Applied Numerical Mathematics 5 (1989) 145–149
North-Holland

THE SECOND-ORDER BACKWARD DIFFERENTIATION FORMULA IS UNCONDITIONALLY ZERO-STABLE *

Robert D. SKEEL
Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

Previous studies of the stability of the second-order backward differentiation formula have concluded that stability is possible only if restrictions are placed on the stepsize ratios, for example, limiting the ratio to some value less than $1+\sqrt{2}$. However, actual implementations of the BDFs differ from the usual theoretical models of such methods; in particular, practical codes use scaled derivatives (e.g. EPISODE) or backward differences to represent current information about the solution. The representation makes no difference to truncation errors, but it has an important effect of the propagation of roundoff errors and of errors in the solution of the implicit equations. In this paper it is shown that the divided difference implementation of the variable coefficient (variable stepsize extension of the) second-order BDF is zero-stable for unrestricted stepsize ratios.

1. Statement of result

The second-order backward differentiation formula (BDF) is of great practical importance due to its simplicity, its efficiency, and its excellent stability properties for stiff ODEs and PDEs. Efficiency, for stiff problems especially, requires the use of variable stepsize. There are two well-known ways of extending the BDFs to variable stepsize that have been used in ODE software. The natural *variable coefficient* extension used in EPISODE is known to have superior stability properties to the interpolatory Gear–Nordsieck *fixed coefficient* extension used in LSODE. This paper shows that the variable coefficient BDF of order two is, in fact, 0-stable regardless of how the stepsize is varied. In Section 3 it is noted that the fixed coefficient BDF is only conditionally 0-stable.

Let $y(x)$, $a \leqslant x \leqslant b$, be the solution to

$$y'(x)=f(x,\ y(x))$$

which satisfies

$$y(a)=\eta,$$

where it is assumed that $f(x,\ y)$ is Lipschitz continuous in y with Lipschitz constant L. For simplicity we assume that the numerical solution is computed by one step of backward Euler with infinitesimal stepsize followed by $N-1$ steps of second-order BDF. (This is equivalent to one step of trapezoid followed by N-steps of second-order BDF. However, the former point of view is notationally more convenient and slightly closer to what is done in practice.) Hence the mesh is

$$a=x_{-1}=x_0<x_1<x_2< \cdots <x_N=b$$

* This research was partially supported by the U.S. Department of Energy under grant DEAC0276ERO2383.

0168-9274/89/$3.50 © 1989, Elsevier Science Publishers B.V. (North-Holland)

with stepsizes

$$h_n := x_n - x_{n-1}.$$

The variable coefficient extension of the second-order BDF is given by

$$y_n - (1+\rho_n h_n/h_{n-1})y_{n-1} + (\rho_n h_n/h_{n-1})y_{n-2} = (1-\rho_n)h_n f(x_n, y_n) \tag{1}$$

for $n \geqslant 2$ where

$$\rho_n := h_n/(h_{n-1}+2h_n), \quad n \geqslant 1,$$

and $\rho_0 := 0$. Previous studies of 0-stability (Gear and Tu [5], Gear [4], Zlatev [7]) have considered the effect of small perturbations to the above equation. As a consequence, 0-stability is possible only if restrictions are placed on the stepsize ratio, for example, the restriction

$$h_n/h_{n-1} \leqslant \text{constant} < 1+\sqrt{2},$$

given by Crouzeix and Lisbona [2], is sufficient to ensure 0-stability.

Most practical implementations use, instead of the second-order difference equation (1), a pair of first-order difference equations; for example, EPISODE uses scaled derivatives and others use backward differences. For the purpose of discussion let us assume the use of divided differences. Letting z_n be the computed value of $(y_n - y_{n-1})/h_n$, we have

$$y_0 = \eta + \delta_0, \qquad z_0 = f(x_0, y_0) + \delta_0',$$

and for $n = 1, 2, \ldots, N$

$$y_n = y_{n-1} + h_n z_n + \delta_n, \qquad z_n = \rho_n z_{n-1} + (1-\rho_n) f(x_n, y_n) + (1-\rho_n)\delta_n'$$

where δ_n, δ_n' are perturbations due to roundoff errors and the solution of the implicit equations. The factor $1-\rho_n$ in front of δ_n' is for convenience; and because

$$0 \leqslant \rho_n \leqslant \tfrac{1}{2},$$

it should cause no concern. It is expected that both δ_n and δ_n' are $O(u)$ where u is the unit roundoff error. It is *not* to be expected that δ_n' is $O(u/h_n)$, and *that is what makes real methods superior to theoretical models of methods.*

The result that we prove compares the computed solution y_n to the analytical solution $y(x_n)$, which satisfies

$$y[x_n, x_{n-1}] = \rho_n y[x_{n-1}, x_{n-2}] + (1-\rho_n) f(x_n, y(x_n)) - (1-\rho_n)\tau_n$$

where the truncation error (per unit step)

$$\tau_n = h_n(h_n + h_{n-1}) y[x_n, x_n, x_{n-1}, x_{n-2}], \quad n \geqslant 1,$$

and $\tau_0 = 0$. This is bounded by $\frac{1}{3}h_{\max}^2$ times a bound on the norm of the third derivative of $y(x)$. In Section 2 it is proved that

$$|y_n - y(x_n)| \leqslant \omega(D_n + (x_n - a)D_n')\, e^{\omega L(x_n - a)}, \tag{2}$$

where

$$D_n := \max_{0 \leqslant j \leqslant n} \left\| \sum_{i=0}^{j} \delta_i \right\|, \qquad D_n' := \max_{0 \leqslant j \leqslant n} \| \tau_j + \delta_j' \|,$$

and

$$\omega = \left(1 - \tfrac{2}{3} h_{\max} L\right)^{-1}.$$

Note that the factor ω is essentially 1, and thus the result is quantitatively very satisfactory.

After the proof in Section 2 we conclude in Section 3 with a discussion of possible extensions of this result.

2. Proof of result

If we define

$$d_n := y_n - y(x_n), \qquad e_n := z_n - y[x_n, x_{n-1}],$$

and

$$g_n := f(x_n, y_n) - f(x_n, y(x_n)),$$

then after subtracting the appropriate equations we get

$$d_0 = \delta_0, \qquad e_0 = g_0 + \varepsilon_0,$$

and

$$d_n = d_{n-1} + h_n e_n + \delta_n, \tag{3}$$

$$e_n = \rho_n e_{n-1} + (1 - \rho_n) g_n + (1 - \rho_n) \varepsilon_n \tag{4}$$

where

$$\varepsilon_n := \delta'_n + \tau_n.$$

Solving (4) for e_n we get

$$e_n = \sum_{j=0}^{n} (1 - \rho_j) \rho_{j+1} \cdots \rho_n (g_j + \varepsilon_j).$$

Putting this into (3) and solving for d_n gives a double summation for $g_j + \varepsilon_j$ plus a single summation for δ_j. Introducing a term for $h_0 = 0$ into the double summation and interchanging the order of summation gives

$$d_n = \sum_{j=0}^{n} h_{jn} (g_j + \varepsilon_j) + \sum_{j=0}^{n} \delta_j$$

where

$$h_{jn} := \sum_{i=j}^{n} (1 - \rho_j) \rho_{j+1} \cdots \rho_i h_i.$$

Using $\| g_j \| \leqslant L \| d_j \|$, we get

$$\| d_n \| \leqslant \sum_{j=0}^{n-1} h_{jn} L \| d_j \| + (1 - \rho_n) h_n L \| d_n \| + \sum_{j=0}^{n} h_{jn} D'_n + D_n.$$

By interchanging the order of summation and using $\rho_0 = 0$, it can be shown that

$$\sum_{j=0}^{n} h_{jn} = x_n - a. \tag{5}$$

By separately considering $h_n \leqslant h_{n-1}$ and $h_n \geqslant h_{n-1}$ it can be argued that

$$(1 - \rho_n) h_n \leqslant \tfrac{2}{3} h_{\max}.$$

Putting this together we get

$$\| d_n \| \leqslant \sum_{j=0}^{n-1} h_{jn} \omega L \| d_j \| + M_n$$

where

$$M_n = \omega (D_n + (x_n - a) D_n').$$

By induction on n we can show

$$\| d_n \| \leqslant (1 + h_{0n} \omega L) \cdots (1 + h_{n-1,n} \omega L) M_n.$$

(One must use the fact that h_{jn} and M_n are nondecreasing functions of n.) Then we have

$$\| d_n \| \leqslant M_n \exp(h_{0n} \omega L) \cdots \exp(h_{n-1,n} \omega L) \leqslant M_n \exp\left(\sum_{j=0}^{n} h_{jn} \omega L \right)$$

and (2) follows from a second application of the identity (5).

3. Extensions of the result

The fixed coefficient extension of the second order BDF is given by Skeel [6] to be

$$z_n = \tfrac{1}{3}(h_n/h_{n-1}) z_{n-1} + \tfrac{2}{3} f(x_n, y_n) + \tfrac{1}{3}((h_{n-1} - h_n)/h_{n-1}) f(x_{n-1}, y_{n-1}),$$

where again z_n denotes an approximation to the divided difference of the solution. Clearly, an unlimited number of consecutive steps with stepsize ratios exceeding 3 would lead to instability. At the same time it would seem that keeping the stepsize ratios limited to some value less than 3 would suffice for stability. This is much more generous than the restriction to a value less than $\sqrt{3}$ obtained for the ordinate form of the method, and it helps to explain the success of codes such as LSODE.

For the higher order variable coefficient BDFs we cannot expect unconditional 0-stability, but we can expect that conditions cited in the literature can be relaxed. For example, if we consider a long sequence of steps with a fixed stepsize ratio, then a divided difference implementation of the third order BDF is 0-stable provided that $h_n/h_{n-1} < 3.44\ldots$. This is far less restrictive than the value $1.6180\ldots$ quoted by Dahlquist [3] for theoretical implementations.

For the test problem $y' = \lambda y$ it is observed by Brayton and Conley [1] that the principal root of the characteristic polynomial exceeds 1 in modulus for some $h_{n-1}\lambda$ in the left half-plane unless $h_n = h_{n-1}$, and hence some form of variable-stepsize A-stability does not seem possible except under severe restrictions. This situation is unchanged for the divided difference implementation because the principal root of the characteristic polynomial is unaffected by the implementation. Because the roots of the characteristic polynomial vary from one step to the next, one cannot draw precise conclusions about stability from a study of these roots. Nonetheless, it is interesting to note that for the divided difference implementation these roots are of modulus at most one as long as

$$|\operatorname{Arg}(-\lambda)| \leqslant \arctan\left(\tfrac{5}{3}\sqrt{15}\right) \doteq 81°$$

A rigorous analysis is needed to confirm that the second-order BDF is unconditionally A(α)-stable for some α close to 81°. Note that for stiff problems it may no longer be valid to assume that $\delta_n' = O(u)$, but a more appropriate scaling of the perturbation may be indicated.

References

[1] R.K. Brayton and C.C. Conley, Some results on the stability and instability of the backward differentiation methods with non-uniform time steps, in: *Topics in Numerical Analysis* (Academic Press, New York, 1972) 13–33.

[2] M. Crouzeix and F.J. Lisbona, The convergence of variable-stepsize, variable-formula, multistep methods, *SIAM J. Numer. Anal.* *21* (1984) 512–534.

[3] G. Dahlquist, Some comments on stability and error analysis for stiff nonlinear differential systems, in: R.C. Aiken, ed., *Stiff Computation* (Oxford University Press, New York, 1985) 265–294.

[4] C.W. Gear, Stability of variable step methods for ordinary differential equations, REP. UIUCDCS-R-78-938, Department Computer Science, University of Illinois, Urbana, IL (1978).

[5] C.W. Gear and K.W. Tu, The effect of variable mesh size on the stability of multistep methods, *SIAM J. Numer. Anal.* *11* (1974) 1025–1043.

[6] R.D. Skeel, Construction of variable stepsize multistep formulas, *Math. Comp.* *47* (1986) 503–510, S45–S52.

[7] Z. Zlatev, Zero-stability properties of the three-ordinate variable stepsize variable formula methods, *Numer. Math.* *37* (1981) 157–166