STABILITY OF IMPLICIT AND IMPLICIT–EXPLICIT MULTISTEP METHODS FOR NONLINEAR PARABOLIC EQUATIONS

GEORGIOS AKRIVIS

Abstract. We analyze the discretization of nonlinear parabolic equations in Hilbert spaces by both implicit and implicit–explicit multistep methods and establish local stability under best possible and best possible linear stability conditions, respectively. Our approach is based on suitable quantifications of the non-self-adjointness of linear elliptic operators and a discrete perturbation argument.

1. Introduction

Let $T > 0$ and $u^0 \in H$, and consider the initial value problem, for a possibly nonlinear abstract parabolic equation,

\begin{equation}
\begin{aligned}
\frac{du}{dt}(t) + A(t)u(t) &= B(t, u(t)), \quad 0 < t < T; \\
u(0) &= u^0,
\end{aligned}
\end{equation}

in a usual triple of separable complex Hilbert spaces $V \subset H = H' \subset V'$, with $V$ densely and continuously imbedded in $H$. Here $A(t) : V \to V'$, $t \in [0, T]$, is a linear operator, whereas the operator $B(t, \cdot) : V \to V'$, $t \in [0, T]$, may be nonlinear. Cf., e.g., [20, Chapter 3], [30, Chapters 3 and 4], [25, Sections 6.8.1, 6.9.4, 7.11.2, 9.3], [29, Chapter 4]. We assume that (1.1) possesses a unique, smooth solution.

Let $(\alpha, \beta)$ and $(\alpha, \gamma)$ be implicit and explicit $q$-step methods, respectively, generated by three polynomials $\alpha$, $\beta$ and $\gamma$,

$$
\alpha(\zeta) = \sum_{i=0}^{q} \alpha_i \zeta^i,
\beta(\zeta) = \sum_{i=0}^{q} \beta_i \zeta^i,
\gamma(\zeta) = \sum_{i=0}^{q-1} \gamma_i \zeta^i,
$$

with real coefficients $\alpha_i$, $\beta_i$ and $\gamma_i$.

Let $N \in \mathbb{N}, k := T/N$ be the constant time step, and $t^n := nk, n = 0, \ldots, N$, be a uniform partition of the interval $[0, T]$. Since we consider $q$-step schemes, we assume that starting approximations $U^0, \ldots, U^{q-1} \in V$ are given. We recursively define a sequence of approximations $U^m \in V$ to the nodal values $u^m := u(t^m)$ of the solution $u$ of the initial value problem (1.1) by discretizing the differential equation in (1.1) either by the
implicit scheme \((\alpha, \beta)\),

\[(1.2) \quad \sum_{i=0}^{q} (\alpha_i I + k \beta_i A(t^{n+i})) U^{n+i} = k \sum_{i=0}^{q} \beta_i B(t^{n+i}, U^{n+i}), \]

\(n = 0, \ldots, N - q\), with \(I\) the identity operator on \(H\), or by the implicit–explicit scheme \((\alpha, \beta, \gamma)\),

\[(1.3) \quad \sum_{i=0}^{q} (\alpha_i I + k \beta_i A(t^{n+i})) U^{n+i} = k \sum_{i=0}^{q-1} \gamma_i B(t^{n+i}, U^{n+i}), \]

\(n = 0, \ldots, N - q\). The implicit–explicit scheme \((\alpha, \beta, \gamma)\) results by employing the implicit scheme \((\alpha, \beta)\) for the discretization of the linear part and the explicit scheme \((\alpha, \gamma)\) for the discretization of the nonlinear part of the differential equation; see \([1, 6, 11]\). The implicit scheme \((1.2)\) has in general more advantageous stability properties than its implicit–explicit counterpart \((1.3)\); however, if \(B(t, \cdot)\) is nonlinear, \((1.2)\) is a nonlinear equation in the unknown \(U^{n+q}\). To approximate \(U^{n+q}\), we need to linearize, for instance by the Newton method; we shall not study linearizations of \((1.2)\) here. In contrast to \((1.3)\), the unknown \(U^{n+q}\) appears in \((1.3)\) only linearly, since \(\gamma_q = 0\); therefore, to advance with the implicit–explicit scheme \((1.3)\) in time, we need to solve, at each time level, just one linear equation, which reduces to a linear system of equations, if we discretize also in space.

We assume that the implicit method \((\alpha, \beta)\) is strongly \(A(0)\)-stable and denote by \(\theta, 0 < \theta \leq 90^\circ\), the largest angle for which the method \((\alpha, \beta)\) is \(A(\theta)\)-stable.

1.1. Abstract setting. We denote by \((\cdot, \cdot)\) both the inner product on \(H\) and the antiduality pairing between \(V'\) and \(V\), and by \(|\cdot|\) and \(\|\cdot\|\) the norms on \(H\) and \(V\), respectively. The space \(V'\) may be considered the completion of \(H\) with respect to the dual norm \(\|\cdot\|_*\),

\[
\|v\|_* := \sup_{\tilde{v} \in V \atop \|\tilde{v}\|^2 = 1} \frac{|(v, \tilde{v})|}{\|\tilde{v}\|} = \sup_{\tilde{v} \in V \atop \|\tilde{v}\|=1} |(v, \tilde{v})|.
\]

We assume that the operator \(A(t) : V \rightarrow V'\) is uniformly coercive and bounded, i.e.,

\[(1.4) \quad \text{Re}(A(t)v, v) \geq \kappa(t)\|v\|^2 \quad \forall v \in V,\]

and

\[(1.5) \quad \|A(t)v\|_* \leq \nu(t)\|v\| \quad \forall v \in V,\]

respectively, with two positive functions \(\kappa, \nu : [0, T] \rightarrow \mathbb{R}\), uniformly bounded from below by a positive constant and from above, respectively. Operators satisfying \((1.4)\) and \((1.5)\) are sectorial in the sense that their numerical range \(\{(A(t)v, v)/(v, v), v \in V, v \neq 0\}\) is contained in a sector \(S_\varphi, S_\varphi := \{z \in \mathbb{C} : z = \rho e^{i\varphi}, \rho \geq 0, |\varphi| \leq \varphi\}\), of half-angle \(\varphi < 90^\circ\).

High order multistep schemes are not \(A\)-stable. To take advantage of the \(A(\theta)\)-stability of the implicit scheme \((\alpha, \beta)\), in case \(\theta < 90^\circ\), we need to quantify the non-self-adjointness of the operator \(A(t)\). There are several equivalent measures of the non-self-adjointness of linear operators. Which measure or estimate of the non-self-adjointness of an operator is more suitable depends also on the employed stability technique. In
the case of the energy technique, the ratio $\nu(t)/\kappa(t)$ is commonly used as an estimate of the non-self-adjointness of $A(t)$; notice, however, that this ratio depends on the norm $\| \cdot \|$ on $V$ and may be a crude estimate of the non-self-adjointness of $A(t)$; see Section 4. On the other hand, the smallest half-angle of a sector containing the numerical range of $A(t)$ is a measure rather than an estimate of the non-self-adjointness of $A(t)$, that has been used in the stability analysis of multistep methods for linear parabolic equations by alternative stability techniques, like Cauchy integral representations or spectral and Fourier techniques; see, e.g., [14, 15, 26]. As we shall see later on, sharp sufficient stability conditions for nonlinear parabolic equations are nonlinear in this measure of non-self-adjointness of $A(t)$; see (1.14).

It will prove advantageous for our purposes to introduce alternative equivalent measures of the non-self-adjointness of $A(t)$. To this end, we first decompose $A(t)$ in its self-adjoint and anti-self-adjoint parts $A_s(t) := (A(t) + A(t)^*)/2$ and $A_a(t) := (A(t) - A(t)^*)/2$, respectively,

$$A(t) = A_s(t) + A_a(t), \quad t \in [0, T].$$

Then, with the bounded linear operator $A(t) : H \to H$ and its anti-self-adjoint part $A_a(t)$,

$$A(t) := A_s^{-1/2}(t) A(t) A_s^{-1/2}(t) = I + A_a(t), \quad A_a(t) = A_s^{-1/2}(t) A_s(t) A_s^{-1/2}(t),$$

with $I$ the identity operator on $H$, for the operator $A(t)$ we have the equivalence

$$\forall v \in V \quad (A(t)v, v) \in S_\varphi \iff |A_a(t)| \leq \tan \varphi \iff |A(t)| \leq \frac{1}{\cos \varphi}.$$ 

Indeed, for $v \in V$, with $\tilde{v} := A_s^{1/2}(t)v \in H$, we have

$$(A(t)v, v) = (A(t)A_s^{-1/2}(t)\tilde{v}, A_s^{-1/2}(t)\tilde{v}) = (A_s^{-1/2}(t)A(t)A_s^{-1/2}(t)\tilde{v}, \tilde{v})$$

$$= (A(t)\tilde{v}, \tilde{v}) = ((I + A_a(t))\tilde{v}, \tilde{v}) = |\tilde{v}|^2 + (A_a(t)\tilde{v}, \tilde{v}),$$

i.e.,

$$A(t)v, v = |\tilde{v}|^2\left[1 + \frac{(A_a(t)\tilde{v}, \tilde{v})}{|\tilde{v}|^2}\right] \quad \forall v \in V, v \neq 0.$$ 

Due to the anti-self-adjointness of $A_a(t)$, the second term in brackets is purely imaginary and (notice that $iA_a(t)$ is self-adjoint)

$$\sup_{\tilde{v} \in H \setminus \{0\}} \frac{|(A_a(t)\tilde{v}, \tilde{v})|}{|\tilde{v}|^2} = |A_a(t)|;$$

see, e.g., [27, Theorem 2.2.11]; the first equivalence in (1.8) follows immediately from the last relation and (1.9). Furthermore, it is easily seen that the norms of $A_a(t)$ and $A(t)$ are related by

$$|A(t)|^2 = 1 + |A_a(t)|^2;$$

cf. (2.6) in the sequel; this yields the second equivalence in (1.8). In view of (1.8), the norms of both $A_a(t)$ and $A(t)$ are alternative measures of the non-self-adjointness of the linear operator $A(t)$. 
Obviously, any stability condition on $|A_a(t)|$ can be equivalently written as a stability condition on $|A(t)|$ or on the smallest half-angle $\varphi(t)$ of the sector $S_{\varphi(t)}$ containing the numerical range of $A(t)$ and vice versa. Usually, stability analyses lead directly to linear sufficient stability conditions; indirectly, one may reformulate such conditions as nonlinear sufficient stability conditions on other quantities. It turns out that the best possible sufficient stability condition for the implicit scheme (1.12) under which we shall indeed establish stability, is linear in $|A_a(t)|$ (and in the Lipschitz bound $\lambda_2(t)$ of $B(t, \cdot)$, see (1.12) in the sequel) and nonlinear in $|A(t)|$ or $\varphi(t)$. Thus, our crucial assumption on the linear operator $A(t)$ is

\[(1.11) \quad |A_a(t)v| \leq \lambda_1(t)|v| \quad \forall v \in H \quad \forall t \in [0, T],\]

with a non-negative continuous stability function $\lambda_1$. Our main assumption on the nonlinear operator $B(t, \cdot) : V \rightarrow V'$ is that it satisfies the local Lipschitz condition

\[(1.12) \quad |A_s^{-1/2}(t) (B(t, v) - B(t, \tilde{v}))| \leq \lambda_2(t)|A_s^{1/2}(t)(v - \tilde{v})| + \mu_2(t)|v - \tilde{v}| \quad \forall t \in [0, T],\]

for all $v, \tilde{v}$ in a tube $T_u, T_\nu := \{v \in V : \min \|v - u(t)\| \leq 1\}$, around the solution $u$, defined in terms of the norm $\|\cdot\|$ on $V$, with a non-negative continuous stability function $\lambda_2$ and a bounded function $\mu_2$. Clearly, (1.12) holds as an equality with $\mu_2(t) = 0$ for $B(t, \cdot) := \lambda_2(t)A_s(t); \lambda_2(t)$ is positive. It then follows immediately from the Lax–Milgram lemma that, given a $w \in V'$, the linear equation

\[(1.16) \quad \alpha_q v + k\beta_q A(t)v = w\]

possesses a unique solution $v \in V$, for any fixed $t \in [0, T]$. Therefore, given the starting approximations $U^0, \ldots, U^{q-1} \in V$, the approximations $U^q, \ldots, U^N \in V$ are well defined.
by the implicit–explicit scheme (1.3). We refer to [§3 and Remark 6.1] for a discussion concerning existence and local uniqueness of approximate solutions $U^q, \ldots, U^N \in V$ for the implicit scheme (1.2).

### 1.2. The stability result

We first introduce two constants, $K_{(\alpha, \beta)}$ and $K_{(\alpha, \beta, \gamma)}$, by

\[ K_{(\alpha, \beta)} := \sup_{x > 0} \max_{\zeta \in \mathcal{X}} \frac{|x^\beta(\zeta)|}{|\alpha + x^\beta(\zeta)|}, \quad K_{(\alpha, \beta, \gamma)} := \sup_{x > 0} \max_{\zeta \in \mathcal{X}} \frac{|x^\gamma(\zeta)|}{|\alpha + x^\beta(\zeta)|}, \]

with $\mathcal{X}$ denoting the unit circle in the complex plane, $\mathcal{X} := \{z \in \mathbb{C} : |z| = 1\}$. Under our hypotheses, the constants $K_{(\alpha, \beta)}$ and $K_{(\alpha, \beta, \gamma)}$ are finite; cf. [§3 and 4]. Actually, with $\vartheta$ the largest angle for which the scheme $(\alpha, \beta)$ is $A(\vartheta)$-stable, we have

\[ K_{(\alpha, \beta)} = \frac{1}{\sin \vartheta}; \]

see [§3]. Moreover, for some implicit–explicit multistep schemes the constants $K_{(\alpha, \beta, \gamma)}$ are explicitly given in [§3 and §4].

The main result of this article is:

**Theorem 1.1** (Stability of the schemes (1.2) and (1.3)). Let $\lambda_1$ and $\lambda_2$ be the stability functions of the boundedness condition (1.11) and of the local Lipschitz condition (1.12), respectively. Then, under the linear conditions

\[ (\cot \vartheta) \lambda_1(t) + K_{(\alpha, \beta)} \lambda_2(t) < 1 \quad \forall t \in [0, T] \]

and

\[ (\cot \vartheta) \lambda_1(t) + K_{(\alpha, \beta, \gamma)} \lambda_2(t) < 1 \quad \forall t \in [0, T], \]

respectively, on the stability functions $\lambda_1$ and $\lambda_2$, the implicit multistep scheme (1.2) and the implicit–explicit multistep scheme (1.3) are locally stable in the following sense: If $U^0, \ldots, U^N \in T_u$ satisfy (1.2) and (1.3), respectively, and $V^0, \ldots, V^N \in T_u$ satisfy the corresponding perturbed equations

\[ \sum_{i=0}^{q} \left( \alpha_i I + k \beta_i A(t^{n+i}) \right) V^{n+i} = k \sum_{i=0}^{q} \beta_i B(t^{n+i}, V^{n+i}) + kE^n \]

and

\[ \sum_{i=0}^{q} \left( \alpha_i I + k \beta_i A(t^{n+i}) \right) V^{n+i} = k \sum_{i=0}^{q-1} \gamma_i B(t^{n+i}, V^{n+i}) + kE^n, \]

$n = 0, \ldots, N - q$, respectively, then, with $\vartheta^m := V^m - U^m$, for sufficiently small time step $k$, we have

\[ |\vartheta^m|^2 + k \sum_{\ell=0}^{n} ||\vartheta^\ell||^2 \leq C \left\{ \sum_{j=0}^{q-1} (|\vartheta|)^2 + k||\vartheta||^2 \right\} + k \sum_{\ell=0}^{n-q} ||E^\ell||^2, \]

$n = q, \ldots, N$, with a constant $C$ independent of the time step $k$, the approximations $U^n, V^n$, and the perturbations $E^n$. 
Actually, (1.19) and (1.20) are best possible linear sufficient stability conditions on the functions $\lambda_1$ and $\lambda_2$ in the sense that none of the three coefficients $\cot \phi, K_{(\alpha, \beta)}$, and $K_{(\alpha, \beta, \gamma)}$ can be replaced by a smaller coefficient, if we want the schemes (1.2) and (1.3), respectively, to be stable for all initial value problems (1.1) with linear operators $A(t)$ of the form (1.6) satisfying the boundedness condition (1.11) and nonlinear operators $B(t, \cdot)$ satisfying the local Lipschitz condition (1.12); see [3]. Indeed, as we shall see, in the case of the implicit scheme (1.2), a necessary stability condition is linear, namely of the form of (1.19) with the strict inequality replaced by non-strict inequality; therefore, the linear sufficient stability condition (1.19) is sharp, even among possibly nonlinear conditions on the stability functions $\lambda_1$ and $\lambda_2$.

Stability of implicit–explicit multistep methods under the sufficient stability condition (1.20) was recently established in [3] for a particular subclass of equations of the form (1.1), namely with $A(t) = [1 + ia(t)] A_s$, where $A_s$ is a time-independent, positive definite, self-adjoint operator, $i$ is the imaginary unit, and $a$ is a real-valued function.

Combining the stability result of Theorem 1.1 with the easily established consistency of the implicit–explicit scheme (1.3), we are led to optimal order a priori error estimates; see Proposition 2.2 in the sequel. These results extend easily to fully discrete schemes, if we discretize in space, for instance, by the finite element method; cf., e.g., [6, 4]. The error analysis for the implicit schemes (1.2) as well as the analysis of linearizations of these schemes are left to future work.

We assumed that the stability functions $\lambda_1$ and $\lambda_2$ are continuous for simplicity; it suffices to assume that the functions on the left-hand sides of the stability conditions (1.19) and (1.20) are uniformly bounded away from 1. The local stability estimates of Theorem 1.1 are valid in any tube $T_u \subset V$ of $u$, defined in terms of other norms, provided the local Lipschitz conditions (1.12) and (1.14) hold in $T_u$.

An advantage of the implicit schemes is that $K_{(\alpha, \beta)}$ may be much smaller than $K_{(\alpha, \beta, \gamma)}$; see [4, Table 2.1] for the case of the backward difference formula (BDF) methods. A drawback of the implicit schemes, on the other hand, is that to advance in time we need to solve, at each time level, a nonlinear equation. Thus, to implement such schemes, we need to linearize, for instance by the Newton method.

1.3. An example. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with smooth boundary $\partial \Omega$, and consider the following initial and boundary value problem, subject to homogeneous Dirichlet boundary conditions,

\[
\begin{cases}
    u_t - \sum_{i,j=1}^d ((a_{ij}(x, t) + \tilde{a}_{ij}(x, t))u_{x_j})_{x_i} = B(t, u) & \text{in } \Omega \times [0, T], \\
    u = 0 & \text{on } \partial \Omega \times [0, T], \\
    u(\cdot, 0) = u^0 & \text{in } \Omega,
\end{cases}
\]

(1.24)

with $T$ positive and $u^0 : \Omega \to \mathbb{C}$ a given initial value. Here, $\mathcal{A}, \tilde{\mathcal{A}} : \Omega \times [0, T] \to \mathbb{C}^{d \times d}$ are uniformly positive definite and Hermitian, and antihermitian matrices, respectively, with smooth entries $a_{ij}(x, t)$ and $\tilde{a}_{ij}(x, t)$, respectively, and $B(t, \cdot)$ are suitable, possibly nonlinear, operators. We assume that (1.24) possesses a smooth solution.
We consider complex-valued functions in $\Omega$ and, with standard notation for Sobolev spaces, let $H := L^2 := L^2(\Omega), V := H_0^1 := H_0^1(\Omega)$, and $V' = H^{-1} := H^{-1}(\Omega)$. We then let the time-dependent, linear operators $A_s(t), A_a(t) : V \rightarrow V'$ be defined by

$$
(1.25) \\
A_s(t)v := -\sum_{i,j=1}^{d} (a_{ij}(\cdot,t)v_{x_j})_{x_i}, \quad A_a(t)v := -\sum_{i,j=1}^{d} (\tilde{a}_{ij}(\cdot,t)v_{x_j})_{x_i},
$$

and write \(1.24\) in the form \(1.1\) with $A(t) := A_s(t) + A_a(t)$ and $B(t,\cdot) : V \rightarrow V'$, $t \in [0,T]$, (possibly) nonlinear operators. Notice that the operators $A_s(t)$ and $A_a(t)$ are positive definite and self-adjoint, and anti-self-adjoint, respectively. We shall elaborate on the boundedness condition \(1.11\) for this example in Section 3.

Some early references for multistep methods for parabolic equations are \[31, 14, 15, 32, 18, 19\]. The analysis originated in \[31\], where linear parabolic equations with self-adjoint elliptic operators were considered; spectral techniques were employed. A-stable two-step methods for nonlinear parabolic equations are investigated by the energy technique in \[32\]. Multistep methods for linear parabolic equations with general time-independent operators are analyzed in \[14, 15\]; a Cauchy integral representation is used and stability bounds, containing a logarithmic factor, for linear equations with maximally sectorial operators are established. The analysis of \[14, 15\] is extended via a discrete perturbation argument to linear parabolic equations with time-dependent operators as well as to quasilinear parabolic equations, respectively, in \[18, 19\]; spectral theory is employed in the case of self-adjoint operators whereas the von Neumann inequality is used in the case of A-stable methods; in the case of strongly A-stable two-step methods for nonlinear parabolic equations are investigated by the energy method is employed; the drawback of the specific analysis is that it does not lead to quantified sufficient stability conditions on the stability functions $\lambda_1$ or $\lambda_2$. Energy methods for high-order multistep schemes that do lead to quantified sufficient stability conditions were only recently applied to BDF schemes of order up to five, first in \[22\] for implicit BDF methods for linear parabolic equations on evolving surfaces, and subsequently in \[24\] and \[25\] for implicit and implicit–explicit BDF schemes for quasilinear and nonlinear parabolic equations, respectively.

Multistep and in particular, BDF methods in Banach spaces are analyzed, e.g., in \[24, 16, 17, 11, 10\]; the analysis in \[24, 16\] and \[17, 11, 10\] relies on semigroup theory and discrete maximal parabolic regularity, respectively; more precisely, in \[11\] the discrete maximal parabolic regularity is combined with the energy technique.

The efficiency of implicit–explicit BDF methods for nonlinear parabolic equations has been investigated by extensive numerical experiments in, e.g., \[7\] with very satisfactory results; see also relevant references therein.

For a variety of time-stepping schemes for parabolic equations and their properties we refer to the classical monograph in this field, namely \[28\].
An outline of the article is as follows: In Section 2, we first recall the main stability result of [1]; then, combining this result with a suitable decomposition of the operator $A(t)$ in (1.6), we prove Theorem 1.1 and establish optimal order error estimates. In Section 3, we give necessary stability conditions for the schemes (1.2) and (1.3), and discuss the sharpness of the sufficient stability conditions (1.19) and (1.20); the stability condition for the implicit schemes is sharp even among possibly nonlinear sufficient stability conditions on the stability functions $\lambda_1$ and $\lambda_2$; whether this is the case also for the implicit–explicit schemes remains open. In Section 4, we present additional sufficient stability conditions. For BDF schemes up to order 5, stability has also been established by the energy technique, under more stringent sufficient stability conditions than the corresponding present ones in the case of the three-, four- and five-step methods. In Section 5 we elaborate on some special cases of parabolic equations of the form (1.24).

2. Proof of the stability result

We first present a stability result of [1] and then prove Theorem 1.1. Our approach combines the main idea of [3] with a discrete perturbation argument, and concerns a much wider class of nonlinear parabolic equations. Combining stability and consistency, we establish optimal order error estimates for the implicit–explicit schemes.

2.1. A known stability result. In this subsection we assume that the operators $A$ and $B$ in (1.1) are time-independent, and the linear operator $A$ is of the form

$$A = A_s + \tilde{A}$$

with $A_s : V \rightarrow V'$ a time-independent, positive definite, self-adjoint linear operator, and a general bounded linear operator $\tilde{A} : V \rightarrow V'$, not necessarily anti-self-adjoint,

$$|A_s^{1/2}\tilde{A}v| \leq \lambda_1|A_s^{1/2}v| + \mu_1|v| \quad \forall v \in V,$$

with a stability constant $\lambda_1$ and a constant $\mu_1$.

It is shown in [1] that the implicit scheme (1.2) and the implicit–explicit scheme (1.3), respectively, are locally stable in the tube $T_\psi$ for the initial value problem (1.1), with operator $A$ as described in this section, provided the stability constants $\lambda_1$ and $\lambda_2$ in the boundedness condition (2.2) and in the local Lipschitz condition (1.12), with the operator $A_s(t)$ replaced by $A_s$, are small enough such that

$$K_{(\alpha,\beta)}\lambda_1 + K_{(\alpha,\beta)}\lambda_2 < 1$$

and

$$K_{(\alpha,\beta)}\lambda_1 + K_{(\alpha,\beta,\gamma)}\lambda_2 < 1,$$

respectively. Furthermore, (2.3) and (2.4) are best possible linear sufficient stability conditions on the constants $\lambda_1$ and $\lambda_2$ in the sense that none of the coefficients $K_{(\alpha,\beta)}$ and $K_{(\alpha,\beta,\gamma)}$ can be replaced by a smaller coefficient, if we want the schemes (1.2) and (1.3), respectively, to be stable for all equations (1.1) satisfying (2.2) and (1.12); see [1]. Notice that both constants $K_{(\alpha,\beta)}$ and $K_{(\alpha,\beta,\gamma)}$ are, for consistent schemes, larger than or equal to 1.

The only difference between the sufficient stability conditions (2.3) and (1.19), as well as between (2.4) and (1.20), is that $\lambda_1$ in (2.3) and (2.4), respectively, is replaced by $(\cos \theta)\lambda_1$ in (1.19) and (1.20). A geometric interpretation of this difference is that
\( \lambda_1 \) in (2.3) and (2.4) accounts for the operator \( \tilde{A} \) of (2.1), on the “direction” of which we did not impose any restriction; in contrast, \( \lambda_1 \) in (1.19) and (1.20) accounts for the anti-self-adjoint operator \( A_a \); the fact that \( A_a \) is the operator analogue of a purely imaginary number, i.e., a perturbation of the self-adjoint operator \( A_s \) in the “direction” of the imaginary axis, is reflected in (1.19) and (1.20) through the length \((\cos \vartheta)\lambda_1 \) of the projection of \( i\lambda_1 \) in the direction perpendicular to the boundary of the stability sector \( S_\vartheta \); see Figure 2.1.

\[ \text{Figure 2.1. Geometric interpretation of the stability conditions (1.19) and (1.20): } (\cos \vartheta)\lambda_1 \text{ is the length of the component (projection) of } i\lambda_1 \text{ in the direction perpendicular to the boundary of the stability sector } S_\vartheta. \]

We now present the main stability result of [1] for the implicit–explicit scheme (1.3); see [1, (6.6)] and [6, (2.12)]. The proof of the corresponding stability result for the implicit scheme (1.2) is completely analogous.

**Proposition 2.1** (The main stability result of [1]). Assume that the linear operator \( A : V \to V' \) is of the form (2.1), and let \( \lambda_1 \) and \( \lambda_2 \) be the stability constants of the boundedness condition (2.2) and of the local Lipschitz condition (1.12) for the time-independent operator \( B : V \to V' \), with the operator \( A_s(t) \) replaced by \( A_a \), respectively. Then, under the linear stability condition (2.4) on \( \lambda_1 \) and \( \lambda_2 \), the implicit–explicit multistep scheme (1.3) is locally stable in the following sense: if \( U^0, \ldots, U^N, V^0, \ldots, V^N \in T_u \) satisfy (1.3) and the corresponding perturbed equations (1.22), respectively, then the stability estimate (1.23) holds true for the differences \( \vartheta^m := V^m - U^m \), with a constant \( C \) independent of the time step \( k \), the approximations \( U^n, V^n \), and the perturbations \( E^n \).

2.2. **Proof of Theorem 1.1.** In contrast to subsection 2.1, here we consider the initial value problem (1.1) with linear operator \( A(t) \) of the form (1.4). In this case, now with \( \lambda_1 \) as in (1.11), the sufficient stability conditions (2.3) and (2.4) can be relaxed to (1.19) and (1.20), respectively, i.e., the first terms on the left-hand sides of (2.3) and (2.4) can be multiplied by \( \cos \vartheta \). See the relevant comments in [1].

We will see that Theorem 1.1 for the implicit–explicit scheme (1.3) follows from Proposition 2.1 by using a more favorable decomposition of the operator \( A(t) \) and a discrete perturbation argument. The proof of the corresponding stability result for the implicit scheme (1.2) is again completely analogous and is omitted.
The key point in the proof of Theorem 1.1 is the following choice of a decomposition of the operator $A(t)$,
\begin{equation}
(2.5) \quad A(t) = \tilde{A}_s(t) + \tilde{A}(t), \quad \text{with} \quad \tilde{A}_s(t) := (1 + \eta)A_s(t) \quad \text{and} \quad \tilde{A}(t) := A_s(t) - \eta A_s(t),
\end{equation}
with \( \eta \) a non-negative quantity that may depend on \( \lambda_1(t) \) and \( \lambda_2(t) \). We shall see that a suitable choice of \( \eta \) is 
\( \eta := (\tan \vartheta)\lambda_1(t) \), with \( \lambda_1 \) as in (1.1).

Now, in analogy to the anti-self-adjoint operator $A_a(t) = A_s^{-1/2}(t)A_a(t)A_s^{-1/2}(t)$, see (1.7), with the notation of the decomposition (2.5), we consider the operator
\[ \tilde{A}(t) := \tilde{A}_s^{-1/2}(t)\tilde{A}(t)\tilde{A}_s^{-1/2}(t); \]
notice that, in contrast to $A_a(t)$, the operator $\tilde{A}(t)$ is not anti-self-adjoint for positive \( \eta \), since $\tilde{A}(t)$ is not anti-self-adjoint. We have the following crucial relation between $\tilde{A}(t)$ and $A_a(t)$,
\begin{equation}
(2.6) \quad |(A_a(t) - \hat{\eta}I)v|^2 = |A_a(t)v|^2 + \hat{\eta}^2|v|^2 \quad \forall v \in H,
\end{equation}
and infer that the norms of the operators $\tilde{A}(t)$ and $A_a(t)$ are related as follows
\begin{equation}
(2.7) \quad |\tilde{A}(t)|^2 = \frac{1}{(1 + \hat{\eta})^2}(|A_a(t)|^2 + \eta^2).
\end{equation}
From (2.7) and (1.11), we obtain for the norm of $\tilde{A}(t)$ the central estimate
\begin{equation}
(2.8) \quad |\tilde{A}(t)v| \leq \frac{\sqrt{\lambda_1^2(t) + \eta^2}}{1 + \eta}|v| \quad \forall v \in H \forall t \in [0, T].
\end{equation}

Furthermore, it is easily seen that the operators $\tilde{A}_s(t)$ and $B(t, \cdot)$ satisfy the estimate
\begin{equation}
(2.9) \quad |\tilde{A}_s^{-1/2}(t)(B(t, v) - B(t, \tilde{v}))| \leq \lambda_2(t)|\tilde{A}_s^{1/2}(t)(v - \tilde{v})| + \tilde{\mu}_2(t)|v - \tilde{v}| \quad \forall v, \tilde{v} \in T_a,
\end{equation}
with
\begin{equation}
(2.10) \quad \lambda_2(t) := \frac{\lambda_2(t)}{1 + \eta} \quad \text{and} \quad \tilde{\mu}_2(t) := \frac{\mu_2(t)}{\sqrt{1 + \eta}}.
\end{equation}
Compare (2.8) with (1.11), and (2.9) with (1.12), respectively.

2.2.1. Time-independent operators. We first assume that the operators $A$ and $B$ in (1.1) are time-independent. From (2.4) and (2.8), (2.9), (2.10), we then infer that the scheme (1.3) is locally stable for (1.1), with a time-independent operator $A$ of the form (1.6), if \( \lambda_1 \) and \( \lambda_2 \) are such that
\begin{equation}
(2.11) \quad \frac{1}{\sin \vartheta} \sqrt{\frac{\lambda_1^2 + \eta^2}{1 + \eta}} + K_{(\alpha, \beta, \gamma)} \frac{\lambda_2}{1 + \eta} < 1,
\end{equation}
for some non-negative \( \eta \). Now, for \( \eta = (\tan \vartheta) \lambda_1 \), condition (2.11) reduces to the desired sufficient stability condition (1.20) and the proof is complete. The motivation for this choice of \( \eta \) is that (2.11) can be written in the form

\[
\frac{1}{\sin \vartheta} \sqrt{\lambda_1^2 + \eta^2} = K_{(\alpha, \beta, \gamma)} \lambda_2 < 1
\]

and the expression in brackets attains its minimum at \( \eta = (\tan \vartheta) \lambda_1 \).

In the case of A-stable implicit methods \((\alpha, \beta)\), i.e., when \( \vartheta = 90^\circ \), it suffices to choose the parameter \( \eta \) in (2.11) large enough.

Let us mention that the sufficient stability condition in (2.10) is linear; see (2.4). However, since the bound of the norm of \( A(t) \) in (2.8) is nonlinear in \( \lambda_1 \), we are led to the nonlinear stability condition (2.11).

2.2.2. Time-dependent operators. We shall now utilize a discrete perturbation argument to extend our previous stability result to the case of time-dependent operators \( A(t), B(t, \cdot) : V \to V' \), \( t \in [0, T] \), assuming that they satisfy (1.11) and (1.12) as well as the Lipschitz conditions (1.13) and (1.14), respectively, with respect to \( t \). Although this argument is well known, see, for instance, [1, 2, 3, 4, 5, 6, 7, 8, 9], it is, to our best knowledge, for the first time employed in the analysis of implicit and implicit–explicit multistep methods for nonlinear parabolic equations by spectral and Fourier techniques; in particular, it was an essential requirement up to now that the self-adjoint part of the linear operator was time-independent.

First, subtracting (1.3) from (1.2), and letting \( \vartheta^m := V^m - U^m \), we obtain

\[
(2.12) \quad \sum_{i=0}^{q} \left( \alpha_i I + k \beta_i A(t^{n+i}) \right) \vartheta^{n+i} = k \sum_{i=0}^{q-1} \gamma_i [B(t^{n+i}, V^{n+i}) - B(t^{n+i}, U^{n+i})] + kE^n,
\]

\( n = 0, \ldots, N - q \). For a fixed \( n, q \leq m \leq N \), and for \( n = 0, \ldots, m - q \), we rewrite (2.12) in the following form

\[
(2.13) \quad \sum_{i=0}^{q} \left( \alpha_i I + k \beta_i A(t^m) \right) \vartheta^{n+i} = k \sum_{i=0}^{q-1} \gamma_i [B(t^m, V^{n+i}) - B(t^m, U^{n+i})] + k\mathcal{E}^n,
\]

\( n = 0, \ldots, m - q \), where

\[
\mathcal{E}^n := E^n + \mathcal{E}_A^n + \mathcal{E}_B^n,
\]

\[
\mathcal{E}_A^n := \sum_{i=0}^{q} \beta_i [A(t^m) - A(t^{n+i})] \vartheta^{n+i},
\]

\[
\mathcal{E}_B^n := -\sum_{i=0}^{q-1} \gamma_i \left\{ [B(t^m, V^{n+i}) - B(t^m, U^{n+i})] - [B(t^{n+i}, V^{n+i}) - B(t^{n+i}, U^{n+i})] \right\}.
\]

Since the time \( t \) is frozen at \( t^m \) in both operators \( A(t^m) \) and \( B(t^m, \cdot) \) in (2.13), and conditions (1.11) and (1.12) are satisfied for \( t = t^m \), we can apply the already established
stability estimate for time-independent operators, and infer that
\begin{equation}
|q|^2 + k \sum_{\ell=0}^{m} \|q^\ell\|^2 \leq C \left\{ \sum_{j=0}^{q-1} \left( |q^j|^2 + k \|q^j\|^2 \right) + k \sum_{\ell=0}^{m-q} \|q^\ell\|^2 \right\};
\end{equation}
see (1.23) for the case of time-independent operators. Notice also that the constant \(C\) in (2.15) is independent of \(t\); it only depends on \(T\), the suprema of the function on the left-hand side of the sufficient stability condition (1.20), and the function \(\mu_2\) in (1.12).

Our task is now the estimation of the last term on the right-hand side of (2.15) in a suitable way; to this end we shall use the Lipschitz conditions (1.13) and (1.14) with respect to \(t\):

First, we have
\begin{equation}
\|q^\ell\|_\ast \leq \sum_{i=0}^{q} \beta_i \| (A(t^m) - A(t^{\ell+i})) q^{\ell+i} \|_\ast,
\end{equation}
and thus, using the Lipschitz continuity of \(A\) in time, see (1.13),
\begin{equation}
\|q^\ell\|_\ast \leq L_A \sum_{i=0}^{q} \beta_i \| (t^m - t^{\ell+i}) q^{\ell+i} \|_\ast, \quad \ell = 0, \ldots, m - q.
\end{equation}

Analogously, using (1.14), we obtain
\begin{equation}
\|q^\ell\|_\ast \leq L_B \sum_{i=0}^{q-1} \gamma_i \| (t^m - t^{\ell+i}) q^{\ell+i} \|_\ast, \quad \ell = 0, \ldots, m - q.
\end{equation}

Let
\begin{equation}
\delta^m := \sum_{\ell=0}^{n} \|q^\ell\|^2, \quad \Theta^n := k \sum_{\ell=0}^{n} \|q^\ell\|^2.
\end{equation}

In view of (2.16) and (2.17), we have
\begin{equation}
\delta^{m-q} \leq C \sum_{\ell=0}^{m-q} \left( \sum_{i=0}^{q} \| (t^m - t^{\ell+i}) q^{\ell+i} \| \right)^2 \leq C \sum_{\ell=0}^{m-1} (t^m - t^\ell)^2 \|q^\ell\|^2,
\end{equation}
whence
\begin{equation}
\delta^{m-q} \leq C k \sum_{\ell=0}^{m-1} (m - \ell)^2 k \|q^\ell\|^2,
\end{equation}
i.e.,
\begin{equation}
\delta^{m-q} \leq C k \sum_{\ell=0}^{m-1} (m - \ell)^2 (\Theta^\ell - \Theta^{\ell-1}),
\end{equation}
with \(\Theta^{-1} := 0\). Now, by summation by parts, we have
\begin{equation}
\sum_{\ell=0}^{m-1} (m - \ell)^2 (\Theta^\ell - \Theta^{\ell-1}) = \sum_{\ell=0}^{m-1} \left[ 2(m - \ell) - 1 \right] \Theta^\ell,
\end{equation}
and (2.19) yields

\[ \mathcal{E}_{m-q} \leq C k \sum_{\ell=0}^{m-1} (m-\ell) \Theta^\ell \leq C \sum_{\ell=0}^{m-1} \Theta^\ell. \]

From (2.15) and (2.20), we easily infer that

\[ \Theta^m \leq C \left\{ \sum_{j=0}^{q-1} (||\varphi^j||^2 + k||\varphi^j||^2) + k \sum_{\ell=0}^{m-q} ||E^\ell||^2_\star \right\} + C k \sum_{\ell=0}^{m-1} \Theta^\ell. \]

Therefore, with a discrete Gronwall inequality we obtain

\[ \Theta^m \leq C \left\{ \sum_{j=0}^{q-1} (||\varphi^j||^2 + k||\varphi^j||^2) + k \sum_{\ell=0}^{m-q} ||E^\ell||^2_\star \right\}, \]

\[ m = q, \ldots, N. \]

Combining (2.15) with (2.22), see also (2.20), we get the desired stability estimate (4.23) and the proof of Theorem 1.1 is complete.

**Remark 2.1** (Relaxation of the Lipschitz conditions (1.13) and (1.14) to bounded variation conditions). The Lipschitz conditions (1.13) and (1.14), with respect to \( t \), can be easily relaxed to bounded variation conditions,

\[ \| (A(t) - A(s)) v \|_* \leq [\sigma_A(t) - \sigma_A(s)] \| v \|, \quad 0 \leq s \leq t \leq T, \quad \forall v \in V, \]

and

\[ \| [B(t, v) - B(t, v_\star)] - [B(s, v) - B(s, v_\star)] \|_* \leq [\sigma_B(t) - \sigma_B(s)] \| v - v_\star \|, \]

\[ 0 \leq s \leq t \leq T, \quad \forall v, v_\star \in T_v, \]

with two increasing functions \( \sigma_A, \sigma_B : [0, T] \to \mathbb{R} \). Indeed, it is easily seen that the analogue of (2.19) is in this case

\[ \mathcal{E}_{m-q} \leq C \sum_{\ell=0}^{m-1} \left[ \sigma(t^m) - \sigma(t^\ell) \right]^2 (\Theta^\ell - \Theta^{\ell-1}), \]

with \( \sigma(t) := \sigma_A(t) + \sigma_B(t) \). Then, by summation by parts, we obtain

\[ \mathcal{E}_{m-q} \leq \frac{C}{k} \sum_{\ell=0}^{m-1} a_\ell \Theta^\ell, \]

with \( a_\ell := [\sigma(t^m) - \sigma(t^\ell)]^2 - [\sigma(t^m) - \sigma(t^{\ell+1})]^2 \). From (2.15) and (2.26), we infer that

\[ \Theta^m \leq C \left\{ \sum_{j=0}^{q-1} (||\varphi^j||^2 + k||\varphi^j||^2) + k \sum_{\ell=0}^{m-q} ||E^\ell||^2_\star \right\} + C \sum_{\ell=0}^{m-1} a_\ell \Theta^\ell. \]

Since the sum \( \sum_{\ell=0}^{m-1} a_\ell \) is uniformly bounded by a constant independent of \( m \) and the time step \( k \),

\[ \sum_{\ell=0}^{m-1} a_\ell = [\sigma(t^m) - \sigma(0)]^2 \leq [\sigma(T) - \sigma(0)]^2, \]

a discrete Gronwall-type argument applied to (2.27) leads again to (2.22), and the stability proof is completed as before.
Remark 2.2 (Slight relaxation of the boundedness condition (1.11)). The boundedness condition (1.11) can be slightly relaxed to

\[(2.28) \quad |A_n(t)v| \leq \lambda_1(t)|v| + \mu_1(t)\|v\| \quad \forall v \in H \ \forall t \in [0,T],\]

with the non-negative continuous stability function \(\lambda_1\) as in (1.11) and a bounded function \(\mu_1; \mu_1\) accounts for lower order terms and does not enter in the sufficient stability conditions (1.19) and (1.20). For instance, if \(A_n\) is a positive definite self-adjoint operator and \(A_n := i\lambda_1 A_n + i\mu_1 A_n^{-1/2},\) then

\[
\begin{align*}
A_n := A_n^{1/2} A_n A_n^{1/2} = i\lambda_1 I + i\mu_1 A_n^{-1/2},
\end{align*}
\]

whence

\[(2.29) \quad |A_n v| \leq \lambda_1 |v| + \mu_1 \|v\| \quad \forall v \in H.\]

2.3. Error estimates. Here, combining the easily established consistency of the implicit–explicit multistep scheme (1.3) with the local stability result of Theorem 1.1, we derive optimal order error estimates. These results extend easily to fully discrete schemes, if we discretize in space, for instance, by the finite element method; cf., e.g., [6, 14].

For simplicity, we assume that the order of both \(q\)-step methods, the implicit \((\alpha, \beta)\) and the explicit \((\alpha, \gamma)\), is \(p\), i.e.,

\[(2.30) \quad \sum_{i=0}^{q} \ell^i \alpha_i = \ell \sum_{i=0}^{q} \ell^i \beta_i = \ell \sum_{i=0}^{q-1} \ell^{i-1} \gamma_i, \quad \ell = 0, 1, \ldots, p.\]

The consistency error \(E^n\) of the scheme (1.3) for the solution \(u\) of (1.1), i.e., the amount by which the exact solution misses satisfying (1.3), is given by

\[(2.31) \quad k E^n = \sum_{i=0}^{q} (\alpha_i I + k \beta_i A(t^{n+i})) u^{n+i} - k \sum_{i=0}^{q-1} \gamma_i B(t^{n+i}, u^{n+i}),\]

\(n = 0, \ldots, N - q;\) we recall that \(u^\ell = u(t^\ell)\) are the nodal values of the solution \(u\) of (1.1). First, letting

\[
E_1^n := \sum_{i=0}^{q} [\alpha_i u^{n+i} - k \beta_i u'(t^{n+i})], \quad E_2^n := k \sum_{i=0}^{q} (\beta_i - \gamma_i) B(t^{n+i}, u^{n+i}),
\]

with \(\gamma_q := 0,\) and using the differential equation in (1.1), we infer that

\[(2.32) \quad k E^n = E_1^n + E_2^n.\]

Now, by Taylor expanding about \(t^n\), we see that, due to the order conditions of the implicit method \((\alpha, \beta)\), i.e., the first equality in (2.30), and the second equality in (2.30), respectively, the leading terms of order up to \(p\) cancel, and we obtain

\[
\begin{cases}
E_1^n = \frac{1}{p!} \sum_{i=1}^{q} \int_{t^n}^{t^{n+i}} (t^{n+i} - s)^{p-1} [\alpha_i (t^{n+i} - s) - pk \beta_i] u^{(p+1)}(s) \, ds, \\
E_2^n = \frac{k}{(p-1)!} \sum_{i=1}^{q} (\beta_i - \gamma_i) \int_{t^n}^{t^{n+i}} (t^{n+i} - s)^{p-1} \frac{d^p}{ds^p} B(s, u(s)) \, ds.
\end{cases}
\]
Thus, under obvious regularity requirements, we obtain the consistency estimate

\[
\max_{0\leq n\leq N-q} \|E^n\|_* \leq Ck^p.
\]

**Remark 2.3** (A basic difference in the error analysis of implicit and implicit–explicit schemes). For a specific \( n \geq q \), the local stability estimate (1.23) is valid for the implicit–explicit scheme (1.3), provided that \( U^\ell, V^\ell \in T_u, \ell = 0, \ldots, n-1 \). Indeed, since \( \gamma_q \) vanishes, no estimate of the difference \( B(t^n, U^n) - B(t^n, V^n) \) is needed. This fact plays an important role in the error analysis, since it allows us to show inductively that the approximations \( U^m_n \) belong to \( T_u \); see Proposition 2.2. In contrast, in the case of the implicit scheme (1.2), to establish the stability estimate (1.23), for a specific \( n \), we need to assume that \( U^\ell, V^\ell \in T_u, \ell = 0, \ldots, n \).

**Proposition 2.2** (Optimal order error estimates). Assume that the stability functions \( \lambda_1, \lambda_2 \) of Theorem 1.1 satisfy the stability condition (1.20), that the solution \( u \) of (1.1) is sufficiently smooth such that the consistency estimate (2.34) be valid, and that we are given starting approximations \( U^0, U^1, \ldots, U^{n-1} \in V \) to \( u' \), \ldots, \( u^{q-1} \) such that

\[
\max_{0\leq j\leq q-1} (|u^j - U^j| + k^{1/2} \|u^j - U^j\|) \leq Ck^p.
\]

Let \( U^q, \ldots, U^N \in V \) be recursively defined by the implicit–explicit \( q \)-step scheme (1.3). Then, there exists a constant \( C \), independent of \( k \), such that, for \( k \) sufficiently small,

\[
\max_{0\leq n\leq N} |u(t^n) - U^n| \leq Ck^p.
\]

**Proof.** Let \( e^n := u^n - U^n, n = 0, \ldots, N \), and subtract (1.31) from (2.31). According to the consistency estimate (2.34) and our assumption (2.35) on the accuracy of the starting approximations \( U^0, U^1, \ldots, U^{q-1} \), there exists a constant \( C \), such that

\[
C\left\{ \sum_{j=0}^{q-1} (|e^j|^2 + k\|e^j\|^2) + k\sum_{\ell=0}^{N-q} \|E^\ell\|_*^2 \right\} \leq C^2 k^{2p}.
\]

Next, we shall inductively show that

\[
|e^m|^2 + k\sum_{\ell=0}^{m} \|e^\ell\|^2 \leq C\left\{ \sum_{j=0}^{q-1} (|e^j|^2 + k\|e^j\|^2) + k\sum_{\ell=0}^{m-q} \|E^\ell\|_*^2 \right\},
\]

\( m = q - 1, \ldots, N \). This stability estimate completes the proof, since (2.36) is an immediate consequence of (2.38) and (2.37).

Now, the estimate (2.38) is clearly valid for \( m = q - 1 \). Assume inductively that it holds for \( m = q - 1, \ldots, n + q - 1, 0 \leq n \leq N - q \). Then, according to (2.37) and the induction hypothesis, we have, for \( k \) small enough,

\[
\max_{0\leq j\leq n+q-1} \|e^j\| \leq C^*k^{p-1/2} \leq 1,
\]

and thus \( U^j \in T_u, j = 0, \ldots, n + q - 1 \). Therefore, according to our local stability estimate (1.23) with \( n \) replaced by \( n + q \), see also Remark 2.3, the estimate (2.38) holds indeed for \( m = n + q \) as well, and the proof is complete. \( \square \)
3. On the Sharpness of the Stability Conditions \((1.19)\) and \((1.20)\)

In this section, we first prove that the sufficient stability condition \((1.19)\) for the implicit scheme \((1.2)\) is sharp by explicitly constructing suitable parabolic equations. Then, we give necessary stability conditions for the implicit scheme \((1.2)\) as well as for the implicit–explicit scheme \((1.3)\). In the case of the implicit scheme \((1.2)\), the necessary stability condition is linear, and leads us again to the conclusion that the sufficient stability condition \((1.19)\) is sharp. For the implicit–explicit scheme \((1.3)\), the sufficient condition \((1.20)\) is sharp, if the implicit method \((\alpha, \beta)\) is A-stable.

3.1. Sharpness of the sufficient stability condition \((1.19)\) for the implicit scheme \((1.2)\). In view of \((1.18)\), the sufficient stability condition \((1.19)\) for the implicit scheme \((1.2)\) can be written as

\[
(\cos \vartheta)\lambda_1(t) + \lambda_2(t) < \sin \vartheta \quad \forall t \in [0, T].
\]

Notice that the bound \(\sin \vartheta\) on the right-hand side is the distance of the number 1 from the boundary of the stability sector \(S_\vartheta\) of the method. The positive definite self-adjoint operator \(A_s\) is the operator analogue of a positive number; thinking of it as normalized, we can view it as the analogue of 1. On the left-hand side of \((3.1)\), the stability function \(\lambda_2\) is simply added, since it accounts for the perturbation \(B\), on the “direction” of which we did not impose any restriction. On the other hand, the stability function \(\lambda_1\) accounts for the anti-self-adjoint operator \(A_s\); the fact that this is a perturbation of the self-adjoint operator \(A_s\) in the “direction” of the imaginary axis, in the sense that anti-self-adjoint operators are the analogues of purely imaginary numbers, is reflected in \((3.1)\) through the coefficient \(\cos \vartheta\); see Figure 2.1. The stability condition \((3.1)\) can also be written as \(\lambda_2(t) < \sin \vartheta - (\cos \vartheta)\lambda_1(t)\) with the expression on its right-hand side being the distance of the point \(1 + i\lambda_1(t)\) from the boundary of the stability sector \(S_\vartheta\) of the method; see Figure 2.1.

In the simplest case \(\vartheta = 90^\circ\), condition \((3.1)\) reads \(\lambda_2(t) < 1\); the last condition is necessary for the parabolicity of \(u' + A_s u = \lambda_2(t) A_s u\), with a positive definite self-adjoint operator; in particular, condition \((3.1)\) cannot be relaxed in this case.

We now focus on the interesting case \(\vartheta < 90^\circ\) and, assuming that the left-hand side of \((1.19)\) exceeds 1, shall explicitly construct examples of parabolic equations, for which the implicit scheme \((1.2)\) is unstable. Indeed, in that case we will have \((\cos \vartheta)\lambda_1(t) + \lambda_2 = \sin \vartheta\), for some \(\vartheta < \hat{\vartheta} < 90^\circ\) (see \((3.1)\)). Consider the ray \(\ell_{\hat{\vartheta}} := \{\rho e^{i\hat{\vartheta}}, \rho \geq 0\}\), and let \(\hat{z}_2\) denote the orthogonal projection of \(z_1 := 1 + i\lambda_1\) on \(\ell_{\hat{\vartheta}}\). First, \(\lambda_2 = |\hat{z}_2 - z_1|\); see Figures 2.1 and 4.1. Let now \(A_s\) be a positive definite self-adjoint operator with unbounded spectrum and consider the “rotated” operator \(\hat{z}_2 A_s\). The eigenvalues of \(\hat{z}_2 A_s\) lie on the ray \(\ell_{\hat{\vartheta}}\), which is outside of the stability sector \(S_\vartheta\); therefore, according to the von Neumann criterion, the method \((\alpha, \beta)\) is unstable for the equation

\[
u' + \hat{z}_2 A_s u = 0.
\]

More precisely, by definition, for \(\hat{\vartheta}\) sufficiently close to \(\vartheta\), the ray \(\ell_{\hat{\vartheta}}\) is not entirely contained in the stability region \(S\) of the method; if \(\lambda\) is an eigenvalue of \(A_s\), then the method is unstable for equation \((3.2)\) for all time steps \(k\) such that \(k \hat{z}_2 \lambda \not\in S\); since there exists an unbounded sequence of positive eigenvalues of \(A_s\), it is impossible to
find a positive $k_0$ such that the method be stable for equation (3.2) for all time steps $0 < k < k_0$.

Now, we write equation (3.2) in the form

$$u' + A_s u + i \lambda_1 A_s u = B(u)$$

with $B(u) := -(\dot{z}_2 - z_1) A_s u$. It is easily seen that the boundedness condition (1.11) and the Lipschitz condition (1.12) are satisfied for (3.3) with the given constants $\lambda_1$ and $\lambda_2$. We infer that if the left-hand side in the stability condition (1.19) exceeds the bound on its right-hand side, the method $(\alpha, \beta)$ is in general unstable.

3.2. Necessary stability conditions. Here, we shall give necessary stability conditions for the implicit scheme (1.2) as well as for the implicit–explicit scheme (1.3). The necessary stability conditions are expressed in terms of two suitable functions. In the case of the implicit method (1.2), the corresponding function has a simple form and the necessary stability condition for the implicit scheme (1.2) will lead us to the same conclusion as in the previous subsection, namely that the sufficient stability condition (1.19) is sharp.

3.2.1. The implicit scheme. Assume that the implicit method $(\alpha, \beta)$ is $A(\vartheta)$-stable, with $\vartheta < 90^\circ$ as large as possible. Extending the definition (1.17) of the constant $K_{(\alpha, \beta)}$, let

$$K_{(\alpha, \beta)}(y) := \sup_{x > 0} \max_{\zeta \in \mathcal{X}} \frac{|x \beta(\zeta)|}{|\alpha(\zeta) + x(1 + iy)\beta(\zeta)|}, \quad -\tan \vartheta < y < \tan \vartheta.$$ 

Notice that $K_{(\alpha, \beta)}(0) = K_{(\alpha, \beta)}$; see (1.17). Furthermore, since

$$\frac{|x \beta(\bar{\zeta})|}{|\alpha(\zeta) + x(1 + iy)\beta(\zeta)|} = \frac{|x \beta(\zeta)|}{|\alpha(\zeta) + x(1 - iy)\beta(\zeta)|},$$

with $\bar{\zeta}$ denoting the complex conjugate of $\zeta$, $K_{(\alpha, \beta)}$ is an even function of $y$,

$$K_{(\alpha, \beta)}(y) = K_{(\alpha, \beta)}(-y), \quad |y| < \tan \vartheta.$$ 

Our task now is to simplify the presentation (3.4) of $K_{(\alpha, \beta)}(y)$. Since the function is even, it suffices to consider the case of non-negative arguments $y$. Let $d(\zeta) := \alpha(\zeta)/\beta(\zeta), \quad \zeta \in \mathcal{X}$, represent the points of the root locus curve of the method $(\alpha, \beta)$. The root locus curve is symmetric with respect to the real axis. Since the method is $A(\vartheta)$-stable, the root locus curve lies outside of the sector $-S_\vartheta$.

We introduce the parts $\mathcal{X}_y^+$ and $\mathcal{X}_y^-$ of the unit circle $\mathcal{X}$ according to the sign of $\text{Re}\left((1 - iy)d(\zeta)\right)$,

$$\mathcal{X}_y^+ := \{\zeta \in \mathcal{X} : \text{Re}\left((1 - iy)d(\zeta)\right) \geq 0\}, \quad \mathcal{X}_y^- := \{\zeta \in \mathcal{X} : \text{Re}\left((1 - iy)d(\zeta)\right) < 0\},$$

with $d(\zeta)$ the points of the root locus curve; with $0 \leq \varphi < \vartheta$ such that $y = \tan \varphi$, and $d(\zeta) = \rho(\zeta)e^{i\varphi(\zeta)}, \rho(\zeta) \geq 0, -\pi \leq \psi(\zeta) < \pi$, we have $(1 - iy)d(\zeta) = \frac{\bar{\rho}(\zeta)}{\cos \vartheta} e^{i(\psi(\zeta) - \vartheta)}$, whence $\zeta \in \mathcal{X}_y^+$ and $\zeta \in \mathcal{X}_y^-$, respectively, simply means that $d(\zeta)$ belongs to the rotated by the angle $\varphi$ non-negative and negative complex half-planes, respectively, in the positive direction. For $\tilde{x} > 0$, we have

$$|1 + \tilde{x}(1 - iy)d(\zeta)|^2 = 1 + 2\tilde{x} \text{Re} \omega(\zeta; y) + \tilde{x}^2 |\omega(\zeta; y)|^2$$
with \( \omega(\zeta; y) := (1 - iy)d(\zeta) \). Now, for \( \zeta \in \mathcal{K}^+ \), we have \( \text{Re} \omega(\zeta; y) \geq 0 \), and the infimum of the right-hand side of (3.5), for positive \( \bar{x} \), is 1. For \( \zeta \in \mathcal{K}^- \), on the other hand, the quadratic polynomial on the right-hand side of (3.5) attains its minimum at
\[
\bar{x}^*(\zeta; y) := \frac{-\text{Re} \omega(\zeta; y)}{|\omega(\zeta; y)|^2}.
\]
Thus, we have
\[
(3.6) \quad \sup_{\bar{x} > 0} \frac{1}{|1 + \bar{x}(1 - iy)d(\zeta)|} = \begin{cases} 1 & \forall \zeta \in \mathcal{K}^+, \\ \frac{|\omega(\zeta; y)|}{\text{Im} \omega(\zeta; y)} & \forall \zeta \in \mathcal{K}^-.
\end{cases}
\]

Now, from (3.6) and the definition of \( K_{(\alpha, \beta)}(y) \) we infer that
\[
(3.7) \quad K_{(\alpha, \beta)}(y) = \sup_{\zeta \in \mathcal{K}^-} \frac{|d(\zeta)|}{\text{Im} ((1 - iy)d(\zeta))}.
\]
Let \( d(\zeta) = -\rho(\zeta)e^{i\psi(\zeta)} \), with \( \rho(\zeta) \geq 0 \) and \(-\pi \leq \psi(\zeta) < \pi \). Since the method is \( A(\vartheta) \)-
stable, we have \( |\psi(\zeta)| \geq \vartheta \), and the infimum of \( |\psi(\zeta)| \) is \( \vartheta \). Now, with \( 0 \leq \varphi < \vartheta \) such that \( y = \tan \varphi \), we have
\[
(1 - iy)d(\zeta) = -\frac{\rho(\zeta)}{\cos \varphi}e^{i(\psi(\zeta) - \varphi)},
\]
whence (3.7) takes the form
\[
K_{(\alpha, \beta)}(y) = (\cos \varphi) \sup_{\zeta \in \mathcal{K}^-} \frac{1}{|\sin(\psi(\zeta) - \varphi)|}.
\]
Since \( \cos(\psi(\zeta) - \varphi) > 0 \) for \( \zeta \in \mathcal{K}^- \), and either \(-\pi - \varphi \leq \psi(\zeta) - \varphi \leq -\vartheta - \varphi \) or \( \vartheta - \varphi \leq \psi(\zeta) - \varphi \leq \pi - \varphi \), it is easily seen that the infimum of \( |\sin(\psi(\zeta) - \varphi)| \) is \( \sin(\vartheta - \varphi) \). Thus, the previous relation yields the desired simplified representation
\[
(3.8) \quad K_{(\alpha, \beta)}(y) = \frac{\cos \varphi}{\sin(\vartheta - \varphi)}, \quad y = \tan \varphi.
\]
Notice that this relation can also be written in the form
\[
(3.9) \quad K_{(\alpha, \beta)}(y) = \frac{1}{\sin \vartheta - \cos \vartheta \tan \varphi} = \frac{1}{\sin \vartheta - (\cos \vartheta)y}, \quad 0 \leq y < \tan \vartheta;
\]
the denominator is the distance of the point \( 1 + iy \) from the boundary of the stability sector \( S_{\vartheta} \) of the method \((\alpha, \beta)\); see Figure 2.1.

We shall see that
\[
(3.10) \quad K_{(\alpha, \beta)}(\lambda_1(t)) \lambda_2(t) \leq 1 \quad \forall t \in [0, T]
\]
is a necessary stability condition for the implicit scheme (1.2). As before, (3.10) is necessary, if we want (1.2) to be locally stable for all equations satisfying our assumptions with the given stability functions \( \lambda_1 \) and \( \lambda_2 \).

Assume that (3.10) is not satisfied for a certain value \( t^* \) of \( t \); for notational simplicity we drop the dependence of \( \lambda_1 \) and \( \lambda_2 \) on \( t^* \). Then, for the function \( k \),
\[
k(x, \zeta) := \frac{\lambda_2x\beta(\zeta)}{\alpha(\zeta) + x(1 + i\lambda_1)\beta(\zeta)}, \quad x > 0, \ |\zeta| \geq 1,
\]
we have
\begin{equation}
\exists z \in \mathcal{H}, \quad x > 0 \quad |k(x, z)| > 1.
\end{equation}
Since
\begin{equation}
\lim_{|k| \to \infty} |k(x, \zeta)| = \frac{\lambda_2 x \beta_q}{|\alpha_q + x(1 + i\lambda_1)\beta_q|} \leq \frac{\lambda_2 x \beta_q}{\alpha_q + x \beta_q} < \lambda_2 < 1,
\end{equation}
we infer that there exists a $\zeta^* \in \mathbb{C}$ with $|\zeta^*| > 1$ such that $|k(x, \zeta^*)| = 1$, i.e.,
\begin{equation}
\frac{\lambda_2 x \beta(\zeta^*)}{\alpha(\zeta^*) + x(1 + i\lambda_1)\beta(\zeta^*)} = e^{-i\psi},
\end{equation}
for a $\psi \in [0, 2\pi)$. Therefore,
\begin{equation}
\alpha(\zeta^*) + x(1 + i\lambda_1)\beta(\zeta^*) - \lambda_2 x e^{i\psi}\beta(\zeta^*) = 0.
\end{equation}
Then, with $A_s$ a positive definite self-adjoint operator, choosing the anti-self-adjoint operator $A_s := i\lambda_1 A_s$ and the linear operator $B := \lambda_2 e^{i\psi} A_s$, we readily see that the boundedness condition (3.11) and the Lipschitz condition (3.12) are satisfied. According to the von Neumann criterion, a necessary stability condition is that, if $\nu$ is an eigenvalue of $A_s$, the solutions of
\begin{equation}
\sum_{j=0}^q \left[\alpha_j + k\nu \left((1 + i\lambda_1)\beta_j - \lambda_2 e^{i\psi}\beta_j\right)\right] v^{n+j} = 0
\end{equation}
are bounded; for $k\nu = x$ this is not the case, since in view of (3.13) the root condition is not satisfied. Therefore, the scheme (1.2) is not unconditionally stable.

In view of the presentation (3.9) of $K_{(\alpha, \beta)}(y)$, the necessary stability condition (3.10) for the implicit scheme (1.2) is linear, it takes the form
\begin{equation}
(\cos \vartheta)\lambda_1(t) + \lambda_2(t) \leq \sin \vartheta \quad \forall t \in [0, T].
\end{equation}
Notice that the only difference between the sufficient and necessary stability conditions (1.19) and (3.15), respectively, for the implicit scheme (1.2) is that the former is a strict inequality and the latter a non-strict inequality.

3.2.2. The implicit–explicit scheme. In the case of an A-stable implicit method $(\alpha, \beta)$, the sufficient stability condition (1.20) for the implicit–explicit scheme (1.3) is sharp. Therefore, as in subsection 3.2.1, we assume that the implicit method $(\alpha, \beta)$ is A($\vartheta$)-stable, with $\vartheta < 90^\circ$ as large as possible.

In analogy to the function $K_{(\alpha, \beta)}$, see (3.4), for $-\tan \vartheta < y < \tan \vartheta$, we define the even function $K_{(\alpha, \beta, \gamma)}$ by
\begin{equation}
K_{(\alpha, \beta, \gamma)}(y) := \sup_{\alpha > 0} \max_{\zeta \in \mathcal{H}} \frac{|x\gamma(\zeta)|}{|\alpha(\zeta) + x(1 + iy)\beta(\zeta)|}.
\end{equation}
From (3.6) and the definition of $K_{(\alpha, \beta, \gamma)}(y)$, we easily infer that
\begin{equation}
K_{(\alpha, \beta, \gamma)}(y) = \max \left\{ \frac{1}{\sqrt{1 + y^2}} \max_{\zeta \in \mathcal{H}^+} \frac{|\gamma(\zeta)|}{|\beta(\zeta)|}, \sup_{\zeta \in \mathcal{H}^+} \frac{|d(\zeta)|}{\Im((1 - iy)d(\zeta))} \right\};
\end{equation}
compare to (3.7).

As in the case of the implicit method (1.2), see (3.10),
\begin{equation}
K_{(\alpha, \beta, \gamma)}(\lambda_1(t)) \lambda_2(t) \leq 1 \quad \forall t \in [0, T]
\end{equation}
is a necessary stability condition for the implicit–explicit scheme (1.3). A simple representation of \( K_{(\alpha, \beta, \gamma)} \) is unfortunately not available. Also, in contrast to \( K_{(\alpha, \beta)} \), the function \( K_{(\alpha, \beta, \gamma)} \) is in general not increasing for positive \( y \). Therefore, since
\[
K_{(\alpha, \beta, \gamma)}(y)\lambda_2(t) \leq 1 \quad \forall t \in [0, T],
\]
for all \( 0 \leq y \leq \lambda_1(t) \), is clearly also a necessary stability condition, we modify (3.18) as follows: With
\[
\tilde{K}_{(\alpha, \beta, \gamma)}(y) := \sup_{0 \leq s \leq y} K_{(\alpha, \beta, \gamma)}(s), \quad 0 \leq y < \tan \vartheta,
\]
a necessary stability condition for the implicit–explicit scheme (1.3) is
\[
\tilde{K}_{(\alpha, \beta, \gamma)}(\lambda_1(t))\lambda_2(t) \leq 1 \quad \forall t \in [0, T]. \tag{3.19}
\]

The left-hand sides of (1.20) and (3.19) do not coincide, in general; consequently, in contrast to the implicit schemes (1.2), it remains open, whether the sufficient stability condition (1.20) is best possible also among possibly nonlinear sufficient stability conditions for the implicit–explicit schemes (1.3) when the implicit method \((\alpha, \beta)\) is not A-stable. The discrepancy between the best possible linear sufficient stability condition (1.20) and the corresponding necessary stability condition (3.19) for concrete implicit–explicit schemes (1.3) can be studied at least computationally.

4. Additional sufficient stability conditions

In this section, we discuss additional sufficient stability conditions for the schemes (1.2) and (1.3). Here, the non-self-adjointness of \( A(t) \) is measured either by the smallest half-angle \( \varphi(t) \) of a sector containing its numerical range or by the norm of \( A(t) \), and is estimated by the ratio \( \nu(t)/\kappa(t) \).

4.1. Using the norm of \( A(t) \) as a measure of the non-self-adjointness of \( A(t) \).

As an alternative to (1.11), let us now use the norm of the operator \( A(t) : H \to H \) as a measure of the non-self-adjointness of the operator \( A(t) \),
\[
|A(t)v| \leq \hat{\lambda}_1(t)|v| \quad \forall v \in H \forall t \in [0, T]. \tag{4.1}
\]

In view of the relation (1.10) between the norms of \( A(t) \) and \( A_a(t) \), it is easily seen that if (1.11) holds, then (4.1) is valid with \( \hat{\lambda}_1(t) = \sqrt{1 + [\hat{\lambda}_1(t)]^2} \), and, conversely, if (4.1) holds, then (1.11) is valid with \( \lambda_1(t) = \sqrt{[\lambda_1(t)]^2 - 1} \). Thus, without loss of generality, we may assume that \( \lambda_1(t) \) and \( \hat{\lambda}_1(t) \) are related as follows
\[
[\hat{\lambda}_1(t)]^2 = 1 + [\lambda_1(t)]^2 \quad \forall t \in [0, T]. \tag{4.2}
\]

Then, the linear sufficient stability conditions (1.19) and (1.20) for the implicit scheme (1.2) and the implicit–explicit scheme (1.3), respectively, on \( \lambda_1(t) \) and \( \lambda_2(t) \) transform into the following, nonlinear in \( \hat{\lambda}_1(t) \) and linear in \( \lambda_2(t) \), sufficient stability conditions
\[
(cot \vartheta)\sqrt{[\hat{\lambda}_1(t)]^2 - 1 + K_{(\alpha, \beta)}\lambda_2(t)} < 1 \quad \forall t \in [0, T] \tag{4.3}
\]
and
\[
(cot \vartheta)\sqrt{[\hat{\lambda}_1(t)]^2 - 1 + K_{(\alpha, \beta, \gamma)}\lambda_2(t)} < 1 \quad \forall t \in [0, T], \tag{4.4}
\]
respectively. Now,

$$(\cot \vartheta) \sqrt{[\hat{\lambda}_1(t)]^2 - 1} < (\cos \vartheta) \hat{\lambda}_1(t) \iff (\cos \vartheta) \hat{\lambda}_1(t) < 1,$$

and we infer from (4.3) and (4.4), respectively, that

$$(\cos \vartheta) \hat{\lambda}_1(t) + K_{(\alpha, \beta)} \lambda_2(t) < 1 \quad \forall t \in [0, T]$$

and

$$(\cos \vartheta) \hat{\lambda}_1(t) + K_{(\alpha, \beta, \gamma)} \lambda_2(t) < 1 \quad \forall t \in [0, T]$$

are linear sufficient stability conditions on $\hat{\lambda}_1(t)$ and $\lambda_2(t)$ for the implicit scheme (1.2) and for the implicit–explicit scheme (1.3), respectively. However, as we shall now see, (4.5) and (4.6) are not best possible linear sufficient stability conditions for the implicit scheme (1.2) and for the implicit–explicit scheme (1.3), respectively, if the method $(\alpha, \beta)$ is not A-stable. First, in view of (1.8), using the von Neumann stability criterion, it is easily seen that the implicit scheme $(\alpha, \beta)$ is in general unstable for the linear part of

the differential equation in (1.1) (with vanishing nonlinear part $B$), if $(\cos \vartheta) \hat{\lambda}_1$ exceeds 1. Thus, the coefficient $\cos \vartheta$ of $\hat{\lambda}_1(t)$ in (4.5) and (4.6) cannot be replaced by a smaller coefficient; but the coefficients $K_{(\alpha, \beta)}$ and $K_{(\alpha, \beta, \gamma)}$ of $\lambda_2(t)$ in (4.5) and (4.6) can be replaced by $(\sin^2 \vartheta) K_{(\alpha, \beta)}$ and $(\sin^2 \vartheta) K_{(\alpha, \beta, \gamma)}$, respectively; see (4.8) and (4.10) in the sequel; notice that, since the nonlinear sufficient stability condition (4.3) for the implicit scheme (1.2) is sharp, (4.8) is one (among infinitely many) best possible linear sufficient stability condition on $\lambda_1$ and $\lambda_2$ for the implicit scheme (1.2).

First, we rewrite the sharp sufficient stability condition (4.3) in the form

$$(\sin^2 \vartheta) K_{(\alpha, \beta)} \lambda_2(t) < 1 =: f(\hat{\lambda}_1(t)) \quad \forall t \in [0, T]$$

and notice that the function $f$ is decreasing and strictly convex in the interval $[1, 1/\cos \vartheta]$. Replacing $f$ on the right-hand side of (4.7) by its linear Taylor polynomial, about some point $\tilde{x} \in [0, 1/\cos \vartheta]$, we are led to a linear sufficient stability condition for the implicit scheme (1.2), which is best possible in the sense that the corresponding coefficients of $\hat{\lambda}_1(t)$ and $\lambda_2(t)$ cannot be replaced by smaller coefficients. For instance, by Taylor expanding about $\tilde{x} := 1/\cos \vartheta$, we obtain the best possible linear sufficient stability condition

$$K_{(\alpha, \beta)} \lambda_2(t) < \frac{1}{\sin^2 \vartheta} \left[ 1 - (\cos \vartheta) \hat{\lambda}_1(t) \right],$$

i.e.,

$$(\cos \vartheta) \hat{\lambda}_1(t) + (\sin^2 \vartheta) K_{(\alpha, \beta)} \lambda_2(t) < 1 \quad \forall t \in [0, T];$$

compare to (4.5). This linear stability condition can also be written as

$$(\cos \vartheta) \hat{\lambda}_1(t) + (\sin \vartheta) \lambda_2(t) < 1 \quad \forall t \in [0, T].$$

The linear sufficient stability condition for the implicit–explicit scheme (1.3) corresponding to (4.8) reads

$$(\cos \vartheta) \hat{\lambda}_1(t) + (\sin^2 \vartheta) K_{(\alpha, \beta, \gamma)} \lambda_2(t) < 1 \quad \forall t \in [0, T];$$
compare to \((4.6)\). Notice that, due to the fact that \(\tilde{\lambda}_1(t) \geq 1\) in \((4.8)\) and \((4.10)\), whereas \(\lambda_1(t) \geq 0\) in \((1.19)\) and \((1.20)\), the coefficients of \(\lambda_2(t)\) in \((4.8)\) and \((4.10)\) are smaller than the corresponding coefficients in \((1.19)\) and \((1.20)\).

4.2. Using the smallest angle of a sector containing the numerical range of \(A(t)\) as a measure of its non-self-adjointness. As an alternative to the boundedness conditions \((1.11)\) or \((4.1)\), let us now use the smallest half-angle of a sector containing the numerical range of \(A(t)\) as a measure of its non-self-adjointness,

\[ (A(t)v, v) \in S_{\varphi(t)} \quad \forall v \in V \quad \forall t \in [0, T], \]

with \(0 \leq \varphi(t) \leq \vartheta\). Then, the boundedness condition \((1.11)\) is satisfied with \(\lambda_1(t) = \tan \varphi(t)\); see \((1.8)\). Therefore, the sufficient stability conditions \((1.19)\) and \((1.20)\), for the implicit multistep scheme \((1.2)\) and the implicit–explicit multistep scheme \((1.3)\), respectively, take the form

\[ (\cot \vartheta) \tan \varphi(t) + K(\alpha, \beta)\lambda_2(t) < 1 \quad \forall t \in [0, T] \]

and

\[ (\cot \vartheta) \tan \varphi(t) + K(\alpha, \beta, \gamma)\lambda_2(t) < 1 \quad \forall t \in [0, T], \]

respectively. Let us also note that \((4.12)\) can be rewritten as

\[ \lambda_2(t) < \frac{1}{\cos \varphi(t)} [\sin \vartheta \cos \varphi(t) - \sin \varphi(t) \cos \vartheta], \]

i.e., as

\[ \lambda_2(t) < \frac{\sin(\vartheta - \varphi(t))}{\cos \varphi(t)} \quad \forall t \in [0, T]; \]

see Figure 4.1 for the geometric interpretation. Due to the equivalence \((1.8)\), the sufficient stability condition \((4.12)\) for the implicit multistep scheme \((1.2)\) is also sharp.

4.3. Estimating the non-self-adjointness of \(A(t)\) by the ratio \(\nu(t)/\kappa(t)\). The commonly used ratio \(\nu(t)/\kappa(t)\) of the boundedness and coercivity functions, see \((1.4)\) and \((1.5)\), is also an estimate of the non-self-adjointness of the operator \(A(t)\); notice, however, that this ratio may be a crude estimate of the non-self-adjointness of \(A(t)\) since it depends on the choice of the specific norm \(\| \cdot \|\) on \(V\); see [2, Example 2.1]. According to \((1.4)\) and \((1.5)\), the time-dependent norms \(\| \cdot \|_t\),

\[ \|v\|_t := (A_s(t)v, v)^{1/2} = |A_s(t)^{1/2}v| \quad \forall v \in V, \]

first used in the stability analysis of BDF methods in [2, 12], are uniformly equivalent to \(\| \cdot \|\),

\[ \sqrt{\kappa(t)} \|v\| \leq \|v\|_t \leq \sqrt{\nu(t)} \|v\| \quad \forall v \in V. \]

We denote by \(\| \cdot \|_{*, t}\) the corresponding time-dependent dual norm on \(V'\),

\[ \forall v \in V' \quad \|v\|_{*, t} := \sup_{\tilde{v} \in V, \tilde{v} \neq 0} \frac{|(v, \tilde{v})|}{\|\tilde{v}\|_t} = |A_s(t)^{-1/2}v|. \]
Now, for \( \hat{v} \in V, \hat{v} \neq 0 \), using the first inequality in (4.16) and (1.5), with \( v := A_s(t)^{-1/2}\hat{v} \), we have
\[
\frac{|A(t)\hat{v}|}{|\hat{v}|} = \frac{\|A(t)v\|_{*,t}}{\|v\|_t} = \sup_{\tilde{v} \in V \setminus \{0\}} \frac{|(A(t)v, \tilde{v})|}{\|\tilde{v}\|_t \|v\|_t} \leq \frac{1}{\kappa(t)} \sup_{\tilde{v} \in V \setminus \{0\}} \frac{|(A(t)v, \tilde{v})|}{\|\tilde{v}\| \|v\|} = \frac{\|A(t)v\|_*}{\kappa(t)\|v\|} \leq \frac{\nu(t)}{\kappa(t)},
\]
i.e., \( |A(t)| \leq \nu(t)/\kappa(t) \). Consequently, replacing \( \lambda_1(t) \) in (4.3) and (4.4), respectively, by \( \nu(t)/\kappa(t) \), we obtain sufficient stability conditions
\[
(4.17) \quad (\cot \vartheta) \frac{1}{\kappa(t)} \sqrt{[\nu(t)]^2 - [\kappa(t)]^2} + K_{(\alpha,\beta)} \lambda_2(t) < 1 \quad \forall t \in [0, T]
\]
and
\[
(4.18) \quad (\cot \vartheta) \frac{1}{\kappa(t)} \sqrt{[\nu(t)]^2 - [\kappa(t)]^2} + K_{(\alpha,\beta,\gamma)} \lambda_2(t) < 1 \quad \forall t \in [0, T]
\]
for the implicit scheme (1.2) and the implicit–explicit scheme (1.3), respectively; these conditions are never more advantageous than (4.3) and (4.4), respectively, with \( \lambda_1(t) = |A(t)| \), and may be much more restrictive than (4.3) and (4.4). Notice, also, that with the norm \( \| \cdot \|_t \) on \( V \), (1.4) and (1.5) are satisfied with \( \kappa(t) = 1 \) and \( \nu(t) = |A(t)| \), that is, the inequality \( |A(t)| \leq \nu(t)/\kappa(t) \) holds as an equality for this norm.

Remark 4.1 (Comparison with the energy technique). Stability of implicit as well as of implicit–explicit BDF schemes of order up to 5 can be established also by the energy technique under linear sufficient stability conditions on \( \lambda_1(t) \) and \( \lambda_2(t) \); see [2, 9, 12]. In the interesting case of the three-, four-, and five-step methods, the stability conditions of [2, 9, 12] are more stringent than the corresponding conditions (4.5) and (4.6). In
particular, the conditions of \cite{2, 9, 12} are, for high order BDF schemes, not best possible linear sufficient stability conditions.

5. Application to Example \cite{1, 24}

In this section we briefly discuss the boundedness condition \eqref{1.11} in the case of the initial and boundary value problem \eqref{1.24}. We will see that the bound $\lambda_1(t)$ is the spectral radius of a matrix; in two simple cases more explicit forms of $\lambda_1(t)$ can be derived. To save space, we do not discuss the Lipschitz condition \eqref{1.12} here, since it is a reformulation of corresponding conditions used in, e.g., \cite{6, 12, 13, 3}; we refer to these articles and references therein for examples of operators satisfying \eqref{1.12} or modifications thereof in tubes defined in terms of $L^\infty$-based Sobolev norms.

Let $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_*$ denote the $L^2$ inner product and the product $L^2$ inner product in $L^2 := (L^2)^d := (L^2(\Omega))^d$, respectively; we denote by $| \cdot |$ both corresponding norms. With the notation of subsection \cite{1.3}, let us introduce in $V, V = H_0^1 = \mathcal{D}(A^{1/2}(t)), 0 \leq t \leq T$, the time-dependent norms $| \cdot |_t := |A_s^{1/2}(t)v|$. We identify $H$ with its dual, and denote by $V'$ the dual of $V, V' = H^{-1}$, and by $| \cdot |_{*,t}$ the time-dependent dual norms on $V'$, $|v|_{*,t} := |A_s^{-1/2}(t)v|$. We use the notation $\langle \cdot, \cdot \rangle$ also for the duality pairing between $V'$ and $V$; then $|v| := (A_s(t)v, v)^{1/2}$ and $|v|_{*,t} := (v, A_s^{-1}(t)v)^{1/2}$.

We shall see that the boundedness condition \eqref{1.11} is satisfied with

\begin{equation}
\lambda_1(t) := \max_{x \in \Omega} \rho(S(x, t)), \quad t \in [0, T],
\end{equation}

where $\rho(\cdot)$ is the spectral radius of the antihermitian matrices $S(x, t) := \mathcal{O}^{-1/2}(x, t) \mathcal{O}(x, t)\mathcal{O}^{-1/2}(x, t)$.

First, clearly,

\begin{equation}
(A_s v, \tilde{v}) = \sum_{i,j=1}^d (a_{ij} v_{x_j}, \tilde{v}_{x_i}) = \sum_{i=1}^d \left( \sum_{j=1}^d a_{ij} v_{x_j}, \tilde{v}_{x_i} \right) = \sum_{i=1}^d \left( (\mathcal{O} \nabla v)_i, \tilde{v}_{x_i} \right),
\end{equation}

i.e.,

\begin{equation}
(A_s(t)v, \tilde{v}) = \langle \mathcal{O}(\cdot, t) \nabla v, \nabla \tilde{v} \rangle.
\end{equation}

In particular, $(A_s(t)v, v) = \langle \mathcal{O}(\cdot, t) \nabla v, \nabla v \rangle$, whence

\begin{equation}
|v|_t = |A_s^{1/2}(t)v| = |\mathcal{O}^{1/2}(\cdot, t) \nabla v|.
\end{equation}

Now,

\begin{equation}
(A_a(t)v, \tilde{v}) = \langle \mathcal{O}(\cdot, t) \nabla v, \nabla \tilde{v} \rangle = \langle S(\cdot, t) \mathcal{O}^{1/2}(\cdot, t) \nabla v, \mathcal{O}^{1/2}(\cdot, t) \nabla \tilde{v} \rangle,
\end{equation}

whence

\begin{equation}
|(A_a(t)v, \tilde{v})| \leq \max_{x \in \Omega} |S(x, t)|_2 |\mathcal{O}^{1/2}(\cdot, t) \nabla v| |\mathcal{O}^{1/2}(\cdot, t) \nabla \tilde{v}|,
\end{equation}

with $| \cdot |_2$ denoting the spectral (Euclidean) norm of a matrix. Therefore,

\begin{equation}
|(A_a(t)v, \tilde{v})| \leq \max_{x \in \Omega} |S(x, t)|_2 |A_s^{1/2}(t)v| A_s^{1/2}(t) \tilde{v}, \quad t \in [0, T],
\end{equation}
and thus, since the spectral norms of the antihermitian matrices $S(x,t)$ are equal to their spectral radii,

$$|A_s^{-1/2}(t)A_a(t)v| \leq \lambda_1(t)|A_s^{-1/2}(t)v| \quad \forall v \in V \ \forall t \in [0,T],$$

i.e.,

$$|A_s^{-1/2}(t)A_a(t)A_s^{-1/2}(t)v| \leq \lambda_1(t)|v| \quad \forall v \in H \ \forall t \in [0,T],$$

with $\lambda_1$ as given in (5.1). Thus (1.11) is satisfied with the function $\lambda_1$ of (5.1).

Next we consider two special cases. In the first case, (5.1) takes a very simple form; in the second, we give an explicit form of the matrices $S(x,t)$; their spectral radii can then be easily computed.

5.1. Two special cases. First case: Assume first that

$$\tilde{\mathcal{O}}(x,t) = ia(x,t)\mathcal{O}(x,t), \quad x \in \Omega, \quad 0 \leq t \leq T,$$

with $a$ a smooth real-valued function, $a : \Omega \times [0,T] \to \mathbb{R}$. Then, $S(x,t) = ia(x,t)I_d$ with $I_d$ the $d \times d$ unit matrix; thus,

$$\lambda_1(t) = \max_{x \in \Omega} |a(x,t)| \quad \forall t \in [0,T].$$

The analysis of implicit–explicit multistep methods in this particular case, with a matrix $\mathcal{O}$ independent of $t$ and a function $a$ independent of $x$, was the subject of [3]. If

$$\mathcal{O}(x,t) = a(x,t)I_d, \quad \tilde{\mathcal{O}}(x,t) = i\tilde{a}(x,t)I_d, \quad x \in \Omega, \quad 0 \leq t \leq T,$$

with $a$ and $\tilde{a}$ smooth real-valued functions, $a, \tilde{a} : \Omega \times [0,T] \to \mathbb{R}$, then, the parabolic equation in (1.24) takes the form

$$u_t - \nabla \cdot ((a(x,t) + i\tilde{a}(x,t))\nabla u) = B(t,u).$$

In this case

$$\tilde{\mathcal{O}}(x,t) = i\frac{\tilde{a}(x,t)}{a(x,t)}\mathcal{O}(x,t), \quad x \in \Omega, \quad 0 \leq t \leq T,$$

and, according to (5.6),

$$\lambda_1(t) = \max_{x \in \Omega} \frac{|\tilde{a}(x,t)|}{a(x,t)} \quad \forall t \in [0,T].$$

Second case: Here we shall consider the general case in two space dimensions, $d = 2$. It is well known that

$$\mathcal{O}^{1/2} = \frac{1}{\sqrt{\text{tr} \mathcal{O} + 2\sqrt{\det \mathcal{O}}}}(\mathcal{O} + \sqrt{\det \mathcal{O}} I_2),$$

with $\text{tr} \mathcal{O} := a_{11} + a_{22}$ the trace of $\mathcal{O}$. An easy way to check this is by means of the Cayley–Hamilton theorem, which in the case of $2 \times 2$ matrices yields $\mathcal{O}^2 + (\det \mathcal{O})I_2 = (\text{tr} \mathcal{O})\mathcal{O}$; then,

$$(\mathcal{O} + \sqrt{\det \mathcal{O}} I_2)^2 = (\text{tr} \mathcal{O} + 2\sqrt{\det \mathcal{O}}) \mathcal{O},$$

and (5.9) follows. Furthermore, multiplying (5.9) by $\mathcal{O}^{-1/2}$, we easily see that

$$\mathcal{O}^{-1/2} = \frac{1}{\sqrt{\det \mathcal{O} \sqrt{\text{tr} \mathcal{O} + 2\sqrt{\det \mathcal{O}}}}}
\left( (\text{tr} \mathcal{O} + \sqrt{\det \mathcal{O}}) I_2 - \mathcal{O} \right),$$
Therefore, the antihermitian matrices $S(x,t) = \mathcal{O}^{-1/2}(x,t)\tilde{\mathcal{O}}(x,t)\mathcal{O}^{-1/2}(x,t)$ take in this case the form

$$S = \frac{1}{\det \mathcal{O} \left( \text{tr} \mathcal{O} + 2\sqrt{\det \mathcal{O}} \right)}\left(c_\mathcal{O}^2 \tilde{\mathcal{O}} - c_\mathcal{O}(\mathcal{O}\tilde{\mathcal{O}} + \tilde{\mathcal{O}}\mathcal{O}) + \mathcal{O}\tilde{\mathcal{O}}\mathcal{O}\right),$$

with the constant $c_\mathcal{O} := \text{tr} \mathcal{O} + \sqrt{\det \mathcal{O}}$.

REFERENCES


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

E-mail address: akrivis@cse.uoi.gr