An application of the Monte Carlo method to the evaluation of some molecular integrals

By D. R. Cowdrey and C. M. Reeves

The Monte Carlo method is reviewed and applied to the evaluation of a class of molecular integrals for which no closed method is available. Three algorithms are developed in attempts to reduce the sample variances. Extensions on more powerful computing machinery could provide viable means of evaluation of the large numbers of integrals that arise in quantum-mechanical studies of molecular structure.

1. Introduction

The Monte Carlo method has been described as a method of last resort in that it is often necessary to employ very large sample sizes in order to obtain the desired accuracy. Calculations can thus be time consuming and the results correspondingly expensive. Where applicable, the standard methods of numerical analysis are usually preferred.

The values of very large numbers of integrals are needed in any attempt to make quantum-mechanical predictions of molecular structure. The absence of reliable methods of evaluation has been an obstacle for thirty years since the foundations of the quantum theory. No closed formulae have been found for large classes of these integrals, and although the introduction of automatic computers has greatly facilitated the development of methods of approximation, there is still a need for further work in this field. It is therefore appropriate to investigate the use of the Monte Carlo method in the evaluation of these so-called molecular integrals.

In Section 2 we define the integrals of interest. Section 3 reviews the principles of the Monte Carlo method in the context of evaluation of definite integrals. The following three sections describe algorithms for the evaluation of the simplest class of three-centre potential integrals. These are arranged in order of increasing complexity and arise from attempts to reduce the variance of the estimated integral values. The methods are compared in Section 7 in the light of the results obtained for a variety of typical integrals. The extensions necessary to cater for other classes of integral are indicated in Section 8, and a parallel organization of the computation is proposed whereby the calculation of a group of related integrals could be performed with little more working than would be required for one. Development of these ideas on more powerful computing equipment than the Ferranti Pegasus which was available to us could thus compensate for the inherent slowness of the Monte Carlo method in the evaluation of the full set of integrals required for a molecular structure investigation.

2. Molecular integrals

In this section we define the various classes of integral which arise in the course of calculating an electronic wavefunction of a molecule or radical. All properties of a physical system can in principle be predicted from a knowledge of its wavefunction. For some properties, such as the electric dipole moment, it is necessary to evaluate additional classes of integral. A concise description of the different aspects of a molecular structure calculation has been given by Boys and Cook (1960). Our purpose in this section is merely to indicate the extent and nature of the integral problem.

The region of integration is in each case the whole of physical space. The variables of integration are the positional co-ordinates of volume elements dV in this space. The integrands are compounded of two types of quantity; orbitals and operators. An orbital is a differentiable function of position, say $\eta(\mathbf{r})$, where $\mathbf{r}$ is the position vector of the volume element dV. The orbitals $\eta_i(\mathbf{r})$, $1 \leqslant i \leqslant N$, for a calculation in a basis of N orbitals are drawn from a complete system of functions, thus ensuring that as N increases, predictions of properties converge to their correct values. Operators are of several types, those necessary in the calculation of a wavefunction being as follows.

(*a*) Single-particle operators:

(i) the identity	1
(ii) kinetic energy	$-\frac{1}{2}\nabla^2$
(iii) Coulomb potential	$\frac{1}{r_c}$

(*b*) Two-particle operator:

(iv) electrostatic interaction	$\frac{1}{r_{12}}$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplacian operator,

$r_c = |\mathbf{r} - \mathbf{C}|$ denotes the distance of the volume element dV at position $\mathbf{r}$ from an atomic nucleus at position $\mathbf{C}$, and $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ denotes the distance between the volume elements dV_1 and dV_2 at positions $\mathbf{r}_1$ and $\mathbf{r}_2$.

The relevant molecular integrals may now be defined as follows.

$$(\eta_i|G|\eta_j) = \int \eta_i^*(\mathbf{r})\eta_j(\mathbf{r})dV$$

$$(\eta_i|T|\eta_j) = \int \eta_i^*(\mathbf{r})(-\tfrac{1}{2}\nabla^2)\eta_j(\mathbf{r})dV$$

$$(\eta_i|V_c|\eta_j) = \int \eta_i^*(\mathbf{r})\frac{1}{r_c}\eta_j(\mathbf{r})dV$$

and

$$(\eta_i\eta_j|M|\eta_k\eta_l) = \iint \eta_i^*(\mathbf{r}_1)\eta_j^*(\mathbf{r}_2)\frac{1}{r_{12}}\eta_k(\mathbf{r}_1)\eta_l(\mathbf{r}_2)dV_1dV_2,$$

where η^* denotes the complex conjugate of η.

Physical significance can be attached to the various individual integrals only in certain special cases. The orbital functions η_l then have the interpretation that $\eta_i^*(\mathbf{r})\eta_i(\mathbf{r})dV$ is proportional to the probability that an electron is within the volume element dV at $\mathbf{r}$. We speak of an electron being in an orbital just as, classically, we might say that it is in a particular orbit in the attracting field of the nuclei. An integral such as $(\eta_i|T|\eta_i)$ then corresponds to the mean kinetic energy of an electron in the orbital η_i. It must be emphasized that this is an over-simplified picture. There is, for example, no classical counterpart of the interactions such as $(\eta_i|T|\eta_j)$, when $i \neq j$. In the sequel, this probabilistic description should not be confused with the probabilistic element in the use of the Monte Carlo method for evaluation of the integrals.

Clearly, in the evaluation of the integrals, everything depends upon the form of the orbitals. This should ideally be both physically realistic to ensure rapid convergence, and tractable to mathematical analysis. Unhappily, these requirements seem to be largely incompatible. The compromise with which we shall be concerned is the set of so-called Slater orbitals (Slater, 1933). These are of the general form

$$\eta(\mathbf{r}) = S_{lm}(\theta, \phi)r_A^n e^{-\alpha r_A},$$

where (r_A, θ, ϕ) are spherical polar co-ordinates relative to an atomic nucleus at position A as origin, $S_{lm}(\theta, \phi)$ is a spherical harmonic, n a non-negative integer, and α a positive so-called screening parameter. This form is based upon the analytical solutions for the wavefunctions of the Hydrogen atom, and is quite realistic physically for general systems. Some of the the integrals can be evaluated from simple formulae, but for others no closed formulae have been reported. In particular, the evaluation of the three- and four-centre integrals has only been possible by approximate methods.

The present investigation has been confined to the simplest class of three-centre integrals, namely

$$\int e^{-\alpha r_A}\frac{1}{r_c}e^{-\beta r_B}dV,$$

where A, B and C are three general positions. We shall return in the final section to comment upon the possible extention of the Monte Carlo method to the remaining integrals.

3. The Monte Carlo method

The concept of a probability density is basic to the Monte Carlo method. Thus let x be a random variable such that $P(a, b)$ is the probability that a random value of x lies in the range $a \leqslant x \leqslant b$. Then, on the assumption that x is continuous, the probability density $\rho(x)$ of its distribution may be defined as

$$\rho(x) = \lim_{\delta x \to 0} \frac{1}{\delta x}P(x, x + \delta x).$$

Thus, for small δx, $\rho(x)\delta x$ is the probability that a random value of x lies in the range δx at x.

The application of the Monte Carlo method to the evaluation of definite integrals rests upon two theorems in statistics (Kahn, 1960). The first establishes the method to be used for estimating the value of the integral.

Theorem 1. *The Strong Law of Large Numbers*

Let x_i, $1 \leqslant i \leqslant n$, be independent random numbers drawn from a distribution with probability density $\rho(x)$. Suppose that

$$\bar{z} = \int_{-\infty}^{\infty} z(x)\rho(x)dx$$

exists and define

$$\hat{z}_n = \frac{1}{n}\sum_{i=1}^{n} z(x_i).$$

Then as $n \to \infty$, $\hat{z}_n \to \bar{z}$ with probability 1.

In order, therefore, to evaluate an integral such as

$$\int_a^b f(x)dx,$$

we first express the integrand as the product of a function $z(x)$ and a density $\rho(x)$, satisfying

$$\rho(x) = 0 \qquad \text{for } x < a \text{ or } x > b$$
$$\rho(x) \geqslant 0 \qquad \text{for } a \leqslant x \leqslant b$$

and $$\int_a^b \rho(x)\,dx = 1.$$

We then generate random numbers x_i from the distribution with probability density $\rho(x)$ and accept

$$\frac{1}{n}\sum_{i=1}^{n} z(x_i)$$

as an estimate of the integral for some suitable n.

The second theorem establishes the likely error in this procedure. We first define, in the notation of Theorem 1, the variance of $z(x)$, denoted by σ^2. Thus

$$\sigma^2 = \int_{-\infty}^{\infty}[z(x) - \bar{z}]^2\rho(x)dx$$

whence it is easily shown that

$$\sigma^2 = \overline{[z(x) - \bar{z}]^2} = \overline{(z^2)} - (\bar{z})^2.$$

The positive square root σ of the variance is termed the standard deviation of $z(x)$. $\bar{}$ is termed the mean or the expected value of $z(x)$.

Theorem 2. *The Central Limit Theorem*

Let x_i, $1 \leqslant i \leqslant n$, be random values drawn respectively from distributions with means m_i and variances σ_i^2. Then the distribution of the sum

$$x = \sum_{i=1}^{n} x_i$$

is asymptotically normal with mean

$$m = \sum_{i=1}^{n} m_i$$

and variance $$\sigma^2 = \sum_{i=1}^{n} \sigma_i^2.$$

Corollary

In the special case that all the x_i come from a common distribution with mean m and standard deviation σ, their average value

$$x = \frac{1}{n} \sum_{i=1}^{n} x_i$$

is asymptotically distributed normally with the same mean m and standard deviation $\sigma/\sqrt{n}$.

The normal distribution is well known and has probability density

$$\rho(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad -\infty \leqslant x \leqslant \infty$$

where m and σ are the mean and standard deviation of x.

The required error estimate follows. For sufficiently large n we have

$$\text{Prob}\,[|\hat{z}_n - \bar{z}| > \epsilon] = 2\int_{\bar{z}+\epsilon}^{\infty} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(z-\bar{z})^2}{2\sigma^2/n}\right) dz.$$

Writing $\epsilon = \lambda(\sigma/\sqrt{n})$, we have more conveniently

$$\text{Prob}\,[|\hat{z}_n - \bar{z}| > \lambda(\sigma/\sqrt{n})] = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \int_{\lambda}^{\infty} e^{-\frac{1}{2}t^2} dt.$$

In particular, for $\lambda = 3$,

$$\text{Prob}\,[|\hat{z}_n - \bar{z}| > 3\sigma/\sqrt{n}] = 0{\cdot}0027,$$

so that for all practical purposes we may write

$$|\hat{z}_n - \bar{z}| \leqslant 3\sigma/\sqrt{n}$$

as our estimate of the error in accepting $\hat{z}_n$ as an estimate of $\bar{z}$.

This dependence of the error bound upon the sample size n is both the strength and weakness of the Monte Carlo method. On the one hand convergence of the result to the correct value is assured as n increases, and yet on the other hand the accuracy only increases as the square root of the sample size. In practice it is therefore important that the resolution of the integrand $f(x)$ into the product $z(x)\rho(x)$ should lead to as small a value as possible for the standard deviation σ of the factor $z(x)$. Since, however, it is strictly speaking the cost of the calculation that is to be minimized rather than the sample size, it may on occasions be economic to accept a resolution which gives a rather larger standard deviation if thereby the task of generating a random value from the density $\rho(x)$ and subsequent evaluation of $z(x)$ is eased significantly. We are thus led to consider briefly, in the remainder of this section, techniques for the generation of random samples from different probability distributions, and for the reduction of sample variances.

3.1 *The generation of random numbers*

This topic has been well covered by a number of authors. Economic considerations dictate the use on an electronic digital computer of pseudo-random numbers generated internally as required by means of determinate recurrence relations. It is worth-while repeating Taussky and Todd's quotation (Meyer, 1956) of D. H. Lehmer's definition of a pseudo-random sequence as "a vague notion embodying the idea of a sequence in which each term is unpredictable to the uninitiated and whose digits pass a certain number of tests, traditional with statisticians, and depending somewhat upon the uses to which the sequence is to be put".

Commonly, pseudo-random numbers are generated from the so-called standard uniform distribution having density $\rho(x) = 1$, $0 \leqslant x \leqslant 1$. The multiplicative congruential method seems the most satisfactory, having a precisely determined cycle structure with a long period. (Edmonds, 1960; Barnett, 1962). In the remainder of this paper we shall denote by u a pseudo-random number drawn from the standard uniform distribution.

By suitably manipulating samples from the standard uniform distribution it is possible to generate random numbers drawn from any other desired distribution. Reviews of methods are given by Butler (Meyer, 1956) and by Votaw and Rafferty (1951). Particularly powerful are the various rejection techniques in which the application of logical criteria to a set $\{u_i\}$ largely replaces arithmetical manipulation in forming the required random variables. By this means a high operating speed is attained, since in computers logical operations are generally very fast in comparison with arithmetic operations. As an example we give the following ALGOL procedure declaration.

real procedure *sample root*;
begin comment *At each activation "sample root" is assigned a random value drawn from the distribution with density* $\frac{3}{2}\sqrt{x}$, $0 \leqslant x \leqslant 1$. *It is assumed that a function designator "sud" is available which is assigned similarly a random value drawn from the standard uniform distribution. Two values are generated by means of "sud". If the first is greater than the square of the second, the first is accepted as the value for "sample root," otherwise the pair are rejected and a fresh start made*;

```
      real u;
reject: u := sud;
      if u ⩽ sud² then go to reject;
accept: sample root := u
  end sample root;
```

We give below a further procedure declaration which is required in subsequent sections in connection with the algorithms for the molecular integrals.

```
real procedure sample exp;
begin comment This function designator assumes a
      random value drawn from the distribution with
      density e^-x, 0 ⩽ x ⩽ ∞. A random sequence u_i
      is formed such that n is the least value for which
      u_1 + u_2 + ... + u_n ⩾ u_0. The whole sequence is
      rejected and a fresh start made if n is even, other-
      wise t + u_0 is accepted as the required random
      number where t is the number of rejected sequences.
      A proof of the method is given in the papers by
      Butler, and by Votaw and Rafferty quoted above;

      real S, u; integer t; Boolean even;
      for t := 0, t + 1 while even do
          begin even := false;
                S := u := sud;
                u := u + t;
                for S := S - sud while S > 0 do
                  even := not even
          end;
      sample exp := u
end   sample exp;
```

3.2 *Variance reduction*

We have seen that the choice of sampling technique must be such that the variance is as low as possible in order to obtain an efficient procedure. Kahn (1956, 1960) gives a full discussion of this problem. Two conclusions that are particularly relevant in the procedures that we have employed are first that one should pursue analytical methods as far as is feasible. Thus, where it is possible to integrate out certain variables, this should be done. Any symmetry present in the problem should be exploited to reduce the region of integration. Secondly, it is shown that in the factorization $f(x) = z(x)\rho(x)$, the choice of density $\rho(x)$ should leave the factor $z(x)$ as uniform as possible. It has been shown that the variance is minimized by the choice

$$\rho(x) = \frac{|f|}{\int_a^b |f| dx}.$$

When $f(x)$ is non-negative this gives $\rho(x) = f(x) \Big/ \int_a^b f(x)dx$

and the estimate has in fact zero variance. This is not surprising since the determination of the optimum density presupposes a knowledge of the value of the integral.

In practice, therefore, one seeks to make the sampling density reflect the main characteristics of the integrand. In particular, any singularities of the integrand must be absorbed into the density rather than left in the quantity $z(x)$ which is to be averaged.

4. A single centre density function

This first algorithm is relatively crude. Our primary aim was to cater adequately for the singularity of the Coulomb potential in the integral

$$I = \int e^{-\alpha r_A} \frac{1}{r_c} e^{-\beta r_B} dV.$$

Let (r_c, θ, ϕ) denote spherical polar co-ordinates centred on C. The volume element is then

$$dV = r_c^2 \sin\theta dr_c d\theta d\phi.$$

The original integrand is symmetric about the plane of A, B, C and so by choosing the axis $\theta = 0$ normal to this plane the integral may be written

$$I = 2\int_0^\infty dr_c \int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi e^{-\alpha r_A - \beta r_B} r_c \sin\theta.$$

We chose a sampling scheme in which values of r_c, θ and ϕ were drawn independently from the following distributions.

For r_c, density $\rho(r_c) = \lambda e^{-\lambda r_c}$, $0 \leqslant r_c \leqslant \infty$.

For θ, density $\rho(\theta) = \dfrac{2}{\pi}$, $0 \leqslant \theta \leqslant \dfrac{\pi}{2}$.

For ϕ, density $\rho(\phi) = \dfrac{1}{2\pi}$, $0 \leqslant \phi \leqslant 2\pi$.

Here λ is a free parameter, $\lambda > 0$, whose value was to be determined by trial and error.

Writing

$$z = \frac{2\pi^2}{\lambda} r_c \sin\theta \exp(\lambda r_c - \alpha r_A - \beta r_B),$$

we thus obtain

$$\hat{I}_n = \frac{1}{n} \Sigma z$$

as an estimate of the integral. For the error analysis, we have no accurate theoretical value for the variance, but accept as an order-of-magnitude estimate, the value

$$\hat{\sigma}_n^2 = \frac{1}{n}(\Sigma z^2) - (\hat{I}_n)^2.$$

Thus the sampling procedure is used to accumulate simultaneously Monte Carlo estimates of both the integral and the variance.

The program is conveniently described in ALGOL, making use of a self-evident notation for input and output procedures, and incorporating the procedures "*sud*" and "*sample exp*" previously introduced.

```
begin comment This block estimates the integral I and
      standard deviation σ from a sample of a size n
      which is pre-set. From this the sample size necessary
      for a required accuracy can be determined. For
      our exploratory calculations this is a more appro-
      priate organization than permitting the calculation
      to continue until it estimates that it has achieved a
      desired accuracy;
    real λ, Ax, Ay, Bx, By, α, β, S, SS, rA, rB, rC, θ, φ,
                                        X, Y, Z, term, I, σ;
    integer n, i;
    set (λ, n);
    read (Ax, Ay, Bx, By, α, β); comment The
      co-ordinates are given relative to a pair of
      rectangular axes in the plane ABC with C as
      origin;
    S := SS := 0;
    for i := 1 step 1 until n do
      begin rC := sample exp/λ;
            θ := (π/2)* sud; comment We use * to
                 denote multiplication and assume
                 π to be a non-local variable with
                 value 3·14159 . . . ;
            φ := (2*π)*sud;
            X := rC* sin (θ)* cos (φ);
            Y := rC* sin (θ)* sin (φ);
            Z := rC* cos (θ);
           rA := sqrt ( (X − Ax)² + (Y − Ay)² +
                                        Z²);
           rB := sqrt ( (X − Bx)² + (Y − By)² +
                                        Z²);
         term := rC* sin (θ) * exp
                              (λ*rC − α*rA − β*rB);
            S := S + term;
           SS := SS + term²
      end;
    I := (2*π²/λ)*S/n;
    σ := (2*π²/λ)*sqrt (SS/n − (S/n)²);
    print (I, σ)
end;
```

We defer all discussion of the results obtained from the three algorithms until each has been described.

5. A two-centre density function

In this second algorithm we tried both to represent the Coulomb singularity in the sampling distribution and to give a more realistic coverage of the behaviour of the orbitals when C was well removed from A and B. We were attracted to two-centre and elliptic co-ordinates because the relevant analysis leading to the form of the volume element is well known.

We denote by E the point with position vector

$$E = \frac{\alpha A + \beta B}{\alpha + \beta}.$$

Thus E divides AB in the ratio $\alpha : \beta$. We assume that E and C do not coincide and take E and C as the foci of an elliptic co-ordinate system. Thus a general point P has co-ordinates (ξ, η, ϕ) where $\xi + r_E + r_C$, $\eta = r_E - r_C$ and ϕ is the angle between the plane ECP and some fixed plane through EC, say ABC. The volume element is then

$$dV = \frac{\xi^2 - \eta^2}{8EC} d\xi d\eta d\phi,$$

where $EC \leqslant \xi \leqslant \infty$, $-EC \leqslant \eta \leqslant EC$ and $0 \leqslant \phi \leqslant 2\pi$, and the integral is written

$$I = \frac{1}{4EC} \int_{EC}^{\infty} d\xi \int_{-EC}^{EC} d\eta \int_0^{2\pi} d\phi (\xi + \eta) e^{-\alpha r_A - \beta r_B}.$$

At large distances $e^{-\alpha r_A - \beta r_B}$ and $e^{-(\alpha+\beta) r_E}$ are roughly comparable, and are indeed equal on the exterior segments of the line AB produced. It is therefore indicated that we should incorporate $e^{-\mu r_E}$ into the sampling density where $\mu = \alpha + \beta$ is a likely choice.

For our sampling scheme then, we drew values of ξ, η, ϕ independently from distributions with densities as follows.

$$\rho(\xi) = \frac{1}{2}\mu e^{-\frac{1}{2}\mu(\xi - EC)}, \qquad EC \leqslant \xi \leqslant \infty$$

$$\rho(\eta) = \frac{1}{4}\mu \operatorname{cosech}\left(\frac{1}{2}\mu EC\right) e^{-\frac{1}{2}\mu\eta}, \quad -EC \leqslant \eta \leqslant EC$$

$$\rho(\phi) = \frac{1}{2\pi}, \qquad 0 \leqslant \phi \leqslant 2\pi.$$

As before, after n cycles,

$$\hat{I}_n = \frac{1}{n}\Sigma z$$

and

$$\hat{\sigma}_n^2 = \frac{1}{n}(\Sigma z^2) - (\hat{I}_n)^2,$$

where now

$$z = \frac{2\pi}{\mu^2 EC}(1 - e^{-\mu EC})(\xi + \eta) e^{\mu r_E - \alpha r_A - \beta r_B}.$$

In order to implement this scheme, it is necessary to convert from the elliptic co-ordinate system (ξ, η, ϕ) back into the rectangular Cartesians which were used initially to specify A, B and C. For this it is convenient to define an auxiliary set of right-handed orthogonal unit vectors e_1, e_2, e_3. Thus, let N be the foot of the perpendicular from P on to EC, so that

$$EN = (r_E^2 + EC^2 - r_C^2)/2EC$$

and

$$PN = (r_E^2 - EN^2)^{1/2}.$$

We now assume that A, B and C are linearly independent and define

$$e_3 = (C - E)/EC$$

$$e_2 = (V^2)^{-1/2} V \quad \text{where } V = e_3 \times (A - B)$$

and

$$e_1 = e_2 \times e_3.$$

It now follows readily that

$$\boldsymbol{P} = \boldsymbol{E} + PN\cos\phi\,\boldsymbol{e}_1 + PN\sin\phi\,\boldsymbol{e}_2 + EN\,\boldsymbol{e}_3.$$

The ALGOL description of the program follows.

```
begin comment This second algorithm has the same
      general specification as the first;

    procedure sample trig (sine, cosine);
    real sine, cosine;
    begin comment It has been stated that the angular
          co-ordinates φ are drawn from the density
          1/(2π), 0 ⩽ φ ⩽ 2π and it will be noted that
          these values appear in the algorithm only as
          arguments of sin and cos in the expression
          for P. This present procedure achieves this
          effect without explicit calculation of trigono-
          metric functions. Thus if t = tan ½φ, where
          0 ⩽ φ ⩽ π/2 then t has probability density
          4/(π(1 + t²)) for 0 ⩽ t ⩽ 1, and sin φ = 2t/(1 + t²),
          cos φ = (1 − t²)/(1 + t²). The remaining three quad-
          rants for φ are catered for by attaching
          random independent signs to sine and cosine;
          real t, D;

    reject : t := sud; D := 1 + t²;
             if sud > 1/D then go to reject;
    accept: sine := (if sud ⩾ 0·5 then 1 else − 1)
             *2*t/D; cosine := (if sud ⩾ 0·5 then 1
             else − 1) * (1 − t²)/D
    end sample trig;

    procedure vector product (U, V, P);
    array U, V, P;
    begin comment The vector product U × V
          is assigned to P;
      P[1] := U[2] * V[3] − U[3] * V[2];
      P[2] := U[3] * V[1] − U[1] * V[3];
      P[3] := U[1] * V[2] − U[2] * V[1]
    end vector product;

    real μ, α, β, EC, k, S, SS, ξ, η, sin φ, cos φ,
      rE, rC, EN, PN, rA, rB, term, I, σ;
    integer n, p, i;
    real array A, B, C, E, e1, e2, e3, W1, W2,
      P[1 : 3]; set (μ, n);
      read vectors (A, B, C); read (α, β);
    for p := 1, 2, 3 do
        E[p]:= (α* A[p]+β*B[p])/(α +β);
      EC := sqrt ((C[1] − E[1])² + (C[2] − E[2])² +
          (C[3] − E[3])²);
    for p := 1, 2, 3 do
      begin e3[p] := (C[p]−E[p])/EC;
            W1[p] := A[p] − B[p]
      end;
      vector product (e3, W1, W2);
      k := sqrt (W2[1]² + W2[2]² + W2[3]²);
      for p := 1, 2, 3 do e2[p] := W2[p]/k;
      vector product (e2, e3, e1);
      S := SS := 0;
      for i := 1 step 1 until n do
        begin ξ := sample exp * 2/μ + EC;
      reject: η := sample exp * 2/μ − EC;
              if η > EC then go to reject;
              sample trig (sin φ, cos φ); comment
                   sin φ and cos φ are identifiers
                   not procedure calls with argu-
                   ment φ;
            rE :=(ξ + η)/2;
            rC := (ξ − η)/2;
            EN := (rE² + EC² − rC²)/(2 * EC);
            PN := sqrt (rE² − EN²);
            for p := 1, 2, 3 do
              P[p] := E[p] + PN * cos φ * e1[p] +
                   PN * sin φ * e2[p] + EN * e3[p];
            rA := sqrt ( (P[1] − A[1])² + (P[2] − A[2])²
                       + (P[3] − A[3])²);
            rB := sqrt ( (P[1] − B[1])² + (P[2] − B[2])²
                       + (P[3] − B[3])²);
            term := (ξ + η) * exp (μ*rE − α*rA −
                       β*rB);
             S := S + term;
            SS := SS + term²
        end;
      k := (2*π/(μ²*EC))*(1 − exp (− μ*EC));
      I := k*S/n;
      σ := k*sqrt (SS/n − (S/n)²);
      print (I, σ)
end;
```

6. A three-centre density function

This third algorithm is a logical extension of the ideas contained in the other two. We suppose that A, B and C are distinct and not collinear, and we formulate the integral in terms of the three distances r_A, r_B, r_C as a co-ordinate system. We first establish the appropriate ranges for the co-ordinate values, then derive an expression for the element of volume, and finally develop the algorithm.

6.1 *The range of co-ordinate values*

Every position P determines a unique set (r_A, r_B, r_C). The converse situation is more complex. Some sets of values do not determine a position at all, and each valid set determines a pair of points P which are mirror images in the plane ABC. From consideration of the triangle PAB we may take

$$0 \leqslant r_A \leqslant \infty$$
$$|r_A - AB| \leqslant r_B \leqslant r_A + AB$$

and it remains to consider the consequent permitted range of r_C. The set of points having a given r_A and r_B

lie on a circle with centre 0 on AB and in a plane normal to AB. The positions of maximum and minimum r_C are the intersections U and L respectively of this circle with the plane ABC.

Let N be the foot of the perpendicular from C on to AB. We use

$$AN = (AB^2 + CA^2 - BC^2)/2AB$$
$$CN = (CA^2 - AN^2)^{1/2}$$
$$AO = (AB^2 + r_A^2 - r_B^2)/2AB$$

and $UO = LO = (r_A^2 - AO^2)^{1/2}$

to obtain $UC = [(AN - AO)^2 + (CN + UO)^2]^{1/2}$

and $LC = [(AN - AO)^2 + (CN - LO)^2]^{1/2}$,

where $LC \leqslant r_C \leqslant UC$.

6.2 *The element of volume*

Let (x, y, z) be the rectangular Cartesian co-ordinates of the general point P in some frame. Denote the position vectors of A, B, C and P relative to the origin of this frame by $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{r}$. We require the Jacobian

$$\frac{\partial(x, y, z)}{\partial(r_A, r_B, r_C)}$$

for then

$$dV \equiv dx\,dy\,dz = \left|\frac{\partial(x, y, z)}{\partial(r_A, r_B, r_C)}\right| dr_A\,dr_B\,dr_C.$$

Now $$\frac{\partial(x, y, z)}{\partial(r_A, r_B, r_C)} = \left(\frac{\partial(r_A, r_B, r_C)}{\partial(x, y, z)}\right)^{-1}$$

and $$\nabla r_A = (\boldsymbol{r} - \boldsymbol{A})/r_A, \text{ etc.}$$

so that

$$\frac{\partial(r_A, r_B, r_C)}{\partial(x, y, z)} = \frac{1}{r_A r_B r_C}\begin{vmatrix} x - A_x & x - B_x & x - C_x \\ y - A_y & y - B_y & y - C_y \\ z - A_z & z - B_z & z - C_z \end{vmatrix}$$

$$= D/r_A r_B r_C, \text{ say.}$$

Now, by multiplying the matrix of which D is the determinant on the left by its transpose, and by taking the determinant of the product, we obtain

$$D^2 = \begin{vmatrix} (\boldsymbol{r} - \boldsymbol{A})^2 & (\boldsymbol{r} - \boldsymbol{B}).(\boldsymbol{r} - \boldsymbol{A}) & (\boldsymbol{r} - \boldsymbol{C}).(\boldsymbol{r} - \boldsymbol{A}) \\ (\boldsymbol{r} - \boldsymbol{A}).(\boldsymbol{r} - \boldsymbol{B}) & (\boldsymbol{r} - \boldsymbol{B})^2 & (\boldsymbol{r} - \boldsymbol{C}).(\boldsymbol{r} - \boldsymbol{B}) \\ (\boldsymbol{r} - \boldsymbol{A}).(\boldsymbol{r} - \boldsymbol{C}) & (\boldsymbol{r} - \boldsymbol{B}).(\boldsymbol{r} - \boldsymbol{C}) & (\boldsymbol{r} - \boldsymbol{C})^2 \end{vmatrix}.$$

We now use $$(\boldsymbol{r} - \boldsymbol{A})^2 = r_A^2$$

and $$(\boldsymbol{r} - \boldsymbol{A}).(\boldsymbol{r} - \boldsymbol{B}) = \tfrac{1}{2}(r_A^2 + r_B^2 - AB^2), \text{ etc.,}$$

and, after writing

$$r_A^2, r_B^2, r_C^2 = \lambda, \mu, \nu$$

and $$BC^2, CA^2, AB^2 = l, m, n, \text{ respectively,}$$

we obtain after some algebraic manipulation of an elementary though tedious kind

$$dV = 2r_A r_B r_C K\,dr_A dr_B dr_C$$

where

$$K = \{(\lambda l + \mu\nu)(-l + m + n) + (\mu m + \nu\lambda)(l - m + n) + (\nu n + \lambda\mu)(l + m - n) - l\lambda^2 - m\mu^2 - n\nu^2 - lmn\}^{-1/2}.$$

6.3 *The algorithm*

In order to work with co-ordinates which were less interdependent than r_A and r_B, we transformed to ξ and η, where $\xi = r_A + r_B$ and $\eta = r_A - r_B$. We then have

$$AB \leqslant \xi \leqslant \infty \text{ and } -AB \leqslant \eta \leqslant AB.$$

The differential elements are related by

$$dr_A dr_B = \tfrac{1}{2}d\xi d\eta.$$

Considering the integral

$$I = \int e^{-\alpha r_A}\frac{1}{r_C}e^{-\beta r_B}dV,$$

we obtain $$I = \int \frac{1}{r_C}\exp\left(-\frac{\alpha + \beta}{2}\xi - \frac{\alpha - \beta}{2}\eta\right)dV$$

whence, making use of the symmetry about the plane ABC,

$$I = 2\int_{AB}^{\infty} e^{-\frac{\alpha+\beta}{2}\xi}d\xi \int_{-AB}^{AB} e^{-\frac{\alpha-\beta}{2}\eta}d\eta \int_{LC}^{UC} dr_C r_A r_B K.$$

We chose to sample ξ, η and r_C from distributions with probability densities as follows:

for ξ, density $\rho(\xi)$

$$= \frac{\alpha + \beta}{2}e^{-\frac{\alpha+\beta}{2}(\xi - AB)}, \qquad AB \leqslant \xi \leqslant \infty,$$

for η, density $\rho(\eta)$

$$= \tfrac{1}{4}(\alpha - \beta)\,\text{cosech}\,\frac{\alpha - \beta}{2}AB.e^{-\frac{\alpha-\beta}{2}\eta} \text{ if } \alpha \neq \beta$$

$$= \frac{1}{2AB} \qquad \text{if } \alpha = \beta, \qquad -AB \leqslant \eta \leqslant AB,$$

then for r_C, density $\rho(r_C)$

$$= \frac{1}{UC - LC}, \qquad LC \leqslant r_C \leqslant UC.$$

After N cycles, we have again

$$\hat{I}_N = \frac{1}{N}\Sigma z$$

and $$\hat{\sigma}_N^2 = \frac{1}{N}(\Sigma z^2) - (\hat{I}_N)^2,$$

where $$z = \frac{4}{\alpha + \beta}e^{-\frac{\alpha+\beta}{2}AB}F(UC - LC)r_A r_B K,$$

where $$F = \frac{4}{\alpha - \beta}\sinh\frac{\alpha - \beta}{2}AB \quad \text{if } \alpha \neq \beta$$

$$= 2AB \qquad \text{if } \alpha = \beta.$$

An ALGOL description of the calculation is given below.

```
begin comment The same general specification applies
      here as previously with N replacing n as the number
      of points;
      real BC, CA, AB, α, β, w, l, m, n, AN, CN, S, SS,
           ξ, η, rA, rB, AO, UO, UC, LC, rC, λ, μ, ν,
           K, term, k, I, σ;
      integer N, i;
      set (N);
      read (BC, CA, AB, α, β);
      if α < β then begin w := α; α := β; β := w;
                          w := CA; CA := AB; AB := w
                    end this relabels the data if neces-
                        sary to ensure α ⩾ β for future
                        convenience;
      l := BC²; m := CA²; n := AB²;
      AN := (− l + m + n)/(2*AB);
      CN := sqrt (CA² − AN²);
      S := SS := 0;
      for i := 1 step 1 until N do
          begin ξ := sample exp *2/(α + β) + AB;
              if α > β then begin reject : η := sample
                                    exp *2/(α − β) − AB;
                                if η > AB then go to reject
                                end
                       else η := AB*(2*sud − 1);
                  rA := (ξ + η)/2; rB := (ξ − η)/2;
                  AO := (AB² + rA² − rB²)/(2*AB);
                  UO := sqrt (rA² − AO²);
                  UC := sqrt ((AN − AO)² +
                                          (CN + UO)²);
                  LC := sqrt ((AN − AO)² +
                                          (CN − UO)²);
                  rC := LC + (UC − LC)*sud;
                   λ := rA²; μ := rB²; ν := rC²;
                   K := 1/sqrt ((λ*l + μ*ν)
                        *(− l + m + n) + (μ*m + ν*λ)
                        *(l − m + n) + (ν*n + λ*μ)
                        *(l + m − n) − l*λ² − m*μ²
                        − n*ν² − l*m*n);
                term := (UC − LC)*rA*rB*K;
                   S := S + term;
                  SS := SS + term²
          end;
      k := (4/(α + β)) * exp (− (α + β)*AB/2)
           *(if α = β then 2*AB else (4/(α − β))
           *(exp ((α − β)*AB/2) − exp (−(α−β)
           *AB/2))/2);
      I := k*S/N;
      σ := k*sqrt (SS/N − (S/N)²);
      print (I, σ)
end;
```

7. A comparison of methods

In this section we discuss the results obtained for a number of specific integrals. The number of such applications of the three algorithms is not as great as would perhaps be desirable but is sufficient to establish a qualitative assessment of the three methods, and to encourage a more thorough investigation when improved facilities become available.

7.1 *A single-centre integral*

The effect of varying the sampling parameter λ in the first algorithm (Section 4) was examined in the special case

$$I = \int e^{-r} \frac{1}{r} dV.$$

With the stated sampling scheme it may be shown by analytical integration that the choice which minimizes the standard deviation is $\lambda = \frac{1}{2}$. In fact

$$\sigma/I = \left(\frac{4}{27}\pi^2 - 1\right)^{1/2} = 0{\cdot}680 \quad \text{for } \lambda = \tfrac{1}{2}$$

$$= \left(\frac{1}{4}\pi^2 - 1\right)^{1/2} = 1{\cdot}211 \quad \text{for } \lambda = 1.$$

Thus for the apparently more natural choice $\lambda = 1$, the Central Limit Theorem shows that to achieve a given accuracy, a sample size some three times greater will be needed than in the optimum case $\lambda = \frac{1}{2}$.

These conclusions were verified numerically on the computer. Thus 600 points gave $\sigma/I = 0{\cdot}67$ with $\lambda = \frac{1}{2}$ and $1{\cdot}36$ with $\lambda = 1$. Similarly 8,600 points with $\lambda = \frac{1}{2}$ gave $\sigma/I = 0{\cdot}68$. We infer that for the general three-centre integral the choice $\lambda = \frac{1}{2}(\alpha + \beta)$ may well be preferred to $\lambda = \alpha + \beta$. Clearly there are implications here affecting the choice of the parameter μ in the second algorithm (Section 5). In fact all trials reported below were made with $\lambda = \mu = \alpha + \beta$. In explanation and excuse we would say that the time available was limited, and as attention became concentrated upon the third algorithm (Section 6), so the significance of the choice of λ and μ in the other methods was not fully appreciated at the time.

7.2 *Numerical results*

The algorithms were applied to four integrals, including both two- and three-centre dependence. The values of the two-centre integrals were available from explicit formulae, and approximations to the three-centre integrals from an independent method. All four were taken from a calculation on the water molecule by Boys and Reeves (to be published). The values quoted all refer to so-called normalized orbitals. That is to say, a numerical constant is absorbed into each orbital η_i as a factor in order to give $(\eta_i|G|\eta_i) = 1$. The value of the constant is of course $(\eta_i|G|\eta_i)^{-1/2}$ which can be evaluated explicitly. The integrals are as follows.

(i) $AB = AC = 4{\cdot}0$, angle $BAC = 120°$, $\alpha = 3{\cdot}0$, $\beta = 0{\cdot}5$
independent approximation $= 0{\cdot}0179$

(ii) $AC = BC = 4{\cdot}00$, angle $BCA = 120°$, $\alpha = \beta = 0{\cdot}5$
independent approximation $= 0{\cdot}0763$

Table 1

Tabulation of results for four integrals

		METHOD 1	METHOD 2	METHOD 3
Integral (i)	$\hat{I}$	$0{\cdot}0270 \pm 0{\cdot}0102$	$0{\cdot}0204 \pm 0{\cdot}0038$	$0{\cdot}0172 \pm 0{\cdot}0007$
	$\hat{\sigma}/\hat{I}$	14·6	7·2	1·5
Integral (ii)	$\hat{I}$	$0{\cdot}0773 \pm 0{\cdot}0036$		$0{\cdot}0806 \pm 0{\cdot}0050$
	$\hat{\sigma}/\hat{I}$	1·8		2·4
Integral (iii)	$\hat{I}$		$0{\cdot}2655 \pm 0{\cdot}0226$	$0{\cdot}2431 \pm 0{\cdot}0100$
	$\hat{\sigma}/\hat{I}$		3·3	1·6
Integral (iv)	$\hat{I}$		$0{\cdot}2461 \pm 0{\cdot}0140$	
	$\hat{\sigma}/\hat{I}$		2·2	

(iii) A and B coincident, $AC = BC = 4{\cdot}0$, $\alpha = \beta = 3{\cdot}0$
exact value = 0·2500
(iv) A and B coincident, $AC = BC = 4{\cdot}0$, $\alpha = \beta = 0{\cdot}5$
exact value = 0·2363.

In order to apply method three to the two-centre integrals, it was necessary to invent a fictitious third centre. Thus for integral (iii) we took $BC = CA = AB = 4{\cdot}0$, $\alpha = 6{\cdot}0$, $\beta = 0$. This device is not possible if C coincides with A or B, and such two-centre integrals could not be treated by method three. In practice this is irrelevant since all two-centre integrals can be evaluated from explicit formulae, integrals (iii) and (iv) being included here only for demonstration purposes.

The results obtained are set out in Table 1. The entries against $\hat{I}$ are in the form "estimate $\pm$ standard deviation" and relate to a sample of 1,500 points. It is seen that the difference between the estimate and the independently obtained values lies in each case well within three standard deviations; in fact none are more than one standard deviation in error.

The general superiority of method three is evident, though we may note the effectiveness of method one where, as in integral (ii), the integrand is dominated by the Coulomb potential. It is interesting to observe the comparative invariance in method three of the relative accuracy as shown by the small variation of $\hat{\sigma}/\hat{I}$.

8. Discussion

Method three was programmed in Pegasus basic code in fixed-point form with some double-length working. A running speed of about 100 points per minute was achieved. It will be noted that the only function evaluations required in the main cycle are square roots, so that not too many drum transfers were required.

With the ranges of parameter values likely to arise in a practical molecular configuration, the three-centre integrals are small. The integrals are all required to the same absolute accuracy, and so the three-centre integrals require the least relative accuracy. Nevertheless, none of the three methods in their present form seems sufficiently powerful for the evaluation of individual integrals as fast and accurately as required.

A feature of a molecular structure calculation, however, is the large number of integrals required. We noted in Section 2 that Slater orbitals have the general form $\eta = S_{l,m}(\theta, \phi) r_A^n e^{-\alpha r}$. The orbitals for a calculation can be grouped into classes according to their centre A and parameter α. These classes frequently have four or more members. For example the $2s$ and $2p$ orbitals on any nucleus may have a common parameter α and we obtain the class $(r\,e^{-\alpha r},\ x\,e^{-\alpha r},\ y\,e^{-\alpha r},\ z\,e^{-\alpha r})$. Suppose there are two such classes, one on centre A and the other on centre B. These give rise to sixteen distinct integrals of the form $(\eta_A|V_c|\eta_B)$, and, with method three, use the same sampling scheme. The only difference lies in the factors $r_A x_B$, $y_A x_B$, etc., which appear in the quantity to be averaged. It is thus possible to estimate all the integrals simultaneously in parallel at a cost negligibly more than that for a single integral.

The three- and four-centre electrostatic repulsion integrals present a much more serious problem. The economies to be gained from simultaneous estimations are correspondingly greater. Thus the $2s$ and $2p$ orbitals on each of four centres give rise to 256 distinct integrals which could be treated in parallel. It was, however, judged not feasible to attempt the development of suitable algorithms for these integrals on Pegasus, bearing in mind the advent of the much faster so-called second-generation computers with their powerful programming languages.

One of us (D.R.C.) is indebted to D.S.I.R. for the award of an Advanced Course Studentship under whose tenure the work described was carried out.

References

BARNETT, V. D. (1962). "The Behaviour of Pseudo-Random Sequences Generated on Computers by the Multiplicative Congruential Method," *Maths. of Comp.*, Vol. 16, p. 63.

BOYS, S. F., and COOK, G. B. (1960). "Mathematical Problems in the Complete Quantum Predictions of Chemical Phenomena," *Rev. Mod. Phys.*, Vol. 32, p. 285.

EDMONDS, A. R. (1960). "The Generation of Pseudo-Random Numbers on Electronic Digital Computers," *Comp. J.*, Vol. 2, p. 181.

KAHN, H. (1956). *See* Meyer, H. A., below.

KAHN, H. (1960). *See* Ralston, A., and Wilf, H. S., below.

MEYER, H. A. (1956), Ed. *Symposium on Monte Carlo Methods*, New York: Wiley.

RALSTON, A., and WILF, H. S. (1960), Eds. *Mathematical Methods for Digital Computers*, New York: Wiley.

SLATER, J. C. (1930). "Atomic Shielding Constants," *Phys. Rev.*, Vol. 36, p. 57.

VOTAW, D. F., and RAFFERTY, J. A. (1951). "High Speed Sampling," *M.T.A.C.*, Vol. 5, p. 1.

Book Review (*continued from p.* 276)

some problems in mathematical physics", formulates the Schrödinger equations for a system of particles in a potential field, in such a manner that if we make Planck's constant tend to zero in the equations, then these equations tend to those for classical mechanics. K. A. Bezhanov writes on "*The interaction between a shock wave and the free surface of a liquid*", and I. I. Nochevkina writes on "*Supersonic flow around conical bodies in an ideal liquid at different angles of attack*". V. V. Martynyuk's paper on "*The division of an algorithm scheme into networks*" is concerned with the transformation of computing algorithms into forms which are suitable for automatic programming. I. L. Sobel'man, in "*The relation between computing time, stage length and frequency of random machine errors*", examines the problem of optimizing the length of the stages into which a program should be divided in order to guard against machine faults.

Next there are three short communications on "*A method for partitioning a high order matrix into blocks in order to find its eigenvalues*" by V. A. Shishov, "*On some infinite systems of equations*" by E. G. Deĭch, and "*An approximate method of solution of Schrödinger's equation*" by V. B. Uvarov and A. F. Nikiforov. Finally, this issue lists the papers published in the remaining five parts of Volume 1 (1961), which include the Report on ALGOL 60 by Naur *et al.* Other papers are devoted to numerical solution of polynomials, theory of automata, error analysis in linear algebra, reactor kinetics, automatic programming, aerodynamics, queueing theory, finite-difference techniques for solving differential equations, and many other branches of numerical analysis and mathematical physics.

This translated journal should be a most useful addition to any mathematical library. But if any prospective reader hesitates to pay £50 per year, he is advised to acquire a smattering of the Russian language, buy a good Russian-English mathematical dictionary, and subscribe to the Russian journal at £4. 6s. 0d. per year. G. J. TEE

Correspondence

The Editor,
The Computer Journal.

Sir,

A Hardware Representation for ALGOL 60 using Creed Teleprinter Equipment

Regarding the paper by Gerard and Sambles (this *Journal*, Vol. 5, p. 338, 1963) and subsequent correspondence, I am surprised that no one has suggested designing a 5-hole tape code quite specifically for the representation of ALGOL. By using two non-feed characters, ___ and | , it is possible to have a code that gives printed texts that look like ALGOL, the main restriction being that only one size of letter is available. Such a code has, in fact, been implemented on conventional teleprinter equipment for R.R.E., Malvern, but as far as I know has not been publicized.

I have yet to be convinced that the advantages of 7- and 8-hole tape are so overwhelming that they justify the considerable price difference between the editing equipment required and that available for 5-hole tape. I suggest that it is worth our while to consider and agree on such a 5-hole code as this gives us the advantage of compatibility of our ALGOL representations and also allows the use of cheaper equipment. I would welcome the views of your readers.

Yours faithfully,

PAUL A. SAMET.

Computation Laboratory,
The University, Southampton.
17 June 1963.