up with N_1 as a free parameter which can be typed into the machine as occasion demands, no further information being needed. This elaboration of the program makes it possible to operate a lengthy sieve problem as a backlog workload without the need of a specially trained operator.

D. H. L.

[1] D. H. LEHMER, "The mechanical combination of linear forms," *Amer. Math. Monthly*, v. 35, 1928, p. 114–121, "A photo-electric number sieve," *Amer. Math. Monthly*, v. 40, 1933, p. 401–406, "A machine for combining sets of linear congruences," *Math. Annalen*, v. 109, 1934, p. 661–667.

[2] L. E. DICKSON, *History of the Theory of Numbers*, v. 2, Washington, 1923; New York, 1934, p. 57.

The Use of Large Intervals in Finite-Difference Equations

In a recent article SIR RICHARD SOUTHWELL[1] has challenged the general theory of finite differences and in particular the use of it in connection with the solution of differential equations by relaxation methods.[2] Opinions differ on the method of treatment of a numerical differentiation formula such as

$$h^2y_0'' = \left(\delta^2 - \frac{1}{12}\delta^4 + \frac{1}{90}\delta^6 - \cdots\right)y_0,$$

where h is the interval between pivotal points and where $\delta^{2n}y_0$ is the $2n^{\text{th}}$ central difference of y_0. This equation is replaced for the purpose of numerical solution by its equivalent

$$h^2y_0'' = (y_1 + y_{-1} - 2y_0) + \Delta\,(y_0),$$

where y_1, y_0, and y_{-1} refer to the values of y at $x_0 + h$, x_0, and $x_0 - h$ respectively and

$$\Delta(y) = \left(-\frac{1}{12}\delta^4 + \frac{1}{90}\delta^6 - \cdots\right)y.$$

I advocate the use of as large an interval as possible (consistent with convergence of the finite-difference equations and convenience of computation) calculating $\Delta(y)$ and incorporating it into the computation. Southwell prefers to use such a small interval that Δ is negligible. He states: "*Accuracy is not predictable from quantities computed from a finite-difference approximation unless the interval is small enough to justify belief in the convergence of the Taylor series:* in general the radius of convergence, though it exceeds one or two of the smallest intervals that are practicable, will not exceed many such intervals, and consequently *what have been claimed as closer approximations may in fact be less accurate.*"

The purpose of this note is to point out that, in the writer's opinion, this statement is unduly cautious and that its conclusion cannot be true if finite-difference equations are properly used.

In the first place it is not clear what role the Taylor series plays in finite-difference theory. If the Taylor series is not convergent, care must certainly be exercised; but finite-difference methods do not always break down in this case.[3] The point to be emphasized, however, is that an examination of the differences will tell us, in all practical cases, whether the finite-difference

equations we wish to use are valid and how many differences are significant at the particular interval used. It is stressed by Fox[2] that in all the finite-difference formulae used the differences must be convergent and that otherwise a smaller interval must be used.

By "convergence of differences" I mean that the differences of some order should oscillate about zero, with a maximum amplitude determined in the usual way by the binomial coefficients of that order.[4] The tabulated function can then be regarded as a polynomial, slightly perturbed by rounding errors of not more than half a unit at any point. The degree of the polynomial is the same as the order of the last significant difference, and the function can be interpolated, integrated, and differentiated by finite-difference formulae based on this polynomial representation, the error of the process being of the order of the first neglected term in the finite-difference equation.

The degree of the approximating polynomial varies with the number of figures to which the given tabular entries are rounded off and with the magnitude of the interval between successive pivotal points. The function $(1 + x^2)^{-1}$, for example, can be represented near the origin, at an interval of .1, by a quadratic to two decimals, a sextic to four, and so on. In each case interpolation and integration can be performed effectively to the same precision as that of the tabular entries, though derivatives suffer from the loss of significant figures in successive differences. In particular no estimate can be given for derivatives of the true function of orders greater than the degree of the approximating polynomial.

The fact that the true higher derivatives may be large and ever increasing as for example in the function $(1 + x^2)^{-1}$, is of no significance with regard to the use of finite-difference formulae as applied to the approximating polynomial. If the derivatives increase fast enough such formulae may be asymptotic, but again the error is of the order of the first term neglected. This is analogous to the treatment, in a recent paper by Fox & Miller[3], of a function whose Taylor series had zero radius of convergence at the origin; the series was in fact asymptotic.

Southwell does not use differences and bases his argument on the fact that increased accuracy is not guaranteed by increasing the order of the approximating polynomial, citing as one example the interpolation at unit interval of the function $(1 + x^2)^{-1}$. The first table gives the differences of this function, tabulated to three decimals at unit interval (the central differences at $x = 0$ are filled in from consideration of symmetry). Examination shows that the differences are diverging rapidly, that interpolation formulae can hardly give even one-figure accuracy in the function near $x = 0$, and that a much smaller interval is necessary.

x	10^3y					x	10^6y				
0	1000		−1000		2400	12	6897		285		
		−500		1200				−1015		−76	
1	500		200		−1200	13	5882		209		22
		−300		0				−806		−54	
2	200		200		−141	14	5076		155		16
		−100		−141				−651		−38	
3	100		59			15	4425		117		12
		−41						−534		−26	
4	59					16	3891		91		

The divergence of differences near $x = 0$, however, does not prevent the accurate use of finite-difference formulae at points sufficiently remote from this point. Central-difference formulae, for example, are based on polynomials whose origin is at the particular pivotal point considered, and the function $(1 + x^2)^{-1}$ can be quite correctly interpolated, differentiated, or integrated to at least four figures, at unit interval, in the range shown in the second of the tables.

Turning now to the rest of the quotation, the second part suggests that the largest interval at which the full finite-difference equation can give accuracy will never greatly exceed an interval for which its first term alone is adequate. This will depend, of course, on the particular function involved, but, even if the ratio of intervals is only two, the use of Δ is worth-while not only for economy of effort but also from the point of view of accuracy.

Computers know that the use of small intervals in step-by-step integration is dangerous, the tendency for rounding-off errors to accumulate being roughly proportional to the square root of the number of intervals used. In relaxation methods, similarly, a small interval involves a large number of simultaneous equations, and their accurate solution is a matter of considerable difficulty; often one or even two extra figures have to be retained in the coefficients and residuals. Practical views on this have been reinforced by a recent theoretical paper by TODD.[5]

In illustration of these points let us try to evaluate to four decimals the pathological function $(1 + x^2)^{-1}$ from the facts that it satisfies the differential equation

$$y'' + 4x\,(1 + x^2)^{-1}\,y' + 2\,(1 + x^2)^{-1}\,y = 0,$$

is symmetric about $x = 0$, and has the value $y\,(\frac{1}{2}) = .8000$. The finite-difference equation is given by

$$y_1\,(1 + 2x_0 f_0) + y_{-1}\,(1 - 2x_0 f_0) - y_0\,(2 - 2hf_0) + \Delta(y_0) = 0,$$

where

$$\Delta(y) = -4xf\left(\frac{1}{6}\mu\delta^3 - \frac{1}{30}\mu\delta^5 + \cdots\right)y + \left(-\frac{1}{12}\delta^4 + \frac{1}{90}\delta^6 - \cdots\right)y$$

and

$$f = h\,(1 + x^2)^{-1}.$$

At an interval $h = .25$ there are only five pivotal points, giving insufficient information about the differences. The next smallest convenient interval is .1 (assuming we wish to have $x = 0$ as a pivotal point), and at this interval we find the following first approximation $y^{(0)}$, with Δ neglected. Extra figures have been retained in the residuals and coefficients to ensure that this approximation is everywhere an accurate solution (to four figures)

to the given simultaneous equations. The evaluation is as follows:

x	$10^4 y^{(0)}$					Δ	R	$10^4 y^{(1)}$	Error in $10^4 y^{(0)}$
.0	10020		−200		22	−1.8	−1.4	10001	20
		−100		11					
.1	9920		−189		21	−1.9	−2.1	9902	19
		−289		32					
.2	9631		−157		12	−1.5	−1.5	9616	16
		−446		44					
.3	9185		−113		2	−1.0	− .8	9175	11
		−559		46					
.4	8626		−67		(−9)	− .2	− .2	8621	5
		−626		(37)					
.5	8000							8000	0

The differences look quite satisfactory, fifth and higher orders being apparently negligible, and the recorded values of Δ include contributions from third and fourth differences only. If these values are now inserted in the finite-difference equations, the remaining residuals are as shown in the column headed R, and their liquidation gives a better approximation $y^{(1)}$, whose maximum error is one unit. The skilled computer would notice further that the fifth and sixth differences, though small, have a definite flow and are not yet oscillating about zero. Their total contribution is only about $-.2$, but their effect is to reduce the maximum error in $y^{(1)}$ to half a unit in the last figure.

This work involved the solution of nine simultaneous equations with two sets of residuals and with a little intermediate computation. If Δ were ignored and the computer solved 19 linear equations at an interval of .05, he would still have maximum errors of five units.

It is very doubtful whether intuition alone could here decide the question of the size of the "ultimate interval" for which Δ is negligible, and, if the computer, still ignoring Δ, wished to make sure of his result by taking an interval of .025, he would still have errors of one unit, having solved no less than 39 linear equations in addition to computing the coefficients thereof. The labour of this process is excessive; the accurate solution of these 39 equations would involve keeping at least two extra figures in the coefficients and residuals.

I therefore maintain, with regard to the end of the statement, that a computer familiar with both the theory and practice of finite differences will never be in danger of claiming more accuracy than he has achieved. Either he has got the desired accuracy, or he will know that he has not got it (and perhaps cannot get it, as for example near a discontinuity or singular point). No question of prior knowledge of the convergence of the Taylor series arises, and such knowledge has certainly not been sought in any of my published examples. Examination of the differences, after a tentative solution has been found, provides the necessary information and is of course essential.

In this paper I have concentrated on functions of one variable, but the remarks apply with greater force to partial differential equations, particularly in the desirability, both for the sake of accuracy and a minimum of labour, of using conveniently large intervals.

L. Fox

National Physical Laboratory
Teddington, England
and NBSCL

[1] *Numerical Methods of Analysis in Engineering.* Edited by L. E. GRINTER, Macmillan Company, New York, 1949, chapter 4.

[2] L. Fox, "Some improvements in the use of relaxation methods for the solution of ordinary and partial differential equations," Royal Soc., London, *Proc.*, v. 190A, 1947, p. 31–59.

[3] L. Fox & J. C. P. MILLER, "Table-making for large arguments. The exponential integral," *MTAC*, v. 4, 1951, p. 163–167.

[4] See, for example, J. C. P. MILLER, "Checking by differences—I," *MTAC*, v. 4, 1950, p. 3–11.

[5] J. TODD, "The conditions of a certain matrix," Cambridge Phil. Soc., *Proc.*, v. 46, 1950, p. 116–118.

Monte Carlo Matrix Inversion and Recurrent Events

1. Introduction. Recently WASOW[1] has given a necessary and sufficient condition that one of two unbiased estimators of the inverse element of a given matrix has a smaller variance. Using the theory of recurrent events, we extend a result by FELLER[2] in order to generalize and reinterpret Wasow's condition. To do this let us consider a simple discrete Markov process with a finite number of states. Of these $m + 1$ states denoted by $0, 1, 2, \cdots, m$, we prescribe that the state named "0" is the only death state or sink—in the sense that the random walk ends when this state is reached. Let p_{ik} be the one-step transition probabilities for $i, k = 1, \cdots, m$, and let

$$p_{i0} = 1 - \sum_{k=1}^{m} p_{ik} \quad \text{for} \quad i = 1, \cdots, m.$$

Further, we assume that each $p_{i0} > 0$. When $n > 0$ define $p_{ik}^{(n)}$ as the probability that in a random walk starting at state i, state k is visited on the n-th step; and define $p_{ik}^{(0)} = \delta_{ik}$, where δ_{ik} is the Kronecker delta.

Now if $(q_{ik}) = (\delta_{ik} - p_{ik})$ is the m by m matrix whose inverse (q^{ik}) is desired, we may estimate each element q^{ik} by Monte Carlo methods as in WASOW[2], FORSYTHE & LEIBLER[3], and CURTISS[4], since

$$q^{ik} = \sum_{n=0}^{\infty} p_{ik}^{(n)}.$$

Two estimators have been investigated. The first estimator is the sample mean of the random variable s_{ik} where $s_{ik} = p_{ko}^{-1}$ if the random walk having started at state i stops just after visiting state k, and $s_{ik} = 0$ if otherwise. The second estimator is the sample mean of the random variable v_{ik} where v_{ik} is the number of visits to state k in a random walk having started at