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

\[
\frac{dY}{dx} = F(x, Y),
\]

\[Y(a) = Y_0, \quad x \in [a, b].\]

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

\[x_0 = a, \quad x_{k+1} = \min (b, x_k + h_0v(x_k)), \quad k = 0, 1, 2, \ldots,\]

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

\[Y_{k+1} = Y_k + h_0v(x_k)I(x_k, Y_k; h_0v(x_k))\]

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

\[E_k = Y_k - Y(x_k).\]
Let \( G(x) = (g_{ij}(x)) \) be a matrix with components
\[
g_{ij}(x) = \frac{\partial f_i(x, Y(x))}{\partial y_j}, \quad i, j = 1, 2, \ldots, n,
\]
where \( f_i \) is the \( i \)th 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
\[
E_k = E(x_k)h_0^p + O(h_0^{p+1})
\]
where \( E(x) \), called the magnified error function, is the solution of the initial value problem
\[
E'(x) = G(x)E(x) + [v(x)]^pQ(x, Y(x)),
E(a) = 0.
\]

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
\[
Y_n - Y(x_n) = E(x_n)h_0^p + O(h_0^{p+1}),
Z_n - Y(x_n) = E(x_n)(h_0/i)^p + O(h_0^{p+1}).
\]
Hence
\[
Y_n - Z_n = (1 - 1/i^p)E(x_n)h_0^p + O(h_0^{p+1}).
\]
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 \( x_n \).
\[
Y(x_n) = \frac{i^pZ_n - Y_n}{i^p - 1} + O(h_0^{p+1}).
\]
The right sides of the relations
\[
E_n = \frac{i^p}{i^p - 1} (Y_n - Z_n)
\]
and
\[
Y(x_n) = \frac{i^pZ_n - Y_n}{i^p - 1}
\]
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 make 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

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

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

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

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

a. The initial value problem

$$y' = -32xy \ln 2,$$
$$y(-1) = 2^{-10}, \quad x \in [-1, 1],$$

has the “peaked” solution

$$y(x) = 2^{p-16x^2}.$$
was used. The computational results are presented in Table 1, Table 2 and Table 3.

b. The initial value problem

\[ y'' + (16 \pi e^{-2x} - \frac{1}{4})y = 0, \]

\[ y(0) = 1, \quad y'(0) = \frac{1}{2}, \quad x \in [0, 20], \]

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

\[ y' = y/x - (1/x) \cos (1/x), \]

\[ y(-1) = \sin 1, \quad x \in [-1, -2^{-5}], \]

which has the highly oscillatory solution

\[ y(x) = x \sin (1/x). \]

The Heun method with

\[ v(x) = \begin{cases} 
1 & -1 \leq x < -\frac{3}{4}, \\
\frac{1}{2} & -\frac{3}{4} \leq x < -\frac{1}{2}, \\
\frac{1}{4} & -\frac{1}{2} \leq x < -\frac{1}{4}, \\
\frac{1}{8} & -\frac{1}{4} \leq x < -\frac{1}{8}, \\
\frac{1}{32} & -\frac{1}{8} \leq x < -2^{-4}, \\
\frac{1}{32} & -2^{-4} \leq x \leq -2^{-5}.
\end{cases} \]
THE USE OF RICHARDSON EXTRAPOLATION IN ONE-STEP METHODS

Table 4

<table>
<thead>
<tr>
<th>$x_n$</th>
<th>$E_n$</th>
<th>$P_n$</th>
<th>$T_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0</td>
<td>$-0.194 \times 10^{-4}$</td>
<td>$-0.201 \times 10^{-4}$</td>
<td>$0.657 \times 10^{-6}$</td>
</tr>
<tr>
<td>20.0</td>
<td>$-0.286 \times 10^{-2}$</td>
<td>$-0.296 \times 10^{-2}$</td>
<td>$0.975 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 5

<table>
<thead>
<tr>
<th>$x_n$</th>
<th>$E_n$</th>
<th>$P_n$</th>
<th>$T_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.750$</td>
<td>$-0.1237 \times 10^{-4}$</td>
<td>$-0.1237 \times 10^{-4}$</td>
<td>$0.8728 \times 10^{-8}$</td>
</tr>
<tr>
<td>$-0.500$</td>
<td>$-0.1954 \times 10^{-4}$</td>
<td>$-0.1955 \times 10^{-4}$</td>
<td>$0.1248 \times 10^{-7}$</td>
</tr>
<tr>
<td>$-0.250$</td>
<td>$0.8301 \times 10^{-5}$</td>
<td>$0.8297 \times 10^{-5}$</td>
<td>$0.3001 \times 10^{-8}$</td>
</tr>
<tr>
<td>$-0.125$</td>
<td>$-0.4321 \times 10^{-4}$</td>
<td>$-0.4323 \times 10^{-4}$</td>
<td>$0.1366 \times 10^{-7}$</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>$-0.8438 \times 10^{-4}$</td>
<td>$-0.8438 \times 10^{-4}$</td>
<td>$0.5011 \times 10^{-8}$</td>
</tr>
</tbody>
</table>

Table 6

<table>
<thead>
<tr>
<th>$x_n$</th>
<th>$E_n$</th>
<th>$P_n$</th>
<th>$T_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6250</td>
<td>$-0.2459 \times 10^{-2}$</td>
<td>$-0.2456 \times 10^{-2}$</td>
<td>$-0.2734 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.8750</td>
<td>$-2.130$</td>
<td>$-2.128$</td>
<td>$-0.2304 \times 10^{-2}$</td>
</tr>
<tr>
<td>1.0</td>
<td>$-30.78$</td>
<td>$-30.74$</td>
<td>$-0.3330 \times 10^{-1}$</td>
</tr>
</tbody>
</table>

was used to integrate (4.3) over the interval $[-1, -2^{-5}]$. Table 5 contains the results for this example.

d. The initial value problem

$$y' = 20y/x,$$

$$(4.4)$$

$y\left(\frac{1}{2}\right) = \frac{1}{2}, \quad x \in [\frac{1}{2}, 1],$$

has the rapidly increasing solution

$$y(x) = 2^{10}x^{20}.$$ 

The Heun method with

$$v(x) = \begin{cases} 
1 & \frac{1}{2} \leq x < \frac{5}{8}, \\
\frac{1}{4} & \frac{5}{8} \leq x < \frac{3}{4}, \\
\frac{1}{8} & \frac{3}{4} \leq x < \frac{7}{8}, \\
\frac{1}{8} & \frac{7}{8} \leq x \leq 1 
\end{cases}$$

was used to obtain the results given in Table 6.

The use of the Runge-Kutta method on (4.4) proved satisfactory for constant step lengths as small as $2^{-15}$. However, for $h_0 = 2^{-16}$ the results obtained from (2.7) and (2.8) were unreliable. The actual and predicted accumulated truncation errors had opposite signs and the extrapolated solution was worse than the nonextrapolated
solution. The magnitude of the round-off error seems to be comparable with the truncation error for step sizes 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

$$y' = 2xe^{-y},$$

$$y(1) = 0, \quad x \in [2^{-4}, 1],$$

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

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

Let \( A_n \) be defined as

\[ A_n = h_n^3 \cdot E(x_n) \]

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 \( v(x) = 1 \).

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

\[ v(x) = \begin{cases} 
1 & \frac{3}{4} \leq x \leq 1, \\
\frac{1}{2} & \frac{1}{2} \leq x < \frac{3}{4}, \\
\frac{1}{4} & \frac{1}{4} \leq x < \frac{1}{2}, \\
\frac{1}{8} & \frac{1}{8} \leq x < \frac{1}{4}, \\
\frac{1}{16} & \frac{1}{16} \leq x < \frac{1}{8}.
\end{cases} \]

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