VETENSKAP OCH KONST

THE ROYAL INSTITUTE OF TECHNOLOGY
STOCKHOLM
SWEDEN

DEPARTMENT OF
INFORMATION PROCESSING
COMPUTER SCIENCE

Inst för Informationsbehandling
KTH
100 44 STOCKHOLM 70

Dept of Information Processing
The Royal Institute of Technology
S-100 44 STOCKHOLM 70, Sweden

A STABILITY PROPERTY OF THE 3-STEP BACKWARDS DIFFERENTIATION METHOD FOR STIFF NON-LINEAR PROBLEMS

by

Germund Dahlquist

TRITA-NA-7617

ABSTRACT. It is shown how to construct a Liapunov function for proving the stability of the third order BDF method, when the differential system and the step size satisfy the condition,

$$h\langle f(u) - f(v), u-v \rangle \leq -\frac{1}{12}\langle u-v, u-v \rangle .$$

The constant $-\frac{1}{12}$ is best possible, even for linear systems with constant coefficients.

1. INTRODUCTION

The k-step Backwards Difference Method (BDF) is defined by the polynomials

$$\begin{cases} \sigma(\zeta) = \zeta^k \\ \rho(\zeta) = \zeta^k \cdot \sum_{j=1}^{k} (1-\zeta^{-1})^j / j \end{cases} \qquad (1.1)$$

After the transformation,

$$\zeta = \frac{z+1}{z-1} . \quad r(z) = \rho(\zeta)\left(\frac{z-1}{2}\right)^k , \quad s(z) = \sigma(\zeta) \cdot \left(\frac{z-1}{2}\right)^k$$

we obtain,

$$s(z) = \left(\frac{z+1}{2}\right)^k$$

$$r(z) = \sum_{j=1}^{k} \frac{1}{j}\left(\frac{z+1}{2}\right)^{k-j} \qquad (1.2)$$

$$\frac{r(z)}{s(z)} = \sum_{j=1}^{k} \frac{1}{j} \nabla^j \qquad \nabla = \frac{2}{z+1} = 1 - \zeta^{-1} .$$

Let

$$x = \text{Re } z, \quad D_k = -\inf_{x>0} \text{Re } r(z)/s(z) \tag{1.3}$$

It is well known that the interior of the instability region $\mathbb{C}/S$ is equal to the numerical range of $r(z)/s(z)$ for $x>0$, and hence $-D_k$ may be called the stability abscissa of the k-step BDF method. It is well known that $D_1 = D_2 = 0$, since the methods are A-stable for $k \le 2$, and that $D_k > 0$ for $k>2$.

2. DETERMINATION OF D_k.

The image of the imaginary axis under the mapping $q = r(z)/s(z)$ will be tangent to the line $\text{Re } q = -D_k$ at at least two points, corresponding to $z = iy_k^*$, $y_k^* \ge 0$. It follows that

$$\text{Re } r(iy_k^*)/s(iy_k^*) = -D_k$$

$$\text{Re} \frac{d}{dy} r(iy)/s(iy)\Big|_{y=y_k^*} = 0$$

Note that $1+iy = 2\nabla^{-1}$; it follows that $idy = -2\nabla^{-2}d\nabla$. Then, the latter equation is equivalent to

$$\text{Im} \left(\nabla^2 \frac{d}{d\nabla} \sum_{j=1}^{k} \frac{1}{j} \nabla^j \right) = 0,$$

for $y(\nabla) = y_k^*$, and hence

$$0 = \text{Im} \sum_{j=1}^{k} \nabla^{j+1} = \text{Im} \frac{\nabla^2(1-\nabla^k)}{1-\nabla} = 0$$

Note that $(1-\nabla)/\nabla^2 = \nabla^{-1}(1-\nabla^{-1}) = \frac{iy+1}{2} \cdot \frac{iy-1}{2} = -\frac{y^2+1}{4}$ is real. It follows that $\text{Im}(1-\nabla^k) = 0$ and hence ∇^k is real. It follows that, for the k-step BDF method,

$$s(iy^*) = \left(\frac{1+y_k^*}{2} \right)^k \text{ is real} \tag{2.1}$$

For small values of k the general shape of the stability region is known, and it follows that (at least for $2 < k \le 6$) y_k^* is the smallest strictly positive value of y such that s(iy) is real, and hence

$$y_k^* = \tan(\pi/k) \tag{2.2}$$

$$\nabla(iy_k^*) = \frac{2}{1+iy_k^*} = 2\cos(\pi/k)e^{-i\pi/k}$$

$$D_k = -\sum_{j=1}^{k} \frac{1}{j}\,(2\cos\tfrac{\pi}{k})^j \cdot \cos\frac{j\pi}{k} \tag{2.3}$$

In particular,

$$D_3 = \frac{1}{12}, \quad D_4 = \frac{2}{3} \tag{2.4}$$

$$y_3^* = \sqrt{3}, \quad y_4^* = 1.$$

3. CONSTRUCTION OF A LIAPUNOV FUNCTION FOR THE BDF METHOD FOR k = 3.

Let

$$r^*(z) = r(z) + D_k \cdot s(z).$$

It follows from (1.3) that

$$\inf_{x>0} \mathrm{Re}\, r^*(z)/s(z) = 0.$$

For k = 3, let

$$8r^*(z) = r_0(z) + r_1(z)$$
$$8s(z) = s_0(z) + s_1(z)$$

where r_0, s_0 are even functions, r_1, s_1 are odd functions. It follows from (1.2) and (2.4) that, for k = 3,

$$s_0(z) = 3z^2 + 1, \qquad s_1(z) = z(z^2 + 3)$$

$$r_0(z) = \frac{9}{4}(z^2 + 3), \qquad r_1(z) = \frac{1}{12}z(z^2 + 75).$$

Note that $r_0(z)$ and $s_1(z)$ have the common factor z^2+3, which vanishes for $z = iy_3^*$.

We shall now seek a representation [1] ,

$$\mathrm{Re}\, r^*(z) \cdot s(\bar{z}) = A(z)x + B(z), \qquad (z = x + iy)$$

$$A(z) = \sum_{j=1}^{k} |\varphi_j(z)|^2$$

$$B(z) = \sum_{j=1}^{k'} |\Psi_j(z)|^2$$

where the φ_j are k linearly independent polynomials (of degree less than k), and the Ψ_j are k' polynomials (of degree less than k+1), which need not be linearly independent. The relation of this representation to the construction of a Liapunov function is described at the end of this paper.

We shall try the algorithm suggested in [1], and consider therefore,

$$\mathrm{Re}\, r_0(z)s_1(\bar{z}) = \tfrac{9}{4}x\, |z^2+3|^2$$

$$\frac{r_1(z)}{s_0(z)} = \frac{1}{36}z + \frac{56}{9} \cdot \frac{z}{3z^2+1}$$

$$\mathrm{Re}(3z^2+1)\bar{z} = x(3|z|^2 + 1)$$

$$\therefore \quad \mathrm{Re}\, s_0(z)r_1(\bar{z}) = \frac{1}{36}x\, |3z^2+1|^2 + \frac{56}{9}x\, (3|z|^2+1)$$

$$A_0(z) = \frac{9}{4}|z^2+3|^2 + \frac{1}{36}\, |3z^2+1|^2 + \frac{56}{3}\, |z|^2 + \frac{56}{9}$$

which can be written in the form,

$$A_0(z) = (\bar{z}^2,\ 1,\ \bar{z})\cdot A_0 \cdot (z^2,\ 1,\ z)^T$$

with the matrix

$$A_0 = \begin{bmatrix} 10/4 & 41/6 & 0 \\ 41/6 & 53/2 & 0 \\ 0 & 0 & 56/3 \end{bmatrix}$$

In order to construct B(z), we consider

$$r_0(iy)s_0(-iy) + r_1(iy)s_1(-iy) = (-y^2+3)\Big(\tfrac{9}{4}(-3y^2+1) + \tfrac{1}{12}y^2(-y^2+75)\Big) =$$

$$= \frac{1}{12}(-y^2+3)^2(y^2+9)$$

(The divisibility by $(-y^2 + (y^*)^2)^2$ is foreseen by the theory in [1].)

We therefore try

$$B(z) = \frac{1}{12}|z^2+3|^2(|z|^2+9)$$

and calculate

$$x\cdot\Delta A(z) = \mathrm{Re}\,[r_0(z)s_0(\bar{z}) + r_1(\bar{z})s_1(z)] - B(z) =$$

$$= \left(\frac{27}{4}|z|^4 + \frac{90}{4}\mathrm{Re}z^2 + \frac{27}{4} + \frac{1}{12}|z|^6 + \frac{78}{12}|z|^2\mathrm{Re}z^2 + \frac{225}{12}|z|^2\right) -$$

$$- \left(\frac{1}{12}|z|^6 + \frac{6}{12}|z|^2\,\mathrm{Re}z^2 + \frac{9}{12}|z|^2 + \frac{9}{12}|z|^4 + \frac{54}{12}\mathrm{Re}z^2 + \frac{81}{12}\right) =$$

$$= 6|z|^4 + 18\mathrm{Re}z^2 + 6|z|^2\mathrm{Re}z^2 + 18|z|^2 =$$

$$= 2x^2(6|z|^2 + 18).$$

i.e.,

$$\Delta A(z) = (z+\bar{z})\cdot(6|z|^2+18) = 6(z^2\bar{z} + z\bar{z}^2) + 18(z+\bar{z})$$

The construction is successful if the quadratic form $A(z) = A_0(z) + \Delta A(z)$ with the matrix,

$$A = \begin{bmatrix} 10/4 & 41/6 & 6 \\ 41/6 & 53/2 & 18 \\ 6 & 18 & 56/3 \end{bmatrix},$$

is positive definite. This is the case, since we have the factorization,

$$A = R^T D R$$

$$R = \begin{bmatrix} 1 & \frac{41}{15} & \frac{36}{15} \\ 0 & 1 & \frac{144}{704} \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \frac{1}{6}\begin{bmatrix} 15 & 0 & 0 \\ 0 & \frac{704}{15} & 0 \\ 0 & 0 & \frac{16640}{704} \end{bmatrix},$$

In order to obtain the Liapunov function we apply the transformation

$$\rho(\zeta) = (\zeta-1)^k r(z), \qquad z = \frac{\zeta+1}{\zeta-1}$$

i.e.

$$z^2 = (\zeta+1)^2 = \zeta^2 + 2\zeta + 1$$

$$1 = (\zeta-1)^2 = \zeta^2 - 2\zeta + 1$$

$$z = (\zeta+1)(\zeta-1) = \zeta^2 - 1$$

to the quadratic form

$$(\bar{z}^2 \quad 1 \quad \bar{z}) \; A \; (z^2 \quad 1 \quad z^T),$$

We obtain

$$(\bar{\zeta}^2 \quad \bar{\zeta} \quad 1) \cdot \begin{bmatrix} 1 & 1 & 1 \\ 2 & -2 & 0 \\ 1 & 1 & -1 \end{bmatrix} \; A \begin{bmatrix} 1 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} \zeta^2 \\ \zeta \\ 1 \end{bmatrix} =$$

$$= \frac{16}{6}(\bar{\zeta}^2 \quad \bar{\zeta} \quad 1) \begin{bmatrix} 41 & -27 & 9 \\ -27 & 23 & -9 \\ 9 & -9 & 5 \end{bmatrix} \begin{bmatrix} \zeta^2 \\ \zeta \\ 1 \end{bmatrix} .$$

The conclusion of the theory in [1] is that the quadratic vector form,

$$G(Y_n) = \sum_{i=0}^{2} g_{ij} \langle y_{n+i}, y_{n+j} \rangle ,$$

where $g_{11} = 41$, $g_{12} = -27$ etc., is a Liapunov function for the third order BDF method in the following sense:

Consider a differential system $dy/dt = f(y)$ and suppose that the function f and the step-size h satisfy the condition,

$$h \langle u-v, f(u) - f(v) \rangle < -\frac{1}{12} \langle u-v, u-v \rangle ,$$

where $\langle \cdot, \cdot \rangle$ is an innner product in R^s. Let u_n, v_n be two vector sequences obtained by the application of the third order BDF method with different initial conditions to this differential system, and put

$$U_n = \begin{bmatrix} u_{n+k-1} \\ u_{n+k-2} \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} \qquad V_n = \begin{bmatrix} v_{n+k-1} \\ v_{n+k-2} \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix}$$

Then, for $n = 0,1,2,\ldots$

$$G(U_{n+1} - V_{n+1}) \leq G(U_n - V_n).$$

The error bounds given for G-stable methods in [2] or [3] are easily modified to this case. For example, in Theorem 1 of [3], one need only put $\gamma = -2h\mu - h\eta - \frac{1}{12}$ and assume that $\gamma > 0$.

REFERENCES

[1] G. Dahlquist, *"On the Relation of G-stability to Other Stability Concepts for Linear Multistep Methods"*, Report TRITA-NA-7618.

[2] G. Dahlquist, *"Error Analysis for a Class of Methods for Stiff Non-linear Initial Value Problems"*, Numerical Analysis Dundee 1975. Springer Lecture Notes in Mathematics nr 506, 60-74.

[3] O. Nevanlinna, *"On Error Bounds for G-stable Methods"*, BIT 16 (1976), 79-84.