General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

(NASA-CR-155169) ERROR BEHAVIOR OF MULTISTEP METHODS APPLIED TO UNSTABLE DIFFERENTIAL SYSTEMS (Virginia Univ.) 16 p HC A02/MF A01 CSCL 12B

N77-33941

Unclas

G3/66 49496

ERROR BEHAVIOR OF MULTISTEP METHODS APPLIED TO UNSTABLE DIFFERENTIAL SYSTEMS

R. Leonard Brown

Dept. of Applied Mathematics and Computer Science
University of Virginia
Charlottesville, Virginia 22903 USA

Synopsis:

The problem of modeling a dynamic system described by a system of ordinary differential equations which has unstable components for limited periods of time is discussed. It is shown that the global error in a multistep numerical method is the solution to a difference equation initial value problem, and the approximate solution is given for several popular multistep integration formulas. Inspection of the solution leads to the formulation of four criteria for integrators appropriate to unstable problems. A sample problem is solved numerically using three popular formulas and two different stepsizes to illustrate the appropriateness of the criteria.

OCT 1977
RECEIVED
NASA STI FACILITY

1. Introduction

When a dynamic physical system is modeled by a system of ordinary differential equations

$$y' = f(y,t) \tag{1a}$$

$$y(t_0) = y_0 \tag{1b}$$

it is possible that the system will be physically unstable for some interval of the independent variable t; an example is an aircraft approaching a spin configuration. This situation is usually described by linearizing the system and saying that it is mathematically unstable at t_n if the Jacobian matrix $\left.\frac{\partial f}{\partial y}\right|_{y(t_n)} = f_y$ has eigenvalues with positive real parts. The usual analysis [1,2] of a numerical method for computing an approximate solution to the system (1) describes the behavior of the numerical solution throughout the stable region in which f_y has all negative eigenvalues. A fundamental result is that no multistep formula

$$\sum_{i=0}^{k} \alpha_i y_{n-i} + h\sum_{i=0}^{k} \beta_i f(y_{n-i}, t_{n-i}) = 0 \tag{2}$$

for $t_j = t_0 + jh$, y_j the approximate solution at t_j, can be stable wherever the system (1) is stable and have an error order r greater than 2. The error order of a formula is r if the one step truncation error is $O(h^{r+1})$, as described in standard texts [2].

This paper addresses the question of what characteristics a multistep formula (2) should have in order to behave well in t-intervals

when the model equations (1) are unstable. Section 2 gives three examples that illustrate the analysis involved in studying typical multistep methods. Section 3 states criteria which should be applied when choosing a formula to simulate an occasionally unstable system and gives an example of an inappropriate choice. Section 4 reports a numerical example with three of the formulas analyzed in the paper.

2. Example Formulas

The Milne corrector formula

$$y_n = y_{n-2} + \frac{h}{3}\left(f(y_n,t_n) + 4f(y_{n-1},t_{n-1}) + f(y_{n-2},t_{n-2})\right) \quad (3)$$

has been analyzed[2]. If the formula is iterated to convergence with any appropriate predictor, the stability depends only on the corrector. Applying (3) to the linear test equation

$$y' = \lambda y \quad (4)$$

with two different numerical solution sequences z_n, y_n, yields an error $e_n = z_n - y_n$ that looks like

$$e_n = a\xi_1^n + b\xi_2^n$$

where the ξ_i are the roots of the characteristic equation

$$\rho(\xi) + h\lambda\sigma(\xi) = (1-\frac{1}{3}h\lambda)\xi^2 - \frac{4}{3}h\lambda\xi - (1+\frac{1}{3}h\lambda) = 0. \quad (5)$$

The constants a and b depend on z_0, y_0, and their first two differences. In the Milne case, $\xi_1 = e^{h\lambda} + \frac{1}{180}(h\lambda)^5 + O(h^6)$, and $\xi_2 = -e^{-h\lambda/3} + O(h^3)$. Thus, when $\lambda = f_y$ has a negative real part, ξ_2^n is growing while ξ_1 is decreasing and the leading constant term of ξ_2 is the same size as that of ξ_1, so when the physical system is stable, the numerical system

is in danger of becoming unstable if $h\lambda$ is large enough and $b \neq 0$. It is easily seen that the leading constant terms 1, -1, are the roots of $\rho(z) = z^2 - 1 = 0$. When λ has a positive real part, ξ_1 is growing at the same rate as the correct solution, ξ_2 is being highly damped, and the relative error will be almost constant. Thus, the Milne method behaves well in unstable regions but is potentially unreliable in stable regions of the exact solution.

A similar analysis can be made of two Adams-Bashforth explicit (predictor only) forumulas:

$$\text{AB2:} \quad y_n = y_{n-1} + \frac{h}{2}\left(3f(y_{n-1},t_{n-1}) - f(y_{n-2},t_{n-2})\right) \qquad (6)$$

$$\text{AB3:} \quad y_n = y_{n-1} + \frac{h}{12}\left(23f(y_{n-1},t_{n-1}) - 16f(y_{n-2},t_{n-2}) + 5f(y_{n-3},t_{n-3})\right) \qquad (7)$$

The error in AB2 is $e_n = a\xi_1^n + b\xi_2^n$, where $e_n = e_{n-1} + \lambda h(3e_{n-1} - e_{n-2})/2$ which has the characteristic equation $\xi^2 - (1 + \frac{3}{2}h\lambda)\xi + \frac{1}{2}h\lambda = 0$. By assuming distinct roots that are polynomials in $h\lambda$ and equating the undetermined coefficients to $(\xi-\xi_1)(\xi-\xi_2)$, the following system of equation must be satisfied. If $\xi_1 = \sum_{i=0}^{\infty} a_i (h\lambda)^i$, and $\xi_2 = \sum_{i=0}^{\infty} b_i (h\lambda)^i$, then $a_o + b_o = 1$, $a_o b_o = 0$, $a_1 + b_1 = 3/2$, $a_2 + b_2 = 0$, $a_1 b_o + a_o b_1 = 1/2$, $a_o b_2 + a_2 b_o + a_1 b_1 = 0$, $a_3 + b_3 = 0$, $a_o b_3 + b_o a_3 + a_1 b_2 + a_2 b_1 = 0$, which is solved by $\xi_1 = 1 + h\lambda + (h\lambda)^2/2 - (h\lambda)^3/4 + O(h^4)$, $\xi_2 = \frac{1}{2}(h\lambda - (h\lambda)^2 + (h\lambda)^3/2) + O(h^4)$. These can be rewritten as $\xi_1 = e^{h\lambda} - (h\lambda)^3/12 + O(h^4)$, $\xi_2 = (e^{h\lambda/2} - 1) - 3(h\lambda)^2/8 + O(h^3)$. So when $\mathrm{Re}(\lambda) < 0$, both error components are decreasing, and ξ_2^n is oscillating. When $\mathrm{Re}(\lambda) > 0$, the principal error ξ_1^n due to the root[1] of $\rho(\xi) = 0$ is following the true solution, maintaining the same relative accuracy. The parasitic error ξ_2^n due to the non-unity root of

$\rho(\xi) = 0$ is increasing at a much slower rate due to a zero constant term and the exponent $h\lambda/2$ being smaller than $h\lambda$. Thus, relative error should remain nearly constant in the unstable region of the physical system. The relative error will also be good in a restricted region of the negative complex half plane called the stability region. This consists of all complex $h\lambda$ such that the numerical method produces a decreasing solution. If the whole negative half plane is in the stability region, the formula is called A-stable.[3] Sample stability regions for AB2 and AB3 appear in Figure 1.

The same analysis for AB3 for three separate roots $\prod_{i=1}^{3} (\xi-\xi_i)$ and for a repeated parasitic root in $(\xi-\xi_1)(\xi-\xi_2)(\xi-n\xi_3)$ yield contradictions in the equations for the undetermined coefficients. Instead of a full analysis, then, the worst case can be assumed, that of a multiple root corresponding to the multiple root of $\rho(z) = z^2(z-1)$ which exists for $h\lambda = 0$. This means the error looks like $e_n = a\xi_1^n + (b+cn)\,\xi_2^n$, with a linearly growing parasitic error term. However, the leading terms of ξ_2 can be estimated and are found to be $\xi_2 = \frac{11}{12} h\lambda + O(h\lambda^2) = (e^{\frac{11}{12}h\lambda} - 1) + O(h^2)$. Thus, even in the worst case of large coefficients a,b,c due to large initial errors, the parasitic error is growing at most linearly while the solution is growing exponentially, so the largest contribution to the error will be the principal error term.

3. Selection Criteria

Observation of the three formulas above leads to the following

criteria for methods that do well both in the stable and unstable regions of the physical model.

1. All roots of $\rho(z) = 0$ should be very small, and preferably zero, except the principal root, which is always 1. In general, a nonzero root z_i leads to a root of the characteristic equation $\xi_i = z_i e^{h\lambda\kappa}$; a zero z_i yields $\xi_i = e^{h\lambda\kappa} - 1$, for some real κ.

2. The coefficients κ should be positive, so that the parasitic error is decreasing for $Re(\lambda)<0$, and unless the corresponding $z_i = 0$, $|\kappa| < 1$ is also necessary so the relative error doesn't grow faster than the principal error term for $Re(\lambda) > 0$.

3. Multiple roots should be avoided, since the error term will include factors of n. An exception occurs for $z_i = 0$, since only linear error growth occurs in that case.

The above three criteria all attempt to keep the error in following an exponentially increasing solution from growing faster than the solution grows. The following example shows the other extreme. The 2-step Implicit Backward Differentiation Formula (BDF2) is

$$y_n = \frac{4}{3} y_{n-1} - \frac{1}{3} y_{n-2} + \frac{2}{3} h\ f\ (y_n, t_n) \qquad (7)$$

and $\rho(z) = 0$ has the roots 1, $\frac{1}{3}$. The error terms look like $\xi_1 = e^{h\lambda} + O(h^3)$, $\xi_2 = \frac{1}{3} e^{-h\lambda/3} + \frac{2}{27}(h\lambda)^2 + O(h)^3$, so this method meets criteria 1 and 3, and the small factor 1/3 in ξ_2 lessens the problem with the sign of κ. However, as can be seen in Fig. 2, the stability region is such that the numerical solution, as well as the error, is only increasing for a very small portion of the positive

half plane. So using BDF2 on an unstable system will likely generate a completely spurious decreasing numerical solution, except for small stepsize h, when the error is increasing fast. A fourth criterion has been suggested[4] and will be added here.

4. The stability region of the formula must not make unstable physical solutions numerically stable.

4. Numerical Examples

The three methods AB2, AB3, and BDF2 were applied to the unstable complex problem

$$y' = (1 + j/2)y, \tag{8a}$$

$$y(0) = 1, \tag{8b}$$

where $j = (-1)^{\frac{1}{2}}$, with two fixed positive stepsizes h and both exact and inexact back starting values. The AB2 and BDF2 methods are second order methods and thus have comparable errors that are larger than those for the third order AB3 method, given the same stepsize. The implicit BDF2 method used an Euler predictor and 10 corrector iterations. The starting values given at n=0 were $y_{n-i} = \exp(-i\lambda h)$ for the exact values and $y_{n-i} = \exp(-i\lambda h) + .01i$ for inexact starting values, where $\lambda = (1 + j/2)$. All the elements in this example are exaggerated so that the results may be easily seen to agree with the predictions of the previous sections. The more realistic case of the eigenvalues of the Jacobian being close to the imaginary axis, as in a spinning aircraft, would require a great deal more computation yet

would reach the same conclusions.

The predicted relative errors can be expressed as $(a_1 e^{nh\lambda} + b_1(e^{h\lambda/2}-1)^n)e^{-nh\lambda}$, $(a_2 e^{nh\lambda} + (b_2 + c_2 n)(e^{11h\lambda/12}-1)^n)e^{-nh\lambda}$, $(a_3 e^{nh\lambda} + b_3(\frac{1}{3})^n e^{-nh\lambda/3})e^{-nh\lambda}$, for AB2, AB3, and BDF2 respectively.

When $h < .25$ is chosen, the numerical solution is outside the stability region of all three methods. The logarithm of the principal relative error is $\ln(a_i e^{nh\lambda}/e^{nh\lambda}) = \ln a_i$, a constant in all cases. If the logarithm of the relative error is not a constant when plotted against the independent variable $t = nh$, then the parasitic errors must account for most of this. For the Adams methods and $c_1 = 0$, these are $\ln(b_i + c_i n)(e^{\kappa h\lambda}-1)^n/e^{nh\lambda}) = \ln(b_i + c_i n) + n\ln(e^{\kappa h\lambda}-1) - nh\lambda$, and for BDF2 it is $\ln b_3 + n\ln(\frac{1}{3}) - \frac{4}{3}nh\lambda$.

Figure 3 shows the logarithm of the relative error for all three methods with exact starting values and $h = .05$. Note that the BDF2 error stays less than the AB2 error throughout, indicating only small contributions from the parasitic error terms. Figure 4 shows the same results for inexact starting values with a corresponding increase in b_2, b_3 resulting in a large initial error in BDF2. As n increases, the term $b_3(\frac{1}{3})^n e^{-nh\lambda/3}$ is damped until at $t = 4(n=80)$, this term is again negligible compared to the $b_1(e^{h\lambda/2}-1)^n$ term. Comparison of Figure 3 and Figure 4 indicates the $c_2 n(e^{11h\lambda/12}-1)^n$ term does not significantly affect the error in the AB3 method, except for the initial perturbation, although the principal error term a_2 is much larger.

Figure 5 shows the logarithm of the relative error when $h = .5$, which is still outside the stability region of the Adams methods but

inside that of BDF2. The error in BDF2 is growing faster than that of the other methods since it is due to a decreasing numerical solution modelling an increasing analytic solution. Figure 5 comes from exact starting conditions but similar results arise from inexact starting values.

5. Conclusions

When simulating a problem which is unstable over certain periods of the independent variable, one's choice of method affects the relative accuracy of the numerical solution. When an explicit method is used, as is required in real time simulation and is often the case in simple simulation programs, a multistep method of the Adams-Bashforth type will accurately follow the exact unstable solution despite large initial errors. The AB2 method has less computational pitfalls than higher order formulas due to the lack of multiple roots, but computational experience shows the lack of a leading constant term in these roots causes all such methods to be very accurate. Multistep methods are preferable to multistage methods, e.g. Runge-Kutta, since a k-step method is as accurate with only one function evaluation as a k-stage method requiring k evaluations of $f(y,t)$.

If an implicit corrector is to be employed, formulas with a simple characteristic polynomial $\rho(z) = (1-z)z^{k-1}$, i.e. Adams-Moulton methods, should behave better than even the (possibly) A-stable BDF methods because they have only one non-zero root and produce increasing numerical sequences wherever the solution is increasing, at least for

methods of less than five steps. The best implicit method from a stability/instability standpoint is the trapezoidal formula

$$y_n = y_{n-1} + \frac{h}{2}(f(y_n,t_n) + f(y_{n-1},t_{n-1}))$$

which has no parasitic roots and is increasing (decreasing) only where the exact solution is unstable (stable). However, this method is only second order. All of the implicit methods take on their optimal stability properties only when iterated to convergence, which usually requires at least two evaluations of $f(y,t)$. Therefore, from the standpoint of efficiency and stability, AB2 seems to be a practical integration method for occasionally unstable systems of ODE's.

6. Acknowledgements.

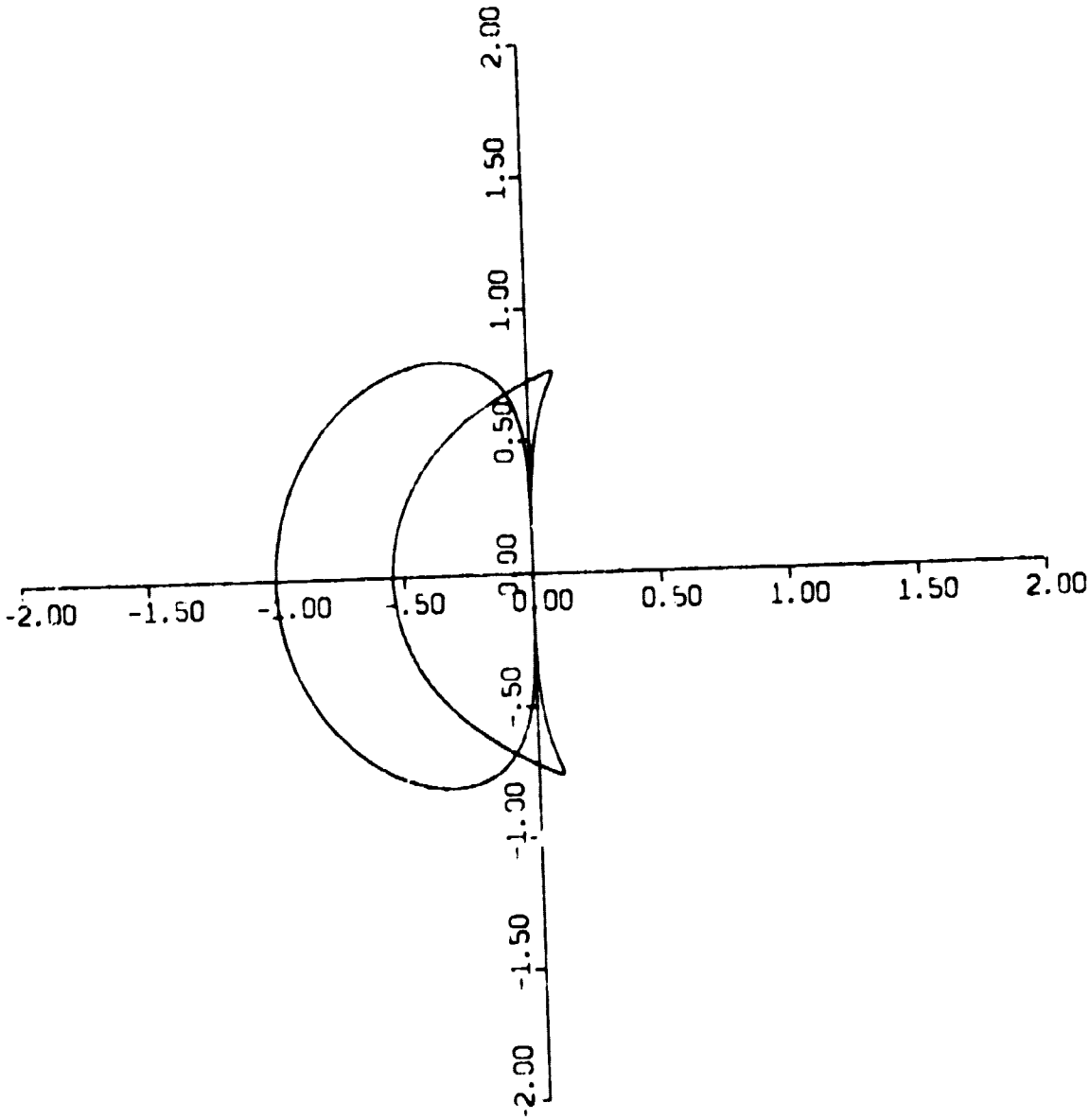
This work was supported in part by the National Aeronautics and Space Administration through its grant NSG 1335.

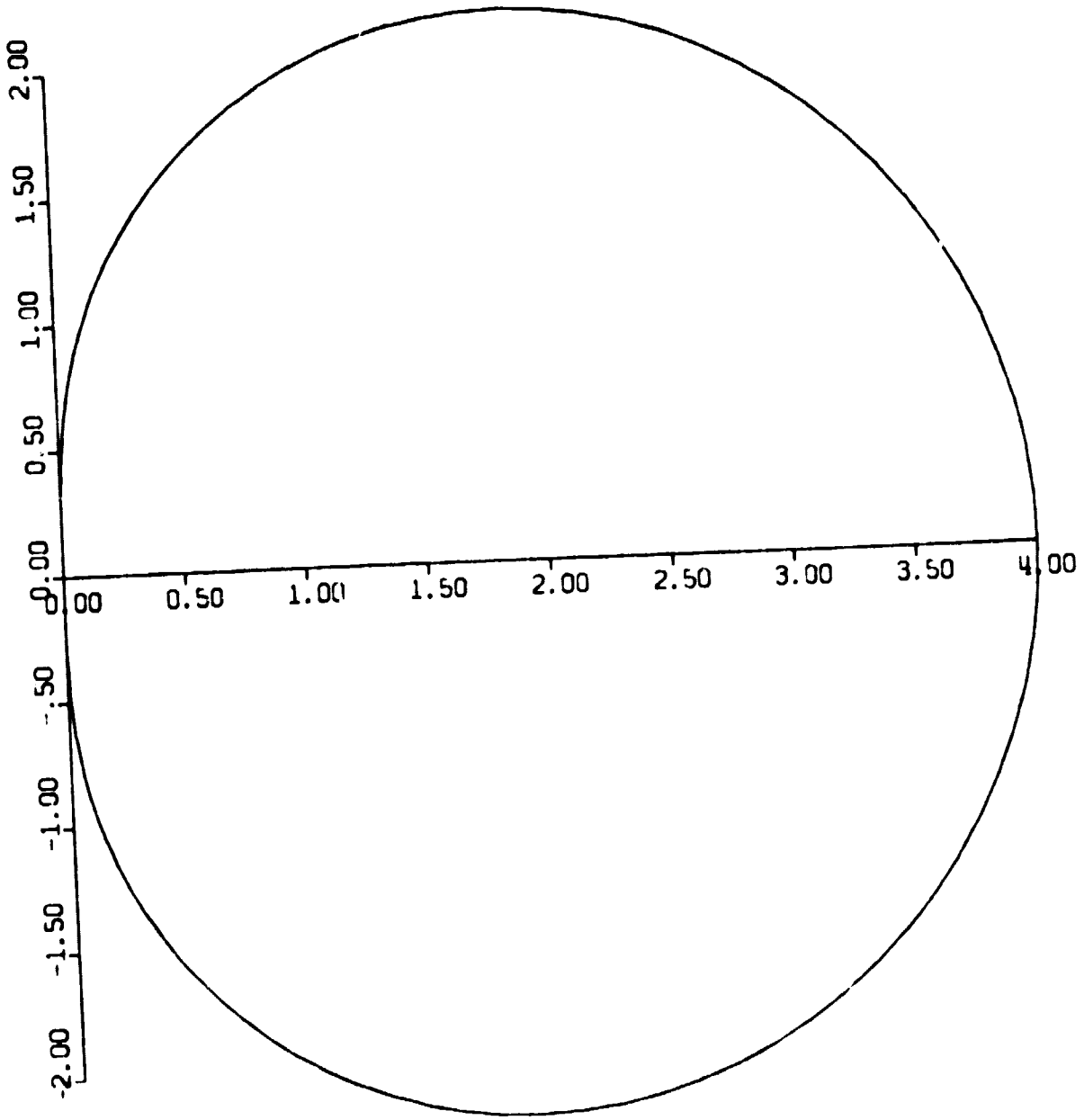
References

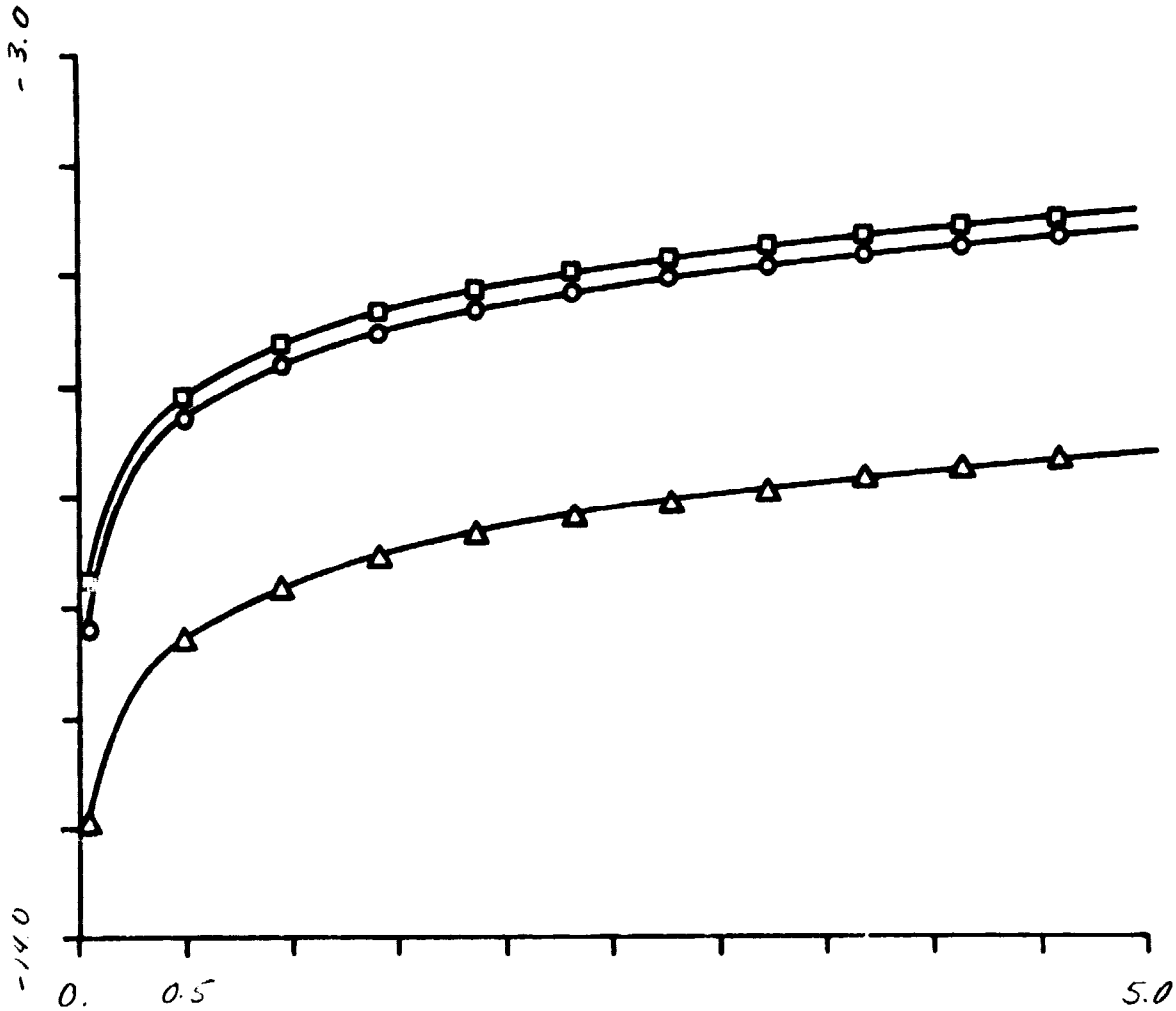
1. Henrici, P., 'Discrete Variable Methods for Ordinary Differential Equations,' Wiley, New York, 1963.
2. Gear, C. W. 'Numerical Initial Value Problems in Ordinary Differential Equations,' Prentice-Hall, Englewood Cliffs, N. J., 1971.
3. Dahlquist, G., "A Special Stability Problem for Linear Multistep Methods," BIT, 1963, 3, pp. 27-43.
4. Lindberg, B., "On a dangerous Property for Methods for Stiff Differential Equations," BIT, 1974, pp. 430-436.

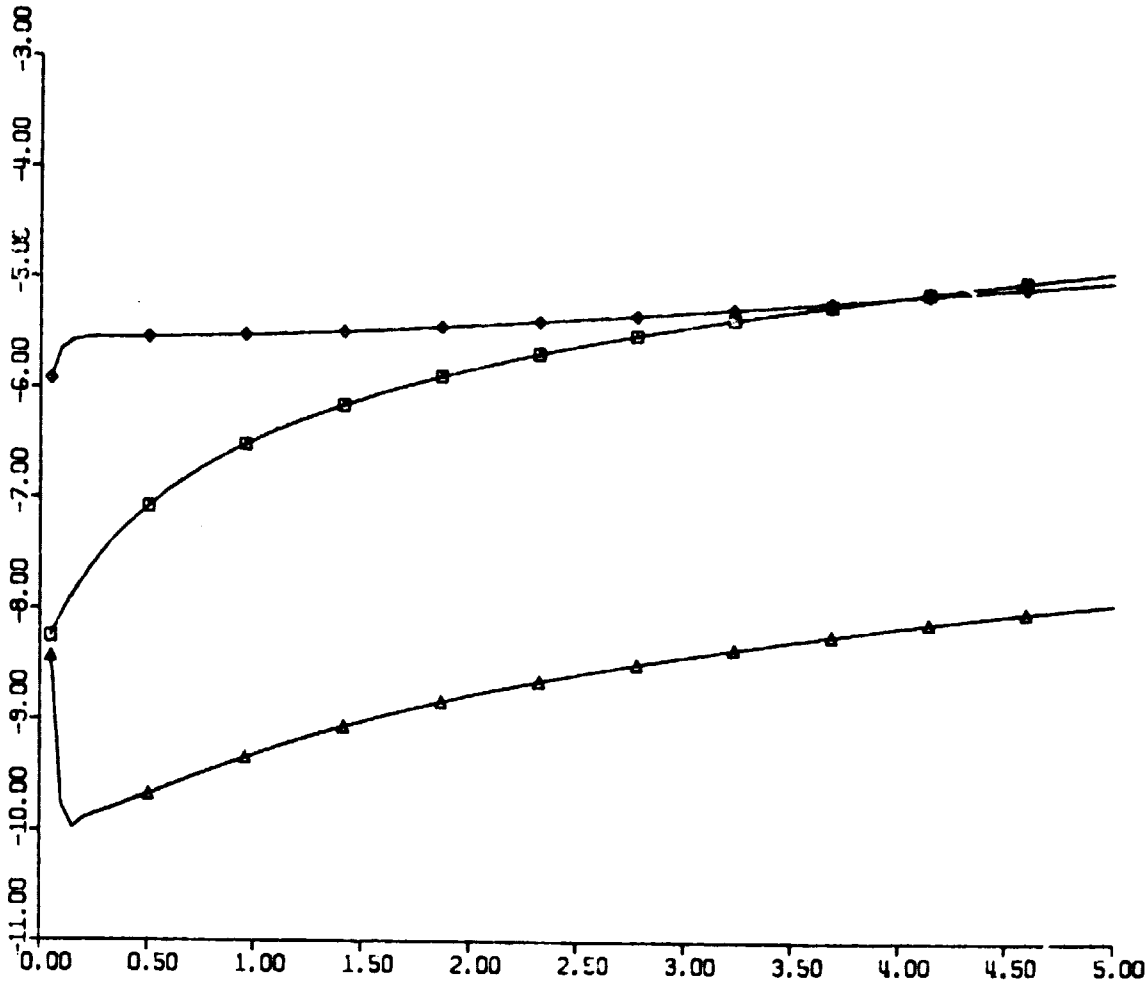
Captions to figures

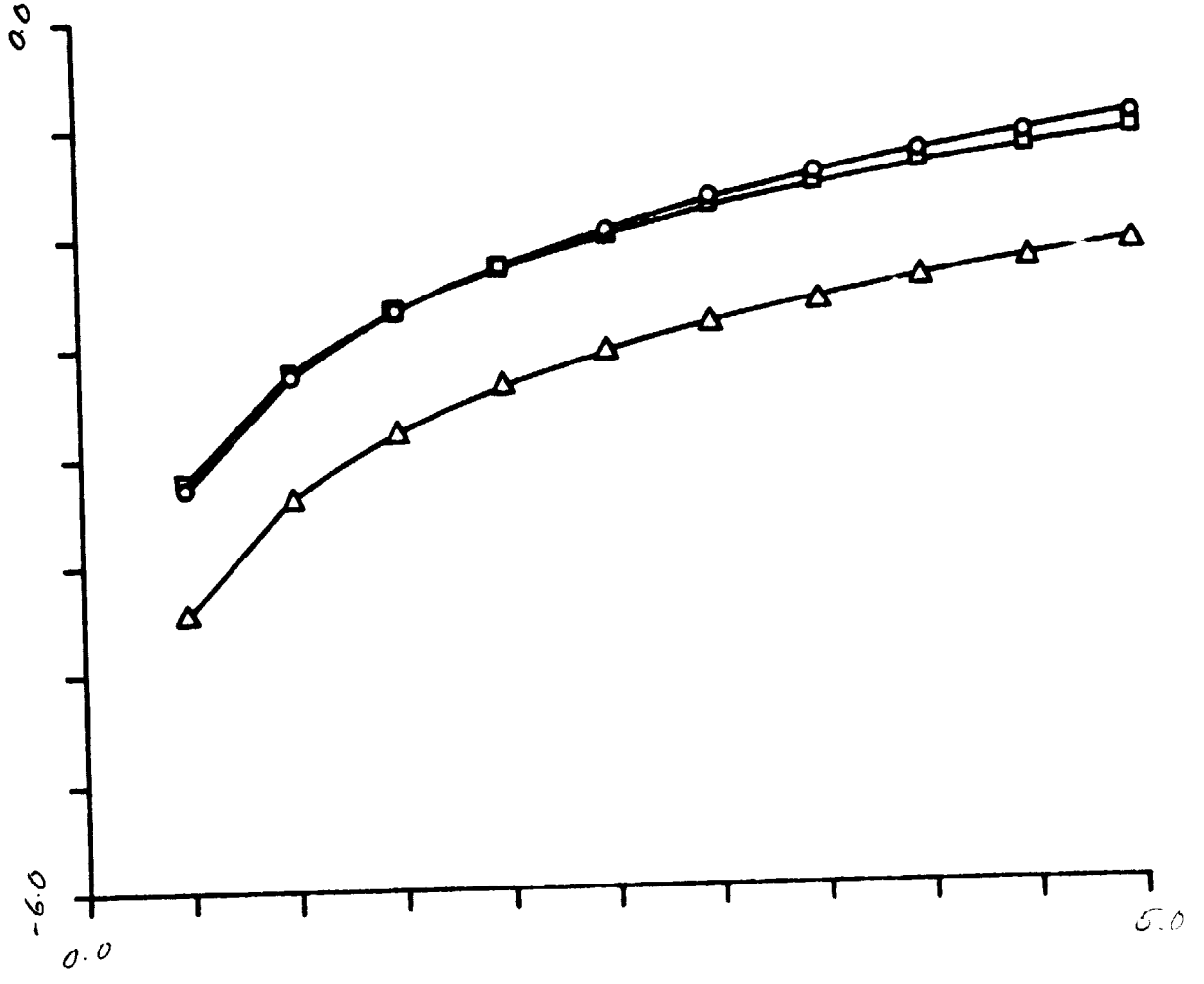
1. Stability regions for AB2 (larger figure) and AB3.
2. Stability region for BDF2 (method is unstable only inside the region indicated).
3. Logarithm of relative error for h = .05, exact starting values - AB2 (□), AB3 (△), BDF2 (◇).
4. Logarithm of relative error for h = .05, inexact starting values - AB2 (□), AB3 (△), BDF2 (◇).
5. Logarithm of relative error for h = .5 - AB2 (□), AB3 (△), BDF2 (◇).


2.00
1.50
1.00
0.50
0.00
-.50
-1.00
-1.50
-2.00
-2.00
-1.50
-1.00
-.50
0.00
0.50
1.00
1.50
2.00


2.00
1.50
1.00
0.50
0.00
-.50
-1.00
-1.50
-2.00
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00


-3.0
-14.0
0.
0.5
5.0


-3.00
-4.00
-5.00
-6.00
-7.00
-8.00
-9.00
-10.00
-11.00
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00


0.0
-6.0
0.0
5.0