combination; LCAO = Linear combination of atomic orbitals; $\mathbf{R}_A \equiv$ (R_{Ax}, R_{Ay}, R_{Az}) or (x_A, y_A, z_A) = 3D position (spatial) vector of (fixed) nucleus A; $R_{AB} \equiv |\mathbf{R}_A-\mathbf{R}_B|$ = Nucleus-nucleus distance; $R_{Ai} \equiv |\mathbf{R}_A-\mathbf{r}_i|$ = Nucleus-electron distance; $\mathbf{r}_i \equiv$ (x_i, y_i, z_i) = 3D position (spatial) vector of (moving) electron i; $r_{ij} \equiv |\mathbf{r}_i-\mathbf{r}_j|$ = Electron-electron distance; $\rho(i) \equiv \rho(\mathbf{r}_i)$: Real$^3 \rightarrow$ Real, the one-electron density, the case (n, m)=(0,0) in the main title reduces to the normalization as $(\int \rho(1)d\mathbf{r}_1)^i = N^i$ for i = 1, 2 and 3, respectively, where N is the number of electrons and M is the number of nuclei in the system (with atomic masses m_1, m_2, ..., m_M); STO = Slater-type atomic orbital, the G_{Ai}(a, nx, ny, nz) in Equation 1 with n = 1; ZPE = Zero point energy.

Acknowledgement

Disclosure statement DS

Conflict of interests: The author declares that they have no conflict of interest. Ethical approval: All ethical guidelines have been adhered.

ORCID iD and Email ✉

Sandor Kristyan

✉ kristyan.sandor@ttk.mta.hu

✉ kristyan.sandor@ttk.hu

iD https://orcid.org/0000-0003-4169-2392

References

[1]. Parr, R. G.; Weitao, Y. *Density-functional theory of atoms and molecules*; Oxford University Press: New York, NY, 1994.
[2]. Koch, W.; Holthausen, M. C. *A chemist's guide to density functional theory*; Wiley, 2001.
[3]. Kristyan, S. Variational calculation with general density functional to solve the electronic Schrödinger equation directly for ground state: a recipe for self-consistent field solution. *J. Theor. Appl. Phys.* **2013**, *7*, 61.
[4]. Kristyan, S. Properties of the multi-electron densities "between" the Hohenberg–Kohn theorems and variational principle. *Theochem* **2008**, *858*, 1–11.
[5]. Szabó, A. *Modern quantum chemistry: Introduction to advanced electronic structure theory*; Collier Macmillan: New York, 1982.
[6]. Klopper, W.; Manby, F. R.; Ten-No, S.; Valeev, E. F. R12 methods in explicitly correlated molecular electronic structure theory. *Int. Rev. Phys. Chem.* **2006**, *25*, 427–468.
[7]. Kristyán, S. Immediate estimation of correlation energy for molecular systems from the partial charges on atoms in the molecule. *Chem. Phys.* **1997**, *224*, 33–51.
[8]. Kristyán, S.; Pulay, P. Can (semi)local density functional theory account for the London dispersion forces? *Chem. Phys. Lett.* **1994**, *229*, 175–180.
[9]. Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, *88*, 2547–2553.
[10]. Lebedev, V. I. Über Quadraturen auf der Sphäre. *Zh. Vychisl. Mat. Mat. Fiz* **1976**, *16*, 293–306.
[11]. Lebedev, V. I. Quadrature formulas for the sphere of 25th to 29th order accuracy. *Sibirsk. Mat. Zh.* **1977**, *18*, 132–142.

[12]. Kristyan, S. Estimating correlation energy and basis set error for Hartree–Fock-SCF calculation by scaling during the SCF subroutine with inserting only a few program lines, using analytical integration and no extra CPU time and no extra disc space. *Comput. Theor. Chem.* **2011**, *975*, 20–23.
[13]. Kristyan, S. Solving the non-relativistic electronic Schrödinger equation with switching the electron-electron Coulomb integrals off and on. In *International Conference Of Numerical Analysis And Applied Mathematics ICNAAM 2019*; AIP Publishing: https://doi.org/10.1063/5.0026476, 2293, 420005, 2020.
[14]. Kristyan, S. Analytic evaluation for integrals of product Gaussians with different moments of distance operators (RC1−nRD1−m, RC1−nr12−m and r12−n r13−m with n, m=0,1,2), useful in Coulomb integrals for one, two and three-electron operators. In *AIP Conference Proceedings*; Author(s): https://doi.org/10.1063/1.5044100, 1978, 470030, 2018.
[15]. Reine, S.; Helgaker, T.; Lindh, R. Multi-electron integrals. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 290–303.
[16]. Kristyan, S. Theory of variational calculation with a scaling correct moment functional to solve the electronic schrödinger equation directly for ground state one-electron density and electronic energy. *Int. J. Quantum Chem.* **2013**, *113*, 1479–1492.
[17]. Kristyan, S. Numerical evaluation of Coulomb integrals for 1, 2 and 3-electron distance operators, RC1-Nrd1-M, RC1-Nr12-M and R12-Nr13-M with real (N, M) and the Descartes product of 3 dimension common density functional numerical integration scheme. In *AIP Conference Proceedings*; AIP Publishing: https://doi.org/10.1063/1.5114496, 2116, 450029, 2019.
[18]. Kristyán, S. Note on the choice of basis set in density functional theory calculations for electronic structures of molecules (test on the atoms from the first three rows of the periodic table (2 ⩽ N ⩽ Z ⩽ 18), water, ammonia and pyrrole). *Chem. Phys. Lett.* **1995**, *247*, 101–111.
[19]. Peng, Z.; Kristyan, S.; Kuppermann, A.; Wright, J. S. Excited electronic potential-energy surfaces and transition moments for theH3system. *Phys. Rev. A* **1995**, *52*, 1005–1023.
[20]. Kristyán, S. Statistical analysis of the dependence of correlation energy and zero point energy on the nuclear frame and number of electrons in molecular systems. *Theochem* **2004**, *712*, 153–158.
[21]. Kristyán, S. How can (semi)local density functional theory account for the ground-state total energy of highly ionized atoms of the first three periods in the periodic table? *J. Chem. Phys.* **1995**, *102*, 278–284.
[22]. Ruzsinszky, A.; Kristyán, S.; Margitfalvi, J. L.; Csonka, G. I. Rapid estimation of zero-point energies of molecules using Hartree–Fock atomic partial charges. *J. Phys. Chem. A* **2003**, *107*, 1833–1839.
[23]. Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 theory for molecular energies of first- and second-row compounds. *J. Chem. Phys.* **1991**, *94*, 7221–7230.
[24]. Davidson, E. R.; Hagstrom, S. A.; Chakravorty, S. J.; Umar, V. M.; Fischer, C. F. Ground-state correlation energies for two- to ten-electron atomic ions. *Phys. Rev. A* **1991**, *44*, 7071–7083.
[25]. Kristyan, S. Interesting properties of Thomas–Fermi kinetic and Parr electron–electron-repulsion DFT energy functional generated compact one-electron density approximation for ground-state electronic energy of molecular systems. *J. Comput. Chem.* **2009**, *30*, 1445–1453.
[26]. Ratuchne, F.; Celeste, R. Neutral atoms and ion energies, accurate ionization potential, and electron affinities by polynomial generator coordinate Hartree–Fock method. *Turk. J. Chem.* **2018**, *42*, 1678–1686.
[27]. Irikura, K. K. Experimental vibrational zero-point energies: Diatomic molecules. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389–397.
[28]. Milgrom, M. A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *Astrophys. J.* **1983**, *270*, 365–370.
[29]. Mordell, L. J. On the Rational Solutions of the Indeterminate Equations of the Third and Fourth Degrees. *Proc. Cambridge Philos. Soc.* **1922**, *21*, 179–192.
[30]. Faltings, G. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Inventiones Mathematicae* **1983**, *73*, 349–366.
[31]. Weizsacker, C. F. v. Zur Theorie der Kernmassen. *Eur. Phys. J. A* **1935**, *96*, 431–458.