

*Dedicated to the memory of our colleague  
Josef Schneid (1955–2005)*

Winfried Auzinger, Wolfgang Herfort

## A UNIFORM QUANTITATIVE STIFF STABILITY ESTIMATE FOR BDF SCHEMES

**Abstract.** The concepts of stability regions,  $A$ - and  $A(\alpha)$ -stability – albeit based on scalar models – turned out to be essential for the identification of implicit methods suitable for the integration of stiff ODEs. However, for multistep methods, knowledge of the stability region provides no information on the quantitative stability behavior of the scheme. In this paper we fill this gap for the important class of Backward Differentiation Formulas (BDF). Quantitative stability bounds are derived which are uniformly valid in the stability region of the method. Our analysis is based on a study of the separation of the characteristic roots and a special similarity decomposition of the associated companion matrix.

**Keywords:** BDF schemes, stiff ODEs, stability, companion matrix, univalence.

**Mathematics Subject Classification:** Primary 65L06, 65L20; Secondary 15A21, 30C35.

### 1. INTRODUCTION

We consider linear multistep methods applied to the stiff test equation

$$y' = \lambda y, \quad \text{with } \operatorname{Re} \lambda < 0. \quad (1.1)$$

Let  $\Delta t$  denote the stepsize, assumed to be constant,  $t_\nu := \nu \Delta t$ , and

$$\mu := \lambda \Delta t. \quad (1.2)$$

A  $k$ -step linear multistep method is characterized by its coefficients  $\alpha_j$  and  $\beta_j$ ,  $j = 0 \dots k$ .

Starting from given  $k$  initial values  $y_0, \dots, y_{k-1}$ , applying the method to (1.1) generates a sequence of approximations

$$y_\nu \approx y(t_\nu), \quad \nu = k, k+1, \dots \quad (1.3)$$

defined via the difference equation

$$(\alpha_k - \mu\beta_k)y_\nu + (\alpha_{k-1} - \mu\beta_{k-1})y_{\nu-1} + \dots + (\alpha_0 - \mu\beta_0)y_{\nu-k} = 0. \quad (1.4)$$

The stability behavior of the scheme is determined by the roots of the so-called characteristic equation

$$\rho(\zeta) - \mu\sigma(\zeta) = 0, \quad \text{with} \quad \rho(\zeta) := \sum_{j=0}^k \alpha_j \zeta^j, \quad \sigma(\zeta) := \sum_{j=0}^k \beta_j \zeta^j, \quad (1.5)$$

cf. e.g. [5].

Implicit multistep methods of Backward Differentiation type (BDF) play a prominent role in the numerical integration of stiff ODEs. BDF uses only one (implicit) evaluation of the right hand side of the ODE. Application to model equation (1.1) yields the difference scheme

$$\sum_{j=1}^k \frac{1}{j} (\nabla^j y)_\nu = \mu y_\nu, \quad (1.6)$$

where  $\nabla^j$  denotes the  $j$ -th backward difference. The associated characteristic polynomials  $\rho$  and  $\sigma$  are given by

$$\rho(\zeta) = \sum_{j=1}^k \frac{1}{j} \zeta^{k-j} (\zeta - 1)^j, \quad \sigma(\zeta) = \zeta^k, \quad (1.7)$$

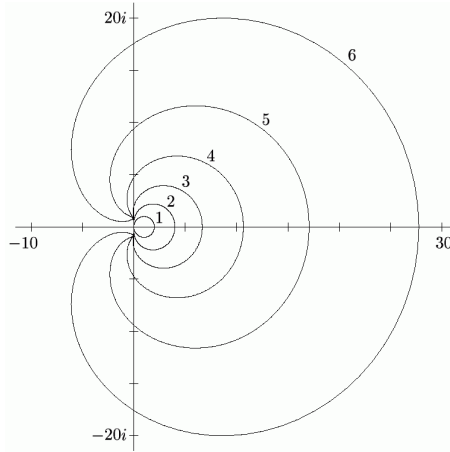
see [4]. For orders up to  $k = 6$  the BDF schemes are well known to be  $A(\alpha)$ -stable, and their convergence has been studied in detail for various classes of stiff problems.

In this paper we take a closer look at the stability properties of BDF in the stiff case. For the scalar model (1.1) the solution of the difference equation (1.4) remains uniformly bounded for all  $\mu \in S_k$  where  $S_k$  denotes the stability region of the scheme. (The BDF stability regions are the sets outside the closed curves shown in Fig. 1.) A proof of uniform boundedness can be found in [5]. It applies to general linear multistep schemes and is valid for the case that  $S_k$  is a compact subset of  $\bar{C}$ . However, it does not yield an explicit bound.

The point is that, by definition of the stability region  $S_k$ , for  $\mu \in S_k$  all roots of (1.5) (characteristic roots) are contained in the closed unit disk and all multiple roots<sup>1)</sup> are located in the open unit disk. However, the precise behavior of the sequence  $(y_\nu)$  depends on the particular distribution of these roots. If, for instance, for some  $\mu \in S_k$  a cluster of roots occurs on or close to the boundary of the unit disk, then the discretization error  $y_\nu - y(t_\nu)$  might grow strongly in a certain transient phase.

Note that the stability results for the model problem play also a role in the analysis of more general classes of stiff problems, e.g. for nonautonomous stiff systems and for nonlinear equations of singular perturbation type, see [5] and [6].

<sup>1)</sup> Polynomial roots are counted according to their multiplicity.



**Fig. 1.** Stability regions  $S_k$  of BDF schemes for  $k = 1 \dots 6$

For obtaining *quantitative* stability bounds in our situation, uniform estimates for the location of the characteristic roots are required. In the sequel we perform such an analysis for the BDF methods of orders<sup>2)</sup>  $k = 2 \dots 5$ . In Section 2 we study the distribution of these roots, in particular how they are separated. This analysis is based on a study of the Riemann surface of the mapping defined by the solutions of a reduced bivariate algebraic equation of degree  $k - 1$ . In this way, explicit but very cumbersome estimation of the characteristic roots can be avoided.

The results obtained are used in Section 3 to derive quantitative stability bounds. To this end, we use a special similarity decomposition of the companion matrix  $C(\mu)$  of the stability function  $\rho(\cdot) - \mu \sigma(\cdot)$ , which has been introduced in [3].

Appendices A and B contain auxiliary lemmas and a visualization of our results, respectively.

## 2. LOCATION OF CHARACTERISTIC ROOTS

For the  $k$ -step BDF scheme, the characteristic equation reads<sup>3)</sup>

$$\pi(\zeta; \mu) := \rho(\zeta) - \mu \sigma(\zeta) = \zeta^k \left( \sum_{j=1}^k \frac{1}{j} \left(1 - \frac{1}{\zeta}\right)^j - \mu \right) = 0, \tag{2.1}$$

cf. (1.5),(1.7).

<sup>2)</sup>  $k = 1$  yields a one-step scheme (Backward Euler).

For  $k = 6$ , BDF is still  $A(\alpha)$ -stable, but with a very tight angle, see Figure 1 and [5], and it is rarely implemented. This case is not covered by our analysis (cf. Section 2) and will not be considered here. See, however, the remark at the end of Section 3.

For  $k \geq 7$  the BDF schemes are not even 0-stable, i.e.,  $0 \notin S_k$ .

<sup>3)</sup> Throughout the paper, we suppress the dependence of various polynomials on  $k$  to simplify notation.

It will turn out to be convenient to use transforms

$$z := \frac{1}{\zeta}, \quad \hat{z} := 1 - z. \tag{2.2}$$

Since  $\rho(0) = (-1)^k/k$  and  $\sigma(0) = 0$ , no characteristic root  $\zeta$  vanishes, and therefore (2.1) is equivalent to

$$p(z; \mu) := z^k \pi(1/z; \mu) = \sum_{j=1}^k \frac{1}{j} (1-z)^j - \mu = 0, \tag{2.3}$$

i.e.,  $p$  is a reflected version of  $\pi$ . Equivalently,

$$\hat{p}(\hat{z}; \mu) := p(1 - \hat{z}; \mu) = \sum_{j=1}^k \frac{1}{j} \hat{z}^j - \mu = 0. \tag{2.4}$$

The derivative of  $\hat{p}$  does not depend on  $\mu$ , as it is just the cyclotomic polynomial of degree  $k - 1$ ,

$$\hat{p}'(\hat{z}; \mu) \equiv \sum_{j=0}^{k-1} \hat{z}^j. \tag{2.5}$$

Thus, the roots of  $\hat{p}'$  are precisely the nontrivial  $k$ -th roots of unity

$$\omega_k^j = e^{2\pi i j/k}, \quad j = 1 \dots k - 1. \tag{2.6}$$

This observation enables us to localize the double roots of  $\pi(\zeta; \mu)$ . (Note that due to Lemma A.1,  $\pi$  cannot have roots of higher multiplicity.)

**Proposition 1 (Annuli containing simple roots only).** *For  $k = 2 \dots 5$  and arbitrary  $\mu \in S_k$ , each root of  $\pi(\zeta; \mu)$  which is contained in the annulus*

$$A_k := \{ \zeta \in \mathbb{C} : \varrho_k < |\zeta| \leq 1 \}, \quad \text{with } \varrho_k := \frac{1}{|1 - \omega_k|} < 1, \tag{2.7}$$

*is simple.*

*Proof.* Actually, the assertion is even true when  $\zeta$  merely satisfies the inequality  $|\zeta| > \varrho_k$ : For such a root, there is

$$|\zeta| > \varrho_k = \frac{1}{|1 - \omega_k|} = \max_{j=1 \dots k-1} \frac{1}{|1 - \omega_k^j|},$$

because  $\omega_k$  is at the minimum distance to 1 among all  $\omega_k^j$ ,  $j = 1 \dots k - 1$ . Therefore, due to Lemma A.1,  $\zeta$  cannot be a double root. □

For  $k = 2 \dots 5$ , the inner radii  $\varrho_k$  are given in Table 1.<sup>4)</sup>

---

<sup>4)</sup> The argument in the proof of Proposition 1 is also valid for  $k > 5$ ; however, we have  $\varrho_6 = 1$  (and  $\varrho_k > 1$  for  $k > 6$ ).

**Table 1.** Inner radius  $\varrho_k$  of  $A_k$ , for  $k = 2 \dots 5$  (see Proposition 1)

$k$	$\varrho_k$	float( $\varrho_k$ )
2	$\frac{1}{2}$	0.5000
3	$\frac{1}{\sqrt{3}}$	0.5774
4	$\frac{1}{\sqrt{2}}$	0.7071
5	$\frac{\sqrt{5+\sqrt{5}}}{\sqrt{10}}$	0.8507

Proposition 2 below, our main result, is a sharpened version of Proposition 1: We shall show that if  $\zeta$  is a root contained in  $A_k$ , then  $\zeta$  is simple and  $A_k$  contains no other root. To this end, we consider the characteristic equation (2.4) in the variable  $\hat{z} = 1 - 1/\zeta$ . In the sequel, we consider an arbitrary fixed  $\mu \in \mathbb{C}$ .

The bivariate polynomial  $\hat{p}(\hat{w}; \mu) - \hat{p}(\hat{z}; \mu)$  does not depend on  $\mu$  and can be factored as

$$\hat{p}(\hat{w}; \mu) - \hat{p}(\hat{z}; \mu) = (\hat{w} - \hat{z})\hat{q}(\hat{z}, \hat{w}), \tag{2.8}$$

with

$$\hat{q}(\hat{w}, \hat{z}) = \sum_{j=1}^k \frac{1}{j} \sum_{\ell=0}^{j-1} \hat{z}^\ell \hat{w}^{j-1-\ell}. \tag{2.9}$$

Note that  $\hat{q}(\hat{z}, \hat{w}) = 0$  if and only if  $\exists \mu \in \mathbb{C}$  with  $\hat{p}(\hat{z}; \mu) = \hat{p}(\hat{w}; \mu) = 0$  and either  $\hat{z} \neq \hat{w}$  (i.e.,  $\hat{w}$  is a companion root of  $\hat{p}$  to  $\hat{z}$ ), or  $\hat{z} = \hat{w}$  is a double root of  $\hat{p}(\hat{z}; \mu)$ . Later we shall make use of the symmetries

$$\hat{q}(\hat{z}, \hat{w}) = \hat{q}(\hat{w}, \hat{z}) \quad \text{and} \quad \hat{q}(\bar{\hat{z}}, \bar{\hat{w}}) = \overline{\hat{q}(\hat{z}, \hat{w})}. \tag{2.10}$$

For  $k = 2 \dots 5$ , one can rewrite (2.9) as follows:

—  $k = 2$ :

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \left( \frac{\hat{z} + \hat{w}}{2} \right), \tag{2.11}$$

—  $k = 3$ :

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \left( \frac{\hat{z} + \hat{w}}{2} \right) + \left( \frac{\hat{z} + \hat{w}}{2} \right)^2 + \frac{1}{12} (\hat{z} - \hat{w})^2, \tag{2.12}$$

—  $k = 4$ :

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \left(\frac{\hat{z} + \hat{w}}{2}\right) + \left(\frac{\hat{z} + \hat{w}}{2}\right)^2 + \left(\frac{\hat{z} + \hat{w}}{2}\right)^3 + \frac{1}{24}(2 + 3(\hat{z} + \hat{w}))(\hat{z} - \hat{w})^2, \quad (2.13)$$

—  $k = 5$ :

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \left(\frac{\hat{z} + \hat{w}}{2}\right) + \left(\frac{\hat{z} + \hat{w}}{2}\right)^2 + \left(\frac{\hat{z} + \hat{w}}{2}\right)^3 + \left(\frac{\hat{z} + \hat{w}}{2}\right)^4 + \frac{1}{240}(20 + 30(\hat{z} + \hat{w}) + 3(11\hat{z}^2 + 18\hat{z}\hat{w} + 11\hat{w}^2))(\hat{z} - \hat{w})^2. \quad (2.14)$$

Thus, for  $k = 2 \dots 5$ , equation (2.9) can be written as

$$\hat{q}(\hat{w}; \hat{z}) \equiv \sum_{j=1}^k \left(\frac{\hat{z} + \hat{w}}{2}\right)^{j-1} + \gamma_{k-3}(\hat{w}; \hat{z})(\hat{z} - \hat{w})^2, \quad (2.15)$$

where  $\gamma_{k-3}(\hat{w}; \hat{z}) = \gamma_{k-3}(\hat{z}; \hat{w})$  is a polynomial in  $\hat{w}$  and  $\hat{z}$ .

We are now ready to formulate our main result.

**Proposition 2 (Annulus  $A_k$  contains at most a simple, solitary root).** *For  $k = 2 \dots 5$  and arbitrary  $\mu \in S_k$ , any root of  $\pi(\zeta; \mu)$  contained in the annulus  $A_k$  defined in Proposition 1, (2.7) is simple and solitary within  $A_k$ , i.e.,  $A_k$  contains no other root.*

*Proof.* Fix  $k \in \{2, \dots, 5\}$  and any root  $\zeta$  of  $\pi(\zeta; \mu)$  satisfying  $|\zeta| > \varrho_k$ . By Proposition 1,  $\zeta$  is a simple root.

It is now convenient to use  $z = 1/\zeta = 1 - \hat{z}$  as an independent variable. Our assumption  $\zeta \in A_k$  implies that  $z := 1/\zeta$  is a root of  $p(z; \mu) = z^k \pi(1/z; \mu)$  (cf. (2.3)) satisfying  $|z| > 1/\varrho_k =: r_k > 1$ , see Table 2.

**Table 2.** Radius  $r_k = 1/\varrho_k$  of  $K_k$ , for  $k = 2 \dots 5$  (see proof of Proposition 2)

$k$	$r_k$	float( $r_k$ )
2	2	2.0000
3	$\sqrt{3}$	1.7320
4	$\sqrt{2}$	1.4142
5	$\frac{\sqrt{10}}{\sqrt{5+\sqrt{5}}}$	1.1756

Now it suffices to check that for  $z$  inside the open disk

$$K_k := \{z \in \mathbb{C} : |z| < r_k\}, \tag{2.16}$$

any solution of

$$q(z, w) := \hat{q}(1 - w, 1 - z) = 0 \tag{2.17}$$

lies outside  $K_k$ .

For the bivariate polynomial  $q(z, w)$ , the following observations hold:

- $q(z, w)$  is monic in  $w$  of degree  $k - 1$ , and since  $q$  has real coefficients, we have  $q(z, w) = 0 \Leftrightarrow q(\bar{z}, \bar{w}) = 0$ .
- Solving the equation  $q(0, w) = 0$  yields  $k - 1$  distinct points  $w \notin \bar{K}_k$  (cf. Fig. 2, see Appendix B).
- Lemma A.2 shows that the discriminant set

$$D_k := \{z \in \mathbb{C} : \exists w \in \mathbb{C} : q(z, w) = q_w(z, w) = 0\} \tag{2.18}$$

is finite and disjoint from  $K_k$ .

- For any  $z_0 \notin D_k$ , Lemma A.4 implies the existence of an analytic function element  $(w(z), K_R, z_0, R)$  with  $K_R(z_0) = \{z : |z - z_0| < R\}$  and  $R$  the distance from  $z_0$  to  $D_k$ , i.e., a power series  $w(z) = \sum_{j=0}^{\infty} a_j(z - z_0)^j$  convergent on  $K_R(z_0)$ .

Following the description in [7, Vol. III, p. 308] one may construct a Riemann surface by considering  $k - 1$  replicas of the extended  $z$ -plane (the Riemann sphere), in each of them perform cuts from the point  $\infty$  to the points in  $D_k$  and consider a certain gluing along the cuts. This at hand, analytic continuation along paths on the surface gives rise to a  $(k - 1)$ -valued function  $w(z) = (w_1(z), \dots, w_{k-1}(z))$  on the Riemann surface with values in  $\mathbb{C}$  satisfying  $q(z, w_j(z)) = 0$ .

In our situation we can assume each of the  $k - 1$  cuts to be a half-ray emanating from a point in  $D_k$  parallel to the  $\text{Re } z$ -axis to the right, because they connect to  $\infty$  in the extended plane without intersecting each other.

- Lemma A.4 implies that on  $K_k$  one can define a function element  $(f_0(z), K_k, 0, r_k)$  by fixing its value  $f_0(0) \notin \bar{K}_k$ . The special form of  $q(z, w)$  shows that poles of  $f_0$  cannot occur, since  $w = \infty$  implies  $z = \infty$ . Then,  $f_0$  has an analytic continuation to the selected sheet (with cuts). Since  $r_k$  is smaller than the distance from  $0$  to  $D_k$  and no poles can occur within  $K_R(0)$ , we conclude that  $f_0$  has a continuous extension, say  $f$ , to the closure  $\bar{K}_k$ .
- To prove that the restriction of  $f$  to  $\partial K_k$  is injective, we consider arbitrary points  $z, \tilde{z} \in \partial K_k$  and  $w \in \mathbb{C}$  with  $f(z) = f(\tilde{z}) = w$ . Then,  $q(z, w) = q(\tilde{z}, w) = 0$ . Now suppose  $\tilde{z} \neq z$  and assume w.l.o.g.  $w \neq z$  (else exchange the roles of  $z$  and  $\tilde{z}$ ). Together with the definition of  $q(\cdot, \cdot)$  this implies

$$\frac{z - \tilde{z}}{z - w} q(z, \tilde{z}) \equiv q(z, w) - \frac{\tilde{z} - w}{z - w} q(\tilde{z}, w) = 0, \tag{2.19}$$

hence  $q(z, \tilde{z}) = 0$ , and since  $|z| = |\tilde{z}| = r_k$ , Lemma A.3 yields  $z = \tilde{z}$ , a contradiction. Thus,  $f$  restricted to  $\partial K_k$  is indeed injective.

By the Theorem of Darboux–Picard, it follows that  $f$  is univalent on  $\bar{K}_k$  and in fact maps  $K_k$  to the interior of  $f(K_k)$  (see [2, p. 310, Corollary 9.16]).

— In order to show that

$$K_k \cap f(K_k) = \emptyset \quad (2.20)$$

holds, we first consider  $z_0 \in \partial K_k \cap f(\partial K_k)$ . Then,  $q(z_0, f(z_0)) = 0$  and  $|z_0| = |f(z_0)| = r_k$ . Therefore, Lemma A.3 implies  $z_0 = f(z_0) \in \{1 - \omega_k, 1 - \bar{\omega}_k\}$  is a double root of  $p$ . As can be seen from the proof of Lemma A.1,  $z_0$  indeed has multiplicity 2, thus

$$0 \neq p''(z_0) \equiv \frac{d}{dz} q(z, f(z)) \Big|_{z=z_0} = q_z(z_0, z_0) + q_w(z_0, z_0) \cdot 1 = 2q_z(z_0, z_0), \quad (2.21)$$

and together with  $0 \equiv \frac{d}{dz} q(z, f(z)) = q_z(z, f(z)) + q_w(z, f(z))f'(z)$  this implies  $f'(z_0) = -1$ . This means that the curve  $f(\partial K_k)$  is differentiable and tangent to  $\partial K_k$  at  $z_0$ .

Hence  $f$  certainly maps points from  $K_k$  to its complement. Since  $\partial K_k \cap f(\partial K_k)$  consists of at most two points,  $f(\partial K_k)$  lies in the complement of  $K_k$ . As noted above, the interior of  $f(\bar{K}_k)$  has boundary  $f(\partial K_k)$  and thus, (2.20) indeed holds.

This completes the proof of Proposition 2.  $\square$

It can also be shown that for  $k = 3, 4, 5$ , two different algebraic solution branches  $f(z)$ ,  $\tilde{f}(z)$  of equation  $q(z, w) = 0$  satisfy

$$f(\bar{K}_k) \cap \tilde{f}(\bar{K}_k) = \emptyset, \quad (2.22)$$

see Figure 2. Since  $f$  and  $\tilde{f}$  are both analytic on a domain containing  $\bar{K}_k$  and univalent on  $\bar{K}_k$ , (2.22) will follow from  $f(\partial K_k) \cap \tilde{f}(\partial K_k) = \emptyset$ .

Suppose there exists  $w \in f(\partial K_k) \cap \tilde{f}(\partial K_k)$ . Then there are  $z, \tilde{z} \in \partial K_k$  with  $w = f(z) = \tilde{f}(\tilde{z})$ , hence  $q(z, w) = q(\tilde{z}, w) = 0$ . The same argument as in the above proof for injectivity (cf. (2.19)) shows that  $\tilde{z} = z$ , where  $z \in \{1 - \omega_k, 1 - \bar{\omega}_k\}$  is a double root of  $p$ . Now there are two possibilities:

- (i) If  $w = z$ , then we may repeat an earlier argument (cf. (2.21)) for both  $f$  and  $\tilde{f}$  and conclude

$$f'(z) = \tilde{f}'(z) = -1.$$

This means that both curves  $f(\partial K_k)$  and  $\tilde{f}(\partial K_k)$  are differentiable and tangent to  $\partial K_k$  at the point  $z$ . But this would imply that  $z$  has multiplicity  $> 2$ , which contradicts Lemma A.1.

- (ii) If  $w \neq z$ , then  $w = f(z) = \tilde{f}(z)$  implies that  $w$  is a double root of  $p$ . Thus,  $z$  must be contained in the discriminant set  $D_k$ . But this together with  $|z| = r_k$  contradicts Lemma A.2.



This proves (2.22).

Proposition 2 shows that, for  $k = 2 \dots 5$  and  $\mu \in S_k$ , there are two possibilities concerning the distribution of the characteristic roots  $\zeta_j$ ,  $j = 1 \dots k$ :

- (i) If the annulus  $A_k$  contains a root  $\zeta_k$  (we may call it the principle root), then  $\zeta_k$  is simple, and all other roots  $\zeta_j$  satisfy  $|\zeta_j| \leq \varrho_k$ ,  $j = 1 \dots k - 1$ .
- (ii) Otherwise, all roots satisfy  $|\zeta_j| \leq \varrho_k$ ,  $j = 1 \dots k$ .

In Appendix B the situation is visualised. For this purpose, the reduced equation  $q(z, w) = 0$  was solved using Maple 10 for an appropriately large sample of  $z$ -values. In Figure 2, the  $k - 1$  outer loin-shaped zones – let us call them *kloings* – in each case corresponds to the range of solutions of  $q(z, w) = 0$  for a given  $z \in K_k$ . The positions of double roots of  $p$  are marked by diamonds. The contact points between  $K_k$  and a certain kloing correspond to the special roots  $z = w$  which are identified in Lemma A.3.

Figure 3 (see Appendix B) refers to the original variables: For a given characteristic root  $\zeta$  outside the disk of radius  $\rho_k$ , the companion roots  $\eta$  are separated and located within the  $k - 1$  interior kloings.

### 3. UNIFORM STABILITY ESTIMATE

Proposition 2 enables us to derive a uniform stability estimate. To this end, it is convenient to formulate multistep method (1.4) in a common single-step fashion, namely

$$\begin{pmatrix} y_{\nu-k} \\ \vdots \\ y_{\nu} \end{pmatrix} = C(\mu) \begin{pmatrix} y_{\nu-1-k} \\ \vdots \\ y_{\nu-1} \end{pmatrix}, \tag{3.1}$$

with the companion matrix

$$C = C(\mu) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -\gamma_0 & -\gamma_1 & \dots & -\gamma_{k-2} & -\gamma_{k-1} \end{pmatrix} \in \mathbb{C}^{k \times k}. \tag{3.2}$$

Here,  $\gamma_j \equiv (\alpha_j - \mu\beta_j)/(\alpha_k - \mu\beta_k)$ , and the characteristic polynomial of  $C$  reads

$$\tilde{\pi}(\zeta) = \tilde{\pi}(\zeta; \mu) = \sum_{j=0}^k \gamma_j \zeta^j. \tag{3.3}$$

This is just the monic version of the characteristic polynomial  $\pi$  of the given multistep method (cf. (1.5)). Each matrix of the form (3.2) is nonderogatory, i.e., all eigenvalues

$\zeta_j$  of  $C$  have geometric multiplicity 1, and the algebraic multiplicity of  $\zeta_j$  equals the multiplicity of  $\zeta_j$  as a root of  $\pi$  or  $\tilde{\pi}$ , respectively.

We shall quantify the stability behavior in the BDF case by providing a uniform bound for  $\|C^\nu(\mu)\|$  where  $\nu \in \mathbb{N}$  and  $\mu \in S_k$ . The single-step formulation enables us to apply linear algebra techniques. Estimation of the powers of  $C$  can be gained in various ways. We shall derive an estimate using a special similarity decomposition of  $C(\mu)$  and applying the Kreiss Matrix Theorem (cf., e.g., [5]) to the transformed matrix. The resulting bounds are certainly not sharp but easy to obtain.<sup>5)</sup>

In [3] and the forthcoming report [1], a decomposition of the companion matrix  $C(\mu)$  is considered which, in contrast to the Jordan decomposition, depends continuously on the parameter  $\mu$ . To describe it, let us fix some notation. For an arbitrary  $\zeta \in \mathbb{C}$ , we denote

$$x(\zeta) := (1, \zeta, \zeta^2, \dots, \zeta^{k-1})^T. \tag{3.4}$$

Now, let  $\zeta_1, \dots, \zeta_k$  be given complex numbers (not necessarily distinct). In the sequel,  $f[\zeta_j, \dots, \zeta_\ell]$  denotes a — possibly confluent — divided difference of a function  $f$ , i.e.,  $f[\zeta_j] := f(\zeta_j)$ , and

$$f[\zeta_j, \dots, \zeta_\ell] := \begin{cases} \frac{f[\zeta_{j+1}, \dots, \zeta_\ell] - f[\zeta_j, \dots, \zeta_{\ell-1}]}{\zeta_\ell - \zeta_j}, & \zeta_j \neq \zeta_\ell, \\ \lim_{\epsilon \rightarrow 0} \frac{f[\zeta_{j+1}, \dots, \zeta_\ell + \epsilon] - f[\zeta_j, \dots, \zeta_{\ell-1}]}{\epsilon}, & \zeta_j = \zeta_\ell. \end{cases} \tag{3.5}$$

For the powers  $f(\zeta) = \zeta^n$  and for  $f(\zeta) = \tilde{\pi}(\zeta)$  we abbreviate

$$\zeta_{[j..l]}^n := \zeta^n[\zeta_j, \dots, \zeta_l], \quad \tilde{\pi}_{[j..l]} := \tilde{\pi}[\zeta_j, \dots, \zeta_l]. \tag{3.6}$$

**Proposition 3 (Bidiagonal-Frobenius canonical form).**

— For arbitrary  $\zeta_1, \dots, \zeta_k$ , the matrix  $C$  from (3.2) can be written as

$$C = L B L^{-1}, \tag{3.7}$$

with the lower triangular matrix

$$L = \left( \begin{array}{c|c|c|c} x[\zeta_1] & & & \\ \hline x[\zeta_1, \zeta_2] & & & \\ \hline \dots & & & \\ \hline x[\zeta_1, \dots, \zeta_k] & & & \end{array} \right) = \left( \begin{array}{cccc} 1 & & & \\ \zeta_1 & 1 & & \\ \zeta_1^2 & \zeta_{[1..2]}^2 & 1 & \\ \vdots & \vdots & \ddots & \ddots \\ \zeta_1^{k-1} & \zeta_{[1..2]}^{k-1} & \dots & \zeta_{[1..k-1]}^{k-1} & 1 \end{array} \right), \tag{3.8}$$

<sup>5)</sup> One may also apply the Kreiss Matrix Theorem directly to  $C(\mu)$ . However, the algebra involved would be somewhat more laborious. Alternatively, one may proceed as in [3], where a scaled  $\infty$ -norm is used.

and

$$B = \begin{pmatrix} \zeta_1 & 1 & & & \\ & \zeta_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \zeta_{k-1} & 1 \\ -\tilde{\pi}_{[1]} & -\tilde{\pi}_{[1..2]} & \dots & -\tilde{\pi}_{[1..k-1]} & -\tilde{\pi}_{[1..k]} + \zeta_k \end{pmatrix} \tag{3.9}$$

— When  $\zeta_j$  runs through the roots of  $\tilde{\pi}$ ,  $B$  in (3.9) takes bidiagonal shape where the main diagonal contains the  $k$  roots of  $\tilde{\pi}$  counted with multiplicity:

$$B = \begin{pmatrix} \zeta_1 & 1 & & & \\ & \zeta_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \zeta_{k-1} & 1 \\ & & & & \zeta_k \end{pmatrix}. \tag{3.10}$$

*Proof.* In [3], a proof is given for the special case (3.10). An analogous proof for (3.9) is to be found in [1]. □

The inverse of  $L$  is lower triangular with unit diagonal, and the  $(i, j)$ -th element below the diagonal can be written as

$$(L^{-1})_{i,j} = (-1)^{i+j} \sum_{i,j} \prod_{i-j} \zeta_\ell, \tag{3.11}$$

where  $\sum_{i,j}$  stands for summation of all possible  $\binom{i-1}{j-1}$  products  $\prod_{i-j} \zeta_\ell$  of  $i-j$  values  $\zeta_\ell$  with pairwise distinct indices  $1 \leq \ell < i$ .

Now we consider bidiagonal form (3.8),(3.10) for the stable BDF case  $k = 2 \dots 5$ , with  $\mu \in S_k$ . We assume that  $\zeta_k$  has maximal modulus, i.e.,  $1 \geq |\zeta_k| \geq |\zeta_j|$  for  $j = 1 \dots k-1$ . To apply the Kreiss Matrix Theorem, we compute the resolvent  $(\zeta I - B)^{-1}$ , where  $\zeta \in \mathbb{C}$  is not a characteristic root:

$$(\zeta I - B)^{-1} = \begin{pmatrix} \zeta_{1,1}^{-1} & \zeta_{1,2}^{-1} & \dots & \dots & \zeta_{1,k}^{-1} \\ & \zeta_{2,2}^{-1} & \dots & \dots & \zeta_{2,k}^{-1} \\ & & \ddots & & \vdots \\ & & & \zeta_{k-1,k-1}^{-1} & \zeta_{k-1,k}^{-1} \\ & & & & \zeta_{k,k}^{-1} \end{pmatrix}, \quad \text{with } \zeta_{i,j} := \prod_{\ell=i}^j (\zeta - \zeta_\ell). \tag{3.12}$$

Consider an arbitrary  $\zeta \in \mathbb{C}$  with  $|\zeta| > 1$ . Since for  $\mu \in S_k$  all characteristic roots  $\zeta_j$  satisfy  $|\zeta_j| \leq 1$ , one has

$$|\zeta - \zeta_j| \leq |\zeta| - 1, \quad j = 1 \dots k. \tag{3.13}$$

Moreover, due to Proposition 2, one has

$$|\zeta - \zeta_j| \leq 1 - \varrho_k, \quad j = 1 \dots k-1, \tag{3.14}$$

(for the principal root  $\zeta_k$  this is not true in general). From (3.13),(3.14) we immediately obtain bounds for the entries in (3.12):

$$\left| (\zeta I - B)_{j,j+\ell}^{-1} \right| \leq \frac{1}{(|\zeta| - 1)(1 - \varrho_k)^\ell}, \quad j = 1 \dots k, \quad \ell = 0 \dots k - j. \quad (3.15)$$

Using (3.15), one finds the norm estimate

$$\|(\zeta I - B)^{-1}\|_2 \leq \sqrt{\|(\zeta I - B)^{-1}\|_1 \|(\zeta I - B)^{-1}\|_\infty} \leq \frac{\mathcal{K}_k}{|\zeta| - 1}, \quad \forall |\zeta| > 1, \quad (3.16)$$

with

$$\mathcal{K}_k = \left( \frac{1 - \varrho_k}{\varrho_k} \right) \left( \left( \frac{1}{1 - \varrho_k} \right)^k - 1 \right). \quad (3.17)$$

Taking (3.16) into account, application of the Kreiss Matrix Theorem (cf. [5]) to the matrix  $B$  yields

$$\|B^\nu\|_2 \leq e k \mathcal{K}_k, \quad \nu = 1, 2, \dots, \quad \forall \mu \in S_k. \quad (3.18)$$

**Table 3.** Stability bounds  
for  $k = 2 \dots 5$

$k$	float( $\rho_k$ )	float( $\mathcal{C}_k$ )
2	0.5000	4.0e1
3	0.5774	4.6e2
4	0.7071	1.6e4
5	0.8507	4.5e6

Moreover, using (3.11) we can estimate the condition number  $\kappa(L) = \|L\| \|L^{-1}\|$  of the transformation matrix  $L$  from (3.8) by

$$\kappa_2(L) \leq \sqrt{\kappa_1(L) \kappa_\infty(L)} \leq (1 + \varrho_k)^{2(k-1)}. \quad (3.19)$$

Combining the above estimates, we arrive at the desired uniform stability estimate for the powers of  $C = C(\mu)$ :

$$\|C^\nu\|_2 \leq \kappa_2(L) \|B^\nu\|_2 \leq \mathcal{C}_k, \quad \nu = 1, 2, \dots, \quad \forall \mu \in S_k, \quad (3.20)$$

with

$$\mathcal{C}_k = (1 + \varrho_k)^{2(k-1)} \left( \frac{1 - \varrho_k}{\varrho_k} \right) \left( \left( \frac{1}{1 - \varrho_k} \right)^k - 1 \right). \quad (3.21)$$

Numerical values for the  $\mathcal{C}_k$  are given in Table 3.

These bounds are certainly not sharp, but they seem to indicate that, especially for  $k = 5$ , significant growth cannot be precluded; typically this might occur in mildly stiff situations, i.e., for moderate values of  $\mu$  near the imaginary axis.

It would be interesting to compare the above estimates with numerical experiments, including also the case  $k = 6$ . Furthermore, other methods of estimating the powers of  $C(\mu)$  could be investigated, which may lead to sharper estimates. In this paper, these topics are not addressed further.

Let us conclude with a remark concerning  $k = 6$ . As mentioned in Section 2, this case is not covered by our analysis, because  $\pi(\zeta; \mu)$  may have multiple roots of modulus 1. A straightforward calculation shows that these roots do actually occur. Namely,

$$\zeta = \frac{1}{2} \pm i \frac{1}{2} \sqrt{3}, \quad |\zeta| = 1, \tag{3.22}$$

is a double root of  $\pi(\zeta; \mu)$  for some  $\mu \in S_6$ . However, this occurs for

$$\mu = \frac{7}{2} \pm i \frac{63\sqrt{3}}{2}, \quad \text{with } \operatorname{Re} \mu > 0, \tag{3.23}$$

i.e., in that part of the stability region  $S_6$  where the BDF scheme behaves ‘artificially stable’ (cf. fig. 1): For  $\lambda \Delta t = \mu \in S_k$  with  $\operatorname{Re} \mu > 0$ , the roots of  $p(\zeta; \mu)$  satisfy  $|\zeta| \leq 1$ , although the ODE  $y' = \lambda y$  has exponentially growing solutions. A reasonable quantitative stability analysis for  $k = 6$  would have to observe the additional condition  $\operatorname{Re} \mu \leq 0$ .

### APPENDIX A. FOUR LEMMAS

This section contains auxiliary results used in the proofs of Propositions 1 and 2. The notation is as in Section 2; in particular, we recall the definition of the reduced equation  $q(z, w) \equiv \hat{q}(\hat{z}, \hat{w}) = 0$ , cf. (2.9), (2.17). In certain parts of the proofs we have used computer algebra (Maple 10).

**Lemma A.1 (Double roots of  $\pi$ ).** *For arbitrary  $k \geq 2$  and  $\mu \in \mathbb{C}$ , any double root of  $\pi(\zeta; \mu)$  is of the form*

$$\zeta = \frac{1}{1 - \omega_k^\ell} \quad \text{for some } \ell \in \{1, \dots, k - 1\}. \tag{A.1}$$

Moreover, there exists no  $\mu \in \mathbb{C}$  for which a root with multiplicity  $> 2$  occurs.

*Proof.* Let  $k \geq 2$  and  $\mu \in \mathbb{C}$  be fixed. Assume  $\zeta$  is a root of  $\pi(\zeta; \mu)$  of multiplicity  $\geq 2$  for some  $\mu \in \mathbb{C}$ , i.e.,  $\pi(\zeta; \mu) = \pi'(\zeta; \mu) = 0$ . Since  $\pi(0; \mu) = (-1)^k \neq 0$  and  $\pi'(1; 0) = 1$ , we know  $\zeta \notin \{0, 1\}$ . Hence we may transform the equations according to (2.2)–(2.4). Now, elementary calculation shows that  $\pi(\zeta; \mu) = \pi'(\zeta; \mu) = 0$  implies

$$\hat{p}'(\hat{z}) = -p'(z) = 0 \quad \text{for } \hat{z} = 1 - z = 1 - \zeta^{-1}.$$

Here,  $\hat{p}(\hat{z})$  is the cyclotomic polynomial (2.5) and thus,  $\zeta = 1/(1 - \hat{z})$  must be of the form (A.1).

Let  $k \geq 3$  and assume that  $\zeta$  has multiplicity  $> 2$ , i.e.,  $\pi''(\zeta; \mu) = 0$ . Elementary calculation shows that  $\pi(\zeta; \mu) = \pi'(\zeta; \mu) = \pi''(\zeta; \mu) = 0$  would imply  $\hat{p}''(\hat{z}) = p''(z) = 0$  for  $\hat{z} = 1 - \zeta^{-1}$ , which is not the case, since  $\hat{p}'$  is the cyclotomic polynomial which has simple roots only. Thus, roots of multiplicity  $> 2$  do not occur. □

**Lemma A.2 (Discriminant set of the reduced equation).** *For  $k = 3 \dots 5$ , the discriminant set*

$$D_k = \{z \in \mathbb{C}: \exists w \in \mathbb{C}: q(z, w) = q_w(z, w) = 0\}$$

is finite, and each  $z \in D_k$  satisfies  $|z| > r_k$ . The associated  $w$ -values are of the form

$$w = 1 - \omega_k^j, \quad j \in \{1, \dots, k\}. \tag{A.2}$$

*Proof.* The proof is based on a routine calculation; we show that the determination of  $D_k$  amounts to solving a certain set of polynomial equations of degree  $k - 2$ . We do not write down all details but just show how to proceed.

For  $w = z$  we have  $q(z, w) = p'(w)$ ; for  $w \neq z$  we obtain

$$q_w(z, w) = \frac{\partial}{\partial w} \frac{p(w) - p(z)}{z - w} = \frac{p'(w)(z - w) + (p(w) - p(z))}{(z - w)^2} = \frac{p'(w) + q(z, w)}{z - w}.$$

Therefore, for arbitrary  $z, w$ , one has

$$q(z, w) \equiv (z - w)q_w(z, w) - p'(w).$$

Thus,  $q(z, w) = q_w(z, w) = 0$  is equivalent to

$$q_w(z, w) = 0, \quad p'(w) = 0.$$

Now,  $p'(w) = 0$  implies that  $w$  is of the form  $1 - \omega_k^j$ ,  $j \in \{1 \dots k - 1\}$ . Substituting these values into the equation  $q_w(z, w) = 0$  and solving for  $z$ , yields the desired values for  $z$ . It turns out that all these  $z \in D_k$  satisfy  $|z| > r_k$  for  $k = 3, 4, 5$ , and none of them is of the form  $1 - \omega_k^j$ . □

Some of the  $(z, w)$  with  $z \in D_k$  are visible in Figure 2; the  $z$ -values appear as cusps of the various kloings.

**Lemma A.3 (Zeros of the reduced equation satisfying  $|z| = |w| = r_k$ ).** *For  $k \in \{2, \dots, 5\}$ , the solutions of*

$$q(z, w) = 0, \quad |z| = |w| = r_k \tag{A.3}$$

are given by<sup>6)</sup>

$$z = w = 1 - \omega_k, \quad z = w = 1 - \bar{\omega}_k. \tag{A.4}$$

*Proof.* The proof uses elimination techniques exploiting symmetry properties (2.10) of the bivariate polynomial  $q(z, w)$ . We shall present the material in a way indicating the essential steps, so that an interested reader will be able to reproduce missing technical details and calculations. For higher values of  $k$ , we used computer algebra (Maple 10) for the more laborious calculations. In particular, for the case  $k = 5$ , we include Maple code.

Let us introduce auxiliary quantities

$$d := \frac{z - w}{2} = \frac{\hat{w} - \hat{z}}{2}, \quad h := \frac{z + w}{2}, \quad \hat{h} := \frac{\hat{z} + \hat{w}}{2} = 1 - h. \tag{A.5}$$

To prove (A.4), we first note that our assumption  $|z| = |w| = r_k$  implies  $h\bar{h} \leq r_k^2$ . Moreover, there must hold  $h \neq 0$ , i.e.  $\hat{w} \neq -\hat{z}$ : Namely, substitution of  $\hat{w} = -\hat{z}$  into (2.11)–(2.14) in each case yields a polynomial equation in  $z = 1 - \hat{z}$  which is simple to solve and which has no zero satisfying  $|z| = r_k$ . Thus we know

$$h\bar{h} = s, \quad \text{with } 0 < s \leq r_k^2. \tag{A.6}$$

Furthermore,  $|z| = |w| = r_k$  together with (A.5),(A.6) imply

$$d = i\sqrt{\frac{r_k^2 - s}{s}}h. \tag{A.7}$$

The cases  $k = 2 \dots 5$  are now considered separately. The eliminations leading to (A.4) are performed in terms of the variables  $h \in \mathbb{C}$  and  $s \in \mathbb{R}$ .

- $k = 2$  (with  $r_2 = 2$ )

This case is very straightforward. The reduced equation (degree 1) is precisely the cyclotomic equation of degree 1 in the variable  $\hat{h}$  (cf. (2.11)):

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \hat{h} = 0.$$

We conclude

$$h = 1 - \hat{h} = 1 - \omega_2 = 2,$$

and thus,

$$\frac{z + w}{2} = h = 2 = r_2.$$

Now,  $|z| = |w| = r_2$  implies

$$w = z = 1 - \omega_2 = 1 - \bar{\omega}_2 = 2,$$

as asserted.

---

6) For  $k = 2$  these solutions coincide:  $1 - \omega_2 = 1 - \bar{\omega}_2 = 2$ .

- $k = 3$  (with  $r_3 = \sqrt{3}$ )

In terms of  $\hat{h}$  and  $d$ , the reduced equation (degree 2) reads (cf. (2.12))

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \hat{h} + \hat{h}^2 + \frac{d^2}{6} = 0, \quad (\text{A.8})$$

or equivalently,

$$q(z, w) = 3 - 3h + h^2 + \frac{d^2}{6} = 0.$$

Using (A.7) and the equality  $r_3^2 = 3$ , this can be rewritten as an equation in terms of  $h$  and  $s$ ,

$$3 - 3h + \frac{1}{6} \left(7 - \frac{3}{s}\right) h^2 = 0. \quad (\text{A.9})$$

Note that  $h^2 = s$  cannot hold, because otherwise (A.6) would imply that  $h = \bar{h}$  is real – but substituting  $s = h^2$  into (A.9) results in a quadratic equation for  $h$  with no real solution.

Since (A.9) is an equation for  $h$  with real coefficients, it is also valid for  $\bar{h} = s/h$  (cf. (A.6)), i.e.,

$$3 - 3 \left(\frac{s}{h}\right) + \frac{1}{6} \left(7 - \frac{3}{s}\right) \left(\frac{s}{h}\right)^2 = 0. \quad (\text{A.10})$$

The linear combination

$$s \cdot (\text{A.9}) - h^2 \cdot (\text{A.10})$$

results in

$$\frac{7}{6}(h^2 - s)(s - 3) = 0.$$

Since  $h^2 \neq s$ , we conclude  $s = r_3^2 = 3$ , hence  $d = 0$  (cf. (A.7)), and (A.8) reduces to the cyclotomic equation of degree 2 in the variable  $\hat{h}$  (cf. (2.5)):

$$1 + \hat{h} + \hat{h}^2 = 0.$$

Thus,  $h$  is necessarily of the form

$$h = 1 - \hat{h} = 1 - \omega_3^j = 1 - e^{2\pi ij/3}, \quad j \in \{1, 2\},$$

and both solutions satisfy  $h\bar{h} = s = r_3^2 = 3$ . Eventually, we obtain

$$\frac{z+w}{2} = h = 1 - e^{2\pi ij/3}, \quad j \in \{1, 2\}, \quad \left| \frac{z+w}{2} \right| = \sqrt{3} = r_3.$$

Now,  $|z| = |w| = r_3$  implies

$$w = z = 1 - \omega_3, \quad \text{or} \quad w = z = 1 - \omega_3^2 = 1 - \bar{\omega}_3,$$

as asserted.



- $k = 4$  (with  $r_4 = \sqrt{2}$ )

In terms of  $\hat{h}$  and  $d$ , the reduced equation (degree 3) reads (cf. (2.13))

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \hat{h} + \hat{h}^2 + \hat{h}^3 + \frac{d^2}{3}(1 + 3\hat{h}) = 0, \tag{A.11}$$

or equivalently,

$$q(z, w) = 4 - 6h + 4h^2 - h^3 + \frac{d^2}{3}(4 - 3h) = 0.$$

Using (A.7) and the equality  $r_4^2 = 2$ , this can be rewritten as an equation in terms of  $h$  and  $s$ ,

$$4 - 6h + \frac{8}{3}\left(2 - \frac{1}{s}\right)h^2 - 2\left(1 - \frac{1}{s}\right)h^3 = 0. \tag{A.12}$$

Note that  $h^2 = s$  cannot hold, because otherwise (A.6) would imply that  $h = \bar{h}$  is real – but substituting  $s = h^2$  into (A.12) results in a cubic equation for  $h$  with an invalid real solution  $h \approx 1.735\dots > r_4 = \sqrt{2}$  and two strictly complex solutions.

Since (A.12) is an equation for  $h$  with real coefficients, it is also valid for  $\bar{h} = s/h$  (cf. (A.6)), i.e.,

$$4 - 6\left(\frac{s}{h}\right) + \frac{8}{3}\left(2 - \frac{1}{s}\right)\left(\frac{s}{h}\right)^2 - 2\left(1 - \frac{1}{s}\right)\left(\frac{s}{h}\right)^3 = 0. \tag{A.13}$$

Now we proceed by staggered elimination.

- The linear combination

$$s \cdot (\text{A.12}) - h^2 \cdot (\text{A.13})$$

results in

$$\frac{2}{3h}(h^2 - s)((3s - 3)h^2 - (8s - 10)h + (3s - 3)s) = 0.$$

Since  $h^2 \neq s$ , we conclude

$$(3s - 3)h^2 - (8s - 10)h + (3s - 3)s = 0. \tag{A.14}$$

- The linear combination

$$s^2 \cdot (\text{A.12}) - h^4 \cdot (\text{A.13})$$

results in

$$-2(h^2 - s)(2h^2 + (s^2 - 4s)h + 2s) = 0.$$

Since  $h^2 \neq s$ , we conclude

$$2h^2 + (s^2 - 4s)h + 2s = 0. \tag{A.15}$$

— Now, the linear combination

$$2 \cdot (\text{A.14}) - (3s - 3) \cdot (\text{A.15})$$

results in

$$-h(3s^2 - 9s + 10)(s - 2) = 0.$$

Here, the quadratic factor has no real zero.

Therefore,  $s = r_2^2 = 2$  must hold, hence  $d = 0$  (cf. (A.7)), and (A.11) reduces to the cyclotomic equation of degree 3 in the variable  $\hat{h}$  (cf. (2.5)):

$$1 + \hat{h} + \hat{h}^2 + \hat{h}^3 = 0.$$

Thus,  $h$  is necessarily of the form

$$h = 1 - \hat{h} = 1 - \omega_4^j = 1 - e^{2\pi ij/4}, \quad j \in \{1, 2, 3\},$$

where for  $j = 1, 3$  this solution satisfies  $h\bar{h} = s = r_4^2 = 2$ . Eventually, we obtain

$$\frac{z+w}{2} = h = 1 - e^{2\pi ij/4}, \quad j \in \{1, 3\}, \quad \left| \frac{z+w}{2} \right| = \sqrt{2} = r_4.$$

Now,  $|z| = |w| = r_4$  implies

$$w = z = 1 - \omega_4, \quad \text{or} \quad w = z = 1 - \omega_4^3 = 1 - \bar{\omega}_4,$$

as asserted.

- $k = 5$  (with  $r_5 = \frac{\sqrt{10}}{\sqrt{5+\sqrt{5}}}$ )

In terms of  $\hat{h}$  and  $d$ , the reduced equation (degree 4) reads (cf. (2.14))

$$\hat{q}(\hat{z}, \hat{w}) = 1 + \hat{h} + \hat{h}^2 + \hat{h}^3 + \hat{h}^4 + \frac{d^2}{3} \left( 1 + 3\hat{h} + 6\hat{h}^2 \right) + \frac{d^4}{5} = 0, \quad (\text{A.16})$$

or equivalently,

$$q(z, w) = 5 - 10h + 10h^2 - 5h^3 + h^4 + \frac{d^2}{3} (10 - 15h + 6h^2) + \frac{d^4}{5} = 0.$$

Using (A.7), this can be rewritten as an equation in terms of  $h$  and  $s$  (let  $r := r_5$ ; we do not insert the value of  $r_5$  for the moment),

$$5 - 10h + \frac{10}{3} \left( 4 - \frac{r^2}{s} \right) h^2 - 5 \left( 2 - \frac{r^2}{s} \right) h^3 + \frac{1}{5} \left( 16 - \frac{12r^2}{s} + \frac{r^4}{s^2} \right) h^4 = 0. \quad (\text{A.17})$$

Note that  $h^2 = s$  cannot hold, because otherwise (A.6) would imply that  $h = \bar{h}$  is real – but substituting  $s = h^2$  and  $r = r_5$  into (A.17) results in a quartic equation for  $h$  with four strictly complex solutions.

Since (A.17) is an equation for  $h$  with real coefficients, it is also valid for  $\bar{h} = s/h$  (cf. (A.6)), i.e.,

$$5 - 10 \left(\frac{s}{h}\right) + \frac{10}{3} \left(4 - \frac{r^2}{s}\right) \left(\frac{s}{h}\right)^2 - 5 \left(2 - \frac{r^2}{s}\right) \left(\frac{s}{h}\right)^3 + \frac{1}{5} \left(16 - \frac{12r^2}{s} + \frac{r^4}{s^2}\right) \left(\frac{s}{h}\right)^4 = 0. \quad (\text{A.18})$$

Now we proceed by staggered elimination, which we present in the form of a Maple code.

```
> '(A.17) ': = 5 - 10*h + (10/3)*(4 - r^2/s)*h^2 - 5*(2 - r^2/s)*h^3 + (1/5)*(16 - 12*r^2/s + r^4/s^2)*h^4;
```

$$(\text{A.17}) := 5 - 10h + \frac{10 \left(4 - \frac{r^2}{s}\right) h^2}{3} - 5 \left(2 - \frac{r^2}{s}\right) h^3 + \frac{\left(16 - \frac{12r^2}{s} + \frac{r^4}{s^2}\right) h^4}{5}$$

```
> '(A.18) ': = collect(subs(h=s/h, '(A.17) '), h);
```

$$(\text{A.18}) := 5 - \frac{10s}{h} + \frac{10 \left(4 - \frac{r^2}{s}\right) s^2}{3h^2} - \frac{5 \left(2 - \frac{r^2}{s}\right) s^3}{h^3} + \frac{\left(16 - \frac{12r^2}{s} + \frac{r^4}{s^2}\right) s^4}{5h^4}$$

— Since  $h^2 \neq s$ , the linear combination

$$s \cdot (\text{A.17}) - h^2 \cdot (\text{A.18})$$

implies (A.19) = 0:

```
> factor(s*(A.17) - h^2*(A.18));
```

$$\begin{aligned} & (-s + h^2)(48h^4s^2 - 36h^4r^2s + 3h^4r^4 - 150h^3s^2 + 75h^3sr^2 - 50h^2sr^2 \\ & - 75h^2s + 200h^2s^2 + 48h^2s^3 - 36h^2s^2r^2 + 3h^2sr^4 - 150hs^3 \\ & + 75hs^2r^2 + 48s^4 - 36s^3r^2 + 3s^2r^4)/(15h^2s) \end{aligned}$$

```
> op(3,%);
```

$$\begin{aligned} & 48h^4s^2 - 36h^4r^2s + 3h^4r^4 - 150h^3s^2 + 75h^3sr^2 - 50h^2sr^2 - 75h^2s \\ & + 200h^2s^2 + 48h^2s^3 - 36h^2s^2r^2 + 3h^2sr^4 - 150hs^3 \\ & + 75hs^2r^2 + 48s^4 - 36s^3r^2 + 3s^2r^4 \end{aligned}$$

```
> '(A.19) ': = collect(%, h);
```

$$\begin{aligned} (\text{A.19}) := & (48s^2 - 36r^2s + 3r^4)h^4 + (-150s^2 + 75r^2s)h^3 \\ & + (-50r^2s - 75s + 200s^2 + 48s^3 - 36s^2r^2 + 3sr^4)h^2 \\ & + (-150s^3 + 75s^2r^2)h + 48s^4 - 36s^3r^2 + 3s^2r^4 \end{aligned} \quad (\text{A.19})$$

— Since  $h^2 \neq s$ , the linear combination

$$s^2 \cdot (\text{A.17}) - h^4 \cdot (\text{A.18})$$

implies (A.20) = 0, which is quadratic in  $h$ :

```
> factor(s^2*(A.17) - h^4*(A.18));
```

$$(-s + h^2)(h^2r^4 - 25h^2 + 16h^2s^2 - 12h^2sr^2 - 50hs^2 + 25hr^2s + 50hs + sr^4 - 25s + 16s^3 - 12s^2r^2)/5$$

```
> op(3,%);
```

$$h^2r^4 - 25h^2 + 16h^2s^2 - 12h^2sr^2 - 50hs^2 + 25hr^2s + 50hs + sr^4 - 25s + 16s^3 - 12s^2r^2$$

```
> '(A.20) := collect(% , h);
```

(A.20)

$$(\text{A.20}) := (r^4 - 25 + 16s^2 - 12r^2s)h^2 + (-50s^2 + 25r^2s + 50s)h + sr^4 - 25s + 16s^3 - 12s^2r^2$$

— Since  $h^2 \neq s$ , the linear combination

$$s^3 \cdot (\text{A.17}) - h^6 \cdot (\text{A.18})$$

implies (A.21) = 0:

```
> factor(s^3*(A.17) - h^6*(A.18));
```

$$(-s + h^2)(-75h^4 + 150h^3s + 3h^2sr^4 + 48h^2s^3 - 36h^2s^2r^2 + 50h^2sr^2 - 200h^2s^2 - 75h^2s + 150hs^2 - 75s^2)/15$$

```
> op(3,%);
```

$$-75h^4 + 150h^3s + 3h^2sr^4 + 48h^2s^3 - 36h^2s^2r^2 + 50h^2sr^2 - 200h^2s^2 - 75h^2s + 150hs^2 - 75s^2$$

```
> '(A.21) := collect(% , h);
```

(A.21)

$$(\text{A.21}) := -75h^4 + 150h^3s + (3sr^4 + 48s^3 - 36s^2r^2 + 50r^2s - 200s^2 - 75s)h^2 + 150hs^2 - 75s^2$$

— Next we perform a linear combination of (A.19) and (A.21) eliminating the terms which are quartic in  $h$ . The result contains a factor  $h$ , and we end up with (A.22) = 0, which is quadratic in  $h$ :

```
> factor(coeff('(A.21)', h, 4) * (A.19) - coeff('(A.19)', h, 4) * (A.21));
```

$$\begin{aligned}
& -3hs(-1875h - 1250hr^2 + 1875r^2h^2 + 5000hs - 3750h^2s \\
& - 1800h^2sr^2 + 3200hs^2r^2 - 800sr^4h - 3200hs^3 - 3750s^2 \\
& + 2400s^3 + 1875r^2s - 1800s^2r^2 + 2400h^2s^2 + 528r^4hs^2 \\
& - 1152hs^3r^2 - 72hr^6s + 150h^2r^4 + 3hr^8 + 768hs^4 + 50hr^6 + 150sr^4)
\end{aligned}$$

$$> \text{'(A.22)'} := \text{collect}(\% / (-3*\mathbf{h}*s), \mathbf{h}); \quad (\text{A.22})$$

$$\begin{aligned}
(\text{A.22}) & := (1875r^2 - 3750s - 1800r^2s + 2400s^2 + 150r^4)h^2 \\
& + (-1250r^2 - 1152s^3r^2 + 5000s - 1875 + 3r^8 + 3200s^2r^2 - 800sr^4 \\
& - 72r^6s - 3200s^3 + 528s^2r^4 + 768s^4 + 50r^6)h \\
& - 3750s^2 + 2400s^3 + 1875r^2s - 1800s^2r^2 + 150sr^4
\end{aligned}$$

- Eventually, we eliminate the terms which are quadratic in  $h$  from equations (A.20) and (A.22). The result contains a factor  $h$ , and we obtain an equation of degree 6 for  $s$ . After substituting the value  $r = r_5$ , the latter factors into polynomials of degree 5 and 1 in  $s$ . For the quintic factor it can be shown by an analysis based on Sturm sequences that it has no zero in the interval  $[0, r_5^2] = [0, (5 - \sqrt{5})/2]$ . We do not present this argument here but simply resort to numerical approximation. We obtain an invalid real solution  $s \approx 2.410 \dots > r_5^2$  and four strictly complex solutions:

$$\begin{aligned}
& > \text{factor}(\text{coeff}(\text{'(A.22)'} , \mathbf{h}, 2) * \text{'(A.20)'} - \text{coeff}(\text{'(A.20)'} , \mathbf{h}, 2) * \text{'(A.22)'}); \\
& - h(46875 - 125000s + 117500s^2r^2 + 31250r^2 - 71250r^2s - 14375sr^4 \\
& - 121200s^3r^2 + 39300s^2r^4 + 89600s^4r^2 - 54400s^3r^4 - 1950r^6s \\
& + 157500s^2 - 147500s^3 + 100800s^4 - 51200s^5 + 13600r^6s^2 - 1400r^8s \\
& + 23040r^4s^4 + 1440r^8s^2 - 8640r^6s^3 - 108r^{10}s - 27648s^5r^2 + 3r^{12} \\
& + 50r^{10} + 12288s^6 - 75r^8 - 2500r^6 - 1875r^4)
\end{aligned}$$

$$> \text{collect}(\text{op}(3, \%), s);$$

$$\begin{aligned}
& 12288s^6 + (-27648r^2 - 51200)s^5 + (89600r^2 + 23040r^4 + 100800)s^4 \\
& + (-147500 - 121200r^2 - 54400r^4 - 8640r^6)s^3 \\
& + (117500r^2 + 1440r^8 + 157500 + 39300r^4 + 13600r^6)s^2 \\
& + (-125000 - 108r^{10} - 71250r^2 - 1950r^6 - 1400r^8 - 14375r^4)s \\
& + 46875 + 31250r^2 + 3r^{12} + 50r^{10} - 75r^8 - 2500r^6 - 1875r^4
\end{aligned}$$

$$> \mathbf{r5} := \text{sqrt}(10) / \text{sqrt}(5 + \text{sqrt}(5));$$

$$r5 := \frac{\sqrt{10}}{\sqrt{5 + \sqrt{5}}}$$

$$> \text{factor}(4*\text{subs}(\mathbf{r}=\mathbf{r5}, \%));$$

$$(24576s^5 - 179200s^4 + 15360s^4\sqrt{5} + 508800s^3 - 76800s^3\sqrt{5}$$

```

- 685 000s2 + 119 600s2√5 + 412 500s - 52 500s√5 - 69 375
- 20 675√5)(2s - 5 + √5)
> op(1,%);
24 576s5 - 179 200s4 + 15 360s4√5 + 508 800s3 - 76 800s3√5
- 685 000s2 + 119 600s2√5 + 412 500s - 52 500s√5 - 69 375
- 20 675√5
> solve(% ,s);
RootOf(%1, index = 1), RootOf(%1, index = 2), RootOf(%1, index = 3),
RootOf(%1, index = 4), RootOf(%1, index = 5)
%1 := 24 576_Z5 + (-179 200 + 15 360√5)_Z4 + (508 800 - 76 800√5)_Z3
+ (-685 000 + 119 600√5)_Z2 + (-52 500√5 + 412 500)_Z
- 69 375 - 20 675√5
> evalf [20] (%);
2.4101358391847271384,
1.3264349804676543239 + 0.69600508061974223202 I,
0.41555919030463116003 + 0.83494679402459839215 I,
0.41555919030463116003 - 0.83494679402459839215 I,
1.3264349804676543239 - 0.69600508061974223202 I
> %[1], evalf [20] (r5^2);
2.4101358391847271384, 1.3819660112501051518

```

Therefore,  $s$  must be the solution of the linear equation

$$2s - 5 + \sqrt{5} = 0,$$

which is precisely  $s = r_5^2 = (5 - \sqrt{5})/2$ . Hence, one has  $d = 0$  (cf. (A.7)), and (A.16) reduces to the cyclotomic equation of degree 4 in the variable  $\hat{h}$  (cf. (2.5)):

$$1 + \hat{h} + \hat{h}^2 + \hat{h}^3 + \hat{h}^4 = 0.$$

Thus,  $h$  necessarily is of the form

$$h = 1 - \hat{h} = 1 - \omega_5^j = 1 - e^{2\pi ij/5}, \quad j \in \{1, 2, 3, 4\},$$

where for  $j = 1, 4$  this solution satisfies  $h\bar{h} = s = r_5^2 = (5 - \sqrt{5})/2$ . Eventually, we obtain

$$\frac{z+w}{2} = h = 1 - e^{2\pi ij/5}, \quad j \in \{1, 4\}, \quad \left| \frac{z+w}{2} \right| = \frac{\sqrt{10}}{\sqrt{5+\sqrt{5}}} = r_5.$$

Now,  $|z| = |w| = r_5$  implies

$$w = z = 1 - \omega_5, \quad \text{or} \quad w = z = 1 - \omega_5^4 = 1 - \bar{\omega}_5,$$

as asserted. □

**Lemma A.4 (Analytic continuation for  $w = w(z)$ ).** *Let  $F(z, w)$  be a polynomial with complex coefficients. Assume that the discriminant set*

$$D = \{z \in \mathbb{C} : \exists w \in \mathbb{C} : F(z, w) = F_w(z, w) = 0\}$$

*is finite and that  $f$  has no zeros of the form  $(z, \infty)$ .*

*Pick a point  $z_0 \notin D$  and let  $R$  denote the distance from  $z_0$  to  $D$ . Then, for any solution  $(z_0, w_0)$  of  $F(z, w) = 0$ , there exists an analytic function  $w = w(z)$  defined on  $K_R(z_0) = \{z : |z - z_0| < R\}$  with  $F(z, w(z)) = 0$  for  $z \in K_R(z_0)$  and  $w(z_0) = w_0$ .*

*Proof.* We may assume w.l.o.g. that  $z_0 = w_0 = 0$  and  $R = 1$  in order to simplify notation. Application of the implicit function theorem yields a positive number  $\delta$ , a disk  $K_\delta := \{z : |z| < \delta\}$  and an analytic function  $f(z)$  defined on  $K_\delta$  and representing a local solution of  $F(z, w(z)) = 0$ .

If  $\delta = 1$ , we are done. Suppose not. Due to our assumptions, any local solution  $g$  on any disk  $K$  inside  $K_R(z_0)$  cannot have singularities, because  $g$  is algebraic, with no poles (they correspond to  $F(z, \infty) = 0$ ) nor branch points (when  $F = F_w = 0$ ). Thus,  $f$  possesses an analytic continuation  $w$  living on  $K_R(z_0)$ . Then  $F(z, w(z))$  is an analytic continuation of  $0 = f(z, f(z))$  and hence vanishes. □

APPENDIX B. VISUALIZATION

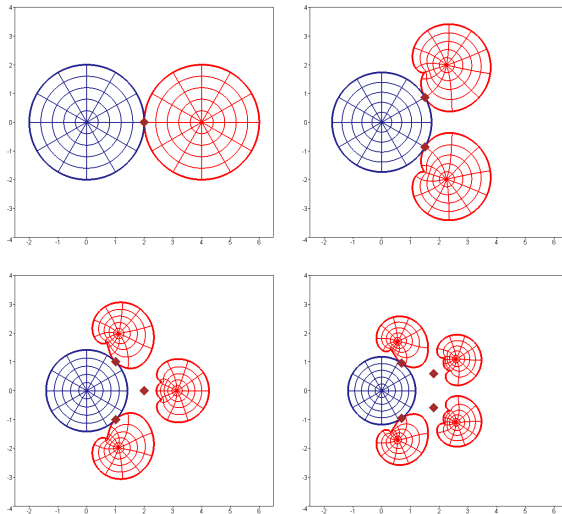
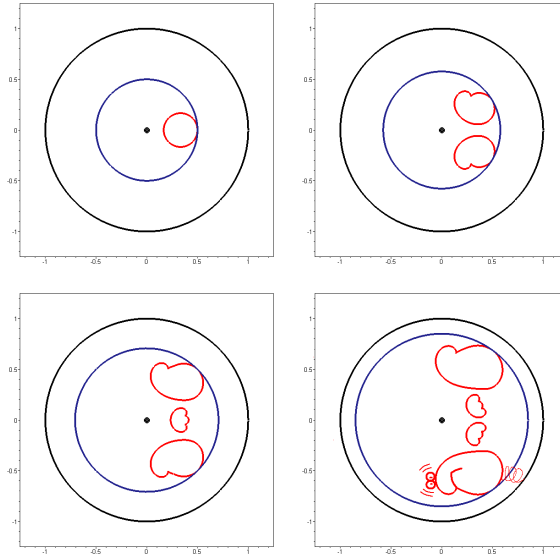


Fig. 2.  $z$ - $w$ -plots for  $k = 2 \dots 5$



**Fig. 3.**  $\zeta$ - $\eta$ -plots for  $k = 2 \dots 5$

## REFERENCES

- [1] Auzinger W., Kirlinger G., *Canonical forms for companion matrices and applications*, in preparation.
- [2] Burckel R. B., *An Introduction to Classical Complex Analysis*, Vol. 1, Pure and Applied Mathematics, 82. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1979.
- [3] Eder A., Kirlinger G., *A normal form for multistep companion matrices*, Math. Models and Methods in Applied Sciences, **11** (2001) 1, 57–70.
- [4] Hairer E., Nørsett S. P., Wanner G., *Solving Ordinary Differential Equations I. Nonstiff Problems*, 2nd ed., Springer-Verlag, 1987.
- [5] Hairer E., Wanner G., *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, Springer-Verlag, 1996.
- [6] Kirlinger G., *On the convergence of backward differentiation formulas for non-autonomous stiff initial value problems*, BIT 41, **5** (2001), 1039–1048.
- [7] Markushevich A. I., *Theory of Functions of a Complex Variable*, Vol. I, II, III, translated and edited by R. A. Silverman, Second English edition, Chelsea Publishing Co., New York, 1977.



Winfried Auzinger  
w.auzinger@tuwien.ac.at

Vienna University of Technology,  
Institute for Analysis and Scientific Computing,  
Wiedner Hauptstrasse 8–10/101, A-1040 Wien, Austria, EU.

Wolfgang Herfort  
w.herfort@uwien.ac.at

Vienna University of Technology,  
Institute for Analysis and Scientific Computing,  
Wiedner Hauptstrasse 8–10/101, A-1040 Wien, Austria, EU.

*Received: October 4, 2005.*