COMPUTER SIMULATION OF STOCHASTIC PROCESSES THROUGH MODEL-SAMPLING (MONTE CARLO) TECHNIQUES *

C. W. SHEPPARD
Memphis, Tennessee

A simple Monte Carlo simulation program is outlined which can be used for the investigation of random-walk problems, for example in diffusion, or the movement of tracers in the blood circulation. The results given by the simulation are compared with those predicted by well-established theory, and it is shown how the model can be expanded to deal with drift, and with reflexion from or adsorption at a boundary.

1. Introduction

When we observe natural phenomena we often find ourselves measuring some numerical quantity and watching it vary randomly with time as though there were some progressive underlying process which governs its evolution. Such a variable is called in statistics a *random variable* and the underlying process which we are conjecturing from our observations is called a *stochastic process.* In biochemistry the system we are observing may be an ion-exchange column and the variable may be the ultraviolet absorption or radioactivity of nucleotides or other chemical substances at its outflow[1–3]. In physiology we may be interested in the kinetics of transport of indicator substances through the circulation [4,5]. As Hamming has asserted "the purpose of computing is insight" and our progress in this direction is often aided by the simulation of stochastic processes by the use of high-speed stored-program digital computers. The techniques which are employed have been variously termed model sampling or "Monte Carlo" [6].

In throwing well-balanced pennies, theorems in the science of probability tell us that the ratio of heads to tails must converge to unity with increasing numbers of throws. We are also told how many throws we should expect to make before the ratio reaches within, say, 1 percent of this ultimate *expected value.* Since there is a long tradition among biologists that we prefer the results of experiment to theory we should soon find ourselves testing these theorems by throwing pennies. Since, as we make successive throws, the number of heads in a given run can have only integral values we call this run a sample from a *discrete* stochastic process. In the two tracer-kinetic examples mentioned above there are so many particles in the effluent, i.e., samples are so large that deviations from the expected value are not the result of sampling error, i.e., fluctuations which occur when only a finite number of elements are present in a given sample. At the same time, the concentration of material at the outflow, itself a measure of probability, varies uniformly as a function of time and we are now dealing with a *continuous* process. Primitive techniques such as coin tossing are inadequate to test our theory and the simulation of such a process can not readily be performed without the use of a digital computer.

The simulations which will be discussed here are all in relatively simple systems for which the basic mathematics has been comprehensively worked out [7]. Our interest has been the result of efforts to develop the convection-diffusion (random-walk) model of tracer kinetics in the blood circulation. It is not to be suggested that Monte Carlo simulations are frequently as simple as these. We will say nothing about other very important uses of the method such as for numerically evaluating cumbersome integrals [8] or inversion of matrices [9].

* Supported by Research Grant HE-09495 from the National Institutes of Health. The author was aided in a portion of this paper by a discussion with Dr. Z.Pechan.

North-Holland Publishing Company – Amsterdam

2. Jumps between lattice points

Fig. 1 shows the model which we have used in most of our simulation studies. Given a set of discrete points numbered with positive and negative integers and extending both ways to infinity, let us follow the position of a point or "particle" which is initially at zero.

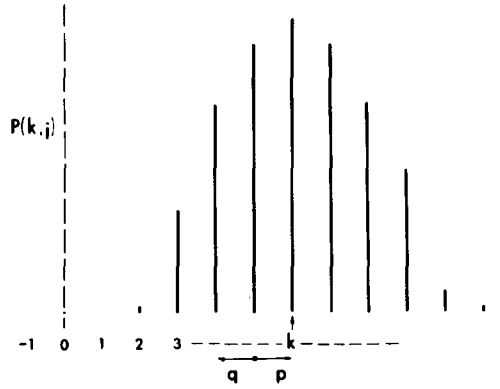

Fig. 1. Model system for simulation studies.

A coin is thrown and if it lands "heads" the particle jumps to the right and if "tails" comes up the jump is to the left. If k is the number of the point where the particle is found after j throws, what is the probability $P(k, j)$ that the particle is at k? We call the function which describes these outcomes the *probability density function.* More generally, then, if the probability of a jump to the right is p and to the left q, where $p+q=1$, what is the probability density function? From the binomial law [10] the probability of m successes of an event of probability p in j trials is

$$P(m,j) = \binom{j}{m} p^m q^{(j-m)} ,$$

where $\binom{j}{m}$ is the binomial coefficient $j!/[m!(j-m)!]$. In the model, however, the variable k is the difference between the number of successes, i.e., jumps to the right and failures, i.e., jumps to the left. That is, if n is the number of failures, $k=m-n$ and the total number of jumps $j=m+n$. Thus solving the two relations yields

$$m = (j+k)/2 ,$$

$$j-m = (j-k)/2 ,$$

and

$$P(k,j) = \binom{j}{(j+k)/2} p^{(j+k)/2} q^{(j-k)/2} . \qquad (1)$$

Because all values are integers, $j \pm k$ must be even, so for even j, $P(k, j) = 0$ for all odd k and vice versa. As a series of trials is made, jumps will occur randomly to the left or right and the particle will execute a "random walk". We may then follow the individual histories of particles as we simulate the process and inquire for many histories, and for a given j what the density will be as a function of k. We may also wish to know for a given k what will be the value of j for which a particle arrives for the first time at k. Thus k becomes a kind of space variable and j a kind of time variable. We may suspect that when $p>q$ the density function will drift to the right and, as j increases, it will broaden out.

3. Simulation program

Fig. 2 shows a summary flow chart of the program employed to simulate the above stochastic process using the IBM 1620 computer. In order to achieve speed the instructions were written in symbolic language and assembled in machine language using fixed-point arithmetic. This particular program yields the fraction of particles found at each k for a given j.
A table is set up in machine memory with addresses corresponding to the lattice points in fig. 1. At the beginning of the program the values of j and p (J and P in the program) are read in, together with the desired number of histories NH. The address ORIG of the initial position of the particle (zero in fig. 1) is also read in. The program also has the feature that addresses for an absorbing (ABS) and a reflecting (REFL) barrier may be set. If the particle wanders far enough to the left it will encounter the latter barrier and be reflected to the right. On reaching the absorbing barrier on the right, the particle is lost from the system by "absorption" without being tallied. The unrestricted walk is programmed by setting the barrier addresses at the extreme limits of the table and the origin well away from them.

Following the read-in operation a history begins. In the first operation a temporary variable NJ is set

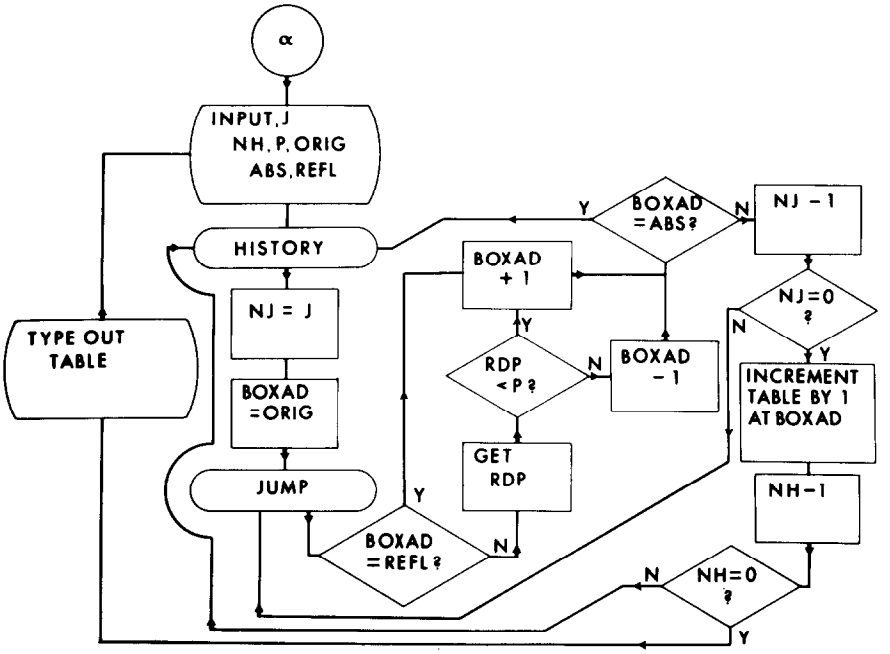

Fig. 2. Flow chart of program to simulate discrete random walk.

equal to J since J must be saved for further histories. The temporary address BOXAD of the particle which changes with the particle's movement is equated to the initial address ORIG. The next step is to execute a particle jump. If BOXAD=REFL, i.e., whatever has occurred previously, if the particle's address is that of the reflecting boundary, the program takes the "yes" branch (Y in fig. 2) to the instruction which increments the table address by one unit, thus stimulating a reflection to the right by one lattice unit. In the absence of reflexion there is a branch (N) to a subroutine which generates a random digit pair (RDP in fig. 2). If the resulting two-digit number is less than p then BOXAD is incremented by 1, otherwise BOXAD is decreased by 1. A test is then made to see if the address thus generated is now equal to ABS in which event the program returns to begin a new history. If absorption did not occur the value of NJ is decremented by 1 and if NJ is not then zero, the program returns to execute another jump. If NJ is zero, the history is complete and the total number of jumps has been executed. To indicate where the particle is now located, i.e., its k address, the value 1 is added to the contents of the memory location corresponding to k. In other words a "marble" is dropped into the appropriate "box" of the table of lattice addresses. After decrementing NH, the number of histories, and testing for zero the machine returns to execute a new history or, if NH = 0, types out the table. From this one may determine how many "marbles" are in each box, and, dividing by the total number of histories, the percentage distribution of histories terminating at each k can be determined.

Fig. 3 shows the result of one simulation for the case $p = 0.75$. The probabilities estimated by the simulation are shown as the length of the vertical lines. The circles show the theoretical expected results obtained by the use of equation (1). By simulating varying numbers of histories we have noted that the percentage scatter resulting from sampling error of the simulated results around the theoretical points tends to decrease with increasing NH.

The appearance of the distribution suggests a quasi-normal curve. Obviously this cannot be exact since if j is odd, all values of k which are even contain no points, and if j is even, odd-valued addresses are not occupied. Furthermore there can be no points for k values beyond the total number of jumps. Nevertheless as j becomes greater, one may imagine that these objections will be removed in the limit $j \to \infty$. We

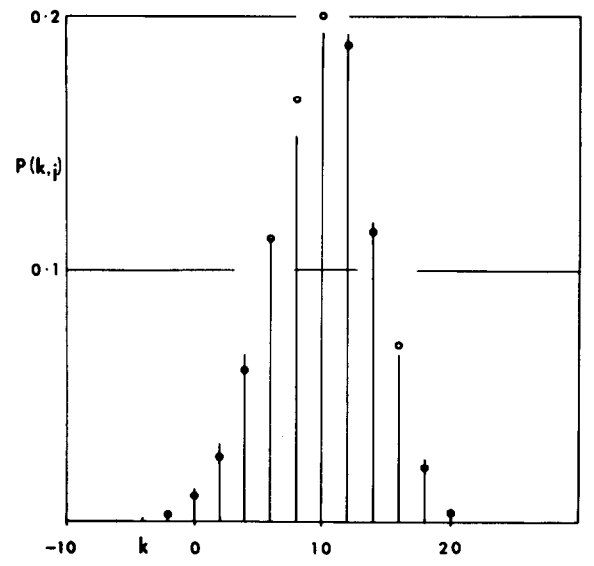

Fig. 3. Unrestricted random walk, $p = 0.75$, 20 jumps/history showing results of a 4000-history simulation (vertical bars) and theoretical points (circles). Ordinate is probability, abscissa is number of points from zero.

thus inquire into the nature of the distribution for very large j.

Insert into the binomial coefficient of equation (1) the Stirling approximation for the factorials, $N! \approx (2\pi N)^{1/2}(N/e)^N$ and simplify, obtaining

$$P(k,j) = (2\pi jpq)^{-\frac{1}{2}}\left[\frac{2p}{(1+k/j)}\right]^{\frac{1}{2}(j+k+1)} \times \left[\frac{2q}{(1-k/j)}\right]^{\frac{1}{2}(j-k+1)}$$

Let $2p = 1 + E/j$, and $2q = 1 - E/j$, and take logarithms of both sides, thus

$$\log[P(k,j)\sqrt{2\pi jpq}] = -\tfrac{1}{2}(j+k+1)\log\left[\frac{1+k/j}{1-E/j}\right] - \tfrac{1}{2}(j-k+1)\log\left[\frac{1-k/j}{1-E/j}\right]$$

For large enough j we can approximate $\log(1+k/j)$ and equivalent terms with the first two terms of a power series expansion and, after considerable simplification obtain

$$\log[(2\pi jpq)^{\frac{1}{2}}P(k,j)] \approx -[(E-k)^2/2j]$$

or
$$P(k,j) \approx (2\pi jpq)^{-\frac{1}{2}} \exp[-(E-k)^2/2j] \; . \qquad (2)$$

This is the DeMoivre-Laplace approximation. It is fairly good even for relatively small j if p and q remain in the neighbourhood of 0.5 and provided we consider only integer values of j and even values of $(j \pm k)$. The equation illustrates a special case of a more general theorem, the "central-limit" theorem [10] which asserts generally that, subject to rather lenient restrictions, the sum of R independent random variables tends toward a normal distribution in the limit $R \to \infty$. The probability is normally distributed and is symmetrical about the value E which is the centroid of the distribution. In statistical language E is called the *expectation* of k. Solving in terms p and q we find that

$$E = (p-q)j \; . \qquad (3)$$

Thus in this approximation the distribution moves as though it had a velocity $(p-q)$ and as though j were the time. From equation (2) the variance of the distribution is simply j, the number of trials. Thus, in this sense, the standard deviation increases in proportion to the square root of the "time" j. For a fine enough grained lattice and a large enough number of trials, i.e., $R \to \infty$, we can simulate with a discrete lattice something which closely resembles diffusion with simultaneous convection. The basic theory of the merging of the discrete into the continuous model is well-developed (see, for example, ref. [11]) but we will not refer to it further. The analogous equation for the probability of finding a particle between x and $x + dx$ at time t under simultaneous diffusion and drift with velocity v is

$$P(x,t) = (4\pi Dt)^{-\frac{1}{2}} \exp[-(x-vt)^2/4Dt] \; , \qquad (4)$$

where D is the diffusion coefficient. The analogy between this equation and equation (2) is apparent.

Fig. 4 shows the effect of placing a reflecting boundary at the origin of the random walk. When $p > q$ and the distribution drifts to the right it seems

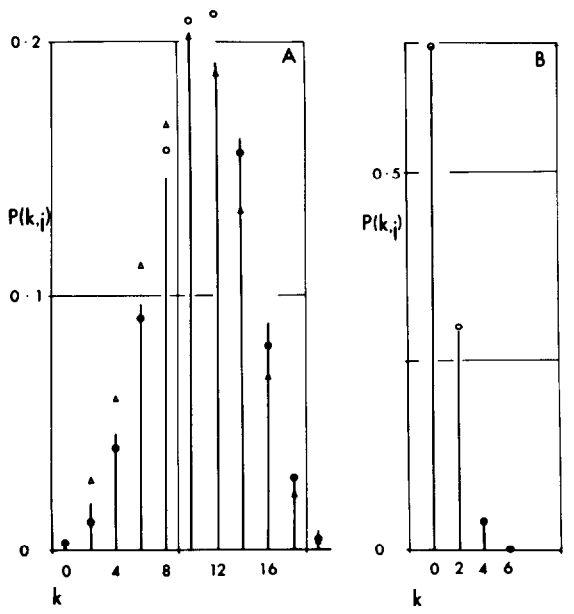

Fig. 4. Walks with reflecting barrier at the origin. Each simulation contained 4000 histories of 20 jumps each. Vertical bars are simulations and circles are theoretical points. The triangles in A are for an unrestricted walk. In A $p = 0.75$ and in B $p = 0.25$.

at first as though there were no boundary effect. Careful tracing of several histories stepwise on the computer showed that many reflections were indeed occurring. Furthermore the points computed from equation (1) (triangles in fig. 4) do not fit the data. Kac [11] has derived an expression for the discrete random walk in the presence of reflexion. Because he defined p as the probability of jumps to the left we have interchanged p and q and also include the case $p > q$ in his result, yielding for the boundary at the origin

$$P(k,j) = \frac{[|p-q|+(q-p)]}{4pq}\left(\frac{p}{q}\right)_*^k [1+(-1)^{j+k}]$$

$$+\frac{2}{\pi}\left(\frac{p}{q}\right)_*^{\frac{1}{2}k}(2\sqrt{pq})^j$$

$$\times\int_0^{\pi}\cos^j\theta\,\{\tan^2\theta/[(p-q)^2+\tan^2\theta]\}\,f_k(\theta)\,d\theta\,,$$

where

$$f_k(\theta) = \cos(k\theta)+(p-q)\sin(k\theta)/\tan\theta\,,$$

$$\left(\frac{p}{q}\right)_*^k = \left(\frac{p}{q}\right)^k \quad \text{if} \quad k>0\,,$$

$$= p \quad \text{if} \quad k=0\,. \qquad (5)$$

The circles in the figure show the comparison of the simulation results with computed results from numerical evaluation of equation (5) with the computer. It was found that the argument of the integral in the equation oscillates rather strongly necessitating the use of Filon's method to avoid the unnecessary expenditure of computer time (see, for example, ref. [12]). The results of this study indicated that, if the distribution is strongly convected away from the barrier a good representation is obtained by considering that the distribution originated from an effective origin displaced somewhat from the real one. That this approximation cannot be generally accepted is seen in fig. 4B where the case for $p = 0.25$ is shown. Here the convection is to the left against the barrier and the normality of the distribution is rapidly destroyed as particles repeatedly bounce against the barrier.

According to the central-limit theorem as $R \to \infty$ the expectation of a sum of independent random variables approaches the sum of the expectations of the individual variables and a similar relation holds for the variances. For the case of equation (2) the variance for an individual trial is the mean of the squares of the jump sizes to the right and left which is unity. Multiplying this unit mean by the number of trials yields the value j for the variance which is equivalent to the $2Dt$ in equation (4). Thus it is possible to set any desired value for the diffusion coefficient which we wish to simulate by choosing an appropriate scale of the lattice. Similarly, by choosing the appropriate values of p and q we may obtain any desired drift rate because of the analogy between the expectation $(p-q)j$ in equation (3) and the corresponding value vt in equation (4). In cases where we wish to simulate variations in D, however, an alternative procedure has proved useful. We now employ a

program which compares the random digit pair with 50. If it is less the particle jumps to the left, otherwise it jumps to the right. Appropriate values of D are simulated by varying the jump size and a table of jump sizes is created in machine memory. The table address which is scanned moves together with the simulated particle movement and, at each jump, the address in the table corresponding to the lattice point address of the particle is consulted to determine the size of the jump. Drift is produced by adding a second biasing jump of magnitude U to all particles jumping to the right. Since, on the average, jumps to the right occur half the time the expectation for single jumps is $\frac{1}{2}U$ and the amount of drift at time j in this analogy is $\frac{1}{2}jU$. In the case of the diffusion coefficient, if L is the jump size read from the table, the mean-square jump size is evidently $\frac{1}{2}[L^2+(L+U)^2]$ and thus

$$\tfrac{1}{2}j\,[L^2+(L+U)^2] \qquad \text{is analogous to } 2Dt\,.$$

The derivation of the probability that a particle will reach an absorbing barrier (or arrive for the first time at a non-absorbing boundary) is given elsewhere [13]. The appropriate relation is

$$P(t) = a(4\pi Dt^3)^{-\frac{1}{2}} \exp\left[-(a-vt)^2\right]/4Dt\,, \qquad (6)$$

where a is the distance between the origin and the barrier.

It is convenient to employ a mean transit time $\bar{t}$ between origin and barrier where $\bar{t}=a/v$. Results of a simulation of this system are shown in fig. 5 where the vertical bars are obtained from simulation and the circles are computed from equation (6).

A question which may arise in the case of random walks with varying D is the probability density function which might be expected in a composite medium where an interface separates two regions of diffusion coefficients D_1 and D_2. For the case of diffusion without drift with an initial delta function at the origin, the analytical solution is obtained in the classical literature on conduction of heat in solids [14]. Converting to analogous diffusion variables we obtain

$$P(t) = A\,\exp\left[-x^2/4D_1t\right] \qquad \text{for } x \leqslant 0$$
$$= A\,\exp\left[-x^2/4D_2t\right] \qquad \text{for } x \geqslant 0$$
$$\text{where} \quad A = \left[\sqrt{\pi t}\left(\sqrt{D_1}+\sqrt{D_2}\right)\right]^{-1}. \qquad (7)$$

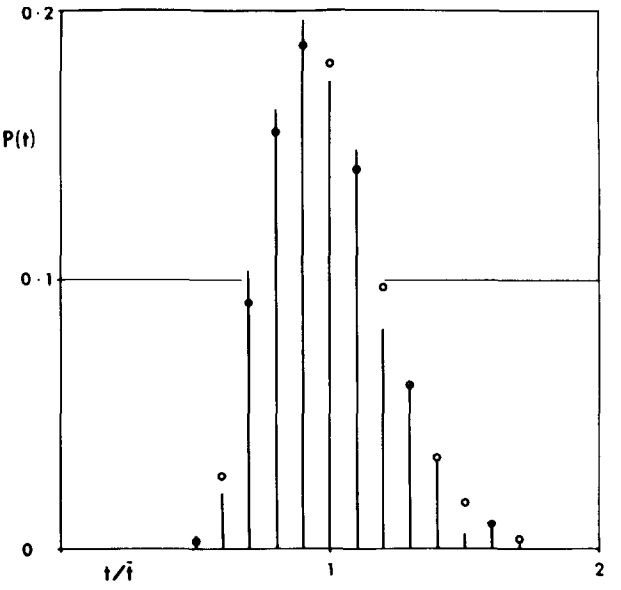

Fig. 5. Probability of arrival at an absorbing barrier 100 lattice points to the right of the origin. $U=2, L=1$ (see text). Vertical bars are simulation results, circles are theoretical points. Ordinates are probability, abscissas relative time.

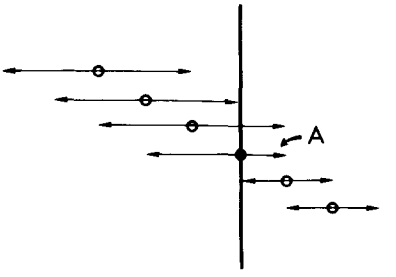

Fig. 6. Jumps in the vicinity of a discontinuous boundary. Reversing arrow at A corrects defect.

Fig. 6 shows a kind of defect which arises in the matching of jumps across an interface in a finite lattice system. Jumps and their magnitudes are indicated by the horizontal arrows for the case where $D_2 = 4D_1$. At the boundary (vertical heavy line in fig. 6) there is one jump to the right too many and one jump too few to the left. By reversing one jump (A in fig. 6) the situation is corrected. The result of a simulation of the 2-component system is shown in fig. 7 and, with it, a comparison of the simulation with the values of the corresponding expression in equation (7),

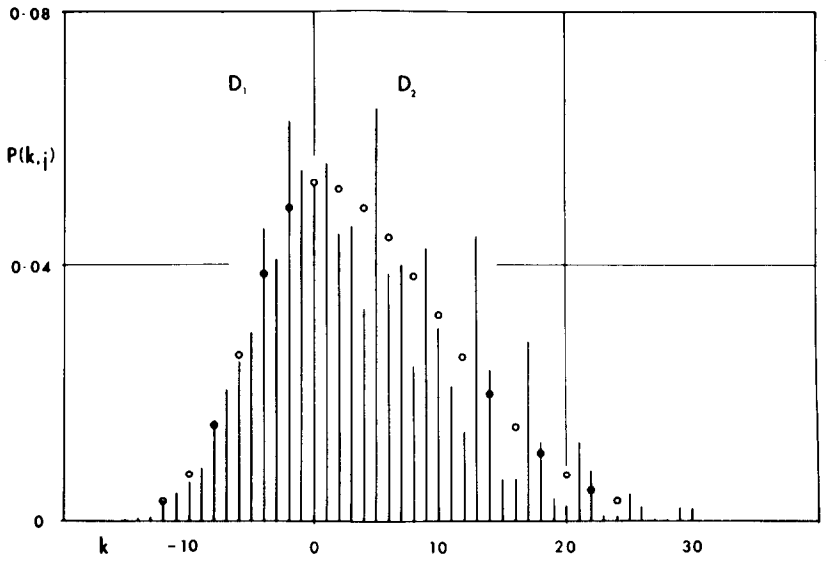

Fig. 7. Simulation in a composite system without drift. On the left $L = 1$ and on the right $L = 4$. Initial distribution is a delta function at the boundary. Vertical bars are simulation, circles are theoretical results.

Best results were obtained by expanding the scale of the simulation system 10-fold. Instead of a sudden jump at the boundary there was a progressive increase in the values over a space of 10 expanded lattice points in the boundary region. Finally the results were compressed 10-fold by adding together the sum of values in groups of 10 lattice points combined. Though rather fluctuant the simulation results appear to follow the theory in a satisfactory manner.

4. Practical considerations in Monte-Carlo simulation and computation

Many simulations are more complex than the simple ones which we have chosen here for illustration. Often many histories must be run and in many cases one must work in several dimensions. An increasing demand is thus set on the available memory space in the machine and the speed needed to obtain the requisite number of histories in any reasonable time. Economy is sometimes obtained by the use of special "variance reduction" procedures [8]. The term "random number" is in a sense a misnomer because, given the starting value and the algorithm for generating the sequence all members of the sequence are then determined. We have obtained random-digit pairs by generating a 20-digit decimal number and removing the digits pair-by-pair from one end to the other. The mid-squaring method for obtaining tandom numbers is hazardous as can be readily seen by squaring 625000 and examining the 6 middle digits. A popular algorithm for generating such numbers by the "power-residue" method has gained recent acceptance (see ref. [8]). Other methods have been suggested [15]. There is an important practical hazard which arises in the simulation of random processes which inherently contain fluctuations. Program errors can readily hide within these fluctuant results. They can be found only through thorough step-by-step program tracing.

References

[1] A.J.P.Martin and R.L.M.Synge, A new form of chromatogram employing two liquid phases. 1. A theory of chromatography, Biochem. J. 35 (1941) 1358.
[2] S.W.Mayer and E.R.Tompkins, Ion exchange as a separations method. IV. A theoretical analysis of the column-separation process, J. Am. Chem. Soc. 69 (1947) 2866.
[3] S.Ohashi and K.Kogushi, Calculation of peak position in gradient elution chromatography with an ion exchanger, J. Chromatog. 27 (1967) 214.

[4] W.Perl and F.P.Chinard, A convection-diffusion model of indicator transport through an organ. Circulation Res. 22 (1968) 273.
[5] C.W.Sheppard and L.J.Savage, The random walk problem in relation to the physiology of circulatory mixing, Phys. Rev. 83 (1951) 489.
[6] A.W.Marshall, Introductory note to H.A.Meyer (ed.), Symp. on Monte Carlo Methods (Wiley, New York, 1956) p. 1.
[7] J.B.H.Kemperman, The Passage Problem for a Stationary Markov Chain (University of Chicago Press, 1961).
[8] H.Kahn, Multiple quadrature by Monte Carlo methods, in: Mathematical Methods for Digital Computers, vol. 1, eds. A.Ralston and H.S.Wilf (Wiley, New York, 1965) p. 249.
[9] F.J.Oswald, Matrix inversion by Monte Carlo methods, in: Mathematical Methods for Digital Computers, vol. 1, eds. A.Ralston and H.F.Wilf (Wiley, New York, 1965) p. 78.
[10] J.R.McCord and R.M.Moroney, Probability Theory (MacMillan, New York, 1964).
[11] M.Kac, Random walk and the theory of Brownian motion, Am. Math. Monthly 54 (1947) 295.
[12] C.J.Tranter, Integral Transforms in Mathematical Physics (Wiley, New York, 1951).
[13] C.W.Sheppard, Basic Principles of the Tracer Methods (Wiley, New York, 1962).
[14] H.S.Carslaw and J.C.Jaeger, Conduction of Heat in Solids (Oxford University Press, London, 1947).
[15] O.Taussky and J.Todd, Generation and testing of pseudo-random numbers, in: Symp. on Monte Carlo Methods, ed. H.A.Meyer (Wiley, New York, 1956) p. 15.