THE WEIGHTED AND SHIFTED SEVEN-STEP BDF METHOD FOR PARABOLIC EQUATIONS

GEORGIOS AKRIVIS, MINGHUA CHEN, AND FAN YU

ABSTRACT. Stability of the BDF methods of order up to five for parabolic equations can be established by the energy technique via Nevanlinna–Odeh multipliers. The nonexistence of Nevanlinna–Odeh multipliers makes the six-step BDF method special; however, the energy technique was recently extended by the authors in [Akrivis et al., SIAM J. Numer. Anal. **59** (2021) 2449–2472] and covers all six stable BDF methods. The seven-step BDF method is unstable for parabolic equations, since it is not even zero-stable. In this work, we construct and analyze a stable linear combination of two non zero-stable schemes, the seven-step BDF method and its shifted counterpart, referred to as WSBDF7 method. The stability regions of the WSBDF$q, q \leqslant 7$, with a *weight* $\vartheta \geqslant 1$, increase as ϑ increases and are larger than the stability regions of the classical q-step BDF methods, corresponding to $\vartheta = 1$. We determine novel and suitable multipliers for the WSBDF7 method and establish stability for parabolic equations by the energy technique. The proposed approach is applicable for mean curvature flow, gradient flows, fractional equations and nonlinear equations.

1. Introduction

Let $T > 0, u^0 \in H$, and consider the initial value problem of seeking $u \in C((0,T];\mathscr{D}(A)) \cap C([0,T];H)$ satisfying

$$
\begin{cases} u'(t) + Au(t) = f(t), & 0 < t < T, \\ u(0) = u^0, \end{cases} \tag{1.1}
$$

with A a positive definite, selfadjoint, linear operator on a Hilbert space $(H,(\cdot,\cdot))$ with domain $\mathscr{D}(A)$ dense in H and $f : [0,T] \to H$ a given forcing term. We shall analyze the discretization of (1.1) by the *weighted* and *shifted* q-step backward difference formula (WSBDFq) with $q = 7$, described by a *weight* $\vartheta > 0$ and the corresponding characteristic polynomials α and β,

$$
\alpha(\zeta) := \vartheta a(\zeta) + (1-\vartheta)\tilde{a}(\zeta) = \sum_{j=0}^{q} \alpha_j \zeta^j, \quad \beta(\zeta) := \vartheta \zeta^q + (1-\vartheta)\zeta^{q-1},\ q \leqslant 7, \tag{1.2}
$$

Date: December 11, 2024.

2010 *Mathematics Subject Classification.* Primary 65M12, 65M60; Secondary 65L06.

Key words and phrases. Weighted and shifted seven-step BDF method, multipliers, parabolic equations, stability estimate, energy technique.

The work was supported by the Science Fund for Distinguished Young Scholars of Gansu Province under Grant No. 23JRRA1020 and National Natural Science Foundation of China under Grant No. 12471381.

with a and $\tilde a$ the characteristic polynomials of the q-step BDF method and the shifted q-step BDF method, respectively, for $1 \leqslant q \leqslant 7$,

$$\begin{cases} a(\zeta) = \displaystyle\sum_{j=1}^{q} \frac{1}{j}\zeta^{q-j}(\zeta-1)^j = \sum_{j=0}^{q} a_j\zeta^j,\\ \tilde a(\zeta) = a(\zeta) - \displaystyle\sum_{j=2}^{q} \frac{1}{j-1}\zeta^{q-j}(\zeta-1)^j = \sum_{j=0}^{q} \tilde a_j\zeta^j. \end{cases} \tag{1.3}$$

In particular, $\tilde a(\zeta) = a(\zeta)$ for $q=1$.

Let $N \in \mathbb{N}, \tau := T/N$ be the time step, and $t_n := n\tau, n = 0,\dots,N$, be a uniform partition of the interval $[0,T]$. We recursively define a sequence of approximations u^m to the nodal values $u(t_m)$ of the exact solution by the WSBDF7 method,

$$\sum_{i=0}^{7} \alpha_i u^{n+i} + \vartheta\tau A u^{n+7} + (1-\vartheta)\tau A u^{n+6} = \vartheta\tau f^{n+7} + (1-\vartheta)\tau f^{n+6}, \quad \vartheta \geqslant 2.6, \tag{1.4}$$

for $n = 0,\dots,N-7$, with $f^m := f(t_m)$, assuming that starting approximations $u^0,\dots,u^6$ are given. For convenience, we suppressed the dependence of α and of its coefficients on ϑ. We are particularly interested in the WSBDF7 method (1.4) for $\vartheta = 3$,

$$\sum_{i=0}^{7} \alpha_i u^{n+i} + 3\tau A u^{n+7} - 2\tau A u^{n+6} = 3\tau f^{n+7} - 2\tau f^{n+6}, \quad n = 0,\dots,N-7. \tag{1.5}$$

Let $P_7 \in \mathbb{P}_7$ be the Lagrange interpolating polynomial of a function y at the nodes $t_n, t_{n+1},\dots,t_{n+7}$. We recall that the seven-step BDF method,

$$\sum_{i=0}^{7} a_i y^{n+i} = \tau f(t_{n+7}, y^{n+7}), \tag{1.6}$$

for an o.d.e. $y' = f(t,y)$, is constructed by approximating the derivative of y at the node t_{n+7} in the relation $y'(t_{n+7}) = f\big(t_{n+7}, y(t_{n+7})\big)$ by the derivative $P_7'(t_{n+7})$ of the interpolating polynomial. Analogously, the shifted seven-step BDF method,

$$\sum_{i=0}^{7} \tilde a_i y^{n+i} = \tau f(t_{n+6}, y^{n+6}), \tag{1.7}$$

is constructed by approximating $y'(t_{n+6})$ in the relation $y'(t_{n+6}) = f\big(t_{n+6}, y(t_{n+6})\big)$ by $P_7'(t_{n+6})$. Notice that $\vartheta P_7'(t_{n+7}) + (1-\vartheta)P_7'(t_{n+6})$ is, in general, different from $P_7'(t_{n+6}+\vartheta\tau)$.

Multiplying (1.6) and (1.7) by ϑ and $1-\vartheta$, respectively, and adding the results, we obtain the weighted and shifted seven-step BDF (WSBDF7) method,

$$\vartheta\sum_{i=0}^{7} a_i y^{n+i} + (1-\vartheta)\sum_{i=0}^{7} \tilde a_i y^{n+i} = \vartheta\tau f(t_{n+7}, y^{n+7}) + (1-\vartheta)\tau f(t_{n+6}, y^{n+6}). \tag{1.8}$$

In particular, for $\vartheta = 1$ and $\vartheta = 0$, the WSBDF7 method reduces to the standard seven-step BDF method and to the corresponding shifted seven-step BDF method.

It is well known that both methods, seven-step BDF and shifted seven-step BDF methods, are not zero-stable; for seven-step BDF method, see, e.g., [12, Theorem 3.4]; concerning shifted seven-step BDF method, it is easily seen that $\tilde a(-13)\tilde a(-12) < 0$, whence $\tilde a$ has a root in the interval $(-13,-12)$. The order of both methods is 7. Here, we show that their combination, WSBDF7, is $A(\varphi)$-stable for $\vartheta = 3$, stable even for parabolic equations.

The explicit form of the polynomials a and $\tilde{a}$ in (1.3) with $q = 7$ is

$$a(\zeta) = \frac{1}{420}\big(1089\zeta^7 - 2940\zeta^6 + 4410\zeta^5 - 4900\zeta^4 + 3675\zeta^3 - 1764\zeta^2 + 490\zeta - 60\big),$$
$$\tilde{a}(\zeta) = \frac{1}{420}\big(60\zeta^7 + 609\zeta^6 - 1260\zeta^5 + 1050\zeta^4 - 700\zeta^3 + 315\zeta^2 - 84\zeta + 10\big).$$

Let $|\cdot|$ denote the norm on H induced by the inner product $(\cdot,\cdot)$, and introduce on $V, V := \mathscr{D}(A^{1/2})$, the norm $\|\cdot\|$ by $\|v\| := |A^{1/2}v|$. We identify H with its dual, and denote by V' the dual of V, and by $\|\cdot\|_\star$ the dual norm on V', $\|v\|_\star = |A^{-1/2}v|$. We shall use the notation $(\cdot,\cdot)$ also for the antiduality pairing between V' and V. For simplicity, we denote by $\langle\cdot,\cdot\rangle$ the inner product on V, $\langle v, w\rangle := (A^{1/2}v, A^{1/2}w)$.

Our stability results are established by the energy technique utilizing suitable multipliers, and are given in the following two theorems and in a corollary.

Theorem 1.1 (Stability of method (1.5))**.** *Let $u^0, u^1, \dots, u^6 \in V$. The WSBDF7 method (1.5) is stable in the sense that*

$$|u^n|^2 + \tau\|u^n\|^2 \leqslant C\sum_{j=0}^{6}\big(|u^j|^2 + \tau\|u^j\|^2\big) + C\tau\sum_{\ell=6}^{n}\|f^\ell\|_\star^2, \quad n = 7,\dots,N. \tag{1.9}$$

Here C denotes a generic constant, independent of T and the operator A as well as of f, τ, and n.

Theorem 1.2 (Second stability estimate)**.** *Let $u^0, u^1, \dots, u^6 \in V$, and let us indicate by a dot the application of the seven-step weighted and shifted backward difference operator,*

$$\dot{v}^m := \frac{1}{\tau}\sum_{i=0}^{7}\alpha_i v^{m-7+i}, \quad m = 7,\dots,N. \tag{1.10}$$

The WSBDF7 method (1.5) is stable in the sense that

$$\|u^n\|^2 + \tau|\dot{u}^n|^2 \leqslant C\sum_{j=0}^{6}\|u^j\|^2 + C\tau\sum_{\ell=6}^{n}|f^\ell|^2, \quad n = 7,\dots,N. \tag{1.11}$$

Here C denotes a generic constant as in Theorem 1.1.

Corollary 1.1 (Third stability estimate)**.** *Let $u^0, u^1, \dots, u^6 \in V$, and let us denote by $\partial_\tau u^n := (u^n - u^{n-1})/\tau, n = 1,\dots,N$, the backward difference quotients. The WSBDF7 method (1.5) is stable in the sense that*

$$\|u^n\|^2 + \tau|\partial_\tau u^n|^2 \leqslant C\sum_{j=0}^{6}\|u^j\|^2 + C\tau\sum_{j=1}^{6}|\partial_\tau u^j|^2 + C\tau\sum_{\ell=6}^{n}|f^\ell|^2, \quad n = 7,\dots,N, \tag{1.12}$$

with a generic constant C.

The application of the energy technique to establish stability of high order multistep methods for parabolic equations relies on suitable multipliers. The multiplier technique was introduced by Nevanlinna and Odeh in [18] and is based on Dahquist's equivalence between A- and G-stability; see also [13, §V.8, pp. 342–347]. In [18], suitable multipliers for the three-, four- and five-step BDF methods were determined; see also [3] for optimal Nevanlinna–Odeh multipliers for these methods, i.e., multipliers with minimal sum of absolute values.

The multiplier technique became widely known and popular after its first actual application to the stability analysis for parabolic equations by Lubich, Mansour, and Venkataraman in 2013; see [17]. In recent years, the energy technique has been frequently used in the analyses of various variants of BDF methods of order up to 5, such as fully implicit, linearly implicit or implicit–explicit, for a series of linear and nonlinear equations of parabolic type. Nonexistence of Nevanlinna–Odeh multipliers for the six-step BDF method was established in [1]; there, to overcome this difficulty, the notion of multipliers was slightly modified, and multipliers for the six-step BDF method were determined, which, in combination with the Grenander–Szegő theorem for symmetric banded Toeplitz matrices, made the energy technique applicable also to this method for parabolic equations with self-adjoint elliptic part.

Here, focusing on the discretization of parabolic equations with self-adjoint elliptic part by multistep methods, we first extend the notion of multipliers, and then determine suitable multipliers for the WSBDF7 method (1.5) and prove Theorems 1.1 and 1.2 by the energy technique. The new, more general notion of multipliers reduces to the corresponding notion in [1] in the case of the BDF methods; however, both the proofs and the stability results here and in [1] are different. The present approach is shorter and simpler but it yields weaker stability results, in the sense that (1.9) leads to optimal order error estimates in the discrete $\ell^\infty(H)$ norm but to suboptimal by half-an-order error estimates in the discrete $\ell^\infty(V)$ norm; see (7.2); in contrast, the stability estimates in [1] lead to optimal order error estimate in the discrete $\ell^\infty(H)$ as well as in the discrete $\ell^2(V)$ norms. Of course, (1.11) leads to optimal order error estimates in the discrete $\ell^\infty(V)$ norm. Let us also mention that the stability approach in [1] is restricted to BDF methods, in which case banded Toeplitz matrices enter. In the case of the WSBDF7 method (1.5), the corresponding matrices reduce to banded Toeplitz matrices only if we disregard their last row and column; this fact prevents us from using the Grenander–Szegő theorem.

In 1991, linear multistep methods of orders from 2 to 7 for ordinary differential equations, with stability regions larger than the stability regions of the BDF method of the same order, were constructed in [16]. In particular, the seven-step methods of order 7 of [16] are the WSBDF7 method (1.4) with a parameter $\vartheta \geqslant 2.6$. High order implicit-explicit multistep methods were constructed and analyzed in [9]. Then, in 1995, implicit-explicit multistep schemes of orders from 1 to 4 were constructed in [4]; the method of order 2 in [4] coincides with the WSBDF2 method, but the methods of order 3 and 4 are different from the WSBDF$q, q = 3, 4$, methods. The construction techniques in [16] and for the WSBDF methods are significantly different. The point of departure in [16] is the polynomial β in (1.2); the corresponding polynomial α is then determined via the order conditions. Here, the WSBDF7 method (1.8) is constructed by the simpler, direct, and more flexible weighted and shifted technique, which immediately extends also to variable time step schemes. For example, the variable time-step WSBDF3 methods for parabolic equations are presented in [7]. However, there is no published work on variable or even uniform time-step WSBDF$q, q \geqslant 4$, methods.

The proposed methods, for $q \leqslant 6$, including the classical case $\vartheta = 1$, have been recently widely used to various applied scientific phenomena, such as mean curvature flow [11, 15], gradient flows [14], and fractional equations [8].

An outline of the paper is as follows: In the short Section 2, we briefly comment on the stability regions of the WSBDF7 method (1.8) for various values of the parameter ϑ. In Section 3, we make precise the requirements on the multipliers for multistep methods that are suitable for our stability approach, and show that our notion extends the multiplier notion of [1] for BDF methods. The remaining part of the article is devoted to the WSBDF7 method (1.5). First, in Section 4 we give a suitable multiplier for this method and in Section 5 comment on the determination of such multipliers; in particular, we provide information about the range of such multipliers with up to the first four nonvanishing components. Sections 6 and 7 are devoted to the proof of the stability estimates (1.9), (1.11) and (1.12), and to the derivation of error estimates. We conclude in Section 8 with numerical results.

2. Stability regions

The q-step BDF methods are $A(\varphi_q)$-stable with $\varphi_1 = \varphi_2 = 90°$, $\varphi_3 \approx 86.03°, \varphi_4 \approx 73.35°, \varphi_5 \approx 51.84°$, and $\varphi_6 \approx 17.84°$; see [13, Section V.2]. The WSBDFq methods are $A(\tilde\varphi_q)$-stable with $\tilde\varphi_1 = \tilde\varphi_2 = 90°$, $\tilde\varphi_3 \approx 89.55°, \tilde\varphi_4 \approx 85.32°, \tilde\varphi_5 \approx 73.2°$, and $\tilde\varphi_6 \approx 51.23°$, for the weights $\vartheta_3 = 20, \vartheta_4 = 60, \vartheta_5 = 48, \vartheta_6 = 50$, respectively; see [16, Table 2]. Notice that $\varphi_q < \tilde\varphi_q < \varphi_{q-1}$ for $q = 3, 4, 5, 6$. It can also be shown that $\tilde\varphi_q \to \varphi_{q-1}$ as $\vartheta \to \infty$ for $q = 3, \dots, 7$; see Remark 2.1. It seems that the value $\tilde\varphi_7 \approx 18.32° > \varphi_6 \approx 17.84°$ for $\vartheta_7 = 200/7$ given in [16] is incorrect. We numerically computed the approximations $\tilde\varphi_1 = \tilde\varphi_2 = 90°, \tilde\varphi_3 \approx 89.99°, \tilde\varphi_4 \approx 85.93°, \tilde\varphi_5 \approx 73.2°$, $\tilde\varphi_6 \approx 51.63°$, and $\tilde\varphi_7 \approx 17.47°$ for the weight $\vartheta = 100$.

Remark 2.1 (The limit of the stability angles $\tilde\varphi_q$ of WSBDFq)**.** Nørsett, [19, p. 263], established an A(φ_q)-stability criterion for the q-step BDF methods, namely, in his notation,

$$\tan(\varphi_q) = \min_{x\in D_q}\left(-\sqrt{1-x^2}\cdot\frac{I_q(x)}{R_q(x)}\right), \qquad D_q = \{x \in [-1, 1], R_q(x) < 0\};$$

here $I_q(x)$ and $R_q(x)$ are related to the imaginary and real parts of points on the root locus curve of the method, and are expressed in terms of the Chebyshev polynomials. Using this criterion and results in [16, p. 7], we have

$$\tan(\tilde\varphi_q) = \min_{x\in \tilde D_q}\left(-\sqrt{1-x^2}\cdot\lim_{\vartheta\to\infty}\frac{\tilde I_q(x)}{\tilde R_q(x)}\right), \qquad \tilde D_q = \{x \in [-1, 1], \tilde R_q(x) < 0\}$$

for the stability angle $\tilde\varphi_q$ of the WSBDFq method. Now, we can easily check that

$$\lim_{\vartheta\to\infty}\frac{\tilde I_q(x)}{\tilde R_q(x)} = \frac{I_{q-1}(x)}{R_{q-1}(x)}, \qquad q = 3, \dots, 7,$$

and infer that

$$\tilde\varphi_q \to \varphi_{q-1} \quad \text{as} \quad \vartheta \to \infty, \qquad q = 3, \dots, 7.$$

Here, we examine the stability regions of the WSBDF7 method (1.8). For the test equation $y'(t) = \lambda y(t), \lambda \in \mathbb{C}$, the method reads

$$\sum_{i=0}^{7} \alpha_i y^{n+i} = \lambda\tau\vartheta y^{n+7} + \lambda\tau(1-\vartheta)y^{n+6}, \qquad n = 0, \dots, N-7. \tag{2.1}$$

To study the stability regions, we set $z = \lambda\tau$ in (2.1) and consider the characteristic polynomial

$$p(\zeta) := \alpha(\zeta) - z\beta(\zeta), \qquad \zeta \in \mathbb{C};$$

again, for convenience, we suppressed the dependence of the polynomials α and β on ϑ; see (1.2). Then, the stability region of the method is the set of all $z \in \mathbb{C}$ such that the characteristic polynomial p satisfies the root condition. We plot the stability regions of

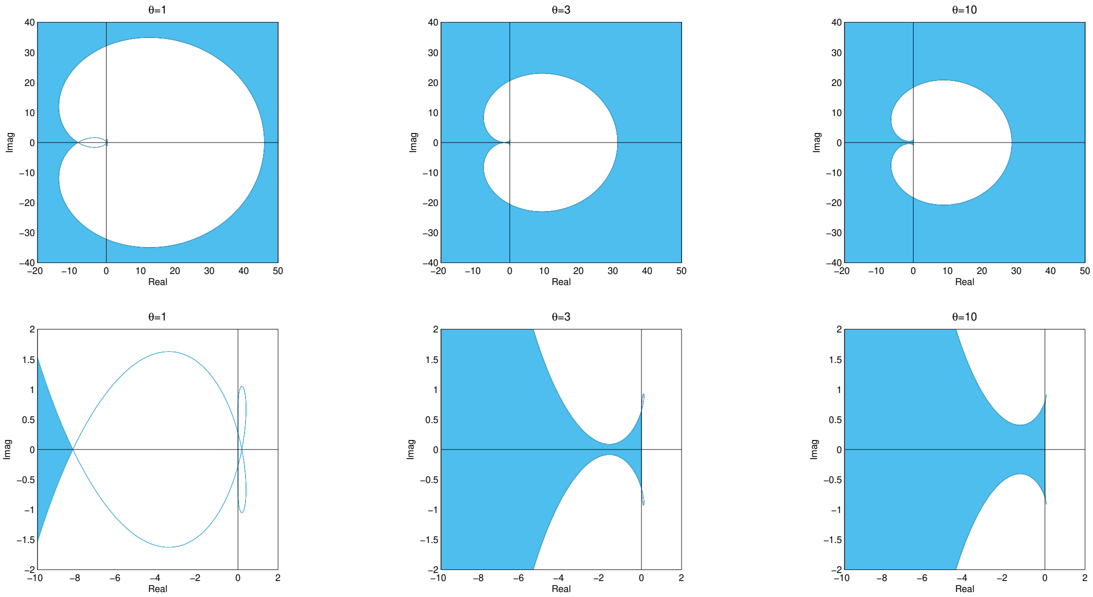

FIGURE 2.1. The stability regions (upper panel) as well as zoom in around the origin (lower panel), in light blue, for $\vartheta = 1, 3, 10$, respectively.

the WSBDF7 method (2.1) for $\vartheta = 1, 3, 10$ in Figure 2.1. Notice that the stability regions increase as ϑ increases.

3. Multipliers for multistep methods

Here, we extend the notion of multipliers for multistep methods applied to parabolic equations with self-adjoint elliptic part. The present, more general notion of multipliers reduces to the corresponding notion in [1] in the case of the BDF methods. The multiplier technique hinges on the celebrated equivalence of A- and G-stability for multistep methods by Dahlquist.

Lemma 3.1 ([10]; see also [5] and [13, Section V.6])**.** *Let $\rho(\zeta) = \rho_q\zeta^q + \cdots + \rho_0$ and $\sigma(\zeta) = \sigma_q\zeta^q + \cdots + \sigma_0$ be polynomials of degree q, with real coefficients, that have no common divisor. Let $(\cdot,\cdot)$ be a real inner product with associated norm $|\cdot|$. If*

$$\text{(A)} \qquad \operatorname{Re}\frac{\rho(\zeta)}{\sigma(\zeta)} > 0 \quad \text{for } |\zeta| > 1,$$

then there exist a positive definite symmetric matrix $G = (g_{ij}) \in \mathbb{R}^{q,q}$ and real $\gamma_0, \dots, \gamma_q$ such that for $v^0, \dots, v^q$ in the inner product space,

$$\text{(G)} \qquad \Big(\sum_{i=0}^q \rho_i v^i, \sum_{j=0}^q \sigma_j v^j\Big) = \sum_{i,j=1}^q g_{ij}(v^i, v^j) - \sum_{i,j=1}^q g_{ij}(v^{i-1}, v^{j-1}) + \Big|\sum_{i=0}^q \gamma_i v^i\Big|^2.$$

Let us briefly comment on the assumption that ρ and σ are polynomials of the same degree; obviously, if (A) were satisfied for polynomials of different degrees, then A-stable explicit methods would exist. First, nonconstant polynomials cannot satisfy (A) since they cannot retain the sign of their real part as $|z| \to \infty$. This shows also that the degree of ρ cannot exceed the degree of σ, since then we could write ρ/σ as the sum of a nonconstant polynomial and a rational function R such that the degree of its numerator is lower than the degree of its denominator, whence, in particular, $\lim_{|z|\to\infty} \operatorname{Re} R(z) = 0$. Finally, (A) is symmetric with respect to ρ and σ since $\operatorname{Re}[\rho(z)/\sigma(z)]\operatorname{Re}[\sigma(z)/\rho(z)] = \cos^2\varphi > 0$ for $\rho(z)/\sigma(z) = r\mathrm{e}^{\mathrm{i}\varphi}$ not purely imaginary.

Next, we specify our requirements on the multipliers; in Section 4, we shall provide motivation for these requirements.

Definition 3.1 (Multipliers)**.** Let α and β be the characteristic polynomials of an A(0)-stable q-step method; then, $\alpha_q\beta_q > 0$, and thus we can assume that the leading coefficients α_q and β_q are positive. For a q-tuple of real numbers $(\mu_1,\dots,\mu_q)$ consider the polynomial $\mu(\zeta) := \zeta^q - \mu_1\zeta^{q-1} - \cdots - \mu_q$. Then, $(\mu_1,\dots,\mu_q)$ is called a *multiplier* of the method if it satisfies three properties, namely, if the pairs of polynomials (α,μ) and (β,μ) have no common divisors, except possibly of a common factor of the form ζ^ℓ for the pair (β,μ), and satisfy the A-stability condition (A) in Lemma 3.1 and a slightly more restrictive version of it, respectively, that is,

$$\operatorname{Re}\frac{\alpha(\zeta)}{\mu(\zeta)} > 0 \quad \text{for } |\zeta| > 1, \tag{3.1}$$

i.e., the method described by the coefficients of the polynomials α and μ is A-stable, and

$$\operatorname{Re}\frac{\beta(\zeta)}{\mu(\zeta)} > c_\star \quad \text{for } |\zeta| > 1, \tag{3.2}$$

for some positive constant $c_\star$, and the polynomial μ does not have unimodular roots.

The A-stability condition (3.1) is symmetric with respect to the polynomials α and μ, since for any not purely imaginary complex number $z = r\mathrm{e}^{\mathrm{i}\varphi}$, we have $\operatorname{Re} z \operatorname{Re}\frac{1}{z} = \cos^2\varphi > 0$. This property is crucial because otherwise the A-stability condition (A) would not be equivalent to the obviously symmetric condition (G). On the other hand, the more stringent condition (3.2) cannot be symmetric in case μ has unimodular roots, since it implies that β does not have unimodular roots. For instance for $\mu(\zeta) = \zeta - 1$ and $\beta(\zeta) = \zeta$, i.e., for the characteristic polynomials of the implicit Euler method, we obviously have

$$\lim_{\zeta\to 1}\frac{\mu(\zeta)}{\beta(\zeta)} = 0,$$

while, with $\zeta = r\mathrm{e}^{\mathrm{i}\varphi}$,

$$\operatorname{Re}\frac{\beta(\zeta)}{\mu(\zeta)} = \frac{r^2 - r\cos\varphi}{r^2 - 2r\cos\varphi + 1} \quad\text{and}\quad \frac{r^2 - r\cos\varphi}{r^2 - 2r\cos\varphi + 1} \geqslant \frac{1}{2} \iff r \geqslant 1.$$

The additional condition that μ does not have unimodular roots makes condition (3.2) symmetric; cf. also Lemma 3.2.

Let us note that the third property of multipliers, i.e., that μ does not have unimodular roots, is not needed for the proof of Theorem 1.1; we will use it only in the proof of Theorem 1.2.

Remark 3.1 (Equivalent version of the conditions on the pair (β,μ))**.** Let $\ell \geqslant 0$ be the largest integer such that ζ^ℓ is a common factor of β and μ. We factor ζ^ℓ out, and write β and μ in the form

$$\beta(\zeta) = \zeta^\ell\delta(\zeta) \quad\text{and}\quad \mu(\zeta) = \zeta^\ell\kappa(\zeta). \tag{3.3}$$

Then, the conditions on the pair (β,μ) in the previous Definition can be equivalently formulated in the form: the pair of polynomials (δ,κ) has no common divisor and satisfies the condition

$$\operatorname{Re}\frac{\delta(\zeta)}{\kappa(\zeta)} > c_\star \quad \text{for } |\zeta| > 1, \tag{3.4}$$

for some positive constant $c_\star$.

The following result is an immediate consequence of the maximum principle for harmonic functions.

Lemma 3.2 (Equivalent versions of conditions (3.1) and (3.2))**.** *With the notation in Definition 3.1 and $\mu_0 := -1$, conditions (3.1) and (3.2) are equivalent to*

$$-\sum_{j,\ell=0}^{q} \alpha_j \mu_\ell \cos((j+\ell-q)\varphi) \geqslant 0 \quad \forall \varphi \in \mathbb{R} \tag{3.5}$$

and, if the polynomial μ does not have unimodular roots,

$$-\sum_{j,\ell=0}^{q} \beta_j \mu_\ell \cos((j+\ell-q)\varphi) \geqslant \tilde{c}_\star \quad \forall \varphi \in \mathbb{R}, \tag{3.6}$$

for some positive constant $\tilde{c}_\star$, respectively.

Proof. The rational functions α/μ and β/μ are holomorphic outside the unit disk in the complex plane and

$$\lim_{|z|\to\infty} \frac{\alpha(z)}{\mu(z)} = \alpha_q > 0, \qquad \lim_{|z|\to\infty} \frac{\beta(z)}{\mu(z)} = \beta_q > 0.$$

Therefore, according to the maximum principle for harmonic functions, (3.1) and (3.2) for $c_\star \leqslant \beta_q$ are equivalent to

$$\operatorname{Re}\big[\alpha(\zeta)\mu(\bar{\zeta})\big] \geqslant 0 \quad \forall \zeta \in \mathscr{K}, \quad \text{and} \quad \operatorname{Re}\frac{\beta(\zeta)}{\mu(\zeta)} \geqslant c_\star \quad \forall \zeta \in \mathscr{K},$$

respectively, with $\mathscr{K}$ the unit circle in the complex plane, $\mathscr{K} := \{\zeta \in \mathbb{C} : |\zeta| = 1\}$, i.e., equivalent to

$$\operatorname{Re}\big[\alpha(\mathrm{e}^{\mathrm{i}\varphi})\mu(\mathrm{e}^{-\mathrm{i}\varphi})\big] \geqslant 0 \quad \text{and} \quad \operatorname{Re}\big[\beta(\mathrm{e}^{\mathrm{i}\varphi})\mu(\mathrm{e}^{-\mathrm{i}\varphi})\big] \geqslant c_\star |\mu(\mathrm{e}^{\mathrm{i}\varphi})|^2 \quad \forall \varphi \in \mathbb{R}. \tag{3.7}$$

Now, $\mu(\zeta) = -\sum_{\ell=0}^{q} \mu_\ell \zeta^{q-\ell}$, and thus

$$\beta(\mathrm{e}^{\mathrm{i}\varphi})\mu(\mathrm{e}^{-\mathrm{i}\varphi}) = -\sum_{j=0}^{q} \beta_j \mathrm{e}^{j\mathrm{i}\varphi} \sum_{\ell=0}^{q} \mu_\ell \mathrm{e}^{(\ell-q)\mathrm{i}\varphi} = -\sum_{j,\ell=0}^{q} \beta_j \mu_\ell \mathrm{e}^{(j+\ell-q)\mathrm{i}\varphi},$$

whence

$$\operatorname{Re}\big[\beta(\mathrm{e}^{\mathrm{i}\varphi})\mu(\mathrm{e}^{-\mathrm{i}\varphi})\big] = -\sum_{j,\ell=0}^{q} \beta_j \mu_\ell \cos((j+\ell-q)\varphi),$$

and, since $|\mu(\mathrm{e}^{\mathrm{i}\varphi})|^2$ is bounded from above and below by positive constants, (3.6) is equivalent to the second relation in (3.7). Analogously, the first relation in (3.7) is equivalent to (3.5). ☐

Remark 3.2 (BDF methods)**.** In the case of the standard q-step BDF method, we have $\beta(\zeta) = \zeta^q$, whence (3.6) takes the form

$$-\sum_{\ell=0}^{q} \mu_\ell \cos(\ell\varphi) \geqslant \tilde{c}_\star \quad \forall \varphi \in \mathbb{R},$$

i.e., since $\mu_0 = -1$,

$$1 - \mu_1 \cos\varphi - \cdots - \mu_q \cos(q\varphi) \geqslant \tilde{c}_\star \quad \forall \varphi \in \mathbb{R}. \tag{3.8}$$

Since $\tilde{c}_\star$ can be chosen arbitrarily small, (3.8) can be written as a positivity condition,

$$1 - \mu_1 \cos\varphi - \cdots - \mu_q \cos(q\varphi) > 0 \quad \forall \varphi \in \mathbb{R}. \tag{3.9}$$

Conditions (3.1), with α the characteristic polynomial of the q-step BDF method, and (3.9) were used in [1] to establish stability of BDF methods for parabolic equations with self-adjoint elliptic part by the energy technique. The motivation for the positivity condition (3.9) in [1] was that $1 - \mu_1 \cos\varphi - \cdots - \mu_q \cos(q\varphi)$ is the generating function of symmetric banded Toeplitz matrices of arbitrary dimension entering into the stability analysis, and (3.9) ensured the positive definiteness of these matrices.

4. Multipliers for the WSBDF7 method (1.5)

From now on, we consider the WSBDF7 method (1.5), i.e., (1.4) with $\vartheta = 3$.

As we shall see,

$$\mu_1 = 1.6, \quad \mu_2 = -1.6, \quad \mu_3 = 1.1, \quad \mu_4 = -0.3, \quad \mu_5 = \mu_6 = \mu_7 = 0, \tag{4.1}$$

are multipliers of the method with $\tilde{c}_\star := 0.01$ in (3.6).

Let us now use these specific multipliers to motivate our requirements in Definition 3.1.

To prove the first stability result for method (1.5) by the energy technique, we subtract and add the term $c_\star A\Big(u^{n+7} - \sum_{j=1}^4 \mu_j u^{n+7-j}\Big)$ from and to its left-hand side, and subsequently test by $u^{n+7} - \mu_1 u^{n+6} - \cdots - \mu_7 u^n = u^{n+7} - \mu_1 u^{n+6} - \cdots - \mu_4 u^{n+3}$ to obtain

$$\Big(\sum_{i=0}^{7} \alpha_i u^{n+i}, u^{n+7} - \sum_{j=1}^{4} \mu_j u^{n+7-j}\Big) + \tau A_{n+7} + \tau c_\star \Big\| u^{n+7} - \sum_{j=1}^{4} \mu_j u^{n+7-j} \Big\|^2 = \tau F_{n+7}, \tag{4.2}$$

$n = 0, \dots, N-7$, with

$$\begin{cases} A_{n+7} := \Big\langle 3u^{n+7} - 2u^{n+6} - c_\star\Big(u^{n+7} - \displaystyle\sum_{j=1}^{4} \mu_j u^{n+7-j}\Big), u^{n+7} - \displaystyle\sum_{j=1}^{4} \mu_j u^{n+7-j} \Big\rangle, \\ F_{n+7} := \Big(3f^{n+7} - 2f^{n+6}, u^{n+7} - \displaystyle\sum_{j=1}^{4} \mu_j u^{n+7-j}\Big). \end{cases}$$

The term F_{n+7} in (4.2) can be easily estimated from above via elementary inequalities. Subsequently, the term involving the approximate solutions will be absorbed in the third term on the left-hand side of (4.2); this is the motivation for the use of a positive constant $c_\star$ in (3.4) or in (3.2).

To estimate the first term in (4.2) from below, we shall first prove that the pair of polynomials α and μ, with α given in (1.2) for $\vartheta = 3$,

$$420\alpha(\zeta) = 3147\zeta^7 - 10038\zeta^6 + 15750\zeta^5 - 16800\zeta^4 + 12425\zeta^3 - 5922\zeta^2 + 1638\zeta - 200, \tag{4.3}$$

and μ the polynomial associated to the multipliers in (4.1),

$$\mu(\zeta) = \zeta^7 - 1.6\zeta^6 + 1.6\zeta^5 - 1.1\zeta^4 + 0.3\zeta^3, \tag{4.4}$$

satisfy the conditions in Definition 3.1; see Proposition 4.1. This fact in combination with Lemma 3.1 will enable us to utilize a relation of the form (G).

Analogously, to estimate A_{n+7} from below, in view of its specific form and, in particular, the fact that it depends only on five consecutive approximations, namely on $u^{n+3}, \dots, u^{n+7}$, it suffices to use polynomials of degree 4. Thus, we factor ζ^3 out of the polynomials β in (1.2) for $\vartheta = 3$ and μ and consider the polynomials δ and κ,

$$\delta(\zeta) := 3\zeta^4 - 2\zeta^3, \tag{4.5}$$

and

$$\kappa(\zeta) := \zeta^4 - \frac{8}{5}\zeta^3 + \frac{8}{5}\zeta^2 - \frac{11}{10}\zeta + \frac{3}{10} = \big(\zeta - \frac{3}{5}\big)\big(\zeta^3 - \zeta^2 + \zeta - \frac{1}{2}\big); \tag{4.6}$$

cf. (3.3). Now, to take advantage of a relation of the form (G), given that the polynomials $\delta - c_\star\kappa$ and κ enter into the first and second arguments in the inner product in A_{n+7}, we need to prove that the pair of polynomial $(\delta - c_\star\kappa, \kappa)$ satisfies the conditions in Lemma 3.1; obviously, these conditions can be reformulated in the form that the polynomials δ and κ have no common divisor and satisfy condition (3.4). We shall prove these properties in Proposition 4.2.

Proposition 4.1 (Polynomials α and μ satisfy condition (3.1))**.** *The polynomials α for $\vartheta = 3$ of (4.3) and μ of (4.4) do not have common divisor and satisfy condition (3.1).*

Proof. First, $\mu(\zeta) = \zeta^3(\zeta - 3/5)\tilde\kappa(\zeta)$ with $\tilde\kappa(\zeta) := \zeta^3 - \zeta^2 + \zeta - 1/2$; see (4.4) and (4.6). Thus, to show that the roots of μ are inside the unit disk, it suffices to show that this is the case for $\tilde\kappa$. Now,

$$\tilde\kappa\big(\frac{1}{2}\big) = -\frac{1}{8} < 0 \quad \text{and} \quad \tilde\kappa(1) = \frac{1}{2} > 0,$$

and thus $\tilde\kappa$ has a real root $\zeta_1 \in (1/2, 1)$. Actually, this is the only real root of $\tilde\kappa$, since $\tilde\kappa$ is strictly increasing on the real axis,

$$\tilde\kappa'(x) = 3x^2 - 2x + 1 = 2x^2 + (x-1)^2 > 0.$$

Let ζ_2, ζ_3 be the complex conjugate roots of $\tilde\kappa$. Then, according to Vieta's formulas,

$$\zeta_1\zeta_2\zeta_3 = \zeta_1|\zeta_2|^2 = \frac{1}{2},$$

which, in combination with $\zeta_1 > 1/2$, implies $|\zeta_2| < 1$. Thus, $|\zeta_1|, |\zeta_2|, |\zeta_3| < 1$. We infer that all roots of μ are inside the unit disk.

The generating polynomial α of the WSBDF7 method (1.5) is

$$420\alpha(\zeta) = 3147\zeta^7 - 10038\zeta^6 + 15750\zeta^5 - 16800\zeta^4 + 12425\zeta^3 - 5922\zeta^2 + 1638\zeta - 200;$$

see (4.3). First, $\alpha(0) = -10/21$, $\alpha(3/5) = 53/11822$. Furthermore, $420\alpha(\zeta) = (\zeta - 1)\tilde\alpha(\zeta)$ with

$$\begin{aligned}\tilde\alpha(\zeta) &:= 3147\zeta^6 - 6891\zeta^5 + 8859\zeta^4 - 7941\zeta^3 + 4484\zeta^2 - 1438\zeta + 200\\ &= \Big(3147\zeta^3 - 3744\zeta^2 + 1968\zeta - \frac{1311}{2}\Big)\tilde\kappa(\zeta) - \frac{1}{4}\nu(\zeta)\end{aligned}$$

and $\nu(\zeta) := 46\zeta^2 - 806\zeta + 511$, and it is easy to check that none of the roots of the quadratic polynomial ν is a root of $\tilde\kappa$; consequently, the polynomials $\tilde\alpha$ and $\tilde\kappa$ do not have common divisor. We then easily infer that the polynomials α and μ do not have common divisor.

Now, the rational function α/μ is holomorphic outside the unit disk in the complex plane and

$$\lim_{|z|\to\infty} \frac{\alpha(z)}{\mu(z)} = \alpha_7 = \frac{1049}{140} > 0.$$

Therefore, according to the maximum principle for harmonic functions, the A-stability property (A) is equivalent to

$$\operatorname{Re}\frac{\alpha(\zeta)}{\mu(\zeta)} \geqslant 0 \quad \forall \zeta \in \mathscr{K},$$

with $\mathscr{K}$ the unit circle in the complex plane, i.e., equivalent to

$$\operatorname{Re}\big[\alpha(\mathrm{e}^{\mathrm{i}\varphi})\mu(\mathrm{e}^{-\mathrm{i}\varphi})\big] \geqslant 0 \quad \forall \varphi \in \mathbb{R}.$$

In view of (4.4), this property takes the form

$$\operatorname{Re}\big[420\alpha(\mathrm{e}^{\mathrm{i}\varphi})\mathrm{e}^{-\mathrm{i}3\varphi}\kappa(\mathrm{e}^{-\mathrm{i}\varphi})\big] \geqslant 0 \quad \forall \varphi \in \mathbb{R}. \tag{4.7}$$

Now, it is easily seen that

$$\begin{aligned}420\alpha(\mathrm{e}^{\mathrm{i}\varphi})\mathrm{e}^{-\mathrm{i}3\varphi} &= 3147\cos(4\varphi) - 10238\cos(3\varphi) + 17388\cos(2\varphi) - 22722\cos\varphi + 12425\\ &\quad + \mathrm{i}\big[3147\sin(4\varphi) - 9838\sin(3\varphi) + 14112\sin(2\varphi) - 10878\sin\varphi\big].\end{aligned}$$

With $x := \cos\varphi$, recalling the elementary trigonometric identities

$$\begin{aligned}&\cos(2\varphi) = 2x^2 - 1, \quad &&\cos(3\varphi) = 4x^3 - 3x, \quad &&\cos(4\varphi) = 8x^4 - 8x^2 + 1,\\ &\sin(2\varphi) = 2x\sin\varphi, \quad &&\sin(3\varphi) = (4x^2 - 1)\sin\varphi, \quad &&\sin(4\varphi) = (8x^3 - 4x)\sin\varphi,\end{aligned}$$

we see that

$$\begin{aligned}420\alpha(\mathrm{e}^{\mathrm{i}\varphi})\mathrm{e}^{-\mathrm{i}3\varphi} &= 8(1-x)(-3147x^3 + 1972x^2 + 772x - 227)\\ &\quad + \mathrm{i}4(6294x^3 - 9838x^2 + 3909x - 260)\sin\varphi.\end{aligned} \tag{4.8}$$

Notice that the factor $1 - x$ in the real part of $\alpha(\mathrm{e}^{\mathrm{i}\varphi})\mathrm{e}^{-\mathrm{i}3\varphi}$ is due to the fact that $\alpha(1) = 0$. Similarly,

$$\begin{aligned}\kappa(\mathrm{e}^{-\mathrm{i}\varphi}) = &\ \cos(4\varphi) - 1.6\cos(3\varphi) + 1.6\cos(2\varphi) - 1.1\cos\varphi + 0.3\\ &- \mathrm{i}\big[\sin(4\varphi) - 1.6\sin(3\varphi) + 1.6\sin(2\varphi) - 1.1\sin\varphi\big]\end{aligned}$$

and

$$\kappa(\mathrm{e}^{-\mathrm{i}\varphi}) = \frac{1}{10}(80x^4 - 64x^3 - 48x^2 + 37x - 3) - \mathrm{i}\frac{1}{10}(80x^3 - 64x^2 - 8x + 5)\sin\varphi. \tag{4.9}$$

In view of (4.8) and (4.9), the desired property (4.7) can be written in the form

$$\frac{2}{5}(1 - x)P(x) \geqslant 0 \quad \forall x \in [-1, 1]$$

with

$$\begin{aligned}P(x) := &\ 2(-3147x^3 + 1972x^2 + 772x - 227)(80x^4 - 64x^3 - 48x^2 + 37x - 3)\\ &+ (1 + x)(6294x^3 - 9838x^2 + 3909x - 260)(80x^3 - 64x^2 - 8x + 5),\end{aligned}$$

i.e.,

$$P(x) = 32000x^6 - 124640x^5 + 173872x^4 - 104870x^3 + 24891x^2 - 1105x + 62. \tag{4.10}$$

Now, P is positive in the interval $[-1, 1]$, and thus (4.7) is valid. Indeed, first, all terms are positive for $-1 \leqslant x < 0$, whence P is positive in $[-1, 0)$. For $0 \leqslant x \leqslant 1$, see Figure 4.1. □

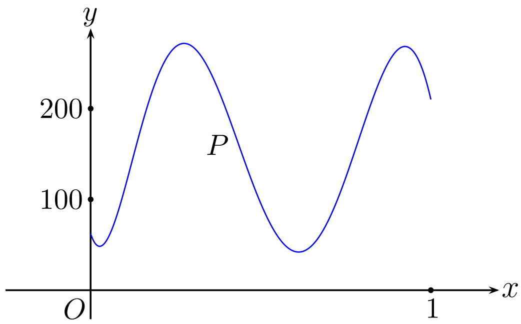

FIGURE 4.1. The graph of polynomial P of (4.10) in the interval $[0,1]$.

Proposition 4.2 (Polynomials δ and κ satisfy condition (3.4))**.** *The polynomials δ and κ of (4.5) and (4.6) do not have common divisor and satisfy the analogue of condition (3.6) with $\tilde c_\star = 0.01$, and thus condition (3.4) for some positive constant $c_\star$. In addition, the dual to (3.4) is also valid,*

$$\operatorname{Re}\frac{\kappa(\zeta)}{\delta(\zeta)} > \hat c_\star \quad \text{for } |\zeta|>1, \tag{4.11}$$

for some positive constant $\hat c_\star$.

Proof. First, $\kappa(0)=3/10$ and $\kappa(2/3)=1/810$, whence the polynomials δ and κ have no common divisor.

Since the roots of κ are inside the unit disk, the rational function δ/κ is holomorphic outside the unit disk in the complex plane; see the proof of Proposition 4.1. Furthermore,

$$\lim_{|z|\to\infty}\frac{\delta(z)}{\kappa(z)} = 3 \geqslant c_\star.$$

Therefore, according to the maximum principle for harmonic functions, the A-stability property (3.4) is equivalent to

$$\operatorname{Re}\frac{\delta(\zeta)}{\kappa(\zeta)} \geqslant c_\star \quad \forall \zeta\in\mathscr{K},$$

that is, equivalent to

$$\operatorname{Re}\big[(3\mathrm{e}^{\mathrm{i}2\varphi}-2\mathrm{e}^{\mathrm{i}\varphi})\mathrm{e}^{\mathrm{i}2\varphi}\kappa(\mathrm{e}^{-\mathrm{i}\varphi})\big] \geqslant c_\star|\kappa(\mathrm{e}^{\mathrm{i}\varphi})|^2 \quad \forall\varphi\in\mathbb{R}.$$

Thus, it suffices to show that

$$\operatorname{Re}\big[(3\mathrm{e}^{\mathrm{i}2\varphi}-2\mathrm{e}^{\mathrm{i}\varphi})\mathrm{e}^{\mathrm{i}2\varphi}\kappa(\mathrm{e}^{-\mathrm{i}\varphi})\big] \geqslant \tilde c_\star = 0.01 \quad \forall\varphi\in\mathbb{R}. \tag{4.12}$$

With $x := \cos\varphi$, recalling the elementary trigonometric identities

$$\cos(2\varphi) = 2x^2-1, \quad \sin(2\varphi) = 2x\sin\varphi,$$

we easily see that

$$\begin{aligned} 3\mathrm{e}^{\mathrm{i}2\varphi}-2\mathrm{e}^{\mathrm{i}\varphi} &= 3\cos(2\varphi)-2\cos\varphi+\mathrm{i}\big(3\sin(2\varphi)-2\sin\varphi\big)\\ &= 6x^2-2x-3+\mathrm{i}2(3x-1)\sin\varphi. \end{aligned} \tag{4.13}$$

Similarly,

$$\begin{aligned} \mathrm{e}^{\mathrm{i}2\varphi}\kappa(\mathrm{e}^{-\mathrm{i}\varphi}) &= 1.3\cos(2\varphi)-2.7\cos\varphi+1.6-\mathrm{i}\big(0.7\sin(2\varphi)-0.5\sin\varphi\big)\\ &= 2.6x^2-2.7x+0.3-\mathrm{i}(1.4x-0.5)\sin\varphi. \end{aligned} \tag{4.14}$$

In view of (4.13) and (4.14), the desired property (4.12) can be written in the form

$$\begin{aligned} g(x) &:= (6x^2-2x-3)(2.6x^2-2.7x+0.3)+2(1-x^2)(3x-1)(1.4x-0.5)\\ &= 7.2x^4-15.6x^3+6.8x^2+1.7x+0.1 \geqslant \tilde{c}_\star, \quad x\in[-1,1]. \end{aligned} \tag{4.15}$$

Now, g attains its minimum 0.01379862357 in $[-1,1]$ at -0.09331476; this value of the minimum is the motivation for choosing $\tilde{c}_\star = 0.01$ in (3.6) for the multiplier $(1.6,-1.6,1.1,-0.3,0,0,0)$. Thus, (4.12) is valid. See also Figure 4.2.

Let us provide also a complete theoretical proof of (4.15). For negative and positive x, we write g in the form $g(x) = 7.2x^4-15.6x^3+0.01+g_1(x) = x^2g_2(x)+1.7x+0.1$, with

$$g_1(x) = 6.8x^2+1.7x+0.09 \quad \text{and} \quad g_2(x) = 7.2x^2-15.6x+6.8.$$

The roots $x_{1,2}$ and $x_{3,4}$ of g_1 and g_2, respectively, are

$$x_1 = -0.17388461, \quad x_2 = -0.07611538, \quad x_3 = 0.60461977, \quad x_4 = 1.56204688.$$

Therefore, g_1 and g_2, respectively, are positive outside the intervals $[x_1,x_2]$ and $[x_3,x_4]$, and we easily see that $g(x)\geqslant 0.01$ in $[-1,-0.17388461]$ and in $[-0.07611538, 0.60461977]$.

Furthermore,

$$g'(x) = 28.8x^3-46.8x^2+13.6x+1.7 \quad \text{and} \quad g''(x) = 86.4x^2-93.6x+13.6.$$

The roots of g'' are $x_5 = 0.17289116$ and $x_6 = 0.91044216$. Therefore, g'' is negative in $[0.60461977, x_6]$ and positive in $[x_6,1]$, whence g' is decreasing in $[0.60461977, x_6]$ and increasing in $[x_6,1]$. Since $g'(0.60461977) = -0.82001368$ and $g'(1) = -2.7$, we see that g' is negative in $[0.60461977,1]$, whence g is decreasing in $[0.60461977,1]$. Therefore, $g(x)\geqslant g(1) = 0.2$ for $x\in[0.60461977,1]$. Summarizing, up to now, we proved that (4.15) is valid in $[-1,-0.17388461]$ and in $[-0.07611538,1]$.

Finally, let us write g in the form $g(x) = x^2g_3(x)+g_4(x)+0.01$ with

$$g_3(x) = 7.2x^2-15.6x-1.228 \quad \text{and} \quad g_4(x) = 8.028x^2+1.7x+0.09.$$

The function g_3 is obviously decreasing in $[-1,-0.07611538]$. Since $g_3(-0.07611538) = 0.00111349$, and the discriminant of g_4 is $-8\cdot 10^{-5}$, we see that g_3 and g_4 are positive in $[-1,-0.07611538]$. Therefore, $g(x)>0.01$ for $x\in[-1,-0.07611538]$. This completes the proof of (4.15).

The roots of δ are $\zeta_1 = 0$ and $\zeta_2 = 2/3$, whence the rational function κ/δ is holomorphic outside the unit disk in the complex plane. Therefore, (4.11), for $\hat{c}_\star\leqslant 1/3$, is equivalent to

$$\operatorname{Re}\frac{\kappa(\bar\zeta)}{\delta(\bar\zeta)} \geqslant \hat{c}_\star \quad \forall \zeta\in\mathscr{K},$$

that is, equivalent to

$$\operatorname{Re}\big[(3\mathrm{e}^{\mathrm{i}2\varphi}-2\mathrm{e}^{\mathrm{i}\varphi})\mathrm{e}^{\mathrm{i}2\varphi}\kappa(\mathrm{e}^{-\mathrm{i}\varphi})\big] \geqslant \hat{c}_\star|\delta(\mathrm{e}^{\mathrm{i}\varphi})|^2 \quad \forall\varphi\in\mathbb{R}.$$

Obviously, $|\delta(\mathrm{e}^{\mathrm{i}\varphi})|\leqslant 5$. We have already seen in (4.12) that the function on the left-hand side is strictly positive, and easily infer that this inequality is indeed valid for some positive constant $\hat{c}_\star$. □

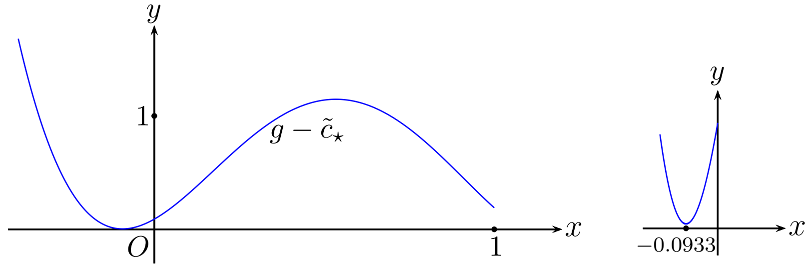

FIGURE 4.2. The graph of the polynomial $g-\tilde c_\star$ of (4.15) in the interval $[-0.4,1]$, left, and zoom in in the interval $[-0.17,0]$, right.

5. On the determination of multipliers

In this section, the objective is the derivation of necessary conditions for multipliers for the WSBDF7 method with $\vartheta=3$ such that $\mu_5=\mu_6=\mu_7=0$; we utilized these conditions to determine the multipliers (4.1). Let us mention that multipliers with $\mu_4=\mu_5=\mu_6=\mu_7=0$ do not exist; see Remark 5.1.

In the case $\mu_5=\mu_6=\mu_7=0$, we have $\mu(\zeta)=\zeta^3\kappa(\zeta)$ with $\kappa(\zeta)=\zeta^4-\mu_1\zeta^3-\mu_2\zeta^2-\mu_3\zeta-\mu_4$, and, provided that the roots of κ lie in the unit disk, the A-stability condition (3.1) takes the form

$$4(1-x)P(x)\geqslant 0\quad \forall x\in[-1,1]$$

with

$$\begin{aligned}P(x):=\ &2(-3147x^3+1972x^2+772x-227)\cdot\\ &\cdot\big(8x^4-8x^2+1-\mu_1(4x^3-3x)-\mu_2(2x^2-1)-\mu_3x-\mu_4\big)\\ &+(1+x)(6294x^3-9838x^2+3909x-260)\big(8x^3-4x-\mu_1(4x^2-1)-2\mu_2x-\mu_3\big),\end{aligned}$$

i.e.,

$$\begin{aligned}P(x)=\ &3200x^6-9904x^5+8184x^4+2990x^3-7020x^2+2584x-454\\ &+(-1600x^5+4952x^4-4492x^3-257x^2+2287x-260)\mu_1\\ &+(-800x^4+2476x^3-2446x^2+2064x-454)\mu_2\\ &+(-400x^3+4385x^2-3195x+260)\mu_3\\ &+(6294x^3-3944x^2-1544x+454)\mu_4\geqslant 0\quad \forall x\in[-1,1].\end{aligned}\tag{5.1}$$

Analogously, condition (3.4), for some positive constant $c_\star$, leads to the strict inequality condition

$$g(x)>0\quad \forall x\in[-1,1]$$

with

$$\begin{aligned}g(x):=\ &\big(3(8x^4-8x^2+1)-2(4x^3-3x)\big)\cdot\\ &\cdot\big(8x^4-8x^2+1-\mu_1(4x^3-3x)-\mu_2(2x^2-1)-\mu_3x-\mu_4\big)\\ &+(1-x^2)\big(3(8x^3-4x)-2(4x^2-1)\big)\big(8x^3-4x-\mu_1(4x^2-1)-2\mu_2x-\mu_3\big),\end{aligned}$$

i.e., to

$$\begin{aligned}g(x)=\ &-2x+3+(-3x+2)\mu_1+(-6x^2+2x+3)\mu_2\\ &+(-12x^3+4x^2+9x-2)\mu_3+(-24x^4+8x^3+24x^2-6x-3)\mu_4>0\end{aligned}\tag{5.2}$$

for all $x\in[-1,1]$. Notice that the strict inequality in (5.2) implies that (3.6) is satisfied with $\tilde c_\star:=\min_{-1\leqslant x\leqslant 1}g(x)$; consequently, (3.4) is satisfied for some positive constant $c_\star$.

Necessary conditions for (5.1) and (5.2) can be derived by evaluating P and g at certain points. For instance, we obtain the following necessary condition, which we utilized to determine the multipliers (4.1).

Lemma 5.1 (Range of multipliers)**.** *If $(\mu_1,\mu_2,\mu_3,\mu_4,0,0,0)$ is a multiplier of the WSBDF7 method with $\vartheta=3$, then there holds*

$$\begin{aligned}&1.5561 \leqslant \mu_1 < 2.3133, \quad && -2.2024 < \mu_2 < -1.4259,\\ &0.5394 < \mu_3 < 1.3955, \quad && -0.6518 < \mu_4 < -0.0504.\end{aligned}$$

Proof. First,

$$\begin{aligned}&g(1) = 1-\mu_1-\mu_2-\mu_3-\mu_4, \qquad && 2g(-1/2) = 8+7\mu_1+\mu_2-8\mu_3+7\mu_4,\\ &g(0) = 2(\mu_1-\mu_3)-3(\mu_4-\mu_2-1).\end{aligned}$$

Furthermore,

$$g(0)+2g(1) = 5+\mu_2-4\mu_3-5\mu_4 > 0, \tag{5.3}$$

$$7g(1)+2g(-1/2) = 15-6\mu_2-15\mu_3 > 0. \tag{5.4}$$

Also,

$$10^{-3}P(0.999999)/0.21 = -2+3\mu_1+4\mu_2+5\mu_3+6\mu_4,$$

whence

$$3g(1)+10^{-3}P(0.999999)/0.21 = 1+\mu_2+2\mu_3+3\mu_4 > 0. \tag{5.5}$$

Multiplying (5.5) by 2 and adding the result to (5.3), we get $7+3\mu_2+\mu_4>0$. Adding the positive quantities

$$\begin{aligned}g(0.214929)/1.3552 &= 1.8965+\mu_1+2.3264\mu_2-2.3264\mu_4,\\ g(0.941785)/0.8254 &= 1.3526-\mu_1-0.5309\mu_2+0.5309\mu_4,\end{aligned}$$

we obtain $3.2491+1.7955\mu_2-1.7955\mu_4>0$, which, in combination with $7+3\mu_2+\mu_4>0$, yields $15.8176+7.1820\mu_2>0$, i.e., $\mu_2>-2.2024$. Similarly, adding the positive quantities

$$\begin{aligned}g(0.214929)/3.1527 &= 0.8152+0.4299\mu_1+\mu_2-\mu_4,\\ g(0.941785)/0.4382 &= 2.5477-1.8836\mu_1-\mu_2+\mu_4,\end{aligned}$$

we obtain $3.3629-1.4537\mu_1>0$, i.e., $\mu_1<2.3133$.

Summation of the nonnegative quantities

$$\begin{aligned}P(0.68481813)/431.5990 &= -0.8759+1.3696\mu_1+\mu_2-\mu_4,\\ P(0.09319637)/280.9437 &= -0.9653-0.1864\mu_1-\mu_2+\mu_4,\end{aligned}$$

leads to $-1.8412+1.1832\mu_1\geqslant 0$, i.e., $\mu_1\geqslant 1.5561$.

Adding the positive quantities

$$\begin{aligned}g(-0.264464)/2.7934 &= 1.2633+\mu_1+0.7344\mu_2-1.3884\mu_3,\\ g(0.962266)/0.8868 &= 1.2128-\mu_1-0.7118\mu_2-0.3699\mu_3,\end{aligned}$$

we obtain $2.4761+0.0226\mu_2-1.7583\mu_3>0$, which, in combination with (5.4), yields $15.1956-10.8888\mu_3>0$, i.e., $\mu_3<1.3955$.

Also, summation of the nonnegative quantities

$$\begin{aligned}P(-1/9)/510.3402 &= -1.6273-\mu_1-1.4050\mu_2+1.3122\mu_3+1.1134\mu_4,\\ P(1/2)/517.2500 &= -0.5631+\mu_1+0.4369\mu_2-0.5631\mu_3-\mu_4,\end{aligned}$$

yields $-2.1904 - 0.9681\mu_2 + 0.7491\mu_3 + 0.1134\mu_4 \geqslant 0$, which, together with (5.3), yields $-5.0152 - 3.1232\mu_2 - 3.2924\mu_4 > 0$. Similarly, adding the nonnegative quantities

$$P(0.68481813)/591.1336 = -0.6395 + \mu_1 + 0.7301\mu_2 - 0.7301\mu_4,$$
$$P(0.09319637)/52.3659 = -5.1786 - \mu_1 - 5.3650\mu_2 + 5.3650\mu_4,$$

we have $-5.8181 - 4.6349\mu_2 + 4.6349\mu_4 \geqslant 0$, which, together with $-5.0152 - 3.1232\mu_2 - 3.2924\mu_4 > 0$, yields $-42.4005 - 29.7357\mu_2 > 0$, i.e., $\mu_2 < -1.4259$.

From the positivity of $g(1)$ and

$$g(-0.26)/2.7800 = 1.2662 + \mu_1 + 0.7462\mu_2 - 1.3880\mu_3 - 0.0244\mu_4 > 0,$$

we have $2.2662 - 0.2538\mu_2 - 2.3880\mu_3 - 1.0244\mu_4 > 0$, which, together with (5.5), yields $6.9204+1.8804\mu_2+5.1152\mu_4 > 0$. Combining the latter condition with $-5.8181-4.6349\mu_2+4.6349\mu_4 \geqslant 0$, we obtain $21.1350 + 32.4239\mu_4 > 0$ i.e., $\mu_4 > -0.6518$.

Adding the nonnegative quantities

$$P(0.81865385)/611.3646 = -0.6632 + \mu_1 + 0.9741\mu_2 + 0.5950\mu_3,$$
$$10^{-3}P(-0.407988)/0.7755 = -3.2576 - \mu_1 - 2.4417\mu_2 + 2.9924\mu_3,$$

we get $-3.9208 - 1.4676\mu_2 + 3.5874\mu_3 \geqslant 0$. Similarly, adding the nonnegative and positive, respectively, quantities

$$P(-1/9)/510.3402 = -1.6273 - \mu_1 - 1.4050\mu_2 + 1.3122\mu_3 + 1.1134\mu_4,$$
$$g(1/33)/1.9091 = 1.5397 + \mu_1 + 1.6003\mu_2 - 0.9030\mu_3 - 1.6550\mu_4,$$

we have $-0.0876 + 0.1952\mu_2 + 0.4092\mu_3 - 0.5416\mu_4 > 0$, which, together with (5.5), yields $0.2788+1.1272\mu_2+2.3108\mu_3 > 0$. Combining the latter relation with the already established relation $-3.9208 - 1.4676\mu_2 + 3.5874\mu_3 \geqslant 0$, we get $-4.0104 + 7.4350\mu_3 > 0$, i.e., $\mu_3 > 0.5394$.

Combining $-2.1904 - 0.9681\mu_2 + 0.7491\mu_3 + 0.1134\mu_4 \geqslant 0$ with $-0.0876 + 0.1952\mu_2 + 0.4092\mu_3 - 0.5416\mu_4 > 0$, we obtain $-0.5124+0.5424\mu_3-0.5022\mu_4 > 0$. From the positivity of $g(1)$ and the nonnegativity of

$$P(0.7)/594.2772 = -0.6437 + \mu_1 + 0.7563\mu_2 + 0.0588\mu_3 - 0.6740\mu_4,$$

we get $0.3563 - 0.2437\mu_2 - 0.9412\mu_3 - 1.6740\mu_4 > 0$. Combined with (5.3), the latter relation yields $1.5748-1.9160\mu_3-2.8925\mu_4 > 0$, which together with $-0.5124+0.5424\mu_3-0.5022\mu_4 > 0$ leads to $-0.1276 - 2.5311\mu_4 > 0$, i.e., $\mu_4 < -0.0504$. The proof is complete. □

Remark 5.1 (Nonexistence of multipliers of the form $(\mu_1, \mu_2, \mu_3, 0, 0, 0, 0)$)**.** Our first attempt was to determine multipliers of the form $(\mu_1, \mu_2, \mu_3, 0, 0, 0, 0)$. That such multipliers do not exist follows immediately from Lemma 5.1 since μ_4 must be negative.

6. Stability

Here we prove the stability estimates, Theorems 1.1 and 1.2, and Corollary 1.1.

6.1. **Proof of Theorem 1.1.** According to Propositions 4.1 and 4.2, respectively, in combination with Lemma 3.1, there exist two positive definite symmetric matrices $G = (g_{ij}) \in \mathbb{R}^{7,7}$ and $\widetilde{G} = (\tilde{g}_{ij}) \in \mathbb{R}^{4,4}$ such that with the notation $\mathcal{U}^n := (u^{n-6}, \dots, u^n)^\top$ and $\boldsymbol{U}^n := (u^{n-3}, \dots, u^n)^\top$, and the norms $|\mathcal{U}^n|_G$ and $\|\boldsymbol{U}^n\|_{\widetilde{G}}$ given by

$$|\mathcal{U}^n|_G^2 = \sum_{i,j=1}^{7} g_{ij}(u^{n-7+i}, u^{n-7+j}), \quad \|\boldsymbol{U}^n\|_{\widetilde{G}}^2 = \sum_{i,j=1}^{4} \tilde{g}_{ij}\langle u^{n-4+i}, u^{n-4+j}\rangle, \tag{6.1}$$

there holds

$$\Big(\sum_{i=0}^{7}\alpha_i u^{n+i}, u^{n+7}-\sum_{j=1}^{4}\mu_j u^{n+7-j}\Big) \geqslant |\mathcal{U}^{n+7}|_G^2 - |\mathcal{U}^{n+6}|_G^2 \tag{6.2}$$

and

$$A_{n+7} \geqslant \|\boldsymbol{U}^{n+7}\|_{\widetilde{G}}^2 - \|\boldsymbol{U}^{n+6}\|_{\widetilde{G}}^2. \tag{6.3}$$

Utilizing (6.2) and (6.3), we infer from (4.2) that

$$|\mathcal{U}^{n+7}|_G^2 - |\mathcal{U}^{n+6}|_G^2 + \tau\|\boldsymbol{U}^{n+7}\|_{\widetilde{G}}^2 - \tau\|\boldsymbol{U}^{n+6}\|_{\widetilde{G}}^2 + c_\star\tau\Big\|u^{n+7}-\sum_{j=1}^{4}\mu_j u^{n+7-j}\Big\|^2 \leqslant \tau F_{n+7}. \tag{6.4}$$

Furthermore, the terms involving the forcing term can be estimated by elementary inequalities in the form

$$\begin{aligned} F_{n+7} &\leqslant \big\|3f^{n+7}-2f^{n+6}\big\|_\star \Big\|u^{n+7}-\sum_{j=1}^{4}\mu_j u^{n+7-j}\Big\| \\ &\leqslant \frac{1}{4c_\star}\|3f^{n+7}-2f^{n+6}\|_\star^2 + c_\star\Big\|u^{n+7}-\sum_{j=1}^{4}\mu_j u^{n+7-j}\Big\|^2. \end{aligned}$$

Using this estimate in (6.4) and summing over n, from $n=0$ to $n=m-7$, we obtain

$$|\mathcal{U}^{m}|_G^2 - |\mathcal{U}^{6}|_G^2 + \tau\|\boldsymbol{U}^{m}\|_{\widetilde{G}}^2 - \tau\|\boldsymbol{U}^{6}\|_{\widetilde{G}}^2 \leqslant \frac{\tau}{4c_\star}\sum_{n=7}^{m}\|3f^n-2f^{n-1}\|_\star^2. \tag{6.5}$$

Now, with c_1 and c_2 the smallest eigenvalues of the matrices G and $\widetilde{G}$, respectively, we have

$$|\mathcal{U}^{m}|_G^2 \geqslant c_1|u^m|^2,\ \|\boldsymbol{U}^{m}\|_{\widetilde{G}}^2 \geqslant c_2\|u^m\|^2,\ |\mathcal{U}^{6}|_G^2 \leqslant C\sum_{j=0}^{6}|u^j|^2,\ \|\boldsymbol{U}^{6}\|_{\widetilde{G}}^2 \leqslant C\sum_{j=0}^{6}\|u^j\|^2.$$

Thus, (6.5) yields

$$|u^m|^2+\tau\|u^m\|^2 \leqslant C\sum_{j=0}^{6}\big(|u^j|^2+\tau\|u^j\|^2\big) + C\tau\sum_{n=7}^{m}\|3f^n-2f^{n-1}\|_\star^2.$$

The asserted result (1.9) is an obvious consequence of this estimate.

6.2. **Proof of Theorem 1.2.** With the notation (1.10), we write the WSBDF7 method (1.5) in the form

$$\dot u^n + 3Au^n - 2Au^{n-1} = 3f^n - 2f^{n-1},\quad n=7,\dots,N. \tag{6.6}$$

Let us introduce the notation

$$v^m := 3u^m - 2u^{m-1} \quad\text{and}\quad g^m := 3f^m - 2f^{m-1},\quad m=1,\dots,N, \tag{6.7}$$

and write (6.6) as

$$\dot u^n + Av^n = g^n,\quad n=7,\dots,N. \tag{6.8}$$

For $n \geqslant 11$, to take advantage of the properties of the multiplier (4.1), we consider method (6.8) with n replaced by $n-j$, multiply it by $\mu_j, j = 1,2,3,4$, and subtract the resulting relations from (6.8), to obtain

$$\dot{u}^n - \sum_{j=1}^{4} \mu_j \dot{u}^{n-j} + A\Big(v^n - \sum_{j=1}^{4} \mu_j v^{n-j}\Big) = g^n - \sum_{j=1}^{4} \mu_j g^{n-j}, \quad n = 11, \dots, N. \tag{6.9}$$

Now, we subtract and add the term $\hat{c}_\star(3\dot{u}^n - 2\dot{u}^{n-1})$ from and to the left-hand side of (6.9), and subsequently test the relation by

$$\tau(3\dot{u}^n - 2\dot{u}^{n-1}) = \sum_{i=0}^{7} \alpha_i v^{n-7+i},$$

to obtain

$$\tau I_n + \hat{c}_\star \tau |3\dot{u}^n - 2\dot{u}^{n-1}|^2 + \Big\langle \sum_{i=0}^{7} \alpha_i v^{n-7+i}, v^n - \sum_{j=1}^{4} \mu_j v^{n-j} \Big\rangle = \tau G_n, \quad n = 11, \dots, N, \tag{6.10}$$

with

$$\begin{cases} I_n := \Big(\dot{u}^n - \sum_{j=1}^{4} \mu_j \dot{u}^{n-j} - \hat{c}_\star(3\dot{u}^n - 2\dot{u}^{n-1}), 3\dot{u}^n - 2\dot{u}^{n-1}\Big), \\ G_n := \Big(g^n - \sum_{j=1}^{4} \mu_j g^{n-j}, 3\dot{u}^n - 2\dot{u}^{n-1}\Big). \end{cases}$$

In view of (4.11), the pair of polynomials $(\kappa - \hat{c}_\star \delta, \delta)$ satisfies the A-stability condition (A) in Lemma 3.1; let us denote by $\widehat{G} = (\hat{g}_{ij}) \in \mathbb{R}^{4,4}$ the corresponding positive definite symmetric matrix entering into the analogue to (G) for this pair of polynomials.

With the notation $\mathcal{V}^n := (v^{n-6}, \dots, v^n)^\top$, $\dot{\boldsymbol{U}}^n := (\dot{u}^{n-3}, \dots, \dot{u}^n)^\top$, and the norms $\|\cdot\|_G$ and $|\cdot|_{\widehat{G}}$, given, in analogy to (6.1), by

$$\|\mathcal{V}^n\|_G^2 = \sum_{i,j=1}^{7} g_{ij} \langle v^{n-7+i}, v^{n-7+j} \rangle, \quad |\dot{\boldsymbol{U}}^n|_{\widehat{G}}^2 = \sum_{i,j=1}^{4} \hat{g}_{ij} (\dot{u}^{n-4+i}, \dot{u}^{n-4+j}),$$

we have

$$\Big\langle \sum_{i=0}^{7} \alpha_i v^{n-7+i}, v^n - \sum_{j=1}^{4} \mu_j v^{n-j} \Big\rangle \geqslant \|\mathcal{V}^n\|_G^2 - \|\mathcal{V}^{n-1}\|_G^2 \tag{6.11}$$

and

$$I_n \geqslant |\dot{\boldsymbol{U}}^n|_{\widehat{G}}^2 - |\dot{\boldsymbol{U}}^{n-1}|_{\widehat{G}}^2; \tag{6.12}$$

cf. (6.2).

Now, in view of (6.11) and (6.12), relation (6.10) yields

$$\|\mathcal{V}^n\|_G^2 - \|\mathcal{V}^{n-1}\|_G^2 + \tau\big(|\dot{\boldsymbol{U}}^n|_{\widehat{G}}^2 - |\dot{\boldsymbol{U}}^{n-1}|_{\widehat{G}}^2\big) + \hat{c}_\star \tau |3\dot{u}^n - 2\dot{u}^{n-1}|^2 \leqslant \tau G_n. \tag{6.13}$$

Furthermore, the terms involving the forcing term can be estimated by elementary inequalities in the form

$$\begin{aligned} G_n &\leqslant \Big| g^n - \sum_{j=1}^4 \mu_j g^{n-j} \Big| \, |3\dot u^n - 2\dot u^{n-1}| \\ &\leqslant \frac{1}{4\hat c_\star} \Big| g^n - \sum_{j=1}^4 \mu_j g^{n-j} \Big|^2 + \hat c_\star |3\dot u^n - 2\dot u^{n-1}|^2. \end{aligned}$$

Using this estimate in (6.13) and summing over n, from $n = 11$ to $n = m$, we obtain

$$\|\mathcal V^m\|_G^2 - \|\mathcal V^{10}\|_G^2 + \tau\big(|\dot{\boldsymbol U}^m|_{\widehat G}^2 - |\dot{\boldsymbol U}^{10}|_{\widehat G}^2\big) \leqslant \frac{\tau}{4\hat c_\star} \sum_{n=11}^m \Big| g^n - \sum_{j=1}^4 \mu_j g^{n-j} \Big|^2,$$

and easily see that

$$\|\mathcal V^m\|_G^2 + \tau|\dot{\boldsymbol U}^m|_{\widehat G}^2 \leqslant \|\mathcal V^{10}\|_G^2 + \tau|\dot{\boldsymbol U}^{10}|_{\widehat G}^2 + C\tau \sum_{n=6}^m |f^n|^2, \quad m = 11, \dots, N.$$

From this estimate, we infer that

$$\|v^m\|^2 + \tau|\dot u^m|^2 \leqslant C\sum_{j=4}^{10} \|v^j\|^2 + C\tau \sum_{j=7}^{10} |\dot u^j|^2 + C\tau \sum_{n=6}^m |f^n|^2, \quad m = 11, \dots, N,$$

and thus

$$\|v^m\|^2 + \tau|\dot u^m|^2 \leqslant C\sum_{j=3}^{10} \|u^j\|^2 + C\tau \sum_{j=7}^{10} |\dot u^j|^2 + C\tau \sum_{n=6}^m |f^n|^2, \quad m = 11, \dots, N. \tag{6.14}$$

Let us denote by E_m the square root of the quantity on the right-hand side of (6.14). Then, for $\ell = 11, \dots, m \leqslant N$, (6.14) yields $\|v^\ell\| \leqslant E_m$, whence

$$\|u^\ell\| \leqslant \frac{2}{3}\|u^{\ell-1}\| + \frac{1}{3}E_m, \quad \ell = 11, \dots, m.$$

Iterating from $\ell = 11$ to $\ell = m$, we obtain

$$\|u^m\| \leqslant \Big(\frac{2}{3}\Big)^{m-10} \|u^{10}\| + \frac{1}{3}\sum_{j=0}^{m-11} \Big(\frac{2}{3}\Big)^j E_m,$$

and thus

$$\|u^m\| \leqslant \|u^{10}\| + E_m, \quad m = 11, \dots, N. \tag{6.15}$$

Now, (6.14) and (6.15) yield

$$\|u^m\|^2 + \tau|\dot u^m|^2 \leqslant C\sum_{j=3}^{10} \|u^j\|^2 + C\tau \sum_{j=7}^{10} |\dot u^j|^2 + C\tau \sum_{n=6}^m |f^n|^2, \quad m = 11, \dots, N. \tag{6.16}$$

In view of (6.16), to complete the proof of (1.11), it suffices to show that

$$\|u^m\|^2 + \tau|\dot u^m|^2 \leqslant c\sum_{j=0}^{6} \|u^j\|^2 + c\tau \sum_{\ell=6}^m |f^\ell|^2, \quad m = 7, 8, 9, 10. \tag{6.17}$$

This can be done via elementary inequalities; cf. [2, Appendix] and [1]. Testing (6.6) for $n = 7$ by $\dot u^7$ and using (1.10), we have

$$|\dot u^7|^2 + \frac{3\alpha_7}{\tau}\|u^7\|^2 = -\frac{3}{\tau}\sum_{i=0}^{6}\alpha_i\langle u^7, u^i\rangle + \frac{2}{\tau}\sum_{i=0}^{7}\alpha_i\langle u^6, u^i\rangle + (3f^7 - 2f^6, \dot u^7).$$

Estimating the terms on the right-hand side by the Cauchy–Schwarz and the weighted arithmetic–geometric mean inequalities, we can hide the terms involving $|\dot u^7|^2$ and $\|u^7\|^2/\tau$ to the left-hand side, and easily obtain (6.17) for $m = 7$. Then, using (6.17) for $m = 7$, we similarly obtain the asserted result for $m = 8$, and subsequently also for $m = 9, 10$.

6.3. **Proof of Corollary 1.1.** We write the characteristic polynomial α of the WSBDF7 method (1.5) in the form $\alpha(\zeta) = (\zeta - 1)\tilde\alpha(\zeta)$ with $\tilde\alpha(\zeta) = \tilde\alpha_6\zeta^6 + \cdots + \tilde\alpha_0$. The difference quotients $\partial_\tau u^m := (u^m - u^{m-1})/\tau$ satisfy then the equations

$$\sum_{i=0}^{6}\tilde\alpha_i\partial_\tau u^{n+i} = \dot u^{n+6}, \quad n = 1, \dots, N-6. \tag{6.18}$$

Since the roots of the polynomial $\tilde\alpha$ lie in the open unit disk, the rational function

$$\frac{1}{\zeta^6\tilde\alpha(1/\zeta)} = \frac{1}{\tilde\alpha_6 + \tilde\alpha_5\zeta + \cdots + \tilde\alpha_0\zeta^6}$$

is holomorphic in a disk of radius larger than 1 centered at the origin. Thus, Taylor expansion about 0 yields

$$\frac{1}{\zeta^6\tilde\alpha(1/\zeta)} = \sum_{n=0}^{\infty}\gamma_n\zeta^n, \quad |\zeta| \leqslant 1, \quad \text{where} \quad |\gamma_n| \leqslant c\gamma^n \quad \text{with } \gamma < 1.$$

An obvious consequence of this expansion are the relations

$$\tilde\alpha_6\gamma_0 = 1 \quad \text{and} \quad \sum_{j=0}^{6}\tilde\alpha_{6-j}\gamma_{n-j} = 0, \quad n \in \mathbb{N}, \quad \text{where} \quad \gamma_{-6} = \cdots = \gamma_{-1} = 0. \tag{6.19}$$

To solve the linear difference equation (6.18), we consider the corresponding equations with n replaced by $n - 6 - j, j = 0, \dots, n-7$, multiply them by γ_j and sum over j. In view of (6.19), this leads to

$$\partial_\tau u^n = -\sum_{i=1}^{6}\sum_{j=7}^{6+i}\tilde\alpha_{6-i}\gamma_{n-j}\partial_\tau u^{j-i} + \sum_{j=0}^{n-7}\gamma_j\dot u^{n-j}, \quad n = 7, \dots, N. \tag{6.20}$$

Since $|\gamma_j| \leqslant c\gamma^j$ with $\gamma < 1$, from (6.20) and (6.19), we obtain

$$|\partial_\tau u^n| \leqslant C\sum_{i=1}^{6}|\partial_\tau u^i| + C\max_{7\leqslant\ell\leqslant n}|\dot u^\ell|, \quad n = 7, \dots, N. \tag{6.21}$$

The asserted estimate (1.12) is an immediate consequence of (6.21) and (1.11).

7. Error estimates

Error estimates are easily established by combining the stability and consistency of the method.

Proposition 7.1 (Error estimates)**.** *Assume that the solution u of* (1.1) *is sufficiently smooth and that the starting approximations $u^i \in V$ to $u(t_i), i = 0, \dots, 6$, are sufficiently accurate, namely,*

$$|u(t_i) - u^i| + \tau^{1/2} \|u(t_i) - u^i\| \leqslant C\tau^7, \quad i = 0, \dots, 6. \tag{7.1}$$

Then, we have the error estimate

$$|u(t_n) - u^n| + \tau^{1/2} \|u(t_n) - u^n\| \leqslant C\tau^7, \quad n = 0, \dots, N, \tag{7.2}$$

with a constant C independent of the time step τ.

Proof. Let $d^\ell, \ell = 7, \dots, N$, denote the consistency error of the WSBDF7 method (1.5) for the initial value problem (1.1), the amount by which the exact solution misses satisfying (1.5),

$$\tau d^{n+7} = \sum_{i=0}^{7} \alpha_i u(t_{n+i}) + 3\tau A u(t_{n+7}) - 2\tau A u(t_{n+6}) - 3\tau f^{n+7} + 2\tau f^{n+6}, \tag{7.3}$$

$n = 0, \dots, N-7$, that is,

$$\tau d^{n+7} = \sum_{i=0}^{7} \alpha_i u(t_{n+i}) - 3\tau u'(t_{n+7}) + 2\tau u'(t_{n+6}). \tag{7.4}$$

An immediate consequence of the fact that the WSBDF7 method is a linear combination of two methods of order 7, namely the seven-step BDF method and the shifted seven-step BDF method, is that its order is 7, i.e.,

$$\sum_{i=0}^{7} i^\ell \alpha_i = \ell(3 \cdot 7^{\ell-1} - 2 \cdot 6^{\ell-1}), \quad \ell = 0, 1, \dots, 7; \tag{7.5}$$

actually, the consistency error of the WSBDF7 method is a linear combination of the consistency errors of the seven-step BDF and shifted seven-step BDF methods. Therefore, by Taylor expanding about t_n in (7.4), we see that, due to the order conditions (7.5), leading terms of order up to 7 cancel, and we obtain

$$\begin{aligned}\tau d^{n+7} = \frac{1}{7!}\Bigg[\sum_{i=0}^{7} \alpha_i \int_{t_n}^{t_{n+i}} (t_{n+i} - s)^7 u^{(8)}(s)\, \mathrm{d}s - 21\tau \int_{t_n}^{t_{n+7}} (t_{n+7} - s)^6 u^{(8)}(s)\, \mathrm{d}s \\ + 14\tau \int_{t_n}^{t_{n+6}} (t_{n+6} - s)^6 u^{(8)}(s)\, \mathrm{d}s \Bigg],\end{aligned}$$

$n = 0, \dots, N-7$. Thus, under obvious regularity requirements, we obtain the desired optimal order consistency estimate

$$\max_{7 \leqslant \ell \leqslant N} \|d^\ell\|_\star \leqslant C\tau^7. \tag{7.6}$$

Subtracting the numerical method (1.5) from the consistency relation (7.3), we see that the errors $e^\ell := u(t_\ell) - u^\ell, \ell = 0, \dots, N$, satisfy the error equation

$$\sum_{i=0}^{7} \alpha_i e^{n+i} + 3\tau A e^{n+7} - 2\tau A e^{n+6} = \tau d^{n+7}, \quad n = 0, \dots, N-7. \tag{7.7}$$

Now, the stability estimate (1.9) for the error equation (7.7) in combination with the consistency estimate (7.6) and our assumption (7.1) on the starting approximations lead to the asserted error estimate (7.2). ∎

8. Numerical results

We consider initial and boundary value problems for the equation

$$u_t - \Delta u + u = f \quad \text{in} \quad \Omega \times [0,T], \tag{8.1}$$

with $\Omega = (-1,1)^2$ and $T = 1$, subject to periodic boundary conditions. We used the WSBDF7 method, the standard six- and seven-step BDF methods, as well as two algebraically stable four-stage Runge–Kutta methods [6, 13], namely the Radau IIA and Gauss methods, of classical orders 7 and 8, respectively, for time stepping. In space, we discretized by the spectral collocation method with the Chebyshev–Gauss–Lobatto points.

We express the space discrete approximation u_I^n in terms of its values at the Chebyshev–Gauss–Lobatto points,

$$u_I^n(x,y) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} u_{ij}^n \ell_i(x)\ell_j(y), \qquad \ell_i(x) = \prod_{\substack{j=0\\ j\neq i}}^{N_x} \frac{x - x_j}{x_i - x_j},$$

where $u_{ij}^n := u_I^n(x_i, y_j)$ at the mesh points (x_i, y_j). Here, $-1 = x_0 < x_1 < \cdots < x_{N_x} = 1$ and $-1 = y_0 < y_1 < \cdots < y_{N_y} = 1$ are nodes of Lobatto quadrature rules. In order to test the temporal error, we fix $N_x = N_y = 30$; the spatial error is negligible since the spectral collocation method converges exponentially; see, e.g., [20, Theorem 4.4, §4.5.2].

Example 8.1. We chose the initial value and the forcing term such that the exact solution of equation (8.1) is

$$u(x,y,t) = (t^8 + 1)\sin(\pi x)\sin(\pi y).$$

We present in Table 8.1 the L^2-norm of the errors as well as the corresponding convergence orders (rates) for the WSBDF7 method with $\vartheta = 3, 5, 10$. In Tables 8.2 and 8.3, respectively, we present the L^2-norm of the errors as well as the corresponding convergence orders (rates) for the (unstable) seven-step BDF (BDF7) and the (stable) six-step BDF (BDF6) methods, and for the four-stage Radau IIA and Gauss methods.

Table 8.1. WSBDF7 methods for $\vartheta = 3, 5, 10$. The discrete L^2-norm errors and numerical convergence orders with $N_x = N_y = 30$. The CPU times are given in seconds.

τ	$\vartheta = 3$	Rate	CPU	$\vartheta = 5$	Rate	CPU	$\vartheta = 10$	Rate	CPU
1/30	4.3994e-08		1.50	7.5004e-08		1.53	1.4354e-07		1.49
1/40	6.1509e-09	6.83	2.13	1.0200e-08	6.93	2.07	2.0078e-08	6.83	2.04
1/50	1.2572e-09	7.11	2.38	2.1259e-09	7.02	2.44	4.2649e-09	6.94	2.45
1/60	3.4785e-10	7.04	2.87	5.9202e-10	7.01	3.05	1.1960e-09	6.97	2.88

Let us mention that the four-stage Radau IIA and Gauss methods achieve here their full orders 7 and 8, respectively, due to the periodic boundary conditions. As is well known, in the case of other boundary conditions these methods suffer from order reduction; see, for instance, [21, Chapter 8]. High-order Runge–Kutta methods achieve their full order for parabolic equations only under unnatural compatibility conditions, which are satisfied in the case of periodic boundary conditions and smooth solutions. The order reduction is due to the fact that the consistency error of high-order Runge–Kutta methods

TABLE 8.2. BDF7 and BDF6 methods. The discrete L^2-norm errors and numerical convergence orders with $N_x = N_y = 30$. The CPU times are given in seconds.

τ	BDF7	Rate	CPU	BDF6	Rate	CPU
1/30	6.3047e-08		1.51	4.0286e-07		1.26
1/40	1.9600e-09	12.06	1.95	7.3621e-08	5.90	1.68
1/50	3.5199e-10	7.69	2.44	1.9579e-08	5.93	2.09
1/60	7.4116e-10	-4.08	3.18	6.6204e-09	5.94	2.52

TABLE 8.3. Four-stage Radau IIA and Gauss methods. The discrete L^2-norm errors and numerical convergence orders with $N_x = N_y = 30$. The CPU times are given in seconds.

τ	Radau IIA	Rate	CPU	Gauss	Rate	CPU
1/30	1.6172e-10		122.72	7.4790e-12		119.73
1/40	2.1984e-11	6.93	160.38	7.5111e-13	7.98	156.01
1/50	4.6588e-12	6.95	199.37	1.2423e-13	8.06	202.60
1/60	1.3099e-12	6.95	244.02	2.6920e-14	8.38	234.54

cannot be expressed in terms of the exact solution only; the underlying equation enters into their consistency errors. Multistep methods, on the other hand, do not suffer from order reduction: for smooth solutions, they achieve their full order of convergence since, in contrast to Runge–Kutta methods, their consistency errors can be expressed in terms of the exact solution only.

REFERENCES

1. G. Akrivis, M. H. Chen, F. Yu, and Z. Zhou, *The energy technique for the six-step BDF method*, SIAM J. Numer. Anal. **59** (2021) 2449–2472. DOI 10.1137/21M1392656. MR4316580
2. G. Akrivis, M. Feischl, B. Kovács, and C. Lubich, *Higher-order linearly implicit full discretization of the Landau–Lifshitz–Gilbert equation*, Math. Comp. **90** (2021) 995–1038. DOI 10.1090/mcom/3597. MR4232216
3. G. Akrivis and E. Katsoprinakis, *Backward difference formulae: New multipliers and stability properties for parabolic equations*, Math. Comp. **85** (2016) 2195–2216. DOI 10.1090/mcom3055. MR3511279
4. U. M. Ascher, S. J. Ruuth, and B. T. R. Wetton, *Implicit-explicit methods for time-dependent partial differential equations*, SIAM J. Numer. Anal. **32** (1995) 797–823. DOI 10.1137/0732037. MR1335656
5. C. Baiocchi and M. Crouzeix, *On the equivalence of A-stability and G-stability*, Appl. Numer. Math. **5** (1989) 19–22. DOI 10.1016/0168-9274(89)90020-2. Recent theoretical results in numerical ordinary differential equations. MR979543
6. J. C. Butcher, *Implicit Runge-Kutta processes*, Math. Comp. **18** (1964) 50–64. DOI 10.2307/2003405. MR159424
7. M. H. Chen, F. Yu, Q. Zhang, and Z. Zhang, *Variable step-size BDF3 method for Allen-Cahn equation*, *J. Comput. Math.* **42** (2024) 1380–1406.
8. M. H. Chen, F. Yu, and Z. Zhou, *Backward difference formulae: the energy technique for subdiffusion equation*, J. Sci. Comput. **87** (2021), Paper No. 94. DOI 10.1007/s10915-021-01509-9. MR4257068
9. M. Crouzeix, *Une méthode multipas implicite-explicite pour l'approximation des équations d'évolution paraboliques*, Numer. Math. **35** (1980) 257–276 DOI 10.1007/BF01396412. MR592157
10. G. Dahlquist, *G-stability is equivalent to A-stability*, BIT **18** (1978) 384–401. DOI 10.1007/BF01932018. MR520750

11. C. M. Elliott, H. Garcke, and B. Kovács, *Numerical analysis for the interaction of mean curvature flow and diffusion on closed surfaces*, Numer. Math. **151** (2022) 873–925. DOI 10.1007/s00211-022-01301-3. MR4453294
12. E. Hairer, S. P. Nørsett, and G. Wanner: *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Series in Computational Mathematics v. 8, Springer–Verlag, Berlin, 2nd revised ed., 1993, corr. 2nd printing, 2000. DOI 10.1007/978-3-540-78862-1. MR1227985
13. E. Hairer and G. Wanner, *Solving ordinary differential equations II: stiff and differential-algebraic problems*, volume 14 of Springer Series in Computational Mathematics. Springer–Verlag, Berlin, second rev. edition, 2010, pages xvi+614. DOI 10.1007/978-3-642-05221-7. MR2657217
14. F. K. Huang and J. Shen, *A new class of implicit-explicit BDFk SAV schemes for general dissipative systems and their error analysis*, Comput. Methods Appl. Mech. Engrg. **392** (2022), Paper No. 114718. DOI 10.1016/j.cma.2022.114718. MR4383075
15. B. Kovács and B. Y. Li, *Maximal regularity of backward difference time discretization for evolving surface PDEs and its application to nonlinear problems*, IMA J. Numer. Anal. **43** (2023) 1937–1969. DOI 10.1093/imanum/drac033. MR4621836
16. Q. Li and J. Xie, *A linear multistep method for solving stiff ordinary differential equations*, J. of Tsinghua University **31** (1991) (in Chinese, with English abstract). DOI 10.16511/j.cnki.qhdxxb.1991.06.001
17. C. Lubich, D. Mansour, and C. Venkataraman, *Backward difference time discretization of parabolic differential equations on evolving surfaces*, IMA J. Numer. Anal. **33** (2013) 1365–1385. DOI 10.1093/imanum/drs044. MR3119720
18. O. Nevanlinna and F. Odeh, *Multiplier techniques for linear multistep methods*, Numer. Funct. Anal. Optim. **3** (1981) 377–423. DOI 10.1080/01630568108816097. MR636736
19. S. P. Nørsett, *A criterion for $A(\alpha)$-stability of linear multistep methods*, BIT **9** (1969) 259–263. DOI 10.1007/bf01946817. MR0256571
20. J. Shen, T. Tang, and L. Wang, *Spectral Methods: Algorithms, Analysis and Applications*, Springer–Verlag, Berlin, 2011. DOI 10.1007/978-3-540-71041-7. MR1311481
21. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer–Verlag, Berlin, 2006. DOI 10.1007/3-540-33122-0. MR2249024

Department of Computer Science and Engineering, University of Ioannina, 451 10 Ioannina, Greece, and Institute of Applied and Computational Mathematics, FORTH, 700 13 Heraklion, Crete, Greece

Email address: akrivis@cse.uoi.gr

School of Mathematics and Statistics, Gansu Key Laboratory of Applied Mathematics and Complex Systems, Lanzhou University, Lanzhou 730000, P.R. China

Email address: chenmh@lzu.edu.cn

Institute for Math & AI, Wuhan University, Wuhan 430072, P.R. China

Email address: yufan24@whu.edu.cn