Molecular Integral Evaluation

The 10th Sostrup Summer School
Quantum Chemistry and Molecular Properties
June 29 – July 11, 2008

Trygve Helgaker
Centre for Theoretical and Computational Chemistry,
Department of Chemistry, University of Oslo, Norway
Molecular integral evaluation

- one-electron interactions
  - overlap, multipole-moment, momentum, and kinetic-energy integrals
  - Coulomb attraction integrals
    \[ O_{\mu\nu} = \int \chi_\mu(\mathbf{r})\hat{O}(\mathbf{r})\chi_\nu(\mathbf{r}) \, d\mathbf{r} \]

- two-electron interactions
  - Coulomb integrals
    \[ g_{\mu\nu\lambda\sigma} = \int \int \frac{\chi_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)\chi_\lambda(\mathbf{r}_2)\chi_\sigma(\mathbf{r}_2)}{r_{12}} \, d\mathbf{r}_1 \, d\mathbf{r}_2 \]
  - Coulomb potential
    \[ J_{\mu\nu} = \sum_{\lambda\sigma} g_{\mu\nu\lambda\sigma} D_{\lambda\sigma} \]

- one-electron functions
  - primitive Cartesian GTOs

- integration schemes
  - McMurchie–Davidson, Obara–Saika and Rys schemes
Overview

- Cartesian and Hermite Gaussians
  - properties of Cartesian Gaussians
  - properties of Hermite Gaussians
- Simple one-electron integrals
  - Gaussian product rule and overlap distributions
  - overlap distributions expanded in Hermite Gaussians
  - overlap and kinetic-energy integrals
- Coulomb integrals
  - Gaussian electrostatics and the Boys function
  - one- and two-electron Coulomb integrals over Cartesian Gaussians
- Sparsity and screening
- Coulomb potential
  - early density matrix contraction
  - density fitting
Primitive Cartesian Gaussians centered at $\mathbf{A}$:

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

where $a > 0$ is the orbital exponent, $r_A = r - \mathbf{A}$ are the electronic coordinates, $i \geq 0$, $j \geq 0$, $k \geq 0$ are the quantum numbers, and $l = i + j + k \geq 0$ is the total angular-momentum quantum number.

The full set of Gaussians of a given $a$ and $l$ constitutes a *shell* of GTOs:

- Each shell contains $(l + 1)(l + 2)/2$ GTOs
  - **s shell**: $G_{000}$
  - **p shell**: $G_{100}, G_{010}, G_{001}$
  - **d shell**: $G_{200}, G_{110}, G_{101}, G_{020}, G_{011}, G_{002}$
We only discuss integration over primitive Cartesian GTOs:

\[ G_{ijk}(r_A, a) = x_A^i y_A^j z_A^k \exp(-ar_A^2) \]

Final integrals are needed over contracted spherical-harmonic GTOs:

\[ \chi_{lm}(r_A) = S_{lm}(r_A) \sum_\mu d_\mu \exp(-a_\mu r_A^2) \]

This may be accomplished by a simple linear transformation

\[ \chi_{lm}(r_A) = \sum_{ijk} C_{ijk}^{lm} \sum_\mu d_\mu G_{ijk}(r_A, a_\mu) \]

The number of contracted spherical-harmonic GTOs is usually much smaller than the number of primitive Cartesian GTOs.

Properly implemented, the transformation to a contracted spherical-harmonic basis reduces CPU time (relative to that required for the corresponding primitive Cartesian basis).
Properties of Cartesian Gaussians

- Definition of a primitive Cartesian Gaussian:
  \[ G_{ijk}(a, \mathbf{r}_A) = x^i_A y^j_A z^k_A \exp(-a r^2_A), \quad \mathbf{r}_A = \mathbf{r} - \mathbf{A} \]

- Each Gaussian factorizes in the Cartesian directions:
  \[ G_{ijk}(a, \mathbf{r}_A) = G_i(a, x_A) G_j(a, y_A) G_k(a, z_A) \]
  \[ G_i(a, x_A) = x^i_A \exp(-a x^2_A) \]
  – note: this is not true for solid-harmonic Gaussians nor for STOs

- Differentiation of Gaussians:
  \[ \frac{\partial G_i(a, x_A)}{\partial x} = - \frac{\partial G_i(a, x_A)}{\partial x} = 2a G_{i+1}(a, x_A) - i G_{i-1}(a, x_A) \]
  – a linear combination two undifferentiated Gaussians

- Gaussian recurrence relation:
  \[ x_A G_i(a, x_A) = G_{i+1}(a, x_A) \]
Hermite Gaussians

- The Hermite Gaussians are defined as
  \[
  \Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) = \left( \frac{\partial}{\partial P_x} \right)^t \left( \frac{\partial}{\partial P_y} \right)^u \left( \frac{\partial}{\partial P_z} \right)^v \exp \left( -pr_P^2 \right), \quad \mathbf{r}_P = \mathbf{r} - \mathbf{P}
  \]

- Like Cartesian Gaussians, they also factorize in the Cartesian directions:
  \[
  \Lambda_t(x_P) = \left( \frac{\partial}{\partial P_x} \right)^t \exp \left( -px_P^2 \right) \leftarrow \text{times a polynomial of degree } t
  \]

- The Hermite Gaussians are not much used as basis functions by themselves
  - rather, they are used as intermediates in the evaluation of Gaussian integrals

  - useful for integration
    \[
    \int_{-\infty}^{\infty} \Lambda_t(x) \, dx = \delta_{t0} \sqrt{\frac{\pi}{p}}
    \]
    - McMurchie and Davidson (1978)

- We shall make extensive use of the McMurchie–Davidson scheme for integration
Hermite recurrence relation

- An important property of the Hermite Gaussians is the recurrence relation

\[ x_P \Lambda_t = \frac{1}{2p} \Lambda_{t+1} + t \Lambda_{t-1} \]

- The proof is simple:
  - from the definition of Hermite Gaussians, we have

\[ \Lambda_{t+1} = \left( \frac{\partial}{\partial P_x} \right)^t \frac{\partial}{\partial P_x} \exp(-px^2_P) = 2p \left( \frac{\partial}{\partial P_x} \right)^t x_P \Lambda_0 \]

  - inserting the identity

\[ \left( \frac{\partial}{\partial P_x} \right)^t x_P = x_P \left( \frac{\partial}{\partial P_x} \right)^t - t \left( \frac{\partial}{\partial P_x} \right)^{t-1} \]

  we then obtain

\[ \Lambda_{t+1} = 2p \left[ x_P \left( \frac{\partial}{\partial P_x} \right)^t - t \left( \frac{\partial}{\partial P_x} \right)^{t-1} \right] \Lambda_0 = 2p \left( x_P \Lambda_t - t \Lambda_{t-1} \right) \]
Integration over Hermite Gaussians

- From the definition of Hermite Gaussians, we have

\[
\int_{-\infty}^{\infty} \Lambda_t(x) \, dx = \int_{-\infty}^{\infty} \left( \frac{\partial}{\partial P_x} \right)^t \exp \left( -px^2 \right) \, dx
\]

- We now change the order of differentiation and integration (Leibniz’ rule):

\[
\int_{-\infty}^{\infty} \Lambda_t(x) \, dx = \left( \frac{\partial}{\partial P_x} \right)^t \int_{-\infty}^{\infty} \exp \left( -px^2 \right) \, dx
\]

- The Gaussian integral is given by

\[
\int_{-\infty}^{\infty} \exp \left( -px^2 \right) \, dx = \sqrt{\frac{\pi}{p}}
\]

- Since the integral is independent of \( P \), differentiation with respect to \( P \) gives zero:

\[
\int_{-\infty}^{\infty} \Lambda_t(x) \, dx = \delta_{t0} \sqrt{\frac{\pi}{p}}
\]

- only integrals over \( s \) functions do not vanish
## Comparison of Cartesian and Hermite Gaussians

<table>
<thead>
<tr>
<th>Cartesian Gaussians</th>
<th>Hermite Gaussians</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>definition</strong></td>
<td>$G_i = x^i_A \exp(-ax^2_A)$</td>
</tr>
<tr>
<td><strong>recurrence</strong></td>
<td>$x_A G_i = G_{i+1}$</td>
</tr>
<tr>
<td><strong>differentiation</strong></td>
<td>$\frac{\partial G_i}{\partial A_x} = 2aG_{i+1} - iG_{i-1}$</td>
</tr>
</tbody>
</table>
The Gaussian product rule

- The product of two Gaussians is another Gaussian centered somewhere on the line segment connecting the original Gaussians; its exponent is the sum of the original exponents:

\[
\exp(-ax_A^2) \exp(-bx_B^2) = \exp(-\mu X_{AB}^2) \exp(-px_P^2)
\]

where

- \( P_x = \frac{aA_x + bB_x}{p} \) \( \leftarrow \) “center of mass”
- \( p = a + b \) \( \leftarrow \) total exponent
- \( X_{AB} = A_x - B_x, \) \( \leftarrow \) relative separation
- \( \mu = \frac{ab}{a + b} \) \( \leftarrow \) reduced exponent

- The Gaussian product rule greatly simplifies the calculation of integrals
  - two-center integrals are reduced to one-center integrals
  - four-center integrals are reduced to two-center integrals
Overlap distributions

- The product of two Cartesian Gaussians is known as an overlap distribution:
  \[ \Omega_{ij}(x) = G_i(x, a, A_x)G_j(x, b, B_x) \]
- The Gaussian product rule reduces two-center integrals to one-center integrals
  \[
  \int \Omega_{ij}(x) \, dx = K_{AB} \int x_A^i x_B^j \exp(-px_P^2) \, dx
  \]
  - the Cartesian monomials still make the integration awkward
  - we would like to utilize the simple integration properties of Hermite Gaussians
- To simplify integration, we expand overlap distributions in Hermite Gaussians
  overlap distribution \( \rightarrow \) \( \Omega_{ij}(x) = \sum_{t=0}^{i+j} E_t^{ij} \Lambda_t(x_P) \) \( \leftarrow \) in Hermite Gaussians
  - note: \( \Omega_{ij}(x) \) is a single Gaussian times a polynomial in \( x \) of degree \( i + j \)
  - it may be exactly represented as a linear combination of \( \Lambda_t \) with \( 0 \leq t \leq i + j \)
- Recurrence relations may be used for the expansion coefficients
  \[
  E_t^{i+1,j} = \frac{1}{2p} E_{t-1}^{ij} + XPA E_t^{ij} + (t + 1) E_{t+1}^{ij}
  \]
Recurrence relations for Hermite coefficients

- Hermite expansion of $\Omega_{i+1,j}$:
  \[
  \Omega_{i+1,j} = K_{AB} x_A^{i+1} x_B^j \exp \left( -pr_P^2 \right) = \sum_t E_{t}^{i+1,j} \Lambda_t
  \]

- Alternative Hermite expansion of $\Omega_{i+1,j}$ (with use of $x_P \Lambda_t = \frac{1}{2p} \Lambda_{t+1} + t \Lambda_{t-1}$)
  \[
  \Omega_{i+1,j} = x_A \Omega_{ij} = (x - A_x) \Omega_{ij}
  \]
  \[
  = (x - P_x) \Omega_{ij} + (P_x - A_x) \Omega_{ij} = x_P \Omega_{ij} + X_{PA} \Omega_{ij}
  \]
  \[
  = \sum_t E_{t}^{ij} x_P \Lambda_t + X_{PA} \sum_t E_{t}^{ij} \Lambda_t
  \]
  \[
  = \sum_t E_{t}^{ij} \left( \frac{1}{2p} \Lambda_{t+1} + t \Lambda_{t-1} + X_{PA} \Lambda_t \right)
  \]
  \[
  = \sum_t \left[ \frac{1}{2p} E_{t-1}^{ij} + X_{PA} E_{t}^{ij} + (t+1) E_{t+1}^{ij} \right] \Lambda_t
  \]

- A comparison of the two expansions yields the recurrence relations
  \[
  E_{t}^{i+1,j} = \frac{1}{2p} E_{t-1}^{ij} + X_{PA} E_{t}^{ij} + (t+1) E_{t+1}^{ij}
  \]
Overlap integrals

\( S_{ab} = \langle G_a | G_b \rangle \)

\( G_a = G_{ikm}(r, a, A) = G_i(x_A)G_k(y_A)G_m(z_A) \)

\( G_b = G_{jln}(r, b, B) = G_j(x_B)G_l(y_B)G_n(z_B) \)

The overlap integral factorizes in the Cartesian directions:

\( S_{ab} = S_{ij}S_{kl}S_{mn}, \quad S_{ij} = \langle G_i(x_A) | G_j(x_B) \rangle \)

Use the Gaussian product rule and Hermite expansion:

\[
S_{ij} = \int \Omega_{ij}(x) \, dx = \sum_{t=0}^{i+j} E_{t}^{ij} \int \Lambda_t(x_P) \, dx = \sum_{t=0}^{i+j} E_{t}^{i+j} \delta_{t0} \sqrt{\frac{\pi}{p}} = E_{0}^{ij} \sqrt{\frac{\pi}{p}}
\]

– only one term survives

The total overlap integral is therefore given by

\[
S_{ab} = E_{0}^{ij} E_{0}^{kl} E_{0}^{mn} \left( \frac{\pi}{p} \right)^{3/2}
\]
Dipole integrals

- Multipole integrals follow the same scheme as overlap integrals

\[ D_{ij} = \langle G_i(x_A) | x_C | G_j(x_B) \rangle = \int \Omega_{ij}(x_P)x_C dx \]

\[ = \sum_{t=0}^{i+j} E_t^{ij} \int \Lambda_t(x_P)x_c dx \quad \text{(use } x_C = x_P + X_{PC}) \]

\[ = \sum_{t=0}^{i+j} E_t^{ij} \int [x_P \Lambda_t(x_P) + X_{PC} \Lambda_t(x_P)] dx \quad \text{(use } x_P \Lambda_t = \frac{1}{2p} \Lambda_{t+1} + t \Lambda_{t-1}) \]

\[ = \sum_{t=0}^{i+j} E_t^{ij} \int \left[ \frac{1}{2p} \Lambda_{t+1}(x_P) + X_{PC} \Lambda_t(x_P) + t \Lambda_{t-1}(x_P) \right] dx \]

\[ = E_0^{ij} X_{PC} \sqrt{\frac{\pi}{p}} + E_1^{ij} \sqrt{\frac{\pi}{p}} \]

- Final dipole integral

\[ D_{ij} = \left( E_1^{ij} + X_{PC} E_0^{ij} \right) \sqrt{\frac{\pi}{p}} \]
Kinetic-energy integrals

- Kinetic-energy integral

\[ T_{ab} = -\frac{1}{2} \left\langle G_a \left| \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right| G_b \right\rangle \]

\[ T_{ab} = T_{ij} S_{kl} S_{mn} + S_{ij} T_{kl} S_{mn} + S_{ij} S_{kl} T_{mn} \]

\[ S_{ij} = \left\langle G_i(x_A) | G_j(x_B) \right\rangle \]

\[ T_{ij} = -\frac{1}{2} \left\langle G_i(x_A) \left| \frac{\partial^2}{\partial x^2} \right| G_j(x_B) \right\rangle \]

- Differentiation of Cartesian Gaussians

\[ \frac{\partial}{\partial x} G_j(x_B) = -2bG_{j+1} + jG_{j-1} \]

\[ \frac{\partial^2}{\partial x^2} G_j(x_B) = 4b^2G_{j+2} - 2b(2j+1)G_j + j(j-1)G_{j-2} \]

- the second derivative is linear combination of three undifferentiated Gaussians

- Kinetic-energy integrals are linear combinations of overlap integrals

\[ T_{ij} = 4b^2 S_{i,j+2} - 2b(2j+1)S_{i,j} + j(j-1)S_{i,j-2} \]
Review of simple one-electron integrals

- Gaussian product rule and expansion of overlap distributions

\[
\exp(-ax_A^2) \exp(-bx_B^2) = K_{AB} \exp(-px_P^2)
\]

\[
G_i(x_A)G_j(x_B) = \sum_{t=0}^{i+j} E_t^{ij} \Lambda_t(x_P)
\]

- Overlap and dipole integrals

\[
S_{ab} = E_0^{ij} E_0^{kl} E_0^{mn} \left( \frac{\pi}{p} \right)^{3/2}
\]

\[
X_{ab} = \left( E_1^{ij} + X_{PC} E_0^{ij} \right) E_0^{kl} E_0^{mn} \left( \frac{\pi}{p} \right)^{3/2}
\]

- Kinetic-energy integrals by differentiation

- We shall now consider the Coulomb integrals ...
Coulomb integrals

- One-electron nuclear-attraction integrals
  \[ \langle G_a(r_A) | r_C^{-1} | G_b(r_B) \rangle \]

- Two-electron repulsion integrals
  \[ \langle G_a(r_{1A})G_b(r_{1B}) | r_{12}^{-1} | G_c(r_{2C})G_d(r_{2D}) \rangle \]

- Coulomb integrals are not separable in the Cartesian directions

- The Gaussian Coulomb integrals may be reduced to a one-dimensional finite integral: the Boys function
  \[ F_n(x) = \int_0^1 \exp (-xt^2) t^{2n} \, dt \]
  - makes the evaluation of Gaussian Coulomb integrals relatively simple

- We begin by evaluating the simplest Coulomb integral:
  \[ V_p = \int r_C^{-1} \exp (-pr_P^2) \, dr = \frac{2\pi}{p} F_0 \left( pR_{PC}^2 \right) \]
  - we shall then go on to consider the usual Coulomb integrals
Coulomb integral over a spherical Gaussian I

\[ V_p = \int r_C^{-1} \exp (-pr_P^2) \, dr \]

1. The presence of \( r_C^{-1} \) is awkward and is avoided by the substitution

\[ \frac{1}{r_C} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp (-r_C^2 t^2) \, dt \]

yielding the four-dimensional integral

\[ V_p = \frac{1}{\sqrt{\pi}} \int \exp (-pr_P^2) \int_{-\infty}^{\infty} \exp (-t^2 r_C^2) \, dt \, dr \]

2. To prepare for integration over \( r \), we invoke the Gaussian product rule

\[ \exp (-pr_P^2) \exp (-t^2 r_C^2) = \exp \left( -\frac{pt^2}{p + t^2} R_{CP}^2 \right) \exp \left[ - (p + t^2) r_S^2 \right] \]

to obtain (where the exact value of \( S \) in \( r_S \) does not matter):

\[ V_p = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \int \exp \left[ - (p + t^2) r_S^2 \right] \, dr \exp \left( -\frac{pt^2}{p + t^2} R_{CP}^2 \right) \, dt \]
Coulomb integral over a spherical Gaussian II

\[ V_p = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \int [-(p + t^2) r_S^2] \, dr \exp\left(-\frac{pt^2}{p + t^2} R_{CP}^2 \right) \, dt \]

3. Integration over all space \( r \) now yields a one-dimensional integral

\[ V_p = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \left( \frac{\pi}{p + t^2} \right)^{3/2} \exp\left(-pR_{CP}^2 \frac{t^2}{p + t^2} \right) \, dt \]

4. To introduce a finite integration range, we perform the substitution

\[ u^2 = \frac{t^2}{p + t^2} \Rightarrow 1 + pt^{-2} = u^{-2} \Rightarrow pt^{-3} \, dt = u^{-3} \, du \Rightarrow dt = p^{-1} \left( \frac{t^2}{u^2} \right)^{3/2} \, du \]

and obtain

\[ V_p = \frac{2\pi}{p} \int_0^1 \exp\left(-pR_{CP}^2 u^2 \right) \, du = \frac{2\pi}{p} F_0 \left(pR_{PC}^2 \right) \]

- We have here introduced the Boys function

\[ F_n(x) = \int_0^1 \exp \left(-xt^2 \right) t^{2n} \, dt \]

- a 3D integral over all space has been reduced to a 1D integral over \([0, 1]\)
The Boys function $F_n(x)$

- $F_n(x) > 0$ since the integrand is positive:
  \[ F_n(x) = \int_0^1 \exp(-xt^2) t^{2n} \, dt > 0 \]

- $F_n(x)$ is convex and decreasing:
  \[ F'_n(x) = -F_{n+1}(x) < 0 \]
  \[ F''_n(x) = F_{n+2}(x) > 0 \]

- Small and large values:
  \[ F_n(x) = \frac{1}{2n+1} + \sum_{k=1}^\infty \frac{(-x)^k}{k!(2n+2k+1)} \quad (x \text{ small}) \]
  \[ F_n(x) \approx \int_0^\infty \exp(-xt^2) t^{2n} \, dt = \frac{(2n-1)!!}{2^{n+1}} \sqrt{\frac{\pi}{x^{2n+1}}} \quad (x \text{ large}) \]

- Downward recurrence relation:
  \[ F_{n-1}(x) = \frac{2xF_n(x) + \exp(-x)}{2n-1} \]
Gaussian electrostatics

- Spherical-Gaussian integrals are nicely expressed in terms of the error function

\[ \int r_C^{-1} \exp \left(-p r_P^2\right) \, dr = \frac{2\pi}{p} F_0 \left(p R_{PC}^2\right) \]

\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp \left(-t^2\right) \, dt = \frac{2}{\sqrt{\pi}} x F_0 \left(x^2\right) \]

- Introducing unit charge distributions and the reduced exponent as

\[ \rho_p(r_P) = \left(\frac{p}{\pi}\right)^{3/2} \exp \left(-p r_P^2\right), \quad \alpha = \frac{pq}{p + q} \]

we obtain formulas similar to those of electrostatics (but damped)

\[ \int \frac{\rho_p(r_P)}{r_C} \, dr = \frac{\text{erf} \left(\sqrt{p} R_{PC}\right)}{R_{PC}} \]

\[ \int \int \frac{\rho_p(r_{1P}) \rho_q(r_{2Q})}{r_{12}} \, dr_1 dr_2 = \frac{\text{erf} \left(\sqrt{\alpha} R_{PQ}\right)}{R_{PQ}} \]
Cartesian Coulomb integrals

- We have obtained a simple result for one-center spherical Gaussians:

\[
\int r_C^{-1} \exp\left(-pr_C^2\right) \, dr = \frac{2\pi}{p} F_0 \left(pR_{PC}^2\right)
\]

- We shall now consider general two-center one-electron Coulomb integrals:

\[
V_{ab} = \langle G_a \mid r_C^{-1} \mid G_b \rangle = \int \frac{\Omega_{ab}(r)}{r_C} \, dr
\]

- The overlap distribution expanded in Hermite Gaussians

\[
\Omega_{ab}(r) = \Omega_{ij}(x)\Omega_{kl}(y)\Omega_{mn}(z), \quad \Omega_{ij}(x) = \sum_{t=0}^{i+j} E_{ij}^t \Lambda_t(x_P)
\]

- The total overlap distribution may therefore be written in the form

\[
\Omega_{ab}(r) = \sum_{tuv} E_{ij}^t E_{kl}^u E_{mn}^v \Lambda_{tuv}(r_P) = \sum_{tuv} E_{tuv}^{ab} \Lambda_{tuv}(r_P)
\]

- Expansion of two-center Coulomb integrals in one-center Hermite integrals:

\[
\text{two-center integral} \rightarrow V_{ab} = \sum_{tuv} E_{tuv}^{ab} \int \frac{\Lambda_{tuv}(r_P)}{r_C} \, dr \leftarrow \text{one-center integrals}
\]
Coulomb integrals over Hermite Gaussians

- Expansion of two-center Coulomb integrals in one-center Hermite integrals

\[
V_{ab} = \sum_{tuv} E_{tuv}^{ab} \int \frac{\Lambda_{tuv}(r_P)}{r_C} \, dr, \quad \int \frac{\exp(-pr_P^2)}{r_C} \, dr = \frac{2\pi}{p} F_0(pR_{PC}^2)
\]

- Changing the order of integration and differentiation by Leibniz rule, we obtain

\[
\int \frac{\Lambda_{tuv}(r_P)}{r_C} \, dr = \frac{\partial^{t+u+v}}{\partial P_x \partial P_y \partial P_z} \int \frac{\exp(-pR_P^2)}{r_C} \, dr \\
= \frac{2\pi}{p} \frac{\partial^{t+u+v} F_0(pR_{PC}^2)}{\partial P_x \partial P_y \partial P_z} = \frac{2\pi}{p} R_{tuv}(p, R_{PC})
\]

where the one-center Hermite Coulomb integral is given by

\[
R_{tuv}(p, R_{PC}) = \frac{\partial^{t+u+v} F_0(pR_{PC}^2)}{\partial P_x \partial P_y \partial P_z}
\]

- the Hermite integrals are derivatives of the Boys function
Evaluation of Hermite Coulomb integrals

- The Hermite integrals are derivatives of the Boys function
  - can be obtained by repeated differentiation, using $F'_n(x) = -F_{n+1}(x)$
  - recursion is simpler and more efficient
- To set up recursion, we introduce the auxiliary Hermite integrals
  \[ R^n_{tuv}(p, P) = (-2p)^n \frac{\partial^{t+u+v} F_n \left( pR^2_{PC} \right)}{\partial P^t_x \partial P^u_y \partial P^v_z} \]
  which include as special cases the source and target integrals
  \[ R^n_{000} = (-2p)F_n \rightarrow R^0_{tuv} = \frac{\partial^t}{\partial P^t_x} \frac{\partial^u}{\partial P^u_y} \frac{\partial^v}{\partial P^v_z} F_0 \]
- The Hermite integrals can now be generated from the recurrence relations

\[
\begin{align*}
R^n_{t+1,u,v} &= tR^n_{t-1,u,v} + X_{PC} R^n_{tuv} \\
R^n_{t,u+1,v} &= uR^n_{t-1,u,v} + Y_{PC} R^n_{tuv} \\
R^n_{t,u,v+1} &= vR^n_{t,u,v-1} + Z_{PC} R^n_{tuv}
\end{align*}
\]
Final expressions for Coulomb integrals

- One-electron integrals

\[ \langle G_a \mid r_C^{-1} \mid G_b \rangle = \int \frac{\Omega_{ab}}{r_C} \, dr = \sum_{tuv} E_{tuv}^{ab} \int \frac{\Lambda_{tuv}}{r_C} \, dr = \frac{2\pi}{p} \sum_{tuv} E_{tuv}^{ab} R_{tuv}(p, R_{PC}) \]

- Two-electron integrals

\[ \langle G_a(a)G_b(1) \mid r_{12}^{-1} \mid G_c(2)G_d(2) \rangle = \int \int \frac{\Omega_{ab}(1)\Omega_{cd}(2)}{r_{12}} \, dr_1 dr_2 = \sum_{tuv} E_{tuv}^{ab} \sum_{\tau \nu \phi} E_{\tau \nu \phi}^{cd} \int \int \frac{\Lambda_{tuv}(1)\Lambda_{\tau \nu \phi}(2)}{r_{12}} \, dr_1 dr_2 = \frac{2\pi^{5/2}}{pq\sqrt{p + q}} \sum_{tuv} E_{tuv}^{ab} \sum_{\tau \nu \phi} E_{\tau \nu \phi}^{cd} (-1)^{\tau + \nu + \phi} R_{t+\tau, u+\nu, v+\phi}(\alpha, R_{PQ}) \]

- two-electron integrals very similar to one-electron integrals
- Hermit expansion for both electrons, around centers P and Q
- integrals depend on \( R_{PQ} \) and \( \alpha = pq/(p + q) \), with \( p = a + b \) and \( q = c + d \)
Summary one-electron Coulomb integral evaluation

1. Calculate expansion coefficients

\[ E_{00}^0 = \exp \left( -\mu X_{AB}^2 \right) \]
\[ E_{t+1}^{i,j} = \frac{1}{2p} E_t^{ij} + X_{PA} E_t^{ij} + (t + 1) E_t^{ij} \]

2. Calculate the Boys function

\[ F_n(x) = \int_0^1 \exp \left( -x t^2 \right) t^{2n} dt \quad \text{(with } x = pR_{PC}^2) \]
\[ F_n(x) = \frac{2x F_{n+1}(x) + \exp(-x)}{2n + 1} \]

3. Calculate Hermite integrals

\[ R_{000}^n = (-2p)^n F_n \]
\[ R_{t+1,u,v}^n = X_{PC} R_{tuv}^{n+1} + t R_{t-1,u,v}^{n+1} \]

4. Calculate Cartesian integrals

\[ \langle G_a | r_C^{-1} | G_b \rangle = \sum_{tuv} E_t^{ij} E_u^{kl} E_v^{mn} R_{tuv}^0 \]
The general expression for two-electron integrals is given by

\[
g_{abcd} = \frac{2\pi^{5/2}}{pq\sqrt{p+q}} \sum_{tuv} E_{tuv}^{ab} \sum_{\tau\nu\phi} E_{\tau\nu\phi}^{cd} (-1)^{\tau+\nu+\phi} R_{t+\tau,u+\nu,v+\phi}(\alpha, \mathbf{R}_{PQ})
\]

In SCF theories, these integrals make contributions to the Fock/KS matrix:

\[
F_{ab} = \sum_{cd} (2g_{abcd} D_{cd} - g_{acbd} D_{cd})
\]

– in early days, one would write all integrals to disk and read back as required
– in direct SCF theories, integrals are calculated as needed
– this development (1980) made much larger calculations possible

Many developments have since improved the efficiency of direct SCF
– efficient screening of integrals
– early contraction with density matrices (Coulomb)
– density fitting (Coulomb)
– multipole methods (Coulomb)
Scaling properties

- Product of two $s$ functions from Gaussian product rule
  \[
  \exp(-ar_A^2) \exp(-br_B^2) = \exp(-ab/(a+b)R_{AB}^2) \exp(-(a+b)r_P^2)
  \]

- Overlap integral between two $s$ orbitals:
  \[
  S_{ab} = \left( \frac{\pi}{a+b} \right)^{3/2} \exp\left( -\frac{ab}{a+b}R_{AB}^2 \right)
  \]
  - the number of such integrals scales quadratically with system size
  - however, $S_{ab}$ decreases rapidly with $R_{AB}$

- Let us assume that we may neglect all integrals
  \[
  |S_{ab}| < 10^{-k} \quad \leftarrow \text{insignificant integrals}
  \]

- We may neglect all integrals separated by more than
  \[
  R_{AB} > \sqrt{a_{\min}^{-1} \ln\left[\left(\frac{\pi}{2a_{\min}}\right)^3 10^{2k}\right]}
  \]
  - in a large system, most integrals becomes small and may be neglected
  - the number of significant integrals increases linearly with system size
The sparsity of overlap and electron-repulsion integrals

- linear system of 16 Gaussian 1s functions of unit exponent, separated by $1a_0$
Two-electron integral scaling

- Two-electron $ssss$ integrals:
  \[
  g_{abcd} = \text{erf} \left( \sqrt{\alpha R_{PQ}} \right) \frac{S_{ab}S_{cd}}{R_{PQ}}
  \]
  - the total number of integrals scales quartically with system size
  - the number of significant integrals scales quadratically
  \[
  S_{ab}, S_{cd} \to 0 \text{ rapidly}
  \]
  \[
  R_{PQ}^{-1} \to 0 \text{ very slowly}
  \]

- Decompose integral into classical and nonclassical parts:
  \[
  g_{abcd} = \left( \frac{S_{ab}S_{cd}}{R_{PQ}} \right) - \text{erfc} \left( \sqrt{\alpha R_{PQ}} \right) \frac{S_{ab}S_{cd}}{R_{PQ}}
  \]
  \[
  \underbrace{\frac{S_{ab}S_{cd}}{R_{PQ}}} \quad \underbrace{- \frac{S_{ab}S_{cd}}{R_{PQ}}} = 
  \]
  classical
  nonclassical
  - quadratic scaling of classical part, can be treated by multipole methods
  - linear scaling of nonclassical part since $S_{ab} \to 0, S_{cd} \to 0$, erfc $\to 0$ rapidly

\[
\text{erfc}(\sqrt{\alpha R_{PQ}}) \leq \frac{\exp(-\alpha R_{PQ}^2)}{\sqrt{\pi \alpha R_{PO}}}
\]
The scaling properties of molecular integrals

- linear system of up to 16 1s GTOs of unit exponent, separated by 1\(a_0\)

### Graphs

#### Overlap
- \(n^1\)
- \(n^2\)
- \(n^3\)

#### Kinetic Energy
- \(n^1\)
- \(n^2\)
- \(n^4\)

#### Nuclear Attraction
- \(n^1\)
- \(n^2\)
- \(n^3\)

#### Electron Repulsion
- \(n^1\)
- \(n^2\)
- \(n^4\)
The sparsity of overlap and electron-repulsion integrals

- linear system of 16 Gaussian 1s functions of unit exponent, separated by 1\(a_0\)
Sparsity in linear alkanes and alkenes

- Percentage of matrix elements greater than $10^{-6}$ in alkane and alkene chains
  
  - **overlap matrix:** sparse
  
  - **density matrix:**
    - nonsparse for alkenes
    - sparse for alkanes
  
  - **Fock/KS matrices:**
    - KS matrix like overlap
    - Fock matrix intermediate between overlap and density matrices
Integral prescreening

- Small integrals ($< 10^{-10}$) are not needed and should be avoided by some prescreening technique.

- The two-electron integrals

$$g_{abcd} = \int \int \frac{\Omega_{a}(1)\Omega_{c}(2)}{r_{12}} \, dr_{1} \, dr_{2}$$

are the elements of a positive definite matrix with diagonal elements

$$g_{ab,ab} \geq 0$$

- The conditions for an inner product are thus satisfied.

- The Cauchy–Schwarz inequality yields

$$\|g_{ab,cd}\| \leq \sqrt{g_{ab,ab}} \sqrt{g_{cd,cd}}$$

- Precalculate

$$G_{ab} = \sqrt{g_{ab,ab}}$$

and prescreen

$$\|g_{ab,cd}\| \leq G_{ab} G_{cd}$$
Calculation of Coulomb potential

- The traditional Coulomb-potential evaluation is a two-step procedure
  1. evaluate significant two-electron integrals (in batches)

\[
g_{abcd} = \frac{2\pi^{5/2}}{pq\sqrt{p+q}} \sum_{tuv} E_{tuv}^{ab} \sum_{\tau\nu\phi} E_{\tau\nu\phi}^{cd} (-1)^{\tau+\nu+\phi} R_{t+\tau, u+\nu, v+\phi}(\alpha, R_{PQ})
\]

  2. add their contribution to the Fock or Kohn–Sham matrix:

\[
J_{ab} = \sum_{cd} g_{abcd} D_{cd}
\]

- The cost of the different steps ($L$ is ang. mom., $p$ number of primitives):

<table>
<thead>
<tr>
<th>Boys</th>
<th>$E$ coef</th>
<th>$R$ int</th>
<th>$\tau\nu\phi$</th>
<th>cont</th>
<th>$S_{lm}$</th>
<th>$tuv$</th>
<th>cont</th>
<th>$S_{lm}$</th>
<th>$J_{ab}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Lp^4$</td>
<td>$L^3p^2$</td>
<td>$L^4p^4$</td>
<td>$L^{10}p^4$</td>
<td>$L^7p^4$</td>
<td>$L^8p^2$</td>
<td>$L^9p^2$</td>
<td>$L^6p^2$</td>
<td>$L^7$</td>
<td>$L^4$</td>
</tr>
</tbody>
</table>

- for low $L$, evaluation of the Boys function dominates
- for large $L$, the $\tau\nu\phi$ contraction dominates

- It is not necessary to assemble the integrals fully before contraction with $D_{cd}$
Early contraction with density matrix for Coulomb potential

- Two-electron Coulomb contribution to Fock/Kohn–Sham matrix:

\[ J_{ab} = \frac{2\pi^{5/2}}{pq\sqrt{p+q}} \sum_{tuv} E_{tuv}^{ab} \sum_{\tau\nu\phi} \sum_{cd} D_{cd} E_{\tau\nu\phi}^{cd} (-1)^{\tau+\nu+\phi} R_{t+\tau,u+\nu,v+\phi}(\alpha, R_{PQ}) \]

- Early contraction of density matrix with \( E \) coefficients:

\[ K_{\tau\nu\phi} = (-1)^{\tau+\nu+\phi} \sum_{cd \in Q} D_{cd} E_{\tau\nu\phi}^{cd} \]

\[ J_{ab} = \frac{2\pi^{5/2}}{pq\sqrt{p+q}} \sum_{tuv} E_{tuv}^{ab} \sum_{Q} \sum_{\tau\nu\phi} K_{\tau\nu\phi}^{Q} R_{t+\tau,u+\nu,v+\phi}(\alpha, R_{PQ}) \]

- The \( \tau\nu\phi \) \((tuv)\) cost reduced from \( L^{10}p^4 \) to \( L^6p^4 \) (from \( L^9p^2 \) to \( L^7p^2 \))

- Timings (seconds) for benzene:

<table>
<thead>
<tr>
<th></th>
<th>cc-pVDZ</th>
<th>cc-pVTZ</th>
<th>cc-pVQZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>late contraction</td>
<td>27</td>
<td>374</td>
<td>3823</td>
</tr>
<tr>
<td>early contraction</td>
<td>10</td>
<td>90</td>
<td>687</td>
</tr>
</tbody>
</table>
Density fitting

- Traditionally, the electron density is expanded in overlap distributions
  \[ \rho(r) = \sum_{ab} D_{ab} \Omega_{ab}(r) \]
  - the number of terms is \( n^2 \), where \( n \) is the number of AOs

- The Coulomb contribution to the Fock/KS matrix is evaluated as
  \[ J_{ab} = \langle \Omega_{ab}(r_1) | r^{-1}_{12} | \rho(r_2) \rangle \]
  - the formal cost of this evaluation is therefore quartic \( (n^4) \)

- Consider now an approximate density \( \tilde{\rho}(r) \), expanded in an auxiliary basis \( \omega_\alpha(r) \):
  \[ \tilde{\rho}(r) = \sum_\alpha c_\alpha \omega_\alpha(r), \quad \tilde{\rho}(r) \approx \rho(r) \]
  - the size of the auxiliary basis \( N \) increases linearly with system size

- We may now evaluate the Coulomb contribution approximately as
  \[ \tilde{J}_{ab} = \langle \Omega_{ab}(r_1) | r^{-1}_{12} | \tilde{\rho}(r_2) \rangle \]
  - the cost of the evaluation is therefore cubic \( (n^2 N) \)
  - if the \( \tilde{\rho} \) is sufficiently accurate and easy to obtain, this may give huge savings
• Consider the evaluation of the Coulomb potential

\[ J_{ab} = \sum_{cd} (ab|cd) D_{cd} \]

• We now introduce an auxiliary basis and invoke the resolution of identity:

\[ (ab|cd) \approx \sum_{\alpha\beta} (ab|\alpha) (\alpha|\beta)^{-1} (\beta|cd) \]

– in a complete auxiliary basis, the two expressions are identical

• We may now calculate the Coulomb potential in two different ways:

\[ J_{ab} = \sum_{cd} (ab|cd) D_{cd} = (ab|\rho) \]

\[ \approx \sum_{cd} \sum_{\alpha\beta} (ab|\alpha) (\alpha|\beta)^{-1} (\beta|cd) D_{cd} = \sum_{\alpha} (ab|\alpha) c_{\alpha} = (ab|\tilde{\rho}) = \tilde{J}_{ab} \]

where the coefficients \( c_{\alpha} \) are obtained from the linear sets of equations

\[ \sum_{\alpha} (\beta|\alpha) c_{\alpha} = \sum_{cd} (\beta|cd) D_{cd} \quad \Leftrightarrow \quad (\beta|\tilde{\rho}) = (\beta|\rho) \]

– no four-center integrals, the formal cost is cubic or smaller

– note: the solution of the linear equations scales cubically

• This particular approach to density fitting is called **Coulomb density fitting**
Robust density fitting

- In Coulomb density fitting, the exact and fitted Coulomb matrices are given by
  \[ J_{ab} = (ab | \rho) , \quad \tilde{J}_{ab} = (ab | \tilde{\rho}) , \quad (\alpha | \rho) = (\alpha | \tilde{\rho}) \]

- We may then calculate the exact and density-fitted Coulomb energies as
  \[ E = \sum_{ab} J_{ab} D_{ab} = (\rho | \rho) , \quad \tilde{E} = \sum_{ab} \tilde{J}_{ab} D_{ab} = (\rho | \tilde{\rho}) \]

- The true Coulomb energy is an upper bound to the density-fitted energy:
  \[ E - \tilde{E} = (\rho | \rho) - (\rho | \tilde{\rho}) = (\rho | \rho) - (\rho | \tilde{\rho}) - (\tilde{\rho} | \rho) + (\tilde{\rho} | \tilde{\rho}) = (\rho - \tilde{\rho} | \rho - \tilde{\rho}) \]

  - the error in the density-fitted energy is quadratic in the error in the density
  - Coulomb fitting is thus equivalent to minimization of \[ E - \tilde{E} = (\rho - \tilde{\rho} | \rho - \tilde{\rho}) \]

- It is possible to determine the approximate density in other ways
  - quadratic energy error is ensured by using the robust formula
    \[ \tilde{E} = (\tilde{\rho} | \rho) + (\rho | \tilde{\rho}) - (\tilde{\rho} | \tilde{\rho}) = E - (\rho - \tilde{\rho} | \rho - \tilde{\rho}) \]

  - however, the error in the energy is no longer minimized
Density fitting: sample calculations

- no symmetry used; exact energy: $-230.208 E_h$

<table>
<thead>
<tr>
<th></th>
<th>cc-pVDZ</th>
<th>cc-pVTZ</th>
<th>cc-pVQZ</th>
<th>cc-pV5Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>exact density late</td>
<td>27</td>
<td>374</td>
<td>3823</td>
<td>na</td>
</tr>
<tr>
<td>exact density early</td>
<td>10</td>
<td>90</td>
<td>687</td>
<td>5320</td>
</tr>
<tr>
<td>fitted density</td>
<td>1</td>
<td>7</td>
<td>34</td>
<td>166</td>
</tr>
<tr>
<td>exact energy</td>
<td>$-230.09671$</td>
<td>$-230.17927$</td>
<td>$-230.19463$</td>
<td>na</td>
</tr>
</tbody>
</table>

- Clearly, very large gains can be achieved with density fitting

- Formal scaling is cubical
  - for large systems, screening yields quadratic scaling for integral evaluation
  - the cubic cost of solving linear equations remains for large systems

- Linear scaling is achieved by boxed density-fitting and fast multipole methods
  - the density is partitioned into boxes, which are fitted one at a time
  - fast multipole methods for matrix elements