

Stochastic Wave Function for Atomic Helium¹

M. H. KALOS

*Courant Institute of Mathematical Sciences,
New York University*

ABSTRACT

A method is given for random sampling which yields a density equal to the ground-state wave function for the helium atom. As an intermediate step, a Green's function for the diffusion operator containing the repulsive interaction is sampled. A technique for generating a density function proportional to ψ^2 is outlined. Attention is paid to the development of estimators with finite variance. Results are shown for the electron density and for $\langle r^2 \rangle$. The latter agrees well with previous numerical results.

1. INTRODUCTION

This report describes an extension of work described earlier [1] on the numerical solution by Monte Carlo integration of the Schrödinger equation for few body systems. The problem solved is the generation of the "exact" wave function for the ground state of atomic helium. The work is intended as an exploration of method, since the precision of the results is less than that already achieved for helium by a number of existing calculations [2, 3].² Nevertheless the treatment

¹ The work presented in this paper is supported by the AEC Computing and Applied Mathematics Center, Courant Institute of Mathematical Sciences, New York University, under Contract AT(30-1)-1480 with the U.S. Atomic Energy Commission.

² At the time this work was begun it appeared that a significant discrepancy might exist between estimates of $\langle r^2 \rangle$ obtained from wave functions that minimize the energy and those obtained from a different method due to J.-P. Auffray [4]. In addition there seems to be disagreement [5] between experiment and the computational result. However, further development [6] of Auffray's method has removed the deviation between his result and that of Pekeris [2].

of the singular (Coulomb) potential and the development of a method to estimate integrals with ψ^2 as weight function represent necessary advancements toward the solution of more interesting atomic and nuclear problems. For example, it appears that the development of wave functions for ground and first excited states of the hydrogen molecule are straightforward extensions of the helium work.

2. WAVE FUNCTION ITERATION

Let us review the general approach as discussed in the earlier paper [1]. For convenience suppose the energy is fixed as

$$E = -B \quad (1)$$

and replace the physical coordinates \mathbf{x}_i of particle i by \mathbf{r}_i where

$$\mathbf{r}_i = (2M_i B / \hbar^2)^{1/2} \mathbf{x}_i \quad (2)$$

and the potential V by W :

$$V(\mathbf{x}_1, \dots, \mathbf{x}_N) = -\lambda B W(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \equiv -\lambda B W(\mathbf{R}). \quad (3)$$

The Schrödinger equation is then

$$(-\nabla^2 + 1)\psi(\mathbf{R}) = \lambda W(\mathbf{R})\psi(\mathbf{R}). \quad (4)$$

As discussed in [1], the development of a square-integrable ψ requires that we use G_0 , the Green's function of $-\nabla^2 + 1$ which vanishes at infinity, and write

$$\psi(\mathbf{R}) = \lambda \int d\mathbf{R}' G_0(\mathbf{R}', \mathbf{R}) W(\mathbf{R}') \psi(\mathbf{R}'). \quad (5)$$

Explicitly,

$$G_0(\mathbf{R}_0, \mathbf{R}) = (2\pi)^{-M/2} |\mathbf{R} - \mathbf{R}_0|^{1-M/2} K_{M/2-1}(|\mathbf{R} - \mathbf{R}_0|), \quad (6)$$

where M is the number of dimensions ($M = 3N$) and $K_\nu(z)$ is the Bessel function of imaginary argument.

As explained in [1], G_0 may be regarded as the step-length distribution of a random walk. Then if $\psi(\mathbf{R}')$ is used as a density function for a set of points; if these points are multiplied according to the weight function³ $W(\mathbf{R}')$; and if then

new points \mathbf{R} are chosen according to the density function $G_0(\mathbf{R}', \mathbf{R})$, the points \mathbf{R} have the distribution given by the integral on the right side of (5). Repetition of this procedure is equivalent to iteration of the Eq. (5) and converges to that state for which λ is smallest.

For the helium atom

$$W = \frac{e^2}{B} \left[\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{r_{12}} \right] \quad (7)$$

and is sometimes negative. Now this is not necessarily fatal to the integration of the right side of (5). It is certainly possible to use a negative value of W . In the iteration, this means carrying along a sign which is changed when $W < 0$. If this is done, then $|W|$ is used for the multiplication. Now the iteration converges so that the density of points developed—ignoring the sign—is that of the lowest state with $|W|$ as potential. Any component of a higher state, such as the true ground state of W , decays away. Thus although the signed density of points may represent the desired state, it eventually becomes very difficult if not impossible to see it⁴ in the presence of the random fluctuations.

2.1 Green's Function Iteration.

A way around this difficulty would exist if the Green's function were found which contains all of the repulsive interaction. That is, write

$$W = W_+ - W_- \quad (8)$$

with $W_+, W_- \geq 0$.

The Green's function in question is the solution

$$(-\nabla^2 + 1 + W_-)G(\mathbf{R}', \mathbf{R}) = \delta(\mathbf{R}' - \mathbf{R}). \quad (9)$$

Using G , the integral equation for ψ is

³ That is, extra identical points are created, if necessary, so that, on the average, $W(\mathbf{R}')$ points at \mathbf{R}' are treated in the subsequent calculation.

⁴ If the iteration is carried out with an artificial eigenvalue, λ' , which is sufficiently large, then as the sample of points grows, the statistical error will remain smaller than the component of the W ground state. But this means that the computational effort required to improve any estimate by a fixed amount grows exponentially. Since, in addition, it is usually necessary to carry out some preliminary iterations before the ground-state distribution is established, it seems that this possibility does not help much. If it is known that the two eigenvalues (corresponding to $|W|$ and to W) do not differ much, then it may be very useful.

$$\psi(\mathbf{R}) = \int G(\mathbf{R}', \mathbf{R}) W_+(\mathbf{R}') \psi(\mathbf{R}') d\mathbf{R}'. \quad (10)$$

The random walk treatment can be applied provided points with density $G(\mathbf{R}', \mathbf{R})$ for fixed \mathbf{R}' can be supplied and if W_+ is used for the multiplication of points.

We may set

$$W_- = e^2/r_{12}.$$

With this choice, the Green's function (or its analytic continuation) has been discussed by several authors [7–9].

An attempt was made to find a convenient rapid computational scheme based on the analytic Green's function, but without success. Instead the expansion of the Green's function in terms of the unperturbed Green's function was carried out. If we move the term W_-G of (9) to the right side of the equation and use G_0 , the following results:

$$G(\mathbf{R}', \mathbf{R}_0) = G_0(\mathbf{R}', \mathbf{R}_0) - \int W_-(\mathbf{R}'') G(\mathbf{R}', \mathbf{R}'') G_0(\mathbf{R}'', \mathbf{R}) d\mathbf{R}''. \quad (11)$$

Such an equation may be solved by random sampling—if the process converges—to give a set of points having density G . The points may then be used to integrate the equation for ψ . Unfortunately the fact that the integral in (1) has a negative sign means that the points developed in this way frequently have a negative sign, which is precisely what was to be avoided. Since G itself is always positive it should be possible to achieve a large cancellation of opposite contributions. Such cancellation may be carried out conveniently using an iterated form of (11). Apply the operator $(-\nabla^2 + 1 - W_-(\mathbf{R}))$ to both sides of (11). The result is

$$\begin{aligned} (-\nabla^2 + 1)G(\mathbf{R}', \mathbf{R}) &= \delta(\mathbf{R} - \mathbf{R}') - W_-(\mathbf{R})G_0(\mathbf{R}', \mathbf{R}) \\ &+ W_-(\mathbf{R}) \int W_-(\mathbf{R}'') G(\mathbf{R}', \mathbf{R}'') G_0(\mathbf{R}'', \mathbf{R}) d\mathbf{R}''. \end{aligned}$$

The use of G_0 as Green's function for $(-\nabla^2 + 1)$ now gives

$$\begin{aligned} G(\mathbf{R}', \mathbf{R}) &= G_0(\mathbf{R}', \mathbf{R}) - \int G_0(\mathbf{R}', \mathbf{R}_1) W_-(\mathbf{R}_1) G_0(\mathbf{R}_1, \mathbf{R}) d\mathbf{R}_1 \\ &+ \iint G(\mathbf{R}', \mathbf{R}_1) W_-(\mathbf{R}_1) G_0(\mathbf{R}_1, \mathbf{R}_2) W_-(\mathbf{R}_2) G_0(\mathbf{R}_2, \mathbf{R}) d\mathbf{R}_1 d\mathbf{R}_2. \end{aligned} \quad (12)$$

This equation may also be solved by random sampling.

Suppose first that the “source” term

$$S(\mathbf{R}', \mathbf{R}'') = G_0(\mathbf{R}', \mathbf{R}'') - \int G_0(\mathbf{R}', \mathbf{R}_1) W_-(\mathbf{R}_1) G_0(\mathbf{R}_1, \mathbf{R}'') d\mathbf{R}_1 \quad (13)$$

may be sampled to give a point \mathbf{R}'' . Then, on the average $W_-(\mathbf{R}'')$ times, let a step

be taken to \mathbf{R}_2 using $G_0(\mathbf{R}', \mathbf{R}_2)$ as kernel; repeat the process using $W_-(\mathbf{R}_2)$ and $G_0(\mathbf{R}_2, \mathbf{R})$.

If we count up the points arriving at \mathbf{R} either drawn directly from $S(\mathbf{R}', \mathbf{R})$ or else after two steps from some \mathbf{R}_1 , a density results which satisfies Eq. (12). The process given above then describes a sampling procedure for generating points whose density is G . Such a procedure, valid for any \mathbf{R}' , is needed in generating the ψ according to (10).

The development of G converges most rapidly when W_- is made as small as possible. That is, put

$$\begin{aligned} W_-(\mathbf{R}) &= -W(\mathbf{R}), & \text{when } W(\mathbf{R}) < 0 \\ &= 0 & \text{otherwise.} \\ W_+(\mathbf{R}) &= W(\mathbf{R}), & \text{when } W(\mathbf{R}) \geq 0 \\ &= 0 & \text{otherwise.} \end{aligned} \tag{14}$$

The fact that W_- vanishes outside a certain volume in configuration space presents no problem in a calculation of this type.

The advantage of (12) and the sampling procedure that corresponds to it is that the Green's function iteration requires no sign changes.

2.2 Singularities of the Integral Equations.

The preceding discussion requires one essential modification in practice: the functions W_{\pm} are singular, diverging when any one of r_1, r_2 , or r_{12} becomes zero. The procedure of sampling new points whose number is on the average $W(\mathbf{R})$ cannot in general be carried out.

Let us denote by $\mathcal{G}, \mathcal{G}_0, \mathcal{S}$ the integral operators whose kernels are G, G_0 , and S , respectively. Equations (13) and (12) correspond to

$$\mathcal{S} = \mathcal{G}_0 - \mathcal{G}_0 W_- \mathcal{G}_0 \tag{15}$$

and

$$\mathcal{G} = \mathcal{S} + \mathcal{G}_0 W_- \mathcal{G}_0 W_- \mathcal{G}. \tag{16}$$

When the last equation is used with Eq. (10), describes the wave function iteration, the result is

$$\psi = \mathcal{S} W_+ \psi + \mathcal{G}_0 W_- \mathcal{G}_0 W_- \mathcal{G} W_+ \psi. \tag{17}$$

This equation may be operated upon on the left by W_+ and W_- , respectively, to give the equations

$$W_+\psi = W_+\mathcal{S}(W_+\psi) + W_+\mathcal{G}_0W_-\mathcal{G}_0W_-\mathcal{G}(W_+\psi), \quad (18)$$

$$W_-\psi = W_-\mathcal{S}(W_+\psi) + W_-\mathcal{G}_0W_-\mathcal{G}_0W_-\mathcal{G}(W_+\psi). \quad (19)$$

The solution of the first of these corresponds to the expansion

$$W_+\psi = W_+\left(1 + \sum_{n=1}^{\infty} \mathcal{G}_0W_-\right)^{2n} \mathcal{S}(W_+\psi). \quad (20)$$

In this form, it is clear that \mathcal{G}_0 is always preceded by W_{\pm} ; this factor—or at least its singular part—can be combined with the kernel G_0 to give a “Coulomb modified kernel,” Eq. (43) of Appendix A.

$$C^-(\mathbf{R}', \mathbf{R}) = A(r'_{12})G_0(\mathbf{R}', \mathbf{R})/r_{12}, \quad (21)$$

where

$$A(r) = r[1 - \exp(r)]^{-1}. \quad (22)$$

In the same way, the appropriate modified kernel for attractive interactions is

$$C^+ = \frac{A(r'_2)}{A(r'_1) + A(r'_2)} \left[\frac{A(r'_1)G_0(\mathbf{R}', \mathbf{R})}{r_1} \right] + \frac{A(r'_1)}{A(r'_1) + A(r'_2)} \left[\frac{A(r'_2)G_0(\mathbf{R}', \mathbf{R})}{r_2} \right]. \quad (23)$$

The terms in brackets are each of the form taken up in Appendix A, and the coefficients are such that each may be used as a probability for sampling the corresponding kernel.

Let W_- be replaced by

$$W'_- = r_{12}W_-/A(r'_{12}) \quad (24)$$

and W_+ by

$$W'_+ = G_0(\mathbf{R}', \mathbf{R})W_+(\mathbf{R})/C^+. \quad (25)$$

Denote by \mathcal{E}^+ and \mathcal{E}^- the operators whose kernels are C^+ and C^- , respectively. Equation (19) then becomes

$$W_-\psi = W_-\mathcal{S}(W_-\psi) + W'_-\mathcal{E}^- - W'_-\mathcal{E}^-W'_-\mathcal{E}^-(W_+\psi). \quad (26)$$

In developing the solution of this equation W'_- , a bounded function, plays the role of W_- , i.e., is used to determine the extent to which point \mathbf{R} is carried further in the wave function iteration. W_+ is similarly replaced by the bounded W'_+ .

2.3 Singularities of the Source Term

The source term S of Eq. (13) has singularities, both in W_- and in the Green's function. These are evaded in a manner analogous to that described above.

First a point \mathbf{R}'' is chosen (from a kernel $C^+(\mathbf{R}', \mathbf{R}'')$ or $C^-(\mathbf{R}', \mathbf{R}'')$ depending upon which iteration is being carried out). Then a point \mathbf{R} is generated from the kernel

$$D(\mathbf{R} | \mathbf{R}', \mathbf{R}'') = \frac{A(r_1'')}{A(r_1') + A(r_1'')} \left[\frac{A(r_1')G_0(\mathbf{R}', \mathbf{R})}{r_1} \right] + \frac{A(r_1')}{A(r_1') + A(r_1'')} \left[\frac{A(r_1'')G_0(\mathbf{R}'', \mathbf{R})}{r_1} \right]. \quad (27)$$

The weight

$$W'' = \frac{G_0(\mathbf{R}', \mathbf{R}'')}{C^\pm(\mathbf{R}', \mathbf{R}'')} \left[1 - \frac{G_0(\mathbf{R}', \mathbf{R})W_-(\mathbf{R})G_0(\mathbf{R}, \mathbf{R}'')}{D(\mathbf{R} | \mathbf{R}', \mathbf{R}'')} \right] W_\pm(\mathbf{R}'') \quad (28)$$

is evaluated numerically. The last term is W_+ or W_- according as the point is to be used in wave function or Green's function iteration. $|W''|$ is used as the expected number of points assigned to \mathbf{R}'' . Since W'' is not always positive, a sign must be kept, but it was found experimentally that the fraction of points to which a negative sign must be given is of the order of 3×10^{-4} . This is sufficiently small to permit the wave function to be iterated of the order of 10^2 without serious loss of significance.⁵

Note also that the function D used to sample points \mathbf{R}_1 contains the singularities of the integrand of S so that the ratio which occurs inside the bracket of (28) is bounded. This is essential in the procedure, and in particular in keeping small the fraction of negative points. It is worth mentioning that more elaborate procedures for integrating (13) within the framework of Monte Carlo could have been used if further improvement had been necessary.

3. ESTIMATION OF ψ^2

When a function $\psi(\mathbf{R})$ is sampled repeatedly it is straightforward to obtain linear functionals of the form

$$\langle f \rangle_1 = \int \psi(\mathbf{R})f(\mathbf{R})d\mathbf{R}. \quad (29)$$

⁵ It is likely that points with negative signs contribute, on the average, less than others in the development of the wave function iteration. After 80 iterations in the numerical work, the negative strength grew to a fraction of only 0.002 of the total.

Let $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_i$ be the points that turn up in I independent samples. Then, under rather general conditions [10] the sample average

$$\bar{f} = \sum_{l=1}^i f(\mathbf{R}_l)/I$$

converges in probability to the integral (29). In certain applications (e.g., linear transport theory), this is the kind of result which is generally sought.

For computations in which quantum mechanical wave functions are developed, primary interest is in functionals of the form

$$\langle f \rangle_2 = \int |\psi(\mathbf{R})|^2 f(\mathbf{R}) d\mathbf{R} \quad (30)$$

or in ratios like

$$\int |\psi(\mathbf{R})|^2 f(\mathbf{R}) d\mathbf{R} / \int |\psi(\mathbf{R})|^2 d\mathbf{R}. \quad (31)$$

If, as we suppose, a Monte Carlo solution is most attractive for many dimensional problems, then quadratic functionals present something of a difficulty: it is not possible to tabulate the density ψ over a mesh in many dimensions and then square each result. A method is required, analogous to that for linear functionals, in which points are used in the estimation of ψ^2 as they appear in the sampling for ψ .

Fortunately, this is possible when ψ satisfies an integral equation—as, for example Eq. (5):

$$\psi(\mathbf{R}) = \int G_0(\mathbf{R}', \mathbf{R}) W(\mathbf{R}') \psi(\mathbf{R}') d\mathbf{R}'.$$

Suppose that the sampling procedure is carried out twice to give independent sets $\{\mathbf{R}^{(1)}\}$ and $\{\mathbf{R}^{(2)}\}$. Take one point from each set, $\mathbf{R}_j^{(1)}, \mathbf{R}_k^{(2)}$, and for any \mathbf{R} , form the product

$$F = W(\mathbf{R}_j^{(1)}) G_0(\mathbf{R}_j^{(1)}, \mathbf{R}) W(\mathbf{R}_k^{(2)}) G_0(\mathbf{R}_k^{(2)}, \mathbf{R}). \quad (32)$$

Let this product be formed for I pairs of independent points. The expected value of the sum of the products is obtained by averaging separately over $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$, using for each the density function ψ :

$$\begin{aligned} \langle \sum F \rangle &= I \int G_0(\mathbf{R}^{(1)}, \mathbf{R}) W(\mathbf{R}^{(1)}) \psi(\mathbf{R}^{(1)}) d\mathbf{R}^{(1)} \\ &\quad \times \int G_0(\mathbf{R}^{(2)}, \mathbf{R}) W(\mathbf{R}^{(2)}) \psi(\mathbf{R}^{(2)}) d\mathbf{R}^{(2)} \\ &= \lambda^{-2} I \psi^2(\mathbf{R}). \end{aligned} \quad (33)$$

In order to obtain integrals over \mathbf{R} , we only need to allow \mathbf{R} to be varied over

the entire space. For example, if \mathbf{R} is selected from some density function $T(\mathbf{R})$, then the expected value of the sum of I evaluation of

$$F = W(\mathbf{R}^{(1)})G_0(\mathbf{R}^{(1)}, \mathbf{R})W(\mathbf{R}^{(2)})G_0(\mathbf{R}^{(2)}, \mathbf{R})f(\mathbf{R})/T(\mathbf{R}) \quad (34)$$

is

$$\begin{aligned} \langle \sum F \rangle &= \lambda^{-2} I \int T(\mathbf{R}) d\mathbf{R} \int d\mathbf{R}^{(1)} \int d\mathbf{R}^{(2)} F \\ &= \lambda^{-2} I \int f(\mathbf{R}) \psi^2(\mathbf{R}) d\mathbf{R} \end{aligned} \quad (35)$$

from which (30) can be inferred. By making pairs of estimates of this kind in one of which $f = 1$, the integrals required for (31) may be obtained. The quotient of the sums is a biased estimator for the quotient, but the bias is expected to become small as the sampling proceeds.

This treatment can also be carried out using (10), the integral equation containing the perturbed Green's function. As a practical matter, this did not seem necessary since the unperturbed Green's function is so readily computed and the difficulty with negative points does not propagate itself.

In fact, points were generated in the random walk having densities $W_+\psi$ and $W_-\psi$, as given by Eqs. (18) and (19). For points from both of these, estimator (34) is replaced by

$$F' = [\pm G_0(\mathbf{R}^{(1)}, \mathbf{R})] [\pm G_0(\mathbf{R}^{(2)}, \mathbf{R})] f(\mathbf{R})/T(\mathbf{R}). \quad (36)$$

The sign used for each factor is the sign in the W_\pm associated with the point $\mathbf{R}^{(1)}$ or $\mathbf{R}^{(2)}$.

The singularity of G_0 naturally somewhat complicates the procedure. The treatment of this follows the lines of Sections (2.2) and (2.3), and details are deferred to Appendix B along with a discussion of other devices used to improve the ψ^2 calculation.

The machine computation was set up with

$$f(\mathbf{R}) = \frac{1}{2}(r_1^2 + r_2^2) \quad (37)$$

so as to calculate the mean square radius of the helium atom. In addition, the value of (36) was tabulated in a histogram as a function of r_1 . This gives the charge density of the helium atom as a function of electron radial coordinate.

It is, of course, possible to obtain a variety of estimates at the same time; the extra computation is generally small compared with the time required for the random walk and ψ^2 estimation itself.

4. RESULTS

A computer program was written following the scheme described in the preceding sections and in the appendices. A trial function was used which consisted simply of a product of hydrogen-like wave functions adjusted to have the right $\langle r^2 \rangle$. The wave function iteration was carried out, and after 14 such cycles, the generation size appeared approximately to have stabilized. The computation of $\langle r^2 \rangle$ and $\langle 1 \rangle$ was performed in 67 additional iterations.⁶

Because there is no attempt to stabilize the generation size, it drifted, showing correlations in direction for 6–20 iterations. It is reasonable that such long periods would be observed since the average size of a “step” in the random walk is small compared with the size of the equilibrium distribution. For this reason a fluctuation can persist for many generations. At the end of the computation, the generation size was very close to the value after 14 iterations. (This is somewhat fortuitous; in any case this was not the criterion for termination.) From the change in size one might deduce a value of the strength of the interaction required to give the correct binding as 1 part in 10^4 smaller than the correct value. Estimating this is of course not the main aim of the calculation but it is reassuring.

The overall average value of $\langle r^2 \rangle$ was, in the units of Eq. (2),

$$\langle r^2 \rangle = 6.92_{84}$$

or using

$$(\hbar^2/2mB)^{1/2} = 0.41496 \text{ a.u.},$$

$$\langle r^2 \rangle = 1.193_0.$$

This value agrees very well with the result $\langle r^2 \rangle = 1.193483$ obtained by Pekeris [2] or the extrapolated value of 1.1935 given by Scherr and Knight [3].

The statistical error of $\langle r^2 \rangle$ is somewhat difficult to estimate because of long-term correlations in partial average values. By breaking up the run in different ways, estimates of relative error (i.e., standard deviation divided by average value) of 0.0015–0.004 may be obtained; smaller values derive from divisions into two or three separate runs. Much additional computation is needed to clear this up. Comparison of average values of partial quotients with overall quotient

⁶ In the existing program, the preliminary iterations took about 3 minutes each to follow about 20,000 points. Iterations which include the ψ^2 evaluation took about 5 minutes. These times can be reduced by large factors in future applications by avoiding FORTRAN on the CDC 6600 (which loses a factor of 3 to 4 in running time in existing compilers) at least for crucial subroutines, and by replacing certain sampling procedures by more efficient ones.

of $\langle r^2 \rangle / \langle 1 \rangle$ shows that any bias in the computation is less than 10^{-4} on a relative basis.

Figure 1 shows the density of electron (normalized to one) as a function of distance from the nucleus. The change in slope shows the effect of increasing screening by an inner electron upon the density of an electron far from the nucleus.

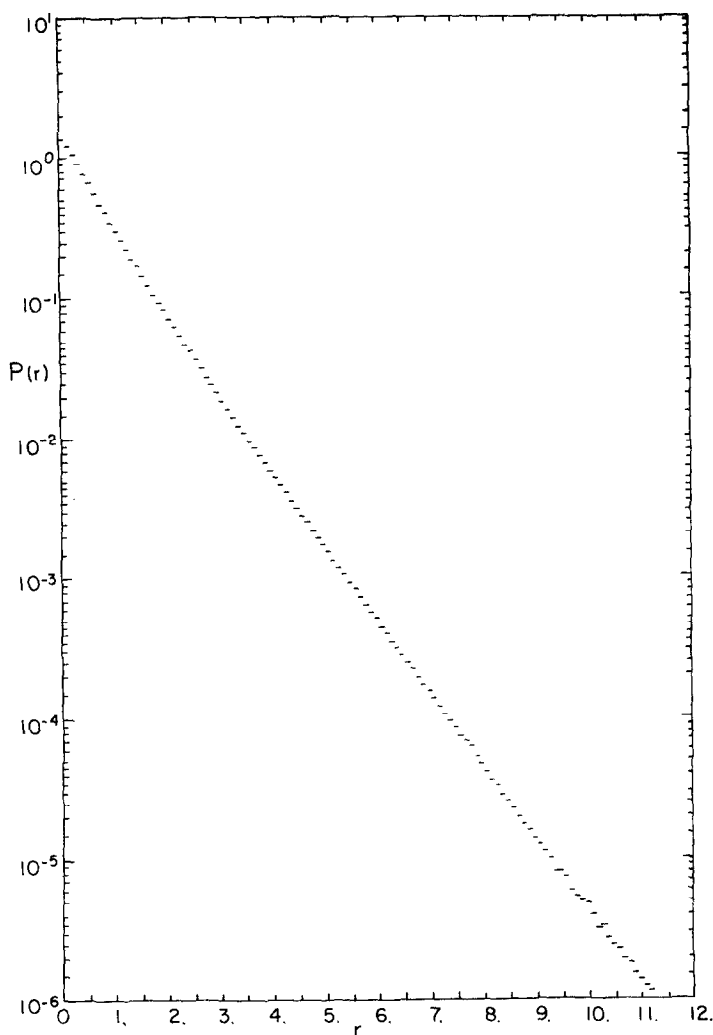


FIG. 1. Electron density for the helium atom as a function of radial distance from the nucleus, normalized to $\int P(r)r^2 dr = 1$. r is given in units of $(\hbar^2/2mB)^{1/2} = 0.415$ atomic units.

The results show that Monte Carlo calculations are indeed feasible for two-electron systems. It remains to be seen whether estimates, having the precision which can be attained by these methods in more complicated systems, will be found useful.

ACKNOWLEDGMENTS

The computations were carried out on the CDC 6600 Computer at New York University during a period when the machine behavior was more than slightly erratic. I should like to thank the operation staff, and particularly T. C. Moore, for their cooperation, without which the work would have been literally impossible. The assistance of W. Ray of CDC was most valuable.

APPENDIX A

Coulomb Modified Kernel

Let \mathbf{R} and \mathbf{R}' be vectors in the full configuration space (six-dimensional for the two-electron problem). Let \mathbf{r}_1 be the vector formed from the first three components of \mathbf{R} , and \mathbf{r}_2 represent the remaining elements. In the same way form $\mathbf{r}_1', \mathbf{r}_2'$ from the components of \mathbf{R}' . As has been shown above, it is necessary to select a point \mathbf{R} from a p.d.f. $C(\mathbf{R}', \mathbf{R})$ having the form

$$C(\mathbf{R}', \mathbf{R}) \propto \frac{1}{r_1} G_0(\mathbf{R}', \mathbf{R}) \equiv C'(\mathbf{R}', \mathbf{R}). \quad (38)$$

We may introduce the integral representation [11] for the Bessel function used in G_0 to obtain⁷

$$C' = (4\pi)^{-N/2} r_1^{-1} \int_0^\infty t^{-N/2} \exp[-t - |\mathbf{R} - \mathbf{R}'|^2/(4t)] dt. \quad (39)$$

Setting

$$|\mathbf{R} - \mathbf{R}'|^2 = |\mathbf{r}_1 - \mathbf{r}_1'|^2 + |\mathbf{r}_2 - \mathbf{r}_2'|^2 \quad (40)$$

we may immediately integrate over the $N-3$ dimensions denoted by \mathbf{r}_2 to find

⁷ See also [1], Eq. (11). There is a misprint in (11) and in (12). The exponents $-3N$ and $-3N+3$ should be $-3N/2$ and $(-3N+3)/2$, respectively.

$$\begin{aligned} \int C'(\mathbf{R}', \mathbf{R}) d\mathbf{r}_2 &= (4\pi)^{-3/2} r_1^{-1} \int_0^\infty t^{-3/2} \exp[-t - |\mathbf{r}_1 - \mathbf{r}_1'|^2/(4t)] dt \\ &= (2\pi)^{-3/2} K_{1/2}(|\mathbf{r}_1 - \mathbf{r}_1'|) / (r_1 |\mathbf{r}_1 - \mathbf{r}_1'|^{1/2}) \\ &= \frac{1}{4\pi} \frac{\exp[-|\mathbf{r}_1 - \mathbf{r}_1'|]}{r_1 |\mathbf{r}_1 - \mathbf{r}_1'|} \end{aligned} \tag{41}$$

$$\int C'(\mathbf{R}', \mathbf{R}) d\mathbf{r}_2 d\mathbf{r}_1 = [1 - \exp(-r_1')]/r_1'. \tag{42}$$

From (42) we see that the normalized density function C is

$$C = \frac{r_1'}{r_1 [1 - \exp(-r_1')]} G_0(\mathbf{R}', \mathbf{R}). \tag{43}$$

Define $C_3(\mathbf{r}_1', \mathbf{r}_1)$ to be the normalized density function that corresponds to the integral (41), that is, the marginal distribution for \mathbf{r}_1 in C . We may sample C_3 as follows. Set $s = |\mathbf{r}_1 - \mathbf{r}_1'|$; using Eqs. (41) and (42), we may write

$$C_3 = \frac{r_1'}{4\pi [1 - \exp(-r_1')]} \frac{\exp(-s)}{r_1 s}. \tag{44}$$

Let

$$\begin{aligned} v &= \frac{1}{2}(r_1 + s - r_1'); & 0 \leq v < \infty \\ u &= \frac{1}{2}(s - r_1 + r_1'); & 0 \leq u \leq r_1' \end{aligned} \tag{45}$$

and use φ (the azimuth of \mathbf{r}_1 about \mathbf{r}_1' as polar axis) as the third variable. Then

$$\begin{aligned} C_3 d\mathbf{r}_1 &= \frac{r_1'}{4\pi [1 - \exp(-r_1')]} \frac{\exp(-s)}{r_1 s} r_1^2 dr_1 d\omega d\varphi \\ &= \frac{\exp(-u - v)}{[1 - \exp(-r_1')]} du dv d\varphi / 2\pi. \end{aligned} \tag{46}$$

Thus u, v, φ are independent; φ is equidistributed over $(0, 2\pi)$, v is exponential on $(0, \infty)$, and u is exponential on $(0, r_1')$. Each of these distributions is easily sampled. The construction of \mathbf{r}_1 given \mathbf{r}_1', φ , and

$$\omega = \mathbf{r}_1' \cdot \mathbf{r}_1 = \frac{r_1'^2 + r_1^2 - s^2}{2r_1 r_1'}$$

is standard in the simulation of scattering in transport Monte Carlo [11].

With \mathbf{r}_1 given, it remains to pick \mathbf{r}_2 , or equivalently

$$\boldsymbol{\rho} = \mathbf{r}_2 - \mathbf{r}_2',$$

from the conditional distribution

$$Q(\boldsymbol{\rho} \mid \mathbf{R}', \mathbf{r}_1) = C(\mathbf{R}', \mathbf{R})/C_3(\mathbf{r}_1', \mathbf{r}_1). \quad (47)$$

Using $N-3$ dimensional polar coordinates, $\varrho, \boldsymbol{\Omega}$ to represent $\boldsymbol{\rho}$, the ϱ dependence of Q has the form

$$\begin{aligned} Q(\varrho) &\propto \varrho^{N-4} \int_0^\infty t^{-N/2} \exp[-t - (s^2 + \varrho^2)/4t] dt \\ &= 2\varrho^{N-4} \int_0^\infty w^{N-3} \exp[-w^{-2} - (s^2 + \varrho^2)w^2/4] dw. \end{aligned} \quad (48)$$

Now if w be drawn from

$$f(w) \propto \exp[-w^{-2} - s^2w^2/4] \quad (49)$$

and z from

$$g(z) \propto z^{N-4} \exp[-z^2/4], \quad (50)$$

then

$$\varrho = z/w \quad (51)$$

has the density function

$$Q(\varrho) = \int_0^\infty wf(w)g(\varrho w) dw$$

as required by (48). Note that (50) is the density function for the length of a vector \mathbf{z} in $N-4$ dimensions each of whose components has a normal distribution. If we sample such a vector then, since $\boldsymbol{\rho}$ is isotropic, we may obtain the components of $\boldsymbol{\rho}$ from the components of \mathbf{z} according to (51).

To sample from $f(w)$ as given by (49), note first that

$$w^{-2} + b^4w^2 = b^2[(bw)^{-2} + (bw)^2]$$

so that bw is a natural variable to use. Secondly,

$$\eta(x - \log x - 1) + 2 < x^{-2} + x^2;$$

$$\eta = 7.2959.$$

(the constant η is obtained numerically). Thus

$$\exp[-b^2(x^{-2} + x^2)] < \exp[-\eta b(x - \log x - 1) - 2b];$$

with

$$\begin{aligned} n &= \text{integer part of } \eta b, \\ \exp[-b^2(x^{-2} + x^2)] &< \exp[-n(x - \log x - 1) - 2b] \\ &\propto x^n \exp(-nx) \end{aligned}$$

If x is drawn⁸ from the p.d.f. $x^n \exp(-nx)/(n!/n^{n+1})$; and if the resulting x is kept with probability

$$\exp[-b^2(x^{-2} + x^2) + n(x - \log x - 1) + 2b] < 1;$$

(and otherwise a new x is generated and tested), the result, when divided by b (i.e., $s^{1/2}$) has the distribution (49).

APPENDIX B

Reduction of Variance of ψ^2 Computation

B.1 Finite Variance

As explained in Section 3, integrals weighted with ψ^2 may be computed if two independent samples of points are available, each having density $W\psi$. By connecting a point \mathbf{R} to one point from each sample (call them now $\mathbf{R}_1, \mathbf{R}_2$), and by computing two values of the Green's function, a density ψ^2 is developed.

For fixed $\mathbf{R}_1, \mathbf{R}_2$

$$\begin{aligned} \int f(\mathbf{R})\psi^2(\mathbf{R})d\mathbf{R} &= \int G_0(\mathbf{R}_1, \mathbf{R})G_0(\mathbf{R}_2, \mathbf{R})f(\mathbf{R})d\mathbf{R} \\ &= \int \left\{ \frac{G_0(\mathbf{R}_1, \mathbf{R})G_0(\mathbf{R}_2, \mathbf{R})}{T(\mathbf{R} | \mathbf{R}_1, \mathbf{R}_2)} f(\mathbf{R}) \right\} T(\mathbf{R} | \mathbf{R}_1, \mathbf{R}_2)d\mathbf{R}. \end{aligned} \quad (52)$$

$T(\mathbf{R} | \mathbf{R}_1, \mathbf{R}_2)$ is a density function governing the selection of \mathbf{R} ; it is conditional upon \mathbf{R}_1 and \mathbf{R}_2 . Accordingly, the quantity in braces in Eq. (52) is evaluated numerically.

Since T is at our disposal, we seek that form which makes the statistical error of the integration as small as possible. The optimum form of T is

⁸ Cf. [1], Eq. (A. 11) with $u = nx$.

$$T_0(\mathbf{R} | \mathbf{R}_1, \mathbf{R}_2) = \frac{G_0(\mathbf{R}_1, \mathbf{R})G_0(\mathbf{R}_2, \mathbf{R})f(\mathbf{R})}{\int G_0(\mathbf{R}_1, \mathbf{R}')G_0(\mathbf{R}_2, \mathbf{R}')f(\mathbf{R}')d\mathbf{R}'}, \tag{53}$$

for with it the estimator in (52) is a constant and there is no variance. Since this constant is the integral to be evaluated we must assume that the form of (53) cannot be used. But we may be guided by this in attempting to make the ratio $G_0(\mathbf{R}_1, \mathbf{R})G_0(\mathbf{R}_2, \mathbf{R})/T(\mathbf{R} | \mathbf{R}_1, \mathbf{R}_2)$ vary as slowly as possible where T itself is large. In particular, we must examine the effect of the singularities of G_0 upon the choice of T .

For small $|\mathbf{R}_1 - \mathbf{R}|$, and in M dimensions,

$$G_0(\mathbf{R}_1, \mathbf{R}) \propto |\mathbf{R}_1 - \mathbf{R}|^{-(M-2)}. \tag{54}$$

The merits of various sampling functions may be measured by the mean square of the estimator:

$$\int \left\{ \frac{G_0(\mathbf{R}_1, \mathbf{R})G_0(\mathbf{R}_2, \mathbf{R})f(\mathbf{R})}{T(\mathbf{R} | \mathbf{R}_1, \mathbf{R}_2)} \right\}^2 T(\mathbf{R} | \mathbf{R}_1, \mathbf{R}_2) d\mathbf{R}. \tag{55}$$

Write

$$\mathbf{R} = \mathbf{R}_1 + \mathbf{R}'.$$

In the neighborhood of $R' = 0$, use $d\mathbf{R}' = (R')^{M-1}d\Omega$.

Then the integrand in (55) has limiting behavior proportional to

$$(R')^{-2(M-2)}G_0(\mathbf{R}_2, \mathbf{R}_1)f(\mathbf{R}_1)(R')^{M-1}/T(R')$$

If T itself is bounded as $R' \rightarrow 0$, the integral of this diverges for $M \geq 4$. Suppose T diverges as $(R')^{-m}$, then

$$m < M$$

ensures that T itself is integrable, and

$$m > M - 4$$

is required in order that the variance converge. Naturally the same behavior must hold for $\mathbf{R} \rightarrow \mathbf{R}_2$.

We have used

$$m = M - 2$$

so that the singularity at $\mathbf{R} = \mathbf{R}_1$ and $\mathbf{R} = \mathbf{R}_2$ matches the behavior of G_0 . The variance also increases as the points \mathbf{R}_1 and \mathbf{R}_2 vary, but this does not lead to any singularity.

Within these constraints, which make the variance finite, a wide choice of T is permitted. The form used in the computation was

$$T = \tau \left\{ \frac{\varrho^2}{R_1^4(R_1^2 + \varrho^2)^2} + \frac{\varrho^2}{R_2^4(R_2^2 + \varrho^2)^2} + \left[\frac{2.5\varrho}{\varrho^2 + 2R_3^2} \right]^6 \right\}, \quad (56)$$

where

$$\begin{aligned} \varrho^2 &= | \mathbf{R}_1 - \mathbf{R}_2 |^2, \\ R_3^2 &= | \mathbf{R} - \frac{1}{2}\mathbf{R}_2 - \frac{1}{2}\mathbf{R}_3 |^2, \end{aligned}$$

and τ is a normalizing constant.

The first two terms of T were arbitrarily picked as easy to sample while having the correct singularities and general scale. The third term was added on the basis of some numerical evaluation of $G_0(\mathbf{R}_1, \mathbf{R})G_0(\mathbf{R}_2, \mathbf{R})/T$ using several possible forms of T . The region around $R_3 = 0$ was given particular attention since it contributes significantly more than the first two terms of T would indicate.

B.2 Rotation of Coordinates.

A second device used to improve the estimate of ψ^2 makes explicit use of the rotational invariance of the wave functions being developed. That is, if \mathbf{R}_1 is a point obtained in the random walk, and if \mathbf{R}_1' is obtained by a three-dimensional rotation (of the coordinates of \mathbf{r}_1 and \mathbf{r}_2 together), then answers may be obtained from \mathbf{R}_1' as well as \mathbf{R}_1 . In particular, \mathbf{R}_1' may be used with any \mathbf{R}_2 to obtain an estimate of ψ^2 .

Naturally, no advantage is gained if the rotation which leads to \mathbf{R}_1' is isotropic. However, for fixed \mathbf{R}_2 , a biased rotation which tends to minimize the variance of the estimator may be used. Let us choose the form of this rotation which is required when the overall answer is large, i.e., when

$$\varrho = | \mathbf{R}_1' - \mathbf{R}_2 |$$

is small. In this limit, ϱ is the only scale factor of length. Then with $m = M - 2$, a normalized T has the limiting form

$$T \sim \varrho^{M-4}/(R' R'')^{M-2}.$$

Let $\Pi(\varrho)$ be a probability density function for ϱ , to be used in biasing R_1' . A weight of Π^{-1} must be used, of course, to ensure correct estimates. Then the mean square answer has the form

$$\begin{aligned} \langle F^2 \rangle &= \int d\mathbf{p} \Pi(\varrho) \cdot \frac{1}{\Pi^2(\varrho)} \int \frac{\{G_0(\mathbf{R}_1', \mathbf{R})G_0(\mathbf{R}_2, \mathbf{R})\}^2}{T(\mathbf{R})} f^2(\mathbf{R}) d\mathbf{R} \\ &\sim \int d\varrho \Pi^{-1}(\varrho) \int \frac{|R'R''|^{-8}(R'R'')^4}{\varrho^2} f^2(\mathbf{R}) R^5 d\mathbf{R}. \end{aligned} \quad (57)$$

Assuming $f(R)$ is regular at $\varrho = 0$, the inner integral goes as ϱ^{-4} .

$$\langle F^2 \rangle \sim f^2(0) \int d\varrho \Pi^{-1}(\varrho) \varrho^{-4}. \quad (58)$$

To minimize $\langle F^2 \rangle$, subject to the constraint that

$$\int \Pi(\varrho) d\mathbf{p} = 1,$$

it is necessary that

$$\Pi(\varrho) \propto \varrho^{-2},$$

which is, in fact, the limiting behavior of the integral $\int G_0(\mathbf{R}', \mathbf{R})G_0(\mathbf{R}'', \mathbf{R}) d\mathbf{R}$ itself.

The (random) rotations upon \mathbf{R}_1 to produce \mathbf{R}_1' were designed, then, to arrange that $|\mathbf{R}_1' - \mathbf{R}_2| = \varrho$ should have a density proportional to ϱ^{-2} . Now each vector (e.g., \mathbf{R}_1) is a juxtaposition of two three-dimensional vectors ($\mathbf{s}_1, \mathbf{t}_1$). The kind of rotation permitted leaves $s_1, t_1, s_1 \cdot t_1$ unchanged. It is impossible to attain $\varrho = 0$, but we attempt to choose parameter of the rotation so as to minimize ϱ within a range governed by the ϱ^{-2} variation. Now

$$\varrho^2 = |\mathbf{s}_1' - \mathbf{s}_2|^2 + |\mathbf{t}_1' - \mathbf{t}_2|^2. \quad (59)$$

It is clear geometrically that ϱ^2 is minimized when the four vectors are coplanar. Let δ be the angle between s_2 and s_1' . Then in the plane the configuration is specified by setting

$$\sin \delta = \frac{\sin(\theta - \theta')}{(\alpha^2 + 2\alpha \cos(\theta - \theta') + 1)^{1/2}} \quad (60)$$

with

$$\alpha = s_2 s_1' / t_2 t_1'; \quad \cos \theta = \mathbf{s}_1 \cdot \mathbf{t}_1; \quad \cos \theta' = \mathbf{s}_2 \cdot \mathbf{t}_2.$$

If we label the vectors so that $\alpha > 1$, δ is completely determined. This gives the direction $\mathbf{\Omega}$ for the optimum \mathbf{s}_1' . The actual direction is computed by selecting ω , the cosine of a polar angle with respect to $\mathbf{\Omega}$ as axis. The density function for ω is arranged to have a width determined from the ϱ^{-2} behavior. Specifically, let

$$g = -\varrho^2 \frac{\partial^2}{\partial \delta^2} \varrho^{-2} \Big|_{\partial \varrho / \partial \delta} = 0. \quad (61)$$

For $g \leq 1.25$ make no rotation. Otherwise set

$$\begin{aligned}\alpha &= g/(1 + g), \\ \omega &= 1 - 2(1 - \alpha)\xi/(1 - \alpha\xi)\end{aligned}\tag{62}$$

with ξ a random number. This choice is arbitrary but is easy to carry out and has the general effect required. A weight equal to $d\omega/d\xi$ is attached to the answer when obtained.

An azimuthal angle—here taken to be equidistributed on $(0, 2\pi)$ —completes the specification [12] of the direction of \mathbf{s}_1' . This done, the only remaining variable is a second azimuthal angle, φ , which fixes \mathbf{t}_1' . Again it is a straightforward matter to find the value of φ which minimizes ϱ and to select a value biased toward that direction. Again, the width of the density function for φ is related to $\varrho^2 \partial^2 \varrho^{-2} / \partial \varphi^2$.

These changes improved the estimations of $\int \psi^2 d\mathbf{R}$ and $\int r^2 \psi^2 d\mathbf{R}$ in a substantial way.

B.3 Multiple Pairing

The last scheme aimed at improving the accuracy of the ψ^2 integration was the simplest. It consisted in pairing a point derived from the random walk with several points derived from the independent set of points. The fifty last such points in each set were saved for this purpose. A new point was then compared with each of the saved independent points and if the pair seemed promising for significant results after rotation, the entire process was carried through. That is, let

$$\begin{aligned}g_{sc} &= 0.2 |W_1' W_2'| / [(s_1 - s_2)^2 + (t_1 - t_2)^2], \\ p_{sc} &= \min[g_{sc}, 1],\end{aligned}\tag{63}$$

where W_1' and W_2' are the weights attached to the points. With probability p_{sc} the rotation and scoring process was carried out. The numerical coefficient in g_{sc} was established experimentally to give a very significant improvement in accuracy without a very large increase in computing time. The value used is undoubtedly not optimum.

REFERENCES

1. M. H. KALOS, *Phys. Rev.* **128**, 1791 (1962).
2. C. L. PEKERIS, *Phys. Rev.* **115**, 1216 (1959).
3. C. W. SCHERR, and R. E. KNIGHT, *J. Chem. Phys.* **40**, 3034 (1964).

4. J.-P. AUFRAY, *Compt. Rend., Acad. Sci., Paris* **258**, 3212 (1964).
5. J.-P. AUFRAY, *Compt. Rend., Acad. Sci., Paris* **259**, 2974 (1964).
6. J.-P. AUFRAY, private communication (1965).
7. E. H. WICHMANN and C. H. WOO, *J. Math. Phys.* **2**, 178 (1961).
8. L. HOSTLER, *J. Math. Phys.* **5**, 591 (1964).
9. J. SCHWINGER, *J. Math. Phys.* **5**, 1606 (1964).
10. B. V. GNEDENKO and A. N. KOLMOGOROV, "Limit Distributions for Sums of Independent Random Variables," translated by K. L. Chung. Addison-Wesley, Cambridge, Massachusetts, (1954).
11. G. N. WATSON, "Treatise on the Theory of Bessel Functions," 2nd edition, p. 183. Cambridge Univ. Press, New York (1945).
12. G. GOERTZEL and M. H. KALOS, in "Progress in Nuclear Energy" (D. J. Hughes, J. E. Sunder, and J. Horowitz, eds.), Series I, Vol. 2, page 368. Pergamon Press, New York, (1958).