In taking $\bar{\xi}_0 = (1, 1)$ we have then

$$|\bar{\xi}_\nu| = \sqrt{2}(9/25)^\nu. \tag{40}$$

Consider on the other hand the over relaxation with the value of $q = 10/9$. Here the components of the approximating vectors are to be computed from the equations

$$x_1^{(\nu+1)} = -x_1^{(\nu)}/9 - 2x_2^{(\nu)}/3, \quad x_2^{(\nu+1)} = -x_2^{(\nu)}/9 - 2x_1^{(\nu+1)}/3.$$

If we put $\xi_\nu = (x_1^{(\nu)}, x_2^{(\nu)})$ and assume again $\xi_0 = (1, 1) = \bar{\xi}_0$, we obtain as is readily verified

$$\xi_\nu = 3^{-2\nu-1}(3 - 24\nu, 3 + 8\nu) \qquad (\nu = 0, 1, \cdots).$$

Here we have

$$|\xi_\nu| \sim 8(10)^{\frac{1}{2}}\nu 9^{-\nu}/3 \qquad (\nu \to \infty).$$

We give in what follows a table of the initial values of $|\bar{\xi}_\nu|$ and $|\xi_\nu|$.

ν	$\lvert\bar{\xi}_\nu\rvert$	$\lvert\xi_\nu\rvert$
1	0.5091	0.8780
2	0.1833	0.2010
3	0.06598	0.03388
4	0.02375	0.005048
5	0.008551	0.0007037

Although the difference between q_0 and 1 is very small, in fact 1/9, the improvement is already observed at ξ_3 and becomes more and more pronounced from there on.

A. OSTROWSKI

American University, Washington, D. C.
University of Basle, Switzerland

This paper was prepared under a contract of the National Bureau of Standards with the American University, Washington, D. C.

[1] S. P. FRANKEL, "Convergence rates of iterative treatments of partial differential equations," *MTAC*, v. 4, 1950, p. 65–75.
[2] A. M. OSTROWSKI, *On the Linear Iteration Procedures for Symmetric Matrices*. NBS *Report* no. 1844, August 1952, p. 23, (68).

The Accuracy of Numerical Solutions of Ordinary Differential Equations

1. Introduction. The present paper describes a general method by which the random and systematic errors may be estimated of numerical solutions of any systems of ordinary differential equations. The errors arise from the accumulation of rounding-off errors, and from the use of erroneous formulas for performing the numerical integrations. The estimation is based on the properties of the solutions of the system of equations adjoint to the variational equations of the problem, and is applicable to any method of integration.

The present method was described by the author at a meeting of the American Astronomical Society on February 1, 1946 at Columbia University. Development of the method resulted from a conversation with CHARLES B. MORREY, JR., who explained to the author properties of the solutions of adjoint equations with which the author was not then familiar. The procedure seems intuitively obvious and straightforward, and it has not been published earlier both for this reason and because it was understood that HANS RADEMACHER was planning to publish a similar and possibly independent treatment. It now appears, however, that Rademacher[1] was concerned with the accuracy of particular methods of integration, while the present method is applicable to any integration procedure that may be employed. There are still other methods for estimating the accuracy of numerical solutions of special types of ordinary differential equations. For example, BROUWER[2] has made special studies of the accuracy of numerical integrations, by the Crommelin-Cowell method, of the orbital differential equations of dynamical astronomy.

A virtue of the present method is its generality; but there are alternative general methods, possibly just as good, which may not have been published. The author understands from conversations with L. H. THOMAS, that he has made use of general procedures not involving the adjoint equations. The main justification for publishing a description of the author's procedure is his hope that it may help others to select computational procedures, for numerical integrations, that will yield results of desired accuracy.

2. The Adjoint Equations. Consider the system of n first-order differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{i,j}x_j + b_i; \quad i,j = 1,2,\cdots,n \tag{1}$$

where the n^2 quantities $a_{i,j}$ and the n quantities b_i may vary with the independent variable, t. Let λ_i be a set of variables satisfying the adjoint system of equations

$$-\dot{\lambda}_i = \sum_{j=1}^{n} a_{j,i}\lambda_j. \tag{2}$$

Since

$$\begin{aligned}(d/dt)\sum_{i=1}^{n} x_i\lambda_i &= \sum_{1}^{n} \dot{x}_i\lambda_i + \sum_{1}^{n} x_i\dot{\lambda}_i \\ &= \sum_{i,j} a_{i,j}x_j\lambda_i - \sum_{i,j} a_{j,i}x_i\lambda_j + \sum_{i} b_i\lambda_i \\ &= \sum_{i} b_i\lambda_i\end{aligned} \tag{3}$$

it follows that

$$\sum_{i=1}^{n} x_i(A)\,\lambda_i(A) = \sum_{i} x_i(0)\,\lambda_i(0) + \int_0^A \sum_{i} b_i\lambda_i\,dt. \tag{4}$$

3. Application. From any system of ordinary differential equations there can be derived a set of variational equations of the form (1) where the $x_i(t)$'s

are differences, between the exact solution of the original equations corresponding to the desired initial conditions, and any neighboring exact solution not subject to the desired initial conditions. If a single error were made in the course of solving the original system of equations by a scheme of stepwise integration that was otherwise perfect, the solution would be exact before the error was made; for later values of the independent variable the solution would still be an exact solution of the original differential equations but would correspond to altered initial conditions. The x_i's would all be zero before the error, and would grow after it in accordance with the equations of the form (1), with b_i's that were zero for all steps except the one in which the error was made.

If, because of defective methods of calculation or for any other reason errors $\epsilon_i(t)$ are introduced into the i'th variable x_i at a particular step ($t - w$ to t, say) in the solution of the original system of differential equations, one may consider the ϵ's to have been introduced by b_i's in (1) such that

$$\epsilon_i(t) = \int_{t-w}^{t} b_i(t)\, dt$$

and that are zero outside of the interval $(t - w, t)$. One may consider approximately that

$$b_i(t) = (1/w)\epsilon_i(t)$$

throughout the interval and therefore approximately that

$$\int_{t-w}^{t} b_i(t)\, \lambda_i(t)\, dt = \epsilon_i(t)\, \lambda_i(t).$$

Thus by (4) the resulting final error (at $t = A$, say) in a particular variable (e.g., x_1 say) is

$$x_1(A) = \sum_{i=1}^{n} \epsilon_i(t)\, \lambda_i(t)$$

and if errors are introduced at all steps

$$x_1(A) = \sum_{\text{all steps}} \sum_{i=1}^{n} \epsilon_i(t)\, \lambda_i(t) \tag{5}$$

provided that the λ's are any solution of (2) satisfying the boundary conditions

$$\lambda_1(A) = 1; \quad \lambda_j(A) = 0, \quad j \neq 1.$$

4. Truncation and Rounding Errors. Equations (4) and (5) provide means for predicting the errors of numerical solutions of systems of ordinary differential equations. Rough solutions of the equations (2), based on a rough solution of the original equations, are in practice adequate. Rounding-off errors of the usual hand-made variety are drawn from populations whose means are zero, and whose individuals are uniformly distributed from $-1/2$ to $1/2$ in units of the last digit. Their variance is thus $1/12$ in such units. The resulting variance of a final value like $x_1(A)$ is given by

$$\sigma^2_{x_1(A)} = \sigma_1^2 \sum \lambda_1^2(t) + \sigma_2^2 \sum \lambda_2^2(t) + \cdots + \sigma_n^2 \sum \lambda_n^2(t) \tag{6}$$

where σ_i^2 is the variance of the rounding-off errors introduced into the variable x_i at any step, and where the sums are taken over all steps. The preceding equation is general; if the rounding-off errors are not of the usual hand-made sort it is still valid. If the rounding-off errors do not come from populations with zero means, then a bias, or systematic error is introduced whose final value has the population mean

$$\overline{x_1}\,(A) = \sum_t \sum_i M_i(t)\,\lambda_i(t)$$

where $M_i(t)$ is the population mean of the error $\epsilon_i(t)$, conceivably a function of t.

Besides rounding-off errors, "truncation" errors are introduced by the circumstance that the formulas employed for integrations are erroneous. Whatever the formulas are, and however they are employed, iterated or not, any particular method of integration applied to a particular system of differential equations always corresponds to the exact solution of a system of difference equations rather than of the original differential equations. The particular method of integration thus corresponds to an exact solution of a system of differential equations somewhat different from the original differential equations. It is always possible to evaluate, approximately, the differences between the original differential equations and those that the scheme is exactly solving, then to find the appropriate b_i's in the variational equations of the form (1), and finally to apply (4) to predict the final errors, thus

$$x_1(A) = \int_0^A \sum_i b_i(t)\,\lambda_i(t)\,dt \tag{7}$$

in which, as usual, the λ's must be chosen to satisfy the boundary conditions

$$\lambda_1(A) = 1; \quad \lambda_j(A) = 0, \quad j \neq 1.$$

Alternatively, one can find appropriate truncation errors $\epsilon_i(t)$ and then apply equation (5).

For planning purposes, no great accuracy in the calculations of accuracy is necessary, and no great accuracy should be sought. Even rough calculations are expected to suffice to decide how many digits, what size of steps, and what scheme of integration to employ.

5. Example. An example, suggested by WERNER LEUTERT, will be given of the use of the preceding method by an application to the non-linear differential equation of the first order

$$\dot{y} = (3/2)\,t\,y^{-1/3} \tag{8}$$

which can be integrated analytically but which will be treated as though it can not be. Suppose that one wishes to integrate this equation numerically, starting with the value

$$y(1) = 1,$$

as far as $y(5)$; and that one wishes the value $y(5)$ to be accurate "to the third place" of decimals. One wishes to use the scheme of integration defined

by the approximate formula

$$\nabla y = \left[1 - \frac{\nabla}{2} - \frac{\nabla^2}{12}\right] w\dot{y} \tag{9}$$

in which the differences are backward, or ascending, and in which w is the length of a step. One wishes to determine w, and the number of decimals to retain in the calculations.

The variational equation corresponding to (1) is

$$\dot{x} = -\frac{1}{2} t\, y^{-4/3}\, x \tag{10}$$

and the adjoint equation corresponding to (2) is thus

$$\dot{\lambda} = \frac{1}{2} t\, y^{-4/3}\, \lambda. \tag{11}$$

A rough integration of (8) must first be accomplished, and then a rough integration of (11) to find λ. Such integrations have been accomplished by the aid of a ten-inch slide rule, and the results appear in the first four columns of the following table:

t	y	2.27λ	λ	$\lambda D^4 y$
1	1	1	.44	.25
2	2.80	1.44	.63	.06
3	5.15	1.76	.78	.03
4	7.95	2.03	.89	.01
5	11.15	2.27	1.00	.01

No attempt has been made to obtain results correct to the second place of decimals, although two places were retained. The integration for λ with the starting value unity led to a value 2.27 at $t = 5$; the fourth column contains λ adjusted to have the value unity at $t = 5$.

Consideration of the difference formula (9) shows that it corresponds substantially to the differential equation

$$Dy - (1/24)\, w^3\, D^4 y + (11/720)\, w^4\, D^5 y + \cdots = (3/2)\, t\, y^{-1/3} \tag{8'}$$

instead of to equation (8), so that approximately

$$b(t) = (1/24)\, w^3\, D^4 y$$

or

$$\epsilon(t) = (1/24)\, w^4\, D^4 y;$$

these results could have been obtained directly from the term of lowest order that has been omitted from the right-hand member of equation (9). By equation (5) or equation (7) the truncation error at $t = 5$ is

$$x(5) = (w^3/24) \int_1^5 D^4 y(t)\, \lambda(t)\, dt.$$

Values of D^4y are obtained from equation (8); values of $D^4y(t)\ \lambda(t)$ are tabulated above; a quadrature furnishes the result

$$x(5) = w^3/120. \tag{12}$$

It is noticed that a unit error in $w\dot{y}$ at any stage introduces a unit error in y when the scheme of integration is (9). Hence by equation (6) the variance of $y(5)$ arising from rounding errors is

$$\begin{aligned}\sigma^2_{y(5)} &= (1/12)\sum_{t=1}^{5}\lambda^2(t)\\ &= (1/12\,w)\int_1^5 \lambda^2(t)\,dt\\ &= 1/5\,w\end{aligned}$$

by quadratures; very nearly, in units corresponding to the last place retained in $w\dot{y}$.

To obtain a value of $y(5)$ accurate "to the third place" of decimals, one equates the right-hand member of equation (12) to 0.0005 and solves for w. It is found that w is .39. One can adopt a value $w = .4$, if one will tolerate a bias error of 0.00053 in $y(5)$; this is tolerated, and the value $w = .4$ is adopted. A number fairly rounded to the nearest 0.001 has a rounding error whose variance is 1/12 in units of the sixth place. It is therefore reasonable to require that the variance of $y(5)$ from the accumulation of rounding errors should be smaller than 1/12 in the sixth place. If only three decimals were retained in the values of $w\dot{y}$ then the variance of $y(5)$ would be $1/5w$ or 1/2 in the sixth place, which is too large to be acceptable. With four decimals, the variance is 1/2 in the eighth place, or 1/200 in the sixth, which is better than is needed. Therefore four places of decimals should be retained in values of $w\dot{y}$.

The definitive integration of equation (8) by the scheme (9) was next accomplished, with steps of 0.4 and with four places of decimals. The value obtained was

$$y(5) = 11.1810.$$

The correct value of $y(5)$, obtained by analytical solution of (8), is

$$\begin{aligned}y(5) &= 5^{3/2}\\ &= 11.18034\cdots\cdots\end{aligned}$$

showing that the error of $y(5)$ obtained by the numerical integration is 0.00066, and that $y(5)$ is substantially as accurate as was desired and as was predicted.

THEODORE E. STERNE

Ballistic Research Laboratories
Aberdeen Proving Ground
Maryland

[1] See, for instance, HANS RADEMACHER, "On the accumulation of errors in numerical integration on the Eniac," lecture 19 of *Theory and Techniques for Design of Electronic Digital Computers*, v. 2, Moore School of Electrical Engineering, University of Penn., 1947.

[2] DIRK BROUWER, "On the accumulation of errors in numerical integration," *Astronomical Jn.*, v. 46, p. 149, 1937.