

## Boundary value problems and dichotomic stability

**Citation for published version (APA):**

England, R., & Mattheij, R. M. M. (1988). Boundary value problems and dichotomic stability. *SIAM Journal on Numerical Analysis*, 25(5), 1037-1054. <https://doi.org/10.1137/0725060>

**DOI:**

[10.1137/0725060](https://doi.org/10.1137/0725060)

**Document status and date:**

Published: 01/01/1988

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

## BOUNDARY VALUE PROBLEMS AND DICHOTOMIC STABILITY\*

R. ENGLAND† AND R. M. M. MATTHEIJ‡

**Abstract.** Since the conditioning of a boundary value problem (BVP) is closely related to the existence of a dichotomic fundamental solution (i.e., where one set of modes is increasing and a complementary set is decreasing), it is important to have discretization methods that conserve this dichotomy property. The conditions this imposes on such a method are investigated in this paper.

They are worked out in more detail for scalar second-order equations (the central difference scheme), and for linear first-order systems as well; for the latter type both one-step methods (including collocation) and multistep methods (those that may be used in multiple shooting) are examined.

**Key words.** boundary value problems, dichotomy, stability

**AMS (MOS) subject classifications.** 65L05, 65L10, 65L15

**1. Introduction.** In the study of boundary value problems (BVP) stability notions, describing and interpreting the effects of (small) local perturbations, play an important role. For initial value problems (IVP) the stability theory of numerical methods is very well developed (cf. [3], [5], [14], [24]); for BVP this stability question is much more complicated and less developed, although there has been rapid progress, in particular, for singularly perturbed problems (cf. [1], [13], [21], [25], [28]).

The basic difficulty for BVPs is that the global errors depend on the data of the entire interval on which the problem is defined, whereas they only depend on data of the past interval in IVPs. Nevertheless there is much similarity. Indeed, thinking of a linear problem, where the solution space of the homogeneous equations can be split into subspaces of decaying modes on the one hand and growing modes on the other, it is known that a condition number mainly depends on the boundary conditions (BCs) imposed; these should be such that the “initial conditions” control the decaying modes, and the “terminal conditions” the growing modes [18]. This question is closely related to the notions of conditional stability (cf. [23]) and dichotomy (cf. [4]) (see also [2], [15], [17], [20]). The latter concept will also play an important role here. In principle dichotomy denotes a splitting of solution spaces into subspaces of solutions with a markedly different growth behaviour, like increasing  $\leftrightarrow$  decreasing, increasing faster than a certain exponential rate  $\leftrightarrow$  increasing slower as compared with this rate. Recent results show that in a well-conditioned BVP, the ordinary differential equations (ODEs) should be dichotomic in the sense that there is a splitting into solutions that do not increase significantly on the one hand and do not decrease significantly on the other (see [11]). Since the BCs also control the modes of the discretized problem, it is clear that it makes sense to aim at discretizations that produce a decaying (growing) approximate mode if the corresponding continuous mode is decaying (growing), in particular, for singular perturbation problems.

Among existing BVP algorithms, it seems that multiple shooting (“stabilized march”) types of methods have hardly been investigated from this point of view, in contrast to global approaches such as collocation (cf. [1], [25]). In [1], for example,

---

\* Received by the editors March 1, 1984; accepted for publication (in revised form) August 11, 1987. This work was supported in part by a grant from the Netherlands Organization for the Advancement of Pure Research (ZWO).

† Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Apdo Postal 20-726, 01000 Mexico D. F. Present address, Faculty of Mathematics, The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom.

‡ Faculteit Wiskunde en Informatica, Technische Universiteit Eindhoven, Postbus 513, 5600 MB Eindhoven, the Netherlands.



$$= \begin{bmatrix} h^2\left(f_0 - \frac{5f_1}{2+\alpha}\right) - (2-\alpha) + \frac{5}{2+\alpha} \\ h^2f_1\left(1 + \frac{2-\alpha}{2+\alpha}\right) - \frac{2-\alpha}{2+\alpha} \\ h^2f_2 \\ \vdots \\ h^2f_{N-2} \\ h^2f_{N-3} + 1 \end{bmatrix}.$$

By choosing  $-\lambda^2$  as the forcing term  $f(t)$  the exact solution to the original BVP is  $y(t) = 1$  for  $t \in [0, 1]$ ; of course both (2.2) and (2.3) should integrate this solution exactly so that the only errors to be expected are those due to rounding. In Table 2.1 the errors  $|y_i - y(ih)|$  are given some values of  $i$  where  $\lambda^2$  and  $h$  were chosen as 800 and  $1/80$ , respectively. The computations were performed on an IBM/MVS 4331 with double precision (relative machine error  $\approx 10^{-16}$ ).

The second choice, that is, eliminating  $y_{N-1}$ , leads to the system

$$(2.5) \quad \begin{bmatrix} -5 & 4 & -1 & & & & \\ 2-\alpha & -5 & 4 & -1 & & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & & 2-\alpha & -5 & 4 & -1 & \\ & & & 2-\alpha & -5 & 4-1/(2+\alpha) & \\ & & & & 2-\alpha & -5+4/(2+\alpha) & \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_{N-2} \end{bmatrix} = \begin{bmatrix} h^2f_0 - (2-\alpha) \\ h^2f_1 \\ \vdots \\ h^2f_{N-5} \\ h^2\left(f_{N-4} - \frac{f_{N-1}}{2+\alpha}\right) + \frac{1}{2+\alpha} \\ h^2\left(f_{N-3} + \frac{4}{2+\alpha}f_{N-1}\right) + 1 - \frac{4}{2+\alpha} \end{bmatrix}.$$

A table corresponding to Table 2.1 only shows error  $\leq 2 \times 10^{-15}$ . Therefore it would appear that (2.5) is stable and (2.4) is unstable. A simple explanation can be given as follows: the characteristic polynomial  $\chi$  of the homogeneous part of (2.2) is

$$(2.6) \quad \chi(r) = r^3 - 4r^2 + 5r - 2 + \alpha.$$

TABLE 2.1

$i$	$t_i$	$ y_i - y(t_i) $
0	0	0
10	0.125	$2.00 \times 10^{-14}$
20	0.25	$1.11 \times 10^{-12}$
30	0.375	$6.41 \times 10^{-11}$
40	0.5	$3.30 \times 10^{-9}$
50	0.625	$1.88 \times 10^{-7}$
60	0.75	$1.10 \times 10^{-5}$
70	0.875	$6.40 \times 10^{-11}$
80	1	0

For  $h = \frac{1}{80}$  and  $\lambda^2 = 800$  ( $2 - \alpha = \frac{15}{8}$ ),  $\chi$  has the zeros

$$(2.7) \quad r_1 = \frac{3}{2}, \quad r_2 = \frac{5}{4} - \frac{1}{4}\sqrt{5}, \quad r_3 = \frac{5}{4} + \frac{1}{4}\sqrt{5},$$

that is, two “unstable” roots ( $|r_j| > 1$ ) and one “stable” root ( $|r_j| < 1$ ). Hence the homogeneous part of (2.2) has two increasing basis solutions and one decreasing basis solution. From an analysis such as that given in [18] it follows that it is necessary to have two terminal conditions and one initial condition. Therefore (2.4) cannot be stable, whereas (2.5) is likely to be stable. By comparing the errors in Table 2.1 at points with distance  $10h$ , it becomes clear that the solution  $\{r_1^i\}$  which is second in growth, is responsible for the error growth (a factor  $\approx 58$  after each 10 steps). It turns out that such a dichotomy (cf. [4]) also holds for  $\alpha \in (0, 12]$ ; hence the stability of (2.5) is uniform in  $h$  if  $h \leq \frac{1}{8}$ , although for values of  $\alpha \geq \frac{4}{27}$  the situation becomes more complicated. To examine it, Fig. 2.1 shows a graph of  $\chi(r)$  for  $\alpha = 0$ . As can be seen there is a double root 1 and a third root 2 and there is a local minimum of  $-\frac{4}{27}$  at  $t = \frac{5}{3}$ . For any nonzero value of  $\alpha$ , it is easy to check the roots by shifting this graph up a distance  $\alpha$ . Hence as long as  $\alpha < \frac{4}{27}$  there are three real (and positive) roots, one smaller than, and the other two larger than one. For  $\frac{4}{27} < \alpha < 12$  there is a real root of modulus less than 1 and two complex conjugate roots of modulus greater than 1. For  $\alpha > 12$ , however, the smallest root becomes smaller than  $-1$ . Therefore, if  $h$  and  $\lambda$  are such that  $h^2\lambda^2 > 12$ , both methods are expected to be unstable.

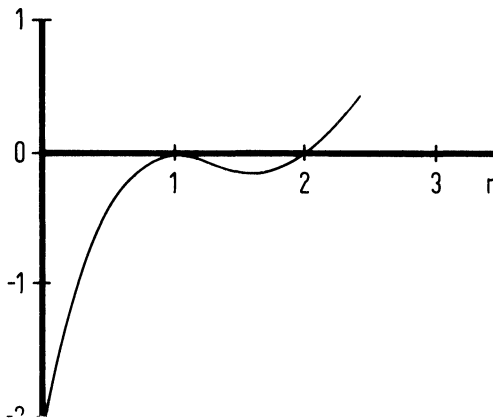


FIG. 2.1

For this problem, the homogeneous differential equation has two modes, a growing mode  $e^{\lambda t}$  controlled by a terminal condition  $y(1) = 1$ , and a decaying mode  $e^{-\lambda t}$  controlled by an initial condition  $y(0) = 1$ . The discretization (2.2) has three modes, one spurious mode being added by the use of a difference equation of higher order than the differential equation. While  $h^2\lambda^2 < 12$  it has two growing modes and one decaying mode. The initial condition controls the decaying mode; the final condition controls the dominant (spurious) growing mode; and to obtain a stable numerical solution, it is necessary to impose the extra boundary condition at the end so as to control the second growing mode. When  $h^2\lambda^2 > 12$  the original decaying mode becomes an oscillatory growing mode and the given initial condition can no longer control it.

It is now appropriate to give the following definition.

DEFINITION 2.8. Let  $\mathcal{D}$  denote a class of linear BVPs on an interval  $[\alpha, \beta]$ , where: (a) the  $n$ th order ( $n \geq 2$ ) ODE,  $Dy = f(t)$ , is such that the homogeneous equation,  $Dy = 0$ , has  $k$  independent solutions  $g_j(t)$  ( $j = 1, \dots, k$ ) which grow in magnitude,

and  $(n - k)$  independent solutions  $g_j(t)$  ( $j = k + 1, \dots, n$ ) with decaying magnitude; and (b) the BC can be written as  $Py(\alpha) = p$  ( $(n - k)$  equations) and  $Qy(\beta) = q$  ( $k$  equations).

Let  $M^h$  be a consistent discretization method defined on an equally spaced grid with mesh size  $h$ , giving an  $m$ th order ( $m \geq n$ ) linear difference equation  $D^h(y^h) = f^h(t)$  and BC  $P^h y^h(\alpha) = p^h$  ( $(m - l)$  equations say),  $Q^h y^h(\beta) = q^h$  ( $l$  equations), where  $y^h$  is defined on the grid. Let the basis solutions of the discrete problem  $D^h(y^h) = 0$ , be denoted by  $r_j(ih)$ ,  $j = 1, \dots, m$ , and ordered in such a way that  $r_j(ih) = g_j(ih) + O(h)$ ,  $j = 1, \dots, n$ . Then  $M^h$  is said to be *dichotomically stable* for  $\tilde{\mathcal{D}} \subset \mathcal{D}$  if for each BVP of this subset and its discretization and each mesh size  $h$  there holds:

- (i)  $|r_j(ih)|$  is an increasing function of  $i$  ( $j = 1, \dots, k$ );
- (ii)  $|r_j(ih)|$  is a decreasing function of  $i$  ( $j = k + 1, \dots, n$ );
- (iii) Of the remaining  $(m - n)$  solutions of the discrete problem,  $r_j(ih)$ ,  $j = n + 1, \dots, m$ ,  $(l - k)$  are such that  $|r_j(ih)|$  increases and  $(m - l - n + k)$  are such that  $|r_j(ih)|$  decreases with  $i$ .  $\square$

*Remark 2.9.* Definition 2.8 requires that to the spurious modes (only if  $m > n$ ) there correspond the *proper type* of BCs (e.g., initial conditions for decaying ones). It does not require, however, that these BCs actually control the discrete solutions at all. Including this requirement would make it necessary to introduce a condition number or some threshold for it, not only for the discretization but also for the original problem. Since the actual determination of the discrete BCs might be uncoupled from the determination of the difference equation, our definition is still a meaningful one, as it gives a necessary condition for a stable or well-conditioned discrete problem. Thus, dichotomic stability is by no means a sufficient condition for a proper discretization. Quite apart from accuracy criteria, the conditioning of a problem (continuous or discrete) depends on the boundary conditions, which are not considered in Definition 2.8.

In what follows, attention will be mainly restricted to difference equations which are of the same order as the differential equations. The following simpler notion, applied to an equation with constant coefficients, may then be used.

**DEFINITION 2.10.** Let the basis solutions of the homogeneous part of an  $n$ th order linear differential equation have components proportional to  $e^{\lambda_j t}$  ( $j = 1, \dots, n$ ). Let  $M^h$  be a discretization method giving an  $n$ th order linear difference equation with corresponding basis solutions of the discrete problem having components with growth  $\{(r_j^h)^i\}$ , where  $r_j^h = 1 + h\lambda_j + O(h^2)$ . Then  $M^h$  is *di(chotomically) stable on a region*  $R \subset \mathbb{C}$  if and only if, for  $h\lambda_j \in R$ ,

- (i)  $\text{Re}(h\lambda_j) \leq 0 \Rightarrow |r_j^h| \leq 1$ ,
- (ii)  $\text{Re}(h\lambda_j) \geq 0 \Rightarrow |r_j^h| \geq 1$ .  $\square$

Obviously, if  $m = n$  in Definition 2.8 condition (iii) is redundant, while, for the constant coefficient case, Definition 2.8(i) and (ii) are equivalent to Definition 2.10(i) and (ii). The class  $\tilde{\mathcal{D}}$  then consists of BVPs for which the homogeneous part of the ODEs has basic solutions that for a given  $h$  satisfy Definition 2.10(i) and (ii). In this sense Definition 2.10 is consistent with Definition 2.8. Of course, Definition 2.10 resembles the more familiar IVP notions of stability such as absolute stability and  $A(\alpha)$  stability [5], [29].

*Example 2.11.* If  $M^h$  is the forward Euler method, then  $R = \{z \in \mathbb{C} \mid \text{Re}(z) \leq 0\} \cup \{z \in \mathbb{C} \mid |z + 1| \leq 1\}$ .

**3. An investigation of the di-stability of some methods.** In this section the di-stability of a number of methods is investigated to demonstrate the use of the concept, rather

than to give a detailed discussion of all its aspects. In § 3.1 consideration is given to the classical central difference scheme for scalar second-order ODEs. Then in § 3.2 one-step difference equations for systems of first-order differential equations are examined. Finally, in § 3.3 a more special one-step equation is considered, viz. multiple shooting relations, and attention is paid to questions to selection of appropriate discretization schemes for integration over a shooting interval.

**3.1. The central difference scheme.** Consider the scalar second-order ODE

$$(3.1) \quad y'' + py' + qy = 0.$$

Perhaps one of the oldest numerical discretization methods uses central differences

$$(3.2) \quad \left(1 + \frac{p}{2}h\right)y_{i+1}^h + (-2 + qh^2)y_i^h + \left(1 - \frac{p}{2}h\right)y_{i-1}^h = 0.$$

As may be seen the basis solutions of (3.1) are  $e^{\lambda_1 t}$ ,  $e^{\lambda_2 t}$  with

$$(3.3) \quad \lambda_1, \lambda_2 = -\frac{p}{2} \pm \frac{1}{2}\sqrt{p^2 - 4q},$$

and those of (3.2) are  $\{(r_1^h)^i\}$ ,  $\{(r_2^h)^i\}$  with

$$(3.4) \quad r_1^h, r_2^h = \left\{1 - \frac{q}{2}h^2 \pm \frac{1}{2}h\sqrt{p^2 - 4q + q^2h^2}\right\} / \left(1 + \frac{p}{2}h\right).$$

Note that Definition 2.10 may be used to investigate the di-stability region  $R$ . The expressions in (3.3) and (3.4) look rather uninviting to use for exploring this  $R$ . It is therefore possible to proceed as follows:

Write for short

$$(3.5) \quad K^h = \frac{-2 + qh^2}{1 + (p/2)h}, \quad L^h = \frac{1 - (p/2)h}{1 + (p/2)h}.$$

It is easy to see that

$$(3.6a) \quad |r_1^h|, |r_2^h| \text{ are both } \leq 1 \text{ iff } |K^h| - 1 \leq L^h \leq 1,$$

$$(3.6b) \quad |r_1^h|, |r_2^h| \text{ are both } \geq 1 \text{ iff } |K^h| + L^h \leq -1 \text{ or } |K^h| - L^h \leq 1 \leq L^h,$$

$$(3.6c) \quad \text{one of } |r_1^h|, |r_2^h| \text{ is } \geq 1 \text{ and the other } \leq 1 \text{ iff } |K^h| \geq |L^h + 1|.$$

In order to find out the values of  $p$ ,  $q$  for which  $r_1^h$  and  $r_2^h$  are smaller or larger than 1 in modulus, it is necessary to investigate these inequalities separately for the case  $1 + ph/2 > 0$  and  $1 + ph/2 < 0$  (cf. (3.5)). In Fig. 3.1 the various regions in which (3.2) has similar dichotomy properties are sketched, using as abscissa  $qh^2$  and as ordinate  $ph$ .

As may be seen, the lines  $qh^2 = 0$ ,  $qh^2 = 4$  and  $ph = 0$  form the boundaries of these regions:

$$(3.7a) \quad |r_1^h|, |r_2^h| \text{ are both } \leq 1 \text{ iff } 0 \leq qh^2 \leq 4, \quad ph \geq 0,$$

$$(3.7b) \quad |r_1^h|, |r_2^h| \text{ are both } \geq 1 \text{ iff } 0 \leq qh^2 \leq 4, \quad ph \leq 0,$$

$$(3.7c) \quad \text{one of } |r_1^h|, |r_2^h| \text{ is } \geq 1 \text{ and the other is } \leq 1 \text{ iff } |qh^2 - 2| \geq 2.$$

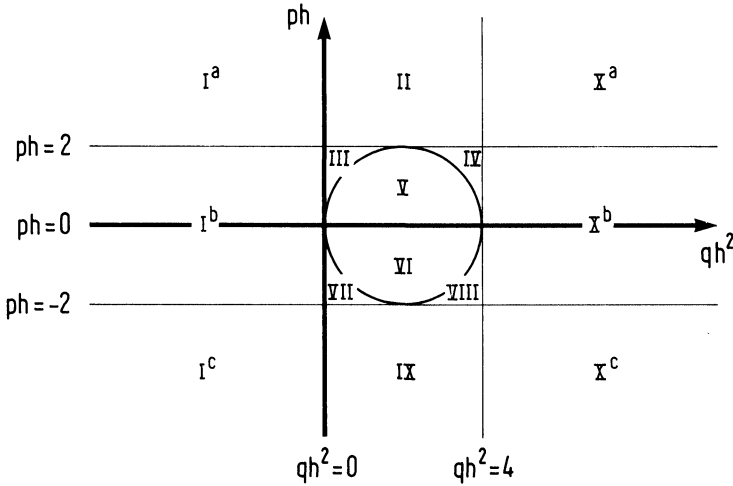


FIG. 3.1

In greater detail the regions may be broken down as shown into

- $I^a$ :  $(qh^2 \geq 0, ph \geq 2): r_1^h \geq 1, 0 \geq r_2^h \geq -1,$
- $I^b$ :  $(qh^2 \leq 0, |ph| < 2): r_1^h \geq 1, 0 \geq r_2^h \geq 1,$
- $I^c$ :  $(qh^2 \leq 0, ph < -2): r_1^h \leq -1, 0 \geq r_2^h \geq 1,$
- II:  $(|qh^2 - 2| \leq 2, ph \geq 2): 0 \geq r_1^h \geq 1, 0 \geq r_2^h \geq -1,$
- III:  $(0 \leq qh^2 \leq 2, 0 \leq ph \leq 2, (ph)^2 + (qh^2 - 2)^2 \geq 4): 0 \geq r_1^h, r_2^h \geq 1,$
- IV:  $(2 \leq qh^2 \leq 4, 0 \leq ph \leq 2, (ph)^2 + (qh^2 - 2)^2 \geq 4): 0 \geq r_1^h, r_2^h \geq -1,$
- V:  $(ph \geq 0, (ph)^2 + (qh^2 - 2)^2 < 4): |r_1^h| = |r_2^h| \leq 1, \text{ complex conjugates,}$
- VI:  $(ph \leq 0, (ph)^2 + (qh^2 - 2)^2 < 4): |r_1^h| = |r_2^h| \geq 1, \text{ complex conjugates,}$
- VII:  $(0 \leq qh^2 \leq 2, 0 \geq ph > -2, (ph)^2 + (qh^2 - 2)^2 \geq 4): r_1^h, r_2^h \geq 1,$
- VIII:  $(2 \leq qh^2 \leq 4, 0 \geq ph > -2, (ph)^2 + (qh^2 - 2)^2 \geq 4): r_1^h, r_2^h \leq -1,$
- IX:  $(|qh^2 - 2| \leq 2, ph < -2): r_1^h \leq -1, r_2^h \geq 1,$
- $X^a$ :  $(qh^2 \geq 4, ph \geq 2): 0 \geq r_1^h \geq 1, r_2^h \leq -1,$
- $X^b$ :  $(qh^2 \geq 4, |ph| < 2): 0 \geq r_1^h \geq -1, r_2^h \leq -1,$
- $X^c$ :  $(qh^2 \geq 4, ph < -2): 0 \geq r_1^h \geq -1, r_2^h \geq 1.$

It is interesting to compare Fig. 3.1 with Fig. 3.2, where regions are shown in which

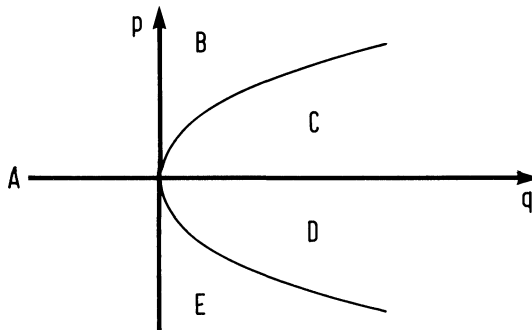


FIG. 3.2



the original differential equation (3.1) has similar dichotomy properties. As before, the lines  $q = 0$  and  $p = 0$  form the boundaries of the regions:

$$(3.8a) \quad \operatorname{Re}(\lambda_1), \operatorname{Re}(\lambda_2) \text{ are both } \leq 0 \text{ iff } q \geq 0, p \geq 0 \quad (B, C),$$

$$(3.8b) \quad \operatorname{Re}(\lambda_1), \operatorname{Re}(\lambda_2) \text{ are both } \geq 0 \text{ iff } q \geq 0, p \leq 0 \quad (D, E),$$

$$(3.8c) \quad \text{one of } \operatorname{Re}(\lambda_1), \operatorname{Re}(\lambda_2) \text{ is } \geq 0 \text{ and the other is } \leq 0 \text{ iff } q \leq 0 \quad (A).$$

In the regions  $C, D$  shown, where  $4q \geq p^2$ ,  $\lambda_1, \lambda_2$  are complex conjugates, but elsewhere they are real. The scales in Fig. 3.2 are somewhat arbitrary and the position of the separatrix  $p^2 = 4q$  does not change if  $qh^2$  and  $ph$  are used instead of  $q$  and  $p$  as abscissa and ordinate. The two figures are thus directly comparable.

If  $q < 0$  (region  $A$ ), the differential equation (3.1) has two modes: a growing mode with  $\lambda > 0$  and a decaying mode with  $\lambda < 0$ . For any positive value of  $h$ , the discretization method (3.2) also has two modes: a growing mode with  $|r| > 1$  and a decaying mode with  $|r| < 1$  (regions  $I^a, I^b, I^c$ ). Thus for all values of  $p$  and  $h$ , the method is di-stable for  $q < 0$ . Results will however deteriorate for  $|ph|$  much greater than 2 as in regions  $I^a, I^c$  where one of the two modes becomes oscillatory ( $r < 0$ ), although it still decays in the appropriate direction.

If  $q > 0, p > 0$  (regions  $B, C$ ) the differential equation (3.1) has two decaying modes with  $\operatorname{Re}(\lambda) < 0$ . If  $qh^2 \leq 4$ , the discretization method also has two decaying modes with  $|r| < 1$  (regions II, III, IV, V). Similarly, if  $q > 0, p < 0$  (regions  $D, E$ ) the differential equation (3.1) has two growing modes with  $\operatorname{Re}(\lambda) > 0$ . If  $qh^2 \leq 4$ , the discretization method also has two growing modes with  $|r| > 1$  (regions VI, VII, VIII, IX). Thus for all values of  $p$ , the method is di-stable for  $qh^2 \leq 4$ . Once again, results will deteriorate for  $|ph|$  greater than about 2, as in regions II, IX where one of the two modes becomes oscillatory, while those of the differential equation do not oscillate in regions  $B, E$ .

Clearly the method is not di-stable for  $qh^2 > 4$ , as the differential equation still has two decaying modes if  $p > 0$ , or two growing modes if  $p < 0$ , but the discretization method has one decaying mode and one growing mode in regions  $X^a, X^b, X^c$ .

**3.2. One-step difference equations for systems.** Most BVP algorithms for systems of first-order differential equations are based on finding (directly or indirectly) a one-step recurrence relation for approximate solutions on a certain grid. Together with the BC this then leads to a linear system from which the (approximate) solution can be found. Examples are the Box scheme [12], higher-order difference schemes [16], collocation [1], [22], and multiple shooting [6], [19]. In this section consideration is given to one-step difference equations that arise from one-step discretization schemes, such as Runge-Kutta or one-step Obrechhoff formulae. It is well known that a large class of collocation methods can be interpreted as implicit Runge-Kutta methods [27]. These methods have been extensively investigated for their IVP stability properties (cf. [1], [3], [26]).

A general Runge-Kutta method for the system of differential equations  $y' = f(t, y)$  may be written

$$(3.9a) \quad q_k = y_i^h + h \sum_{l=1}^m a_{kl} f(t_l + c_l h, q_l), \quad k = 1, 2, \dots, m,$$

$$(3.9b) \quad y_{i+1}^h = y_i^h + h \sum_{l=1}^m b_l f(t_l + c_l h, q_l)$$

where the coefficients  $a_{kl}$  ( $k, l = 1, 2, \dots, m$ ),  $b_l$  ( $l = 1, 2, \dots, m$ ) define the method,  $c_k = \sum_{l=1}^m a_{kl}$ ,  $k = 1, 2, \dots, m$ , and for consistency  $\sum_{l=1}^m b_l = 1$ . Clearly any permutation of the rows of the matrix  $A = \{a_{kl}\}$  applied also to its columns and to the elements of the vectors  $b = \{b_l\}$ ,  $c = \{c_k\}$ , will not change the method. When written in the form (3.9), such a method expresses  $y_{i+1}^h$  in terms of  $y_i^h$ , but for BVPs it is of equal interest to express  $y_i^h$  in terms of  $y_{i+1}^h$ , as though the integration were to be performed in the opposite direction. Thus

$$(3.10a) \quad q_k = y_{i+1}^h - h \sum_{l=1}^m (b_l - a_{kl})f(t_{i+1} - (1 - c_l)h, q_l), \quad k = 1, 2, \dots, m,$$

$$(3.10b) \quad y_i^h = y_{i+1}^h - h \sum_{l=1}^m b_l f(t_{i+1} - (1 - c_l)h, q_l)$$

and if the coefficients  $a_{kl}$ ,  $b_l$ ,  $c_k$  represent a method applied in the forward direction, the same method applied in the backward direction is represented by the coefficients  $(b_l - a_{kl})$ ,  $b_l$ ,  $(1 - c_k)$ .

DEFINITION 3.11. A Runge-Kutta method will be called *symmetric* if it is the same method whether applied in a forward or backward direction.

Thus, letting  $e$  be the vector of  $m$  unit elements, the method (3.9) is symmetric if there is a permutation matrix  $P$  such that

- (i)  $eb^T - A = PAP^T$ ,
- (ii)  $b = Pb$ ,
- (iii)  $e - c = Pc$ .

As defined in Definition 2.10, the concept of di-stability is given in terms of constant coefficient homogeneous linear differential equations. The most general first-order system of this type may be written  $y' = Jy$  where  $J$  is a constant matrix, but since the method (3.9) only uses the function  $f(t, y)$  in a linear manner, it is not affected by a linear change of variables which causes  $J$  to undergo a similarity transformation. For practical purposes, it is sufficient to consider  $J$  a complex diagonal matrix, and thus to examine the effect of the method on one single differential equation of the form  $y' = \lambda y$ , as is normally done for linear stability analysis of methods for initial value problems.

When method (3.9) is applied to this standard test equation, it is found that

$$(3.12) \quad y_{i+1}^h = y_i^h r(h\lambda)$$

where  $r(z)$  is a rational function of degree not exceeding  $m$  in the numerator and the denominator. Explicitly (cf. [3]),

$$(3.13) \quad r(z) = 1 + b^T z (I - Az)^{-1} e$$

the denominator being the determinant of  $I - Az$ .

For a symmetric Runge-Kutta method, and indeed for a larger class of Runge-Kutta methods which could be called linearly symmetric, the rational growth factor  $r(z)$ , which characterizes the fundamental solutions of the difference equation, is such that

$$(3.14) \quad r(-z) = 1/r(z).$$

For one-step Obrechhoff formulae (cf. [14])

$$(3.15) \quad y_{i+1}^h = y_i^h + \sum_{l=1}^m h^l (\beta_{0l} y_{i+1}^{(l)} + \beta_{1l} y_i^{(l)})$$

Downloaded 05/20/15 to 131.155.151.137. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where  $y_i^{(l)}$  is the  $l$ th derivative of  $y(t)$  obtained by repeated differentiation of the differential equation  $y' = f(t, y)$  at  $(t, y) = (t_i, y_i^h)$ . Applying such a method to the standard test equation  $y' = \lambda y$ , it is again found that (3.12) is satisfied, where now

$$(3.16) \quad r(z) = \left(1 + \sum_{l=1}^m \beta_{1l} z^l\right) / \left(1 - \sum_{l=1}^m \beta_{0l} z^l\right).$$

Thus such a method may be defined to be symmetric when  $\beta_{1l} = (-1)^{l-1} \beta_{0l}$  ( $l = 1, 2, \dots, m$ ) which is equivalent to (3.14).

There is clearly a one-to-one relationship between one-step Obrechhoff formulae and their growth factors  $r(z)$ . However, in general, many different one-step methods may give rise to the same rational growth factor  $r(z)$ . Nevertheless, their linear stability properties such as absolute stability,  $A(\alpha)$ -stability, and also di-stability will be entirely determined by the function  $r(z)$ . Thus, by using Definition 2.10, a one-step method is dichotomically stable on a region  $R \subset \mathbb{C}$  if and only if for all  $z \in R$ , its growth factor  $r(z)$  satisfies

$$(3.17a) \quad \operatorname{Re}(z) \leq 0 \Rightarrow |r(z)| \leq 1,$$

$$(3.17b) \quad \operatorname{Re}(z) \geq 0 \Rightarrow |r(z)| \geq 1.$$

**THEOREM 3.18.** *A one-step method that is symmetric and  $A(\alpha)$ -stable is di-stable on the region  $R := \{z \in \mathbb{C} \mid |\operatorname{Re}(z)| \sin \alpha \geq |\operatorname{Im}(z)| \cos \alpha\}$  ( $0 < \alpha \leq \pi/2$ ).*

*Proof.* Since the method is  $A(\alpha)$ -stable, the growth factor  $r(z)$  satisfies  $|r(z)| \leq 1$  on  $R^- := \{z \in \mathbb{C} \mid -\operatorname{Re}(z) \sin \alpha \geq |\operatorname{Im}(z)| \cos \alpha\}$ . Since it is symmetric  $|r(z)| = 1/|r(-z)| \geq 1$  on  $R^+ := \{z \in \mathbb{C} \mid \operatorname{Re}(z) \sin \alpha \geq |\operatorname{Im}(z)| \cos \alpha\}$ . Thus (3.17) is satisfied on  $R = R^- \cup R^+$ .  $\square$

So we may conclude that a stability property in the left half plane plus symmetry as investigated in [1], [25] gives indeed a sufficient criterion for di-stability.

*Example 3.19.* The implicit midpoint rule (which is also the Box scheme, or a collocation method at a single Gauss quadrature point), or the trapezoidal rule (which is a collocation method at two Lobatto points, or a simple one-step Obrechhoff formula), are both implicit Runge–Kutta methods with growth factor

$$(3.20) \quad r(z) = \frac{1 + \frac{1}{2}z}{1 - \frac{1}{2}z}.$$

These methods are  $A$ -stable ( $A(\alpha)$ -stable with  $\alpha = \pi/2$ ) and symmetric. Therefore they are di-stable on the whole of the complex plane  $\mathbb{C}$ .

*Example 3.21.* Any one-step method whose growth factor  $r(z)$  is a diagonal Padé approximant to the exponential  $e^z$  is  $A$ -stable [3]. These Padé approximants also satisfy (3.14), and so such methods are di-stable on the whole of the complex plane  $\mathbb{C}$ .

Since a symmetric method has a growth factor which satisfies (3.14) it also satisfies  $|r(z)|^2 = r(z)\bar{r}(z) = r(z)r(\bar{z}) = r(z)r(-z) = 1$  whenever  $\operatorname{Re}(z) = 0$ . It is thus tempting to suppose that any symmetric method might be di-stable on the whole of the complex plane  $\mathbb{C}$ . The following counterexample shows that this is not the case.

*Example 3.22.* A one-step method (Runge–Kutta or one-step Obrechhoff formula) whose rational growth factor is

$$(3.23) \quad r(z) = \frac{1 + \frac{1}{2}z + \frac{1}{3}z^4}{1 - \frac{1}{2}z + \frac{1}{3}z^4}$$

is symmetric and  $A(\alpha)$ -stable with  $\alpha = \pi/6$ , but is not  $A$ -stable, and hence not di-stable on the whole of the complex plane  $\mathbb{C}$ . It is in fact di-stable on the region

$$R := \{z = x + iy \mid 3y^2 \leq -x^2 + \sqrt{4x^4 + 9}\}$$

which contains the smaller region

$$\hat{R} := \{z \in \mathbb{C} \mid |\operatorname{Re}(z)| \geq \sqrt{3} |\operatorname{Im}(z)|\}.$$

General one-step methods will have their mesh size restricted for stability reasons. This is well known for explicit formulae applied to initial value problems, in which case it is necessary that, for each eigenvalue  $\lambda$  of the Jacobian matrix  $J$ , the product  $h\lambda$  must lie in the absolute stability region  $R := \{z \in \mathbb{C} \mid |r(z)| < 1\}$ . When applied to boundary value problems, implicit formulae also suffer from such a stability restriction upon their step size, as the growing modes must be properly represented. However, a symmetric method which is also  $A$ -stable will not suffer from any mesh size restriction owing to stability, as it is di-stable on the whole of the complex plane  $\mathbb{C}$ . A symmetric  $A(\alpha)$ -stable method will be equally efficient provided the eigenvalues  $\lambda$  of the Jacobian matrix  $J$  lie within the appropriate sectors of di-stability of the complex plane, although these statements are to some extent dependent upon the assumption that the matrix  $J$  is constant. The boundary conditions could also disturb the overall stability of such a discretization, as is well known in the case of the implicit midpoint rule and the trapezoidal rule, which suffer from the same kind of oscillation as the central difference scheme for second-order equations.

**3.3. Multiple shooting.** As illustrated in Example 2.1, the use of a difference equation of order higher than that of the differential system imposes the need for additional boundary conditions. Such a situation most naturally arises when using a central difference scheme of higher order than the differential equation, e.g., fourth-order (five-point) central differences for a second-order differential equation. In those cases, an analysis similar to that performed in Example 2.1 should show where additional boundary conditions are needed.

However, when treating systems of first-order differential equations, shooting or multiple shooting are very natural approaches. One way of viewing multiple shooting is to think of a basic discretization method (one-step or multistep) but to eliminate from the algebraic equations the solution values internal to each shooting interval (internal condensation) or all the internal variables in the case of simple shooting. In practice the difference equations are set up sequentially, and internal variables eliminated as soon as they are not needed. The step size may also be determined and varied dynamically.

If the basic discretization method is a one-step scheme, it is important that it should be dichotomically stable for the problem to which it is applied. Then any decaying (growing) mode of the differential system generates a decaying (growing) mode of the basic difference equation, which is controlled by an initial (terminal) condition. The only difference from direct solution of the one-step difference equation (apart from a reduction in storage requirements) is that sequential block elimination of internal variables may be an unstable process, and lead to swamping of decaying modes by rounding errors in the growing modes. This is one of the principal motivations behind multiple shooting, in which, unlike simple shooting, the sequential block elimination is not carried too far before a new uncontaminated set of fundamental solutions is again restored. The global difference equation, which relates values at the ends of the shooting intervals, must then be solved, by some stable recursion process, involving decoupling of decaying and growing modes and not by direct block elimination.

Since, in each shooting interval, we are solving a number of initial value problems, it is natural to consider multistep schemes, instead of one-step schemes, for use in the basic discretization. As always, the use of a multistep discretization method for a

system of first-order differential equations introduces spurious modes, and imposes the need for additional boundary conditions. In general, since the discretization should be dichotomically stable for the problem to be solved, Definition 2.8 provides a criterion as to how many additional conditions should be imposed at the beginning or at the end of the interval—an initial condition for each spurious decaying mode and a terminal condition for each growing mode.

However, in a multiple shooting context, a multistep discretization scheme would be used to solve initial value problems, and sequentially eliminate internal variables. Thus, for practical reasons the additional boundary conditions should all be extra starting values, which are always required in conjunction with multistep methods for initial value problems. These extra initial conditions could be generated by some one-step method of high order, or by low-order methods of the multistep family at small step size as is done in automatic, variable-order, initial value integrators. The important point is that the multistep discretization method together with the starting procedure (which provides the additional initial conditions) should be dichotomically stable for the problem to be solved. This means that any spurious mode of the multistep method should be of the decaying type.

These considerations lead us to consider new stability properties of discretization formulae for solving initial value problems. In the context of stiff initial value problems, the need to represent decaying modes by decaying approximations has led to the definition of  $A(\alpha)$ -stability (cf. [5], [29]).

Clearly, when considering multiple shooting for boundary value problems, there will normally be growing modes present, and it would be meaningless to require that a discretization represent them by decaying approximations, or be absolutely stable for some  $h\lambda$  with positive real part. The first idea might be to consider some form of relative stability to ensure that numerical approximations do not grow faster than the true growing modes. This leads to the following definition of  $R(\beta)$ -stability.

**DEFINITION 3.24.** If a discretization  $M^h$  applied to the test equation  $y' = \lambda y$  generates a difference equation with basis solutions  $\{(r_j^h)^i\}$ ,  $j = 1, 2, \dots, k$ , it will be called  $R(\beta)$ -stable if  $|r_j^h| < |e^{h\lambda}|$  ( $j = 1, 2, \dots, k$ ) for all step sizes  $h$ , and  $\lambda \in \mathbb{C}$  such that  $h\lambda \in R^+ := \{z \in \mathbb{C} \mid \operatorname{Re}(z) \sin \beta > |\operatorname{Im}(z)| \cos \beta\}$ .  $\square$

This may be interpreted by saying that  $R^+$  lies within the white fingers of the order star of the method (cf. [26]).

Definition 3.24 does not contribute towards the dichotomic stability of the discretization. One important requirement for dichotomic stability is that growing modes should be represented by growing approximations, or that the discretization be absolutely unstable for  $h\lambda$  with positive real part. This idea leads to the definition of the following desirable property.

**DEFINITION 3.25.** If a discretization  $M^h$  applied to the test equation  $y' = \lambda y$  generates a difference equation with fundamental solutions  $\{(r_j^h)^i\}$ ,  $j = 1, 2, \dots, k$ , it will be called  $A(\gamma)$ -unstable if one of the  $r_j^h$ , say  $r_1^h$ , satisfies  $|r_1^h| > 1$  for all step sizes  $h$ , and  $\lambda \in \mathbb{C}$  such that  $h\lambda \in R^+ := \{z \in \mathbb{C} \mid \operatorname{Re}(z) \sin \gamma > |\operatorname{Im}(z)| \cos \gamma\}$ .  $\square$

Most methods for stiff initial value problems are not  $A(\gamma)$ -unstable, as they concentrate on decaying modes, and often require that  $\lim_{|h\lambda| \rightarrow \infty} |r_j^h| < 1$  ( $j = 1, 2, \dots, k$ ). However, at least for boundary value problems, it does seem that  $A(\gamma)$ -instability is a desirable property. Nevertheless, it is still not sufficient to determine a region of dichotomic stability of the global multiple shooting equations. For this it is necessary that, while the dominant approximate mode grows, the spurious modes of the formula decay. Thus there should be a region  $R^+$  in which  $|r_1^h| \geq 1$  and  $|r_j^h| \leq 1$ ,  $j = 2, 3, \dots, k$ . If the discretization satisfies these inequalities in  $R^+$  of Definition 3.25,

and is also  $A(\alpha)$ -stable, then the global equation is dichotomically stable on  $R^+$ , together with the relevant region in the left-hand side of the complex plane.

*Example 3.26.* The methods given in Example 3.19 are  $A(\alpha)$ -stable with  $\alpha = \pi/2$ , and also  $A(\gamma)$ -unstable with  $\gamma = \pi/2$ , since there is only one fundamental solution  $\{(r_1^h)^i\} = \{r(z)^i\}$ , and  $|r(z)| < 1$  when  $\text{Re}(z) < 0$  while  $|r(z)| > 1$  when  $\text{Re}(z) > 0$ . However, these methods are not  $R(\beta)$ -stable since  $|r(z)| \rightarrow \infty$  when  $z \rightarrow 2$ .

*Example 3.27.* The following one-step methods have growth factor

$$(3.28) \quad r(z) = \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}$$

which is the Padé (2.2) approximant to the exponential  $e^z$ :

(i) A one-step Obrechhoff formula:

$$y_{i+1}^h = y_i^h + \frac{1}{2} h(y'_{i+1} + y'_i) - \frac{1}{12} h^2(y''_{i+1} - y''_i).$$

(ii) The two-stage implicit Runge-Kutta method using Gaussian quadrature points:

$$q_1 = y_i^h + h \left[ \frac{1}{4} f\left(t_i + \left(\frac{1}{2} - \frac{\sqrt{3}}{6}\right) h, q_1\right) + \left(\frac{1}{4} - \frac{\sqrt{3}}{6}\right) f\left(t_i + \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right) h, q_2\right) \right],$$

$$q_2 = y_i^h + h \left[ \left(\frac{1}{4} + \frac{\sqrt{3}}{6}\right) f\left(t_i + \left(\frac{1}{2} - \frac{\sqrt{3}}{6}\right) h, q_1\right) + \frac{1}{4} f\left(t_i + \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right) h, q_2\right) \right],$$

$$y_{i+1}^h = y_i^h + \frac{1}{2} h \left[ f\left(t_i + \left(\frac{1}{2} - \frac{\sqrt{3}}{6}\right) h, q_1\right) + f\left(t_i + \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right) h, q_2\right) \right].$$

(iii) The three-stage implicit Runge-Kutta method using Lobatto quadrature points [10]:

$$q_2 = y_i^h + h \left[ \frac{5}{24} y'_i + \frac{1}{3} f\left(t_i + \frac{1}{2} h, q_2\right) - \frac{1}{24} y'_{i+1} \right]$$

$$= \frac{1}{2} (y_{i+1}^h + y_i^h) - \frac{1}{8} h(y'_{i+1} - y'_i),$$

$$y_{i+1}^h = y_i^h + h \left[ \frac{1}{6} y'_i + \frac{2}{3} f\left(t_i + \frac{1}{2} h, q_2\right) + \frac{1}{6} y'_{i+1} \right].$$

These methods are  $A(\sigma)$ -stable with  $\alpha = \pi/2$ , and  $A(\gamma)$ -unstable with  $\gamma = \pi/2$ . In addition they are  $R(\beta)$ -stable with  $\beta = 16.9^\circ$ . In contrast to the previous example, the positive real axis is in a white finger of the order star.

These examples were all of one-step methods with only one fundamental solution. The next examples are of genuine multistep methods.

*Example 3.29.* The following two-step methods, when applied to the test equation  $y' = \lambda y$ , generate a difference equation with the same pair of basis solutions  $(r_1^h)^i, (r_2^h)^i$ :

(i) A two-step Obrechhoff method (Enright's two-step second derivative method) [9]:

$$y_{i+1}^h = y_i^h + h \left( \frac{29}{48} y'_{i+1} + \frac{5}{12} y'_i - \frac{1}{48} y'_{i-1} \right) - \frac{1}{8} h^2 y''_{i+1}.$$

(ii) A hybrid implicit stiffly stable method [7]:

$$\begin{aligned} q_2 &= \frac{21}{32} y_{i+1}^h + \frac{3}{8} y_i^h - \frac{1}{32} y_{i-1}^h - \frac{3}{16} h y_{i+1}' \\ &= \frac{33}{32} y_i^h - \frac{1}{32} y_{i-1}^h + h \left[ \frac{7}{64} y_i' + \frac{7}{16} f \left( t_i + \frac{1}{2} h, q_2 \right) - \frac{5}{64} y_{i+1}' \right], \\ y_{i+1}^h &= y_i^h + h \left[ \frac{1}{6} y_i' + \frac{2}{3} f \left( t_i + \frac{1}{2} h, q_2 \right) + \frac{1}{6} y_{i+1}' \right]. \end{aligned}$$

For both methods

(3.30a)

$$r_1^h = \left[ 1 + \frac{5}{12} h\lambda + \left( 1 + \frac{3}{4} h\lambda + \frac{43}{192} h^2 \lambda^2 - \frac{1}{96} h^3 \lambda^3 \right)^{1/2} \right] / \left[ 2 \left( 1 - \frac{29}{48} h\lambda + \frac{1}{8} h^2 \lambda^2 \right) \right],$$

(3.30b)

$$r_2^h = \left[ 1 + \frac{5}{12} h\lambda - \left( 1 + \frac{3}{4} h\lambda + \frac{43}{192} h^2 \lambda^2 - \frac{1}{96} h^3 \lambda^3 \right)^{1/2} \right] / \left[ 2 \left( 1 - \frac{29}{48} h\lambda + \frac{1}{8} h^2 \lambda^2 \right) \right].$$

These methods are therefore  $A(\alpha)$ -stable with  $\alpha = \pi/2$ , and  $R(\beta)$ -stable with  $\beta = 17.4^\circ$  as may be seen by determining the order star. However, they are not  $A(\gamma)$ -unstable since  $r_1^h, r_2^h \rightarrow 0$  when  $h\lambda \rightarrow \infty$ . If they are used with too large a step size  $h$  then all the basis solutions of the difference equation will decay, while some of those for the original differential equation will grow. Thus, for the global difference equation, the region of di-stability is bounded in the positive half of the complex plane, and the step size will be limited for stability reasons, as is that for the central difference method (3.2) when the coefficient  $q > 0$ .

*Example 3.31.* Consider the three-step Obrechhoff method

$$y_{i+1}^h = y_i^h + h \left( \frac{79}{160} y_{i+1}' + \frac{247}{480} y_i' - \frac{1}{96} y_{i-1}' + \frac{1}{480} y_{i-2}' \right) - \frac{19}{240} h^2 (y_{i+1}'' - y_i'').$$

When applied to the test equation  $y' = \lambda y$ , this method generates three fundamental solutions  $\{(r_1^h)^i\}$ ,  $\{(r_2^h)^i\}$ ,  $\{(r_3^h)^i\}$ ;  $r_1^h$  approximates  $e^{h\lambda}$  for  $|h\lambda|$  small. In fact  $r_1^h - e^{h\lambda} = O(h^6 \lambda^6)$  as  $h\lambda \rightarrow 0$ . The spurious roots  $r_2^h, r_3^h$  are both zero for  $h\lambda = 0$ . Also  $r_1^h \rightarrow 1$  and  $r_2^h, r_3^h \rightarrow 0$  as  $h\lambda \rightarrow \infty$ . It may be checked by determining the absolute stability region and order star, that the method is  $A(\alpha)$ -stable with  $\alpha = 88.3^\circ$ , and  $A(\gamma)$ -unstable with  $\gamma = \pi/2$ . It is also  $R(\beta)$ -stable with  $\beta = \pi/12$  ( $15^\circ$ ) the largest possible value of  $\beta$  for a discretization method with fifth-order accuracy. Furthermore the spurious roots satisfy  $|r_1^h| < 1$ ,  $|r_2^h| < 1$  for  $\text{Re}(h\lambda) \geq 0$  (and indeed for all values of  $h\lambda$ ), and so the global multiple shooting equations have an unbounded region of di-stability containing  $R^- \cup R^+$  where

$$R^- := \{z \in \mathbb{C} \mid -\text{Re}(z) \sin 88.3^\circ > |\text{Im}(z)| \cos 88.3^\circ\},$$

$$R^+ := \{z \in \mathbb{C} \mid \text{Re}(z) > 0\}.$$

With this method, the step size will not be limited for stability reasons, unless the Jacobian matrix  $J$  has eigenvalues very close (within  $2^\circ$ ) to the imaginary axis, or unless stability is disturbed by either the variation of  $J$ , or by boundary conditions which do not actually control the modes of the discrete solution.

Very few (if any) multistep methods of this type are to be found in the literature. A paper has been published [8], which specifies precise families of such methods; we hope to use these for solving boundary value problems in the future.

**4. Examples.** This initial paper introduces the concept of dichotomic stability, and its importance for numerical methods for the solution of boundary value problems. It is not possible, in a limited space, to give an exhaustive analysis of the implications of the concept, or a complete description of a multiple shooting method which takes this analysis into account. For other papers that extend the analysis and implementation aspects, see, e.g., [8].

Nevertheless, this paper would not be complete without an example to illustrate (i) the necessity, for certain “stiff” problems, of a large region of di-stability, and (ii) the failure of other methods, e.g.,  $A$ -stable methods, for such problems. Implicitly, the importance of dichotomic stability for global methods (in particular collocation) was demonstrated in [1]. A shooting approach is considered here. In subsequent papers, details are given as to how an integrator should select appropriate step sizes, which are commensurate with the activity of a desired particular solution. Here, it is assumed that a suitable constant step size can be found, and bounds upon its value are considered.

*Example 4.1.* Consider the following pair of ordinary differential equations:

$$(4.2a) \quad y' = \begin{bmatrix} 0 & \lambda \\ \lambda & 0 \end{bmatrix} y + \begin{bmatrix} (1-\lambda) e^t \\ (1-\lambda) e^t \end{bmatrix}$$

(where  $\lambda$  is a real positive constant), with the boundary conditions

$$(4.2b) \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} y(0) + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad y(1) = \begin{bmatrix} 1 \\ e \end{bmatrix}.$$

The solution of (4.2) is

$$y(t) = (e^t, e^t)^T,$$

while the basis solutions of the homogeneous part form a fundamental set

$$\Phi(t) = \begin{bmatrix} e^{-\lambda t} & e^{\lambda t} \\ -e^{-\lambda t} & e^{\lambda t} \end{bmatrix}.$$

Problem (4.2) is well conditioned, with a condition number independent of  $\lambda$  (cf. [18]), and it might be expected that for any given tolerance (TOL), there should exist a maximum step size  $h$  (independent of  $\lambda$ ) such that the global error is bounded by TOL.

Consider the use of the Backward Euler method, and a large value of  $\lambda$ . Since the step size  $h$  should not depend upon  $\lambda$ , the product  $\lambda h$  may also be very large. The discretization of (4.2a) is

$$(4.3) \quad y_{i+1}^h = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & \lambda h \\ \lambda h & 0 \end{bmatrix} \right\}^{-1} [y_i^h + h f_{i+1}^h]$$

where

$$f_i^h = \begin{bmatrix} (1-\lambda) e^{ik} \\ (1-\lambda) e^{ik} \end{bmatrix},$$

and if  $\lambda h$  is large the discrete fundamental solution set is approximately

$$(4.4) \quad \Phi_i^k \approx \begin{bmatrix} (\lambda h)^{-i} & (-\lambda h)^{-i} \\ -(\lambda h)^{-i} & (-\lambda h)^{-i} \end{bmatrix}.$$

One of the difficulties of shooting is clearly absent, as there is no growth of either of the discrete fundamental modes, which might swamp other modes or the particular



solution. However, at  $t = ih = 1$ , and as  $\lambda \rightarrow \infty$ ,  $\|\Phi_i^h\| \rightarrow 0$ , and it can be seen that the second of the boundary conditions (4.2b) fails to control either of the discrete fundamental modes, unless the step size  $h$  is also reduced so that  $\lambda h$  remains of order unity. For large values of  $\lambda h$ , the discrete problem is ill conditioned, and so the step size  $h$  is highly dependent upon  $\lambda$ .

In this example, the eigenvalues of the Jacobian matrix  $J$  were  $\pm\lambda$ . Methods with  $A$ -stability, such as the Backward Euler method, are very useful for stiff initial value problems, but their treatment of eigenvalues with positive real parts is very bad. This is also true of any one-step method whose growth factor  $r(z)$  is a Padé approximant to the exponential  $e^z$  with the degree of the numerator smaller than that of the denominator. For large positive eigenvalues  $\lambda$ , lack of di-stability forces the step size down in order to recover a well-conditioned discrete problem, in much the same way as lack of stiff stability does for problems with large negative eigenvalues.

*Example 4.5.* Consider the same continuous problem (4.2); but now it is discretized with the trapezoidal rule (3.19) which is di-stable on the whole of the complex plane  $\mathbb{C}$ . The discretization is

$$(4.6) \quad y_{i+1}^h = \left[ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 & \lambda h \\ \lambda h & 0 \end{bmatrix} \right]^{-1} \left\{ \left[ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 & \lambda h \\ \lambda h & 0 \end{bmatrix} \right] y_i^h + h g_i^h \right\}$$

where  $g_i^h = \frac{1}{2}(f_i^h + f_{i+1}^h)$ , and if  $\lambda h$  is large the discrete fundamental solution set is approximately

$$(4.7) \quad \Phi_i^h \approx \begin{bmatrix} (-1)^i & (-1)^i \\ -(-1)^i & (-1)^i \end{bmatrix}.$$

For such large values of  $\lambda h$ , these discrete fundamental modes do not provide a good approximation to the continuous fundamental modes, but they are good enough for the boundary conditions to control them. It is straightforward to see that the resulting discrete problem is well conditioned. Given any tolerance (TOL), the maximum step size  $h$ , such that the global error is bounded by TOL, depends only upon the particular solution of (4.2), and not upon  $\lambda$ . Furthermore, there is no difficulty associated with the solution of the discrete problem by simple shooting.

It may be noted that dichotomic stability is not a sufficient condition to guarantee well conditioning of the discrete problem although it is a necessary condition, at least when the eigenvalues  $\lambda$  of the Jacobian matrix become large, whether with positive or negative real parts. If boundary conditions (4.2b) were replaced with

$$(4.8) \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} y(0) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} y(1) = \begin{bmatrix} 1 \\ e \end{bmatrix},$$

both the solution, and the conditioning of the continuous problem would remain unchanged. However, the boundary conditions (4.8) would not actually control the discrete fundamental modes of (4.7), and the resulting discrete problem would be ill conditioned, in spite of the di-stability of the trapezoidal rule. The difference between (4.2b) and (4.8) is similar to that occurring with the central difference scheme when the number  $N$  of subintervals is changed from an odd to an even value.

**5. Conclusion.** A problem with fast growing and decaying fundamental modes may be, if appropriate boundary conditions are given, a very well-posed problem, and this would be reflected in a small condition number [18]. If a discretization is used which does not approximate the fundamental modes of the continuous problem by

similar (growing or decaying) modes of the discrete problem, then boundary conditions that control the continuous modes cannot control the discrete modes. The resulting discrete problem would then be ill conditioned, with a large condition number [11], and small discretization errors would give rise to large global errors. Even if care is taken in the boundary layers, with accurate approximations obtained by using sufficiently small steps, the use of larger steps outside the boundary layers could change the nature of some fundamental modes, from growing to decaying, or vice versa. The resulting discrete problem would not have the correct dichotomy, and large errors could result.

Dichotomic stability, as defined in Definition 2.8 or 2.10, is a property of the global discretization method for a boundary value problem. It guarantees that the fundamental modes of the continuous problem are approximated by the proper type (growing or decaying) of mode for the discrete problem. It appears to be a necessary condition to ensure that the conditioning of the discrete problem is not worse than that of the continuous problem. It is not, in itself, a sufficient condition, without any consideration of the boundary conditions. It is not impossible for a di-stable discretization to produce an ill-conditioned discrete problem from a well-conditioned continuous problem. This can happen if the fundamental modes, while having the correct type (growth or decay) of behaviour, are nevertheless distorted (in the  $n$ -space of the dependent variables  $y$ ) in such a way that the BCs, while being correct in number, do not actually control the discrete modes at the correct end of the interval.

In the context of marching, or multiple shooting, type methods, the requirement for dichotomic stability of the global discretization implies the same property for the basic discretization scheme, which is used in a marching mode to solve IVPs.  $A(\alpha)$ -stability and  $A(\gamma)$ -instability (3.25) then jointly contribute toward an unbounded region of distability, but are not of themselves sufficient, since  $A(\gamma)$ -instability does not place any condition upon the spurious modes (if any) of the basic, discretization.  $R(\beta)$ -stability is also defined (3.24), but is of interest principally for IVPs, and does not contribute towards the di-stability of a discretization for a BVP.

**Acknowledgments.** The authors thank the Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, and Mathematisch Instituut Katholieke Universiteit Nijmegen, whose funding made possible the contacts which led to this paper. They are also grateful to ZWO, which funded a longer visit by the first author to the Netherlands, thus enabling completion of the work.

#### REFERENCES

- [1] U. ASCHER AND R. WEISS, *Collocation for singular perturbation problems I*, SIAM J. Numer. Anal., 20 (1983), pp. 537–557.
- [2] C. DE BOOR, F. DE HOOG, AND H. KELLER, *The stability of one step schemes for first-order two-point boundary value problems*, SIAM J. Numer. Anal., 20 (1983), pp. 1139–1146.
- [3] J. C. BUTCHER, *Runge–Kutta methods*, in *Modern Numerical Methods for Ordinary Differential Equations*, G. Hall and J. M. Watt eds., Clarendon Press, Oxford, 1976, Chaps. 5, 10.
- [4] W. A. COPPEL, *Dichotomies in stability theory*, Lecture Notes in Mathematics 629, Springer-Verlag, Berlin, 1978.
- [5] G. G. DAHLQUIST, *A special stability problem for linear multistep methods*, BIT, 3 (1963), pp. 17–25.
- [6] R. ENGLAND, *A program for the solution of boundary value problems for systems of ordinary differential equations*, Culham Laboratory Report PDN 3/73, Culham, U.K., 1976.
- [7] ———, *Some hybrid implicit stiffly stable methods for ordinary differential equations*, Numerical Analysis Proc., Cocoyoc, Morelos, J. P. Hennart, ed.; Lecture Notes in Mathematics 909, Springer-Verlag, Berlin, 1982, pp. 147–158.

- [8] R. ENGLAND AND R. M. M. MATTHEIJ, *Discretizations with dichotomic stability for two-point boundary value problems*, in Numerical Boundary Value ODE's, U. M. Ascher and R. D. Russell, eds., Series Progress in Scientific Computing, Vol. 5, Birkhäuser, Boston (1985), pp. 91-106.
- [9] W. H. ENRIGHT, *Second derivative multistep methods for stiff ordinary differential equations*, SIAM J. Numer. Anal., 11 (1974), pp. 321-331.
- [10] J. P. HENNART AND R. ENGLAND, *A comparison between several piecewise continuous one step integration techniques*, Working Papers for the 1979 SIGNUM Meeting on Numerical Ordinary Differential Equations, Champaign, IL, R. D. Skeel, ed.; Report 963, Dept. of Computer Science, Univ. of Illinois at Urbana-Champaign, 1979, pp. 33.1-33.4.
- [11] F. R. DE HOOG AND R. M. M. MATTHEIJ, *On dichotomy and well-conditioning in BVP*, SIAM J. Numer. Anal., 24 (1987), pp. 89-105.
- [12] H. B. KELLER, *Numerical Solution of Two Point BVP*, CBMS-NSF Regional Conference Series in Applied Mathematics 24, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1976.
- [13] B. KREISS AND H.-O. KREISS, *Numerical methods for singular perturbation problems*, SIAM J. Numer. Anal., 18 (1981), pp. 262-276.
- [14] J. P. LAMBERT, *Computational Methods in Ordinary Differential Equations*, John Wiley, London, 1973.
- [15] M. LENTINI, M. R. OSBORNE, AND R. D. RUSSELL, *The close relationships between methods for solving two-point boundary value problems*, SIAM J. Numer. Anal., 22 (1985), pp. 280-309.
- [16] M. LENTINI AND V. PEREYRA, "PASVA4"; *an ODE boundary solver for problems with discontinuous interfaces and algebraic parameters*, Matematica Apl. Comput., 2 (1983), pp. 103-118.
- [17] R. M. M. MATTHEIJ, *Estimates for the errors in the solution of linear boundary value problems, due to perturbation*, Computing, 27 (1981), pp. 299-318.
- [18] ———, *The conditioning of linear boundary value problems*, SIAM J. Numer. Anal., 19 (1982), pp. 963-978.
- [19] R. M. M. MATTHEIJ AND G. W. M. STAARINK, *An efficient algorithm for solving general linear two point BVP*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 745-763.
- [20] R. M. M. MATTHEIJ, *Decoupling and stability of BVP algorithms*, SIAM Rev., 27 (1985), pp. 1-44.
- [21] R. E. O'MALLEY, JR. AND J. E. FLAHERTY, *Analytical and numerical methods for nonlinear singular singularly perturbed initial value problems*, SIAM J. Appl. Math., 38 (1980), pp. 225-248.
- [22] R. D. RUSSELL, *A comparison of collocation and finite differences for two-point BVP*, SIAM J. Numer. Anal., 14 (1977), pp. 19-39.
- [23] R. J. SACKER AND G. R. SELL, *Singular perturbations and conditional stability*, J. Math. Anal. Appl., 76 (1980), pp. 406-431.
- [24] L. F. SHAMPINE AND G. W. GEAR, *A user's view of solving stiff ordinary differential equations*, SIAM Rev., 21 (1979), pp. 1-17.
- [25] M. VAN VELDHIJZEN, *Some results on symmetric discretizations for a stiff model problem*, Report 208, Vrije Univ., Amsterdam, 1982.
- [26] G. WANNER, E. HAIRER, AND S. P. NØRSETT, *Order stars and stability theorems*, BIT, 18 (1978), pp. 475-489.
- [27] R. WEISS, *The application of implicit Runge-Kutta and collocation methods to boundary value problems*, Math. Comp., 28 (1974), pp. 449-464.
- [28] ———, *An analysis of the box and trapezoidal schemes for linear singularly perturbed boundary value problems*, Math. Comp., 42 (1984), pp. 41-67.
- [29] O. B. WIDLUND, *A note on unconditionally stable linear multistep methods*, BIT, 7 (1967), pp. 65-70.