Molecular Integral Evaluation

Trygve Helgaker

Centre for Theoretical and Computational Chemistry
Department of Chemistry, University of Oslo, Norway

The 11th Sostrup Summer School
Quantum Chemistry and Molecular Properties
July 4–16 2010

Molecular integral evaluation

- one-electron interactions
    - overlap, multipole-moment, and kinetic-energy integrals
    - Coulomb attraction integrals

$$O_{ab} = \int \chi_a(\mathbf{r}) \hat{O}(\mathbf{r}) \chi_b(\mathbf{r}) \, d\mathbf{r}$$

- two-electron interactions
    - Coulomb integrals

$$g_{abcd} = \iint \frac{\chi_a(\mathbf{r}_1)\chi_b(\mathbf{r}_1)\chi_c(\mathbf{r}_2)\chi_d(\mathbf{r}_2)}{r_{12}} \, d\mathbf{r}_1 \, d\mathbf{r}_2$$

    - Coulomb and exchange contributions to Fock/KS matrix

$$F_{ab} = \sum_{cd} \left(2g_{abcd}D_{cd} - g_{acbd}D_{cd} \right)$$

- basis functions
    - primitive Cartesian GTOs
- integration schemes
    - McMurchie–Davidson, Obara–Saika and Rys schemes

Overview

1. **Cartesian and Hermite Gaussians**
   - properties of Cartesian Gaussians
   - properties of Hermite Gaussians

2. **Simple one-electron integrals**
   - Gaussian product rule and overlap distributions
   - overlap distributions expanded in Hermite Gaussians
   - overlap and kinetic-energy integrals

3. **Coulomb integrals**
   - Gaussian electrostatics and the Boys function
   - one- and two-electron Coulomb integrals over Cartesian Gaussians

4. **Sparsity and screening**
   - Cauchy–Schwarz screening

5. **Coulomb potential**
   - early density matrix contraction
   - density fitting

Cartesian Gaussian-type orbitals (GTOs)

- We shall consider integration over primitive Cartesian Gaussians centered at **A**:

$$G_{ijk}(\mathbf{r}, a, \mathbf{A}) = x_A^i y_A^j z_A^k \exp(-ar_A^2), \quad \begin{cases} a > 0 & \text{orbital exponent} \\ \mathbf{r}_A = \mathbf{r} - \mathbf{A} & \text{electronic coordinates} \\ i \geq 0, j \geq 0, k \geq 0 & \text{quantum numbers} \end{cases}$$

  - total angular-momentum quantum number $\ell = i + j + k \geq 0$
  - Gaussians of a given ℓ constitutes a shell:
    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}

- Each Gaussian factorizes in the Cartesian directions:

$$G_{ijk}(a, \mathbf{r}_A) = G_i(a, x_A) G_j(a, y_A) G_k(a, z_A), \quad G_i(a, x_A) = x_A^i \exp(-ax_A^2)$$

  - note: this is not true for spherical-harmonic Gaussians nor for Slater-type orbitals

- Gaussians satisfy a simple recurrence relation:

$$x_A G_i(a, x_A) = G_{i+1}(a, x_A)$$

- The differentiation of a Gaussian yields a linear combination of two Gaussians

$$\frac{\partial G_i(a, x_A)}{\partial A_x} = -\frac{\partial G_i(a, x_A)}{\partial x} = 2a G_{i+1}(a, x_A) - i G_{i-1}(a, x_A)$$

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:

$$\Lambda_t(x_P) = (\partial/\partial P_x)^t \exp\left(-px_P^2\right) \quad \leftarrow \text{a Gaussian times a polynomial of degree } t$$

- Hermite Gaussians yields the same spherical-harmonic functions as do Cartesian Gaussians
    - they may therefore be used as basis functions in place of Cartesian functions
- We shall only consider their use as intermediates in the evaluation of Gaussian integrals

$$\text{Cartesian product} \rightarrow G_i(x_A) G_j(x_B) = \sum_{t=0}^{i+j} E_t^{ij} \Lambda_t(x_P) \leftarrow \text{Hermite expansion}$$

    - such expansions are useful because of simple integration properties such as

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

    - McMurchie and Davidson (1978)
- We shall make extensive use of the McMurchie–Davidson scheme for integration

Integration over Hermite Gaussians

▶ From the definition of Hermite Gaussians, we have

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

▶ We now change the order of differentiation and integration by Leibniz' rule:

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

▶ The basic Gaussian integral is given by

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

▶ Since the integral is independent of **P**, differentiation with respect to **P** gives zero:

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

  ▶ only integrals over Hermite s functions do not vanish

Hermite recurrence relation

- Cartesian Gaussians satisfy the simple recurrence relation

$$x_A G_i = G_{i+1}$$

- The corresponding Hermite recurrence relation is slightly more complicated:

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

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

    $$\Lambda_{t+1} = \left(\frac{\partial}{\partial P_x} \right)^t \frac{\partial}{\partial P_x} \exp(-p x_P^2) = 2p \left(\frac{\partial}{\partial P_x} \right)^t x_P \Lambda_0$$
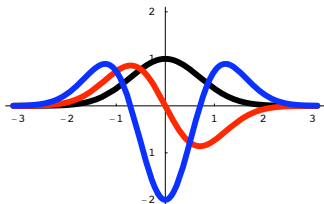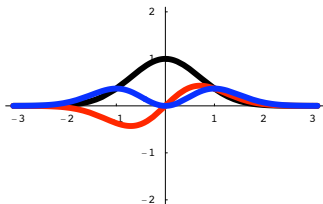
  - inserting the identity

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

    we then obtain

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

Comparison of Cartesian and Hermite Gaussians

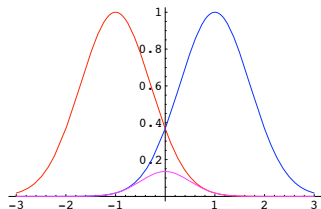|	Cartesian Gaussians	Hermite Gaussians
definition	$G_i = x_A^i \exp\left(-a x_A^2\right)$	$\Lambda_t = \frac{\partial^t}{\partial P_x^t} \exp\left(-p x_P^2\right)$
recurrence	$x_A G_i = G_{i+1}$	$x_P \Lambda_t = \frac{1}{2p}\Lambda_{t+1} + t\Lambda_{t-1}$
differentiation	$\frac{\partial G_i}{\partial A_x} = 2a G_{i+1} - i G_{i-1}$	$\frac{\partial \Lambda_t}{\partial P_x} = \Lambda_{t+1}$

The Gaussian product rule

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

$$\exp\left(-ax_A^2\right)\exp\left(-bx_B^2\right) = \underbrace{\exp\left(-\mu X_{AB}^2\right)}_{\text{exponential prefactor}} \underbrace{\exp\left(-px_P^2\right)}_{\text{product Gaussian}}$$

where

$$P_x = \frac{aA_x + bB_x}{p} \qquad \leftarrow \text{``center of mass''}$$

$$p = a + b \qquad \leftarrow \text{total exponent}$$

$$X_{AB} = A_x - B_x, \qquad \leftarrow \text{relative separation}$$

$$\mu = \frac{ab}{a+b} \qquad \leftarrow \text{reduced exponent}$$

- The Gaussian product rule greatly simplifies integral evaluation
  - 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

$$\int \Omega_{ij}(x) \, dx = K_{AB} \int x_A^i x_B^j \exp\left(-px_P^2\right) \, dx$$

  ▶ the Cartesian monomials still make the integration awkward
  ▶ we would like to utilize the simple integration properties of Hermite Gaussians

▶ We therefore expand Cartesian overlap distributions in Hermite Gaussians

$$\text{overlap distribution} \rightarrow \Omega_{ij}(x) = \sum_{t=0}^{i+j} E_t^{ij} \Lambda_t(x_P) \leftarrow \text{Hermite Gaussians}$$

  ▶ 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 Λ_t with $0 \le t \le i + j$

▶ The expansion coefficients may be evaluated recursively:

$$E_t^{i+1,j} = \frac{1}{2p} E_{t-1}^{ij} + X_{PA} E_t^{ij} + (t+1) E_{t+1}^{ij}$$

  ▶ we shall now prove these recurrence relations

Recurrence relations for Hermite expansion coefficients

- A straightforward Hermite expansion of $\Omega_{i+1,j}$

$$\Omega_{i+1,j} = K_{AB} x_A^{i+1} x_B^j \exp\left(-pr_P^2\right) = \sum_t E_t^{i+1,j} \Lambda_t$$

- An alternative Hermite expansion of $\Omega_{i+1,j}$

$$\begin{aligned}
\Omega_{i+1,j} &= x_A \Omega_{ij} = (x - A_x)\Omega_{ij} \\
&= (x - P_x)\Omega_{ij} + (P_x - A_x)\Omega_{ij} = x_P \Omega_{ij} + X_{PA}\Omega_{ij} \\
&= \sum_t E_t^{ij} x_P \Lambda_t + X_{PA} \sum_t E_t^{ij} \Lambda_t \\
&= \sum_t E_t^{ij} \left(\tfrac{1}{2p}\Lambda_{t+1} + t\Lambda_{t-1} + X_{PA}\Lambda_t\right) \\
&= \sum_t \left[\tfrac{1}{2p}E_{t-1}^{ij} + (t+1)E_{t+1}^{ij} + X_{PA}E_t^{ij}\right]\Lambda_t
\end{aligned}$$

  - we have here used the recurrence $x_P \Lambda_t = \tfrac{1}{2p}\Lambda_{t+1} + t\Lambda_{t-1}$

- A comparison of the two expansions yields the recurrence relations

$$\boxed{E_t^{i+1,j} = \tfrac{1}{2p}E_{t-1}^{ij} + X_{PA}E_t^{ij} + (t+1)E_{t+1}^{ij}}$$

Overlap integrals

► We now have everything we need to evaluate overlap integrals:

$$S_{ab} = \langle G_a | G_b \rangle$$
$$G_a = G_{ikm}(\mathbf{r}, a, \mathbf{A}) = G_i(x_A)G_k(y_A)G_m(z_A)$$
$$G_b = G_{jln}(\mathbf{r}, b, \mathbf{B}) = G_j(x_B)G_l(y_B)G_n(z_B)$$

► 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:

$$S_{ij} = \int \Omega_{ij}(x)\,dx = \sum_{t=0}^{i+j} E_t^{ij} \int \Lambda_t(x_P)\,dx = \sum_{t=0}^{i+j} E_t^{ij} \delta_{t0}\sqrt{\frac{\pi}{p}} = E_0^{ij}\sqrt{\frac{\pi}{p}}$$

  ► only one term survives!

► The total overlap integral is therefore given by

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

Dipole-moment integrals

- Many other integrals may be evaluated in the same manner
- For example, dipole-moment integrals are obtained as

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

$$= \sum_{t=0}^{i+j} E_t^{ij} \int \Lambda_t(x_P) x_c \, dx \quad \text{(expand in Hermite Gaussians)}$$

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

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

$$= E_0^{ij} X_{PC} \sqrt{\frac{\pi}{p}} + E_1^{ij} \sqrt{\frac{\pi}{p}} \quad \text{(s functions integration)}$$

- The final dipole-moment integral is given by

$$D_{ij} = \left(E_1^{ij} + X_{PC} E_0^{ij} \right) \sqrt{\frac{\pi}{p}}$$

Kinetic-energy integrals

► As a final example of non-Coulomb integrals, we consider the kinetic-energy integrals:

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

$$T_{ab} = T_{ij} S_{kl} S_{mn} + S_{ij} T_{kl} S_{mn} + S_{ij} S_{kl} T_{mn}$$

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

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

► Differentiation of Cartesian Gaussians gives

$$\frac{\mathrm{d}}{\mathrm{d}x} G_j(x_B) = -2b G_{j+1} + j G_{j-1}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} G_j(x_B) = 4b^2 G_{j+2} - 2b(2j+1) G_j + j(j-1) G_{j-2}$$

  ► the first derivative is linear combination of two undifferentiated Gaussians
  ► the second derivative is linear combination of three undifferentiated Gaussians

► Kinetic-energy integrals are linear combinations of up to three overlap integrals:

$$T_{ij} = -2b^2 S_{i,j+2} + b(2j+1) S_{i,j} - \tfrac{1}{2} j(j-1) S_{i,j-2}$$

Coulomb integrals

- We shall consider two types of Coulomb integrals:
    - one-electron nuclear-attraction integrals

    $$\left\langle G_a(\mathbf{r}_A) \left| r_C^{-1} \right| G_b(\mathbf{r}_B) \right\rangle$$

    - two-electron repulsion integrals

    $$\left\langle G_a(\mathbf{r}_{1A}) G_b(\mathbf{r}_{1B}) \left| r_{12}^{-1} \right| G_c(\mathbf{r}_{2C}) G_d(\mathbf{r}_{2D}) \right\rangle$$

- Coulomb integrals are not separable in the Cartesian directions
    - however, they may be reduced to a one-dimensional integral on a finite interval:

    $$F_n(x) = \int_0^1 \exp\left(-xt^2\right) t^{2n} \mathrm{d}t \leftarrow \text{the Boys function}$$

    - this makes the evaluation of Gaussian Coulomb integrals relatively simple
- We begin by evaluating the one-electron spherical Coulomb integral:

    $$V_p = \int \frac{\exp\left(-pr_P^2\right)}{r_C} \, \mathrm{d}\mathbf{r} = \frac{2\pi}{p} F_0\left(pR_{PC}^2\right)$$

    - we shall next go on to consider more general Coulomb integrals

Coulomb integral over a spherical Gaussian I

▶ We would like to evaluate the three-dimensional Coulomb-potential integral

$$V_p = \int \frac{\exp\left(-pr_P^2\right)}{r_C}\, d\mathbf{r}$$

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

$$\frac{1}{r_C} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left(-r_C^2 t^2\right) dt \quad \leftarrow \text{ Laplace transform}$$

to yield the four-dimensional integral

$$V_p = \frac{1}{\sqrt{\pi}} \int \exp\left(-pr_P^2\right) \int_{-\infty}^{\infty} \exp\left(-t^2 r_C^2\right) dt\, d\mathbf{r}$$

② To prepare for integration over $\mathbf{r}$, we invoke the Gaussian product rule

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

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

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

– to be continued. . .

Coulomb integral over a spherical Gaussian II

- carried over from previous slide:

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

3 Integration over all space $\mathbf{r}$ 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]

Gaussian electrostatics
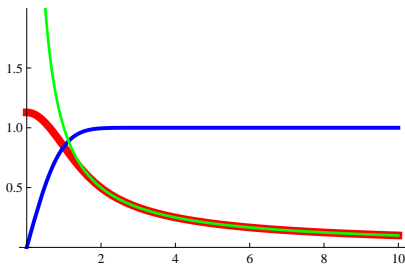
▶ The Boys function is related to the error function

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

▶ Introducing unit charge distributions and the reduced exponent as

$$\rho_P(\mathbf{r}_P) = \left(\frac{p}{\pi}\right)^{3/2} \exp\left(-p r_P^2\right), \quad \alpha = \frac{pq}{p+q}$$

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

$$\int \frac{\rho_p(\mathbf{r}_P)}{r_C} \, d\mathbf{r} = \frac{\text{erf}\left(\sqrt{p} R_{PC}\right)}{R_{PC}}$$

$$\iint \frac{\rho_p(\mathbf{r}_{1P}) \rho_q(\mathbf{r}_{2Q})}{r_{12}} \, d\mathbf{r}_1 d\mathbf{r}_2$$

$$= \frac{\text{erf}\left(\sqrt{\alpha} R_{PQ}\right)}{R_{PQ}}$$

The Boys function $F_n(x)$

- The Boys function is central to molecular integral evaluation
  - it is evaluated by a combination of numerical techniques

- $F_n(x) > 0$ since the integrand is positive:

$$F_n(x) = \int_0^1 \exp\left(-xt^2\right) t^{2n}\,dt > 0$$

- $F_n(x)$ is convex and decreasing:

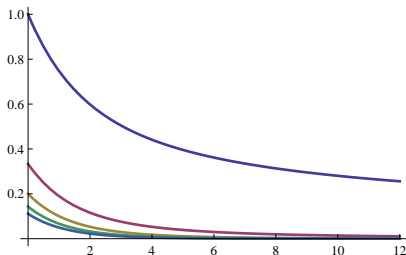$$F_n'(x) = -F_{n+1}(x) < 0$$
$$F_n''(x) = F_{n+2}(x) > 0$$

- Evaluation for small and large values:

$$F_n(x) = \frac{1}{2n+1} + \sum_{k=1}^{\infty} \frac{(-x)^k}{k!(2n+2k+1)} \quad (x \text{ small})$$

$$F_n(x) \approx \int_0^{\infty} \exp\left(-xt^2\right) t^{2n}dt = \frac{(2n-1)!!}{2^{n+1}}\sqrt{\frac{\pi}{x^{2n+1}}} \quad (x \text{ large})$$

- Different n-values are related downward recurrence relations:

$$F_{n-1}(x) = \frac{2xF_n(x) + \exp(-x)}{2n-1}$$

Cartesian Coulomb integrals

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

$$\int \frac{\exp\left(-pr_P^2\right)}{r_C}\,\mathrm{d}\mathbf{r} = \frac{2\pi}{p} F_0\left(pR_{PC}^2\right)$$

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

$$V_{ab} = \left\langle G_a \left| r_C^{-1} \right| G_b \right\rangle = \int \frac{\Omega_{ab}(\mathbf{r})}{r_C}\,\mathrm{d}\mathbf{r}$$

- The overlap distribution is expanded in Hermite Gaussians

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

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

$$\Omega_{ab}(\mathbf{r}) = \sum_{tuv} E_t^{ij} E_u^{kl} E_v^{mn} \Lambda_{tuv}(\mathbf{r}_P) = \sum_{tuv} E_{tuv}^{ab} \Lambda_{tuv}(\mathbf{r}_P)$$

- We now expand two-center Coulomb integrals in one-center Hermite integrals:

$$\text{two-center integral} \rightarrow V_{ab} = \sum_{tuv} E_{tuv}^{ab} \int \frac{\Lambda_{tuv}(\mathbf{r}_P)}{r_C}\,\mathrm{d}\mathbf{r} \leftarrow \text{one-center integrals}$$

- Our next task is therefore to evaluate Hermite Coulomb integrals

Coulomb integrals over Hermite Gaussians

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

$$V_{ab} = \sum_{tuv} E_{tuv}^{ab} \int \frac{\Lambda_{tuv}(\mathbf{r}_P)}{r_C} \, d\mathbf{r}, \quad \int \frac{\exp\left(-p r_P^2\right)}{r_C} \, d\mathbf{r} = \frac{2\pi}{p} F_0\left(p R_{PC}^2\right)$$

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

$$\int \frac{\Lambda_{tuv}(\mathbf{r}_P)}{r_C} \, d\mathbf{r} = \frac{\partial^{t+u+v}}{\partial P_x^t \partial P_y^u \partial P_z^v} \int \frac{\exp(-p R_P^2)}{r_C} \, d\mathbf{r}$$

$$= \frac{2\pi}{p} \frac{\partial^{t+u+v} F_0(p R_{PC}^2)}{\partial P_x^t \partial P_y^u \partial P_z^v} = \frac{2\pi}{p} R_{tuv}(p, \mathbf{R}_{PC})$$

where the one-center Hermite Coulomb integral is given by

$$R_{tuv}(p, \mathbf{R}_{PC}) = \frac{\partial^{t+u+v} F_0\left(p R_{PC}^2\right)}{\partial P_x^t \partial P_y^u \partial P_z^v}$$

▶ The Hermite Coulomb 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

Evaluation of Hermite Coulomb integrals

▶ To set up recursion for the Hermite integrals, we introduce the auxiliary Hermite integrals

$$R_{tuv}^n(p, \mathbf{P}) = (-2p)^n \frac{\partial^{t+u+v} F_n\left(pR_{PC}^2\right)}{\partial P_x^t \partial P_y^u \partial P_z^v}$$

which include as special cases the source and target integrals

$$R_{000}^n = (-2p)F_n \;\rightarrow\; R_{tuv}^0 = \frac{\partial^t}{\partial P_x^t} \frac{\partial^u}{\partial P_y^u} \frac{\partial^v}{\partial P_z^v} F_0$$

▶ The Hermite integrals can now be generated from the recurrence relations:

$$R_{t+1,u,v}^n = tR_{t-1,u,v}^{n+1} + X_{PC} R_{tuv}^{n+1}$$
$$R_{t,u+1,v}^n = uR_{t,u-1,v}^{n+1} + Y_{PC} R_{tuv}^{n+1}$$
$$R_{t,u,v+1}^n = vR_{t,u,v-1}^{n+1} + Z_{PC} R_{tuv}^{n+1}$$

Summary one-electron Coulomb integral evaluation

1. Calculate Hermite expansion coefficients for overlap distributions by recurrence

$$E_0^{00} = \exp\left(-\mu X_{AB}^2\right)$$
$$E_t^{i+1,j} = \frac{1}{2p} E_{t-1}^{ij} + X_{PA} E_t^{ij} + (t+1) E_{t+1}^{ij}$$

2. Calculate the Boys function by a variety of numerical techniques

$$F_n(x) = \int_0^1 \exp\left(-xt^2\right) t^{2n} \mathrm{d}t \quad (x = pR_{PC}^2)$$
$$F_n(x) = \frac{2xF_{n+1}(x) + \exp(-x)}{2n+1}$$

3. Calculate Hermite Coulomb integrals by recurrence

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

4. Obtain the Cartesian Coulomb integrals by expansion in Hermite integrals

$$\left\langle G_a \left| r_C^{-1} \right| G_b \right\rangle = \sum_{tuv} E_t^{ij} E_u^{kl} E_v^{mn} R_{tuv}^0$$

Two-electron Coulomb integrals

▶ Two-electron integrals are treated in the same way as one-electron integrals

▶ Evaluation of one-electron integrals:

$$\left\langle G_a \left| r_C^{-1} \right| G_b \right\rangle = \int \frac{\Omega_{ab}}{r_C} \, d\mathbf{r}$$

$$= \sum_{tuv} E_{tuv}^{ab} \int \frac{\Lambda_{tuv}}{r_C} \, d\mathbf{r} = \frac{2\pi}{p} \sum_{tuv} E_{tuv}^{ab} R_{tuv}(p, \mathbf{R}_{PC})$$

▶ Evaluation of two-electron integrals:

$$\left\langle G_a(a) G_b(1) \left| r_{12}^{-1} \right| G_c(2) G_d(2) \right\rangle = \iint \frac{\Omega_{ab}(1)\Omega_{cd}(2)}{r_{12}} \, d\mathbf{r}_1 d\mathbf{r}_2$$

$$= \sum_{tuv} E_{tuv}^{ab} \sum_{\tau\nu\phi} E_{\tau\nu\phi}^{cd} \iint \frac{\Lambda_{tuv}(1)\Lambda_{\tau\nu\phi}(2)}{r_{12}} \, d\mathbf{r}_1 d\mathbf{r}_2$$

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

  ▶ Hermit expansion for both electrons, around centers $\mathbf{P}$ and $\mathbf{Q}$
  ▶ integrals depend on $\mathbf{R}_{PQ}$ and $\alpha = pq/(p+q)$, with $p = a+b$ and $q = c+d$

Direct SCF and other techniques

- The general expression for two-electron integrals is given by

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

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

$$F_{ab} = \sum_{cd} \big(\underbrace{2g_{abcd} D_{cd}}_{\text{Coulomb}} - \underbrace{g_{acbd} D_{cd}}_{\text{exchange}} \big)$$

    - 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

    - screening of integrals
    - early contraction with density matrices
    - density fitting
    - multipole methods

Screening of overlap integrals

- The product of two s functions is by the Gaussian product rule

$$\exp\left(-ar_A^2\right)\exp\left(-br_B^2\right) = \exp\left(-\frac{ab}{a+b}R_{AB}^2\right)\exp\left(-(a+b)r_P^2\right)$$

- This gives a simple expression for the overlap integral between two such 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 the separation R_{AB}

- Let us assume that we may neglect all integrals smaller than 10^{-k}:

$$|S_{ab}| < 10^{-k} \quad \leftarrow \text{insignificant integrals}$$

- We may then neglect 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

Screening of two-electron integrals

- A two-electron *ssss* integral may be written in the form:

$$g_{abcd} = \text{erf}\left(\sqrt{\alpha}R_{PQ}\right)\frac{S_{ab}S_{cd}}{R_{PQ}}$$

  - the total number of integrals scales quartically with system size
  - the number of significant integrals scales quadratically

$$S_{ab}, S_{cd} \to 0 \quad \text{rapidly}$$
$$R_{PQ}^{-1} \to 0 \quad \text{very slowly}$$

- Let us decompose the integral into classical and nonclassical parts:

$$g_{abcd} = \underbrace{\frac{S_{ab}S_{cd}}{R_{PQ}}}_{\text{classical}} - \underbrace{\text{erfc}\left(\sqrt{\alpha}R_{PQ}\right)\frac{S_{ab}S_{cd}}{R_{PQ}}}_{\text{nonclassical}}$$

  - 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 $1a_0$
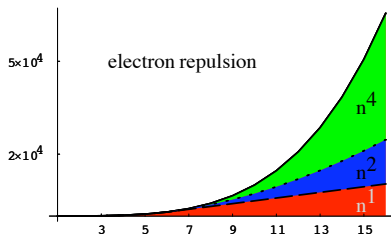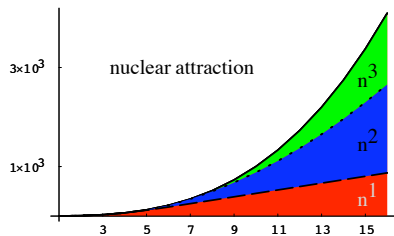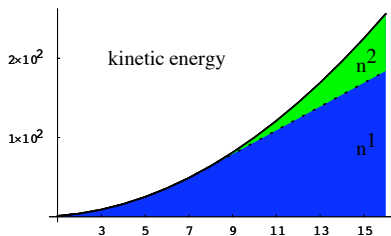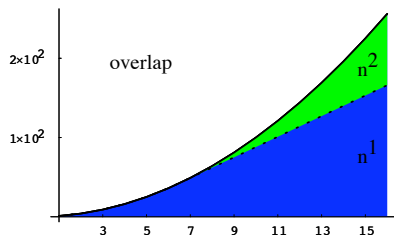
Integral prescreening

- Small integrals ($< 10^{-10}$) are not needed and should be avoided by prescreening
- The two-electron integrals

$$g_{abcd} = \iint \frac{\Omega_{ab}(1)\Omega_{cd}(2)}{r_{12}} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2$$

  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

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

- Precalculate

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

  and prescreen

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

Early contraction with density matrix

- In direct SCF, two-electron integrals are only used for Fock/KS-matrix construction

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

  - only integrals that contribute significantly are evaluated
  - screening is made based on the size of the density matrix

- It is not necessary to evaluate significant integrals fully before contraction with D_{cd}

- An early contraction of D_{cd} with the coefficients $E^{cd}_{\tau\nu\phi}$ is more efficient:

$$K^Q_{\tau\nu\phi} = (-1)^{\tau+\nu+\phi} \sum_{cd \in Q} D_{cd} E^{cd}_{\tau\nu\phi}$$

$$J_{ab} = \frac{2\pi^{5/2}}{pq\sqrt{p+q}} \sum_{tuv} E^{ab}_{tuv} \sum_{Q} \sum_{\tau\nu\phi} K^Q_{\tau\nu\phi} R_{t+\tau,u+\nu,v+\phi}(\alpha, \mathbf{R}_{PQ})$$

- Timings (seconds) for benzene:

	cc-pVDZ	cc-pVTZ	cc-pVQZ
late contraction	27	374	3823
early contraction	10	90	687

Density fitting

- Traditionally, the electron density is expanded in orbital products

$$\rho(\mathbf{r}) = \sum_{ab} D_{ab}\, \Omega_{ab}(\mathbf{r})$$

  - 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

$$J_{ab} = \left\langle \Omega_{ab}(\mathbf{r}_1) \left| r_{12}^{-1} \right| \rho(\mathbf{r}_2) \right\rangle$$

  - the formal cost of this evaluation is therefore quartic (n^4)

- Consider now an approximate density $\widetilde{\rho}(\mathbf{r})$, expanded in an auxiliary basis $\omega_\alpha(\mathbf{r})$:

$$\widetilde{\rho}(\mathbf{r}) = \sum_\alpha c_\alpha \omega_\alpha(\mathbf{r}), \quad \widetilde{\rho}(\mathbf{r}) \approx \rho(\mathbf{r})$$

  - the size of the auxiliary basis N increases linearly with system size

- We may now evaluate the Coulomb contribution approximately as

$$\widetilde{J}_{ab} = \left\langle \Omega_{ab}(\mathbf{r}_1) \left| r_{12}^{-1} \right| \widetilde{\rho}(\mathbf{r}_2) \right\rangle$$

  - the cost of the evaluation is therefore cubic ($n^2 N$)
  - if the $\widetilde{\rho}$ is sufficiently accurate and easy to obtain, this may give large savings

Coulomb density fitting

- 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:

$$J_{ab} = \sum_{cd} (ab|cd) \, D_{cd} = (ab|\rho)$$
$$\approx \sum_{cd} \sum_{\alpha\beta} (ab|\alpha) \, (\alpha|\beta)^{-1} \, (\beta|cd) \, D_{cd} = \sum_{\alpha} (ab|\alpha) \, c_{\alpha} = (ab|\widetilde{\rho}) = \widetilde{J}_{ab}$$

  where the coefficients c_{α} are obtained from the linear sets of equations

$$\sum_{\alpha} (\beta|\alpha) \, c_{\alpha} = \sum_{cd} (\beta|cd) \, D_{cd} \quad \Leftrightarrow \quad (\beta|\widetilde{\rho}) = (\beta|\rho)$$

  - 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

$$J_{ab} = (ab \,|\, \rho), \quad \widetilde{J}_{ab} = (ab \,|\, \widetilde{\rho}), \quad (\alpha \,|\, \rho) = (\alpha \,|\, \widetilde{\rho})$$

▶ We may then calculate the exact and density-fitted Coulomb energies as

$$E = \sum_{ab} J_{ab} D_{ab} = (\rho \,|\, \rho), \quad \widetilde{E} = \sum_{ab} \widetilde{J}_{ab} D_{ab} = (\rho \,|\, \widetilde{\rho})$$

▶ The true Coulomb energy is an upper bound to the density-fitted energy:

$$E - \widetilde{E} = (\rho \,|\, \rho) - (\rho \,|\, \widetilde{\rho}) = (\rho \,|\, \rho) - (\rho \,|\, \widetilde{\rho}) - (\widetilde{\rho} \,|\, \rho) + (\widetilde{\rho} \,|\, \widetilde{\rho}) = (\rho - \widetilde{\rho} \,|\, \rho - \widetilde{\rho})$$

where we have used the relation $(\widetilde{\rho}|\rho) = (\widetilde{\rho}|\widetilde{\rho})$ which follows from $(\alpha \,|\, \rho) = (\alpha \,|\, \widetilde{\rho})$

  ▶ the error in the density-fitted energy is quadratic in the error in the density
  ▶ Coulomb fitting is thus equivalent to minimization of

$$E - \widetilde{E} = (\rho - \widetilde{\rho} \,|\, \rho - \widetilde{\rho}) \geq 0$$

▶ It is possible to determine the approximate density in other ways

  ▶ quadratic energy error is ensured by using the robust formula

$$\widetilde{E} = (\widetilde{\rho} \,|\, \rho) + (\rho \,|\, \widetilde{\rho}) - (\widetilde{\rho} \,|\, \widetilde{\rho}) = E - (\rho - \widetilde{\rho} \,|\, \rho - \widetilde{\rho})$$

  ▶ however, the error in the energy is no longer minimized

Density fitting: sample calculations

- Calculation on benzene

	cc-pVDZ	cc-pVTZ	cc-pVQZ
exact density late	27	374	3823
exact density early	10	90	687
fitted density	1	7	34
exact energy	−230.09671	−230.17927	−230.19463
density-fitted energy	−230.09733	−230.17933	−230.19466

- 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