JOURNAL OF COMPUTATIONAL PHYSICS 9, 528–537 (1972)

Selection of Stepsize in the Variable-Step Predictor–Corrector Method of Van Wyk

ROBERT E. HUDDLESTON

Numerical Applications Division 8321,
Sandia Laboratories, Livermore, California 94550

Received May 5, 1971

Van Wyk has proposed and tested some predictor–corrector methods for ordinary differential equations which will allow a different mesh size for each equation in the system. In this paper we will justify in a more rigorous fashion one aspect of Van Wyk's mesh selection criterion — that of local truncation-error estimation. Van Wyk follows the almost universal practice of numerically justifying his truncation-error estimator by displaying small relative errors in the solution. All the numerical testing in this report will compare the truncation-error estimators against the true local truncation error.

1. INTRODUCTION[1]

We shall be interested in the numerical solution of the initial value problem

$$y'(x) = f(x, y(x)) \tag{1}$$

with initial condition $y(a) = A$. The results developed in this paper are for a generalized form of the Adams–Bashforth–Moulton (ABM) predictor–corrector method of fourth order (globally) [2]. However, it will be clear that similar results can easily be derived for other predictor–corrector methods.

The predictor–corrector formulas for ABM can be derived by the following procedure. From (1), one has

$$\int_{x_n}^{x_{n+1}} y'(x)\, dx = \int_{x_n}^{x_{n+1}} f(x, y(x))\, dx. \tag{2}$$

If we use the Lagrangian form of the interpolation polynomial to fit the data

$$(x_{n-i}, f(x_{n-i}, y(x_{n-i}))), \qquad i = 0, 1, 2, 3$$

[1] I have attempted to rewrite this paper to stay close to Van Wyk's notation. Some superscripting of his variables was, however, necessary.

with $x_{n-i} < x_{n-i+1}$ and then integrate as in (2) we have

$$y(x_{n+1}) = y(x_n) + h_n \sum_{j=0}^{3} b_j f(x_{n-j}, y(x_{n-j})) + \frac{h_n^5}{5!} y^{(5)}(\eta) P_n, \tag{3}$$

where

$$\begin{aligned} b_3 &= \frac{6\alpha_n\beta_n + 4(\alpha_n + \beta_n) + 3}{12\gamma_n(\gamma_n - \alpha_n)(\beta_n - \gamma_n)}, \\ b_2 &= \frac{2 + 3\alpha_n - 6\gamma_n(\gamma_n - \alpha_n) b_3}{6\beta_n(\beta_n - \alpha_n)}, \\ b_1 &= \frac{-(1 + 2\gamma_n b_3 + 2\beta_n b_2)}{2\alpha_n}, \\ b_0 &= 1 - b_3 - b_2 - b_1, \end{aligned} \tag{4}$$

$$\begin{gathered} \alpha_n = h_n^{-1}(x_n - x_{n-1}), \qquad \beta_n = h_n^{-1}(x_n - x_{n-2}), \qquad \gamma_n = h_n^{-1}(x_n - x_{n-3}), \\ h_n = x_{n+1} - x_n, \qquad x_{n-3} \leqslant \eta \leqslant x_n \end{gathered} \tag{5}$$

and

$$P_n = 1 + \frac{5}{4}(\alpha_n + \beta_n + \gamma_n) + \frac{5}{3}(\alpha_n\beta_n + \alpha_n\gamma_n + \beta_n\gamma_n) + \frac{5}{2}\alpha_n\beta_n\gamma_n. \tag{6}$$

Fitting the data $(x_{n-i}, f(x_{n-i}, y(x_{n-i})))$, $i = -1, 0, 1, 2$, and performing the integration of (2) yields the corrector formula

$$y(x_{n+1}) = y(x_n) + h_n \sum_{j=-1}^{2} d_j f(x_{n-j} . y(x_{n-j})) + \frac{h_n^5}{5!} y^{(5)}(\zeta) C_n, \tag{7}$$

where

$$\begin{aligned} d_2 &= \frac{1 + 2\alpha_n}{12\beta_n(1 + \beta_n)(\beta_n - \alpha_n)}, \\ d_1 &= \frac{2\beta_n + 1}{12\alpha_n(1 + \alpha_n)(\beta_n - \alpha_n)}, \\ d_0 &= \tfrac{1}{2} - d_2(1 + \beta_n) - d_1(1 + \alpha_n), \\ d_{-1} &= 1 - d_2 - d_1 - d_0, \qquad x_{n-2} \leqslant \xi \leqslant x_{n+1} \end{aligned} \tag{8}$$

and

$$C_n = -\left(\frac{1}{4} + \frac{5}{6}\alpha_n\beta_n + \frac{5}{12}\alpha_n + \frac{5}{12}\beta_n\right). \tag{9}$$

Formulas (3) and (7) yield the predictor and corrector formulas

$$y^p_{n+1} = y_n + h_n \sum_{j=0}^{3} b_j f_j\,, \qquad y_0 = A \tag{10}$$

and

$$y^c_{n+1} = y_n + h_n \sum_{j=-1}^{2} d_j f_j\,, \tag{11}$$

where

$$f_j = f(x_j\,, y_j).$$

Of course y_{n+1} is just the final value of y^C_{n+1} which we choose to accept in the iteration procedure.

2. Step-Size Control

Let Z_n be the integral curve defined by

$$\begin{aligned} Z_n'(x) &= f(x, Z_n(x)), \\ Z_n(x_n) &= y_n\,. \end{aligned} \tag{12}$$

The local truncation error made by the predictor formula in going from x_n to x_{n+1} is defined to be

$$\tau_n{}^p = Z_n(x_{n+1}) - y^p_{n+1}\,. \tag{13}$$

Similarly for the corrector formula we define

$$\tau_n{}^c = Z_n(x_{n+1}) - y^c_{n+1}\,. \tag{14}$$

It is the quantity $\tau_n{}^c$ that we shall try to estimate and use for step-size control. In this section we shall derive an estimate for $\tau_n{}^c$ which contains Van Wyk's as a special case (though a very accurate and practical special case).

From (3) and (10) one has

$$\begin{aligned} \tau_n{}^p = {} & [Z_n(x_n) - y_n] + h_n \sum_{j=0}^{3} b_j[f(x_{n-j}\,, Z_n(x_{n-j})) - f(x_{n-j}\,, y_{n-j})] \\ & + \frac{1}{5!}\, h_n{}^5 Z_n^{(5)}(\omega_n)\, P_n\,, \qquad \text{where} \quad x_{n-3} \leqslant \omega_n \leqslant x_n\,. \end{aligned}$$

Using a Taylor-series expansion of f in its second variable one has

$$\tau_n{}^p = \frac{1}{5!} h_n{}^5 Z_n^{(5)}(\omega_n) P_n + h_n \sum_{j=1}^{3} b_j f_y(x_{n-j}, \xi_j)[Z_n(x_{n-j}) - y_{n-j}].$$

But the terms $h_n[Z_n(x_{n-j}) - y_{n-j}]$ are of the order $O(h_n\tau_{n-j} + h_n^2 f_y \tau_{n-j})$, which will be small relative to $\tau_n{}^p$. For a proof of the above assertion see [1]. [This assertion is quite often dispensed with in papers by assuming that the numerical formula (10) has the true values for y_j (and thus for f_j) at the back values, i.e., for $j \leqslant n$. This assumption is certainly not justifiable.] Thus we accept the *estimate*

$$\tau_n{}^p = \frac{1}{5!} h_n{}^5 Z_n^{(5)}(\omega_n) P_n . \tag{15}$$

Similarly we arrive at

$$\tau_n{}^c = \frac{1}{5!} h_n{}^5 Z_n^{(5)}(\xi_n) C_n , \qquad \text{where} \quad x_{n-2} \leqslant \xi_n \leqslant x_{n+1} . \tag{16}$$

From (16) it follows that

$$\tau_n{}^c = \left(\frac{h_n{}^5 Z_n^{(5)}(\xi_n) C_n}{h_{n-1}^5 Z_{n-1}^5(\xi_{n-1}) C_{n-1}}\right) \tau_{n-1}^c . \tag{17}$$

From (13) and (14) one has

$$\tau_{n-1}^c = \frac{\tau_{n-1}^c}{\tau_{n-1}^c - \tau_{n-1}^p} (y_n{}^p - y_n{}^c)$$

which, with (15) and (16), implies

$$\tau_{n-1}^c = \frac{Z_{n-1}^{(5)}(\xi_{n-1}) C_{n-1}}{Z_{n-1}^{(5)}(\xi_{n-1}) C_{n-1} - Z_{n-1}^{(5)}(\omega_{n-1}) P_{n-1}} (y_n{}^p - y_n{}^c).$$

Substituting this result into (17) one has

$$\tau_n{}^c = \left(\frac{h_n}{h_{n-1}}\right)^5 \frac{Z_n^{(5)}(\xi_n) C_n}{Z_{n-1}^{(5)}(\xi_{n-1}) C_{n-1} - Z_{n-1}^{(5)}(\omega_{n-1}) P_{n-1}} (y_n{}^p - y_n{}^c).$$

We now make the usual assumption that the fifth-order derivatives are essentially constant over a few steps. (The greatest justification presented for this "usual assumption" is that it seems to be borne out in numerical experimentation with real world problems. As far as comparing derivatives of Z_n with those of Z_{n-1}, it can be shown that $Z_n^{(5)}(x) - Z_{n-1}^{(5)}(x) = O(\tau_n)$; see [1].) Hence, we arrive at the estimate

$$\tau_n^c = \left(\frac{h_n}{h_{n-1}}\right)^5 \frac{C_n}{C_{n-1} - P_{n-1}} (y_n^p - y_n^c), \tag{18}$$

for the local truncation error.

For step-size control let us require that

$$|\tau_n^c / y_n^c| = K \cdot 10^{-B}$$

for each n. (This criterion is of course replaced by an absolute criterion whenever $y_n^c = 0$.) Then, from (18), h_n must be determined such that

$$K \cdot 10^{-B} | y_n^c | = \frac{h_n^5}{h_{n-1}^5} \left| \frac{C_n}{C_{n-1} - P_{n-1}} \right| \cdot | y_n^p - y_n^c |. \tag{19}$$

From the relations for C_n and the definitions of α_n, β_n, and γ_n one arrives at

$$C_n = -\frac{1}{12} [3 + A_1 h_n^{-1} + A_2 h_n^{-2}], \tag{20}$$

where

$$A_1 = 5(2x_n - x_{n-1} - x_{n-2}),$$
$$A_2 = 10(x_n - x_{n-1})(x_n - x_{n-2}).$$

In like manner one arrives at

$$C_{n-1} - P_{n-1} = -(1/12)D, \tag{21}$$

where

$$D = 15 + 20(\alpha_{n-1} + \beta_{n-1}) + 15\gamma_{n-1} + 30\alpha_{n-1}\beta_{n-1} + 20(\alpha_{n-1}\gamma_{n-1} + \beta_{n-1}\gamma_{n-1}) + 30\alpha_{n-1}\beta_{n-1}\gamma_{n-1}.$$

It is important to note that A_1, A_2, and D are dependent only on x_{n-3} through x_n (and not on x_{n+1}) and therefore are known before one needs to calculate h_n.

Substituting (20) and (21) into (19) yields

$$3h_n^5 + A_1 h_n^4 + A_2 h_n^3 + A_3 = 0, \tag{22}$$

where

$$A_3 = -K \cdot 10^{-B} h_{n-1}^5 D \left| \frac{y_n^c}{y_n^p - y_n^c} \right|.$$

This is simply a fifth-degree polynomial in h_n which we must solve at each step for step-size prediction. Applying Newton's algorithm for (22) one has

$$H_{K+1} = H_K - \frac{3H_K^5 + A_1 H_K^4 + A_2 H_K^3 + A_3}{15H_K^4 + 4A_1 H_K^3 + 3A_2 H_K^2}. \tag{23}$$

In practice we iterate N times and take

$$h_n = H_N.$$

In order to start the iterative procedure one needs an initial guess for H_0. If we take $H_0 = h_{n-1}$ then $N = 20$ gives agreement between H_N and H_{N-1} to 5 significant figures. This of course is much too conservative in practice.

Numerical testing indicates, though it has not been proven, that Eq. (22) has one real root and four complex roots for all A_1, A_2, and A_3 arising from differential equations with real coefficients.

If in (19) we accept the further estimate $C_n \cong C_{n-1}$, we arrive at

$$h_n = h_{n-1} \left[K \cdot 10^{-B} \left| \frac{y_n^c}{y_n^p - y_n^c} \right| \cdot \left| \frac{C_{n-1} - P_{n-1}}{C_{n-1}} \right| \right]^{1/5}. \tag{24}$$

Taking the right-hand side of (24) as the initial guess, H_0, for (23) yields agreement of H_N and H_{N-1} to five significant figures for $N = 5$. In fact, with this initial guess, we usually get agreement of H_0 and H_5 to one or two significant figures. This is significant since H_0, as given by the right-hand side of (24), is exactly the estimate which Van Wyk uses to control step size. Hence Van Wyk's estimate agrees very well with the estimate produced by (22) and, in the next section, we give examples in which the estimate produced by (22) is very accurate and reliable.

3. Numerical Results

Most procedures for changing step size depend on an estimate of the local truncation error. However, in testing these procedures, one usually calculates a sample problem with known solution and compares the true solution with the calculated solution. Usually no attempt is made at measuring the difference between the true local truncation error [as defined by Eq. (14)] and the estimate for the local truncation error [which in this paper is given by Eq. (18)]. This is important for two reasons. If the truncation-error estimator is consistently too conservative (i.e., always estimates the error to be larger than the true truncation error) it will produce an accurate solution to a numerically stable problem, but at the cost of too much computation time. Secondly, one may miss entirely the fact that in a numerically unstable problem the estimator may indeed very accurately compute the local truncation error while producing an arbitrarily bad solution. That is, it is possible to have small local truncation errors at every step of a solution while producing an arbitrarily inaccurate solution. This behavior is exhibited in the solution of Eq. (27), the results for which are given in Table III. In this case (see Fig. 1), the integral curves for the differential equation rapidly diverge. At every step truncation errors and rounding errors occur, and even though the errors may be small, the calculation procedure tries to follow a new integral curve which is diverging from the true solution.

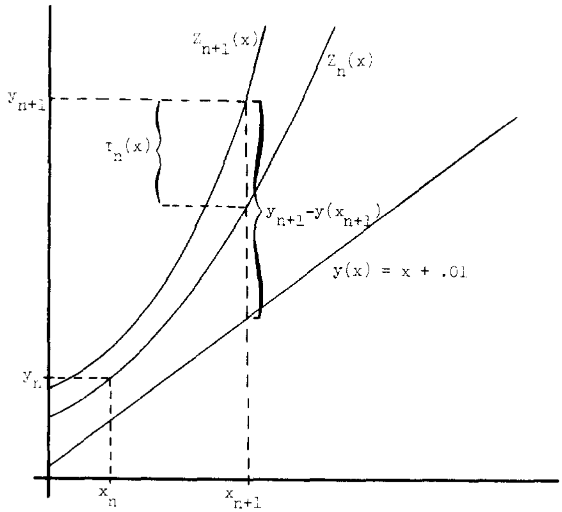

FIGURE 1

Numerical results are given here for the following three initial-value problems:

$$y' = -2xy^2 \qquad \text{Solution: } y(x) = [1 + x^2]^{-1},$$
$$y(0) = 1, \tag{25}$$
Integral curve through (a, b):
$$Z(x) = b[1 + (x^2 - a^2)b]^{-1};$$

$$y' = 5y + 20\pi e^{5x} \cos 20\pi x \qquad \text{Solution: } y(x) = e^{5x} \sin 20\pi x,$$
$$y(0) = 0, \tag{26}$$
Integral curve through (a, b):
$$Z(x) = be^{5(x-a)} + e^{5x}[\sin 20\pi x - \sin 20\pi a].$$

$$y' = 100(y - x) \qquad \text{Solution: } y(x) = x + 1/100,$$
$$y(0) = 1/100, \tag{27}$$
Integral curve through (a, b):
$$Z(x) = x + 1/100 + e^{100(x-a)}(b - a - 1/100).$$

The routine was started with a fourth-order classical Runge–Kutta algorithm using a fixed step size of 0.001. The corrector formula was applied once at each step with two derivative solutions per step. Each time the routine is run a value, T, for the desired local truncation error is preset. The question with which we are concerned is whether the step size predicted by (23) causes the routine to achieve T at each step. Hence, for each step we have also calculated the true local truncation error at each step [as defined by (14)] for comparison. The calculated located local truncation error appearing in the tables is the estimate given by Eq. (18). The "true-solution" column in the tables is the exact (to the number of significant figures listed) solution of the initial-value problem as calculated from the known solution.

TABLE I[a]

x	Calculated solution	True solution	Calculated local truncation error	True local truncation error	Step Size
0.01	0.99990001	0.99990001	1.83 E-15	3.55 E-15	0.002
0.07	0.99512388	0.99512389	6.6 E-9	8.7 E-9	0.032
0.15469	0.97662873	0.97662877	6.0 E-9	6.9 E-9	0.0195
0.41205	0.85485891	0.85485901	4.5 E-9	4.7 E-9	0.0192
0.8938	0.55590377	0.55590376	−4.9 E-9	−5.8 E-9	0.0248
1.508	0.30535907	0.30535900	−4.0 E-9	−4.6 E-9	0.038
2.507	0.13730929	0.13730933	4.7 E-9	6.0 E-9	0.0578
3.585	0.07219506	0.07219512	4.6 E-9	5.6 E-9	0.0785
4.545	0.04617671	0.04617676	1.3 E-9	1.5 E-9	0.08
5.025	0.03809607	0.03809611	6.6 E-10	8.0 E-10	0.08

Note: In the last two steps, the local truncation error fell below the preset value of $T = 5$ E-9 due to the maximum step-size limitation of 0.08.
[a] The notation E-5 represents 10^{-5}.

In Table I we have the information for problem (25) which has a well-behaved solution (i.e., small changes in the data yield only small variations in the solutions). All the calculations in Table I are for $T = 5 \times 10^{-9}$. An upper bound on the step size was set at 8×10^{-2} ($h^4 = 0.5 \times 10^{-4}$), and the step size was not allowed to more than double at each step. The same problem with $T = 10^{-6}$ was run without an upper limit on the step size with a resulting step size of 7.8 being used for the step following $x = 5.597$ with two significant figures of accuracy remaining in the computed solution and local truncation error of 1.89×10^{-7}.

The initial-value problem (26) has a rapidly oscillating solution which grows exponentially like e^{5x} in absolute value. Calculations are given between $x = 0$ and $x = 1$ which covers ten periods of the sine wave oscillation. The values given in Table II are for $T = 5 \times 10^{-7}$.

TABLE II

x	Step Size	Calculated solution	True solution	Calculated local truncation error	True local truncation error	Slope of the true solution
0.0096	0.0018	5.9544 E-1	5.9544 E-1	−5.1 E-7	−5.1 E-7	5.7 E + 1
0.0891	0.0036	−9.9037 E-1	−9.9036 E-1	−1.9 E-6	−3.6 E-6	7.1 E + 1
0.208	0.0015	1.3864	1.3865	−4.2 E-7	−4.2 E-7	1.6 E + 2
0.347	0.0014	7.8233 E-1	7.8248 E-1	7.4 E-7	7.9 E-7	−3.5 E + 2
0.695	0.0016	−8.4169	−8.4157	−8.5 E-7	−9.6 E-7	1.9 E + 3
0.850	0.0007	−2.0961	−2.0931	2.5 E-7	2.5 E-7	−4.4 E + 3
0.978	0.0021	−1.3020 E + 2	−1.3019 E + 2	4.1 E-7	3.3 E-7	1.1 E + 3

TABLE III

x	Step size	Calculated solution	True solution	Calculated local truncation error	True local truncation error
0.103	1.0 E-3	1.13000 E-1	1.13000 E-1	4.1 E-14	−1.8 E-13
0.213	1.0 E-3	2.22999 E-1	2.30000 E-1	2.2 E-9	−9.3 E-9
0.302	1.4 E-3	3.07089 E-1	3.12479 E-1	5.8 E-9	2.4 E-9
0.400	4.6 E-4	−9.48037 E + 1	4.10273 E-1	5.0 E-9	4.1 E-9
0.500	4.6 E-4	−2.06936 E + 6	5.10189 E-1	5.0 E-9	4.1 E-9
0.600	4.6 E-4	−4.5016 E + 10	6.10014 E-1	5.0 E-9	4.1 E-9
0.900	4.6 E-4	−4.8533 E + 23	0.10106 E-1	5.0 E-9	4.1 E-9

Small perturbations of the data in problem (27) cause large changes in the solution and we should, therefore, expect a scheme which progresses step by step to give very poor results. This behavior is certainly illustrated in Table III. A maximum step-size was set at 10^{-2} and T was preset to 5×10^{-9}. The iterative procedure produces step sizes to achieve local truncation errors of T despite the fact that the calculated solution diverges rapidly from the true solution. Examples like (27) and Table III illustrate the care which must be exercised when using any method of step-size control based on local truncation-error estimation.

4. Conclusions

The subject of this paper has dealt with only one aspect of Van Wyk's paper [2], that of truncation-error estimation as it is used for step-size control. I believe that we have established the accuracy of Van Wyk's estimates while pointing out that it is local truncation error which is being estimated and not the true error in the solution.

References

1. R. E. Huddleston, Variable-step truncation error estimates for Runge–Kutta methods of order 4 or less, *J. Math. Anal. Appl.*, to appear.
2. R. Van Wyk, Variable mesh multistep methods for ordinary differential equations, *J. Comput. Phys.* **5** (1970), 244–264.