# The Gaussian Function in Calculations of Statistical Mechanics and Quantum Mechanics

**Isaiah Shavitt**

IBM Watson Scientific Computing Laboratory, Columbia University
New York, New York

I. Introduction ........................................... 1
II. Some Properties of the Gaussian Function .......... 2
   A. Definition ...................................... 2
   B. Products of Gaussians ......................... 3
   C. Integrals of Gaussians ....................... 6
   D. Expansions in Gaussians ...................... 11
III. A Problem in Statistical Mechanics: Virial Coefficients .... 16
    A. Statement of the Problem .................... 16
    B. Introduction of Gaussians .................. 19
    C. Computational Problems and Solutions ...... 23
IV. A Problem in Quantum Mechanics: Multicenter Integrals .... 27
    A. Statement of the Problem .................... 27
    B. First Approach: Gaussian Type Orbitals .... 30
    C. Second Approach: Transformation to Gaussians ... 33
    Acknowledgments .................................. 44
    References ...................................... 44

---

**I. Introduction**

The Gaussian function \( \exp(-ax^2) \) is best known for its application in statistics, as the normal probability distribution. This chapter is concerned with a demonstration of its usefulness in an entirely different field, the evaluation of certain types of multiple integrals. These integrals frequently appear in calculations of interaction energies between atoms or molecules, and are of considerable importance in theoretical chemistry and physics. In the most general form they involve integrands which are products of functions of distances between a number of points, the integrations being over the space coordinates of some or all of these points.

Two specific examples will be treated in detail. The first, from statistical mechanics, is the evaluation of higher virial coefficients, representing the averaged interactions between molecules in a fluid; the other is the
calculation of interelectron repulsion integrals in molecular quantum mechanics.

The unique usefulness of the Gaussian function in the evaluation of interaction integrals is due, in great measure, to its dependence on the square of the argument rather than on the argument itself. When this argument represents an interparticle distance, we can avoid the appearance of square roots in expressions describing relations between functions of different distances (e.g., a shift of the origin of the coordinate system). Coupled with the exponential nature of the Gaussian function, this makes for easy simplification of products of Gaussians involving different distances. In fact, as will be shown, a product of Gaussian functions, each centered on a different origin, can be expressed as a single Gaussian function in terms of a suitable center.

A detailed treatment of the pertinent properties of the Gaussian function is given in the following section; the section also describes some techniques for the useful introduction of Gaussians into calculations which did not involve them originally, for example, integral transforms of a number of functions to Gaussians. Section III is devoted to an example from statistical mechanics, in which the introduction of Gaussians (by means of least-squares fittings) enabled the estimation of integrals involving all pair-distances between four particles interacting with arbitrary pair-potentials; these were required for the theoretical prediction of fourth virial coefficients of gases. Section IV is concerned with an example from molecular quantum mechanics, where an integral transformation from exponentials to Gaussians was used in the calculation of the extremely intractable three- and four-center integrals which appear in polyatomic calculations.

II. Some Properties of the Gaussian Function

A. Definition

We shall mostly be concerned with the three-dimensional Gaussian function, which will be written in any of the equivalent forms

\[ G(r_A) = G(\alpha_i; r_A) = \exp\left( -\alpha_i r_A^2 \right). \]  

(1)

Here \( \alpha_i \) (which must be positive) is a parameter, the index \( i \) referring to its position in an independently specified ordered list of \( \alpha_i \)'s. The argument \( r_A \) is the magnitude of the vector \( r_A \) from a given fixed point \( A \) to the variable point \( P(x, y, z) \). The fixed point, having the Cartesian coordinates \( A_x, A_y, A_z \), will be referred to as the center of the Gaussian.

Different subscripts \( A, B, \ldots \) will be used to distinguish between different centers; thus \( G_A(r_A) \) is a Gaussian centered on the point \( A \) and having the parameter \( \alpha_A \). It can be written as a function of the coordinates of \( P(x, y, z) \),

\[ G_A(r_A) = \exp\left( -\alpha_A[(x - A_x)^2 + (y - A_y)^2 + (z - A_z)^2] \right). \]  

(2)

We shall also use the abbreviations

\[ x_A = x - A_x \quad (x \text{ component of } r_A), \]
\[ y_A = y - A_y \quad (y \text{ component of } r_A), \]
\[ z_A = z - A_z \quad (z \text{ component of } r_A). \]  

(3)

When the center designation \( (A \text{ in } r_A) \) is omitted, it can be assumed that the Gaussian is centered on the origin of the coordinate system.

In some applications both end-points of \( r \) are variable; in such a case, if the end-points are designated 1 and 2, the Gaussian is written \( G(r_{12}) \). In Cartesian coordinates this will appear as

\[ G(r_{12}) = \exp\left( -\alpha_i[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2] \right). \]  

(4)

It should be noted that the expressions in Eqs. (2) and (4) can each be factored into three simple (one-dimensional) Gaussians, one in each of \( x, y, \) and \( z \).

We shall have occasion to refer to the "range" of a Gaussian. This is the maximum distance \( r \) for which \( G(\alpha; r) \) cannot be considered negligible. Obviously, the range is inversely proportional to \( \sqrt{\alpha} \).

B. Products of Gaussians

1. Gaussians with Fixed Centers

Theorem. The product of two Gaussians having different centers \( A \) and \( B \) is itself a Gaussian (apart from a constant factor) with a center somewhere on the line segment \( AB \). Specifically,

\[ G_A(r_A) G_B(r_B) = KG_{AB}(r_C), \]  

(5)

where \( K \) is a constant,

\[ K = \exp\left( -\frac{\alpha_A\alpha_B}{\alpha_i + \alpha_j} AB^2 \right), \]  

(6)

and

\[ \alpha_C = \alpha_A + \alpha_B; \]
\[ C_x = \frac{\alpha_i A_x + \alpha_j B_x}{\alpha_i + \alpha_j}; \]
\[ C_y = \frac{\alpha_i A_y + \alpha_j B_y}{\alpha_i + \alpha_j}; \]
\[ C_z = \frac{\alpha_i A_z + \alpha_j B_z}{\alpha_i + \alpha_j}. \]  

(7)
Proof (see Fig. 1): If the point $C$ is placed in accordance with Eqs. (8), then
\begin{align*}
a &= \overline{AC} = \frac{\alpha_i}{\alpha_i + \alpha_j} R, \\
b &= \overline{CB} = \frac{\alpha_j}{\alpha_i + \alpha_j} R,
\end{align*}
where $R = \overline{AB}$. Using the cosine law, we have
\begin{align*}
r_A^2 &= a^2 + r_c^2 + 2ar_c \cos \theta \\
r_B^2 &= b^2 + r_c^2 - 2br_c \cos \theta \\
b r_A^2 + ar_B^2 &= a^2b + ab^2 + (a + b) r_c^2 \\
&= R(ab + r_c^2).
\end{align*}
Substituting for $a$ and $b$ from Eqs. (9), and multiplying throughout by $(\alpha_i + \alpha_j)/R$, we get
\[a r_A^2 + \alpha_i r_B^2 = \frac{\alpha_i \alpha_j}{\alpha_i + \alpha_j} R^2 + (\alpha_i + \alpha_j) r_c^2,
\]
from which the theorem follows directly.

**Corollary.** The product of any number of Gaussians with arbitrary centers is expressible as a constant factor multiplying a single Gaussian.

2. **Gaussians with Variable Centers**

Consider a product of Gaussians
\[P_{ijk}(x_1, r_2) = G_i(r_1) G_j(r_2) G_k(r_3),\]
where the first two factors involve distances from the origin. Using Eq. (4), and factoring into $x, y,$ and $z$ parts, we have
\[P_{ijk}(r_1, r_2) = P_{ij}(x_1, x_2) P_{ik}(y_1, y_2) P_{jk}(z_1, z_2),\]
where
\[P_{ijk}(x_1, x_2) = \exp \left\{ -\alpha_i x_1^2 - \alpha_j x_2^2 - \alpha_k (x_1 - x_2)^2 \right\}
= \exp \left\{ -[(\alpha_i + \alpha_k) x_1^2 + (\alpha_j + \alpha_k) x_2^2 - 2\alpha_k x_1 x_2] \right\},
\]
with similar equations for the $y$ and $z$ factors. The expression in the square brackets is a positive definite quadratic form, and can be diagonalized by an orthogonal transformation to a new set of coordinates $\tilde{x}_1, \tilde{x}_2$:
\[P_{ijk}(x_1, x_2) = \exp \left\{ -[\tilde{\alpha}_1 \tilde{x}_1^2 + \tilde{\alpha}_2 \tilde{x}_2^2] \right\},\]
where $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ are linear combinations of $\alpha_i, \alpha_j,$ and $\alpha_k$. In most applications we shall require the product $\tilde{\alpha}_1 \tilde{\alpha}_2$, but not each of these individually. It is therefore worth noting that, because the determinant of a quadratic form is not changed in value by an orthogonal transformation,
\[\tilde{\alpha}_1 \tilde{\alpha}_2 = \begin{vmatrix} \alpha_i + \alpha_k & -\alpha_k \\ -\alpha_k & \alpha_j + \alpha_k \end{vmatrix}
= \alpha_i \alpha_j + \alpha_i \alpha_k + \alpha_j \alpha_k.
\]

We have thus been able to factorize Eq. (10) into 6 simple factors,
\[P_{ijk}(r_1, r_2) = \exp (-\tilde{\alpha}_1 \tilde{x}_1^2) \exp (-\tilde{\alpha}_2 \tilde{x}_2^2) \exp (-\tilde{\alpha}_3 \tilde{y}_1^2) \cdots \exp (-\tilde{\alpha}_5 \tilde{z}_2^2).
\]

Obviously, similar considerations apply to products of Gaussians involving distances between any number of points. For example,
\[G_i(r_1) G_j(r_2) G_k(r_3) G_m(r_5) G_n(r_6) G_o(r_7) = \exp (-\tilde{\alpha}_1 \tilde{x}_1^2) \cdots \exp (-\tilde{\alpha}_7 \tilde{x}_7^2),
\]
where
\[\tilde{\alpha}_1 \tilde{\alpha}_2 \tilde{\alpha}_3 \tilde{\alpha}_4 \tilde{\alpha}_5 \tilde{\alpha}_6 \tilde{\alpha}_7 = \begin{vmatrix} \alpha_i + \alpha_j & -\alpha_i & -\alpha_j & -\alpha_j & -\alpha_j & -\alpha_j & -\alpha_j \\ -\alpha_i & \alpha_j + \alpha_k & -\alpha_j & -\alpha_k & -\alpha_k & -\alpha_k & -\alpha_k \\ -\alpha_j & -\alpha_k & \alpha_j + \alpha_m & -\alpha_m & -\alpha_m & -\alpha_m & -\alpha_m \\ -\alpha_j & -\alpha_k & -\alpha_m & \alpha_k + \alpha_m & -\alpha_m & -\alpha_m & -\alpha_m \\ -\alpha_j & -\alpha_k & -\alpha_m & -\alpha_m & \alpha_m + \alpha_n & -\alpha_n & -\alpha_n \\ -\alpha_j & -\alpha_k & -\alpha_m & -\alpha_m & -\alpha_m & \alpha_m + \alpha_n & -\alpha_n \\ -\alpha_j & -\alpha_k & -\alpha_m & -\alpha_m & -\alpha_m & -\alpha_m & \alpha_m + \alpha_n \end{vmatrix}.
\]
C. Integrals of Gaussians

1. One-Dimensional Integrals

In the application of the methods discussed in this chapter we shall require a few definite integrals involving Gaussian functions. The most elementary of these is

$$\int_{0}^{\infty} \exp (-\alpha x^2) \, dx = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}}, \quad (\alpha > 0), \quad (17)$$

for which the following proof is well known.

$$\left( \int_{0}^{\infty} \exp (-\alpha x^2) \, dx \right)^2 = \int_{0}^{\infty} \int_{0}^{\infty} \exp (-\alpha x^2) \exp (-\alpha y^2) \, dx \, dy$$

$$= \int_{0}^{\infty} \int_{0}^{\infty} \exp (-\alpha \rho^2) \, d\rho \, d\rho$$

$$= \frac{\pi}{2} \int_{0}^{\infty} r \exp (-\alpha r^2) \, dr$$

by the substitution \( x = r \cos \theta, \ y = r \sin \theta \). Putting \( r^2 = s \), we get

$$\left( \int_{0}^{\infty} \exp (-\alpha x^2) \, dx \right)^2 = \frac{\pi}{4} \int_{0}^{\infty} \exp (-\alpha s) \, ds = \frac{\pi}{4\alpha},$$

from which Eq. (17) follows.

The more general integral,

$$\int_{0}^{\infty} r^s \exp (-\alpha r^2) \, dr = \frac{1}{2} \alpha^{-\frac{s+1}{2}} \Gamma \left( \frac{s+1}{2} \right), \quad (\alpha > 0), \quad (18)$$

can be derived from the integral representation of the Gamma function (see, e.g., Margenau and Murphy, 1943, p. 91).

If the integrand is a product of a Gaussian and a simple exponential, we have

$$\int_{-\infty}^{\infty} \exp (-\alpha x^2 + 2\beta x) \, dx = \sqrt{\frac{\pi}{\alpha}} \exp \left( \frac{\beta^2}{\alpha} \right), \quad (\alpha > 0). \quad (19)$$

This is easily proved by the substitution \( x = y - \beta/\alpha \), noting that

$$\alpha x^2 - 2\beta x = \alpha y^2 - \beta^2/\alpha.$$ 

Another useful integral is

$$\int_{0}^{\infty} \exp (-\alpha r^2 - \beta r^2) \, dr = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}} \exp \left( -2\sqrt{\alpha\beta} \right), \quad (\alpha > 0; \beta > 0). \quad (20)$$

To prove this we define a function \( I(x) \),

$$I(x) = \int_{0}^{\infty} \exp \left( -\left(t - \frac{x}{2t}\right)^2 \right) \, dt,$$

and obtain its derivative,

$$I'(x) = \int_{0}^{\infty} \left(1 - \frac{x}{2t}\right) \exp \left( -\left(t - \frac{x}{2t}\right)^2 \right) \, dt$$

$$= I(x) - \int_{0}^{\infty} \left(x/2t^2\right) \exp \left( -\left(t - \frac{x}{2t}\right)^2 \right) \, dt.$$

The second integral on the right is also equal to \( I(x) \) by virtue of the substitution \( u = x/2t \). Thus the derivative is identically zero, and \( I(x) \) is a constant independent of \( x \),

$$I(x) = I(0) = \sqrt{\pi}/2.$$

Furthermore,

$$\int_{0}^{\infty} \exp \left( -t^2 - \left(\frac{x}{2t}\right)^2 \right) \, dt = e^{-x/4} I(x),$$

$$e^{-x} = \left(2/\sqrt{\pi}\right) \int_{0}^{\infty} \exp \left( -t^2 - \left(\frac{x}{2t}\right)^2 \right) \, dt.$$

Equation (20) follows easily by the substitutions \( t = \sqrt{\alpha} r, x = 2 \sqrt{\alpha\beta} \).

A more general form can be derived from an integral representation of the modified Bessel function of the second kind \( K_{\nu}(x) \) (Watson, 1952, p. 183),

$$\int_{0}^{\infty} r^{\nu} \exp \left( -\alpha r^2 - \beta r^2 \right) \, dr = (\alpha^{\nu}(\beta^{\nu} - 1)/2) K_{\nu}(2 \sqrt{\alpha\beta}), \quad (\alpha > 0; \beta > 0). \quad (21)$$

We shall now consider the class of integrals

$$F_{m}(t) = \int_{0}^{1} u^{2m} \exp \left( -tu^2 \right) \, du, \quad (t > 0; m = 0, 1, 2, \ldots), \quad (22)$$

which can be considered a reduced form of the incomplete gamma function (Erdélyi et al., 1953, Chapter 1X),

$$F_{m}(t) = \frac{1}{2^{-m+1/2}} \gamma(m + \frac{1}{2}, t).$$

They are closely related to the error function \( \text{erf}(\sqrt{t}) \) and its derivatives, for by making the substitution \( tu^2 = v^2 \), we find

$$F_{0}(t) = \frac{1}{\sqrt{t}} \int_{0}^{\sqrt{t}} \exp (-v^2) \, dv$$

$$= \frac{1}{2} \sqrt{\pi/t} \text{erf}(\sqrt{t}).$$
Furthermore, by differentiating with respect to $t$ under the integral sign, we get

$$F_{m+1}(t) = -\frac{d}{dt} F_m(t). \tag{23}$$

Integration by parts yields the following recurrence relation,

$$F_m(t) = \left[ \frac{u^{2m+1}}{2m+1} e^{-ut^2} \right]_{u=0}^{u=\infty} - \int_0^t \frac{2tu^{2m+2}}{2m+1} e^{-tu^2} \, du$$

$$= \frac{1}{2m+1} e^{-t} + \frac{2t}{2m+1} F_{m+1}(t)$$

$$= \frac{1}{2m+1} \left[ 2tF_{m+1} + e^{-t} \right]. \tag{24}$$

Equation (24) provides one of the best ways of computing the functions $F_m(t)$, especially when they are required for a sequence of orders $m$. If $F_m(t)$ is assumed zero for a sufficiently high order, and Eq. (24) is used for successively lower orders, then, if $t$ is not too great, the accuracy increases rapidly as we proceed. For example, with $t = 1$, starting with $F_{11} = 0$, we obtain $F_3$ with 6-decimals accuracy and $F_0$ with 8-decimals accuracy. Evidently, this recurrence relation should not be used in the opposite direction.

Repeated application of Eq. (24) can be used to obtain a series expansion for $F_m(t)$,

$$F_m(t) = \frac{1}{2m+1} \left[ \frac{2t}{2m+3} + \cdots + e^{-t} \right]$$

$$= e^{-t} \sum_{i=2}^{\infty} \frac{(2t)^i}{(2m+1)(2m+3) \cdots (2m+2i+1)}. \tag{25}$$

It will be convenient to express this in terms of gamma functions of half integral argument (see Dwight, 1947, p. 194):

$$F_m(t) = \frac{1}{\Gamma\left(m + \frac{1}{2}\right)} e^{-t} \sum_{i=1}^{\infty} \frac{t^i}{\Gamma(2m+2i+1)} \tag{25}$$

To obtain an asymptotic expansion for large values of $t$ we define a complementary function $\Phi_m(t)$:

$$\Phi_m(t) = \int_0^\infty u^{2m} e^{-tu^2} \, du, \quad (m = 0, \pm 1, \pm 2, \ldots). \tag{26}$$

Obviously, for $m$ positive or zero,

$$F_m(t) = \int_0^\infty u^{2m} e^{-tu^2} \, du - \Phi_m(t)$$

$$= \frac{1}{\Gamma\left(m + \frac{1}{2}\right)} e^{-t} - \Phi_m(t). \tag{27}$$

Integrating Eq. (26) by parts, we can again obtain a recurrence relation

$$\Phi_m(t) = \frac{1}{2t} \left[ (2m-1) \Phi_{m-1}(t) + e^{-t} \right]. \tag{28}$$

For $m \leq 0$ the first term in the square brackets is negative, so that

$$\Phi_m(t) < \frac{1}{2t} e^{-t}, \quad (m \leq 0). \tag{29}$$

Applying Eq. (28) $i$ times results in

$$\Phi_m(t) = \frac{1}{2t} \left\{ \frac{2m-3}{2t} \cdots \left( (2m-2i+1) \Phi_{m-i}(t) + e^{-t} \right) + e^{-t} \right\}$$

$$= e^{-t} \left\{ 1 + \frac{2m-1}{2t} + \frac{(2m-1)(2m-3)}{(2t)^2} + \cdots + \frac{(2m-1)(2m-2i+3)}{(2t)^{i-1}} \right\}$$

$$+ \frac{(2m-1)(2m-2i+1)}{(2t)^i} \Phi_{m-i}(t) \tag{30}$$

From Eq. (29) we see that, provided $i \geq m$, neglecting the last term in Eq. (30) leads to an error which is less than

$$e^{-t} \frac{(2m-1)(2m-2i+1)}{(2t)^i},$$

which corresponds to what would have been the next term inside the curly brackets. In this case we have applied the recurrence relation once more. We have thus obtained a true asymptotic expansion for $\Phi_m(t)$, in which the error is less than the first neglected term (provided we compute at least $m$ terms):

$$\Phi_m(t) \approx \frac{1}{\Gamma\left(m + \frac{1}{2}\right)} e^{-t} \sum_{i=0}^{\infty} \frac{1}{\Gamma(2m+2i+1)} \frac{1}{t^i} \tag{31}$$
(the notation \([N]\) for the upper limit implies that the summation should not extend beyond the smallest term). Thus,

\[
F_m(l) \approx \frac{\Gamma(m + \frac{1}{2})}{2r^{m+1/2}} - \frac{\Gamma(m + \frac{1}{2})}{2r} e^{-l} \sum_{i=0}^{[N]} \frac{1}{\Gamma(m - i + \frac{1}{2})} e^{-lt}.
\]  

(Equations (23–32) can also be obtained directly from the corresponding relations for the incomplete gamma function.)

2. Multidimensional Integrals

We shall now consider an integral over a large closed volume \(V\) of a product of Gaussians of the type introduced in Section II, B, 2.

\[
I = \iiint_V G_1(r_{12}) G_2(r_{23}) G_3(r_{31}) \, dr_1 \, dr_2 \, dr_3,
\]

where \(dr_i = dx_i dy_i dz_i\). The dimensions of the volume will be assumed very large compared with the ranges of the Gaussians (Section II, A).

The integrand in Eq. (33) is a function of the distances between the points 1, 2, 3; it is independent of the absolute position of the group of points provided their relative positions are not changed. Furthermore, the integrand is negligibly small unless all three distances are smaller than the ranges of the corresponding Gaussians. Thus, ignoring the negligibly thin region within range of the boundary of \(V\), the integral over two of the points, say 1 and 2, is independent of the absolute position of the third point, and can be considered a constant for the integration over \(dr_3\). Therefore, this integration just yields a factor \(V\), and we have

\[
I = V \iiint G_1(r_{12}) G_2(r_{23}) G_3(r_{31}) \, dr_1 \, dr_2 \, dr_3,
\]

where point 3 can be assumed fixed at some point in the interior of \(V\), this point being taken as the origin. The range of the remaining integrations can now be taken as infinite, for the integrand is negligibly small wherever points 1 and 2 are not very close to the origin.

Proceeding as in Section II, B, 2, Eqs. (11-12), the integrand can now be factored into three equal integrals in \(x, y,\) and \(z\):

\[
I = V \left( \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \exp \left\{ -[(\alpha_i + \alpha_k) x_1^2 + (\alpha_i + \alpha_k) x_2^2 - 2\alpha_1 x_1 x_2] \right\} \right)^3,
\]

where we have put \(x_k = 0\). Diagonalizing the quadratic form results in

\[
I = V \left( \int_{-\infty}^{\infty} \exp \left\{ -\sum_1^3 \alpha_i \xi_i^2 \right\} d\xi_1 \right)^3 \left( \int_{-\infty}^{\infty} \exp \left\{ -\sum_1^3 \alpha_i \xi_i^2 \right\} d\xi_2 \right)^3 = \pi^3 V(\sum_1^3 \alpha_i)^{-3/2}.
\]

The Gaussian Function

Again as in Section II, B, 2, Eq. (14), \(\alpha_i \alpha_j\) is the determinant of the diagonalized quadratic form and is, therefore, equal to the determinant of the original quadratic form:

\[
\alpha_i \alpha_j = \left| \begin{array}{ccc} \alpha_i + \alpha_k & -\alpha_i \\ -\alpha_i & \alpha_i + \alpha_k \end{array} \right| = \alpha_i \alpha_k + \alpha_j \alpha_k + \alpha_i \alpha_k.
\]

Therefore,

\[
I = \pi^3 V(\alpha_i \alpha_j + \alpha_k \alpha_j + \alpha_l \alpha_j)^{-3/2}.
\]

Similar procedures can be used for integrals involving more than three points. For example,

\[
\iiint G_1(r_{12}) G_2(r_{23}) G_3(r_{34}) G_4(r_{41}) \, dr_1 \, dr_2 \, dr_3 \, dr_4 = \pi^6 V(\alpha_1 \alpha_2 \alpha_3 + \alpha_2 \alpha_3 \alpha_4 + \alpha_3 \alpha_4 \alpha_1)^{-3/2}.
\]

D. Expansions in Gaussians

The way in which we can take advantage of the unique properties of Gaussians in a particular calculation depends primarily upon the details of the problem. However, three general types of approach can be distinguished.

(1) Formulating the original problem in terms of Gaussians; this is only possible in some special cases, and each case has to be considered separately on its merits. An example in which this is possible, even if not very profitable, is provided by the variational calculations in molecular quantum mechanics, in which a set of somewhat arbitrary expansion functions are used. These can, in principle, be chosen as Gaussians (apart from a factor of the type \(r^{n-3} Y_{lm}(\theta, \phi)\), in which \(Y_{lm}\) is a spherical harmonic) rather than the more usual exponentials, resulting in simplifications in the evaluation of integrals (Boys, 1950). Further details are given in Section IV, B.

(2) Expanding the original functions of the problem in term of Gaussians by means of least squares (or some similar technique). This is discussed in Section II, D, 1.

(3) Applying some analytical transformation—generally an integral transform—from the original functions to Gaussians. This again will depend on the particular functions involved; some examples are discussed in Section II, D, 2.
1. Least-Squares Expansions

If a given, fairly smooth function $f(r)$ vanishes rapidly enough as $r$ increases, it is possible to obtain a reasonably accurate expansion for it in terms of a number of Gaussians (see Boys and Shavitt, 1960a):

$$f(r) = \sum_{i=0}^{n-1} C_i G_i(r) = \sum_{i=0}^{n-1} C_i G(\alpha_i; r)$$  \hfill (37)

The $2n$ parameters $\alpha_i$ and $C_i$ are to be chosen so as to make the approximation as good as possible. The most convenient criterion to use is to minimize the squared deviation,

$$D = \int \left[ f(r) - \sum_{i=0}^{n-1} C_i G_i(r) \right]^2 w(r) \, dr.$$  \hfill (38)

The integration is over the desired range of application of the expansion, usually 0 to $\infty$; it may be replaced by a sum over a discrete set of points $r_p$,

$$D = \sum_p \left[ f(r_p) - \sum_{i=0}^{n-1} C_i G_i(r_p) \right]^2 w(r_p).$$  \hfill (39)

The factor $w(r)$ is a weight function, and its choice is governed by the use to which the expansion is to be put, as well as by the spacing of the points $r_p$ if a sum is used rather than an integral. For example, if the purpose of the expansion, Eq. (37), is to facilitate a three-dimensional integration in which $r$ is the distance from a given center, then, in Eq. (38), we would choose

$$w(r) = r^2,$$  \hfill (40)

which is proportional to the volume of an element of spherical shell of radius $r$. This would improve the fit for large $r$, where great volume is covered, at the expense of accuracy in the relatively small region around the center. In Eq. (39) the weight factor would still be given by Eq. (40) if the points are equally spaced; though in principle this would require an infinite number of points for an infinite range, the sum would usually converge fairly quickly if the functions vanish rapidly enough as $r \rightarrow \infty$. Alternatively, one could use the points and weights of the appropriate Gauss quadrature formula; for the range $0 < r < \infty$ we would use the zeros of the Laguerre polynomials (Salzer and Zucker, 1949), and would multiply the corresponding quadrature weights by the original weight function (such as $r^2$ above) to get $w(r)$ for Eq. (39).

For any given set of exponential parameters $\alpha_i$, the choice of the linear coefficients $C_i$ to satisfy the minimum condition on $D$ requires the vanishing of the partial derivatives $\partial D/\partial C_i$. The following analysis will be in terms of the integral form for $D$, Eq. (38); the corresponding analysis for the sum, Eq. (39), is analogous.

$$\frac{\partial D}{\partial C_i} = -2 \int G_i(r) \left[ f(r) - \sum_{i=0}^{n-1} C_i G_i(r) \right] w(r) \, dr$$

$$= 2 \sum_{j=0}^{n-1} C_j A_{ij}, \quad (i = 0, 1, ..., n - 1),$$  \hfill (41)

where $C_n = 1$ and

$$A_{ij} = \int G_i(r) G_j(r) w(r) \, dr, \quad (i, j = 0, 1, ..., n - 1),$$  \hfill (42)

$$A_{nn} = -\int G_i(r) f(r) w(r) \, dr, \quad (i = 0, 1, ..., n - 1).$$

We thus have a set of $n$ simultaneous linear equations for the $C_i$,

$$\sum_{j=0}^{n} A_{ij} C_j = 0, \quad (i = 0, 1, ..., n - 1).$$  \hfill (43)

These equations are inhomogeneous (for $C_n = 1$ is not an unknown) and symmetric ($A_{ij} = A_{ji}$), and can be solved by any of the standard techniques. The deviation $D$ can then be written as,

$$D = 2 \sum_{i=0}^{n} \sum_{j=0}^{n} C_i C_j A_{ij},$$  \hfill (44)

where

$$A_{nn} = \int f^2(r) w(r) \, dr.$$  \hfill (45)

When Eqs. (43) are satisfied, this simplifies to

$$D = 2 \sum_{i=0}^{n} C_i A_{nn}.$$  \hfill (46)

The value of $D$ when Eqs. (43) are satisfied is called the least-squares deviation and will be denoted by $D$. 
2. Integral Transforms

In the previous section we have considered the approximate expansion of a given function \( f(r) \) in a finite, discrete set of Gaussians,

\[
f(r) = \sum_i C_i \exp\left(-\alpha_i r^2\right).
\]

Alternatively, in some cases, we may be able to make use of an analytically derived continuous expansion of the type

\[
f(r) = \int_{0}^{\infty} C(\alpha) \exp\left(-\alpha^2 r^2\right) d\alpha.
\]

Such transforms are available for many functions \( f(r) \). An obvious example, which will be found useful in the evaluation of some molecular integrals (Section IV, B), is

\[
\frac{1}{r} = \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \alpha^{-1/2} \exp\left(-\alpha r^2\right) d\alpha, \quad (r > 0),
\]

and, in general (compare Eq. (18)),

\[
\frac{1}{r^a} = \frac{1}{\Gamma(a/2)} \int_{0}^{\infty} \lambda^{a/2-1} e^{-\lambda r^2} d\lambda, \quad (\lambda > 0, r > 0).
\]

Many others can be generated by the substitution \( t = r^2 \) in the Laplace transforms

\[
f(t) = \int_{0}^{\infty} \exp\left(-ts\right) g(s) ds.
\]

Numerous transforms of the type (53) have been listed (see, e.g., Erdélyi et al., 1954; particularly page 146). The following is a useful example (compare Eq. 21):

\[
2(a/4t)^{\nu/2} K_{\nu}(\sqrt{at}) = \int_{0}^{\infty} e^{-s} s^{\nu-1} e^{-a^{2}s/4} ds,
\]

\( K_{\nu}(x) \) being the modified Bessel function of the second kind (see Watson, 1952). Noting that \( K_{\nu}(x) = K_{-\nu}(x) \), and making the substitutions

\[ t = r^2, \quad a = \alpha^2, \quad \lambda = -\nu, \]

we find,

\[
h_{\nu}(\alpha r) = (\alpha r)^{\nu} K_{\nu}(\alpha r) = \frac{1}{2} (\alpha^2/2)^{\nu} \int_{0}^{\infty} s^{\nu-1} e^{-\alpha^2 s/4} e^{-sr^2} ds.
\]
The reduced Bessel functions $k_0(\alpha r)$, for $\lambda > 0$, are regular at the origin and decrease exponentially as $r \to \infty$. For half-integral values of $\lambda$ they reduce to polynomials in $\alpha r$ multiplied by $\exp(-\alpha r)$,

$$k_{0.2}(\alpha r) = \sqrt{\pi/2} \exp(-\alpha r),$$  \hspace{1cm} (56)

$$k_{0.4}(\alpha r) = \sqrt{\pi/2} (1 + \alpha r) \exp(-\alpha r),$$  \hspace{1cm} (57)

and, in general,

$$k_{n+1/2}(\alpha r) = \frac{\sqrt{\pi}}{2^{n+1/2}} \sum_{i=0}^n \frac{(2n - i)!}{(n - i)! n!} (2\alpha r)^i e^{-\alpha r}. \hspace{1cm} (58)$$

Thus the functions $k_\lambda(\alpha r)$, with variable $\lambda$, may be used as the radial part of a generalized type of orbital in atomic and molecular calculations. They could be considered as a form of "fractional quantum number" orbitals, somewhat different from the $r^l \exp(-\alpha r)$ orbitals suggested by Parr and Joy (1957), but offering the advantage, for molecules, that the transform of Eq. (55) would be helpful in evaluating the required integrals (Section IV, C).

The special case of Eq. (55) with $\lambda = 1/2$ is of particular interest to us, as it will be used in the evaluation of molecular integrals between ordinary exponential-type orbitals. Using Eq. (56), we obtain the transform (compare Eq. 20)

$$e^{-\alpha r} = (\alpha/2) \sqrt{\pi} \int_0^\infty s^{3/2} e^{-\alpha^2 s^2} e^{-s t} ds. \hspace{1cm} (59)$$

A. Statement of the Problem

The compressibility behavior of a fluid is usually described by means of an "equation of state," which relates the pressure $p$ with the molar volume $V$ and the absolute temperature $T$ (see, e.g., Hirschfelder et al., 1954). The simplest such equation is the well-known perfect gas law,

$$pV = RT,$$ \hspace{1cm} (60)

$R$ being the "gas constant."

Any closed form equation of state is of necessity only approximate, and can represent the actual fluid with reasonable accuracy over at most a limited range of the variables. To construct an equation which could be valid over a very wide range of conditions we resort to an infinite series expansion, the so-called virial equation of state,

$$\frac{pV}{RT} = 1 + \frac{B(T)}{V} + \frac{C(T)}{V^2} + \frac{D(T)}{V^3} + \ldots \hspace{1cm} (61)$$

$B(T), C(T), D(T), \ldots$ are called the second, third, fourth, ... virial coefficients, respectively. Their temperature dependence is not explicitly stated, and is to be determined empirically by fitting Eq. (61) to the available experimental data. (The question of the range of convergence of the virial expansion will not be discussed here.)

From the theoretical point of view, the most important property of the virial equation of state is that, by means of statistical mechanics, the virial coefficients can be directly related to the law of force between the individual molecules of the fluid (Mayer and Mayer, 1940). The law of force is expressed by means of the intermolecular potential $U(r_{ij})$, which is defined as the work performed in bringing the molecules $i$ and $j$ from an initially infinite separation to the relative position described by $r_{ij}$. In general, $r_{ij}$ specifies both the intermolecular distance $r_{ij}$ and the relative orientation, but we shall only be concerned here with molecules having spherical potentials $U(r_{ij})$. In the formulation given here we assume pairwise-additive potentials, namely, that the total intermolecular potential energy in a fluid is equal to the sum of all pair interactions, each of these interactions being independent of the presence of the other molecules. This assumption is probably quite valid except at rather high densities (Hirschfelder et al., 1954).

The equations relating the first few virial coefficients to the intermolecular potential, for spherical molecules, are (Mayer and Mayer, 1940)

$$B(T) = -\frac{N}{2V} \int_0^\infty \int_0^\infty f_{13} dr_1 dr_3, \hspace{1cm} (62)$$

$$C(T) = -\frac{N^2}{3V} \int_0^\infty \int_0^\infty \int_0^\infty f_{13} f_{23} f_{13} dr_1 dr_2 dr_3, \hspace{1cm} (63)$$

$$D(T) = -\frac{N^3}{8V} \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \left(3f_{13} f_{24} f_{31} f_{41} + 6f_{13} f_{24} f_{31} f_{41} + f_{13} f_{23} f_{14} f_{34} f_{13} + f_{13} f_{23} f_{14} f_{34} f_{13} \right) \times dr_1 dr_2 dr_3 dr_4, \hspace{1cm} (64)$$

where

$$f_{ij} = f(r_{ij}) = \exp \left[ -U(r_{ij})/kT \right] - 1, \hspace{1cm} (65)$$
$N$ is Avogadro's number (the number of molecules per mole), and $k$ is the Boltzmann constant, $k = R/N$. The general form of $U(r)$ and $f(r)$ is shown in Fig. 2. Each integration $dr$ is over the three space coordinates of the corresponding molecule, and extends over the volume $V$. The actual value of $V$ is immaterial, as long as it is large compared with the range of the intermolecular forces; for, employing the technique used in reducing the number of variable points in the integrals of Section II, C, 2, we can integrate over the coordinates of one molecule directly, yielding a factor $V$

$$B(T) = -\frac{N}{2} \int f_{12} dr_2 = -2\pi N \int_0^{r_{\min}} f(r) r^2 dr$$

(66)

$$C(T) = -\frac{N^2}{3} \int \int f_{12} f_{23} f_{13} dr_2 dr_3$$

(67)

$$D(T) = -\frac{N^3}{8} \int \int \int (3f_{12} f_{23} f_{14} + 6f_{12} f_{24} f_{13} f_{24} + f_{12} f_{24} f_{13} f_{24} f_{14})$$

(68)

$x \times dx_3 dx_4 dx_5$

The limits of the remaining integrations can be taken as infinite.

The intermolecular potential $U(r)$ is generally expressed in some analytical form having a number of adjustable parameters. The best known of these forms is the Lennard-Jones potential,

$$U(r) = 4\epsilon (r_0/r)^{12} - (r_0/r)^6$$

(69)

its two parameters $\epsilon$ and $r_0$, which satisfy the relations $\epsilon = -U_{\text{min}}(r)$, $U(r_0) = 0$ (see Fig. 2), can be considered as scale factors. More elaborate forms, having more parameters, have also been used (Hirschfelder et al., 1954).

Having chosen a reasonable analytical form for $U(r)$, we would proceed, if possible, to compute the macroscopic properties, such as compressibility and viscosity, as a function of the parameters by the methods of statistical mechanics. Then, by comparing the results with experimental data, we would determine the best values of the parameters for each fluid, and thus obtain information about the intermolecular forces in it.

Compressibility data, in the form of the virial coefficients, is of primary importance in any analysis of the above type. Its application requires the evaluation of the integrals of Eqs. (66 — 68) for various potentials. The expression for $B(T)$, Eq. (66), is fairly simple, and even if an analytical solution for some assumed form of $U(r)$ may not be possible, it would be quite easy to perform the one-dimensional integration numerically. However, the second virial coefficients is markedly insensitive to the detailed shape of the potential, beyond its two over-all scale factors (Boys and Shavitt, 1960b). We have to proceed to the third, and perhaps the fourth, virial coefficients to obtain a closer correlation.

The evaluation of the third coefficient, though considerably more laborious than the second, no longer presents great difficulty with modern computing equipment. The six-dimensional integral, Eq. (67), can be reduced to three dimensions (Kihara, 1948) and then evaluated numerically. This has been done even on very slow and elementary equipment (Bird et al., 1950) for the Lennard-Jones potential. However, any similar technique would probably be too impractical for higher virial coefficients. Some indication of the possible behavior of $D(T)$ was obtained by S. Katsura (1959) by the use of a drastically simplified form of $U(r)$, the "square well" potential. The estimation of $D(T)$ for arbitrary potentials has first been made possible by the introduction of the Gaussian expansion techniques described in Section II. D. I (Boys and Shavitt, 1960a, b). More recently, a new series expansion approach has been used successfully for the calculation of the fourth virial coefficient by Barker and Monaghan (1962).

B. INTRODUCTION OF GAUSSIANS

It is evident from the discussion of integrals involving products of Gaussian functions (Section II, C, 2) that if an expansion of $f(r)$, Eq. (65), in a reasonably small number of Gaussians can be obtained,
then the calculation of \( C(T) \) and \( D(T) \) for any potential would be quite feasible. Thus (Boys and Shavitt, 1960a), if

\[
f(r) \approx \sum_{l=0}^{\infty} C_l G_l(r),
\]

(70)

then, by Eqs. (35), (36), and their analogues (compare Eq. (16)),

\[
C(T) = -\frac{1}{3} \pi^2 N^2 \sum_{i,l,i=0}^{i=0} C_i C_j C_k T_{ijkl},
\]

(71)

\[
D(T) = D_1(T) + D_4(T) + D_6(T)
\]

\[
= -\frac{3}{8} \pi^2 N^2 \sum_{i,j,k,l} C_i C_j C_k C_l T_{ijkl}
\]

\[
- \frac{3}{4} \pi^2 N^2 \sum_{i,j,k,l,m} C_i C_j C_k C_l C_m T_{ijklm}
\]

\[
- \frac{1}{8} \pi^2 N^2 \sum_{i,j,k,l,m,n} C_i C_j C_k C_l C_m C_n T_{ijklmn},
\]

(72)

where,

\[
T_{ijkl} = (\alpha_{ij} + \alpha_{lk} + \alpha_{ik})^{-3/2},
\]

(73)

\[
T_{ijkl} = (\alpha_{ij} \alpha_{lk} + \alpha_{ik} \alpha_{lj} + \alpha_{ik} \alpha_{lj} + \alpha_{ij} \alpha_{lk})^{-3/2},
\]

(74)

\[
T_{ijklm} = (\alpha_{ij} \alpha_{lk} + \alpha_{ik} \alpha_{lj} + \alpha_{ik} \alpha_{lj} + \alpha_{ij} \alpha_{lk})^{-3/2},
\]

(75)

\[
T_{ijklmn} = (\alpha_{ij} \alpha_{lk} \alpha_{km} + \alpha_{ij} \alpha_{lk} \alpha_{mn} + \alpha_{ij} \alpha_{lk} \alpha_{mn} + \alpha_{ij} \alpha_{lk} \alpha_{mn})^{-3/2}.
\]

(76)

The total number of terms in the sums for \( C(T), D_1(T), D_4(T), \) and \( D_6(T) \) is \( \nu^3 \), \( \nu^4 \), \( \nu^5 \), and \( \nu^6 \), respectively. However, many of these terms are equal to each other, because some permutations of the indices (all permutations in the case of \( C \) and \( D_1 \)) do not affect a term's value. The full utilization of these symmetry properties results in very significant savings in the computation; e.g., for \( \nu = 6 \), the total number of \( D_6 \) terms in 46656, while the number of numerically distinct terms is 2451

only. The techniques used in handling the summations automatically and efficiently on the computer are elaborated in the next section.

Turning our attention to the function \( f(r_0) \), Eq. (65), which is to be expanded in Gaussians, we see that it is dependent on the temperature \( T \) as well as on the details of the potential. The use of reduced units can minimize the number of individual expansions (and therefore the number of virial coefficient calculations) that are needed to cover a given range. We define the reduced temperature as

\[
T^* = kT/\epsilon
\]

(77)

(compare Fig. 2), and the reduced virial coefficients,

\[
C^*(T^*) = C(T)/b_0^2,
\]

(78)

\[
D^*(T^*) = D(T)/b_0^3,
\]

(79)

where \( b_0 = \frac{3}{8} \pi N r_0^3 \) (this is equal to the value of \( B(T) \) for hard spheres of radius \( r_0 \)), \( T^*, C^*, \) and \( D^* \) are dimensionless quantities; thus we can use \( r_0 = 1 \) as the unit of length in \( U(r) \) and \( f(r) \). If we have chosen for study an intermolecular potential having \( n \) parameters (including the two scale factors \( \epsilon \) and \( r_0 \)), then \( f(r) \) and the reduced virial coefficients will depend on \( T^* \) and \( n - 2 \) parameters only.

The number \( \nu \) of expansion functions required to obtain a reasonably good fit for \( f(r) \) depends very strongly upon \( T^* \). The function \( f(r) \) is particularly steep around \( r/r_0 = 1 \), and changes direction sharply at
both ends of this steep section. These characteristics are especially marked at small \( T^* \) (see Fig. 2) and make it very difficult to obtain a satisfactory Gaussian expansion. Examples of two 6-term expansions obtained by Boys and Shavitt (1960a) for the Lennard-Jones potential at \( T^* = 1.333 \) and at \( T^* = 20 \) are shown in Fig. 3. These were derived by the methods of Section II, D, 1, employing a weight function \( w(r) = r^2 \), so that accuracy around the origin was sacrificed in order to improve the fit at the more voluminous region of larger \( r \). The values of the parameters for these expansions are given in Table I. The relative least-squares deviations,

\[
d = D|A, \nu,
\]

(compare Eq. (45)), are 0.0626 for the lower temperature, and 0.0066 for the higher. Even though the spurious fluctuations in the expansions might cancel out to some extent in the integrations for the virial coefficients, the quality of the fit at \( T^* = 1.333 \) clearly leaves much to be desired. However, the limitations of the computer which was used in these calculations (EDSAC I, at the Cambridge University Mathematical Laboratory, with a fixed point addition time of 1.5 msec and a storage capacity of 512 35-bit words) made it very difficult to derive expansions of more than 8 terms and quite impractical to use expansions of more than 6 terms for calculations of \( D(T) \). The difficulties were mostly due to the large values of the coefficients \( C_i \) (Table I), causing serious round-

### Table I

**Details of Two Expansions.**

<table>
<thead>
<tr>
<th>( T^* )</th>
<th>( i )</th>
<th>( a_i )</th>
<th>( C_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.333</td>
<td>1</td>
<td>3.52</td>
<td>3210.48652</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.94</td>
<td>-7055.65397</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.08</td>
<td>8368.44455</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6.60</td>
<td>-7919.06846</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>8.28</td>
<td>4131.08282</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>10.90</td>
<td>-737.58270</td>
</tr>
<tr>
<td>20.000</td>
<td>1</td>
<td>4.36</td>
<td>161.40422</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.94</td>
<td>-1941.57541</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>7.22</td>
<td>5913.18649</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>8.50</td>
<td>-7365.09262</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>9.90</td>
<td>3728.78501</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>12.44</td>
<td>-497.92027</td>
</tr>
</tbody>
</table>

off problems both in the derivation of the expansions and in the summations for the virial coefficients, Eqs. (71) and (72). As a result, quadruple-precision arithmetic had to be used in the calculation of the \( T_{ij} \) terms and their sums. These problems are discussed further in the next section.

Some representative results for the virial coefficients of the Lennard-Jones potential are shown in Table II. Values of the relative least squares deviation \( d, \) Eq. (80), and of \( C^*(T^*) \) and \( D^*(T^*) \), are shown for different numbers of expansion terms \( \nu \) at two values of the reduced temperature. The more accurate numerical integration results for \( C^*(T^*) \) of Bird et al. (1950) and the recent results of Barker and Monaghan (1962) for \( D^*(T^*) \) are shown for comparison. Similar calculations were carried out for other temperatures and for other potentials; for complete results the reader is referred to the original papers (Boys and Shavitt, 1960a, b).

### Table II

**Virial Coefficient Calculations at Two Reduced Temperatures.**

(Lennard-Jones Potential)

<table>
<thead>
<tr>
<th>( T^* )</th>
<th>( \nu )</th>
<th>( d )</th>
<th>( C^<em>(T^</em>) )</th>
<th>( D^<em>(T^</em>) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.333</td>
<td>3</td>
<td>0.233</td>
<td>0.696</td>
<td>0.336</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.126</td>
<td>0.708</td>
<td>0.330</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.084</td>
<td>0.664</td>
<td>0.219</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.063</td>
<td>0.639</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.050</td>
<td>0.603</td>
<td></td>
</tr>
<tr>
<td>Bird et al.</td>
<td></td>
<td></td>
<td>0.5823</td>
<td></td>
</tr>
<tr>
<td>Barker and Monaghan</td>
<td></td>
<td></td>
<td>0.292 ± 0.020</td>
<td></td>
</tr>
<tr>
<td>20.000</td>
<td>3</td>
<td>0.056</td>
<td>0.250</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.023</td>
<td>0.245</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.011</td>
<td>0.244</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.007</td>
<td>0.243</td>
<td>0.082</td>
</tr>
<tr>
<td>Bird et al.</td>
<td></td>
<td></td>
<td>0.2464</td>
<td></td>
</tr>
<tr>
<td>Barker and Monaghan</td>
<td></td>
<td></td>
<td>0.084 ± 0.006</td>
<td></td>
</tr>
</tbody>
</table>

### C. Computational Problems and Solutions

1. **The Use of Symmetry in the Summations**

The efficient summation of Eqs. (71) and (72) requires full use of the invariance of the terms \( T_{ij} \) with respect to permutations of the indices. This is fairly simple in the case of \( C(T) \) and \( D_i(T) \), which have full
permutational symmetry. In these cases we can write the sums in the form

\[ C(T) = -\frac{3}{2} \pi^2 N^3 \sum_{i \geq j \geq k \geq l = 0} \sigma_{ijk} C_i C_j C_k T_{ijkl} \]  

\[ D(T) = -\frac{3}{8} \pi^{6/2} N^3 \sum_{i \geq j \geq k \geq l = 0} \sigma_{ijkl} C_i C_j C_k C_l T_{ijkl} \]  

(81)  

(82)

where the “symmetry factors” \( \sigma_{ijk} \) and \( \sigma_{ijkl} \) are equal to the number of distinguishable permutations of the actual values of the indices. For example, for \( C(T) \),

\[ \sigma_{ijk} = 1 \quad \text{if } i = j = k, \]

\[ \sigma_{ijk} = 3 \quad \text{if } i = j \text{ or } j = k, \]

\[ \sigma_{ijk} = 6 \quad \text{if all indices are different}. \]

The case of \( D(T) \) and \( D(T) \) is more complicated, since not all index permutations would leave the corresponding terms invariant. The “allowed” symmetry operations of \( T_{ijklmn} \), Eq. (76) (i.e., those under which the term is invariant), are analogous to the 24 symmetry operations (rotations and reflections) of a regular tetrahedron in which the pairs of indices \( ij, kl \), and \( mn \) have been assigned to pairs of opposite edges (Fig. 4). Reflection of the tetrahedron in a plane of symmetry thus corresponds to the interchange of one of the above index pairs with another, keeping the order within each pair. Rotation around a two-fold axis results in the simultaneous interchange of the two indices within each of two pairs. All combinations of these (including the identity operation) give rise to the 24 allowed permutations. Similarly, the 8 allowed symmetry operations of \( T_{ijklmn} \), Eq. (75), correspond to those rotations and reflections (of the same tetrahedron) which cause the edge marked \( m \) to transform into itself.

The general problem of such a “multi-index” summation can be broken down into two parts:

1. Evolve a scheme for stepping the summation indices in such a way that only essentially different terms will be generated.

2. For each term generated, determine the symmetry factor \( \sigma \), i.e., the number of distinguishable permutations of the index values which give rise to equivalent terms.

To solve the first part, we define the “signature” of a term \( T_{ijklmn} \) as the number obtained by regarding the sequence of single digit values of the indices \( i, j, \ldots \), in any number system with base \( \geq \nu \), as a single integer. Thus, if \( \nu \leq 10 \) we can use the decimal number system, and the signature of \( T_{5,0,2} \) is the integer 502. The arbitrary decision is then made that out of each set of terms which are equal by symmetry we select for computation that one which has the highest valued signature. It is usually fairly easy to write down the conditions that must be satisfied in order that the value of the term’s signature cannot be increased by any allowed permutation. In the case of \( D(T) \) these rules are

a. \( i \geq \) any other index;

b. \( k \geq \) any of the indices \( l, m, n \);

c. if \( i = k \) then \( j \geq l \);

d. if \( k = m \) then \( l \geq n \);

e. if \( l \geq m \) then \( k \geq n \);

f. if \( l < m \) then \( k > n \);

g. if \( i = j \) and/or \( k = l \) then \( m \geq n \).

It is convenient to start the summation at the upper bound, with \( i = j = k = l = m = n = \nu - 1 \), and step the indices down through all combinations which satisfy the above rules until all are zero; for it is easier, on most computers, to test whether a number is zero than to check whether it has reached a given limit.

The symmetry factor \( \sigma \) depends on the number of equalities between the values of the indices, and rules for its determination can usually be formulated without much trouble. For \( D(T) \) it can be verified that

\[ \sigma = 24/(tr), \]
where,

(a) \( t = 4 \) if \( i = j, k = l \), and \( m = n \) simultaneously;
(b) \( t = 2 \) if only two of these equalities hold;
(c) \( t = 1 \) otherwise;
(d) \( r = 6 \) if \( (i, j) = (k, l) = (m, n) \);
(e) \( r = 2 \) if \( (i, j) = (k, l) \) or \( (k, l) = (m, n) \);
(f) \( r = 1 \) otherwise.

The time spent by the computer in applying this type of rule is usually only a minute fraction of the total computing time of the sums. Evidently, the full utilization of symmetry in multi-index summations can make the difference between a feasible calculation and an impractically long one.

2. Round-Off Difficulties

An examination of the fitting parameters in Table I shows that the approximate expansion of \( f(r) \) by Gaussians could only be achieved by the use of very large coefficients with alternating signs. While the individual coefficients of an expansion approach 10,000 in value, their algebraic sum is of the order of unity. This situation results in severe round-off problems, first in the solution of the simultaneous equations for the least-squares process, and then, to a much greater extent, in the final summations.

The main reason for this behavior, and for the relatively low accuracy of the expansions, can probably be found in the nature of \( f(r) \) around the point \( r = r_0 \), as discussed in Section III, B. A comparison of fittings for different temperatures (Fig. 3) and for potentials of different steepness (Boys and Shavitt, 1960b) clearly supports this view. Therefore we may hope that difficulties of the same magnitude will not arise when the Gaussian expansion method is employed to solve many other problems in which smoother functions are involved.

The round-off problems in the least-squares fitting process can be handled fairly easily on present-day computers by the use of multi-precision arithmetic. On EDSAC I, with its relatively low speed and small storage capacity, it was found more practical to derive the expansions in two single-precision stages. A rough solution was obtained first, its deviation from \( f(r) \) was then re-expanded in the same set of Gaussians, and the resulting coefficients were added on to the original ones. This process (in interpretive floating point), including the calculation of the least squares deviation and its gradient in \( \alpha_r \)-space, took about 15 minutes for a single 6-term expansion. It is estimated that a similar calculation on an IBM 7090 computer could be performed in less than one second. Initially over 20 fittings may have to be made in order to locate a reasonably satisfactory set of Gaussians, but as experience is gained only 3 to 5 may be enough for each new temperature and potential. The process of searching for the optimum \( \alpha_r \) set, which was done through operator judgment and intervention on EDSAC, would have to be automated for efficient use of a faster computer.

Extreme round-off difficulties were encountered in the multi-index summations. Because of the size of the coefficients \( C_\alpha \), individual terms could be as large at \( 10^{24} \), while their sum was of the order of \( 10^{-4} \) or \( 10^{-3} \). Quadruple-precision arithmetic had to be employed in order to obtain meaningful results, and this was particularly time consuming in the evaluation of the \( -3/2 \) power required for each term. As a result, the summation process was extremely slow; each calculation of \( D(T) \) with a 6-term expansion (for one temperature) took over two hours. The same process on the IBM 7090 could probably be completed in about half a minute; even the use of a 10-term expansion would only raise this to about 10 minutes, which is certainly quite practical.

IV. A Problem in Quantum Mechanics: Multicenter Integrals

A. Statement of the Problem

The quantum-mechanical prediction of the properties of molecules is one of the primary goals of theoretical chemistry, and one of the most difficult to achieve. With the rapid advances in electronic computers in recent years, an increasing number of theoretical studies have been carried out on diatomic molecules involving first-row atoms (see, e.g. Ransil, 1960; Hijiwata, 1961). Only few comparable investigations of polyatomic systems have been made (e.g. Foster and Boys, 1960). The main stumbling block to \( \alpha \)-priori quantum-mechanical calculations for polyatomic molecules has been the difficulty in the evaluation of the required multicenter integrals; it has been estimated (Boys and Cook, 1960) that these integrals account for about 70 per cent of the total computation time in any such treatment. We shall attempt to show here that the unique properties of the Gaussian function can be of great help in reducing the magnitude of this problem.

Ignoring those integrals whose evaluation is fairly simple, we are concerned mainly with two types: nuclear attraction integrals,

\[
(\Phi_{\alpha A} | V_C | \Phi_{\beta B}) = \int \Phi_{\alpha}^*(\mathbf{r}_{1A}) \Phi_{\beta} \left( \mathbf{r}_{1B} \right) (1/r_{1C}) d\mathbf{r}_{1},
\]

(83)
and interelectron repulsion integrals,
\[
[\phi_{aA}\phi_{aB}; \phi_{cC}\phi_{dD}] = \int \frac{\phi^*_a(r_{1A})\phi^*_b(r_{1B})}{r_{12}} \phi_c(r_{2C})\phi_d(r_{2D}) dr_1 dr_2.
\] (84)

\(A, B, C, D\) designate fixed centers (which usually coincide with positions of nuclei), \(r_{1A}\) is the vector from center \(A\) to electron 1, etc., \(r_{1B}\) is the distance between electrons 1 and 2, and \(a, b, c, d\) specify the functions (orbitals) involved. A product such as \(\phi^*_a(r_{1A})\phi^*_b(r_{1B})\) describes an electron charge distribution. The 3-dimensional integral in Eq. (83) represents the Coulombic interaction of a charge distribution with a point charge at \(C\), while the 6-dimensional one, Eq. (84), corresponds to the interaction between two distributions. Each charge distribution is either single-center or two-center, according to whether its two orbitals share a common center or not.

The set of orbitals used in any one calculation is termed the basis set; the number and nature of the functions in this set is probably the most important factor influencing the results of the treatment, for the total wave function of the molecule is expanded in terms of these orbitals. The most frequently used basis sets (see Allen and Karo, 1960) are made up of Slater type orbitals (STO; also called exponential type orbitals, or ETO). The first few orbital types in this set are

\[
\begin{align*}
(1s) &= \sqrt{\alpha^3/\pi} \exp (-\alpha r), \\
(2s) &= \sqrt{\alpha^4/3\pi} r \exp (-\alpha r), \\
(2p_x) &= \sqrt{\alpha^5/\pi} x \exp (-\alpha r), \\
(3s) &= \sqrt{2\alpha^7/45\pi} r^2 \exp (-\alpha r), \\
(3p_x) &= \sqrt{2\alpha^7/15\pi} x r \exp (-\alpha r), \\
(3d_{xy}) &= \sqrt{2\alpha^7/3\pi} xy \exp (-\alpha r), \\
(3d_{x^2-y^2}) &= \sqrt{\alpha^7/6\pi} (x^2 - y^2) \exp (-\alpha r), \\
(3d_{z^2}) &= \sqrt{\alpha^7/2\pi} (z^2 - r^2/3) \exp (-\alpha r).
\end{align*}
\]

The “exponential parameters” \(\alpha\) differ, in general, for different orbitals; ideally, they should be varied individually until the best set (in terms of total energy and wave function) is found. Among basis sets of simple analytical form, Slater type orbitals have been found to give the most reasonable convergence in wave-function calculations. Their principle disadvantage is the difficulty in the evaluation of the corresponding multicenter integrals. On the other hand, the use of Gaussian type orbitals (GTO), as described in the next section, practically eliminates the integral problems, but this is done at the cost of much poorer wave function convergence.

The relative difficulty in the calculation of the integrals depends, apart from the form of the orbitals, on the number and arrangement of the different centers involved. Single-center integrals and many two-center ones are comparatively easy and have closed-form solutions (see Roothaan, 1951). The interelectron repulsion integrals, Eq. (84), are classified for computational purposes as follows:

(a) Coulomb integrals—both charge distributions are single-center; the integrals can be single-center or two-center, and are relatively easy.

(b) Hybrid integrals—one charge distribution is single-center, the other two-center; the integral can be two-center or three-center. Some of the two-center ones are not too difficult.

(c) Exchange integrals—both charge distributions are two-center; the integral can be two-, three-, or four-center. This is the most difficult type.

Of the nuclear attraction integrals, only the three-center ones are troublesome.

Some of the techniques which have been used for the evaluation of multicenter integrals between STO's are

(a) The \(\zeta\)-function method (Barnett and Coulson, 1951; see also pages 95 ff. and 155 ff., in this Volume); it is based on an expansion of all orbitals about a single center.

(b) Axial expansion method (Boys and Reeves, to be published; see Boys and Shavitt, 1959); each charge distribution is expanded in a series of exponentials with centers distributed along its axis.

(c) Numerical integration (McLean and Sinai, private communication; Magnusson and Zauli, 1961).

(d) Gaussian expansion (Boys and Jones, to be published); the exponential function in each orbital is expanded by least squares in a number of Gaussians.

(e) Gaussian transform (Shavitt and Karplus, 1962); based on the integral transform of Eq. (59). This method, which is described in detail in Section IV, C, seems at present to be the most efficient.
B. First Approach: Gaussian Type Orbitals

Gaussian functions of the type \(x^a y^m z^n \exp (-ar^2)\) and their linear combinations were first proposed for basis set orbitals in molecular calculations by Boys (1950), who derived the necessary integral formulas. The use of a limited GTO basis set, in which each atomic orbital is represented by a single Gaussian function (e.g., Meckler, 1953) has led to rather disappointing results. However, calculations with extended GTO basis sets (Nesbet, 1960) have been more encouraging. A comparison of the relative numbers of exponents and Gaussians required to obtain a reasonable fit for numerical Hartree-Fock orbitals of a few atoms (Allen, 1962) indicates that extended GTO calculations may be quite competitive with extended ETO ones, considering the much easier integrability of the Gaussians. An even more favorable situation may result if each of the basis set orbitals is taken as a linear combination of a few Gaussian functions; such a linear combination could be chosen initially so as to approximate an atomic Hartree-Fock orbital (or any desirable modification of it), but once chosen, its coefficients would not be varied, thus keeping the total number of independent orbitals fairly small. This is analogous to the use of analytical approximations of numerical Hartree-Fock orbitals in terms of exponents for basis set functions in atomic configuration interaction calculations (e.g., Donath, 1961).

A generalization of Boys’ original Gaussian orbitals was proposed by Singer (1960); his many-electron wave function is of the form \(\sum P_k \exp (-Q_k)\), where \(Q_k\) is any positive definite quadratic form in the space coordinates of all the electrons, and \(P_k\) is a polynomial in all or some of these coordinates. The inclusion of cross terms in \(Q_k\) amounts to an explicit dependence of the wave function on interelectron distances, and improves its convergence considerably while retaining the easy integrability of simple Gaussians. An application of this technique to the hydrogen molecule (Longstaff and Singer, 1960) has been quite successful. A simpler version of this approach, the inclusion of Gaussian correlation functions \(\exp (-ar^2)\) together with ordinary GTO basis sets, has been proposed by Boys (1960). Browne and Poshusta (1962) have suggested the use of ellipsoidal Gaussian orbitals \(\exp (-ar^2 - \beta r^2 - \gamma z^2)\); these may be particularly appropriate for “bond orbitals,” with the long axis of the Gaussian connecting the two bound atoms.

The following formulas were derived by Boys (1950) for the integrals between basic (1s-type) Gaussian orbitals, \((aA) = \exp (-ar^2_{1A})\), etc.,

\[
(aA \mid kB) = \int \exp (-ar^2_{1A} - br^2_{1B}) \exp \left(-\frac{ab}{a + b} \frac{r^2}{AB}\right) dr_1, \quad \text{(85)}
\]

\[
(aA \mid V_C \mid bB) = \int \left(1/r_{1C}\right) \exp (-ar^2_{1A} - br^2_{1B}) dr_1
\]

\[
= \frac{2\pi}{a + b} F_0 \left[(a + b) CP^2 \exp \left(-\frac{ab}{a + b} \frac{r^2}{AB}\right)\right], \quad \text{(87)}
\]

\[
[aAbB; cCdD] = \int \left(1/r_{13}\right) \exp (-ar^2_{1A} - br^2_{1B} - cr^2_{2C} - dr^2_{2D}) dr_1 dr_2
\]

\[
= \frac{2\pi a^2}{(a + b)(c + d)} \sqrt{a + b + c + d} \exp \left(-\frac{ab}{a + b} \frac{r^2}{AB} - \frac{cd}{c + d} \frac{r^2}{CD}\right)
\]

\[
\times \exp \left(-\frac{ab}{a + b} \frac{r^2}{AB} - \frac{cd}{c + d} \frac{r^2}{CD}\right), \quad \text{(88)}
\]

The points \(P\) and \(Q\) lie on \(AB\) and \(CD\), respectively, and are given by,

\[
P_x = \frac{aA_x + bB_x}{a + b}, \quad Q_x = \frac{cC_x + dD_x}{c + d}, \quad \text{etc.} \quad \text{(89)}
\]

The function \(F_0\) is the one defined in Section II, C, 1, Eq. (22). All these formulas remain valid when any two or more centers coincide.

As an example, we shall derive the formula for the exchange integral, Eq. (88). The proof given here (compare Singer, 1960) serves as a demonstration of the usefulness of Gaussian transforms in multidimensional integrations. As a first step, we apply the theorem on a product of Gaussians, Eq. (5), to each of the two charge distributions, thus converting the four-center exchange integral into a two-center Coulomb one,

\[
[aAbB; cCdD] = \exp \left(-\frac{ab}{a + b} \frac{r^2}{AB} - \frac{cd}{c + d} \frac{r^2}{CD}\right) \int \left(1/r_{13}\right)
\]

\[
\times \exp \left(-r^2_{1P} - r^2_{3Q}\right) dr_1 dr_2, \quad \text{(90)}
\]

where \(p = a + b, q = c + d\). Then we introduce the transform of Eq. (51),

\[
\frac{1}{r_{12}} = \frac{1}{\sqrt{\pi}} \int_0^\infty s^{-1/2} \exp (-ar^2) ds,
\]

and factor the integral into \(x, y, \) and \(z\) parts

\[
[aAbB; cCdD] = \frac{1}{\sqrt{\pi}} \exp \left(-\frac{ab}{a + b} \frac{r^2}{AB} - \frac{cd}{c + d} \frac{r^2}{CD}\right) I_x I_y I_z s^{-1/2} ds, \quad \text{(91)}
\]
where

\[ I_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left[ -p(x_1 - P_x)^2 - q(x_2 - Q_x)^2 - s(x_1 - x_2)^2 \right] dx_1 dx_2, \quad (92) \]

etc. Putting \( u = x_1 - P_x, v = x_2 - Q_x, X = P_x - Q_x, \) and using Eq. (19) for the \( u \) and \( v \) integrations successively, we get

\[ I_x = \exp \left( -sX^2 \right) \int_{-\infty}^{\infty} du \exp \left[ -(p + s)u^2 - 2Xu \right] \int_{-\infty}^{\infty} dv \]

\[ \times \exp \left[ -(q + s)v^2 + 2(X + u)v \right] \]

\[ = \sqrt{\frac{\pi}{q + s}} \exp \left( -sX^2 \right) \int_{-\infty}^{\infty} \left[ -(p + s)u^2 - 2Xu + (X + u)q^2(q + s) \right] du \]

\[ = \sqrt{\frac{\pi}{q + s}} \exp \left( -\frac{q^2}{q + s} X^2 \right) \int_{-\infty}^{\infty} \exp \left[ -\frac{pq + (p + q)s}{q + s} u^2 - 2 \frac{qs}{q + s} Xu \right] du \]

\[ = \sqrt{\frac{\pi}{pq}} \exp \left( -\frac{q^2}{q + s} X^2 \right) \exp \left[ \frac{pq^2}{q + s} X^2 \right] \]

\[ = \sqrt{\frac{\pi}{pq}} \left( 1 + \frac{p + q}{pq} s \right)^{-3/2} \exp \left[ -\frac{sPQ^2}{1 + s(p + q)/pq} \right]. \quad (93) \]

Thus,

\[ [aAbB; cCdD] = \frac{\pi^{5/2}}{(pq)^{3/2}} \exp \left( -\frac{ab}{a + b} AB^2 - \frac{cd}{c + d} CD^2 \right) \]

\[ \times \int_{0}^{\infty} \left( 1 + \frac{p + q}{pq} s \right)^{-3/2} s^{-1/2} \exp \left[ -\frac{sPQ^2}{1 + s(p + q)/pq} \right] ds. \]

Making the substitution,

\[ 1 + \frac{p + q}{pq} s = \frac{1}{1 - t^2}, \]

we find

\[ \left( 1 + \frac{p + q}{pq} s \right)^{-3/2} s^{-1/2} ds = 2 \sqrt{\frac{pq}{p + q}} dt, \]

\[ [aAbB; cCdD] = \frac{2\pi^{5/2}}{pq\sqrt{p + q}} \exp \left[ -\frac{ab}{a + b} AB^2 - \frac{cd}{c + d} CD^2 \right] \]

\[ \times \int_{0}^{1} \exp \left[ -\frac{pq}{p + q} P^2 \right] dt, \]

which is equivalent to Eq. (88).

The basis set proposed by Boys is made up of the class of functions \( x_a \gamma_a x_a \alpha_a \exp (-ar_a^2), \) where \( i, m, n \) are integers and \( a \) is positive. Integral formulas for orbitals in which not all \( i, m, n \) are zero can be derived from the basic formulas, Eqs. (85 - 88), by differentiations with respect to parameters, for

\[ x_a \exp (-ar_a^2) = \frac{1}{2a} \frac{\partial}{\partial a} \exp (-ar_a^2), \quad \text{etc.} \quad (94) \]

For example,

\[ \left\{ (1/r_{1C}) x_{1A} \exp (-ar_{1A}^2 - br_{1B}^2) dr_1 \right\} \]

\[ = \frac{\pi}{a(a + b)} \frac{\partial}{\partial A_a} \left\{ F_0[(a + b) CP] \exp \left( \frac{-ab}{a + b} AB^2 \right) \right\} \]

\[ = \frac{2\pi}{a + b} \left\{ (C_a - P_a) F_a[(a + b) CP] - \frac{b}{a + b} (A_a - B_a) F_0[(a + b) CP] \right\} \]

\[ \times \exp \left( -\frac{ab}{a + b} AB^2 \right). \quad (95) \]

Implicit formulas for integrals involving higher orbitals have been derived by Browne and Poshusta (1962) for their ellipsoidal Gaussians, and by Harris (to be published) for products of spherical harmonics and Gaussians.

C. SECOND APPROACH: TRANSFORMATION TO GAUSSIANS

In this section we shall consider the application of the integral transform from exponentials to Gaussians, Eq. (59), to the calculation of multicenter integrals between exponential type orbitals (Shavitt and Karplus, 1962). A similar approach was first suggested by Kikuchi (1954). We shall start with integrals involving 1s orbitals only, and then show how the formulas for higher orbitals can be derived from the basic ones by differentiation with respect to parameters.

1. Integrals for 1s Orbitals

Disregarding normalization factors, the four-center exchange integral for 1s ETO’s is

\[ I = \int \int (1/r_{12}) \exp (-\alpha_{1A} - \alpha_{2B} - \alpha_{3C} - \alpha_{4D}) dr_1 dr_2. \quad (96) \]
The application of the transform of Eq. (59) to each of the four orbitals results in

\[
I = \left(\alpha_1\alpha_2\alpha_3\alpha_4/16\pi^3\right) \int_0^\infty \int_0^\infty \int_0^\infty \left(\frac{1}{s_1^2s_2^2s_3^2s_4^2}\right)^{-3/2} \exp\left(-\frac{1}{4} \left(\frac{\alpha_1^2}{s_1^2} + \frac{\alpha_2^2}{s_2^2} + \frac{\alpha_3^2}{s_3^2} + \frac{\alpha_4^2}{s_4^2}\right) \right) [s_1As_2B; s_3Cs_4D] \, ds_1 \, ds_2 \, ds_3 \, ds_4,
\]

(97)

where \([s_1As_2B; s_3Cs_4D]\) stands for a four-center exchange integral between Gaussians, Eq. (88). New variables are introduced by

\[
x = s_1 + s_2 + s_3 + s_4, \quad u = s_1/(s_1 + s_2), \quad v = s_3/(s_3 + s_4),
\]

(98)

so that

\[
s_1 = uwx, \quad s_2 = (1 - u)wx, \quad s_3 = v(1 - w)z, \quad s_4 = (1 - v)(1 - w)x.
\]

(99)

The Jacobian of this transformation is

\[
J = \left(\frac{s_1s_2s_3s_4}{w(1 - w)x}\right) = w(1 - w)x^2.
\]

(100)

When Eq. (88) is applied to the GTO exchange integral in Eq. (97), the points \(P\) and \(Q\) will be given by (see Fig. 5)

\[
P_x = \frac{s_1A_x + s_2B_x}{s_1 + s_2} = B_x + uA_x, \quad \text{etc.}
\]

(101)

\[
Q_x = \frac{s_3C_x + s_4D_x}{s_3 + s_4} = D_x + vC_x, \quad \text{etc.}
\]

(102)

\[
p_x = uA_x - vC_x + e_x, \quad \text{etc.}
\]

(103)

Using the definitions

\[
f = u(1 - u)w^2 + v(1 - v)(1 - w)^2 e^2
\]

(104)

\[
g = \frac{1}{4} \left(\frac{\alpha_1^2}{u} + \frac{\alpha_2^2}{1 - u} + \frac{\alpha_3^2}{v} + \frac{\alpha_4^2}{1 - v}\right)
\]

(105)

\[
g = u(1 - w)p^2
\]

(106)

Eq. (97) now transforms into

\[
I = \sqrt{\frac{\pi}{8}} \alpha_1\alpha_2\alpha_3\alpha_4 \int_0^1 \int_0^1 \int_0^1 \int_0^\infty \frac{F_0(qx) \exp\left(-fx - gzx\right)}{[u(1 - u)w(1 - v)][u(1 - w)]^{3/2}} \, dx \, dw \, dv \, du
\]

(107)

\[
= \sqrt{\frac{\pi}{8}} \alpha_1\alpha_2\alpha_3\alpha_4 \int_0^1 \int_0^1 \int_0^1 \frac{F_0(qx) \exp\left(-fx - gzx\right)}{[u(1 - u)][v(1 - v)][1 - u - v]} \, dx \, dw \, dv
\]

\[
J_{nm}(q, f, g)
\]

(108)

where

\[
J_{nm}(q, f, g) = \int_0^\infty x^{-n-1/2} F_m(qx) \exp\left(-fx - gzx\right) \, dx
\]

(109)

The last equality holds only if \(q \neq 0\); if \(q\) vanishes then, because \(F_m(0) = 1/(2m + 1)\), we have

\[
J_{nm}(0, f, g) = \frac{1}{2m + 1} \int_0^\infty x^{-n-1/2} \exp\left(-fx - gzx\right) \, dx
\]

(110)

by Eq. (55).

For the general case we put

\[
\sigma = \frac{f}{q} = \frac{u(1 - u)}{1 - w} \frac{a^2}{p^2} + \frac{v(1 - v)}{w} \frac{c^2}{p^2},
\]

(111)

\[
\tau = \frac{fg}{4} \left[1 - \frac{w}{u} \frac{\alpha_1^2}{1 - u} + \frac{1 - v}{v} \frac{\alpha_2^2}{1 - v} + \frac{w}{v} \frac{\alpha_3^2}{1 - v} - \frac{1 - v}{v} \frac{\alpha_4^2}{1 - v}\right]
\]

(112)

\[
\rho = 1 + \sigma
\]

(113)

\[
\kappa = 2 \sqrt{\pi r}
\]
and using the series, Eq. (25), for \( F_m(x) \), we get

\[
J_{nm}(1, \sigma, \tau) = \sum_{n=0}^{\infty} \frac{\Gamma(m + 1/2)}{\Gamma(m + i + 3/2)} \left( \frac{\tau}{\rho} \right)^{i-n+1/2} K_{i-n+1/2}(\kappa) \tag{114}
\]

by Eq. (54). In terms of the reduced Bessel function, \( k_i(x) = x^i K_i(x) \), this appears as

\[
J_{nm}(1, \sigma, \tau) = \sum_{n=0}^{\infty} \frac{\Gamma(m + 1/2)}{\Gamma(m + i + 3/2)} (2\rho)^{i-n+1/2} k_{i-n+1/2}(\kappa) \tag{115}
\]

The functions \( k_{i+1/2}(\kappa) \) are expressible as polynomials in \( \kappa \) multiplying the exponential \( e^{-\kappa} \), Eq. (58). They are subject to the recurrence relation

\[
k_{i+1/2}(\kappa) = (2i - 1) k_{i-1/2}(\kappa) + x^2 k_{i-3/2}(\kappa) \tag{116}
\]

furthermore, because \( K_i(x) = K_{-i}(x) \), we have

\[
k_{i-1/2}(x) = x^{2i-1} k_{i+1/2}(x) \tag{117}
\]

\[
k_{i-3/2}(x) = (2i - 1)x^{2i-1} k_{i+1/2}(x) + x^{2i-3} k_{i-3/2}(x) \tag{118}
\]

If we define

\[
R_i(\sigma, \tau) = \frac{1}{\Gamma(i + 1/2)(2\sigma + i + 1/2)} k_{i+1/2}(\kappa) \tag{119}
\]

\[
S_i(\sigma, \tau) = \frac{1}{\Gamma(i - 1/2)(2\sigma - i + 1/2)} k_{i+1/2}(\kappa) \tag{120}
\]

\[
T_{nm}(\sigma, \tau) = \sum_{i=m}^{\infty} S_i(\sigma, \tau) + \sum_{i=0}^{m-1} R_i(\sigma, \tau) \tag{121}
\]

where \( \rho \) and \( \kappa \) are given by Eqs. (112-113), then we see that

\[
J_{nm}(1, \sigma, \tau) = \Gamma(m + 1/2) T_{n+m,m}(\sigma, \tau) \tag{122}
\]

To compute the series \( T_{nm}(\sigma, \tau) \) we start with \( R_{1,0} \), by Eqs. (119) and (56); \( S_{1,0} \) follows from Eqs. (119), (120) as

\[
S_{1,0}(\sigma, \tau) = \frac{(2\rho + 1)}{\kappa} R_{1,0}(\sigma, \tau) \tag{123}
\]

Furthermore, using Eqs. (116), (118), we can verify that

\[
R_{i,1}(\sigma, \tau) = \frac{1 + \kappa}{(2i + 3)\rho} R_{i,0}(\sigma, \tau) \tag{124}
\]

\[
S_{i,1}(\sigma, \tau) = (l - 1/2) \frac{1 + \kappa}{2\tau} S_{i,0}(\sigma, \tau) \tag{125}
\]

\[
R_{i+1,0}(\sigma, \tau) = \frac{1}{(l + i + 3/2)\rho} \left[ (i + 1/2) R_{i,0}(\sigma, \tau) + \frac{\tau}{l + i + 1/2} R_{i-1,0}(\sigma, \tau) \right] \tag{126}
\]

\[
S_{i+1,0}(\sigma, \tau) = \frac{l - i - 1/2}{\tau} \left[ (i + 1/2) S_{i,0}(\sigma, \tau) + (l - i + 1/2) \rho S_{i-1,0}(\sigma, \tau) \right] \tag{127}
\]

For future reference we shall also note the recurrence relations between terms and series of different \( l, m \) values,

\[
R_{i-1,0}(\sigma, \tau) = (l + i + 1/2) R_{i,0}(\sigma, \tau) \tag{128}
\]

\[
S_{i-1,0}(\sigma, \tau) = (l - i - 1/2) S_{i,0}(\sigma, \tau) \tag{129}
\]

\[
T_{1,m-1}(\sigma, \tau) = T_{1,m}(\sigma, \tau) + S_{1,m}(\sigma, \tau) \tag{130}
\]

\[
T_{l+1,m}(\sigma, \tau) = \frac{i}{\tau} (iT_{lm}(\sigma, \tau) + \sigma T_{l,m}(\sigma, \tau) - S_{1,m-1}(\sigma, \tau)) \tag{131}
\]

(There is no round-off problem in Eq. (131), because the negative term is considerably smaller in magnitude than the first positive term.)

To study the convergence of the series for \( T_{1,m}(\sigma, \tau) \), Eq. (121), we define the convergence ratio

\[
d_i = R_i(\sigma, \tau)/R_{i-1,0}(\sigma, \tau) \tag{132}
\]

In particular, from Eqs. (123-124),

\[
d_0 = R_0(\sigma, \tau)/S_0(\sigma, \tau) = \frac{\kappa}{(2l + 1) \rho} \tag{133}
\]

\[
d_1 = \frac{1 + \kappa}{(2l + 3) \rho} \tag{134}
\]

If \( d_0 < 1 \) then \( d_i \) decreases as \( i \) increases (note that \( \rho > 1 \)), and reasonable convergence is obtained. For larger values of \( d_0 \) we make use of the asymptotic expansion of \( F_m(x) \), Eq. (32), and obtain the asymptotic series,

\[
T_{lm}(\sigma, \tau) \approx (2\tau)^{-1} k_l(2\sqrt{\tau\sigma}) - \sum_{i=l-m}^{N} S_i(\sigma, \tau) \tag{135}
\]
The criteria for convergence and the techniques used to improve accuracy are discussed in Section IV, C, 3.

The expression for the exchange integral, Eq. (107), can now be written in the form (note that $I'/(1/2) = \sqrt{\pi}$),

$$I = \frac{\pi}{8} a_1 a_2 a_4 J_0 \int_0^1 \frac{du}{u(1-u)^{3/2}} \int_0^{1/2} \frac{dv}{v(1-v)^{3/2}} \int_0^{\pi/2} d\theta \left[ \int_0^1 dw (1-w)^{3/2} T_{0,0}(\sigma, \tau) \right].$$

(136)

This is computed by three-dimensional numerical integration. The number and distribution of the quadrature points are also discussed in Section IV, C, 3.

Equation (136) is used for two-, three-, and four-center exchange integrals without any changes. In the case of hybrid integrals, which have a single-center distribution, say

$$\exp \left( -\alpha r_{1A} - \alpha r_{1B} \right) = \exp \left[ -(\alpha_1 + \alpha_2) r_{1A} \right],$$

the transform of Eq. (59) is applied to the distribution as a whole rather than to each of its two orbitals, thus reducing the dimensionality of the final integration by one. The resulting expression is

$$I = \frac{\pi^{3/2}}{4} \left( \alpha_1 + \alpha_2 \right) \alpha_2 a_4 \int_0^{1/2} \frac{dv}{v(1-v)^{3/2}} \int_0^{1/2} \frac{dw}{w(1-w)^{3/2}} \int_0^{\pi/2} d\theta \left[ \int_0^1 dw (1-w)^{3/2} T_{0,0}(\sigma, \tau) \right],$$

(137)

where the point $P$ coincides with $A$, and

$$\sigma = \frac{(1-v)}{v} \frac{c^2}{\rho^2},$$

$$\tau = \frac{\rho^2}{4} \left[ (1-w)(\alpha_1 + \alpha_2)^2 + \frac{v}{w} \alpha_3^2 + \frac{w}{1-v} \alpha_4^2 \right].$$

(138)

The half-integral value of the first index of $T_{0,0}$ above is troublesome, for it leads to Bessel functions of integer order in Eqs. (119-120), and these are much more difficult to compute than the half-integral order functions. However, this is more than offset by the elimination of one numerical integration in comparison with the exchange integral formula.

For the three-center nuclear attraction integral,

$$I = \int [1/r_{1A}] \exp \left( -\alpha r_{1A} - \alpha r_{1B} \right) dr_1,$$

the application of the integral transform to each of the two orbitals and the use of Eq. (87) lead to

$$I = \sqrt{\pi \over 2} a_2 a_4 \int_0^{1/2} dw (1-w)^{-3/2} \rho^4 T_{0,0}(\sigma, \tau),$$

(140)

in which the point $Q$ coincides with $C$, and

$$\sigma = u(1-u) a^2/p^2,$$

$$\tau = \frac{p^2}{4} \left( \frac{1}{u} \alpha_1^2 + \frac{1}{1-u} \alpha_4^2 \right).$$

(141, 142)

Two-center nuclear attraction integrals, Coulomb integrals, and all single-center integrals are computed fairly easily by conventional methods (Roothaan 1951), and the application of the Gaussian transform method to them is not suggested.

2. Higher Orbitals

As in the case of GTO integrals (Section IV, B.), formulas for integrals involving higher orbitals (i.e., $2s$, $2p$, etc.) are obtained by the application of differential operators to the basic 1s formula. This procedure is based on the relations,

$$r \exp (-\alpha r) = -\frac{\partial}{\partial \alpha} \exp (-\alpha r),$$

(143)

$$x_A \exp (-\alpha x_A) = -\frac{\partial}{\partial \alpha} \left( \frac{1}{x_A} \frac{\partial}{\partial x_A} \right) \exp (-\alpha x_A),$$

(144)

$$r^2 \exp (-\alpha r) = -\frac{\partial}{\partial \alpha} \left( r \exp (-\alpha r) \right),$$

(145)

$$x_A r^2 \exp (-\alpha r) = -\frac{\partial}{\partial \alpha} \left( x_A r \exp (-\alpha r) \right),$$

(146)

etc., which are easily verified if we note that

$$r_A = [(x-A_0)^2 + (y-A_0)^2 + (z-A_0)^2]^{1/2}.$$

To determine the effect of the differential operators on $J_{nm}(q, f, g)$, Eq. (108), we note that

$$\frac{\partial}{\partial x} J_{nm}(q, f, g) = -J_{n-1,m+1}(q, f, g),$$

(147)

$$\frac{\partial}{\partial q} J_{nm}(q, f, g) = -J_{n-1,m}(q, f, g),$$

(148)

$$\frac{\partial}{\partial q} J_{nm}(q, f, g) = -J_{n+1,m}(q, f, g).$$

(149)

Using the definitions of $q, f, g$ for the exchange integral, Eqs. (104-106), we obtain

$$\frac{\partial}{\partial x_A} J_{nm}(q, f, g) = -\frac{\alpha_1}{2\alpha_0} J_{n+1,m}(q, f, g),$$

(150)

$$\frac{\partial}{\partial \alpha_0} J_{nm}(q, f, g) = -2\alpha_0 \left( 1 - u \right) a_A J_{n-1,m+1}(q, f, g) + (1-u) p_A J_{n-1,m+1}(q, f, g).$$

(151)
As an example, we shall consider the integral

\[
\left[(2p_\nu)(1^2); (1^1)(1^1)\right]
= \int \frac{1}{r_{12}^2} x_{1A} \exp \left(-\alpha_1 r_{1A} - \alpha_2 r_{1B} - \alpha_3 r_{2C} - \alpha_4 r_{2D}\right) dr_1 dr_2
\]

\[
= -\frac{\partial}{\partial A_1} \left\{ \frac{1}{8} \alpha_1 \alpha_2 \alpha_3 \alpha_4 \int_0^1 \frac{du}{u[1 - u]^{3/4}} \right\}
\]

\[
\times \int_0^1 \frac{dv}{v[1 - v]^{3/4}} \int_0^1 \frac{dw}{w[1 - w]^{3/4}} \int_0^1 \frac{dx}{x[1 - x]^{3/4}} J_{l,s}(x, f, g)
\]

\[
= -\frac{\sqrt{\pi}}{8} \alpha_1 \alpha_2 \alpha_3 \int_0^1 \frac{du}{u^{3/2}(1 - u)^{1/2}}
\]

\[
\times \int_0^1 \frac{dv}{v^{3/2}(1 - v)^{1/2}} \int_0^1 \frac{dw}{w[1 - w]^{3/4}} J_{l,s}(x, f, g)
\]

\[
+ \int_0^1 \frac{du}{u^{3/2}(1 - u)^{1/2}} \int_0^1 \frac{dv}{v^{3/2}(1 - v)^{1/2}} \int_0^1 \frac{dx}{x^{3/2}(1 - x)^{1/2}} T_{l,s}(x, f, g)
\]

\[
= -\frac{\sqrt{\pi}}{8} \alpha_1 \alpha_2 \alpha_3 \alpha_4 \int_0^1 \frac{du}{u^{3/2}(1 - u)^{1/2}}
\]

\[
\times \int_0^1 \frac{dv}{v^{3/2}(1 - v)^{1/2}} \int_0^1 \frac{dw}{w[1 - w]^{3/4}} T_{l,s}(x, f, g)
\]

\[
+ \frac{1}{2} \int_0^1 \frac{du}{u^{3/2}(1 - u)^{1/2}} \int_0^1 \frac{dv}{v^{3/2}(1 - v)^{1/2}} \int_0^1 \frac{dx}{x^{3/2}(1 - x)^{1/2}} T_{l,s}(x, f, g)
\]

\[
. (152)
\]

We note that the corresponding integrals for 2p_y and 2p_x with the same value of a_1, lead to formulas which differ from Eq. (152) only in the factors a_p and p_x. All three integrals can thus be computed together in practically the same time that it takes to calculate one of them. Similarly, the 3^4 = 81 different integrals that are obtained by assigning x, y, and π subscripts in turn to each of the four orbitals in [(2p), (2p); (2p), (2p)] can be computed simultaneously.

The derivation of the formulas for higher-orbital integrals is a very tedious and error-prone job, when done manually, but is easily adaptable to automatic computation. The complete procedure, for orbitals through 3d, has been programmed for the IBM 7090 computer. The program starts out with the basic 1s formula, then applies the appropriate differential operators as many times as necessary, and finally collects and sorts all the resulting terms. Each term is represented in the computer by the values of the indices l, m and of the exponents of the various factors, such as u, 1 - u, a_p, p_x, a_v, etc. The differential operators are coded as a linear combinations of "substitution operators," each of which modifies the values of one or more of the indices and exponents. The final formula is printed out as a list of terms, each term being represented by a numerical coefficient and the values of all indices and exponents.

3. Computational Accuracy and Efficiency

The speed and accuracy with which the series T_{lm}(\sigma, \tau) is computed are of crucial importance for the calculation of the integrals. A number of special devices are therefore used in the series evaluation, and some of these will now be described.

First, the numerical coefficients

\[(i + 1/2)/[(l + i + 3/2), 1/[(l + i + 1/2)(l + i + 3/2)], etc.,
\]

used in the recurrence relations, Eqs. (123-129), are precomputed and stored in the program. When T_{lm}(\sigma, \tau) is needed for a sequence of orders l, then only the two lowest orders are computed by series summation, the others being obtained by the recurrence relations, Eq. (131). Even the two low-order series are not calculated independently: as each term of one of these series is computed by recurrence, the corresponding term of the other series is obtained by one multiplication, Eqs. (128-129). A special feature of the 7090 computer is very helpful here: multiplication by a simple factor (e.g., integer or half-integer) can be performed more rapidly than a regular (random) multiplication. Series with different m values, but of the same order l, are easily derived from each other by Eq. (130).

The number of terms R_{lm}(\sigma, \tau) required for any given accuracy in T_{lm}(\sigma, \tau) is reduced by the use of the δ^2 process (Aitken, 1926, 1937). This is in effect an extrapolation of the series, after n terms have been computed, on the assumption that the remaining terms continue in geometric progression, and leads to

\[
\sum_{i=0}^{\infty} a_i \approx \sum_{i=0}^{n-2} a_i + \frac{a_{n-1}^2}{a_{n-1} - a_n}. (153)
\]
The series are not calculated to a constant, predetermined number of terms. Rather, each term is compared with a computed tolerance parameter, and if it is small enough the series is truncated and the δ² process is applied. The tolerance parameter is determined by dividing a specified overall accuracy parameter for the integral by the number of quadrature points used and by the complete factor which multiplies the series in the integral formula. In actual fact this procedure is a little more complicated in the cases in which recurrence relations between series are used. It has been found empirically that the total error in an integral due to series truncation (apart from quadrature errors) is about 100 times smaller than the specified accuracy parameter, by virtue of the δ² process.

When the asymptotic expansion, Eq. (135), is used, terms are computed until the latest is smaller, in absolute value, than a tolerance parameter (which is somewhat smaller than the one used for the direct series), or until the smallest term has been reached. In the latter case, only half of the smallest term is used; this has been found to reduce the error by one order of magnitude, and provides satisfactory accuracy.

Round-off errors in floating-point arithmetic can mount catastrophically when a sequence of small increments are added successively to a fairly large sum. Therefore, we start all series summations with the smallest (i.e., latest computed) term, adding up the previously computed terms in reverse order, through \( R_{1,0} \), \( S_{1,0} \), \( \ldots \) to \( S_{1,1-m+1} \). Thus, as the sum builds up, the terms being added to it increase as well and are of comparable magnitude with it.

The same round-off problem occurs in the quadrature summations, and is successfully dealt with by building up intermediate sums (e.g., the sum of all \( \omega \) contributions for any fixed \( u, v \)), which are later added-up into successively higher sums (e.g., all \( v, \omega \) contributions for a fixed \( u \), etc.). Furthermore, each summation is started with the generally smaller contributions of the points around the quadrature boundaries.

The principal factors in determining the accuracy and speed of the integral calculation are the number and distribution of quadrature points. It has been found that Gauss-Legendre points and weights (Salzer and Zucker, 1949; Gawlik, 1958), without subdivision of the intervals, give much better results than any low-order quadrature formulas applied to subintervals. When the four orbital exponents \( \alpha_i \) are of comparable magnitude (within a factor of two, approximately), then 12 points in each of \( u \) and \( v \), and 16 points in \( \omega \), result in absolute accuracy of about \( 10^{-5} \) or better in the integrals. When the charge distributions are rather asymmetric, considerably more points may be required. Work is still in progress on the selection of quadrature points for such cases. It appears—see Eq. (98) and Fig. 5—that a marked inequality between \( \alpha_4 \) and \( \alpha_5 \) can be handled successfully by an increase in the number of \( u \) points alone, while the number of \( v \) points would depend on the relative magnitude of \( \alpha_4 \) and \( \alpha_5 \). An increase in the number of \( w \) points seems particularly effective when the sums \( \alpha_4 + \alpha_5 \) and \( \alpha_3 + \alpha_4 \) differ greatly. Some simple transformations of the quadrature points, distributing them so that they are more closely spaced at one end of the range than at the other (depending on the values of the \( \alpha_i \)'s),

<table>
<thead>
<tr>
<th>Atom</th>
<th>( x )</th>
<th>( y )</th>
<th>( z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>( H_1 )</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>( H_2 )</td>
<td>1.885 618 08</td>
<td>0.0</td>
<td>0.666 666 67</td>
</tr>
<tr>
<td>( H_3 )</td>
<td>-0.942 809 04</td>
<td>1.632 993 16</td>
<td>0.666 666 67</td>
</tr>
<tr>
<td>( H_4 )</td>
<td>-0.942 809 04</td>
<td>-1.632 993 16</td>
<td>0.666 666 67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Orbits</th>
</tr>
</thead>
<tbody>
<tr>
<td>((1s)_C = \sqrt{1.0/\pi} \exp (-1.0r_C))</td>
</tr>
<tr>
<td>((1s)_C = \sqrt{1.0/\pi} \exp (-1.0r_C))</td>
</tr>
<tr>
<td>((2p)_C = \sqrt{1.625/\pi} x_C \exp (-1.625r_C))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Configuration</th>
<th>12³ x 16 = 2304pts.</th>
<th>16³ x 20 = 5120pts.</th>
<th>20³ x 24 = 9600pts.</th>
</tr>
</thead>
<tbody>
<tr>
<td>((1s)(1s)_C)</td>
<td>0.030 681 96</td>
<td>0.030 681 70</td>
<td>0.030 681 87</td>
</tr>
<tr>
<td>((1s)(1s)_C)</td>
<td>0.035 693 98</td>
<td>0.035 693 54</td>
<td>0.035 693 68</td>
</tr>
<tr>
<td>((1s)(1s)_C)</td>
<td>0.012 806 27</td>
<td>0.012 709 14</td>
<td>0.012 712 93</td>
</tr>
<tr>
<td>((2p)(2p)_C)</td>
<td>-0.007 525 73</td>
<td>-0.007 525 63</td>
<td>-0.007 525 59</td>
</tr>
<tr>
<td>((2p)(2p)_C)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>((2p)(2p)_C)</td>
<td>-0.047 661 45</td>
<td>-0.047 662 74</td>
<td>-0.047 662 77</td>
</tr>
<tr>
<td>((2p)(2p)_C)</td>
<td>0.035 566 92</td>
<td>0.035 568 43</td>
<td>0.035 568 80</td>
</tr>
</tbody>
</table>

The Gaussian Function