The Use of Richardson Extrapolation in One-Step Methods with Variable Step Size*

By F. G. Lether

1. Introduction. One of the objections to the use of a one-step method to integrate a system of ordinary differential equations is that an estimate of the accumulated truncation error is difficult to make. If an attempt is made at appraising the truncation error, it is usually confined to an approximate evaluation of the local truncation error. A scheme for estimating the local truncation error, devised by Richardson [3], is based on the results of numerical integrations with steps h and $h/2$. The use of Richardson's extrapolation is well known (see, for example, [1, p. 81], [2, p. 238]). It is the purpose of this paper to show that it is possible to use the Richardson extrapolation procedure to form a useful estimate of the accumulated truncation error for a general one-step method even when the step size is allowed to vary. By using the estimate for the accumulated truncation error the accuracy of the numerical solution can be increased. Numerical examples to illustrate the estimation procedure are included.

2. Problem Formulation. We use the results of Henrici [1] on the asymptotic behavior of the accumulated truncation error. Unless otherwise noted all capital letters in the following mathematical relations denote vectors or vector-valued functions and lower case letters denote scalars.

Let the system of n differential equations be given by

$$\frac{dY}{dx} = F(x, Y), \qquad Y(a) = Y_0, \quad x \in [a, b]. \tag{2.1}$$

Let $v(x)$ be a piecewise continuous function of x such that $0 < v(x) \leqq 1$, where $x \in [a, b]$. We define the mesh points x_k *by*

$$x_0 = a,$$
$$x_{k+1} = \min\,(b, x_k + h_0 v(x_k)), \qquad k = 0, 1, 2, \cdots,$$

where h_0 is a constant basic stepsize.

The differential equations (2.1) are integrated numerically from a to b using mesh points introduced by the function $v(x)$. The one-step method is defined by

$$Y_{k+1} = Y_k + h_0 v(x_k) I(x_k, Y_k; h_0 v(x_k)) \tag{2.2}$$

where I is the *increment function.*

We denote by $Y(x_k)$ the exact solution of the initial value problem defined by (2.1) at the mesh point x_k. The accumulated truncation error at the mesh point x_k is defined by

$$E_k = Y_k - Y(x_k). \tag{2.3}$$

Received August 24, 1965.
* This research was partially supported by National Science Foundation Grant GP-2312.

Let $G(x) = (g_{ij}(x))$ be a matrix with components

$$g_{ij}(x) = \frac{\partial f_i(x, Y(x))}{\partial y_j}, \qquad i, j = 1, 2, \cdots, n,$$

where f_i is the ith component of F. If the functions involved are sufficiently smooth then it is shown in [1, p. 131] that there exists a *principal error function* $Q(x, Y)$. Henrici [1, p. 136] shows that if p is the *exact order* of the one-step method defined by I then

$$E_k = E(x_k)h_0^p + O(h_0^{p+1}) \tag{2.4}$$

where $E(x)$, called the *magnified error function*, is the solution of the initial value problem

$$E'(x) = G(x)E(x) + [v(x)]^p Q(x, Y(x)),$$
$$E(a) = 0.$$

Let Y_n and Z_n denote respectively the results obtained by integrating (2.1) twice from a to x_n by (2.2) with basic step lengths h_0 and h_0/i, where $i > 1$ is a constant and would for convenience normally be chosen as an integer. We use (2.4) to express the accumulated truncation error as a weighted difference between Y_n and Z_n.

It follows from (2.4) and (2.3) that

$$\begin{aligned} Y_n - Y(x_n) &= E(x_n)h_0^p + O(h_0^{p+1}), \\ Z_n - Y(x_n) &= E(x_n)(h_0/i)^p + O(h_0^{p+1}). \end{aligned} \tag{2.5}$$

Hence

$$Y_n - Z_n = (1 - 1/i^p)E(x_n)h_0^p + O(h_0^{p+1}). \tag{2.6}$$

From the first equation in (2.5) and (2.6) we obtain the equation

$$E_n = \frac{i^p}{i^p - 1}(Y_n - Z_n) + O(h_0^{p+1}).$$

By using (2.3) we obtain the Richardson extrapolation to the true solution at the mesh point x_n.

$$Y(x_n) = \frac{i^p Z_n - Y_n}{i^p - 1} + O(h_0^{p+1}).$$

The right sides of the relations

$$E_n \doteq \frac{i^p}{i^p - 1}(Y_n - Z_n) \tag{2.7}$$

and

$$Y(x_n) \doteq \frac{i^p Z_n - Y_n}{i^p - 1} \tag{2.8}$$

estimate, respectively, the accumulated truncation error and the true solution at x_n with an error whose order exceeds the order of the one-step method by one.

3. Computational Considerations. The above analysis assumed there was no round-off error. If Y_n and Z_n are contaminated by round-off error which is comparable in magnitude to the truncation error then the predicted accumulated truncation error (2.7) may be unreliable. In practice one carries enough digits in the computation to made the round-off error negligible in comparison to the truncation error.

FORTRAN IV programs utilizing double precision floating point arithmetic (approximately 16 decimal digits) were written for the IBM 7040 binary computer. The one-step methods considered were Euler's method [1, p. 9], the Heun method [1, p. 67], and the classical fourth-order Runge-Kutta method [1, p. 68]. The exact orders of the Euler, Heun and Runge-Kutta methods are $p = 1$, $p = 2$, and $p = 4$, respectively. A value of $i = 2$ was used in (2.7) and (2.8). The respective vectors Y_n and Z_n were obtained by simultaneous numerical integration of the system (2.1) with basic step size h_0 and the system

$$\begin{aligned} \frac{dZ}{ds} &= F(s, Z), \\ Z(a) &= Y_0, \end{aligned}$$

with basic step size $h_0/2$.

4. Numerical Results. Five differential equations whose solutions are known were considered. All of the data in the following tables is correct to the number of digits given. We denote the predicted accumulated truncation error by P_n and the actual error in the extrapolated solution by T_n where

$$\begin{aligned} P_n &= \frac{2^p}{2^p - 1}(Y_n - Z_n), \\ T_n &= \frac{2^p Z_n - Y_n}{2^p - 1} - Y(x_n). \end{aligned}$$

In the following examples, $v(x) \equiv 1$ unless stated otherwise.

a. The initial value problem

$$\begin{aligned} y' &= -32xy \ln 2, \\ y(-1) &= 2^{-10}, \qquad x \in [-1, 1], \end{aligned} \tag{4.1}$$

has the "peaked" solution

$$y(x) = 2^{6-16x^2}.$$

The Euler, Heun and Runge-Kutta methods were used to compute the numerical solution of (4.1) on the interval $[-1, 1]$. For the Heun method a variable step size introduced by

$$v(x) = \begin{cases} \frac{1}{8} & -1 \leqq x < -\frac{1}{8}, \\ \frac{1}{16} & -\frac{1}{8} \leqq x < \frac{1}{4}, \\ \frac{1}{4} & \frac{1}{4} \leqq x < \frac{1}{2}, \\ \frac{1}{2} & \frac{1}{2} \leqq x < \frac{3}{4}, \\ 1 & \frac{3}{4} \leqq x \leqq 1 \end{cases}$$

TABLE 1
Euler's method, $h_0 = 2^{-10}$

x_n	E_n	P_n	T_n
0.0	-4.238	-4.142	-0.9533×10^{-1}
1.0	-0.1263×10^{-3}	-0.1220×10^{-3}	$-0.4359 \times 10^{-}$

TABLE 2
Heun method, $h_0 = 2^{-8}$

x_n	E_n	P_n	T_n
0.0	-0.6982×10^{-2}	-0.6884×10^{-2}	$-0.7499 \times 10^{-}$
1.0	0.2277×10^{-4}	0.2324×10^{-4}	$-0.5699 \times 10^{-}$

TABLE 3
Runge-Kutta method, $h_0 = 2^{-10}$

x_n	E_n	P_n	T_n
0.0	-0.4274×10^{-6}	-0.4252×10^{-6}	-0.2253×10^{-9}
1.0	0.2035×10^{-12}	0.2103×10^{-12}	-0.6784×10^{-1}

was used. The computational results are presented in Table 1, Table 2 and Tabl

b. The initial value problem

$$y'' + (16\pi e^{-2x} - \tfrac{1}{4})y = 0, \qquad y(0) = 1, \quad y'(0) = \tfrac{1}{2}, \quad x \in [0, 20], \tag{4.2}$$

was integrated over the interval [0, 20] by the Runge-Kutta method. The solutio (4.2) is

$$y(x) = e^{x/2} \cos(4\pi e^{-x}).$$

The numerical results for this example are listed in Table 4.

c. We consider the initial value problem

$$y' = y/x - (1/x)\cos(1/x), \qquad y(-1) = \sin 1, \quad x \in [-1, -2^{-5}], \tag{4.3}$$

which has the highly oscillatory solution

$$y(x) = x \sin(1/x).$$

The Heun method with

$$v(x) = \begin{cases} 1 & -1 \leqq x < -\frac{3}{4}, \\ \frac{1}{2} & -\frac{3}{4} \leqq x < -\frac{1}{2}, \\ \frac{1}{4} & -\frac{1}{2} \leqq x < -\frac{1}{4}, \\ \frac{1}{8} & -\frac{1}{4} \leqq x < -\frac{1}{8}, \\ \frac{1}{16} & -\frac{1}{8} \leqq x < -2^{-4}, \\ \frac{1}{32} & -2^{-4} \leqq x \leqq -2^{-5}, \end{cases}$$

TABLE 1
Euler's method, $h_0 = 2^{-10}$

x_n	E_n	P_n	T_n
0.0	-4.238	-4.142	-0.9533×10^{-1}
1.0	-0.1263×10^{-3}	-0.1220×10^{-3}	-0.4359×10^{-5}

TABLE 2
Heun method, $h_0 = 2^{-8}$

x_n	E_n	P_n	T_n
0.0	-0.6982×10^{-2}	-0.6884×10^{-2}	-0.7499×10^{-5}
1.0	0.2277×10^{-4}	0.2324×10^{-4}	-0.5699×10^{-6}

TABLE 3
Runge-Kutta method, $h_0 = 2^{-10}$

x_n	E_n	P_n	T_n
0.0	-0.4274×10^{-6}	-0.4252×10^{-6}	-0.2253×10^{-9}
1.0	0.2035×10^{-12}	0.2103×10^{-12}	-0.6784×10^{-14}

was used. The computational results are presented in Table 1, Table 2 and Table 3.

b. The initial value problem

$$\begin{gathered} y'' + (16\pi e^{-2x} - \tfrac{1}{4})y = 0, \\ y(0) = 1, \qquad y'(0) = \tfrac{1}{2}, \qquad x \in [0, 20], \end{gathered} \tag{4.2}$$

was integrated over the interval [0, 20] by the Runge-Kutta method. The solution of (4.2) is

$$y(x) = e^{x/2} \cos (4\pi e^{-x}).$$

The numerical results for this example are listed in Table 4.

c. We consider the initial value problem

$$\begin{gathered} y' = y/x - (1/x) \cos (1/x), \\ y(-1) = \sin 1, \qquad x \in [-1, -2^{-5}], \end{gathered} \tag{4.3}$$

which has the highly oscillatory solution

$$y(x) = x \sin (1/x).$$

The Heun method with

$$v(x) = \begin{cases} 1 & -1 \leqq x < -\tfrac{3}{4}, \\ \tfrac{1}{2} & -\tfrac{3}{4} \leqq x < -\tfrac{1}{2}, \\ \tfrac{1}{4} & -\tfrac{1}{2} \leqq x < -\tfrac{1}{4}, \\ \tfrac{1}{8} & -\tfrac{1}{4} \leqq x < -\tfrac{1}{8}, \\ \tfrac{1}{16} & -\tfrac{1}{8} \leqq x < -2^{-4}, \\ \tfrac{1}{32} & -2^{-4} \leqq x \leqq -2^{-5}, \end{cases}$$

TABLE 7
Heun method, $h_0 = 2^{-4}$

x_n	E_n	P_n	A_n
0.750	0.1255×10^{-2}	0.1242×10^{-2}	0.1332×10^{-2}
0.500	0.6663×10^{-2}	0.6565×10^{-2}	0.7220×10^{-2}
0.250	0.4935×10^{-1}	0.4780×10^{-1}	0.5776×10^{-1}
0.125	0.2408	0.2214	0.3466
2^{-4}	0.8030	0.6452	1.848

TABLE 8
Heun method, $h_0 = 2^{-6}$

x_n	E_n	P_n	A_n
0.750	0.8209×10^{-4}	0.8190×10^{-4}	0.8324×10^{-4}
0.500	0.4433×10^{-4}	0.4420×10^{-3}	0.4513×10^{-3}
0.250	0.3505×10^{-2}	0.3486×10^{-2}	0.3610×10^{-2}
0.125	0.2042×10^{-1}	0.2019×10^{-1}	0.2166×10^{-1}
2^{-4}	0.1000	0.9693×10^{-1}	0.1155

TABLE 9
Heun method, $h_0 = 2^{-4}$

x_n	E_n	P_n	A_n
0.750	0.1255×10^{-2}	0.1242×10^{-2}	0.1332×10^{-2}
0.500	0.3828×10^{-2}	0.3789×10^{-2}	0.4053×10^{-2}
0.250	0.1691×10^{-1}	0.1670×10^{-1}	0.1802×10^{-1}
0.125	0.6637×10^{-1}	0.6515×10^{-1}	0.7251×10^{-1}
2^{-4}	0.2426	0.2324	0.2902

solution. The magnitude of the round-off error seems to be comparable with the truncation error for step sized smaller than 2^{-16} in this example.

e. All equations for which numerical results have been given are linear. We consider a nonlinear equation in this case. The solution of the initial value problem

$$y' = 2xe^{-y},$$
$$y(1) = 0, \qquad x \in [2^{-4}, 1], \tag{4.5}$$

is

$$y(x) = 2\cdot\ln x.$$

For the Heun method, Henrici [1, p. 78] derives an expression for the principal error function. A short calculation yields

$$Q(x, y) = -\frac{2}{3x^3}$$

for the initial value problem (4.5). The corresponding magnified error function is

$$E(x) = \frac{2}{3x^2} \int_x^1 \frac{[v(t)]^2}{t} \, dt. \tag{4.6}$$

Let A_n be defined as

$$A_n = h_0^2 \cdot E(x_n)$$

where $E(x_n)$ is defined by (4.6). In Table 7 and Table 8 we compare E_n, P_n and A_n for the initial value problem (4.5) when $v(x) \equiv 1$.

In Table 9 we exhibit the agreement of A_n with E_n and P_n when

$$v(x) = \begin{cases} 1 & \frac{3}{4} \leqq x \leqq 1, \\ \frac{1}{2} & \frac{1}{2} \leqq x < \frac{3}{4}, \\ \frac{1}{4} & \frac{1}{4} \leqq x < \frac{1}{2}, \\ \frac{1}{16} & \frac{1}{8} \leqq x < \frac{1}{4}, \\ \frac{1}{64} & \frac{1}{16} \leqq x < \frac{1}{8}. \end{cases}$$

5. Acknowledgments. The author wishes to thank Professor Robert E. Barnhill for his suggested improvements of this paper. Appreciation is also expressed to the referee of this paper for his calculations and suggested improvements of example "e".

Computer Center
University of Utah
Salt Lake City, Utah

1. P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York, 1962. MR **24** #B1772.
2. F. B. Hildebrand, *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956. MR **17**, 788.
3. L. F. Richardson, "The deferred approach to the limit, I-single lattice," *Trans. Roy. Soc. London*, v. 226, 1927, pp. 299–349.