Error Bounds for the Runge-Kutta Single-Step Integration Process*

JOHN W. CARR III

University of Michigan, Ann Arbor, Michigan

Although numerous writers have stated that the class of single-step ("Runge-Kutta"—like) methods of numerical integration of ordinary differential equations are stable under calculational or round-off error, no one has given formal equations for the bounds on the propagated error to indicate this stability. Rutishauser [1] justifies the stability by noting that there is only one solution to the approximating difference equation, and Hildebrand [2] calculates a propagated error bound for the simplest (Euler) case. However, the latter bound does not indicate the stability for even that case.

It is the purpose of this paper to investigate this "stability" of the Kutta fourth order procedure for integration of the ordinary differential equation

$$\frac{dy}{dx} = f(x, y), \tag{1}$$

where $f(x, y)$ possesses a continuous first-order partial derivative with respect to y throughout a region D in which the integration is to take place. (By alteration of the proof below, this condition can be replaced by a Lipschitz condition.) Since the Kutta process is the most complicated of such single-step procedures, it should be apparent that similar error bounds can be derived for the various other single-step methods of same or lower order (and in particular the Gill variant method, probably most often used in machine integration because of the storage savings). It is plausible that such error bounds can also be extended to the stable (extrapolation) multi-step methods, such as the Adams method, and to systems of ordinary differential equations.

If the variational equation

$$\frac{d\eta}{dx} = \frac{\partial f(x, y)}{\partial y} \eta \tag{2}$$

for the above ordinary differential equation has $f_y(x, y) < 0$ throughout a region D the ordinary differential equation is termed *stable* in D, and for small enough variations in the initial conditions the absolute value of the propagated error decreases with growing x.

Todd [3], Milne [4], Rutishauser [1], and others have shown that numerous multi-step numerical integration techniques are unstable in that even when the differential equation is stable, the difference equation will introduce spurious solutions for *any* step-size h. For the Kutta fourth order process, as seen below, this is not the case; for the stable differential equation the propagated error in the

difference approximation remains bounded for small enough (but not too small!) step-size h; and for a given value of x, bounds on the propagated error decrease to a minimum given as a function of the round-off error as the step-size is decreased. Similar statements can be proved (but are not proved here) for other single-step processes. For no round-off (an "infinite word-length machine") the process converges as h goes to zero.

In addition, an algorithm for determining the step-size as a function of the partial derivative is given below so as to keep the propagated error within a given bound.

The classical Kutta procedure [5] gives the value of y at the $(i + 1)$th step in terms of the value at the ith step and the step-size h as follows:

$$y_{i+1} = y_i + h/6(k_1 + 2k_2 + 2k_3 + k_4) + O(h^5)$$
$$k_1 = f(x_i, y_i)$$
$$k_2 = f(x_i + h/2, y_i + k_1 h/2) \tag{3}$$
$$k_3 = f(x_i + h/2, y_i + k_2 h/2)$$
$$k_4 = f(x_i + h, y_i + k_3 h)$$

When the $O(h^5)$ term is neglected, the value of y_{i+1}, here called y_{i+1}^t, is only an approximation.

An error bound for the truncation error at each step is given by Bieberbach [6, 7] and Lotkin [8]. This guarantees that for h small enough the truncation error at a single step, starting at y_i, will be bounded by

$$| y_{i+1} - y_{i+1}^t | \leq C_{i+1} h^5, \tag{4}$$

where $C_{i+1} \geq 0$ is a function only of i, the function $f(x, y)$, and its partial derivatives of the first three orders; and y_i^t is the true solution to the equations (3) truncated so that the $O(h^5)$ term does not appear. If the function and its derivaties of the first three orders exist and are bounded throughout a region, then all C_i would be bounded in the region.

Suppose there exists an error ϵ_i in y_i at the ith step, which might be due either to previous truncation or round-off error. Then the calculated value y_{i+1}^* (as opposed to the true value if ϵ_i were zero) at the next step is given, after several applications of the mean value theorem, by

$$y_{i+1}^* = y_i + \epsilon_i + h/6 \left\{ k_1 + \frac{\partial f}{\partial y} \epsilon_i + 2k_2 + 2 \left[\frac{\partial f}{\partial y} + \left(\frac{\partial f}{\partial y} \right)^2 h/2 \right] \epsilon_i \right.$$
$$+ 2k_3 + 2 \left[\frac{\partial f}{\partial y} \left(1 + \frac{\partial f}{\partial y} h/2 \left(1 + \frac{\partial f}{\partial y} h/2 \right) \right) \right] \epsilon_i \tag{5}$$
$$+ k_4 + \left[\frac{\partial f}{\partial y} \left(1 + \frac{\partial f}{\partial y} h \left(1 + \frac{\partial f}{\partial y} h/2 \left(1 + \frac{\partial f}{\partial y} h/2 \right) \right) \right) \right] \epsilon_i \right\}$$

where the various partial derivatives are evaluated within a rectangle given by $x_i \leq x \leq x_i + h$; $| y_i - y | \leq Qh + | \epsilon_i |$ where Q is given below. Consider the

true solution $y_i{}^t$ to the difference equation. During its evaluation, in order that all values of $f(x, y)$ used in its calculation lie within a region D, the true solution should not approach the y-boundaries of D closer than

$$| y_i{}^t - y | > Qh,$$

where if Q is chosen as $Q \geqq \max_{x, y \epsilon D} | f(x, y) | \geqq \max (| k_1/2 |, | k_2/2 |, | k_3 |)$ then certainly only values within D will be involved in the calculations.

The same argument holds for the difference equation containing error ϵ_i at any step. If such a solution does not approach within closer than $Qh + | \epsilon_i |$ to the y-boundary of D (such a region will be called D^* below), then the true solution will not approach closer than Qh to the boundary. In the latter case the rectangle within which the partial derivatives are to be evaluated will always be within the given region D. This gives for the propagated error at the $(i + 1)$th step (if y_{i+1} is the value at step $i + 1$ given no error at step i, and y^*_{i+1} is the value at step $i + 1$ with the error ϵ_i present):

$$\eta_{i+1} = y^*_{i+1} - y_{i+1} = \epsilon_i \left[1 + h/6 \left(\frac{\partial f}{\partial y} + 2 \frac{\partial f}{\partial y} + 2 \frac{\partial f}{\partial y} + \frac{\partial f}{\partial y} \right) \right.$$

$$+ h^2/6 \left(\frac{\partial f}{\partial y} \frac{\partial f}{\partial y} + \frac{\partial f}{\partial y} \frac{\partial f}{\partial y} + \frac{\partial f}{\partial y} \frac{\partial f}{\partial y} \right) \tag{6}$$

$$\left. + h^3/12 \left(\frac{\partial f}{\partial y} \frac{\partial f}{\partial y} \frac{\partial f}{\partial y} + \frac{\partial f}{\partial y} \frac{\partial f}{\partial y} \frac{\partial f}{\partial y} \right) + h^4/24 \left(\frac{\partial f}{\partial y} \frac{\partial f}{\partial y} \frac{\partial f}{\partial y} \frac{\partial f}{\partial y} \right) \right]$$

where the partial derivaties are written out to indicate that each is evaluated at what may be a different point in the rectangle. One can now prove the following:

THEOREM: *If $\partial f/\partial y$ is continuous, negative, and bounded from above and below throughout a region D in the (x, y) plane,*

$$-M_2 < \frac{\partial f}{\partial y} < -M_1 < 0, \qquad \text{where} \qquad M_2 > M_1 > 0,$$

then for a maximum error (truncation, or round-off, or both) E in absolute value at each step the Kutta fourth-order numerical integration procedure has total error at the i-th step, i arbitrary, in the region D^*, of

$$| \epsilon_i | \leqq \frac{2E}{hM_1} \tag{7}$$

where the step-size h is taken to be $h < \min (M_1/M_2{}^2, 4M_1{}^3/M_2{}^4)$. (Obviously if D extends to infinity, the boundaries of D and D^* will coincide at infinity. Better (but more complex) bounds could certainly be obtained on Q and therefore D^*.)

PROOF: For h as given, the multiplier of $| \epsilon_i |$ within the absolute value signs on the right-hand side of (8) below is positive and the inequality does indeed hold, by applying the bounds on f_y of the theorem to (6) above. In fact, for h as given

$$| \eta_{i+1} | = | y^*_{i+1} - y_{i+1} | \leqq | \epsilon_i | \left| 1 - hM_1 + \frac{(hM_2)^2}{2!} - \frac{(hM_1)^3}{3!} + \frac{(hM_2)^4}{4!} \right| \tag{8}$$

$$\leqq | \epsilon_i | | 1 - hM_1/2 |.$$

The total error at the $(i + 1)$th step is, in absolute value,

$$| \epsilon_{i+1} | = | \rho_{i+1} | + | \eta_{i+1} | \leqq E + | \epsilon_i | (1 - hM_1/2) \qquad (9)$$

where ρ_{i+1} is the sum of the round-off and truncation error in that step and η_{i+1} the propagated error, $1 - hM_1/2 < 1$, and $E > | \rho_i |$ for all i. For $i = 0$, $\epsilon_0 = 0$, and therefore certainly

$$| \epsilon_0 | \leqq \frac{2E}{hM_1}. \qquad (10)$$

If for a given i, $| \epsilon_i | \leqq 2E/hM_1$, then from (9)

$$| \epsilon_{i+1} | \leqq E + \frac{2E}{hM_1} (1 - hM_1/2) \leqq \frac{2E}{hM_1}, \qquad (11)$$

and the proof holds by induction[1] on i.

Note that, if ρ_i consists only of truncation error,

$$| \epsilon_{i+1} | \leqq \frac{2Ch^5}{hM_1}, \qquad (12)$$

where $C > C_i$ for all i (if sufficient bounded derivatives on $f(x, y)$ are assumed), and

$$\lim_{h \to 0} | \epsilon_{i+1} | = 0. \qquad (13)$$

On the other hand, if ρ_i consists of round-off error ρ^* also, bounded by $| \rho^* |$ for all i, then

$$\lim_{h \to 0} | \epsilon_{i+1} | \leqq \frac{2(| \rho^* | + Ch^5)}{hM_1} = \infty, \qquad (14)$$

provided $| \rho^* |$ is bounded from below by a positive non-zero value, which is generally true for fixed word-length computers; and the error bound is of no value in the limit. In that case, for h small enough

$$| \eta_{i+1} | = | y_{i+1}^* - y_{i+1} | \geqq | \epsilon_i | (1 - hM_2) \geqq K | \epsilon_i |, \qquad (15)$$

where now $K = (1 - hM_2)$. If, for example, $\rho_k > R > 0$, for all k, one obtains

$$\begin{aligned} | \epsilon_{i+1} | &= \rho_{i+1} + | \eta_{i+1} | \\ &\geqq \rho_{i+1} + K(\rho_i + K(\rho_{i-1} + \cdots + K^{i+1}\rho_0) \cdots) \\ &\geqq R(1 + K + K^2 + \cdots + K^{i+1}) \\ &\geqq R(i + 2)K^{i+1} \end{aligned} \qquad (16)$$

$$\therefore \lim_{h \to 0} | \epsilon_{i+1} | \geqq (i + 2)R \qquad (17)$$

Therefore, in the case $\rho_k > R > 0$, the error can increase without limit when $(i + 1)h$ is held constant. Thus, as perhaps seems obvious from the start, de-

---

[1] The author is grateful to a referee for suggesting this argument.

creasing the step-size will in this case, speaking roughly, improve the propagated error until the point at which the round-off and truncation error have the same order of magnitude. If on the other hand the round-off error can be made a function of the step-size of order h^k, where $k > 1$, then the total error will go to zero as $h \to 0$, or will remain bounded in the case $k = 1$. Users of variable word-length digital computing machines should therefore increase their word-length as h is decreased in order to avoid this lower limit due to round-off error. Users of fixed word-length computers must use multi-precision. This result is not surprising; a specific algorithm for doing it is given below.

If f_y is allowed also to be zero or positive, but bounded throughout D,

$$0 \leqq \left| \frac{\partial f}{\partial y} \right| < M. \tag{18}$$

Then the propagated error in D^* at the $(i + 1)$th step, due to a given (round-off or truncation error) ϵ_i at the ith step, is

$$| y^*_{i+1} - y_{i+1} | \leqq | \epsilon_i | \left(1 + hM + \frac{(hM)^2}{2} + \frac{(hM)^3}{3!} + \frac{(hM)^4}{4!} \right) \tag{19}$$

or

$$| \eta_{i+1} | \leqq | \epsilon_i | e^{hM}, \tag{20}$$

where η_{i+1} is the propagated error and E and h are given as before; and one obtains an error bound similar to most classical ones

$$| \epsilon_{i+1} | \leqq E(1 + K + K^2 + \cdots + K^i) \leqq E \frac{K^{i+1} - 1}{K - 1} = E \frac{e^{(i+1)hM} - 1}{e^{hM} - 1}, \tag{21}$$

where $K = e^{hM}$.

In both the stable and unstable case, ignoring round-off, the process is obviously convergent in D as $h \to 0$, since D^* approaches D as $h \to 0$ and $| \epsilon_{i+1} | \to 0$. For the case $f_y < 0$ equations (7) and (14) above give an algorithm for a bound on the stepsize h; a similar analysis could be based on equation (21). Suppose the round-off error per single-precision step is bounded by $| \rho^{*(1)}(h) |$, per n-precision step by $| \rho^{*(n)}(h) |$. (Here n could just as well be number of digits as number of words.) Then, since

$$| \epsilon_{i+1} | \leqq \frac{2(| \rho^{*(n)}(h) | + Ch^5)}{hM_1}, \tag{22}$$

for $| \epsilon_{i+1} |$ required to be less than $B > 0$, choose $h < M_1/M_2^2$ such that

$$2 Ch^5/hM_1 < B/2 \tag{23}$$

or

$$h^4 < \frac{M_1 B}{4C}. \tag{23a}$$

Then for that value of h, choose the value of n such that

$$| \rho^{*(n)}(h) | < BhM_1/4. \tag{24}$$

This will guarantee $|\epsilon_{i+1}| < B$. Obviously, such an h is not necessarily the best possible. The computation of $|\rho^{*(n)}(h)|$, while intricate, could very well be done by the computer itself, if procedures for forming formal derivatives of machine functions were included in whatever compiling techniques were used with the routine. If the calculated solution approached the boundaries of D^*, the step-size h would have to be decreased, and the values of h accordingly readjusted in order to continue to apply the theorem.

If h and n are picked in this fashion, the numerical machine solution can be made as close to the true solution as desired, and the process will converge. As for "stability" of the solution process under round-off, as would appear intuitive, the propagated error can be guaranteed to go to zero with the step-size h only if the round-off error per step varies with h as $O(h^k)$ with $k > 1$. The Kutta process is stable in the sense that for a stable differential equation, for a given h, the error will be bounded by (7), where E is given *without round-off* error by (12) and Bieberbach's bound (4). In the latter case, as $h \to 0$, the truncation error goes to zero, which certainly is not true for many multi-step methods.

BIBLIOGRAPHY

1. RUTISHAUSER, HEINZ, Uber die Instabilität von Methoden zur Integration gewöhnlicher Differentialgleichungen *Z.A.M.P.*, *III*, 1952, pp. 65–74.
2. HILDEBRAND, F. B. *Introduction to Numerical Analysis*, McGraw-Hill, 1956, pp. 233–239.
3. TODD, JOHN, Solution of Differential Equations by Recurrence Relations, *M.T.A.C. 4*, pp. 39–44 (1950).
4. MILNE, W. E., *Numerical Solution of Differential Equations*, John Wiley, 1953, pp. 21–24.
5. KUTTA, WILHELM, Beitrag zur näherungsweisen Integration totaler Differentialgleichungen, *Zeitschr. für Math. u. Phys.*, *46*, 1901, pp. 435–453.
6. BIEBERBACH, LUDWIG, *Theorie der Differentialgleichungen*, Dover, 1944, pp. 53–55.
7. BIEBERBACH, LUDWIG, On The Remainder of the Runge-Kutta Formula, *Z.A.M.P.*, *II*, 1951, pp. 233–248.
8. LOTKIN, MAX. On the accuracy of Runge-Kutta's method, *M.T.A.C. 5*, 1951, pp. 128–133.