

# NATIONAL BUREAU OF STANDARDS REPORT

1553

A STATISTICAL METHOD FOR DETERMINING THE LOWEST  
EIGENVALUE OF SCHRÖDINGER'S EQUATION

by

Mark Kac

Cornell University  
and  
National Bureau of Standards

and

Michael Cohen

Cornell University  
and  
National Bureau of Standards



U. S. DEPARTMENT OF COMMERCE  
NATIONAL BUREAU OF STANDARDS

U. S. DEPARTMENT OF COMMERCE

Charles Sawyer, *Secretary*

NATIONAL BUREAU OF STANDARDS

A. V. Astin, *Director*



## THE NATIONAL BUREAU OF STANDARDS

The scope of activities of the National Bureau of Standards is suggested in the following listing of the divisions and sections engaged in technical work. In general, each section is engaged in specialized research, development, and engineering in the field indicated by its title. A brief description of the activities, and of the resultant reports and publications, appears on the inside of the back cover of this report.

**Electricity.** Resistance Measurements. Inductance and Capacitance. Electrical Instruments. Magnetic Measurements. Applied Electricity. Electrochemistry.

**Optics and Metrology.** Photometry and Colorimetry. Optical Instruments. Photographic Technology. Length. Gage.

**Heat and Power.** Temperature Measurements. Thermodynamics. Cryogenics. Engines and Lubrication. Engine Fuels. Cryogenic Engineering.

**Atomic and Radiation Physics.** Spectroscopy. Radiometry. Mass Spectrometry. Solid State Physics. Electron Physics. Atomic Physics. Neutron Measurements. Infrared Spectroscopy. Nuclear Physics. Radioactivity. X-Rays. Betatron. Nucleonic Instrumentation. Radiological Equipment. Atomic Energy Commission Instruments Branch.

**Chemistry.** Organic Coatings. Surface Chemistry. Organic Chemistry. Analytical Chemistry. Inorganic Chemistry. Electrodeposition. Gas Chemistry. Physical Chemistry. Thermochemistry. Spectrochemistry. Pure Substances.

**Mechanics.** Sound. Mechanical Instruments. Aerodynamics. Engineering Mechanics. Hydraulics. Mass. Capacity, Density, and Fluid Meters.

**Organic and Fibrous Materials.** Rubber. Textiles. Paper. Leather. Testing and Specifications. Polymer Structure. Organic Plastics. Dental Research.

**Metallurgy.** Thermal Metallurgy. Chemical Metallurgy. Mechanical Metallurgy. Corrosion.

**Mineral Products.** Porcelain and Pottery. Glass. Refractories. Enameled Metals. Concreting Materials. Constitution and Microstructure. Chemistry of Mineral Products.

**Building Technology.** Structural Engineering. Fire Protection. Heating and Air Conditioning. Floor, Roof, and Wall Coverings. Codes and Specifications.

**Applied Mathematics.** Numerical Analysis. Computation. Statistical Engineering. Machine Development.

**Electronics.** Engineering Electronics. Electron Tubes. Electronic Computers. Electronic Instrumentation.

**Radio Propagation.** Upper Atmosphere Research. Ionospheric Research. Regular Propagation Services. Frequency Utilization Research. Tropospheric Propagation Research. High Frequency Standards. Microwave Standards.

**Ordnance Development.** These three divisions are engaged in a broad program of research and development in advanced ordnance. Activities include basic and applied research, engineering, pilot production, field testing, and evaluation of a wide variety of ordnance matériel. Special skills and facilities of other NBS divisions also contribute to this program. The activity is sponsored by the Department of Defense.

**Missile Development.** Missile research and development: engineering, dynamics, intelligence, instrumentation, evaluation. Combustion in jet engines. These activities are sponsored by the Department of Defense.

● Office of Basic Instrumentation

● Office of Weights and Measures.

# NATIONAL BUREAU OF STANDARDS REPORT

NBS PROJECT

NBS REPORT

1101-11-5100

March 17, 1952

1553

## A STATISTICAL METHOD FOR DETERMINING THE LOWEST EIGENVALUE OF SCHRÖDINGER'S EQUATION\*

by

Mark Kac

Cornell University

and

National Bureau of Standards

and

Michael Cohen

Cornell University

and

National Bureau of Standards

PREPRINT

\*The preparation of this paper was sponsored (in part) by the  
Office of Naval Research



The publication, reprint  
unless permission is ob  
25, D. C. Such permis  
cally prepared if that

Approved for public release by the  
Director of the National Institute of  
Standards and Technology (NIST)  
on October 9, 2015

part, is prohibited  
ndards, Washington  
ort has been specifi-  
ort for its own use.

THE UNIVERSITY OF CHICAGO  
DEPARTMENT OF CHEMISTRY

---

THE UNIVERSITY OF CHICAGO  
DEPARTMENT OF CHEMISTRY  
5708 SOUTH CAMPUS DRIVE  
CHICAGO, ILLINOIS 60637

---

## Table of Contents

	Page
I. Summary . . . . .	1
II. Heuristic outline of the theory . . . . .	1
III. Estimation of the error involved in the method . . . . .	9
IV. Extension to the hydrogen atom . . . . .	16
V. The slope method; results for hydrogen, helium, and "double hydrogen" . . . . .	22
VI. Techniques for speeding the sampling process; high speed computing and importance sampling . . . . .	34
Appendices . . . . .	
A: Computation of discretization error for the harmonic oscillator. . . . .	46
B: Further statistical properties of the random walks . . . . .	52
C: Statistical method for finding the lowest eigenfunction . . . . .	58
Bibliography . . . . .	61



A Statistical Method for Determining the Lowest  
Eigenvalue of Schrödinger's Equation\*

by

Mark Kac

Cornell University  
and  
National Bureau of Standards  
and

Michael Cohen\*\*

Cornell University  
and  
National Bureau of Standards

I. SUMMARY

The experiments discussed here are a continuation of those reported by Donsker and Kac in [2]. Atomic potentials, mainly that of hydrogen, are dealt with here. Improved techniques for accumulating and treating the data are presented; coupled with a high speed digital computer, these techniques should make possible the rapid determination, to within a few percent, of the lowest eigenvalue of many potentials. Considerable attention is given to the errors in the method, and an elementary account of the underlying theory is included.

II. HEURISTIC OUTLINE OF THE THEORY

With the introduction of an imaginary time variable, the time-dependent Schrödinger equation for a single particle in a one-dimensional

---

\*The preparation of this paper was sponsored (in part) by the Office of Naval Research.

\*\*Now at the California Institute of Technology





potential field  $V(x)$  can be brought into the form

$$(1) \quad \frac{\partial P(x,t)}{\partial t} = \frac{1}{2} \frac{\partial^2 P(x,t)}{\partial x^2} - V(x)P(x,t) \quad .$$

We shall shortly see that this form of the equation admits of a simple probabilistic interpretation. There is some instructive value in the probability approach, inasmuch as it adds to our feeling for the meaning of the wave-function and the equation. Furthermore, the probability model for equation (1) yields a new computational method for finding the lowest eigenvalue of the time-independent Schrödinger equation. We shall concentrate our attention mainly on this computational method.

In order to find the probabilistic meaning of (1), let us consider the problem of a single particle performing a random walk in one dimension subject to the following conditions:

- A) At equispaced time intervals  $\tau$  the particle takes a step of length  $h$  either to the left or the right, with equal probabilities.
- B) If the particle stays at a given point  $x$  during a time interval  $dt$ , it is subject to destruction there with probability  $V(x)dt + o(dt)$ . (We now restrict ourselves to the case  $V(x) \geq 0$ , but shall later drop the restriction.)

If we let  $hP(x,t)$  represent the probability that the particle is at the point  $x$  at time  $t$  ( $P(x,t)$  is thus a probability "density"), clearly  $P(x,t)$  must satisfy the difference equation



$$P(x,t) = \frac{1}{2}P(x-h, t-\tau)(1 - V(x-h)\tau + o(\tau))$$

$$+ \frac{1}{2}P(x+h, t-\tau)(1 - V(x+h)\tau + o(\tau))$$

(We have merely decomposed the event of being at the point  $x$  at time  $t$  into the two alternative ways in which it could have occurred.) Subtracting  $P(x,t-\tau)$  from both sides and dividing throughout by  $\tau$ , we obtain

$$\frac{P(x,t)-P(x,t-\tau)}{\tau} = \frac{1}{2} \frac{P(x-h,t-\tau)-2P(x,t-\tau)+P(x+h,t-\tau)}{\tau} - \frac{1}{2} P(x-h, t-\tau)V(x-h) - \frac{1}{2} P(x+h, t-\tau)V(x+h) + \frac{o(\tau)}{\tau} .$$

The left side, of course, is a finite difference approximation to  $\frac{\partial P}{\partial t}$ . If we now require that the space and time increments in the random walk be related by the equation  $\frac{h^2}{\tau} = \tau$ , then the first term on the right is a second difference approximation to  $\frac{\partial^2 P}{\partial x^2}$ . Then, in the limit as  $h \rightarrow 0$  (and hence  $\tau \rightarrow 0$ ), the difference equation approaches the differential equation (1).

Thus, equation (1) is the limiting form of the difference equation describing a random walk with destruction, where  $P(x,t)dx$  is the probability that the particle is between  $x$  and  $x+dx$  at time  $t$ . Equivalently, the equation describes a diffusion process with destruction, where  $P(x,t)$  is the particle density at point  $x$  at time  $t$ , and  $V(x)P(x,t)dx$  is the time rate of destruction of particles in the interval  $(x \text{ to } x+dx)$ .

Under the former interpretation, if we start the particle from the origin, we must require that our solution of equation (1) satisfy the subsidiary condition



(1a)

$$\text{for all } t \geq 0, \quad \int_{-\infty}^{\infty} P(x,t) dx = 1$$

and the initial condition

$$\text{As } t \rightarrow 0^+, \quad P(x,t) \rightarrow \delta(x) .$$

(1b)

$$(\delta(x) = 0 \text{ if } x \neq 0 \quad . \quad \int_{-\infty}^{\infty} \delta(x) dx = 1)$$

If  $V(x) \rightarrow \infty$  as  $x \rightarrow \pm \infty$  (we shall later relax this restriction), the solution of (1) can be expanded in a series

$$P(x,t) = \sum_{j=1}^{\infty} c_j e^{-\lambda_j t} \psi_j(x)$$

where the  $\lambda_j (> 0)$  and  $\psi_j(x)$  are the eigenvalues and normalized eigenfunctions of

$$(2) \quad -\lambda \psi(x) = \frac{1}{2} \frac{d^2 \psi(x)}{dx^2} - V(x) \psi(x)$$

subject to the conditions

$$\left| \int_{-\infty}^{\infty} \psi(x) dx \right| < \infty, \quad \int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty$$

and the  $c_j$  are so chosen as to satisfy the initial condition 1b. (If we look for a solution of (1) in the form of a superposition of functions of the type  $\psi(x)T(t)$ , we find that  $T(t) = e^{-\lambda t}$  and that the space-dependent part  $\psi$  satisfies equation (2).) The eigenfunctions  $\psi_j(x)$  form an orthonormal set, and the initial condition (1b) yields  $c_j = \psi_j(0)$ .



We note here that (2) is Schrödinger's equation for stationary states, where  $\lambda$  represents  $E$ , the energy of the state.

Returning to the limiting form of the random walk problem, we now have the result that the probability of survival of the particle till time  $t$  is

$$(3) \quad \int_{-\infty}^{\infty} P(x,t) dx = \sum_{j=1}^{\infty} e^{-\lambda_j t} \Psi_j(0) \int_{-\infty}^{\infty} \Psi_j(x) dx \equiv \sum_{j=1}^{\infty} b_j e^{-\lambda_j t} .$$

If we arrange the  $\lambda$ 's in order of increasing magnitude, the first term of the series will dominate as  $t \rightarrow \infty$  and hence

$$(4) \quad \lambda_1 = - \lim_{t \rightarrow \infty} \frac{\log \int_{-\infty}^{\infty} P(x,t) dx}{t} .$$

We shall employ the original discrete random walk to obtain an estimate of  $\int_{-\infty}^{\infty} P(x,t) dx$  for large finite  $t$ . Then equation (4) enables us to estimate  $\lambda_1$ .

Returning to the discrete random walk, we inquire: what is the probability  $p(x,\theta)$  that a particle survives the time interval  $\theta$  ( $\theta \leq \tau$ ), given that the particle spends the entire interval at the point  $x$ ? It follows directly from our description of the random walk that this probability must satisfy the differential equation

$$\frac{\partial p(x,\theta)}{\partial \theta} = -V(x)p(x,\theta) .$$

Hence

$$p(x,\theta) = e^{-V(x)\theta} .$$

Now, suppose a particle performs the discrete random walk, starting from the origin and taking steps of length  $h = \frac{1}{\sqrt{n}}$  at time intervals





$\tau = \frac{1}{n}$ . Given that the particle follows a particular path  $x(k\tau)$ , the probability that it survives till time  $t (= n\tau)$  on this path is

$$(4a) \quad \prod_{k=0}^{nt-1} p(x(k\tau), \tau) = \exp \left\{ - \sum_{k=0}^{nt-1} V(x(k\tau)) \tau \right\} \\ = \exp \left\{ - \sum_{k=0}^{nt-1} V(x(\frac{k}{n})) \frac{1}{n} \right\} .$$

From the definition of the random walk, it follows that  $x(\frac{k}{n}) = S_k \sqrt{\frac{1}{n}}$ , where  $S_k$  is the sum of  $k$  random variables  $X_1, X_2, \dots, X_k$ , each  $X_i$  taking the value 1 or -1 with equal probabilities.

We have computed the probability of survival till time  $t$  on a particular path. The probability of survival till time  $t$ , irrespective of path, is the average of (4a) taken over all  $2^{nt-1}$  possible paths  $x(k\tau)$ . Since all paths are equally likely, the average we are interested in is the simple arithmetic average. We symbolize this average by

$$(4b) \quad \left\langle \exp \left\{ - \frac{1}{n} \sum_{k=0}^{nt-1} V\left(\frac{S_k}{\sqrt{n}}\right) \right\} \right\rangle .$$

Then, for large  $n$ , we expect

$$(4c) \quad \left\langle \exp \left\{ - \frac{1}{n} \sum_{k=0}^{nt-1} V\left(\frac{S_k}{\sqrt{n}}\right) \right\} \right\rangle \xrightarrow{n \rightarrow \infty} \int_{-\infty}^{\infty} P(x, t) dx .$$

Thus, by equations (4) and (4c), we can estimate  $\lambda_1$  by

$$(5) \quad \lambda_1 \sim \frac{-\log \left\langle \exp \left\{ - \frac{1}{n} \sum_{k=0}^{nt-1} V\left(\frac{S_k}{\sqrt{n}}\right) \right\} \right\rangle}{t} .$$

It is not necessary actually to compute the average (4b) over all possible paths. We can get a statistical estimate of the average by



considering the average over the paths  $m$  independent random walks.

In equation (5) we have now derived the main result which we shall use for computational purposes. The next few paragraphs are intended to render precise some of our remarks in later sections.<sup>1</sup>

It follows from the central limit theorem of probability that in the limit as  $n \rightarrow \infty$  and  $k \rightarrow \infty$ , the distribution of  $S_k/\sqrt{n}$  becomes Gaussian with mean 0 and variance  $\frac{k}{n}$ ,

$$\text{i.e. Prob} \left\{ x < \frac{S_k}{\sqrt{n}} < x + dx \right\} \rightarrow \frac{1}{\sqrt{2\pi k/n}} \exp \left\{ -\frac{x^2}{2k/n} \right\} dx .$$

Then, as  $n \rightarrow \infty$ , the distribution function of the random variable  $\frac{1}{n} \sum_{k \leq nt} V(S_k/\sqrt{n})$  approaches that of the random variable  $\int_0^t V(x(\tau)) d\tau$ ; here the path  $x(\tau)$  is chosen from a space in which the measure of a set of paths is the probability that a particle performing a Gaussian random walk will follow a path in that set. (By a Gaussian random walk we mean a continuous random walk, with independent increments, with the displacement  $x(\tau)$  at time  $\tau$  governed by the law

$$\text{Prob} \left\{ y < x(\tau) < y + dy \right\} = \frac{1}{\sqrt{2\pi\tau}} \left\langle \exp \left\{ -y^2/2\tau \right\} dy \right\rangle .)$$

Thus, as  $n \rightarrow \infty$ ,

$$(5a) \left\langle \exp \left\{ -\frac{1}{n} \sum_{k \leq nt} V(S_k/\sqrt{n}) \right\} \right\rangle \rightarrow \left\langle \exp \left\{ -\int_0^t V(x(\tau)) d\tau \right\} \right\rangle$$

where the latter average represents the average value of  $\exp \left\{ -\int_0^t V(x(\tau)) d\tau \right\}$ , taken over the space of continuous paths with the

---

<sup>1</sup>For the rigorous development of the theory, see [1].



measure described above. This average is just the Wiener integral of the functional  $\exp \left\{ - \int_0^t V(x(\tau)) d\tau \right\}$ , and the left side of (5a) may be thought of as the analogue of a Riemann sum approximating the Wiener integral. By (4c) and (5a)

$$(6) \quad \left\langle \exp \left\{ - \int_0^t V(x(\tau)) d\tau \right\} \right\rangle = \int_{-\infty}^{\infty} P(x, t) dx \quad .$$

For the purpose of computing  $\lambda_1$ , it would clearly be better to get a sampling estimate of the right side of equation (5a) than of the left side. However, in our actual sampling process we cannot deal with the space of continuous paths  $x(\tau)$ . We can, nevertheless, speed the convergence of the distribution of  $\frac{1}{n} \sum_{k < nt} V(S_k/\sqrt{n})$  to that of  $\int_0^t V(x(\tau)) d\tau$  by taking  $S_k/\sqrt{n}$  directly as a Gaussian variable with mean 0 and variance  $k/n$ , rather than waiting for  $k$  and  $n$  to become large enough for the central limit theorem to apply. In practice, we achieve this by letting  $S_k = X_1 + \dots + X_k$ , where each  $X_i$  is Gaussian (rather than discrete) with mean 0 and variance 1.

Returning to equation (5), we can use an IBM computer to find

$$\left\langle \exp \left\{ - \frac{1}{n} \sum_{k < nt} V(S_k/\sqrt{n}) \right\} \right\rangle, \text{ and thus } \lambda_1, \text{ as follows:}$$

- 1) Pick values for  $n$  and  $t$  (errors involved in choosing finite  $n$  and  $t$  will be discussed below);
- 2) Choose a random number  $X_1$  from a card containing Gaussian deviates with mean 0 and variance 1;
- 3) Compute  $V\left(\frac{X_1}{\sqrt{n}}\right)$ ;
- 4) Choose a random number  $X_2$ ;



5) Compute  $V\left(\frac{x_1+x_2}{\sqrt{n}}\right);$

•  
•  
•

Choose a random number  $X_k;$

Compute  $V\left(\frac{x_1+\dots+x_k}{\sqrt{n}}\right);$

Continue as long as  $k \leq nt;$

Compute  $\frac{1}{n} \left\{ V\left(\frac{x_1}{\sqrt{n}}\right) + V\left(\frac{x_1+x_2}{\sqrt{n}}\right) + \dots + V\left(\frac{x_1+x_2+\dots+x_{nt-1}}{\sqrt{n}}\right) \right\} = A$

Compute  $e^{-A}.$

We have now found one term in the average  $\left\langle \exp \left\{ -\frac{1}{n} \sum_{k=1}^{nt} V\left(\frac{S_k}{\sqrt{n}}\right) \right\} \right\rangle,$  namely the term corresponding to the path determined by our particular choice of the  $X_k$ 's. Each choice of a sequence  $X_1, X_2, \dots, X_{nt}$  corresponds to a random walk; and if we take many random walks and average  $e^{-A}$  for all the walks, we should converge to an accurate value of

$$\left\langle \exp \left\{ -\frac{1}{n} \sum_{k=1}^{nt} V\left(\frac{S_k}{\sqrt{n}}\right) \right\} \right\rangle.$$

As an initial test of the method, the potentials  $V(x) = |x|$  and  $V(x) = x^2$  (harmonic oscillator) were studied by means of 100 random walks with  $n = 100, t_1 = 3.75, t_2 = 5.$ <sup>1</sup> The experimental values for  $\lambda_1$  were 0.81 and 0.75, compared with theoretical values 0.81 and 0.71 respectively.

### III. ESTIMATION OF THE ERROR INVOLVED IN THE METHOD

The feasibility of the method for the calculation of eigenvalues of complicated potentials depends, of course, on the accuracy which can be

---

<sup>1</sup>To minimize a numerical error discussed in Section III it proves desirable to consider two different values of  $t$ . For a detailed account of the experiments on these potentials, see [2].





obtained in a reasonable number of random walks. There seem to be three major possible sources of error:

- A) Error involved in discretizing the problem and sampling the values of the finite sum  $\frac{1}{n} \sum_{k=1}^n V\left(\frac{S_k}{\sqrt{n}}\right)$  instead of  $\int_0^t V(x(\tau)) d\tau$ .  
 i.e.  $\left\langle \exp \left\{ -\frac{1}{n} \sum_{k=1}^n V\left(\frac{S_k}{\sqrt{n}}\right) \right\} \right\rangle \neq \left\langle \exp \left\{ -\int_0^t V(x(\tau)) d\tau \right\} \right\rangle$ .
- B) Error due to using equation (4) for finite  $t$ .
- C) Sampling error, due to the fact that we sample only  $m$  paths out of the space of all possible paths.

Errors B and C, which we shall consider first, are related. It will be seen in the consideration of statistical errors that it is impractical to make  $t$  very large. Then the numerical error involved in using equation (4) for finite  $t$  is twofold:

- 1) We neglect the term  $\frac{\log b_1}{t}$  (see eqn.(3)). This error, which may be serious for moderate  $t$ , is easily removed by considering two different times  $t_1$  and  $t_2$ . Then it follows trivially from equations (3) and (6) that

$$(7) \quad \lambda_1 \sim \frac{1}{t_2 - t_1} \log \frac{\left\langle \exp \left\{ -\int_0^{t_1} V(x(\tau)) d\tau \right\} \right\rangle}{\left\langle \exp \left\{ -\int_0^{t_2} V(x(\tau)) d\tau \right\} \right\rangle}$$

and that the influence of  $b_1$  in the ratio becomes negligible even for moderate  $t$ .

- 2) In the expansion of  $\int_{-\infty}^{\infty} P(x,t) dx$  we omit the higher exponentials in  $t$  (see eqn.(3)). Ignoring the ration  $b_1/b_2$ , we see that the



fractional error caused by this omission is approximately  $e^{-(\lambda_2 - \lambda_1)t}$ . For the harmonic oscillator,  $\lambda_j = \frac{2j-1}{\sqrt{2}}$ , and hence the fractional error in estimating the series by its first term is about  $e^{-1.4t}$ . For  $t = 4$ , this is less than 0.5%.

The sampling error C arises from the fact that we try to approximate the exact average  $\left\langle \exp \left\{ - \int_0^t V(x(\tau)) d\tau \right\} \right\rangle$  (henceforth called  $\langle \alpha \rangle$ ) by the average  $\frac{1}{m} \sum_{k=1}^m \exp \left\{ - \int_0^t V(x_k(\tau)) d\tau \right\}$  taken over only  $m$  independent paths  $x_1(\tau), \dots, x_m(\tau)$ .<sup>1</sup> The expected square of the deviation from  $\langle \alpha \rangle$  of a single value  $\exp \left\{ - \int_0^t V(x_k(\tau)) d\tau \right\}$  is given by  $\langle \alpha^2 \rangle - \langle \alpha \rangle^2$ , the usual expression for the mean square deviation of a single observation from the mean. Then the mean square deviation from  $\langle \alpha \rangle$  of the average of  $m$  observations is  $\frac{1}{m} (\langle \alpha^2 \rangle - \langle \alpha \rangle^2)$ . Thus, the relative fluctuation of the average of  $m$  observations is given by

$$\frac{1}{\sqrt{m}} \frac{\sqrt{\langle \alpha^2 \rangle - \langle \alpha \rangle^2}}{\langle \alpha \rangle} = \frac{1}{\sqrt{m}} \sqrt{\frac{\langle \alpha^2 \rangle}{\langle \alpha \rangle^2} - 1} .$$

Fortunately, we are able to estimate  $\langle \alpha^2 \rangle$ , for

$$\langle \alpha^2 \rangle = \left\langle \exp \left\{ - 2 \int_0^t V(x(\tau)) d\tau \right\} \right\rangle .$$

Then, applying the results of equations (3) and (6) to the potential  $2V$ ,

---

<sup>1</sup>At this point we are ignoring the discretization error A. Otherwise, we should properly replace  $\int_0^t V(x_k(\tau)) d\tau$  by a Riemann sum.



we have

$$\langle \alpha^2 \rangle \sim d_1 e^{-\mu_1 t}$$

where  $\mu_1$  is the lowest eigenvalue of the potential  $2V$ ;  $\varphi_1(x)$  is the lowest eigenfunction of the potential  $2V$ ;  $d_1 = \varphi_1(0) \int_{-\infty}^{\infty} \varphi_1(x) dx$ .

The relative fluctuation in the mean is thus

$$(8) \quad \frac{1}{\sqrt{m}} \sqrt{\frac{d_1 e^{-\mu_1 t}}{b_1^2 e^{-2\lambda_1 t}} - 1}$$

which, for large  $t$ , is of the order  $\frac{1}{\sqrt{m}} e^{(2\lambda_1 - \mu_1)t/2}$ . For any positive potential  $V$ ,  $2\lambda_1 - \mu_1 \geq 0$ ,<sup>1</sup> and hence  $t$  must be kept relatively small, even at the price of an increase in error  $B$ .

In the case  $V(x) = x^2$ , we know  $\lambda_1 = \frac{\sqrt{2}}{2}$ ; a simple change of scale yields  $\mu_1 = 1$ . Then the relative fluctuation in  $\langle \alpha \rangle$  when  $m = 100$ ,  $t = 5$  is approximately

$$\frac{1}{10} e^{(\sqrt{2}-1) 5/2} \sim 30\%$$

This result is not so disastrous as it looks, since the method of calculation involves the logarithm of  $\langle \alpha \rangle$  and is not so sensitive to errors in

<sup>1</sup>This fact is easily proved by the variational method. On the other hand, one can convert our reasoning into a probabilistic "proof" that  $2\lambda_1 - \mu_1 \geq 0$  as follows: if  $\mu_1 > 2\lambda_1$ , then for sufficiently large  $t$  we should have  $e^{-\mu_1 t} \ll e^{-2\lambda_1 t}$  and thus  $\langle \alpha^2 \rangle - \langle \alpha \rangle^2 < 0$ ; but this would imply that for sufficiently large  $t$ , the mean square deviation of a single observation of  $\alpha$  from the mean is negative. This is impossible.



$\langle \alpha \rangle$ . Equation (5) says  $\lambda_2 \sim \frac{1}{t} \log \langle \alpha \rangle$ . Then

$$(9) \quad \left| \frac{d\lambda}{\lambda} \right| \sim \left| \frac{1}{t} \frac{d \langle \alpha \rangle}{\langle \alpha \rangle} \right|$$

For  $t = 5$ , the expected percentage error in our determination of  $\langle \alpha \rangle$  for the harmonic oscillator is thus

$$\left| \frac{d\lambda}{\lambda} \right| \sim \frac{1}{5(.71)} 30\% \sim 9\%$$

The observed relative error in  $\lambda_1$  for this case was about 5%.

The error A is not easy to estimate. What we should like to find is the relative error

$$K_V \equiv \left| \frac{\langle \exp \left\{ - \int_0^t V(x(\tau)) d\tau \right\} \rangle - \langle \exp \left\{ - \frac{1}{n} \sum_{k=1}^n V(S_k/\sqrt{n}) \right\} \rangle}{\langle \exp \left\{ - \int_0^t V(x(\tau)) d\tau \right\} \rangle} \right|$$

A method for estimating  $K_V$  is proposed here, but not rigorously justified.

In the case of the harmonic oscillator, however, we are able to calculate  $K_x^2$  exactly (Appendix A). The answer obtained for  $K_x^2$  agrees with that predicted by our more general method. We shall discuss the method briefly; it is worth noting that in practice the error  $K_V$  seems unimportant.

With a given continuous path  $x(\tau)$  we can associate another path  $x_n(\tau)$ , which is a step-function approximation to  $x(\tau)$ . We define

$$x_n(\tau) = x\left(\frac{k}{n}\right) \quad \text{when } \frac{k}{n} \leq \tau < \frac{k+1}{n}$$

$$\text{Then } \left\langle \exp \left\{ - \frac{1}{n} \sum_{k=1}^n V(S_k/\sqrt{n}) \right\} \right\rangle = \left\langle \exp \left\{ - \int_0^t V(x_n(\tau)) d\tau \right\} \right\rangle$$





By definition  $x_n(\tau) = x(\frac{1}{n} [n\tau])^1$  and thus both  $x(\tau)$  and  $x_n(\tau)$  are Gaussian random variables. If we plot the time  $\tau$  versus the difference  $(\tau - \frac{1}{n}[n\tau])$  of the time arguments of the two random variables, we get a sawtooth function which grows linearly from 0 to  $1/n$  with period  $1/n$ . On the average, the step function lags behind the continuous function in time by an amount  $1/2n$ . Thus, with respect to functionals which depend on the behavior of  $x_n(\tau)$  over periods of time large compared to  $1/n$ , we may replace  $x_n(\tau)$  by  $x(\tau - \frac{1}{2n})$ . (This might be invalid if  $V$  is very rapidly varying, but seems permissible for  $V(x) = x^2$ . In fact, it is easily verified that this scheme gives  $\langle \int_0^t x_n^k(\tau) d\tau \rangle$  correctly, to the first order in  $1/n$ , for any  $k$ .)

We therefore conjecture that, to the first order in  $1/n$ ,

$$\left\langle \exp \left\{ - \int_0^t V(x_n(\tau)) d\tau \right\} \right\rangle = \left\langle \exp \left\{ - \int_{1/2n}^t V(x(\tau - \frac{1}{2n})) d\tau \right\} \right\rangle$$

But the latter quantity is the same as  $\left\langle \exp \left\{ - \int_0^{t-\frac{1}{2n}} V(x(\tau)) d\tau \right\} \right\rangle$ .

By equations (3) and (6), for large  $t$

$$\left\langle \exp \left\{ - \int_0^{t-\frac{1}{2n}} V(x(\tau)) d\tau \right\} \right\rangle \approx e^{-\lambda_1(t-\frac{1}{2n})}$$

Thus

$$K_V \approx \left| \frac{e^{-\lambda_1 t} - e^{-\lambda_1(t-\frac{1}{2n})}}{e^{-\lambda_1 t}} \right| \approx \frac{\lambda_1}{2n},$$

to the first order in  $1/n$ . For  $n = 100$ , this error is negligible.

---

<sup>1</sup>[y] = the greatest integer less than or equal to  $y$ .



Our estimate of the error  $K_V$  is considerably less than that of Fortet [3], who proves rigorously for the potential  $V(x) = x^2$ , that

$$K_{x^2} \leq \frac{2}{3} b_1 e^{\lambda_1 t} \frac{t^{3/2}}{\sqrt{n}},$$

and regards the right side as a good estimate of the left. Fortet's estimate follows directly from his preliminary proof that

$$\left| \left\langle \exp \left\{ - \int_0^t x^2(\tau) d\tau \right\} \right\rangle - \left\langle \exp \left\{ - \int_0^t x_n^2(\tau) d\tau \right\} \right\rangle \right| \leq \frac{2}{3} \frac{t^{3/2}}{\sqrt{n}}.$$

This result is true, but the chain of the inequalities used in proving it involves some very wasteful inequalities, which operate to make the right side a very high bound on the error. The highness of the bound is most easily seen by considering the case of fixed  $n$  and large  $t$ . Then the right side approaches  $+\infty$  as  $t \rightarrow \infty$ , while clearly the left side approaches 0 (since each of the terms on the left goes to 0).

We prove in Appendix A that  $K_{x^2} \approx \frac{1}{2\sqrt{2}n}$ , correct to the first order in  $\frac{1}{n}$ . (This result agrees with our conjecture). For the interesting case of  $n = 100$ ,  $t = 5$ , we thus have

$$K_{x^2} \sim \frac{1}{283} \sim 0.3\%.$$

If we take Fortet's bound on  $K_{x^2}$  as an estimate, we obtain for  $t = 5$ ,  $n = 100$

$$K_{x^2} \sim 16 \sim 1600\%$$

Thus the latter estimate appears too high by a factor of 5000 and, if correct, would indicate that the method is totally useless.



IV. EXTENSION TO THE HYDROGEN ATOM

The harmonic oscillator, of course, represents one of the simplest potentials; the results obtained in the study of the hydrogen potential are of more significance in determining the real value of the method. In Cartesian coordinates and in the proper units, Schrödinger's equation for the hydrogen atom is

$$\frac{1}{2} \left( \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} \right) + \frac{1}{\sqrt{x^2 + y^2 + z^2}} \Psi = - \lambda \Psi .$$

We encounter here the following complicating features, which were not present in the case of the harmonic oscillator:

1.  $V(r) = -\frac{1}{r} = -\frac{1}{\sqrt{x^2 + y^2 + z^2}}$  depends on three coordinates. Hence the random walks must be performed in three dimensions rather than one.
2.  $V$  is negative, has a singularity at the origin, and does not approach  $\infty$  as  $r \rightarrow \infty$  (If we want to continue thinking in terms of random walks, we should now think of  $V$  as representing a probability of multiplication, rather than destruction).
3. The spectrum of eigenvalues is partly discrete and partly continuous, the continuous spectrum arising from the fact that  $V$  does not approach  $\infty$  as  $r \rightarrow \infty$ . The eigenvalues are now  $\lambda_n = -\frac{1}{n^2}$  ( $n = 1, 2, 3, \dots$ ) and the entire positive real axis.

Results analogous to those of the one-dimensional case can still be obtained. Corresponding to equation (3) we have



$$\left\langle e^{\int_0^t \frac{d\tau}{r(\tau)}} \right\rangle = \sum_{j=1}^{\infty} e^{-\lambda_j t} \Psi_j(0,0,0) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi_j(x,y,z) dx dy dz + \int_0^{\infty} e^{-\lambda t} C(\lambda) d\lambda$$

where the integral is the contribution of the continuous spectrum. For convenience, we redefine  $\lambda_n$  as  $\frac{1}{2} \frac{1}{n^2}$  and can now write

$$\left\langle e^{\int_0^t \frac{d\tau}{r(\tau)}} \right\rangle \sim C_1 e^{\lambda_1 t}$$

where

$$C_1 = \Psi_1(0,0,0) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi_1(x,y,z) dx dy dz$$

The replacement of  $\int_0^t V(r(\tau)) d\tau$  by a finite sum proceeds as before,

with  $\sqrt{\frac{S_{kx}^2 + S_{ky}^2 + S_{kz}^2}{n}}$  substituted for  $\frac{S_k}{\sqrt{n}}$ . Here  $S_{kx}$  is the sum of  $k$  independent and identically distributed random variables  $X_1, \dots, X_k$ , each with mean zero and variance one (similarly  $S_{ky}$  and  $S_{kz}$ ). Thus the three-dimensional path is determined by three independent one-dimensional random walks.

Then  $-\int_0^t V(r(\tau)) d\tau = \int_0^t \frac{d\tau}{r(\tau)}$  is approximated by

$$\frac{1}{n} k \sum_{k=1}^n t \frac{1}{\sqrt{\frac{S_{kx}^2 + S_{ky}^2 + S_{kz}^2}{n}}} = \frac{1}{\sqrt{n}} k \sum_{k=1}^n t \frac{1}{\sqrt{S_{kx}^2 + S_{ky}^2 + S_{kz}^2}}$$

As before, we approximate  $\lambda_1$  by the relation





$$\lambda_1 \sim \frac{1}{t_2 - t_1} \log \frac{\left\langle e^{\int_0^{t_2} \frac{1}{r(\tau)} d\tau} \right\rangle}{\left\langle e^{\int_0^{t_1} \frac{1}{r(\tau)} d\tau} \right\rangle}$$

Again we use two different values of  $t$  to eliminate the error B.

In practice, some sampling time is saved by the following scheme. We note that  $x(\tau)$ , being Gaussian with mean 0 and variance  $\tau$ , has the same distribution as  $\sqrt{\tau} x(\frac{\tau}{t})$ . Therefore  $r(\tau)$  has the same distribution as  $\sqrt{\tau} r(\frac{\tau}{t})$ . It follows that  $\int_0^t \frac{1}{r(\tau)} d\tau$  has the same distribution as  $\sqrt{t} \int_0^1 \frac{1}{r(\tau)} d\tau$ . Then, what we actually sample is the random variable

$$\alpha = \int_0^1 \frac{1}{r(\tau)} d\tau \sim .1 \frac{\sum_{k=1}^{100} 1}{\sqrt{S_{kx}^2 + S_{ky}^2 + S_{kz}^2}} .$$

We then compute  $\lambda_1$  by the relation

$$(10) \quad \lambda_1 \sim \frac{1}{t_2 - t_1} \log \frac{\left\langle e^{\int_0^{t_2} \frac{1}{r(\tau)} d\tau} \right\rangle}{\left\langle e^{\int_0^{t_1} \frac{1}{r(\tau)} d\tau} \right\rangle} \sim \frac{1}{t_2 - t_1} \log \frac{\left\langle e^{\sqrt{t_2} \alpha} \right\rangle}{\left\langle e^{\sqrt{t_1} \alpha} \right\rangle} .$$

The values of  $t$  used must be large enough to make error B (discussed in section III) small. The magnitude of this error, as we have already seen, is about  $e^{-(\lambda_1 - \lambda_2)t}$  (the sign of the exponent differs from that in the similar expression for the harmonic oscillator because the solutions now increase, rather than decrease, exponentially). For the hydrogen atom,  $\lambda_1 - \lambda_2 = 0.25$ . The corresponding quantity for the harmonic oscillator is 1.4. When  $t_1 = 9$ , the fractional error in



$\left\langle e^{\int_0^{t_1} \frac{1}{r(\tau)} d\tau} \right\rangle$  due to B is thus about 10%, and the consequent fractional error in  $\lambda_1$  about 2%. For  $t_2 = 16$ , the accuracy is better.

What happens to the sampling error C when we use such large values of  $t$ ? If we sample  $m$  paths, we recall that the relative fluctuation of the average  $\frac{1}{m} \sum_{k=1}^m e^{\int_0^t \frac{1}{r_k(\tau)} d\tau}$  is approximately  $\frac{1}{\sqrt{m}} e^{(\mu_1 - 2\lambda_1)t/2}$ ;  $\mu_1$ , the lowest eigenvalue of the potential  $-\frac{2}{r}$ , is easily calculated as 2. Thus  $\mu_1 - 2\lambda_1 = 1$ . The corresponding term in the statistical error computation for the harmonic oscillator was .414. Hence the statistical fluctuation for a given number of samples  $m$  and given  $t$  is considerably larger for the potential  $-\frac{1}{r}$  than for the potential  $x^2$ .

Intuitively, one can explain this increase in the statistical error in the following manner: The eigenvalue of the hydrogen potential depends mainly on the behavior of the potential in the small region near the origin. We see this fact immediately when we note that the main contribution to the average  $\left\langle e^{\int_0^t \frac{1}{r(\tau)} d\tau} \right\rangle$  comes from the few paths  $r(\tau)$  which stay close to the origin. Hence, in order to calculate the eigenvalue accurately we must get a detailed account of the behavior of the potential in the critical small region about the origin. Naturally, many samples are required before we get a sufficient number of paths staying close to the origin to sample that region accurately. Equivalently, if we do not take a very large number of samples, there is a good chance that the few critical paths which stay near the origin will not turn up in the proper frequency, thus causing a large error in the eigenvalue. In the case of the harmonic



oscillator, however, there is no such small critical region which must be carefully sampled. Remarks similar to these about hydrogen may also be made about atomic potentials in general.

Returning to our computations, we see that for 900 samples with  $t = 9$  the expected relative fluctuation of  $\frac{1}{m} \sum_{k=1}^m e^{\int_0^t \frac{1}{r_k(\tau)} d\tau}$  is about

$$\frac{1}{\sqrt{900}} e^{9/2} \sim \frac{100}{30} \sim 300\% .$$

If we compute more carefully, using the exact expression for the relative fluctuation given in equation (8), we obtain 100% instead of 300%. It follows that for  $m = 100$  and  $t = 9$ ,

$$\frac{d\lambda}{\lambda} \sim \frac{1}{t\lambda} \frac{d\langle\alpha\rangle}{\langle\alpha\rangle} \sim \frac{1}{9 \times .5} 100\% \sim 22\% .$$

For  $t = 16$ , the error is of course much greater. For the initial group of 900 hydrogen random walks, with  $t_1 = 9$  and  $t_2 = 16$ , computations according to equation (10) yielded

$$\lambda_1 = .503 \quad (\text{true value } .500) .$$

In view of the preceding remarks, this accuracy was quite surprising. An examination of the data, however, seems to confirm our previous analysis. Of the 900 paths sampled, 4 yielded extremely large values of  $\int_0^1 \frac{1}{r(\tau)} d\tau$ ; without these 4 paths, which clearly cannot be counted on to recur regularly, the average  $\langle e^{4\alpha} \rangle$  would have been 85% smaller, and the computed value for  $\lambda_1$  about 20% smaller (note that the logarithmic method of calculation is not too sensitive even to a large fluctuation in  $\langle e^{4\alpha} \rangle$ ).



Thus, there seem to be good grounds for attributing the high accuracy in this determination of  $\lambda_1$  to luck. Nevertheless, it seemed worthwhile to run another group of samples. This was easily done in connection with experiments on helium (described below), since each helium random walk yields two hydrogen random walks as a by-product. Two more groups of 900 hydrogen walks were thus obtained.

The qualitative appearance of the data was similar to that of the first group of walks. In the second group of 900, more than 80% of the contribution to  $\langle e^{4\alpha} \rangle$  came from three walks yielding high values of  $\alpha$ , and the calculated value of  $\lambda_1$  was 0.36. If a few more high values of  $\alpha$  had been recorded, the calculated value of  $\lambda_1$  would have been much closer to 0.50; but this again would have been purely a matter of luck. In the third group, there were only two large values of  $\alpha$ , with the result that the low value  $\lambda_1 = 0.29$  was obtained.

A more detailed analysis of some aspects of the hydrogen random walks is given in section V and in Appendix B.

The preceding discussion of hydrogen, which may be extended to other atomic potentials, makes one fact rather clear; for  $t$  great enough to make our numerical approximations good, the statistical fluctuation in

$\left\langle e^{\int_0^t \frac{d\tau}{r(\tau)}} \right\rangle$  becomes prohibitively large unless we use a huge number of samples. Equivalently, our computed value for  $\lambda_1$  is determined almost entirely by a very small group out of the 900 random walks. Hence, most of the data is wasted and our large group of samples acts statistically like a small group.





There now arises the question of whether there is any other method of utilizing our data to find  $\lambda_1$ , by means of some computation less subject to statistical fluctuations. In the next section we shall deal with such a method, which we call the slope method, as distinguished from our previous mean value method.

V. THE SLOPE METHOD; RESULTS FOR  
HYDROGEN, HELIUM, AND "DOUBLE HYDROGEN"

The method for finding  $\lambda_1$  which is to be presented here has the following advantages over the mean value method:

- 1) It makes use of the finer structure of the distribution of the random variable  $\int_0^t \frac{d\tau}{r(\tau)}$  and
- 2) It is quite insensitive to erratic individual observations of the random variable.

Consider the random variable  $\int_0^t \frac{d\tau}{r(\tau)}$ . The value which this random variable assumes depends on the particular path  $r(\tau)$  which is chosen for

the path of integration. We have already observed (section IV) that

$\int_0^t \frac{d\tau}{r(\tau)}$  has the same density function as  $\sqrt{t} \int_0^1 \frac{d\tau}{r(\tau)}$ .

Thus, if we let  $\rho(x)$  be the density function of  $\int_0^1 \frac{d\tau}{r(\tau)}$  (i.e.  $\rho(x)dx = \text{Prob} \left\{ x < \int_0^1 \frac{d\tau}{r(\tau)} < x + dx \right\}$ ) we have

$$\left\langle e^{\int_0^t \frac{d\tau}{r(\tau)}} \right\rangle = \left\langle e^{\sqrt{t} \int_0^1 \frac{d\tau}{r(\tau)}} \right\rangle = \int_0^\infty e^{\sqrt{t} x} \rho(x) dx \underset{t \rightarrow \infty}{\sim} b_1 e^{\lambda_1 t}$$



or, writing  $t^2$  for  $t$ , we have

$$(11) \quad \int_0^{\infty} e^{tx} \rho(x) dx \underset{t \rightarrow \infty}{\sim} b_1 e^{\lambda_1 t^2} .$$

One can verify, by a simple integration, that this asymptotic behavior of  $\int_0^{\infty} e^{tx} \rho(x) dx$  can be achieved by setting

$$(12) \quad \rho(x) \underset{x \rightarrow \infty}{\sim} \frac{b_1}{2\sqrt{\pi \lambda_1}} e^{-x^2/4\lambda_1}$$

By an unproved<sup>1</sup> Tauberian theorem which, however, is intuitively reasonable and similar in form to the Hardy-Littlewood Tauberian Theorem, this is the only asymptotic behavior of  $\rho(x)$  which will make equation (11) true. A result similar to this can be obtained for any homogeneous potential function  $V$ .

Now we can determine  $\lambda_1$  as follows:

- 1) Find the density function of  $\int_0^1 \frac{d\tau}{r(\tau)}$  by replacing the integral with a finite sum and performing random walks as before, to get sample values of  $\int_0^1 \frac{d\tau}{r(\tau)}$ .
- 2) Letting  $x = \int_0^1 \frac{d\tau}{r(\tau)}$ , plot  $\log \rho(x)$  vs.  $x^2$  for large  $x$ . It is to be hoped that a fairly straight line will result. The slope of this line should be  $-\frac{1}{4\lambda_1}$ .

First we present uncritically an example of how this technique is applied to the hydrogen potential.

---

<sup>1</sup>The theorem can be proved in the integrated form

$$\int_x^{\infty} \rho(\xi) d\xi \sim \frac{b_1}{2\sqrt{\pi \lambda_1}} \int_x^{\infty} e^{-\xi^2/4\lambda_1} d\xi .$$



Clearly, we cannot experimentally find a continuous density function  $\rho(x)$ . Our experimental data consists of 900 values of  $\int_0^1 \frac{d\tau}{r(\tau)}$  obtained from 900 random walks. The closest we can come to plotting a density function is to plot a histogram, which shows the number of occurrences of  $\int_0^1 \frac{d\tau}{r(\tau)}$  lying within each of a sequence of intervals of length  $\Delta$ ; e.g., if we take  $\Delta = 0.1$ , we obtain the histogram of Graph 1. The bar covering the interval  $[1.5, 1.6]$  at height 65 indicates that 65 values of  $\int_0^1 \frac{d\tau}{r(\tau)}$  lay between 1.5 and 1.6 ( $= 1.5 + \Delta$ ). Of course the choice of  $\Delta$  does much to determine the appearance of the histogram; a small  $\Delta$  gives an irregular curve, while a large  $\Delta$  obscures the structure of the histogram by "smoothing" it too much. Since the purpose of our histogram is only to study the behavior of the tail, it is clear that  $\Delta = 0.1$  makes the tail too irregular for reliable use in our calculations. If we replot the histogram with  $\Delta = 0.2$  (graph 2), the tail is considerably smoother but still retains some fine structure.

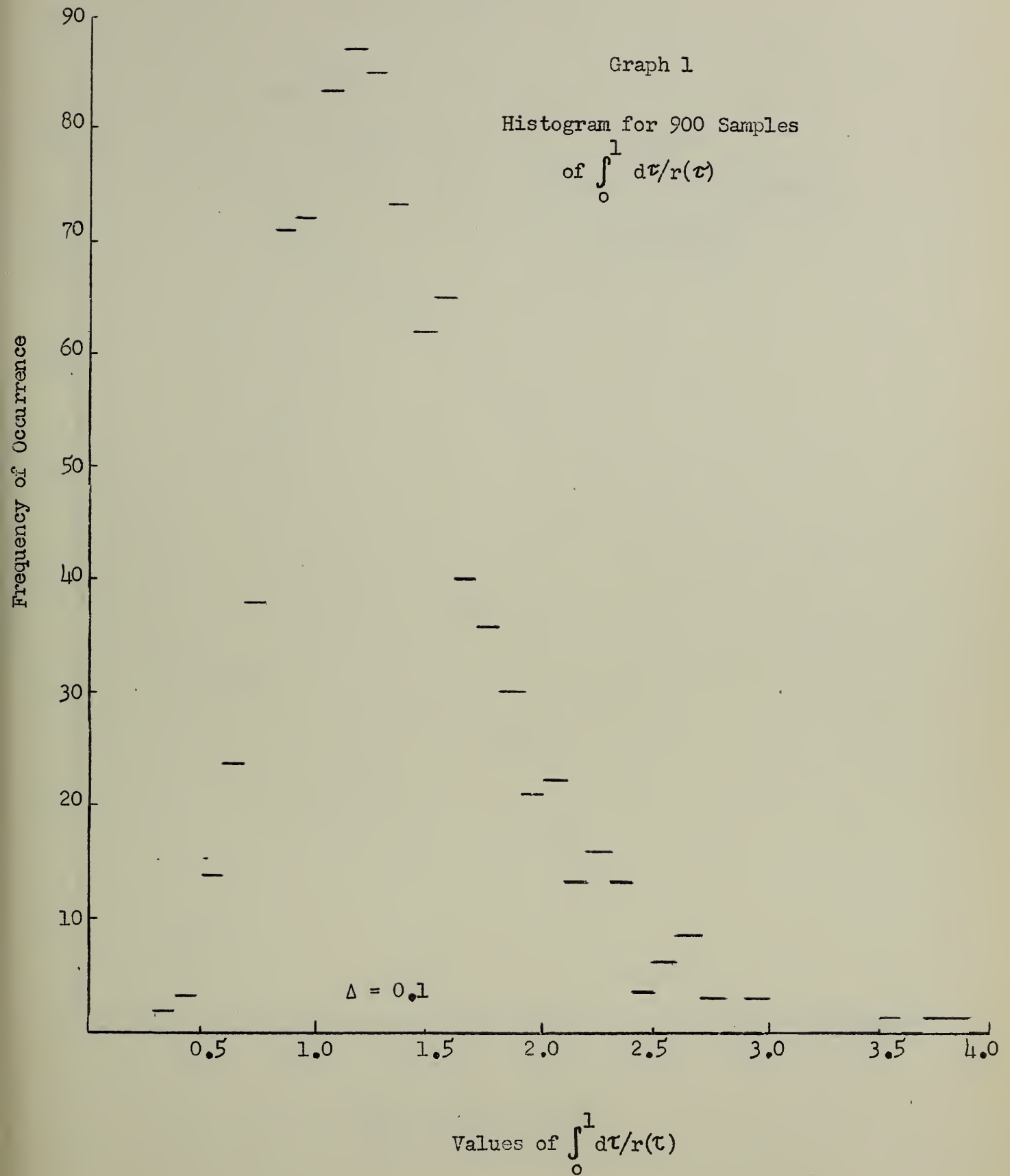
Now, if we plot  $\log \rho(x)$  vs.  $x^2$  for large  $x$ , we expect a straight line with slope  $-\frac{1}{4\lambda_1}$ . Taking values of  $\rho$  from graph 2 and plotting them against  $x^2$  on semi-log paper, we obtain graph 3. Here the "tail" of  $\rho$  is considered as beginning at  $x = 1.2$ , just to the right of the peak of  $\rho$ . The question immediately arises: what is the best line to draw through the points? We reason as follows (this argument and the general character of the graph are typical of any potential):

- 1) The first few points at the left are not really in the tail since  $x$  is not large enough. This is also clear from the



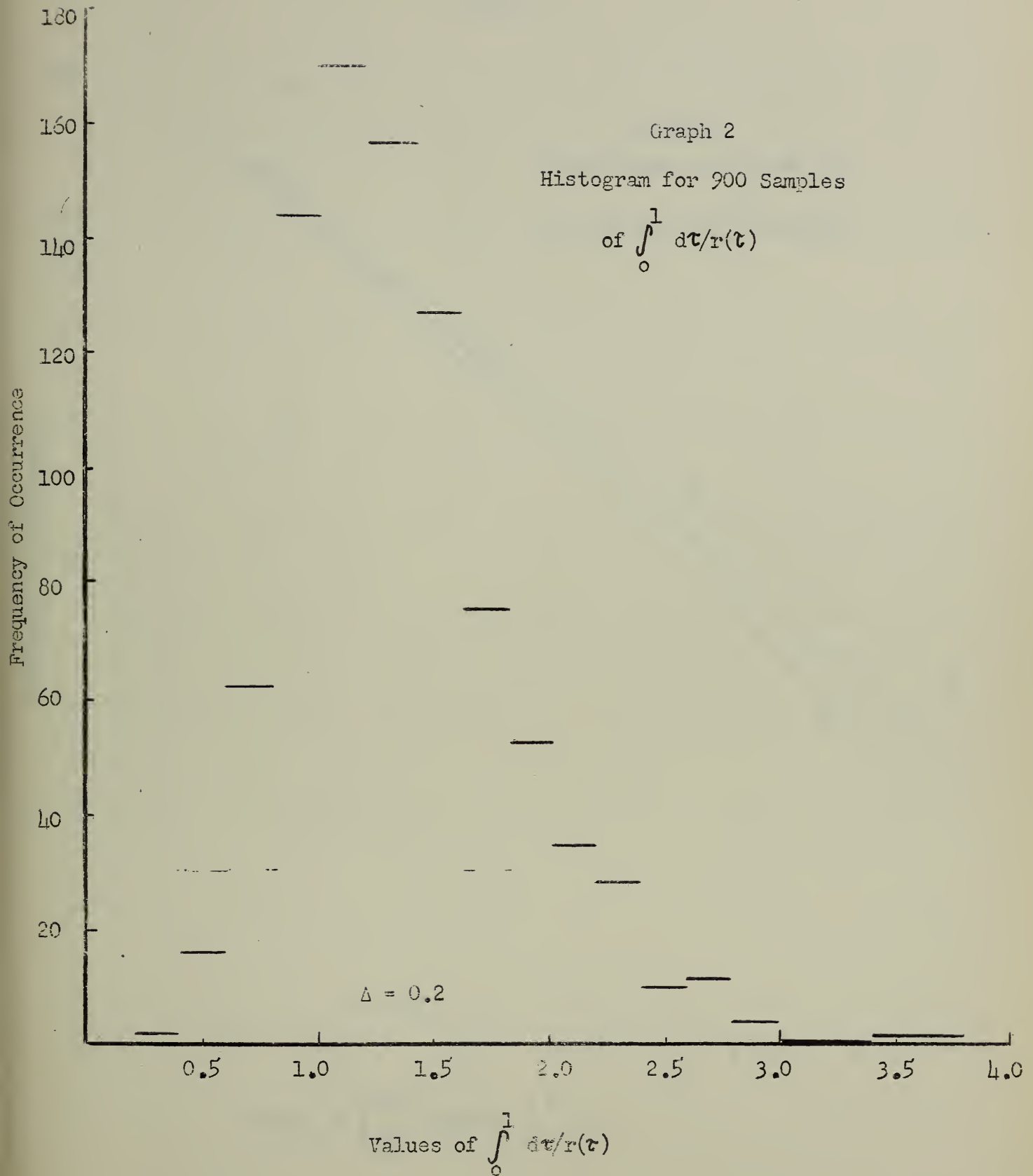
Graph 1

Histogram for 900 Samples  
of  $\int_0^1 d\tau/r(\tau)$

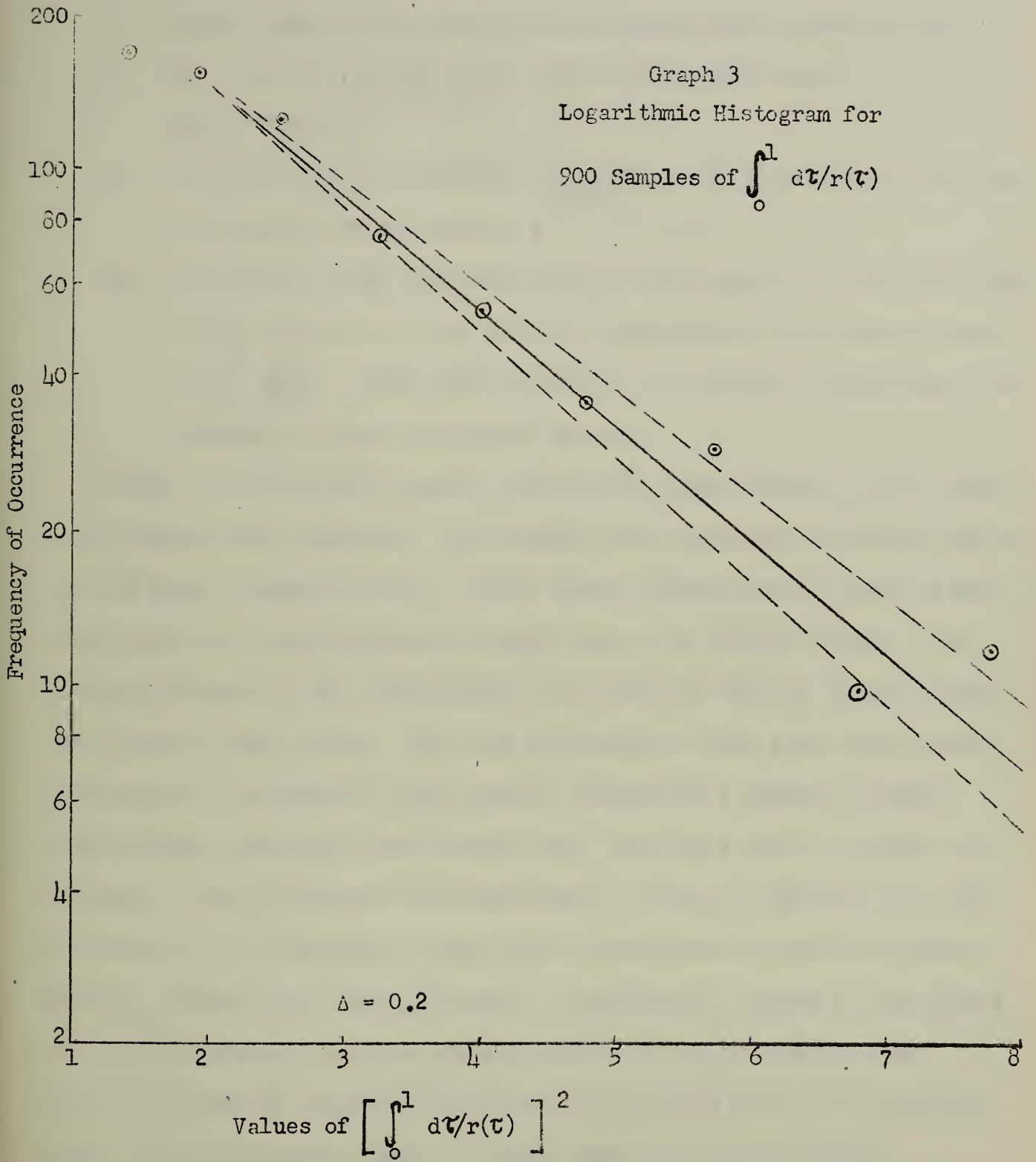














graph, where it is seen that the points cannot be well fitted

by a line, i.e. they do not lie in the region where

$$\rho(x) \sim \text{const.} e^{-x^2/4\lambda_1}.$$

- 2) The larger the x-coordinate of a point, the more valid is the approximation  $\rho(x) \sim \text{const.} e^{-x^2/4\lambda_1}$ . But
- 3) the points which represent really large values of x represent very little data, i.e. there are few occurrences of very large values of  $\int_0^1 \frac{d\tau}{r(\tau)}$ . Hence the points with the largest x-coordinates are subject to large statistical errors.

Thus, the data points become statistically less reliable as the numerical approximation improves. Consequently, we reason that the best line is the one going through a set of points whose x-values are as large as possible, but which still represent enough data to be fairly reliable. Of course, the test of the "reliability" of a point is that it lie on a line with several other points. Near the mid-portion of the graph there generally appears a sequence of such points, followed by a number of "wild" points which represent insufficient data. Drawing a line on graph 3 according to these principles and measuring its slope, we obtain  $\lambda_1 = .48$ . Obviously it is difficult to give a fair estimate of the error in such a method. A good upper bound, however, is obtained by looking at the slopes of the most extreme lines one might attempt to draw through the data points, and seeing how much those slopes differ from that of the original line. In this case the answer is quite certainly less than 10%.



The discussion gains plausibility when we consider the results for the hydrogen potential obtained from 1800 more random walks. The histogram (graph 4) has the same shape as graph 2 and has a peak at the same value of  $x$ . When we plot  $\log \rho(x)$  vs.  $x^2$  for large  $x$ , we get graph 5, which shows a marked improvement over graph 3. With the exception of the last two points on the right, which represent insufficient data, the points quite certainly lie near a line. Calculating the slope of a line fitted by eye, we find  $\lambda_1 = .52$ . Here the error seems less than 5%.

It is worth noting that about 30% of the random walks yielded values of  $\int_0^1 \frac{d\tau}{r(\tau)}$  which are large enough to be considered in the "tail" of the distribution. Thus, in the case of the hydrogen atom, one can say roughly that the slope method utilizes 30% of the data while the mean value method makes use of only about 1% (since the average  $\langle e^{\int_0^t d\tau/r(\tau)} \rangle$  is determined largely by a handful of walks yielding large values of the integral). This difference in efficiency of the two methods seems critical in the case of hydrogen, since the mean value method is infeasible for that potential while the slope method is reasonably accurate. We can get an idea of the limitations of the slope method by considering a more difficult potential.

In the proper units, the stationary Schrödinger equation for helium can be written as

$$(13) \quad \frac{1}{2} \nabla_1^2 \Psi(\vec{r}_1, \vec{r}_2) + \frac{1}{2} \nabla_2^2 \Psi(\vec{r}_1, \vec{r}_2) + \left( \frac{2}{|\vec{r}_1|} + \frac{2}{|\vec{r}_2|} - \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right) \Psi(\vec{r}_1, \vec{r}_2) = -\lambda \Psi(\vec{r}_1, \vec{r}_2) .$$

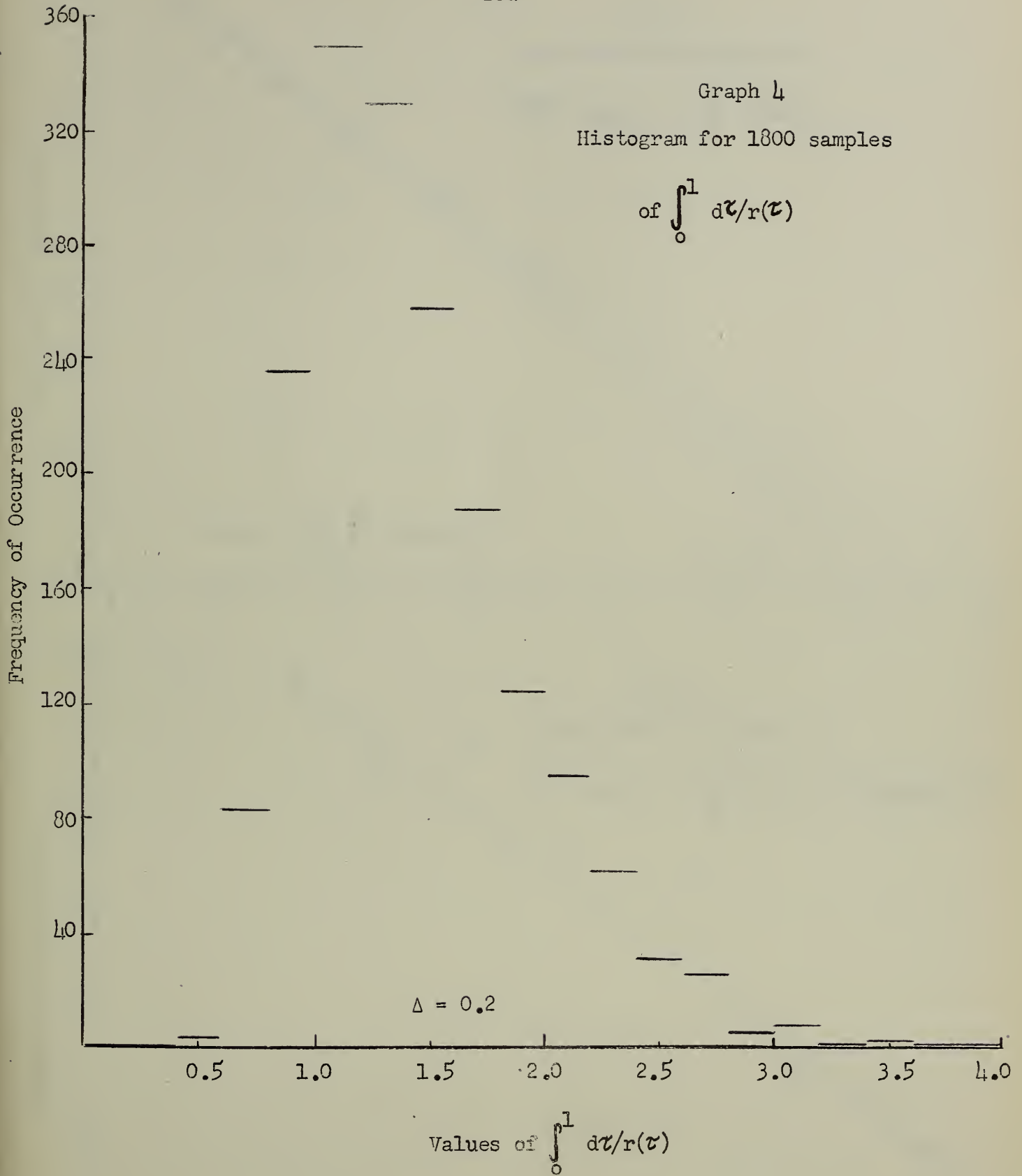




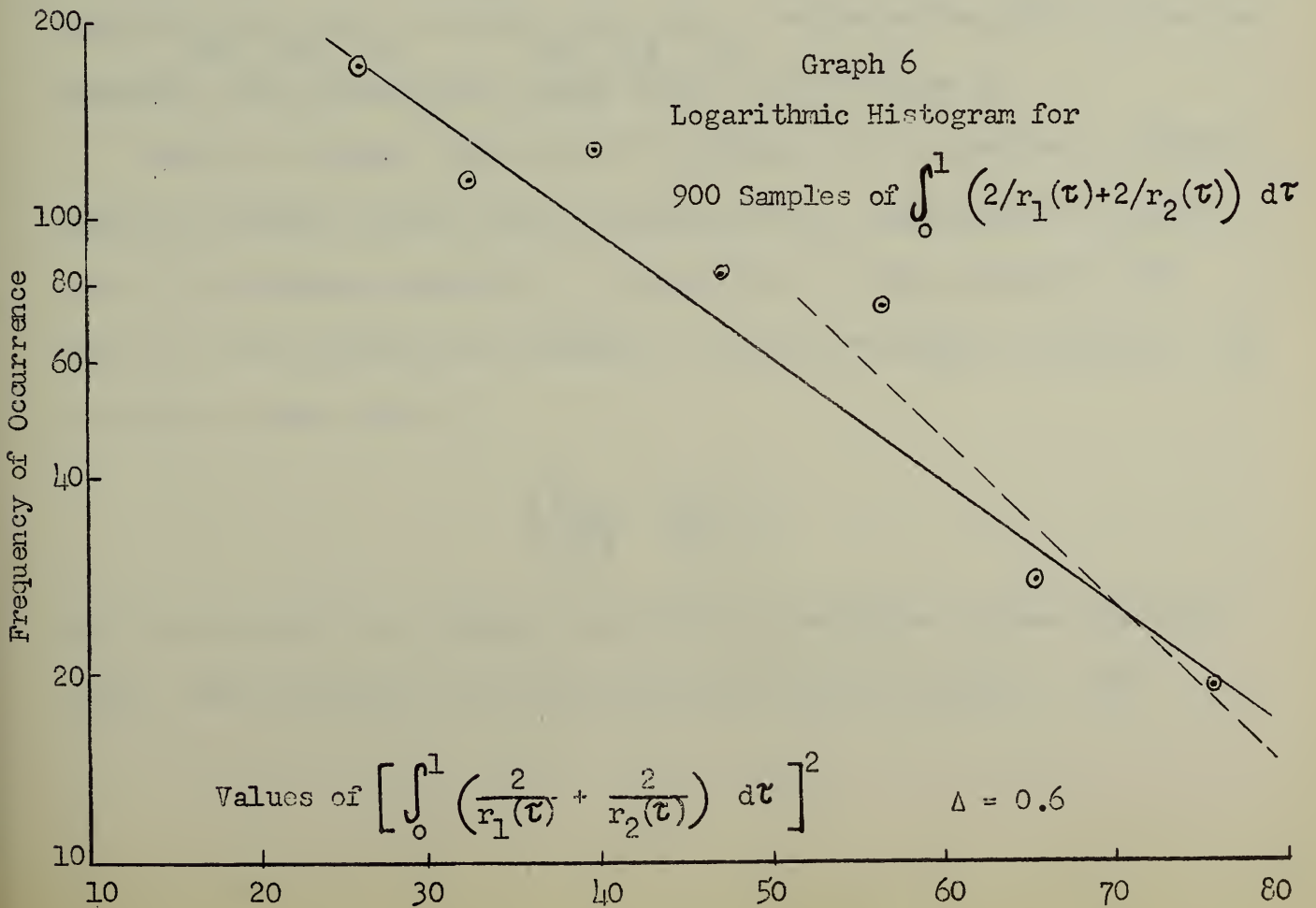
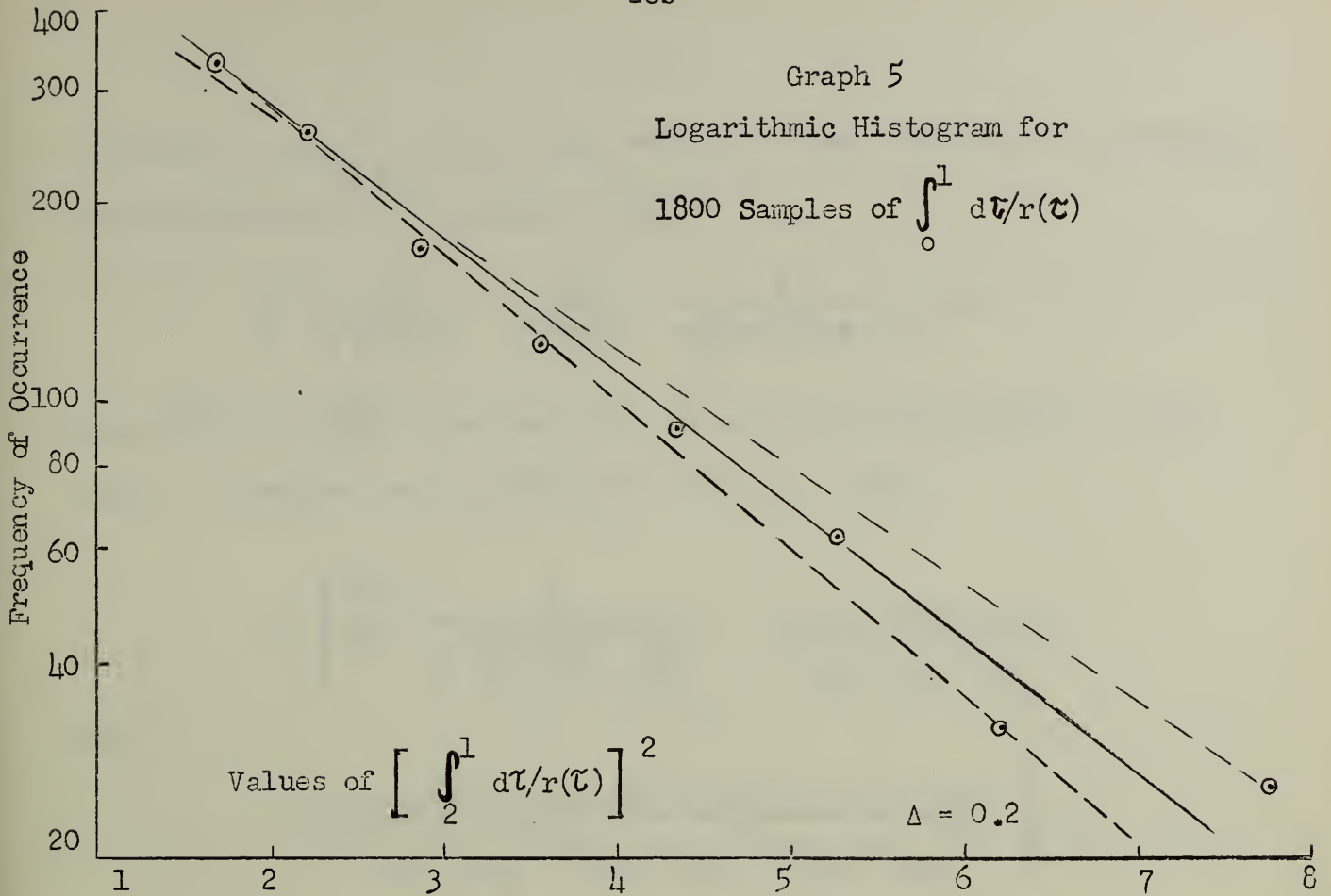
Graph 4

Histogram for 1800 samples

of  $\int_0^1 d\tau/r(\tau)$









In order to find  $\lambda_1$  by the slope method, we must determine, by sampling, the distribution of the random variable

$$\int_0^1 \left( \frac{2}{|\vec{r}_1(\tau)|} + \frac{2}{|\vec{r}_2(\tau)|} - \frac{1}{|\vec{r}_1(\tau) - \vec{r}_2(\tau)|} \right) d\tau$$

where  $\vec{r}_1(\tau)$  and  $\vec{r}_2(\tau)$  are the paths of two independent Gaussian random walks. As before, we discretize the walks and sample

$$(14) \quad \left. \begin{aligned} & .1 \left\{ \sum_{k=1}^{100} \frac{2}{\sqrt{s_{kx_1}^2 + s_{ky_1}^2 + s_{kz_1}^2}} + \frac{2}{\sqrt{s_{kx_2}^2 + s_{ky_2}^2 + s_{kz_2}^2}} \right. \\ & \left. - \frac{1}{\sqrt{(s_{kx_1} - s_{kx_2})^2 + (s_{ky_1} - s_{ky_2})^2 + (s_{kz_1} - s_{kz_2})^2}} \right\} \end{aligned} \right\}$$

where  $(s_{kx_1}, s_{ky_1}, s_{kz_1})$  and  $(s_{kx_2}, s_{ky_2}, s_{kz_2})$  determine the paths of two independent three-dimensional random walks, as in Section IV.

Before we discuss the results for helium, let us consider a slightly simpler potential, whose lowest eigenvalue can be calculated as a by-product of the helium computation. Consideration of this potential will bring out the difficulties involved in the multi-electron potentials. The potential we deal with is

$$- \left( \frac{2}{|\vec{r}_1|} + \frac{2}{|\vec{r}_2|} \right)$$

which corresponds to a helium atom without interaction between the electrons. This potential has been christened "Double Hydrogen". The



Schrödinger equation for double hydrogen is

$$(15) \quad \frac{1}{2} \nabla_1^2 \psi(\vec{r}_1, \vec{r}_2) + \frac{1}{2} \nabla_2^2 \psi(\vec{r}_1, \vec{r}_2) + \left( \frac{2}{|\vec{r}_1|} + \frac{2}{|\vec{r}_2|} \right) \psi(\vec{r}_1, \vec{r}_2) = -\lambda \psi(\vec{r}_1, \vec{r}_2).$$

It is easy to show that the eigenvalues of this equation are

$\lambda = \mu_j + \mu_k$ , where  $\mu_j$  and  $\mu_k$  are eigenvalues (not necessarily distinct) of

$$(16) \quad \frac{1}{2} \nabla^2 \varphi(\vec{r}) + \frac{2}{r} \varphi = \mu \varphi$$

The eigenvalues of (16) are  $\mu_k = -\frac{2}{k^2}$  ( $k = 1, 2, 3, \dots$ ) and the entire positive real axis. The lowest eigenvalue of (15) is thus  $\lambda_1 = -4$ , and the second lowest is  $\lambda_2 = -2.5$ . As before, we now redefine  $\lambda$  so that  $\lambda_1 = 4$ ,  $\lambda_2 = 2.5$ . (It also follows simply from our probabilistic approach that  $\lambda_1 = 2\mu_1$ , since

$$\lambda_1 \sim \frac{\log \left\langle e^{\int_0^t \left( \frac{2}{r_1(\tau)} + \frac{2}{r_2(\tau)} \right) d\tau} \right\rangle}{t};$$

but  $r_1(\tau)$  and  $r_2(\tau)$  are the paths of independent random walks, and hence

$$\left\langle \int_0^t \frac{2}{r_1(\tau)} d\tau \int_0^t \frac{2}{r_2(\tau)} d\tau \right\rangle = \left\langle \int_0^t \frac{2}{r_1(\tau)} d\tau \right\rangle^2.$$

The desired result follows.)

Of course, since we know that  $\lambda_1 = 2\mu_1$ , one statistical method for finding  $\lambda_1$  would be to find  $\mu_1$  by sampling the potential  $\frac{2}{r}$ , and then to double the result. This scheme, however, breaks down as soon as the potential contains an interaction term between  $\vec{r}_1$  and  $\vec{r}_2$ . Hence, since



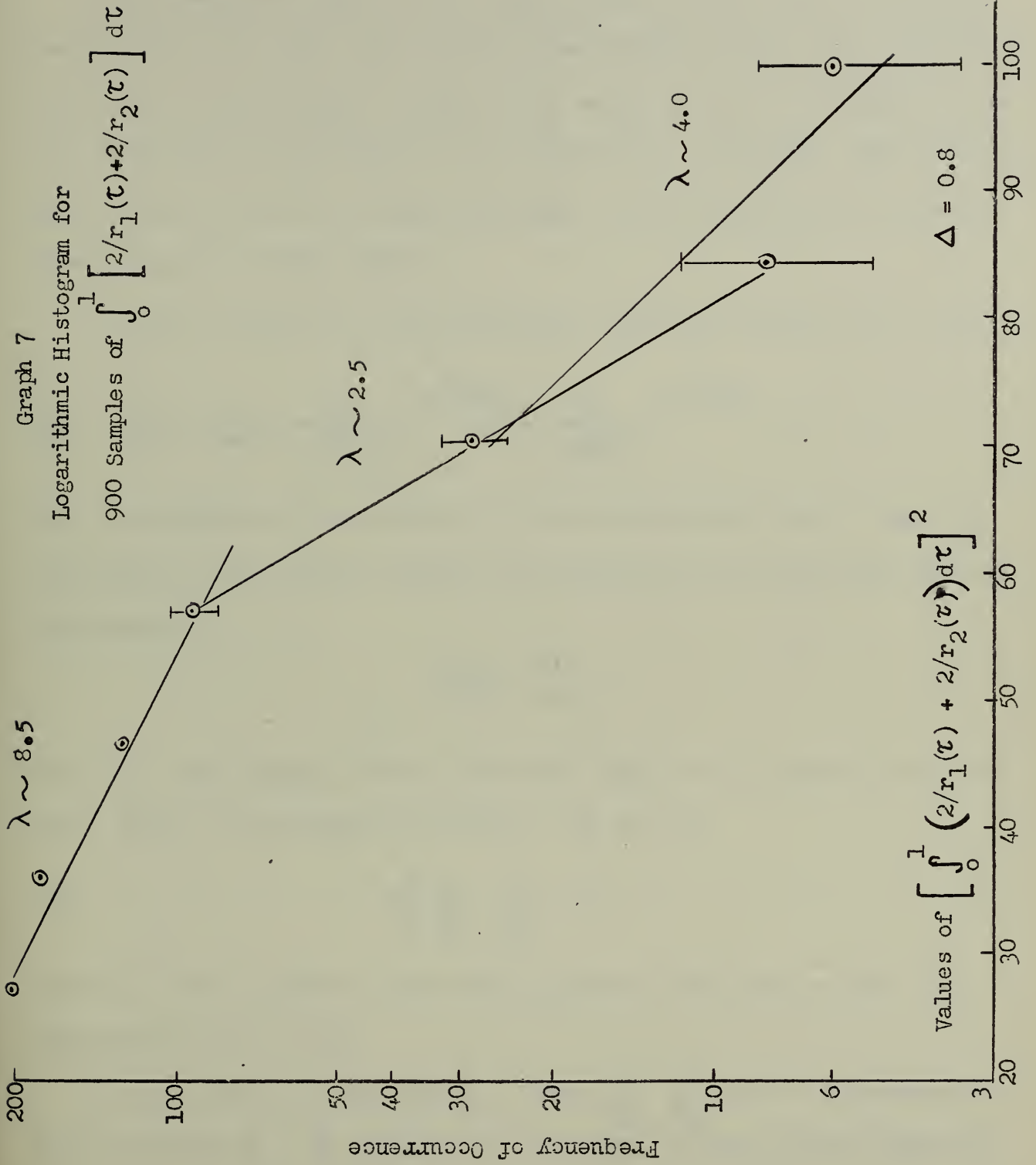


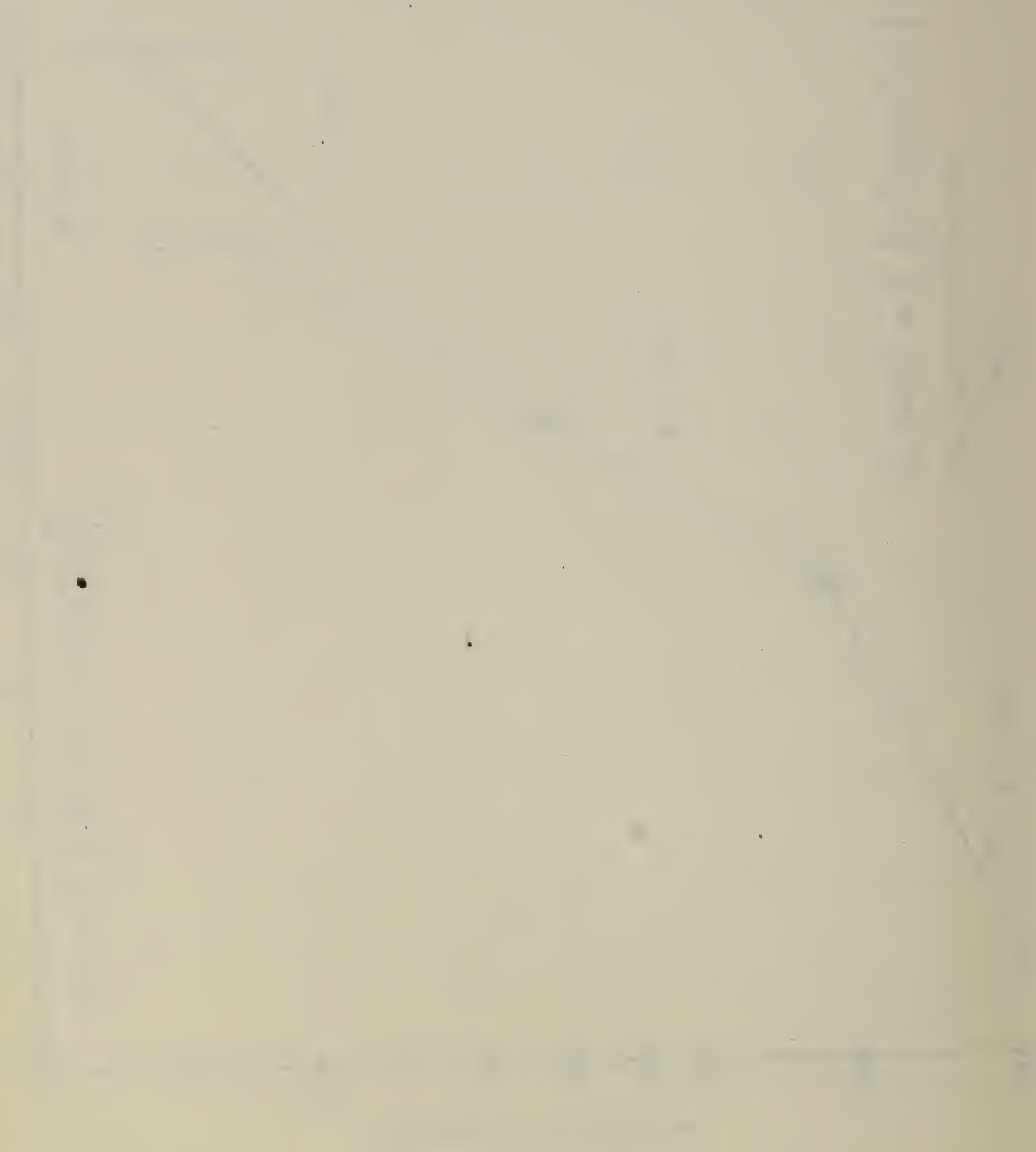
we want to test the general applicability of the method to multi-electron potentials, we do not employ this short-cut, but sample the random variable

$$\int_0^1 \left( \frac{2}{r_1(\tau)} + \frac{2}{r_2(\tau)} \right) d\tau \sim .1 \sum_{k=1}^{100} \left( \frac{2}{\sqrt{s_{kx_1}^2 + s_{ky_1}^2 + s_{kz_1}^2}} + \frac{2}{\sqrt{s_{kx_2}^2 + s_{ky_2}^2 + s_{kz_2}^2}} \right) .$$

Plotting the logarithm of the density function  $\rho(x)$  of this random variable versus  $x^2$ , we obtain graph 6. The graph is based on 900 values of  $\int_0^1 \left( \frac{2}{r_1(\tau)} + \frac{2}{r_2(\tau)} \right) d\tau$ , of which about 500 are represented on the graph (the other values were too small to be in the "tail" of  $\rho(x)$ ). On the basis of the solid line drawn, we find  $\lambda_1 = 5.6$  (true value = 4.0). It is evident that the data points do not clearly determine a line. If we ignore the points on the left, arguing that they are not yet in the tail, and draw the dotted line based on the right points, we get  $\lambda_1 = 4.6$ . This result, however, is hardly reliable. If we increase  $\Delta$  from 0.6 to 0.8, we obtain graph 7, which is considerably smoother. The slope of the line through the first four points on the left yields  $\lambda_1 = 8.5$ . Clearly these points are not yet in the tail. The next three points represent only 29, 8, and 6 walks respectively and are subject to too great statistical fluctuations. The expected statistical fluctuation of a point representing  $n$  walks is about  $\sqrt{n}$ . Hence the limits of inaccuracy of the data points are represented approximately by the vertical bars we have drawn. We see that the points are not inconsistent with the true value  $\lambda_1 = 4$ , but they hardly give us any further information.







The redeeming feature of the slope method, in the case of the hydrogen potential (and thus also for the potential  $\frac{2}{r}$ ) was that the exponential "tail" of the distribution started early enough so that the points in the tail represented enough walks to be statistically reliable. This feature does not seem to hold for double hydrogen. We now proceed to a more precise discussion of this point.

A simple extension of the reasoning leading to equation (12) yields

$$(17) \quad \rho(x) \underset{x \rightarrow \infty}{\sim} \frac{b_1}{2\sqrt{\pi\lambda_1}} e^{-x^2/4\lambda_1} + \frac{b_2}{2\sqrt{\pi\lambda_2}} e^{-x^2/4\lambda_2} + \dots$$

Thus, an estimate of the goodness of the approximation (12) is given by the ratio of the first two terms on the right side of (17); this ratio is approximately

$$e^{x^2 \left( \frac{1}{4\lambda_2} - \frac{1}{4\lambda_1} \right)}$$

When  $x$  is large enough to make this ratio large, we can say we are in the tail. Thus, the criterion for being in the tail is

$$(18) \quad \frac{x^2}{4} \left( \frac{1}{\lambda_2} - \frac{1}{\lambda_1} \right) \geq M$$

where  $M$  is some constant, whose value depends on how good we want the approximation (12) to be.

For the hydrogen potential  $\frac{1}{r}$ , we have  $\frac{1}{\lambda_2} - \frac{1}{\lambda_1} = 8-2 = 6$ , whereas for double hydrogen  $\frac{1}{\lambda_2} - \frac{1}{\lambda_1} = \frac{1}{2.5} - \frac{1}{4} = .15$ . Thus, if the tail for hydrogen starts at  $x_0$ , the tail for double hydrogen starts at  $\sqrt{40} x_0$ . Now we ask



the question: how many random walks for double hydrogen must we perform in order to obtain an accuracy comparable to that obtained for hydrogen?

If we perform  $N_H$  random walks for hydrogen, the number of walks yielding values of  $\int_0^1 \frac{d\tau}{r(\tau)}$  in the tail of  $\rho_H(x)$  is about  $N_H \int_{x_0}^{\infty} \rho_H(x) dx$ .

To achieve a comparable accuracy for double hydrogen, we want to get the same number of points in the tail. Hence we require

$$N_H \int_{x_0}^{\infty} \rho_H(x) dx = N_{DH} \int_{\sqrt{40x_0}}^{\infty} \rho_{DH}(x) dx$$

where  $N_{DH}$  = number of double hydrogen random walks

$$\rho_{DH}(x) = \text{density function of } \int_0^1 \left( \frac{2}{r_1(\tau)} + \frac{2}{r_2(\tau)} \right) d\tau .$$

Ignoring the constant factor in equation (12), we thus have the order-of-magnitude equation

$$\frac{N_{DH}}{N_H} \sim \frac{\int_{x_0}^{\infty} e^{-x^2/2} dx}{\int_{\sqrt{40x_0}}^{\infty} e^{-x^2/16} dx} .$$

Using the asymptotic series for the two integrals, we have, if  $x_0$  is large,

$$\frac{N_{DH}}{N_H} \sim \sqrt{40} \frac{e^{-x_0^2/2}}{e^{-40x_0^2/16}} \sim 6 e^{\frac{1}{2}x_0^2} .$$

Looking at graph 3, we see that the tail seems already to have started at  $x_0^2 = 2$ . The above calculation is inaccurate for such a small value of  $x_0$  (we could calculate  $N_{DH}/N_H$  more exactly, but it hardly seems worthwhile);





it yields the qualitatively correct result that, for a given accuracy in the lowest eigenvalue, we require many more samples for double hydrogen than we do for hydrogen.

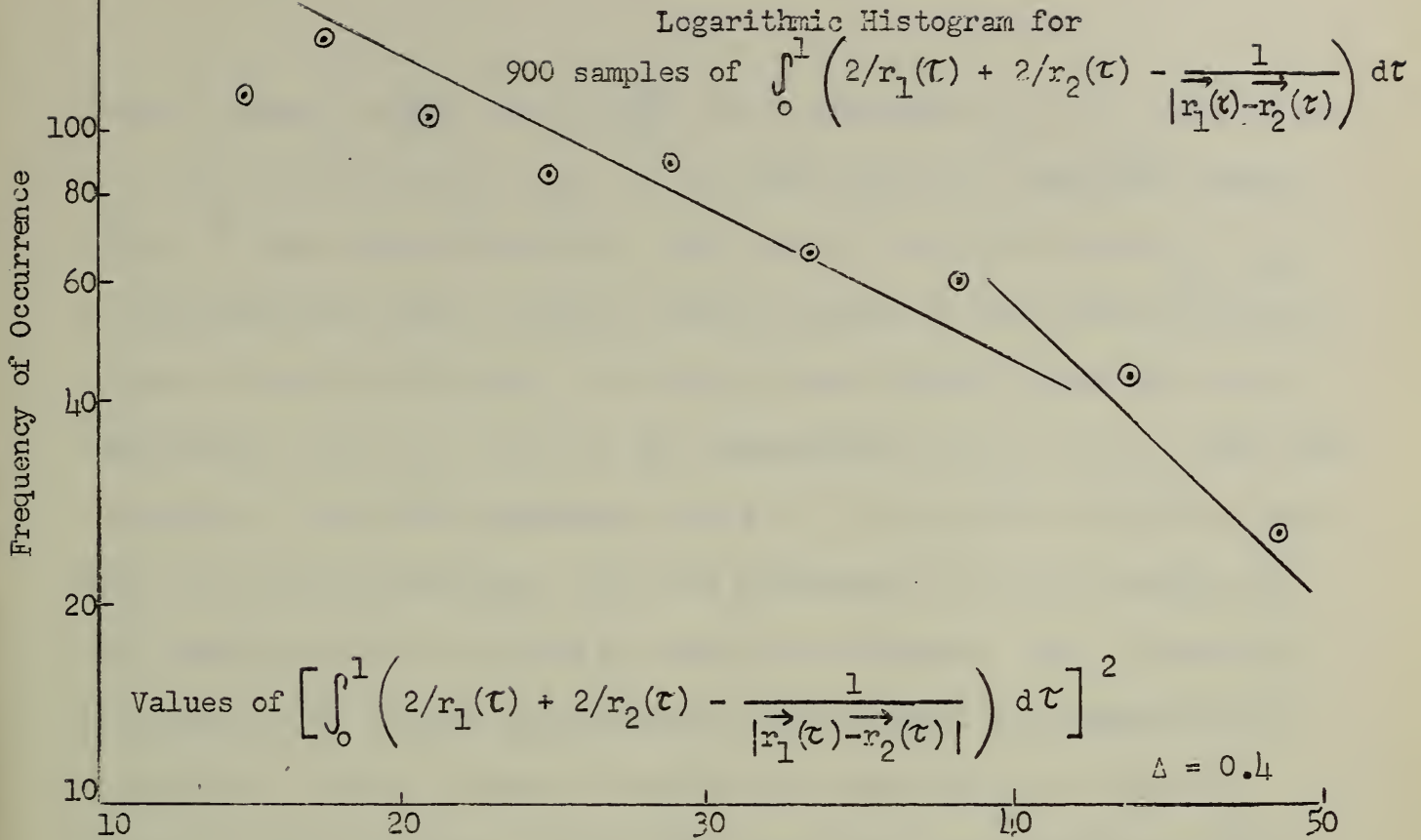
Formally, the reason for the increase in the number of samples required is the creeping together of the reciprocals of the first and second eigenvalues (see eqn. (18)). "Physically", the reason is this: for hydrogen, we sample a single random variable  $X$ ; for double hydrogen we sample the sum  $X + Y$ , where  $X$  and  $Y$  are independent and identically distributed. Roughly speaking, if  $p$  is the probability that a value of  $X$  lie in the tail of the distribution of  $X$ , then the probability that a value of  $X + Y$  lie in the tail of the distribution of  $X + Y$  is only about  $p^2$ . This argument can be made quantitative and leads to our previous result.

We conclude that 900 random walks constitute far too small a sample for the double hydrogen potential or, more generally, for any multi-electron potential. The 900 helium random walks confirm this conclusion. The logarithmic histograms with  $\Delta = 0.4$  and  $\Delta = 0.8$  respectively are given by graphs 8 and 9. The graphs are consistent with the true value  $\lambda = 2.90$ , but not enough points are in the tail to allow an accurate determination.

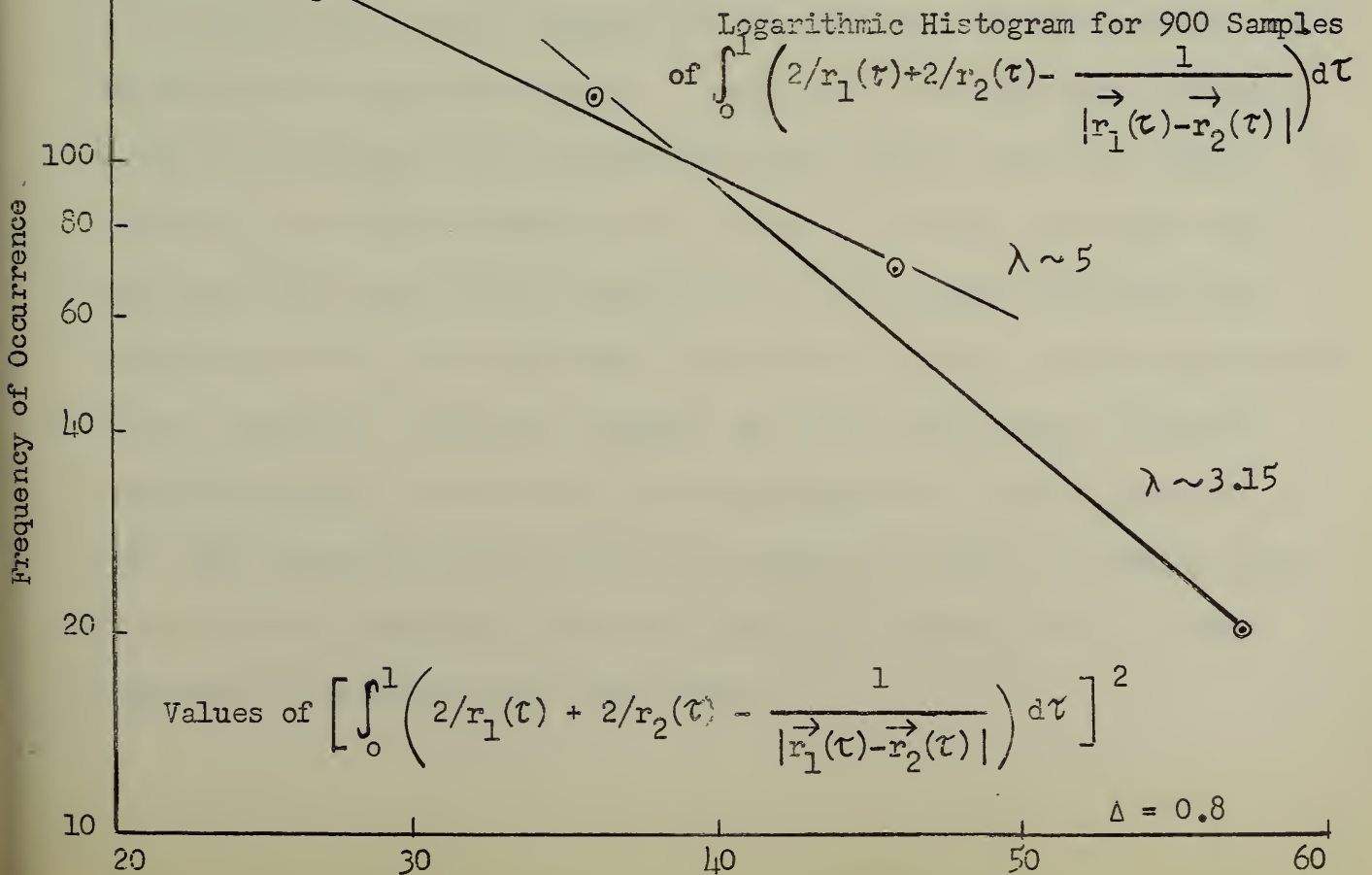
In the next section we discuss some schemes for making our method feasible for the study of more complicated potentials. One more point is worth mentioning, however, in connection with double hydrogen. We plotted a logarithmic histogram (graph 7) for this potential, based on 900 values of the random variable  $X = \int_0^1 \left( \frac{2}{r_1(\tau)} + \frac{2}{r_2(\tau)} \right) d\tau$ . Let us denote these 900 values by  $A_{1_i} + A_{2_i}$  ( $i = 1, \dots, 900$ ) where  $A_{1_i} = \int_0^1 \frac{d\tau}{r_{1_i}(\tau)}$ , etc.



Graph 8



Graph 9





However, since  $\int_0^1 \frac{d\tau}{r_1(\tau)}$  and  $\int_0^1 \frac{d\tau}{r_2(\tau)}$  are independent,  $A_{1_i} + A_{2_j}$  ( $i \neq j$ ) or  $A_{1_i} + A_{1_j}$  ( $i \neq j$ ) or  $A_{2_i} + A_{2_j}$  ( $i \neq j$ ) also constitute legitimate sample values of the random variable X. Thus, out of our 1800 values  $A_{1_i}, A_{2_i}$ , we can manufacture  $1800 \times 1799/2 \sim 160,000$  values of the random variable X, instead of merely 900 values. Of course, these 160,000 values are not all independent; e.g.  $A_{1,1} + A_{1,2}$  is not independent of  $A_{1,1} + A_{1,3}$ . There are a maximum of only 900 independent values of X in the set of 160,000. However, one feels intuitively that the dependence is not too strong and that something might be gained by using the additional data. (The whole situation can be treated quantitatively, but we omit the treatment here. An analogous problem is that of sampling the distribution of the sums rolled on a pair of dice. We roll 900 pairs, but then combine the 1800 single rolls into all possible pairs.)

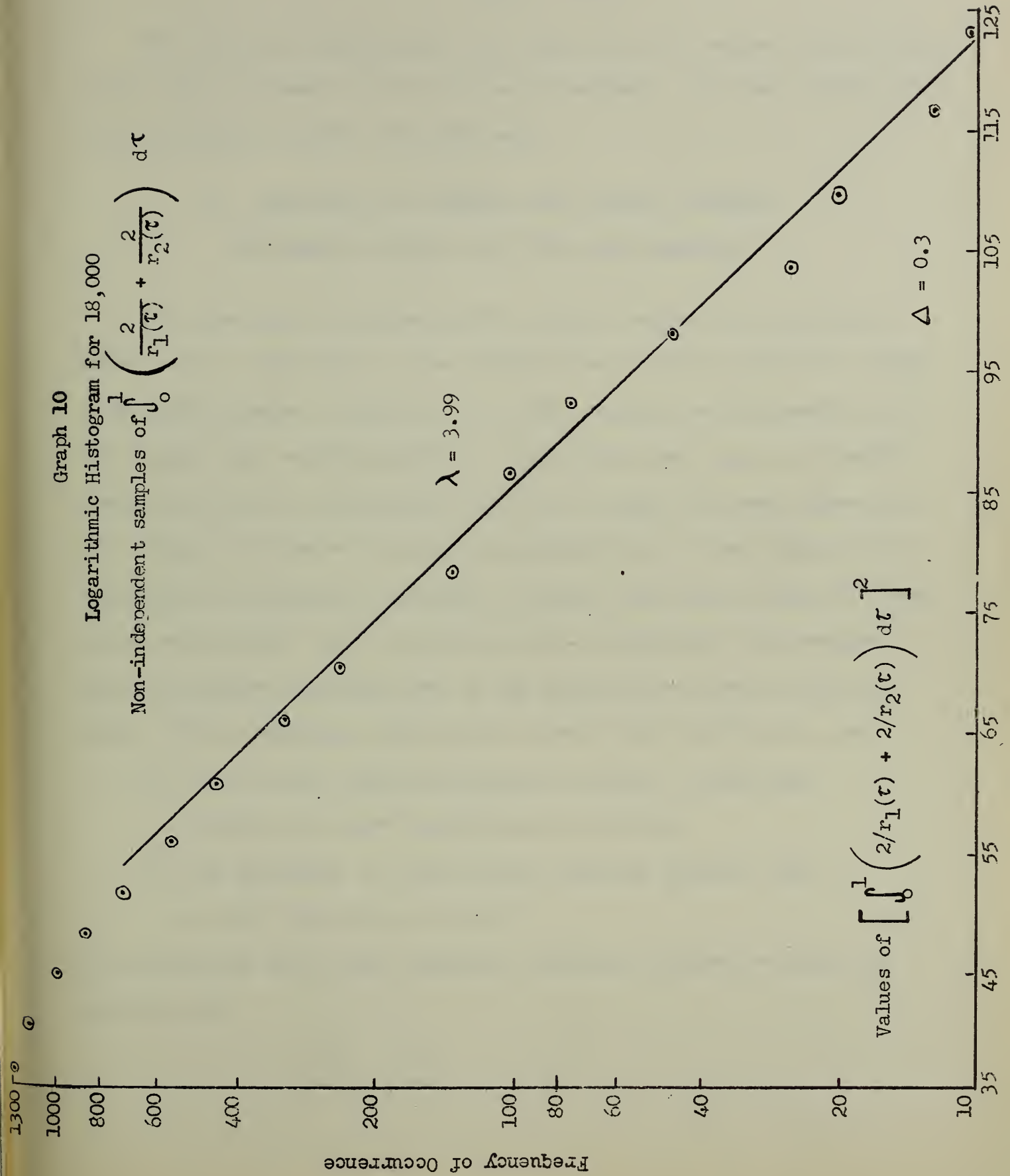
Actually, we did not combine the data into all 160,000 possible pairs, but considered only 18,000 pairs. Graph 10 is the logarithmic histogram based on these pairs. The improvement over graphs 6 and 7 is marked, and we claim about 6% accuracy for the result  $\lambda = 3.99$ . One might tend to challenge this result on the ground that we are getting "something for nothing", and thus violating what some consider to be a fundamental law of nature. However, we should remember that the data already contained enough information to find the lowest eigenvalue of "single" hydrogen to about 5% (graph 5), and thus (by our remarks on p. 28) the lowest eigenvalue of double hydrogen to the same accuracy. Graphs 6 and 7 merely represented inefficient use of the data.



Graph 10

Logarithmic Histogram for 18,000

Non-independent samples of  $\int_0^1 \left( \frac{2}{r_1(\tau)} + \frac{2}{r_2(\tau)} \right) d\tau$







This device of "multiplying" the data fails, of course, when an interaction term is present (as in the case of helium). Our final section deals with procedures feasible for this case.

## VI. TECHNIQUES FOR SPEEDING THE SAMPLING PROCESS;

### HIGH-SPEED COMPUTING AND IMPORTANCE SAMPLING

The experiments described in the previous sections were performed mainly out of curiosity, to see whether equation (5) is verifiable within a reasonable number of samples. At a later stage, one is naturally led to ask whether this sampling method is of practical use, e.g. as a possible competitor with the variational method for finding the lowest eigenvalue. The authors are unaware of just what potentials are of real interest but unamenable to variational treatment. However, there are a number of techniques which should make the present scheme feasible for the treatment, without excessive sampling time, of any system whose dimensionality is small. These techniques, which will be dealt with briefly here, are

- 1) The use of a high-speed digital computer, eliminating entirely the "slow" punched-card operations;
- 2) The employment of sophisticated sampling schemes, which we call "importance sampling".

Both techniques can be used together, effecting a further reduction of sampling time.



For an indication of the times involved in the use of IBM methods, we note that 900 hydrogen random walks of 100 steps each require about 90 IBM hours. To accumulate the same data on the National Bureau of Standards Western Automatic Computer (SWAC), about one hour is required. No hydrogen walks have yet been performed on the SWAC, but harmonic oscillator results will soon be available. One can say, in general, that for our purposes the SWAC is faster than IBM by a factor of at least 50. The programming of these random walk routines is quite elementary, involving mainly the iteration of simple operations. The problem of supplying random numbers to govern the random walks is of some interest. External sources, such as prepared tables, are unfeasible, since the machine must generate its own random numbers if it is to operate at electronic speeds. A number of arithmetic schemes for the internal generation of "random" numbers have been proposed (the numbers so generated are only pseudo-random, since they are completely determined by the numbers inserted into the arithmetic scheme), and a particularly simple method of Dr. J. B. Rosser has been found to produce digits which satisfy the usual tests of randomness.

Regardless of whether SWAC or IBM equipment is being used, one can gain another factor of 10 or more by the use of importance sampling. This topic has been discussed in recent literature (see [5]) and is of enough simplicity and interest to warrant a brief review.

Suppose we have a random variable  $X$  which is continuously distributed over some interval  $(a,b)$  with the probability density  $\rho(x)$  ( $\int_a^b \rho(x) dx = 1$ ),



and suppose we are interested in finding the expected value  $\bar{f}$  of some function  $f(X)$ . Then  $\bar{f} = \int_a^b f(x) \rho(x) dx$ . (Alternatively, and for greater clarity, we sometimes write  $\bar{f} = E_{\rho}(f(X))$ , i.e.  $\bar{f}$  is the expectation of  $f(X)$  over a population in which  $X$  has the density function  $\rho(x)$ .) The integral can, of course, be estimated by sampling: we merely pick many values of  $X$  from a population governed by the density function  $\rho(x)$ , compute  $f(X)$  for each value of  $X$ , and average the results. Suppose we also compute

$$\overline{(f - \bar{f})^2} = \int_a^b (f(x) - \bar{f})^2 \rho(x) dx \equiv \sigma_{f, \rho}^2 .$$

Then it is well known that the expected error in our estimate of  $\bar{f}$ , if we take as our estimate the average of  $n$  sample values, is  $\frac{1}{\sqrt{n}} \sigma_{f, \rho}$ .

Importance sampling is based merely on the recognition of the fact that it is possible to bias the sampling procedure in such a way as to leave  $\bar{f}$  unchanged but to make  $\sigma$  smaller. Specifically, suppose  $\varphi(x)$  is another probability density function defined on the interval  $(a, b)$  with

$$\int_a^b \varphi(x) dx = 1. \text{ We certainly have}$$

$$(19) \quad E_{\rho}(f(X)) = \int_a^b f(x) \rho(x) dx = \int_a^b \left( \frac{f(x) \rho(x)}{\varphi(x)} \right) \varphi(x) dx .$$

Thus, in our notation,

$$E_{\rho}(f(X)) = E_{\varphi} \left( \frac{f(X) \rho(X)}{\varphi(X)} \right) ,$$

i.e. the expectation of  $f(X)$  over a population governed by the density function  $\rho$  is the same as the expectation of  $f \rho / \varphi$  over a population



governed by the density function  $\varphi$ . The interpretation of this fact is simple: we bias the density function of the population by the factor  $\frac{\varphi}{\rho}$ , and hence must "weight" our values of  $f$  by the factor  $\frac{\rho}{\varphi}$  if we are to get the correct answer. Let us compute the expected error in the average of  $n$  samples, if we estimate  $\bar{f}$  by this scheme. The error is  $\frac{1}{\sqrt{n}} \sigma_{f\rho/\varphi, \varphi}$ , where

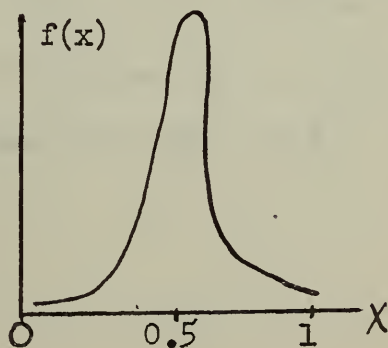
$$\sigma_{f\rho/\varphi, \varphi} = \int_a^b \left( \frac{f(x)\rho(x)}{\varphi(x)} - \bar{f} \right)^2 \varphi(x) dx .$$

By the proper choice of the function  $\varphi(x)$ ,  $\sigma_{f\rho/\varphi, \varphi}$  may be made much smaller than  $\sigma_{f, \rho}$ . In particular, if we knew  $\bar{f}$  (which is of course what we are trying to find), we could let

$$(20) \quad \varphi(x) = \frac{\rho(x)f(x)}{\bar{f}}$$

and thus have no error in our estimate. In practice, we do not know  $\bar{f}$ , and it is not easy to concoct a population governed by an arbitrary density function  $\varphi(x)$ . We can, however, use whatever we know about  $f$  to improve our sampling procedure, for (20) tells us that the error will be decreased if we distort  $\rho$  into a new distribution with more mass centered in the regions where  $f(x)$  is large. In practice, it is not hard to construct a population with this general characteristic.

A simple illustration may clarify this discussion. Suppose  $f(x)$  is a function of the type pictured, and we wish to find  $\int_0^1 f(x)dx$  by a sampling procedure. This is the same as finding  $E\rho(f(X))$  where  $\rho(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$  ;







i.e. we pick values of  $X$  from a rectangular distribution on  $(0,1)$ , and find the average value of  $f(x)$ . It is clear that most of the data are wasted in this procedure, since the major contribution to the expectation comes from the few values of  $X$  which we happen to choose from the neighborhood of  $0.5$ . One easily sees that a better procedure would be to bias the sampling scheme so that most of the points sampled lie in the vicinity of  $0.5$ . It is then necessary to weight the results somehow, so as to compensate for the bias introduced. The previous paragraph shows, quantitatively, how to do this. The term "importance sampling" arises from the fact that the amount of sampling time we spend in studying various parts of the population is proportional to the importance of those parts in determining the average.

The applicability of this technique to our problem is immediate. In the case of the hydrogen atom, we are interested in  $E \left\{ e^{-\int_0^t d\tau/r(\tau)} \right\}$ ; i.e. we sample from a space whose elements are paths  $r(\tau)$ , and compute the average value of the functional  $F(r) \equiv e^{-\int_0^t d\tau/r(\tau)}$ . In practice, the space of paths from which we sample contains only a finite number of elements  $r_i$  ( $i = 1, \dots, N$ ), since the paths are determined by the choice of a finite number (300) of Gaussian deviates, which themselves have a basic increment of  $10^{-3}$ . If  $p(r_i)$  is the probability associated with the  $i^{\text{th}}$  path, we can write



$$E \left\{ e^{\int_0^t d\tau/r(\tau)} \right\} = \sum_{i=1}^N F(r_i) p(r_i) = E_p (F(r)) .$$

We have observed in Section IV that the major contribution to this expectation comes from a few paths  $r_i$  which stay near the origin, i.e. for which  $F(r_i)$  is very large. It follows from the foregoing discussion that we can decrease the error in our estimate by associating a new set of probabilities  $q(r_i)$  with the paths. The  $q(r_i)$  will be an improvement over the  $p(r_i)$  if  $q(r_i) = p(r_i)f(r_i)$  where  $f(r_i)$  is, roughly, proportional to  $F(r_i)$  (it is sufficient that  $f$  be large when  $F$  is large, and small when  $F$  is small). Then, as before, we have

$$(21) \quad E_p(F) = \sum_{i=1}^N F(r_i)p(r_i) = \sum_{i=1}^N \left[ F(r_i) \frac{p(r_i)}{q(r_i)} \right] q(r_i) = E_q \left( F \frac{p}{q} \right) ,$$

but the variance of the  $q$ -scheme is less than that of the  $p$ -scheme.

There arises the practical question of how to find a distribution  $q$  having the desired property, without hopelessly complicating the calculations. The scheme presented here is, perhaps, the simplest possible one. Essentially, we bias the random numbers chosen so that at any step the "particle" is more likely to step toward the origin than away from it. Most of the paths sampled will, therefore, be the interesting ones which stay near the origin. Furthermore, the biasing is conducted in such a fashion that the factor  $p(r_i)/q(r_i)$  is easily computed. We recall from Section IV that in the unbiased walks one step (say the  $k^{\text{th}}$ ) consisted in choosing three numbers  $X_k, Y_k, Z_k$  from the Gaussian deviates with mean 0,



variance 1, and basic increment .001. These three numbers represent the components of the step on each of the Cartesian axes. Then the probability  $p(r_i)$  to be associated with a 100-step path  $r_i$  determined by the 300 choices  $X_{1i}, Y_{1i}, Z_{1i}, \dots, X_{100i}, Y_{100i}, Z_{100i}$  is

$$p(r_i) = \prod_{k=1}^{100} \left[ (2\pi)^{-3/2} \exp \left\{ -\frac{X_{ki}^2}{2} - \frac{Y_{ki}^2}{2} - \frac{Z_{ki}^2}{2} \right\} \cdot 10^{-9} \right] .$$

As far as we are concerned, the only important property of  $p(r_i)$  is that it has the form

$$p(r_i) = \prod_{k=1}^{100} g(X_{ki})g(Y_{ki})g(Z_{ki}) \cdot 10^{-9}$$

where  $g$  is a probability density, and an even function. Suppose now that we bias the walk in the following fashion: let  $\alpha > \beta > 0$  with  $\alpha + \beta = 1$ ; at the  $k^{\text{th}}$  step, pick  $X_k$  as before from a Gaussian distribution, but ignore the sign of  $X_k$ ; then we play some game of chance with two possible outcomes A and B, with probabilities  $\alpha$  and  $\beta$ ; if  $\sum_{j=1}^{k-1} X_j \geq 0$  and the outcome is A, interpret  $X_k$  as negative; if  $\sum_{j=1}^{k-1} X_j \geq 0$  and the outcome is B, interpret  $X_k$  as positive; if  $\sum_{j=1}^{k-1} X_j < 0$  and the outcome is  $\begin{Bmatrix} A \\ B \end{Bmatrix}$ , interpret  $X_k$  as  $\begin{Bmatrix} \text{positive} \\ \text{negative} \end{Bmatrix}$ . Similarly, choose Gaussian deviates and play the auxiliary game for the Y- and Z- coordinates. Clearly, the effect of this scheme is that, at a given step, the absolute value of a particular coordinate decreases with probability  $\alpha$  and increases with probability  $\beta$ . Hence, the walks are systematically biased toward the origin. The two-outcome game involves little extra effort; if we take  $\alpha = .6$  and  $\beta = .4$ , we need



merely pick a digit out of an equidistributed population of zeros, ones, ... , nines. If the digit chosen is not greater than 5, we call the outcome A, otherwise B.

Let us write down  $q(r_i)$ , the probability of following the path  $r_i$  under this biased system. Let

$$\theta(X_k) = \begin{cases} \beta & \text{if } X_k \text{ has the same sign as } \sum_{j=1}^{k-1} X_j \\ \alpha & \text{if } X_k \text{ has opposite sign to } \sum_{j=1}^{k-1} X_j \end{cases} .$$

(Similarly  $\theta(Y_k)$  ,  $\theta(Z_k)$ ) .

It is easily seen then that

$$q(r_i) = \prod_{k=1}^{100} \left[ 2g(X_{ki}) \theta(X_{ki}) \right] \left[ 2g(Y_{ki}) \theta(Y_{ki}) \right] \left[ 2g(Z_{ki}) \theta(Z_{ki}) \right] \cdot 10^{-9} .$$

Then

$$\frac{p(r_i)}{q(r_i)} = 2^{-300} \prod_{k=1}^{100} \frac{1}{\theta(X_{ki})\theta(Y_{ki})\theta(Z_{ki})} = 2^{-300} \alpha^{-N_i} \beta^{N_i-300}$$

where  $N_i$  is the number of times the auxiliary game yielded the outcome A.

We can rewrite this last result as

$$\frac{p(r_i)}{q(r_i)} = \left( \frac{1}{2\alpha} \right)^{N_i} \left( \frac{1}{2\beta} \right)^{300-N_i} .$$

The computation of this factor adds little complexity to the IBM work; we merely keep a tally on  $N_i$ , the number of outcomes A, and read the weight factor  $p(r_i)/q(r_i)$  from a table. The extra IBM time involved in





using this importance sampling scheme is only about 10% or 15% .

As a preliminary experiment, mainly to test the qualitative features of the method, 65 hydrogen random walks with importance sampling were performed.  $\lambda_1$  was computed by means of equations (10) and (21), with  $t_1 = 9$  and  $t_2 = 16$ . The results seem quite promising.

One point concerning our treatment of the data, however, should be mentioned at the outset. The possibility exists that we may sample a path whose weight factor  $p(r_i)/q(r_i)$  is exceptionally large because the auxiliary game yielded the outcome B improbably often. If we sample enough paths, the effect of such aberrations will be negligible (unless, of course, the importance scheme was incorrectly chosen and virtually neglects significant regions of the sample space, which make themselves felt only through such aberrations). However, if we sample only a few paths and happen to include one with a very large weight, the average value  $E_q \left\{ \frac{fp}{q} \right\}$  is determined largely by the single path with high weight, and the rest of the samples are useless. It seems reasonable, in a small sample, to ignore such a path (certainly so if we know, as in the present case, that it represents an unimportant region of the sample space). Consequently, we have neglected a path which occurred with a weight factor 25 times greater than that of any other path, and 100 times greater than that of all but three other paths. These other three paths could have been neglected, with almost no change in our results. (What we have done is statistically quite justifiable, and corresponds to the rejection of measurements which fall several standard



deviations from the mean; the measurements may be correct, but the sample is not large enough to give them meaning.) The following discussion refers to 64 random walks, omitting the one with the large weight.

On the basis of 64 hydrogen random walks with importance sampling,  $\lambda_1 = .518$  (true value .500). We recall from Section IV that three groups of 900 random walks without importance sampling yielded the values  $\lambda_1 = .50, .36, \text{ and } .29$  respectively. The question immediately arises of whether the accurate result yielded by importance sampling may not be mere luck, like the result for the first 900 unbiased walks. To investigate this possibility, the 64 walks were divided into arbitrary subgroups, and  $\lambda_1$  was recomputed on the basis only of the data in each subgroup. If our result were due to luck, the computed values of  $\lambda_1$  should vary greatly from one subgroup to another. Actually, a high degree of consistency is observed in the values of  $\lambda_1$  so obtained. For example, a division of the walks into four arbitrary groups of 16 yielded the values

$$\lambda_1 = 0.49 \quad , \quad 0.39 \quad , \quad 0.58 \quad , \quad 0.53$$

respectively. The corresponding computation over six arbitrary groups each consisting of 300 unbiased random walks, yields

$$\lambda_1 = 0.29 \quad , \quad 0.40 \quad , \quad 0.38 \quad , \quad 0.47 \quad , \quad 0.49 \quad , \quad 0.35$$

respectively. We can therefore safely state that groups of 16 random walks, biased according to our scheme, give at least as good an estimate of  $\lambda_1$  (by the mean value method) as do groups of 300 unbiased walks.



Importance sampling is also applicable to the "slope method" of Section V. In constructing a histogram for unbiased walks, we plotted  $x$  versus the number of observed paths  $r_i$  for which  $x < F(r_i) \leq x + \Delta$ . If enough paths are observed, this latter quantity is proportional to  $\sum_{x < F(r_i) \leq x + \Delta} p(r_i)$  which is the same as  $\sum_{x < F(r_i) \leq x + \Delta} \frac{p(r_i)}{q(r_i)} q(r_i)$ . But the latter quantity can be estimated by the biased walks, and is in fact proportional to the sum of the weights  $p(r_i)/q(r_i)$  of the paths yielding values of  $F(r_i)$  in the range  $(x, x + \Delta)$ . Thus, we can plot a histogram for the unbiased walks on the basis of data obtained from biased walks; for a fixed number of walks, the biased scheme does not give a good picture of the entire histogram, but yields an accurate picture of the tail, where most of the data are concentrated. Such a histogram was plotted on the basis of the 64 biased walks; there was insufficient data, however, to define anything more than the order of magnitude of the slope of a logarithmic histogram, which yielded  $\lambda_1 \sim .4$ .

With very little extra computation, the biased hydrogen walks furnish us with data for biased harmonic oscillator random walks. We recall that the lowest eigenvalue of the harmonic oscillator is computed from the quantity  $E \left\{ e^{-\int_0^t x^2(\tau) d\tau} \right\}$ , which we estimate by random walks. The walks which spend much time near the origin contribute the larger values of this exponential; hence, the biasing scheme for hydrogen, which stresses these same paths, might be expected to be useful for the harmonic oscillator. On the basis of 100 biased harmonic oscillator random walks, we



obtain  $\lambda_1 = 0.81$  (true value 0.71). Division of the data into ten subgroups, each of ten walks, yields

$$\begin{array}{cccccc} \lambda_1 = & 0.66 & 0.85 & 0.86 & 0.78 & 0.66 \\ & 0.85 & 1.15 & 0.96 & 0.81 & 0.71 \end{array}$$

respectively, for each of the subgroups. In this case, our final result for  $\lambda_1$  is less accurate than that obtained from unbiased walks, but our subgroups are more consistent.

One can think of many other types of importance sampling, e.g. where the strength of the biasing is a function of position, or where the biasing scheme is successively improved by some sequential procedure. Even with the present biasing scheme, lowest eigenvalue of helium should be obtainable to within 2% in two hours' time on the SWAC. The authors will welcome correspondence concerning unsolved potentials, to which the sampling method may be applicable.

---

The authors wish to express their thanks to Dr. Wolfgang Wasow, of the National Bureau of Standards, whose valuable comments brought to light many obscurities in the original manuscript; and to Dr. Everett C. Yowell, of the National Bureau of Standards, who not only supervised the IBM work infallibly, but reduced card jams as fast as the junior author could create them.

3/27/52





APPENDIX A

COMPUTATION OF DISCRETIZATION ERROR FOR THE HARMONIC OSCILLATOR

The discretization error  $K_{x^2}$  was defined in Section III as

$$K_{x^2} = \frac{\left\langle \exp \left\{ -\frac{1}{n^2} \sum_{k=1}^{nt} S_k^2 \right\} \right\rangle - \left\langle \exp \left\{ -\int_0^t x^2(\tau) d\tau \right\} \right\rangle}{\left\langle \exp \left\{ -\int_0^t x^2(\tau) d\tau \right\} \right\rangle}$$

We define

$$I(n,t) = \left\langle \exp \left\{ -\frac{1}{n^2} \sum_{k=1}^{nt} S_k^2 \right\} \right\rangle .$$

We shall compute  $I(n,t)$  exactly. Since, by equation (5a)

$$\lim_{n \rightarrow \infty} I(n,t) = \left\langle \exp \left\{ -\int_0^t x^2(\tau) d\tau \right\} \right\rangle ,$$

our calculation of  $I(n,t)$  will yield the exact value of the Wiener integral as a byproduct. Then we can compute  $K_{x^2}$ .

We recall that  $S_k - S_{k-1}$  is a Gaussian random variable with mean 0 and variance 1. Therefore

$$\begin{aligned} \text{Prob} \left\{ x_1 < S_1 < x_1 + dx_1, x_2 < S_2 < x_2 + dx_2, \dots, x_{nt} < S_{nt} < x_{nt} + dx_{nt} \right\} = \\ (2\pi)^{-\frac{nt}{2}} \exp \left\{ -\frac{1}{2} \left[ x_1^2 + (x_2 - x_1)^2 + (x_3 - x_2)^2 + \dots + (x_{nt} - x_{nt-1})^2 \right] \right\} \\ dx_1 dx_2 \dots dx_{nt} . \end{aligned}$$

Consequently,



$$I(n,t) = \left\langle \exp \left\{ -\frac{1}{n^2} \sum_{k=1}^{nt} s_k^2 \right\} \right\rangle$$

$$= (2\pi)^{-nt/2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{n^2} \sum_{k=1}^{nt} x_k^2 - \frac{1}{2} \left[ x_1^2 + (x_2 - x_1)^2 + \dots + (x_{nt} - x_{nt-1})^2 \right] \right\} dx_1 dx_2 \dots dx_{nt}$$

Integrals of the form

$$I = (\pi)^{-m/2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-i \sum_{j=1}^m a_{ij} x_i x_j} dx_1 \dots dx_m$$

(where  $a_{ij}$  is symmetric and positive definite) are easily evaluated by

transformation to principal axes. By a pure rotation of the form

$\vec{y} = B \vec{x}$  (where  $\left| \frac{\partial(y_1, \dots, y_m)}{\partial(x_1, \dots, x_m)} \right| = \det B = 1$ ) we can bring I into the form

$$I = (\pi)^{-m/2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-i \sum_{i=1}^m \lambda_i x_i^2} dx_1 \dots dx_m = \frac{1}{\sqrt{\begin{vmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_m \end{vmatrix}}}$$

The  $\lambda$ 's are the elements of the diagonal matrix  $\Lambda = B A B^{-1}$ .

Hence

$$I = (\det \Lambda)^{-1/2} = \left( \det (a_{ij}) \right)^{-1/2}$$

Applying this result to  $I(n,t)$ , we obtain







Utilizing the identities

$$\cosh (k + 1) \theta + \cosh (k - 1) \theta = 2 \cosh \theta \cosh k \theta$$

$$\sinh (k + 1) \theta + \sinh (k - 1) \theta = 2 \cosh \theta \sinh k \theta$$

we can write

$$D_k (a,b) = A \cosh k \theta + B \sinh k \theta$$

where

$$\cosh \theta = a/2 .$$

To fit the starting conditions on  $D_0$  and  $D_1$  we set

$$A = 1 \quad B = \frac{2b-a}{\sqrt{a^2-4}} .$$

Thus

$$I(n,t) = \left[ \cosh (n t \theta) + \frac{1}{n\sqrt{2+1/n^2}} \sinh (n t \theta) \right]^{-1/2}$$

where  $\cosh \theta = 1 + 1/n^2$ . Solving for  $e^\theta$ , we obtain

$$e^\theta = 1 + 1/n^2 + \frac{1}{n} \sqrt{2 + 1/n^2} = 1 + \frac{\sqrt{2}}{n} + \frac{1}{n^2} + O(n^{-3}) .$$

As  $n \rightarrow \infty$ ,  $e^{n t \theta} \rightarrow e^{t\sqrt{2}}$ . Therefore

$$\lim_{n \rightarrow \infty} I(n,t) = \left\langle \exp \left\{ - \int_0^t x^2(\tau) d\tau \right\} \right\rangle = \left[ \cosh t \sqrt{2} \right]^{-1/2} .$$





In estimating  $K_x^2$ , we shall confine ourselves to the case  $t \gg 1$ ,  $t/n \ll 1$ . This is the case of interest for the harmonic oscillator, where we took  $t = 5$ ,  $n = 100$ . Therefore we neglect terms involving the factor  $e^{-t\sqrt{2}}$  compared with terms involving  $e^{t\sqrt{2}}$ . Thus

$$\left\langle \exp \left\{ - \int_0^t x^2(\tau) d\tau \right\} \right\rangle \sim e^{\frac{-t\sqrt{2}}{2}} \sqrt{2}$$

$$I(n,t) \sim e^{-\frac{nt\theta}{2}} \left[ 1 + \frac{1}{n\sqrt{2+1/n^2}} \right]^{-1/2} \sqrt{2}$$

We shall expand the difference  $e^{\frac{-t\sqrt{2}}{2}} - \frac{I(n,t)}{\sqrt{2}}$  in powers of  $1/n$ , retaining the term of lowest order. It is convenient to deal with logarithms at this point

$$\begin{aligned} \log \frac{I(n,t)}{\sqrt{2}} &= \frac{-nt\theta}{2} - \frac{1}{2} \frac{1}{n\sqrt{2}} + o(n^{-2}) \\ \theta = \log e^\theta &= \left( \frac{\sqrt{2}}{n} + \frac{1}{n^2} \right) - \frac{\left( \frac{1}{n^2} + \frac{\sqrt{2}}{n} \right)^2}{2} + o(n^{-3}) \\ &= \frac{\sqrt{2}}{n} + o(n^{-3}) \end{aligned}$$

The vanishing of terms of order  $n^{-2}$  in the expression for  $\theta$  is significant, since it implies that the expression for  $\log I(n,t)$  contains no error term of order  $t/n$ . This is the computational reflection of our qualitative conjecture that the error, to first order, involves only a shifting of the time origin, rather than a dilation of the time scale.



We obtain, then

$$\log \frac{I(n,t)}{\sqrt{2}} = \frac{-t\sqrt{2}}{2} - \frac{1}{2\sqrt{2}n} + o(t/n^2) ,$$
$$I(n,t) = e^{\frac{-t\sqrt{2}}{2}} \left[ 1 - \frac{1}{2\sqrt{2}n} + o(t/n^2) \right] \sqrt{2}$$

Hence

$$K_{x^2} = \frac{1}{2\sqrt{2}n} + o(t/n^2) .$$



APPENDIX B

FURTHER STATISTICAL PROPERTIES OF THE RANDOM WALKS

Since many properties of the hydrogen random walks are predictable analytically, it seemed worthwhile to check the experimental results against some of these predictions.

The most significant result obtained, and one which may represent a source of error, concerned the distribution of the end points  $r(1)$  of the hydrogen walks. We recall

$$r(1) = .1 \sqrt{\left(\sum_{k=1}^{100} X_k\right)^2 + \left(\sum_{k=1}^{100} Y_k\right)^2 + \left(\sum_{k=1}^{100} Z_k\right)^2} .$$

Since  $\left(\sum_{k=1}^{100} X_k\right)$ , etc., are Gaussian with mean 0 and variance 100, then  $[r(1)]^2$  has the same distribution as  $X^2 + Y^2 + Z^2$  where  $X$ ,  $Y$ , and  $Z$  are independent Gaussian random variables with mean 0 and variance 1.

Hence

$$\begin{aligned} \text{Prob} \left\{ r(1) < \alpha \right\} &= (2\pi)^{-3/2} \int \int \int_{x^2+y^2+z^2 < \alpha^2} e^{-\frac{x^2+y^2+z^2}{2}} dx dy dz \\ &= (2\pi)^{-3/2} \int_0^\alpha e^{-\frac{r^2}{2}} 4\pi r^2 dr \\ &= \text{erf } \alpha - \frac{2\alpha}{\sqrt{2\pi}} e^{-\frac{\alpha^2}{2}} \equiv F(\alpha) \end{aligned}$$

where the last line is obtained through an integration by parts.



If we tabulate the end points  $r(1)$  of the first 900 random walks, the expected number of walks with  $r(1) < \alpha$  is  $900 F(\alpha)$ . Let us define

$$S(\alpha) = \text{number of walks with } r(1) < \alpha \text{ (empirical) } .$$

It is easy to show that

$$\left\langle \left[ S(\alpha) - 900 F(\alpha) \right]^2 \right\rangle \equiv \sigma^2(\alpha) = 900 F(\alpha)(1 - F(\alpha)) .$$

Furthermore, the probability that  $|S(\alpha) - 900F(\alpha)| < \begin{Bmatrix} \sigma(\alpha) \\ 2\sigma(\alpha) \\ 3\sigma(\alpha) \end{Bmatrix}$  is  $\begin{Bmatrix} .68 \\ .95 \\ .997 \end{Bmatrix}$ .

Hence, discrepancies of more than  $2\sigma(\alpha)$  between the empirical and theoretical distributions are to be regarded with suspicion. We compare the two distributions in the table below.

DISTRIBUTION OF END POINTS OF 900 HYDROGEN WALKS

$\alpha$	$F(\alpha)$	$900 F(\alpha)$	$S(\alpha)$	$\sigma(\alpha)$	$\left  \frac{S(\alpha) - 900 F(\alpha)}{\sigma(\alpha)} \right $
0.5	.030	27	28	5.1	0.2
1.0	.197	177	185	11.9	0.7
1.5	.478	430	427	15.0	0.2
2.0	.739	665	645	13.2	1.5
2.5	.900	810	753	9.0	6.3
3.0	.971	874	836	5.0	7.6
3.5	.994	894	867	2.3	11.7
4.0	.999	899	880	1.0	19.0

The tail of the empirical distribution  $S(\alpha)$  evidently falls off much less rapidly than that of a true Gaussian. The magnitude of the





resultant error in  $\lambda_1$ , and possible remedies, have not been investigated, mainly because the SWAC will not employ the Rand Corporation Gaussian deviates. Qualitatively, the effect of the non-Gaussian behavior is to increase the values of  $r(\tau)$ , and thus to decrease the values of  $\int_0^1 d\tau/r(\tau)$ . This may provide a partial explanation of another discrepancy between theory and observation, discussed below.

Comparisons were also made between the theoretical and experimental mean values of integrals arising in the hydrogen and helium computations. Since the operations of averaging and summing (or integrating) commute, we have

$$\left\langle \int_0^1 d\tau/r(\tau) \right\rangle = \int_0^1 \left\langle \frac{1}{r(\tau)} \right\rangle d\tau .$$

But

$$\left\langle \frac{1}{r(\tau)} \right\rangle = (2\pi\tau)^{-3/2} \int_0^\infty e^{-r^2/2\tau} 4\pi \frac{r^2}{r} dr = \sqrt{\frac{2}{\pi}} \sqrt{\frac{1}{\tau}} .$$

Hence

$$\left\langle \int_0^1 d\tau/r(\tau) \right\rangle = \sqrt{\frac{2}{\pi}} \int_0^1 \tau^{-1/2} d\tau = \frac{4}{\sqrt{2\pi}} = 1.59 .$$

The empirical value for 1800 samples was 1.43.

In order to appreciate the significance, if any, of this discrepancy, one must know the variance  $\sigma_0^2$  of a single observation of  $\int_0^1 d\tau/r(\tau)$ . Knowing  $\sigma_0$ , we then know the variance  $\sigma^2$  of the average of 1800 observations, since  $\sigma^2 = \frac{\sigma_0^2}{1800}$ .



By definition

$$\sigma_0^2 = \left\langle \left( \int_0^1 d\tau/r(\tau) - \frac{4}{\sqrt{2\pi}} \right)^2 \right\rangle = \left\langle \left( \int_0^1 d\tau/r(\tau) \right)^2 \right\rangle - \frac{8}{\pi}$$

If we think of  $\int_0^1 d\tau/r(\tau)$  as a Riemann sum of many non-independent random variables  $\frac{1}{r(\tau_i)}$ , it is easily seen that

$$\begin{aligned} \left\langle \left( \int_0^1 d\tau/r(\tau) \right)^2 \right\rangle &= \frac{2}{(2\pi)^3} \int_0^1 \int_0^{\tau_2} \frac{1}{\tau_1^{3/2}} \frac{1}{(\tau_1 - \tau_2)^{3/2}} d\tau_1 d\tau_2 \times \\ &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{x_1^2 + y_1^2 + z_1^2}} \frac{1}{\sqrt{x_2^2 + y_2^2 + z_2^2}} e^{-\frac{x_1^2 + y_1^2 + z_1^2}{2\tau_1}} e^{-\frac{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}{2(\tau_2 - \tau_1)}} dx_1 \dots dz_2 \\ &= 4 \log 2^1 \end{aligned}$$

Thus

$$\sigma_0^2 = 4 \log 2 - \frac{8}{\pi} = .22 \quad (\text{In good agreement with the empirically observed } \sigma_0^2 = .19.)$$

$$\sigma = \frac{\sigma_0}{\sqrt{1800}} = .011$$

---

<sup>1</sup>Evaluation of the integral is left as an exercise for the reader.



The difference of .16 between the calculated and observed value of  $\int_0^1 d\tau/r(\tau)$  is therefore equal to  $15\sigma$ , which is far beyond the limits of statistical inaccuracy. Dr. R. P. Feynman first called attention to a major systematic error, and gave the following method for estimating it.

What we actually sample is not the integral  $\int_0^1 d\tau/r(\tau)$ , but rather the Riemann sum  $.01 \sum_{k=1}^{100} \frac{1}{r(\frac{k}{100})}$ . Thus, the average to be considered is that of the Riemann sum, which is somewhat smaller than that of the integral (the error encountered here is the "discretization error" discussed in Section III).

Using the result already obtained for  $\left\langle \frac{1}{r(\tau)} \right\rangle$ , we have

$$\left\langle .01 \sum_{k=1}^{100} \frac{1}{r(\frac{k}{100})} \right\rangle = \left( \sqrt{\frac{2}{\pi}} \right) \left( .01 \sum_{k=1}^{100} \frac{1}{\sqrt{\frac{k}{100}}} \right).$$

It is geometrically evident that a fairly good estimate of the Riemann sum in parentheses is

$$\int_{.005}^{1.005} \frac{d\tau}{\sqrt{\tau}} = 2 \left[ 1.003 - .071 \right] = 1.86.$$

Thus  $\left\langle .01 \sum_{k=1}^{100} \frac{1}{r(\frac{k}{100})} \right\rangle = 1.48$  (compared with 1.59 for  $\left\langle \int_0^1 d\tau/r(\tau) \right\rangle$ )

and our observed value 1.43 is in error by  $5\sigma$ , instead of  $15\sigma$ . The remaining error may, perhaps, be accounted for by the non-Gaussian behavior of the tail of the distribution of  $r(\tau)$  (discussed above), and the consequent depression of values of  $\int_0^1 d\tau/r(\tau)$ .



One can also calculate the average value of the integral of the helium integration term.

$$\left\langle \int_0^1 \frac{d\tau}{|r_1(\vec{r}) - r_2(\vec{r})|} \right\rangle = \frac{1}{(2\pi)^3} \int_0^1 \frac{d\tau}{\tau^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}} e^{-\frac{x_1^2+y_1^2+z_1^2}{2\tau}} e^{-\frac{x_2^2+y_2^2+z_2^2}{2\tau}} dx_1 \dots dz_2$$

$$= \frac{2}{\sqrt{\pi}} = 1.13 \quad .$$

The observed value was .93. The discrepancy is probably to be explained along the same lines as above.





APPENDIX C

STATISTICAL METHOD FOR FINDING THE LOWEST EIGENFUNCTION

The reasoning in Section II leads easily to a statistical method for finding  $\Psi_1(x)$ , the lowest eigenfunction. From Section II we have

$$P(x, t) dx = \sum C_j \Psi_j(x) e^{-\lambda_j t} dx .$$

Furthermore

$$P(x, t) dx = \text{Prob} \left\{ \begin{array}{l} \text{particle survives till time } t \text{ and is in } (x, x + dx) \text{ at time } t \\ \text{particle picks a path such that } x < x(t) < x + dx \\ \text{particle survives till } t, \text{ given that } x < x(t) < x + dx \end{array} \right\} =$$

But

$$\text{Prob} \left\{ \begin{array}{l} \text{particle picks a part such that} \\ x < x(t) < x + dx \end{array} \right\} = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} dx ;$$

and the probability of survival till  $t$ , given that  $x < x(t) < x + dx$ , is just the average of the survival probability  $\exp \left\{ - \int_0^t V(x(\tau)) d\tau \right\}$  over the set of paths such that  $x < x(t) < x + dx$ . Hence

$$P(x, t) dx = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} dx \left\langle \exp \left\{ - \int_0^t V(x(\tau)) d\tau \right\} \middle| x < x(t) < x + dx \right\rangle .$$

The conditional average in brackets can be found empirically by subdividing the  $x$ -axis into short intervals, and computing the average of



$\exp \left\{ - \int_0^t v(x(\tau)) d\tau \right\}$  over the paths ending in each interval.<sup>1</sup> By sampling  $P(x,t)dx$  for large  $t$ , we can thus determine  $\Psi_1(x)$  (up to a normalizing factor). By considering several different values of  $t$  and waiting until the rate of time decay of  $P(x,t)$  becomes the same at all points  $x$ , we can ascertain when the space distribution has settled down to the correct eigenfunction. This settling-down process can be speeded by starting the walks not from a delta-function distribution, but rather from a distribution which approximates our best guess for  $\Psi_1(x)$ . Then the coefficients  $C_2, C_3, \dots$  will be small.

No experiments have yet been performed with the above scheme, which clearly requires very many random walks.

If we start our walks from a point  $x_0$  instead of the origin, then

$$P(x,t) = \sum \Psi_j(x_0) \Psi_j(x) e^{-\lambda_j t}$$

and hence

$$\Psi_1(x_0) e^{-\lambda_1 t} = \int_{-\infty}^{\infty} P(x,t) \Psi_1(x) dx = \left\langle \exp \left\{ - \int_0^t v(x(\tau)) d\tau \right\} \Psi_1(x(t)) \right\rangle .$$

If we have an empirical estimate of  $\Psi_1$  and  $\lambda_1$  from an earlier set of walks, we may perform a new set of walks starting at  $x_0$  and use the last

---

<sup>1</sup>If  $N$  paths  $x_i(\tau)$  ( $i = 1, \dots, N$ ) are sampled, an equivalent and computationally easier estimate of  $P$  is

$$P(x,t)dx = \frac{1}{N} \sum_{x \leftarrow x_i(t) \leftarrow x+dx} \exp \left\{ - \int_0^t v(x_i(\tau)) d\tau \right\} .$$



equation to improve our guess for  $\Psi_1(x_0)$ . The usefulness of this procedure is doubtful, since it depends on the "approximate orthogonality" of our final  $\Psi_1$  to the true  $\Psi_2, \Psi_3$ , etc., and involves an inordinate amount of effort to improve our guess for  $\Psi_1$  at a single point.



Bibliography

- [1] M. Kac, On the distribution of certain Wiener functionals, Trans. Am. Math. Soc. 65 (1949) pp. 1-13.
- [2] M. Kac and M. D. Donsker, A sampling method for determining the lowest eigenvalue and the principal eigenfunction of Schrödinger's equation, Journal of Research of the National Bureau of Standards, 44 (1950) pp. 551-557.
- [3] R. Fortet, On the estimation of an eigenvalue by an additive functional of a stochastic process, with special reference to the Kac-Donsker method, to appear in Journal of Research of the National Bureau of Standards.
- [4] G. W. King, Stochastic Methods in Quantum Mechanics, Proceedings of IBM Seminar on Scientific Computation, Nov. 1949.
- [5] E. Kahn, Modification of the Monte Carlo Method, Proceedings of IBM Seminar on Scientific Computation, Nov. 1949.





# THE NATIONAL BUREAU OF STANDARDS

## Functions and Activities

The functions of the National Bureau of Standards are set forth in the Act of Congress, March 3, 1901, as amended by Congress in Public Law 619, 1950. These include the development and maintenance of the national standards of measurement and the provision of means and methods for making measurements consistent with these standards; the determination of physical constants and properties of materials; the development of methods and instruments for testing materials, devices, and structures; advisory services to Government Agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; and the development of standard practices, codes, and specifications. The work includes basic and applied research, development, engineering, instrumentation, testing, evaluation, calibration services and various consultation and information services. A major portion of the Bureau's work is performed for other Government Agencies, particularly the Department of Defense and the Atomic Energy Commission. The scope of activities is suggested by the listing of divisions and sections on the inside of the front cover.

## Reports and Publications

The results of the Bureau's work take the form of either actual equipment and devices or published papers and reports. Reports are issued to the sponsoring agency of a particular project or program. Published papers appear either in the Bureau's own series of publications or in the journals of professional and scientific societies. The Bureau itself publishes three monthly periodicals, available from the Government Printing Office: The Journal of Research, which presents complete papers reporting technical investigations; the Technical News Bulletin, which presents summary and preliminary reports on work in progress; and Basic Radio Propagation Predictions, which provides data for determining the best frequencies to use for radio communications throughout the world. There are also five series of nonperiodical publications: The Applied Mathematics Series, Circulars, Handbooks, Building Materials and Structures Reports, and Miscellaneous Publications.

Information on the Bureau's publications can be found in NBS Circular 460, Publications of the National Bureau of Standards (\$1.00). Information on calibration services and fees can be found in NBS Circular 483, Testing by the National Bureau of Standards (25 cents). Both are available from the Government Printing Office. Inquiries regarding the Bureau's reports and publications should be addressed to the Office of Scientific Publications, National Bureau of Standards, Washington 25, D. C.

