NATIONAL BUREAU OF STANDARDS REPORT

2365

"MONTE CARLO" METHODS FOR THE ITERATION OF LINEAR OPERATORS

by

J. H. Curtiss

U. S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

U. S. DEPARTMENT OF COMMERCE
Sinclair Weeks, *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.
Electromechanical Ordnance.
Ordnance Electronics.
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**

1101-10-5100 March 19, 1953 2365

"MONTE CARLO" METHODS FOR THE ITERATION OF LINEAR OPERATORS

by

J. H. Curtiss

Harvard University
The National Bureau of Standards

NBS

"MONTE CARLO" METHODS FOR THE ITERATION OF LINEAR OPERATORS

by J. H. Curtiss
Harvard University and
The National Bureau of Standards

1. Introduction. A purely formal description of the type of problem to be dealt with in this paper is as follows. Let $c = c(x)$ and $u_0 = u_0(x)$ be real-valued functions defined on a coordinate space R, which may be multidimensional. Let $L = L(f)$ be a linear transformation defined on the space of all real-valued functions f whose arguments belong to R. Required, to calculate the sequence of functions $u_1, u_2, \ldots$, defined by the recursion formula

$$(1.1) \qquad u_{N+1} = L(u_N) + c, \qquad N = 0,1,2,\ldots$$

This problem arises in many contexts in both pure and applied mathematics. Much attention has recently been focused on the numerical aspects of it by the nuclear physicists, because the recursion formula (1.1) is obtained when time-dependent diffusion and transport problems are formulated in a discrete form. The methods to be discussed here for estimating the solution are an outgrowth of a novel stochastic attack suggested during the late war by Von Nuemann and Ulam in connection with diffusion problems.* Their idea was to bypass the mathematical formulation (1.1), and set up a computing procedure with various random decisions in it which more or less closely imitated the physical phenomenon under study.

*Various practical aspects of the stochastic estimation are presented in [14].

This type of approach to distribution problems has long been known to statisticians under the name of "model sampling." The physicists have thought up a new name for it that seems likely to stick: the "Monte Carlo Method."

The formal solution of (1.1) is the truncated Neumann series

(1.2) $$u_N = c + L(c) + L^2(c) + \ldots + L^{N-1}(c) + L^N(u_0), \quad N > 0,$$

where L^K means the K-th iterate of L. Especial interest, of course, lies in the case in which the corresponding infinite series converges. If it does, it represents a function u which satisfies the equation

(1.3) $$u = L(u) + c.$$

The error estimate is then provided by

(1.4) $$u_N - u = L^N(u_0 - u).$$

Equation (1.2) shows that it makes no difference in the long run how u_0 was chosen, but Eq. (1.4) suggests that the nearer u_0 is to u, the faster the convergence will be.

The stochastic approach to estimating the solution of (1.1) will now be described in correspondingly general terms. Consider the space R over which the functions c, u_0, $u_1, \ldots,$ are defined. Let x be a point in it at which it is desired to calculate u_N. A probability distribution is now set up on each of the spaces R, $R \times R$, $R \times R \times R, \ldots,$ where $R \times R$ denotes the Cartesian product of R and R. Random variables

$Z_0, Z_1, Z_2, \ldots$ are defined on these respective product spaces. The chain of probability distributions and of random variables is such that the sequence of conditional mean values*

$v_0(x) = E(Z_0 | Z_0 = u_0(x))$ $v_1(x) = E(Z_1 | Z_0 = u_0(x))$, $v_2(x) = E(Z_2 | Z_0 = u_0(x)), \ldots$, satisfies (1.1). Since $v_0 = u_0$, it follows that $v_N(x) = u_N(x)$.

The computational problem then becomes one of calculating repeated realizations of Z_N and combining them into an appropriate statistical estimator of v_N.

The stochastic method is particularly well adapted to the case in which the value of $u_N(x)$ or $u(x)$ is to be estimated at only one point x.

In the case in which the interest lies in estimating the solution of (1.3), it is quite possible to carry the sequence $Z_1, Z_2, \ldots$ on to infinity and define a random variable Z on the infinite product-space $R \times R \times \ldots$ But this has no significance for actual practice, and curiously enough, it turns out to be theoretically disadvantageous. That arises from the fact that mean values in statistics are ordinarily defined through absolutely convergent integrals and sums. This in turn

*We shall use the symbol $Pr(a|b)$ to denote the conditional probability of the event a, given that b has occurred, the symbol $E(Y)$ to denote the mean value of the random variable Y, and the symbol $E(Y|b)$ to denote the mean value of the conditional distribution of Y, given that b has occurred.

imposes some irrelevant restrictions on L if v_{∞} is to be identified with u, at least in the cases to be considered in this paper.* Therefore the attitude here in the "steady state" situation will be that a suitably large but finite N will be chosen once and for all and held fixed during the sampling.

The mathematical material preceding Section 7 is in the main a rearrangement and mathematical formulation of known procedures, presented so as to show up their relationships. The method of error analysis proposed in Section 4 and most of the material in Sections 7, 8, 9, and 10 are believed to be new. However, a good deal of work on the Monte Carlo Method has been "published" in privately circulated, sometimes classified, reports, and one can never be quite sure of a priority under such circumstances.

A word of caution to the reader may be in order. The Monte Carlo method as a computational procedure has had its chief successes in problems which had natural stochastic bases and which were at the same time so complicated that they were inaccessible to ordinary analytic or numerical methods. This paper makes no pretense of putting the method into competition with the standard numerical practices, especially for the

*As applied to the solution of simultaneous linear equations, the methods of the present paper were anticipated by Forsythe and Leibler [10] and Wasaw [16]. Both of these papers deal directly with the infinite product space. See also Curtiss [4].

simpler type of problem for which good methods already exist. The idea here is merely to present some theory which may be of some interest for itself alone, and which unifies and clarifies certain of the Monte Carlo devices which have been proposed, and which lays the groundwork for further numerical experimentation aimed at investigating the limits of usefulness of the method for non-stochastic problems.

We propose to stop short of describing the practical computational details. Except for the process of generating the necessary random elements, they can all be supplied easily by the experts in computing. The process of correctly generating the random elements and making sure that they are appropriately random could be the subject of another article at least as long as this one. If it were well done, it might contain substantial contributions to the philosophy of probability.

2. Specialization of the problem. The foregoing introduction has been very vague as to the nature of the operator L. Actually, the stochastic method to be presented here seems to be automatically confined to operators L of a Lebesgue-Stieltjes type,

(2.1) $$L(u) = \int_R h(x,y)\, u(y)\, d_y k(x,y),$$

in which k is of bounded variation and h has integrability properties which will permit the iteration. We shall not pursue the question of generability any further, however. Instead we shall present the theory for two special cases: that in which L is an ordinary integral transform,*

(2.2) $$L(u) = \kappa \int_R h(x,y)\, u(y)\, dy,$$

and that in which L is a matrix and u is a vector.

The second case is the special case of (2.1) in which k is a step-function. Most of the exposition except for that in Section 4 will be directed toward this case. It is the more fundamental one in numerical analysis and the majority of the results can be carried over so readily to the continuous operator (2.2) that no comment on the matter will be necessary.**

It is convenient to introduce new notation for the matrix case. The space R will be thought of in this case as

*The integral can be taken in either the Riemann or Lebesque sense in the sequel.

**See Cutkosky [5].

consisting of a finite discrete set of points $x_1, x_2, \ldots, x_n$. The function $u = u(x)$ will be represented by the vector $u = (u_1, u_2, \ldots, u_n) = (u(x_1), \ldots, u(x_n))$. Similarly, we write $u_N = (u_N(x_1), u_N(x_2), \ldots, u_N(x_n)) = (u_{N1}, u_{N2}, \ldots, u_{Nn})$, $N = 0, 1, 2, \ldots$, and $c = (c(x_1), c(x_2), \ldots, c(x_n)) = (c_1, c_2, \ldots, c_n)$. Finally, the function $h(x,y)$ is represented by the matrix $H = [h(x_i, y_j)] = [h_{ij}]$.

Equation (1.1) becomes in matrix notation

$$u_{N+1} = Hu_N + c. \tag{2.3}$$

The problem connected with this equation is to calculate one or more components of the vector u_N, given H, $u_0 = (u_{01}, u_{02}, \ldots, u_{0n})$, and c.

The series solution (1.2) becomes

$$u_N = c + Hc + H^2c + \ldots + H^{N-1}c + H^Nu_0 = (I-H^N)(I-H)^{-1}c + H^Nu_0, \quad N = 1, 2, \ldots . \tag{2.4}$$

Here the symbol I stands for the unit matrix. Of course the third member of the equation can be written down only if $I - H$ is non-singular.

It is well known that the necessary and sufficient condition for $\lim_{N \to \infty} H^N = 0$ is that all the eigenvalues of H (that is, the roots of the determinantal equation $|\lambda I - H| = 0$) must lie within the unit circle of the complex plane.* If this

*See [13, pp. 97-98].

is the case—and we shall always assume that it is whenever we are discussing the situation as $N \rightarrow \infty$ — then, $(I-H)^{-1}$ exists, and $u = \lim_{N \rightarrow \infty} u_N = (I - H)^{-1}c$, which satisfies the linear equations

(2.5) $$u = Hu + c.$$

3. <u>The solution of</u> $Au = b$.

We digress for a moment here to note the relationship between (2.3), (2.4), and (2.5), and the important problem of solving the system of linear equations

(3.1) $$Au = b,$$

where $b = (b_1, \ldots, b_n)$ is an arbitrary vector. Choose the matrix H, and also a new one M, so that $H + MA = I$, and choose $c = Mb$. Then (2.5) reduces to $MAu = Mb$. If M is non-singular, then this system is precisely equivalent to (3.1) in the sense that each solution u of (2.5) is a solution of (3.1), and vice versa. If there is more than one solution to (3.1), then A is singular, and in this case it is easily checked that one of the eigenvalues of $H = I - MA$ is unity. This means that u_N in (2.3) and (2.4) cannot converge to the solution. We shall therefore exclude this case and assume that A is non-singular.

With M and A both non-singular, and with H having its eigenvalues all inside the unit circle, (2.3) becomes what is known in the theory of linear algebraic systems as a stationary linear iterative process, or Wittmeyer process, for solving $Au = b$. There are obviously an infinite number of ways of choosing M and H so that the conditions are fulfilled. One standard method is to split up A into the difference of two matrices V and W, where V is easily inverted. That is, let $A = V - W$. Then take $M = V^{-1}$, $H = V^{-1}W$. It is easily checked that $H + MA = I$, and it is not hard to arrange things so that the eigenvalues of H are all sufficiently small in modulus.*

The calculation of A^{-1}, if that is the problem, can of course be accomplished by specializing b appropriately. Alternatively, the Wittmeyer process can be directly adapted to this problem by replacing c by M, and u_0, u_1, u_2,..., by square matrices $\bar{U}_0$, $\bar{U}_1$, $\bar{U}_2$,... .

The importance of these remarks in the present context is merely that they show that with a little preliminary preparation any system $Au = b$ with a non-singular matrix can be solved by the Monte Carlo methods to be described hereinafter.

---

*A method due to Morris (see [7, pp. 132-133]) takes V as a triangular matrix obtained by replacing all elements of A above the principal diagonal by zeros. The iterative method of Jacobi takes V as a principal diagonal matrix whose diagonal is that of A. The well-known method of Seidel is another Wittmeyer process. The so-called relaxation method is not a linear process. So far as the author is aware, the extension of Monte Carlo methods to non-linear processes has not yet been accomplished, and may be impossible. For an interesting and scholarly classification of non-stochastic methods of solving linear equations, see Forsythe [9].

4. <u>The stochastic methods for</u> N = 1. The basic idea in the Monte Carlo attack on problems (1.1) and (1.3) will now be described for the case of the zero-th iteration of the recursion formula. To avoid dealing with a trivial problem, we present the ideas here with L taken as the continuous operator (2.2) rather than as a discrete operator.

The problem then is to estimate the numerical value of

$$u_1(x) = \int_R h(x,y)\ u_0(y)\ dy + c(x).$$

The function c(x) will play no significant role in the present discussion, nor will κ, so we confine ourselves to the estimation of

$$I(x) = \int_R h(x,y)\ u(y)\ dy.$$

Now let functions z(x,y) and p(x,y) be chosen so that*

(1) $zp = h,$

(2) $p \geqq 0,$

(3) $\int_R p(x,y)\ dy = 1.$

Then for each x, p may be regarded as a probability density on R. Let X be a vector random variable with the probability distribution defined by p. Consider the random variable $Z = z(x, X)\ u(X)$. Clearly

$$E(Z) = \int_R z(x,y)\ u(y)\ p(x,y)\ dy$$
$$= \int_R h(x,y)\ u(y)\ dy = I(x).$$

*The stochastic mechanism which is being arranged here has been proposed by various writers. See Kahn [11], and also [14].

A$ an example, let R be one dimensional, and let

$$h(x,y) = \begin{cases} e^{-xy}, & y \geqq 0 \\ 0 \quad , & y < 0 \end{cases}$$

Then $I(x) = \int_0^{\infty} e^{-xy} u(y)\,dy$, the Laplace transform of u(y). A natural choice for p(x,y) would be the Pearson Type III density function, e^{-xy}/x. Then z(x,y) = x, and Z = x u(X).

One of the standard statistical procedures for estimating the mean of the distribution of a random variable Z is to make ν independent observations on the random variable and then take their average $\bar{Z}$ as the estimator. If the standard deviation* σ of the distribution of the random variable is finite, then the probability distribution of $\bar{Z}$ is asymptotically Gaussian, or "normal", with (of course) the same average, and a standard deviation equal to $\sigma/\nu^{1/2}$. (This is a special case of the famous Central Limit Theorem; see Cramer [3, pp.215-217]. The approximation is usually very close for $\nu \geqq 30$.)

More than 99 per cent of a normal distribution lies within the interval: [mean $\pm$ 3 $\times$ standard deviation]. From this we can easily calculate the sample size ν theoretically necessary to achieve with this level of certainty a given statistical accuracy in using $\bar{Z}$ as an estimator of I(x).

*The standard deviation of Y is defined to be $\left\{E[(Y-E(Y))^2]\right\}^{1/2} = \left\{E(Y^2) - [E(Y)]^2\right\}^{1/2}$. The square of the standard deviation is called the variance; we shall write it as Var(Y).

Let us say that we wish to be almost sure that $\bar{Z}$ will lie in the interval $I \pm \Delta I$. That is, we want to arrange things so that $\Pr(|\bar{Z} - I| \leqq \Delta I) \geqq .99$. This means that $\Delta I \geqq 3\sigma/\nu^{1/2}$, from which we get

$$(4.1) \qquad \nu \geqq \frac{9\sigma^2}{(\Delta I)^2} = \frac{9\ \mathrm{Var}(Z)}{(\Delta I)^2} \quad .$$

The formula shows that if ΔI is to be small, say 0.005, then unless Var(Z), is also small, ν will be well up in the hundreds of thousands. The two redeeming traits of the stochastic method are that (1) the sampling error and necessary sample size are independent of dimensionality and (2) they are independent of how locally smooth the integrand of I(x) is. (It will be recalled that the error terms in the standard quadrature formulas involve high-order derivatives.)

But it is clearly worth while to consider methods of reducing the sample size. The statisticians have a number of devices for increasing the accuracy of sampling surveys, and almost all of them are applicable here. In the present paper we shall study only the procedures known as "sampling with probabilities in proportion to size," or "importance sampling".* They take advantage of the fact that we have an infinite number of ways to choose z and p for any given problem, and an astute choice may decrease the variance by a surprising amount.

*See Deming [6 pp. 92-93].

The variance of Z (and we henceforth assume that it is finite) is given by the formula

$$\alpha^2 = E(Z^2) - [E(z)]^2 = \int_R z\ hu^2 dy - I^2.$$

If the answer I were known in advance, and z were chosen to be I_u (p would then have to be hu/I, which would mean that the integrand hu was non-negative), then clearly $\alpha^2 = 0$. This leads us to propose the following arrangement as a guide in choosing z and p. First choose an integrable, non-negative function p' such that the function $\epsilon = \epsilon(x,y)$ defined by

(4.2) $$hu - p' = p'\epsilon$$

is as small as possible in absolute value. The function p' is, moreover, to be chosen so that the integral $J = \int_R p'\,dy$ can be obtained numerically without too much trouble. Let $p = p'/J$ and let $z = h/p = J(1 + \epsilon)/u$. Then it is easily seen that z and p satisfy the conditions (1)-(3) listed above. The estimator Z is given by $Z = J(x)\ (1 + \epsilon(x,X))$.

The variance of $\alpha + \beta Y$, where α and β are constants and Y is a random variable, is simply $\beta^2\ \mathrm{Var}(Y)$. Thus

(4.3) $$\alpha^2 = \mathrm{Var}(Z) = J^2\ \mathrm{Var}\,[\,\epsilon\ (x,X)]\ .$$

If e(x) is the least upper bound of $|\epsilon(x,y)|$ for Y on R, then it might be safe to presume that $\mathrm{Var}(c) \leq e^2/2$. (If the distribution of ϵ were rectangular with range 2e — an unfavorable case — then Var $(\epsilon) = e^2/3$. This appraisal gives us the formula

$$(4.4) \qquad \mathcal{V} \geq \frac{4.5\ J^2 e^2}{(\Delta I)^2}$$

for the sample size theoretically required to achieve an accuracy of $\pm\ \Delta I$ with at least 99 % certainty.*

It might be noted that (4.3) and (4.4) are quite independent of whether or not $|\epsilon|$ remains small for all y. If it does, then there is an implication that hu cannot go very far in the negative direction, since p' cannot be negative. These restrictions on the usefulness of the arrangement (4.2) can be circumvented to some extent, but we shall not go into the matter here.

We have been carrying the somewhat superfluous parameter x throughout the above discussion mainly to emphasize the link between quadrature and other problems. The usual quadrature problem would of course be presented with $h(x,y) = h(y)$, $u(y) \equiv 1$. But it is perhaps worthwhile to observe that if there is a parameter x in the problem, then by choosing p so that it is dependent only on x − y, a set of determinations of X can be gotten once and for all, from which the statistics for Z can be computed over and over again for as many different values of x as may be desired.

One final remark which applies to all the remaining sections of this paper as well as to the present section is this:

*In any case, certainly Var$(\epsilon) \leq e^2$

The statistical error was the only kind of error under consideration here, but of course there would be many other possible sources of error in the actual numerical applications. For example, there would be possible round-off errors, mistakes, and systematic errors of one kind or another. In accordance with the resolution expressed at the end of the Introduction, none of these non-statistical errors will be discussed.

5. <u>The Case $N > 1$</u>. Suppose that the method of the preceding section had been applied to the trivial problem of estimating

$$u_{1i} = \sum_{j=1}^{n} h(x_i,x_j)u(x_j) = \sum_{j=1}^{n} h_{ij}u_j = Hu]_i \ ,$$

for some fixed i, where the notation is that of Section 3*. The procedure would have been to select numbers $z_{ij} = z(x_i,x_j)$ and $p_{ij} = p(x_i,x_j)$ such that

(1) $z_{ij}\, p_{ij} = h_{ij}$, $\quad i,j, = 1,\ldots,n,$

(2) $p_{ij} \geqq 0$, $\quad i,j, = 1,\ldots,n,$

(3) $\sum_j p_{ij} = 1$, $\quad i = 1,\ldots,n;$

then to let X be a vector random variable with the probability distribution given by $Pr(X = x_j) = p_{ij}$, and to form the random variable $Z = z(x_i,X)u(X)$ as before. Its mean value would be

$$E(Z) = \sum_j z(x_i,x_j)u(x_j)\ p(x_i,x_j) = u_{1i}.$$

*By $b]_i$, where b is a vector, we mean the i-th component of b.

We shall now show how to extend this stochastic scheme to the iterations of H.

We continue to define z_{ij} and p_{ij} as above. Consider a random walk on R defined as follows: the starting point X_0 is a random variable with a probability distribution given by $p_i = p(x_i)$, where $p_i > 0$, $\sum_i p_i = 1$, but p_i is otherwise arbitrary.* Thereafter the successive positions or states visited by the random walk are random variables $X_1, X_2, \ldots$, whose distributions are given by the formula $Pr(X_{k+1} = x_j \; X_k = x_i) = p_{ij}$, $i,j = 1,\ldots,n$.

These directions have the effect of unambiguously specifying a probability distribution on the product space $\overbrace{R \times R \times \ldots \times R}^{N}$ for each N. The typical point is assigned the probability $p_i p_{ij_1} p_{j_1 j_2} p_{j_2 j_3} \ldots p_{j_{n-1} j}$. The chain of random variables $X_0, X_1, X_2, \ldots$ so defined is a simple Markov chain.**

Now define a new chain of random variables as follows:

$$Z_0 = Z_0[u] = u(X_0)$$
$$Z_1 = Z_1\,[u,z] = z(X_0,X_1)\; u(X_1)$$
$$Z_2 = Z_2\,[u,z] = z(X_0,X_1)\; z(X_1,X_2)\; u(X_2)$$
$$\vdots$$
$$Z_N = Z_N\,[u,z] = z(X_0,X_1)\ldots z(X_{N-1},X_N)\; u(X_N).$$
$$\vdots$$

We shall call $Z_0, Z_1, Z_2, \ldots,$ an m-chain (m for multiplicative).

*This distribution does not play an intrinsic role in our discussion and is introduced only for logical completeness.

**See Feller [8, Chap. 15]. Actually our specifications uniquely assigns a probability distribution to the infinite product space $R \times R \times \ldots$ (see [12,pp 28-33]),but we shall not make use

Let the vectors v_0, $v_1, \ldots,$ be defined as follows:

$$v_{oi} = v_o(x_i;u) = u(x_i),$$

$$v_{Ni} = v_N(x_i;u,z) = E(Z_N[u,z] \mid X_0 = x_i), \quad N=1,2,\ldots$$

Consider now a typical path of the random walk represented by the sequence of random variables $X_0, X_1, \ldots, X_N$. Let this path be $x_i, x_{j_1}, x_{j_2}, \ldots, x_{j_{N-1}}, x_j$. The conditional probability that the random walk actually takes this path is $p_{ij_1} p_{j_1 j_2} \cdots p_{j_{N-1} j}$. Therefore by the definition of the concept of mean value in the theory of probability,

$$v_{Ni} = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_{N-1}=1}^{n} \sum_{j=1}^{n} z_{ij_1} z_{j_1 j_2} \cdots$$

$$z_{j_{N-1} j} u_j p_{ij_1} p_{j_1 j_2} \cdots p_{j_{N-1} j}$$

$$= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j=1}^{n} h_{ij_1} h_{j_1 j_2} \cdots h_{j_{N-1} j} u_j = H^N u]_i.$$

Thus

(5.1a) $\qquad v_{N+1} = Hv_N, \qquad N = 0,1,2,\ldots,$

(5.1b) $\qquad v_N = H^N u, \qquad N = 0,1,2,\ldots$.

The question of estimating u_N in the recursion relation (2.3) can now be resolved rather easily. It is clear that

$$u_{Ni} = (c + Hc + \ldots + H^{N-1} c + H^N u_0)]_i$$

$$= v_0(x_i,c) + v_1(x_i;c,z) + \ldots$$

$$+ v_{N-1}(x_i;c,z) + v_N(x_i;u_0,z)$$

$$= E(Z_0[c] + Z_1[c,z] + \ldots + Z_{N-1}[c,z] + Z_N[u_0,z] \mid X_0 = x_i),$$

$$= E(\Sigma_{N-1}[c,z] + Z_N[u_0,z] \mid X_0 = x_c),$$

where $\Sigma_N[u,z] = Z_0[u] + Z_1[u,z] + \ldots + Z_N[u,z]$. This shows that if we define vectors $w_0, w_1, w_2, \ldots,$ by

$$w_{0i} = w_0(x_i;u) = u(x_i),$$

$$w_{Ni} = w_N(x_i;c,u,z) = E(\Sigma_{N-1}[c,z] + Z_N[u,z] \mid X_0 = x_i), N=1,2,\ldots,$$

then with $u = u_0$, these vectors satisfy the relations

(5.2a) $$w_{N+1} = Hw_N + c, \qquad N = 0,1,2,\ldots,$$

(5.2b) $$w_N = (I-H^N)(I-H)^{-1}c + H^N u_0, \quad N = 1,2,\ldots,$$

Furthermore, if we assume that the eigenvalues of H are in modulus all less than unity, then the vector $w_\infty = \lim_{N\to\infty} w_N$ exists and is the solution of the system $u = Hu + c$.

Thus the conditional mean values of the chain of random variables $\Sigma_{N-1} + Z_N$, $N = 1,2,\ldots$ provide a solution to our basic problem. The actual computation would consist in making a large number of realizations of the vector random variable $(X_0, X_1, \ldots, X_N)$, calculating $\Sigma_{N-1} + Z_N$ for each realization, and averaging the results, or otherwise combining them into a statistical estimator of the mean value.

Several remarks are now in order.

(1) This method of solving linear equations was proposed by Wasow[16]. He considered only the infinite product space $R \times R \times \ldots$, and was thereby forced to use the restriction that the matrix of the absolute values of the elements of H

had its eigenvalues in the unit circle. The pedigree of Wasow's suggestion goes back to a well-known paper of Courant, Friedrichs, and Lewy [1], in which the Green's function of an elliptic difference equation is identified with the mean number of visits to the points of the lattice made by a particle performing a certain random walk on the lattice. (See the remark (3) below.)

(2) Wasow and other writers describe the process of realizing Z_N as a weighted random walk in which the particle starts at x_i with a "mass" unity, which is then multiplied by the factor $z(x_i, x_{j_1})$, then multiplied by $z(x_{j_1}, x_{j_2})$, and so forth. He deals in [16] with matrix inversion and chooses $u_0 = 0$, so his $c = (\delta_{kj}, j = 1,2,\ldots,n)$, where $\delta_{kj} = 1$, $j = k$, $\delta_{kj} = 0$, $j \neq k$. With this specialization, $\sum_N$ then becomes in Wasow's words the total amount of mass carried through the point x_k during the N steps of the random walk.

(3) If in the function $\sum_N[c,z]$ we take $z_{ij} \equiv 1$, $c = (\delta_{kj})$, then $\sum_N$ is the total number of visits to x_k made by the N-step random walk starting at x_i, counting in one visit for the starting point if $i = k$; and $w_{N+1}(x_i;(\delta_{kj}), 0,1)$ is the mean number of visits to x_k. (In making this observation we are of course no longer tying ourselves down to the requirement that $z_{ij}p_{ij} = h_{ij}$ where h_{ij} is given in advance, but instead we are assuming that the p_{ij} are given a priori.)

(4) The chain Z_0, Z_1, Z_2,... represents a type of branching process. It is a Markov chain which may be of any order from one to infinity, depending on the choice of the z's. Such chains have been studied by Montroll [15], together with the companion type obtained by replacing multiplications by additions. They have numerous applications to theoretical physics and physical chemistry.

(5) In the solution of $u = Hu + c$, or in the problem of inverting of I-H, it should be noted that once a large set of realizations of $(X_0, X_1, X_2, \ldots, X_N)$ has been obtained with all the random walks starting from some fixed x_i, then by proper bookkeeping procedure they can be used a number of times. Not only can they be used to get the statistics for a number of different vectors c— say, all the columns of the identity matrix I, which would be the procedure for inverting I-H— but also, by considering a visit to x_j, $j \neq i$, as starting a new random walk, they can be used to find components of $u_{N'i}$, $N' < N$, or of u, other than the i-th component. Nevertheless a peculiarity of the stochastic method here presented is that it seems to appear to the best advantage in comparison with the standard deterministic methods when the problem is to find only one component of u_N or of u, or one row of $(I-H)^{-1}$. This fact has already been commented upon once before, in the Introduction.

(6) A final remark of minor importance is that the estimator used in the quadrature problem of Section 4 was

analogous to the Z_1 in the present section, not to Z_0. Thus to achieve strict parallelism between the two sections, we should have assigned an appropriate distribution to $R \times R$ in Section 4 rather than just to R, and then considered a conditional distribution in $R \times R$ under a hypothesis relating to R.

6. The m-chain method. The development of the preceding section can easily be modified so that an m-chain does the work of the random sequence $\sum_0 + Z_1 \sum_1 + Z_2, \ldots$. To accomplish this, we adjoin n additional points $x_{n+1}, x_{n+2}, \ldots, x_{2n}$ to R, and set up the 2n x 2n partitioned matrix

$$H^* = \begin{bmatrix} h^*(x_i, x_j) \\ i=1,\ldots,2n \\ j=1,\ldots,2n \end{bmatrix} = \left[\begin{array}{c|c} H & I \\ \hline 0 & I \end{array}\right]$$

Also, we let $(u_0 \vdots c) = (u_{01}, u_{02}, \ldots, u_{0n}, c_1, c_2, \ldots, c_n)$, with $c_i = c(x_i) = c(x_{i+n})$, $i = 1, \ldots, n$. It is to be noticed that

$$H^{*N} = \left[\begin{array}{c|c} H^N & I + H + H^2 + \ldots + H^{N-1} \\ & = (I - H^N)(I - H)^{-1} \\ \hline 0 & I \end{array}\right] . \tag{6.1}$$

Now let z^*_{ij} and p^*_{ij} be chosen so that $z^*_{ij} p^*_{ij} = h^*_{ij}, p^*_{ij} \geq 0$, $\sum_{j=1}^{2n} p^*_{ij} = 1$. Set up the random walk $X^*_0, X^*_1, \ldots,$ on the lattice $x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}$, using the transition probabilities p^*_{ij}. Then construct the new m-chain.

$$Z^*_0 = Z^*_0[(u\vdots c)] = \begin{cases} u(X^*_0), & X^*_0 = x_1,\ldots,x_n, \\ c(X^*_0), & X^*_0 = x_{n+1},\ldots,x_{2n}. \end{cases}$$

$$Z^*_1 = Z^*_1\,[(u\vdots c),\, z^*] = z^*(X^*_0,X^*_1) \times \begin{cases} u(X^*_1), & X^*_1 = x_1,\ldots,x_n, \\ c(X^*_1), & X^*_1 = x_{n+1},\ldots,x_{2n} \end{cases}$$

$$Z^*_N = Z^*_N\,[(u\vdots c),z^*]$$

$$= z^*(X^*_0,X^*_1) \ldots z^*(X_{N-1},X^*_N) \times \begin{cases} u(X^*_N), & X^*_N = x_1,\ldots,x_n, \\ c(X^*_N), & X^*_N = x_{n+1},\ldots,x_{2n}. \end{cases}$$

Let the vectors $v^*_0, v^*_1, \ldots,$ be defined as follows:

$$v^*_{0i} = v^*_0(x_i;(u\vdots c)) = \begin{cases} u_i, & i = 1,2,\ldots,n, \\ c_{i-n}, & i = n+1,\ n+2,\ldots,2n. \end{cases}$$

$$v^*_{Ni} = v^*_N\,(x_i;(u\vdots c),z^*) = E(Z^*_N[(u\vdots c),z^*] \mid X^*_0 = x_i),\ N=1,2,\ldots$$

Then according to (5.1) these vectors satisfy the relations

(6.2a) $\quad v^*_{N+1} = H^*\, v^*_N, \qquad N = 0,1,2,\ldots,$

(6.2b) $\quad v^*_N = H^{*N}(u\vdots c), \qquad N = 0,1,2,\ldots$.

The components of these vectors for $i = n+1,\ldots,2n$ deserve some attention. From the definition of v^*_0, $v^*_{0,n+i} = c_i$, $i = 1,\ldots,n$. As for v^*_1, it will be seen by looking at H^* that $z^*_{ij}\, p^*_{ij} = \delta_{ij}$, $i > n$, $j > n$. This means that $v^*_{1,n+i} = v^*_{2,n+i} = v^*_{N,n+i} = c_i$, $i = 1,\ldots,n$.

Using (6.1) to (6.2) and these facts, we obtain, letting $u = u_0$

$$(6.3a)\quad v^*_{0i} = \begin{cases} u_{0i} & i=1,\ldots,n, \\ c_{i-n}, & i=n+1,n+2,\ldots,2n, \end{cases}$$

$$(6.3b)\quad v^*_{N+1,i} = \begin{cases} Hv^*_N]_i + c_i, & i=1,\ldots,n \\ c_{i-n}, & i=n+1,\ldots,2n \end{cases} \quad N=0,1,\ldots$$

$$(6.3c)\quad v^*_{Ni} = \begin{cases} H^N u_0]_i + (I-H^N)(I-H)^{-1}c]_i & i=1,\ldots,n \\ c_{i-n}, & i=n+1,\ldots,2n. \end{cases} \quad N=1,2,\ldots$$

(In forming the product of the non-conformable factors H and v^*_N, only the first n components of v^*_N are to be used.)

These relations show that the first n components of v^*_N satisfy the recursion relation (2.3) with the initial vector u_0. If $H^N \to 0$, then $v^*_\infty = \lim_{N\to\infty} v^*_N$ exists and its first n components satisfy the linear equations (2.5).

Although the vector w_N of the preceding section and the vector v^*_N of this section must be identical, they are based on quite different stochastic processes. For one thing, in the previous section it was understood that $\sum_{j=1}^{n} p_{ij}$ always equaled unity, which implied that the random walk remained forever on $x_i,\ldots,x_n$, whereas now the corresponding sum of probabilities must be less than unity, and the random walk will not necessarily remain forever on the first n points of the lattice.

In fact, if we choose $p^*_{ij} = \delta_{ij}$ for $i > n$, then the new points $x_{n+1}, x_{n+2},\ldots,x_{2n}$ become trap states or absorbing states for the random walk X^*_0, $X^*_1,\ldots$. It is now a random walk with absorbing barriers.

To increase the comparability of the processes of this section and of the preceding one for solving (2.3) and (2.5) we can set up the solution of the preceding section with H^* replacing H, z^* replacing z, p^* replacing p, $(u_0 \vdots 0) = (u_{01}, u_{02}, \ldots, u_{0n}, 0, \ldots, 0)$ replacing u_0, and $(c \vdots 0) = (c_1, c_2, \ldots, c_n, 0, \ldots, 0)$ replacing c. The vectors $(u_0 \vdots 0)$ and $(c \vdots 0)$ have the effect of annulling the right-hand blocks of H^{*N}. Then the process $\sum^*_{N-1}[(c \vdots 0), z^*] + Z^*_N[(u_0 \vdots 0), z^*]$, where $\sum^*_N = Z^*_0 + \ldots + Z^*_N$, will have the same conditional mean values for $X^*_0 = x_i$, $i \leq n$, as $Z^*_N[(u_0 \vdots c), z^*]$. The random variable $\sum^*_{N-1} + Z^*_N$ is identically equal to zero if the random walk ever visits the new part of the lattice (that is, the point-set $x_{n+1}, x_{n+2}, \ldots, x_{2n}$) in the course of the first N steps.

The m-chain method of inverting matrices was first proposed by Forsythe and Leibler [10], following a suggestion of von Neumann. They considered only the infinite product space $R \times R \times \ldots$ and thus had to make a restriction on H similar to that which Wasow used.

7. <u>Variances</u>. The variance of the random variable $Z_N[u,z]$ of section 5 is easily derived. Let $K = [z_{ij}\, h_{ij}] = [h^2_{ij}/p_{ij}]$. Then by the argument which led to our fundamental equations (5.1), we find that

$$(7.1) \qquad {}_2v_{Ni} = {}_2v_N(x_i; u, z) = E\left\{ (Z_N[u,z])^2 \,\middle|\, X_0 = x_i \right\} = K^N u^2]_i,$$

where u^2 means the vector $(u_1^2, u_2^2, \ldots, u_n^2)$. (We shall frequently use this exponential notation to represent the operation of squaring the components of a vector.)

The conditional variance of Z_N, given that $X_0 = x_i$, is of course the vector ${}_2v_N - v_N^2$.

Turning to the random variable $\sum_{N-1}[c,z] + Z_N[u,z]$, we first find by using the method that led to (5.1) that for $s \leqq t$,

$$E(Z_s[c,z]\, Z_t\, u,z \mid X_0 = x_i) = K^s C H^{t-s} u]_i ,$$

where C is a principal diagonal matrix whose diagonal elements are $c_1, c_2, \ldots, c_n$. We apply this relation and (7.1) to the expansion of the square in the last member of

$$ {}_2w_{Ni} = {}_2w_N(x_i;c,u,z) = E\left\{\left(\sum_{N-1}[c,z] + Z_N[u,z]\right)^2 \middle| X_0 = x_i\right\}$$

$$= E\left\{ (Z_0[c] + Z_1[c,z] + \ldots + Z_{N-1}[c,z] + Z_N[u,z])^2 \middle| X_0 = x_i\right\}.$$

By going through some matrix algebra, the following rather formidable (but equivalent) formulas can be reached:

(7.2a) $${}_2w_N = (I - K)^{-1}c^2 + K^N[u^2 - (I-K)^{-1}c^2] + 2(I-K)^{-1}[C(I-H^{N-1})(I-H)^{-1}Hc - S_{N-1}c] + 2S_{N-1}Hu ,$$

(7.2b) $${}_2w_N = K^{N-1}c^2 + K^N u^2 + (I-K)^{-1}(I-K^{N-1})[2C(I-H)^{-1}c - c^2] - (CH^{N-1} + S_{N-2})(I-H)^{-1}Hc + 2S_{N-1}Hu ,$$

where $S_N = K^N CH + K^{N-1}CH^2 + \ldots + K^2CH^{N-1} + KCH^N$.

The variance is of course to be obtained by subtracting the vector w_N^2 from the vector ${}_2w_N$.

The limiting forms of these formulas will now be derived, under the hypothesis that the eigenvalues of both K and H lie inside the unit circle in the complex plane. For this we need the following simple result.

Lemma. *If* A *is any* n x n *matrix with complex elements, and with the property that all of its eigenvalues lie in the unit circle, then there exist constants* $m = m(A) > 0$ *and* $r = r(A) > 1$ *which are independent of* N, *and are such that* $\left| a_{ij}^{(N)} \right| < m/r^N$, $i,j=1,2,\ldots,n$, $N = 1,2,\ldots$ *where* $a_{ij}^{(N)}$ *is the element in the* i-th *row and the* j-th *column of* A^N.

For the proof, we first observe that the eigenvalues of rA, where r is now any scalar, $r \neq 0$, are those of A multiplied by r, because if y is an eigenvector for the eigenvalue λ, then the equations $Ay = y$ and $rAy = r \lambda y$ are equivalent. This means that if A satisfies the hypothesis of the Lemma, then a real number r exists, $r > 1$, such that the eigenvalues of rA are still less than unity in modulus. Therefore $\lim_{N \to \infty} (rA)^N = 0$. Therefore there is a positive number m, independent of N, such that for all N, $\left| r^N a_{ij}^{(N)} \right| < m$. The result follows at once from this.

We now apply the Lemma to S_N. Let $\bar{c} = \max_i |c_i|$. Then*

*We are using the notation B_{ij} for the i,j-th element of the matrix B.

$$\left|(K^sC)_{ij}\right| < m(K)\bar{c}/[r(K)]^s, \text{ and}$$

$$\left|(K^sCH^{N-s+1})_{ij}\right| < nm(K)m(H)\bar{c}/[r(K)]^s[r(H)]^{N-s+1},\ 0 \leq s \leq N.$$

With $\bar{m} = \max[m(K), m(H)]$, $\bar{r} = \min[r(K), r(H)]$, this becomes $\left|(K^sCH^{N-s+1})_{ij}\right| < n\bar{m}^2\bar{c}/\bar{r}^{N+1}$. From this we get the fact that <u>if the eigenvalue of</u> K <u>and</u> H <u>are all less than unity in modulus</u>, <u>then there exist numbers</u> $\bar{m} = \bar{m}(K,H,C) > 0$ <u>and</u> $\bar{r} = \bar{r}(K,H) > 1$ <u>independent of</u> N <u>such that</u>

(7.3) $\left|(S_N)_{ij}\right| < n\,N\,\bar{m}/\bar{r}^N$; $\quad i,j = 1,\ldots,m;\ N = 1,2,\ldots$.

(This estimate could be refined so that it is independent of n, using for example the methods described in [2, p.16 (footnote)], but it is sufficiently precise as it stands for present purposes.)

The inequality (7.3) implies, of course, that $\lim_{N\to\infty} S_N = 0$.

Therefore the limit of the vector ${}_2w_N$, which we shall denote by ${}_2w_\infty$, is given by the formulas

(7.4a) $$ {}_2w_\infty = (I-K)^{-1}c^2 + 2(I-K)^{-1}C(I-H)^{-1}Hc,$$

(7.4b) $$ {}_2w_\infty = (I-K)^{-1}\,[2C(I-H)^{-1}c - c^2].$$

It is of interest to specialize the m-chain variance given through (7.1) to the augmented 2n x 2n matrix H* discussed in Section 6. Let z^*_{ij} and p^*_{ij} be chosen as previously, but with the proviso now that $p^*_{ij} = 0$ whenever $h^*_{ij} = 0$, $i=1,2,\ldots,n$, $j=n+1,n+2,\ldots,2n$, and also that $p^*_{ij} = \delta_{ij}$, $i > n$, $j > n$. Then the matrix $K^* = [z^*_{ij}\ h^*_{ij}]$ has the appearance

$$K^* = \left[\begin{array}{c|cccc} z^*_{ij}\ h_{ij} & 1/p^*_1 & & & 0 \\ i=1,\ldots n & & 1/p^*_2 & & \\ j=1,\ldots n & 0 & & \ddots & 1/p^*_n \\ \hline 0 & & I & & \end{array}\right]$$

where $p^*_i = 1 - \sum_{i=1}^{n} p^*_{ij}$. It follows as in the derivation of (6.2) that if we define

$${}_2v^*_{0i} = \begin{cases} u_i^2, & i = 1,\ldots,n, \\ c_{i-n}^2, & i = n+1,\ldots,2n, \end{cases}$$

$${}_2v^*_{Ni} = E\left\{ (Z^*_N[(u\vdots c), z^*])^2 \,\middle|\, X^*_0 = x_i \right\}, \quad N = 1,2,\ldots,$$

then letting $u = u_0$,

(7.5a) $${}_2v^*_{N+1,i} = \begin{cases} K({}_2v^*_N)]_i + Qc_i^2, & i = 1,\ldots,n \\ c_{i-n}^2, & i = n+1,\ldots,2n \end{cases} \quad N=0,1,\ldots,$$

(7.5b) $${}_2v^*_{N,i} = \begin{cases} K^N u_0{}^2]_i + (I-K^N)(I-K)^{-1}Qc^2]_i & i=1,\ldots,n \\ c_{i-n}^2, & i=n+1,\ldots,2n \end{cases} \quad N=1,2,\ldots.$$

Here Q denotes the principal diagonal matrix in the upper right-hand corner of K^*, and K means the n x n matrix* $[z^*_{ij}\ h_{ij}]$.

The limiting form of (7.5) is

(7.6a) $${}_2v^*_{\infty i} = K({}_2v^*_\infty)]_i + Qc_i^2, \qquad i = 1,\ldots,n$$

(7.6b) $${}_2v^*_{\infty i} = (I-K)^{-1}\ Qc^2]_i, \qquad i = 1,\ldots,n$$

*To use K in this way is slightly inconsistent with the previous definition of K; perhaps K** should be used; but the author felt that in deference to good taste and the British readers, the page should not be made to look any more star-spangled than it already does.

It was noted in section 6 that the process

$$\sum\nolimits_{N-1}^{*} [(c\vdots 0),z^*] + Z_N^* [(u_0\vdots 0),z^*]$$

has the same conditional mean value as $Z_N^*[(u\vdots c),z^*]$ for $X_0^* = x_i$, $i \leqq n$. With K redefined as above to mean $z_{ij}^* h_{ij}$, (7.2a) (7.2b), (7.4a), and (7.4b) give the components for $i \leqq n$ of the second moment of the random variable $\sum\nolimits_{N-1}^{*} + Z_N^*$ exactly as they stand now. The components with $i > n$ are all zero.

Several remarks will now be made.

(1) The formula (7.6b) was derived by Forsythe and Leibler [16] as the second moment of the conditional distribution of the random variable Z_∞^* defined on the infinite product-space $R \times R \times \ldots$.

(2) The formulas for the second moments suggest that to keep down the variances, the numbers z_{ij} and p_{ij} or z_{ij}^* and p_{ij}^* should be chosen so that on the average (speaking intuitively) the elements of the matrix K should be as small in absolute value as possible. One way <u>not</u> to achieve this end is by letting p_{ij} or p_{ij}^* be positive when $h_{ij} = 0$ or $h_{ij}^* = 0$, because the unnecessary positive values of the p's in a given row could be portioned out to the other elements of K or K* in that row so as to make the elements h_{ij}^2/p_{ij} or h_{ij}^{*2}/p_{ij}^* a little smaller. (We have anticipated this remark to some extent by choosing $p_{ij} = 0$ when $h_{ij} = 0$ for $j > n$.)

(3) A corollary of our results of some slight interest for matrix theory is this. <u>If</u> $[h_{ij}]$ <u>is a real</u> n x m

<u>matrix with all of its eigenvalues inside the unit circle, and if there exists a set of positive numbers</u> p_{ij}, $i,j = 1,\ldots,n$ <u>such that</u> $\sum_j p_{ij} = 1$, $i = 1,\ldots,n$ <u>and the eigenvalues of</u> $[h_{ij}^2/p_{ij}]$ <u>are all inside the unit circle, then the diagonal elements of</u> $(I-H)^{-1}$ <u>are all greater than</u> 1/2. This follows from (7.4b) with c specialized to $c = (\delta_{kj}, j=1,\ldots,n)$, $k = 1,\ldots,n$.

(4) The natural statistical estimation of the mean values v_N, w_N, v_N^*, are of course the arithmetic averages of many determinations of the respective random variables of which these are the theoretical mean values. Clearly the variances of these estimators always exist for finite matrices and finite values of N, no matter where the eigenvalues of H or H* lie. Therefore in particular, the formula (4.1) for the number of random walks which must be performed to attain a given theoretical accuracy are here valid, with ΔI replaced by Δv_N, Δw_N, etc., as the case may be.

8. <u>Comparison of the two methods of inverting matrices</u>. We shall now use the formulas of the preceding section to effect a comparison of the statistical error of the m-chain method of inverting matrices (Forsythe-Leibler) to that of the method based on $\sum_{N-1} + Z_N$ (Wasow) given in Section 5.

It will be impossible to compare these methods unless the z's and p's are chosen comparably. Therefore for both methods we shall suppose that the arrangement described in Section 6 has been set up; that is, the one which uses H^*, z^*_{ij}, p^*_{ij}, X^*_N, and Z^*_N. The comparison will be made by comparing only the limits of the variances as $N \to \infty$, as the finite case seems to be rather intractible.

With $K = [z^*_{ij}\, h_{ij},\ i,j = 1,\ldots,n]$, and Q defined as for (7.5), the formulas to compare are

$$_2W_\infty = (I-K)^{-1}\,[2C(I-H)^{-1}c - c^2]$$

and

$$_2V^*_\infty = (I-K)^{-1}\,Qc^2\ .$$

The specialization to the case of matrix inversion is accomplished by letting $c = (\delta_{kj},\ j = 1,\ldots,n)$; then if the random walks start from $x = x_i$, we shall be estimating $(I-H)^{-1}_{ik}$.

(8.1) $$_2W_{\infty i} = (I-K)^{-1}_{ik}\,[2(I-H)^{-1}_{kk} - 1],$$

$$_2V^*_{\infty i} = (I-K)^{-1}_{ik}/p^*_k\ .$$

The ratio of these second moments is

(8.2) $$\frac{_2W_{\infty i}}{_2V^*_{\infty i}} = p^*_k\,[2\,(I-H)^{-1}_{kk} - 1]$$

Since $(I-H)^{-1}_{kk}$ is fixed by the terms of the problem, we have at our disposal only the numbers p^*_k in this formula. It is clear that for a given H these can always be chosen so that the m-chain method (Forsythe-Leibler) is poorer than the other method (Wasow). And for an H such that the diagonal elements of $(I-H)^{-1}$ are small, the Wasow method might have a smaller variance than the other method, no matter how the random walk was set up.

These negative observations come far from telling the whole story. If the p_k^*'s are adjusted so as to be quite large, in an attempt to make the m-chain method show up favorably in (8.2), then the numbers p_{ij}^*, $i \leq n$, $j \leq n$, will be small and the factors z_{ij}^* will have to be proportionately larger. This will make the elements of K^{*N} and $(I-K)^{-1}$ large, and the values of both ${}_2w_{00}$ and ${}_2v_{00}$ will be large, which of course is undesirable. On the other hand, if the p_{ij}^*'s are chosen so that the values of p_k^* are small, then the mean duration of the random walk becomes very long (see Section 9 below), and this is perhaps more undesirable for the Wasow method then for the m-chain method because the former requires a little more computing per step in the random walk.

The question of an optimum method from the point of view of minimizing computing has not as yet been settled. Probably there is no such thing, because as in the case N = 1 treated in Section 4, it will turn out that the more one knows about the solution, the better one can do. This phenomenon will again be in evidence in Section 10.

9. <u>The duration of the random walk</u>. We shall now apply the formulas of the preceding section to obtain theoretical mean-values and dispersions in two classical problems connected with random walks, which have a bearing on the usefulness of the

Monte Carlo methods discussed in this paper.

We shall be considering the random walk $X_0^*, X_1^*, \ldots,$ on the lattice $x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}$, Section 6. However, in considering the associated matrices $H^* = [z_{ij}^* \; p_{ij}^*]$ and $K^* = [z_{iu}^{*2} \; p_{ij}^*]$, (the latter was introduced in Section 7), the point of view will be that the transition probabilities p_{ij}^* and the factors z_{ij}^* are chosen first, and the values of h_{ij}^* are determined thereby, instead of the other way around. We suppose that $p_{ij}^* = \delta_{ij}$, $i > n$, $j > n$.

In each of the problems to be considered, $z_{ij} \equiv 1$. This means that $K^* = H^* = p_{ij}^*$, $K = H = p_{ij}^*$, $i, j = 1, \ldots, n$. The first of these matrices is a stochastic matrix. We denote the second one by P.

The first problem is that of determining the mean value and variance of the number of visits to x_k if the random walk starts at x_i. In accordance with remark (3) in Section 5, the total number of visits to x_k in an N-step random walk starting at x_i is

$$\sum_N^* [(\delta_{kj}), 1] = \delta_k(X_0^*) + \delta_k(X_1^*) + \delta_k(X_2^*) + \ldots + \delta_k(X_N^*),$$

where $\delta_k(x_j) = \delta_{kj}$, $j = 1, 2, \ldots, 2n$. The answers to the problem are thus given by (5.2) and (7.2a) or (7.2b).

Let the mean number of visits to x_k if the random walk started at x_i be denoted by $G_{ik}^{(N)}$. Then

$$G_{ik}^{(N)} = E\left(\sum_N^* [(\delta_{kj}, 1] \;\middle|\; X_0^* = x_i\right)$$

$$= [(I-P^{N+1})(I-P)^{-1}]_{ik}, \qquad i \leq n,\ k \leq n,$$

$$G_{ik}^{(N)} = \delta_{ik}, \qquad k > n.$$

Assuming that $\lim_{N \to \infty} P^N = 0$, we have

$$(9.1a) \quad G^{(\infty)}_{ik} = \lim_{N \to \infty} G^{(N)}_{ik} = (I-P)^{-1}_{ik}, \qquad i \leq n,\ k \leq n,$$

$$(9.1b) \quad G^{(\infty)}_{ik} = \delta_{ki}, \qquad k > n.$$

Then by substituting into (7.4b), we obtain*

$$\lim_{N \to \infty} \mathrm{Var}\Big(\sum\nolimits^{*}_{N} [(\delta_{kj}),1] \Big| X^*_0 = x_i\Big) = 2G^{\infty}_{ik}\, G^{\infty}_{kk} - G^{\infty}_{ik} - (G^{\infty}_{ik})^2.$$

The first two terms of the formula are of course just a special case of (8.1).

The other problem concerns the duration of the random walk $X^*_0, X^*_1, \ldots$ We define the duration to mean the total number of visits made to each of the transient states $x_1, x_2, \ldots, x_n$ before absorption takes place. In counting the visits, we count in the starting point of the walk as one visit.

If the walk is limited to N + 1 visits (that is, to N steps), then the duration is clearly

$$\sum\nolimits^{*}_{N} [(1\mathbin{\vdots}0),1], \text{ where } (1\mathbin{\vdots}0) = (\underbrace{1,1,\ldots,1}_{n}, \underbrace{0,0,\ldots,0}_{n}).$$

Its conditional mean value for $X^*_0 = x_i$ is the vector d_N whose i-th component is

*By Var(Y|b) we mean the variance of the conditional distribution of the random variable Y, given that the event b has occurred.

$$d_{Ni} = \sum_{j=1}^{n} G_{ij}^{(N)} = \sum_{j=1}^{n} [(I-P^{N+1})(I-P)^{-1}]_{ij}, \qquad i \leqq n,$$

$$d_{Ni} = 1, \qquad i > n.$$

The vector d_N satisfies the recursion relation

$$d_{N+1,i} = Pd_N]_i + 1, \qquad i \leqq n,\ N = 0,1,2,\ldots,$$

with $d_{0i} = 1$. Its limiting value is

$$(9.2) \qquad d_{\infty i} = \sum_{j=1}^{n} (I-P)^{-1}_{ij}, \qquad i \leqq n.$$

The variance of the duration can be obtained by substituting appropriately into (7.2a) or (7.2b). We shall again write out only the limiting case. This is obtained easily from (7.4b); the i-th component is

$$2^{d_{\infty i}} - d^2_{\infty i} = 2 \sum_{j=1}^{n} G^{(\infty)}_{ij} \, d_{\infty j} - d_{\infty i} - d^2_{\infty i} .$$

$$\leqq d_{\infty i} \, (2d - 1 - d_{\infty i}),$$

where $d = \max_i d_{\infty i}$.

The duration of non-truncated random walks in an infinite product space has recently been investigated in very general cases by Wasow [17]. Previously, special cases had been studied at some length by statistical theorists in connection with the problem on the mean length of a sequential test. (See [4] for further results and references.)

The conditional probability that X^*_N falls on one of the states $x_1,\ldots,x_n$, given that $X_0^* = x_i$, is clearly $\sum_{j=1}^{n} (P^N)_{ij}$.

This is the probability that the walk lasts for N steps without falling into a trap state. To assert that $\lim_{N\to\infty} P^N = 0$ is therefore equivalent to saying that the probability that the walk lasts more than N steps approaches zero with N. In fact, if we consider the infinite product-space $R\times R\times\ldots$ for a moment, $\lim_{N\to\infty} P^N = 0$ means that the walk is "almost certainly" of finite duration, and conversely. It is known (see for example Curtiss [4, Section 11]), that if from each one of the states $x_1, x_2,\ldots,x_n$ it is possible to reach a trap state over a path in $R\times R\times\ldots$ with non-zero probability, then the walk is almost certainly of finite duration.

This provides a sufficient condition for the validity of (9.1) and (9.2). If it is satisfied, then all the eigenvalues of P be less than unity in absolute value.

We mentioned in Section 8 that if the number $p_k^* = 1 - \sum_{j=1}^{n} p_{kj}^*$ were small, then the mean duration would tend to be large. This follows intuitively from the fact that the smaller the absorption probabilities are, the longer the walk will go on. A somewhat more rigorous demonstration can be given by considering the dependence of the eigenvalues of P on the row-sums of P. If $p_{ij}^* > 0$, $i = 1,\ldots,n$, $j=1,\ldots,n$, then as the minimum row-sum approaches unity, the eigenvalue of maximum absolute value* approaches unity, and thus the con-

*It happens to be real and positive, since P is a matrix of positive elements.

vergence to zero of P^N becomes slower and slower.

10. Importance sampling. In this section we shall discuss the problem of the control of the statistical error. The treatment will be analogous to the one in the latter part of Section 4. It will pertain only to the estimators Z_N and Z_N^*, and not to the estimator $\sum_{N-1} + Z_N$. We shall assume throughout that the problem is to estimate the solution of (2.3) or (2.5). Therefore the factors z_{ij} and p_{ij}, or z_{ij}^* and p_{ij}^*, will always be related in the usual way to the elements of the matrix H or H*.

It is worth while first to inquire into the conditions under which an m-chain can be a zero-variance estimator. It will suffice here to examine the situation only for the function $Z_N[u,z]$ of Section 5.

If Z_N did indeed have a conditional variance of zero, given that the random walk starts at x_i, then the value assumed on every path with positive probability would depend only on x_i. One way to insure this would be to choose $z_{ij} = \lambda u_i/u_j$, where is a constant; then

$$Z_N = \lambda^N \frac{u(X_0)}{u(X_1)} \cdot \frac{u(X_1)}{u(X_2)} \cdots \frac{u(X_{N-1})}{u(X_N)} u(X_N) = \lambda^N u(X_0).$$

Conversely, it can be shown that this is roughly the most general choice of the factors z_{ij} which will insure the desired result. We shall not try to give an accurate formulation of the theorem here.

Now if $z_{ij} = \lambda u_i/u_j$, then the requirement that $z_{ij}p_{ij} = h_{ij}$ implies that the choice of p_{ij} must be $h_{ij}u_j/u_i$. This in turn requires not only that all of these quotients must be real and non-negative, but also that $\sum_j h_{ij}u_j/\lambda u_i = 1$. The latter relation states that u must be an eigenvector of H corresponding to the eigenvalue λ.

If we stop insisting on stationary transition probabilities for our random walk, and also permit the factors z_{ij} to vary from step to step, then a zero-variance m-chain estimator of $u_N = H^N u_o$ can very easily be constructed for any u_0, provided that $h_{ij} > 0$ and $u_j > 0$, $i,j = 1,\ldots,n$. We simply choose

$$p_{ij}^{(1)} = \frac{h_{ij}\, u_{N-1,j}}{u_{N,i}}, \qquad z_{ij}^{(1)} = \frac{u_{N,i}}{u_{N-1,j}},$$

$$p_{ij}^{(2)} = \frac{h_{ij}\, u_{N-2,j}}{u_{N-1,i}}, \qquad z_{ij}^{(2)} = \frac{u_{N-1,i}}{u_{N-2,j}}$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$p_{ij}^{(N)} = \frac{h_{ij}\, u_{0j}}{u_{1i}}, \qquad z_{ij}^{(N)} = \frac{u_{1i}}{u_{0j}}$$

It is easily checked that Z_N formed with these factors $z_{ij}^{(K)}$ and with $u = u_0$, has the constant value $u_{N,i}$ for any random walk starting at x_i. For each K, $v_{K+1} = H\, v_K$, where $v_{K,i} = E(Z_K | X_0 = x_i)$, but the chain is not a zero-variance chain unless K = N.

The argument can be extended so as to allow zero elements in H and u_0 by giving a little attention to the undefined quotients.

It is not difficult to proceed from here to a practical arrangement whereby given only approximate values of $u_0, u_1, u_2, \ldots, u_N$, a chain can be set up in which the variance can be exhibited as a function of the errors of the approximations, as was done in the case $N = 1$ in Section 4. Because of limitations of space we shall not pursue the matter further. Instead we shall study a special class of m-chains which use stationary transition probabilities and for which the statistical error analysis is usually easy to make. These m-chains, however, have the slight disadvantage that they are connected with a special type of matrix H.

The main problem is, as usual, to estimate the vector u_N in the recursion relation

$$u_{N+1} = Hu_N + c \quad . \tag{10.1}$$

We shall also be interested in the problems of estimating u in the equation

$$u = Hu + c \quad , \tag{10.2}$$

but as elsewhere in this paper this problem will be considered as the limiting case of the first problem as N becomes infinite. Whenever (10.2) is in view, we shall as usual assume that all the eigenvalues of H lie inside the unit circle.

We now impose certain conditions on H. They are that $h_{ij} \geqq 0$, and that a vector u_0 with positive components shall exist such that the components of $(I - H)u_0$ are all positive. The existence of such a vector u_0 follows automatically from known results on matrices if the eigenvalues of H all lie inside the unit circle and H is non-singular. We write $u_0 = Hu_0 + c + \epsilon$; then $c_j + \epsilon_j > 0$, $j = 1,\ldots,n$.

If the problem in view is to solve $Au = b$, then theoretically speaking, matrices H and M can always be selected so that the equivalent system $u = Hu + Mb$, with $H + MA = I$, satisfies the conditions imposed here. For example, choose H to be a principal diagonal matrix with elements lying between 0 and 1. Then its eigenvalues lie in the unit circle, and the existence of a vector u_0 with the required property is assured. (Of course, any particular a priori choice of H such as this usually means that M must be determined from $M = A^{-1}(I-H)$, which requires a knowledge of A^{-1}. In practice therefore it is M that will probably be chosen first, not H.)

We note that ϵ (or rather $-\epsilon$) is the "residual" in the classical theory of solutions of the system $(I - H)u = c$. It is easy to show that if the successive approximations u_1, $u_2, \ldots$ to the solution u of (10.2) are defined by the recursion relation (10.1), then

$$u_N - u = H^N (I - H)^{-1}\epsilon \quad .$$

The second member of this equation is the truncation error (as opposed to the statistical error) in the statistical solution about to be proposed.

At this point we reformulate the problem in terms of the partitioned matrix and vectors introduced in Section 6. The equation (10.1) can be rewritten as

(10.3) $$(u_{N+1} \vdots c) = H^* \times (u_N \vdots c)$$

and the equation (10.2) becomes

(10.4) $$(u \vdots c) = H^* \times (u \vdots c).$$

The problem of solving (10.2) then becomes one of finding an eigenvectorof H^* for the eigenvalue unity, given preassigned values c_1, $c_2, \ldots, c_n$ for the last n of the 2n components of the eigenvector. This formulation of the "steady state" problem as an eigenvalue problem permits us to make an approach to the zero-variance sampling situation discussed earlier in the section.

We define correspondences between the points x_{n+1}, $x_{n+2}, \ldots, x_{2n}$, and the components of ϵ and c respectively by $c(x_{i+n}) = c_i$, $\epsilon(x_{i+n}) = \epsilon_i$, $i = 1, \ldots, n$; and we introduce the vector c^* whose components are given by

$$c^*_i = c^*(x_i) = \begin{cases} u_0(x_i), & i = 1,2,\ldots,n \\ c(x_i) + \epsilon(x_i), & i = n+1,\ldots,2n. \end{cases}$$

That is, $c^* = (u_0 \vdots c + \epsilon)$. The n equations represented by $u_0 = Hu_0 + c + \epsilon$ which define the components of ϵ are equivalent to the 2n equations represented by

(10.5) $$c^* = H^* c^*.$$

We are now ready to set up the basic random walk, and the corresponding m-chain whose mean value is the solution of (10.1).

The transition probabilities of the random walk will be given by

$$p^*_{ij} = h^*_{ij} \frac{c^*_j}{c^*_i}.$$

It is to be noted that $p^*_{ij} \geq 0$, and because of (10.5),

$$\sum_{j=1}^{2n} p^*_{ij} = 1, \qquad i = 1,\ldots,2n$$

Then z^*_{ij} is chosen so that

$$z^*_{ij} = \frac{c^*_i}{c^*_j} .$$

Letting the random variables X^*_0, X^*_1,... denote as usual the successive states visited by a random walk having these transition probabilities, and substituting into $Z^*_N[(u_0 : c), z^*]$, we obtain

$$Z^*_N = \frac{c^*(X^*_0)}{c^*(X^*_1)} \cdot \frac{c^*(X^*_1)}{c^*(X^*_2)} \cdots \frac{c^*(X_{N-1})}{c^*(X^*_N)} \times \begin{cases} u_0(X^*_N), \text{if } X^*_N = x_i, i=1,2,\ldots,n, \\ c(X^*_N), \text{if } X^*_N = x_i, i=n+1, \\ \qquad n+2,\ldots,2n. \end{cases}$$

Cancelling terms and using the definition of c^*, we find that for X^*_0 on $x_1,\ldots,x_n$,

$$(10.6) \qquad Z^*_N = u_0(X^*_0) \left[1 - \delta(X^*_N) \frac{\epsilon(X^*_N)}{c(X^*_N) + \epsilon(X^*_N)}\right],$$

where

$$\delta(x_i) = \begin{cases} 0, & i = 1,\ldots,n \\ 1, & i = n+1,\ldots,2n. \end{cases}$$

It is known from the results of Section 6 that $E(Z^*_N \mid X_0 = x_i) = u_{Ni}$, $i \leqq n$, where u_N is defined by (10.1) above.

The practical procedure implied here for estimating u_N is simply this: Start a random walk at x_i, using the transition probabilities p^*_{ij}. If absorption has not taken place after N steps (that is, if the random walk has not reached any of the points x_{n+1}, x_{n+2},...,x_{2n}), then record u_{0i}. If absorption does take place during the N steps, stop the walk then and there, note the index i of the last point x_i touched before absorption, and record $u_{0i}c_i/(c_i + \epsilon_i)$. Do this for many random walks, and then average up the recorded values.

We now consider the statistical error of the procedure. From (10.6)*,

$$\mathrm{Var}(Z_N^* \mid X_0^* = x_i) = u_{0i}^2 \, \mathrm{Var}\left(\delta(X_N^*) \frac{\epsilon(X_N^*)}{c(X_N^*) + \epsilon(X_N^*)} \middle| X_0^* = x_i\right),$$

$$i = 1,\ldots,n.$$

(We note in passing that if $\epsilon = 0$, then the variance vanishes, as it should). Proceeding as we did in Section 4, and letting

$$e = \max_i \frac{\epsilon_i}{c_i + \epsilon_i},$$

we get

$$\mathrm{Var}(Z_N^* \mid X_0^* = x_i) < \frac{e^2 u_{0i}^2}{2} \qquad i = 1,\ldots,n,$$

which for practical purposes is about as satisfactory as appraisal of the variance as it seems possible to obtain. It contains none of the unknown quantities in the problem.

By using the formulas and methods of Section 7, it is easy to get an explicit formula for the variance. The result is

*By the symbol Var(Y|b), we mean the variance of the conditional distribution of Y, given that the event b has occurred.

$$\mathrm{Var}(Z_N^* \mid X_0^* = x_i)$$
$$= u_{oi}(I-H^N)(I-H)^{-1}\epsilon']_i - \left\{(I-H)^{-1}\epsilon]_i\right\}^2, \qquad i = 1,\ldots,n,$$

where

$$\epsilon'_i = \frac{\epsilon_i^2}{c_i + \epsilon_i}, \qquad i = 1,\ldots,n.$$

One somewhat interesting conclusion that can be drawn from this formula is that if u_0 is chosen so as to have constant components, then the variance will have one standard limiting vector as $N \to \infty$, no matter what the magnitude of the components of u_0 may be. It turns out that this limiting vector is

$$(I - H)^{-1} c' - u^2,$$

where u is the solution of (10.2) and

$$c'_i = \frac{c_i^2}{\sum_{j=1}^{n}(I-H)_{ij}}, \qquad i = 1,\ldots,n.$$

REFERENCES

1. Courant, R., Friedrichs, K., and Lewy, H., Uber die Partiellen Differenzengleichungen der Mathematischen Physik, Math. Annalen, vol. 100(1928), pp. 32-74.

2. Courant, R., and Hilbert, D., Methoden der Mathematischen Physik, Berlin, 1931.

3. Cramer, H., Mathematical Methods of Statistics, Princeton, 1946.

4. Curtiss, J. H., Sampling Methods Applied to Differential and Difference Equations. Proceedings of a Seminar on Scientific Computation, held by the International Business Machines Corporation, Endicott, New York, November, 1949. pp. 87-109.

5. Cutkosky, R. E., A Monte Carlo method for solving a class of integral equations, Journal of Research of the National Bureau of Standards.

6. Demin, W. Edwards, Some Theory of Sampling, New York, 1950.

7. Frazer, R. A., Duncan, W. J., and Collar, A. R., Elementary Matrices, Cambridge, 1950.

8. Feller, W., An Introduction to Probability Theory and its Applications, New York, 1950.

9. Forsythe, G. E., Tentative Classification of methods and bibliography on solving systems of linear equations. To appear in Simultaneous Equations and the Determination of Eigenvalues--Proceedings of a National Bureau of Standards Symposium held in Los Angeles, August, 1951.

10. Forsythe, G. E., and Leibler, R. A., Matrix Inversion by a Monte Carlo Method, Mathematical Tables and Other Aids to Computation, vol. 4(1950), pp. 127-129.

11. Kahn, H., Stochastic (Monte Carlo) Attenuation Analysis, RAND Corp., Santa Monica, California. Report No. R-163, June 14, 1949.

REFERENCES (Continued)

12. Kolmogorov, A. N., Foundations of the Theory of Probability, New York, 1950.

13. MacDuffee, C. C., The Theory of Matrices, New York, 1946.

14. Proceedings of a Symposium on the Monte Carlo Method, National Bureau of Standards Applied Mathematics Series No. 12, Washington, 1951.

15. Montroll, E. W., On the theory of Markov Chains, Annals of Math. Statistics, vol. 18(1947), pp. 18-36.

16. Wasow, W., On the inversion of matrices by random walks, Mathematical Tables and Other Aids to Computation, vol. 6(1951), pp. 78-81.

17. Wasow, W., On the mean duration of random walks, Journal of Research of the National Bureau of Standards, vol. 46(1951), pp. 462-472.

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.

NBS