Stability of IMEX Runge-Kutta methods for delay differential equations

Toshiyuki Koto

Graduate School of Information Science
Nagoya University
Nagoya 464-8601, Japan
email: koto@is.nagoya-u.ac.jp

Abstract

Stability of IMEX (implicit-explicit) Runge-Kutta methods applied to delay differential equations (DDEs) is studied on the basis of the scalar test equation $du/dt = \lambda u(t) + \mu u(t-\tau)$, where τ is a constant delay and λ, μ are complex parameters. More specifically, P-stability regions of the methods are defined and analyzed in the same way as in the case of the standard Runge-Kutta methods. A new IMEX method which possesses a superior stability property for DDEs is proposed. Some numerical examples which confirm the results of our analysis are presented.

Key words: IMEX Runge-Kutta methods, delay differential equations, P-stability regions

1 Introduction

Let us consider ordinary differential equations (ODEs) of the form

$$\frac{du}{dt} = Lu(t) + g\Big(t, u(t)\Big), \tag{1.1}$$

where $u(t)$ is a vector valued unknown function and L is a square matrix. We suppose that $Lu(t)$ on the right hand side gives a stiff term, i.e., L has eigenvalues whose moduli are so large. A typical example of such equations arises after the spatial discretization of a partial differential equation of reaction-diffusion type, and some special numerical methods for solving (1.1) have been proposed along the idea of treating the linear stiff term by an implicit scheme with a superior stability property and the nonlinear term by an explicit scheme. A prototype of such methods, called IMEX (implicit-explicit) method, is the IMEX Θ-method [8] defined by

$$u_{n+1} = u_n + \Delta t(1-\Theta)Lu_n + \Delta t\,\Theta L u_{n+1} + \Delta t\, g(t_n,\, u_n),$$

where Θ is a parameter with $\Theta \geq 1/2$, Δt is the stepsize, $t_n := t_0 + n\Delta t$, and u_n denotes an approximate value of $u(t_n)$. This is a mixture of the standard Θ-method and the Euler method, and of order 1 in accuracy.

This simple method can be improved in terms of accuracy by generalizing the method along the idea of Runge-Kutta methods. Consider a pair of two Runge-Kutta methods represented by the arrays

$$\begin{array}{c|ccccc} 0 & 0 & 0 & & \cdots & 0 \\ c_2 & a_{21} & a_{22} & 0 & \cdots & 0 \\ c_3 & a_{31} & a_{32} & a_{33} & & \vdots \\ \vdots & \vdots & & & \ddots & 0 \\ c_s & a_{s1} & a_{s2} & \cdots & a_{s,s-1} & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_{s-1} & b_s \end{array}\,, \qquad \begin{array}{c|ccccc} 0 & 0 & 0 & & \cdots & 0 \\ c_2 & \widehat{a}_{21} & 0 & & \cdots & 0 \\ c_3 & \widehat{a}_{31} & \widehat{a}_{32} & 0 & & \vdots \\ \vdots & \vdots & & & \ddots & 0 \\ c_s & \widehat{a}_{s1} & \widehat{a}_{s2} & \cdots & \widehat{a}_{s,s-1} & 0 \\ \hline & \widehat{b}_1 & \widehat{b}_2 & \cdots & \widehat{b}_{s-1} & \widehat{b}_s \end{array}\,, \tag{1.2}$$

with the same abscissas c_i. The former corresponds to a diagonally implicit method and the latter corresponds to an explicit method. We assume that a_{ij} and $\widehat{a}_{ij}$ satisfy

$$c_i = \sum_{j=1}^{i} a_{ij} = \sum_{j=1}^{i-1} \widehat{a}_{ij}\,,\ \ i = 1,\, \dots,\, s. \tag{1.3}$$

Then, an s-stage IMEX Runge-Kutta method for (1.1) is defined by

$$\begin{aligned} U_{n,i} &= u_n + \Delta t \sum_{j=1}^{i} a_{ij}\, LU_{n,j} \ + \ \Delta t \sum_{j=1}^{i-1} \widehat{a}_{ij}\, g\Big(t_{n,j}, U_{n,j}\Big),\ \ i = 1,\, \dots,\, s, \\ u_{n+1} &= u_n + \Delta t \sum_{i=1}^{s} b_i\, LU_{n,i} \ + \ \Delta t \sum_{i=1}^{s} \widehat{b}_i\, g\Big(t_{n,i}, U_{n,i}\Big). \end{aligned} \tag{1.4}$$

Here, $t_{n,i} := t_n + c_i\Delta t$, and $U_{n,i}$ are intermediate variables, which are successively computed by solving linear equations. By virtue of (1.3), each $U_{n,i}$ is interpreted as an approximate value of $u(t_{n,i})$.

The accuracy and stability of the IMEX Runge-Kutta methods have been studied by several authors. In particular, methods of order 2 and order 3 are constructed by Ascher, Ruuth and Spiteri [2] (see also [16]), methods of order 4 are constructed by Calvo, de Frutos and Novo [5], Kennedy and Carpenter [11]. A stability property of the methods is also examined in [5].

On the other hand, in some fields of applied mathematics such as mathematical biology and control theory, reaction-diffusion equations with time delay in the reaction term are used for studying the effects of interaction of time delay and spatial diffusion (see, e.g., [14]). In the simplest case, the governing equation becomes a delay differential equation (DDE) of the form

$$\frac{du}{dt} = Lu(t) + g\big(t, u(t), u(t-\tau)\big) \tag{1.5}$$

after the spatial discretization, e.g., by the method of lines (MOL) approach, where τ is a positive constant. As well as many numerical methods for ODEs, IMEX Runge-Kutta methods can be adapted to the DDE (1.5). However, in some numerical experiments, instability phenomena (cf. Sect. 5) are observed which are not easily predictable from numerical results in the case of the usual reaction-diffusion equations without delay (see, e.g., [9], Chap. IV, Sect. 6).

In this paper, we study stability of IMEX Runge-Kutta methods using the scalar test equation

$$\frac{du}{dt} = \lambda u(t) + \mu u(t-\tau), \tag{1.6}$$

proposed by Barwell [3]. Here, $\lambda,\ \mu$ are complex parameters, and if $\lambda,\ \mu$ satisfy

$$|\mu| < -\operatorname{Re}\lambda, \tag{1.7}$$

the zero solution of (1.6) is asymptotically stable for any $\tau > 0$. Until now many studies have been carried out whether numerical methods preserve this asymptotic property or not (see [4] and the references therein). In the context of reaction-diffusion equations, this corresponds to studying whether numerical methods preserve the asymptotic property of solutions to diffusion dominant equations.

This paper is organized as follows. In Sect. 2, we consider the so-called constant step size method, an adaptation of the IMEX method (1.4) for the DDE (1.5) taking advantage of the constancy of the delay τ. We define P-stability regions of the IMEX methods and give a characterization of the regions in the same way as in the case of the standard Runge-Kutta methods. Moreover, we examine P-stability regions of some specific methods in Sect. 3, and modify the definition and the characterization of the regions into those in the case where (1.4) is applied to (1.5) by making use of a continuous extension of

(1.4) in Sect. 4. In Sect. 5, we present some numerical examples which suggest practicality of our stability analysis.

2 Natural IMEX Runge-Kutta methods

Let m be a positive integer and consider a constant stepsize of the form

$$\Delta t = \frac{\tau}{m}\,. \tag{2.1}$$

Then, $t_n - \tau = t_{n-m}$ holds and we can regard $U_{n-m,i}$ as an approximate value of $u(t_{n,i} - \tau)$. An adaptation of (1.4) for (1.5) is given by

$$\begin{aligned} U_{n,i} &= u_n + \Delta t \sum_{j=1}^{i} a_{ij}\, LU_{n,j} \;+\; \Delta t \sum_{j=1}^{i-1} \widehat{a}_{ij}\, g\Big(t_{n,j}, U_{n,j}, U_{n-m,j}\Big), \\ u_{n+1} &= u_n + \Delta t \sum_{i=1}^{s} b_i\, LU_{n,i} \;+\; \Delta t \sum_{i=1}^{s} \widehat{b}_i\, g\Big(t_{n,i}, U_{n,i}, U_{n-m,i}\Big). \end{aligned} \tag{2.2}$$

By the standard argument (see, e.g., [6], II.17), it is verified that this adaptation preserves the order of accuracy of the underlying IMEX Runge-Kutta method (1.4). In particular, if the coefficients satisfy

$$\sum_{i=1}^{s} b_i = 1, \quad \sum_{i=1}^{s} b_i\, c_i = \frac{1}{2}, \quad \sum_{i=1}^{s} \widehat{b}_i = 1, \quad \sum_{i=1}^{s} \widehat{b}_i\, c_i = \frac{1}{2}, \tag{2.3}$$

the IMEX method (2.2) shows second order convergence for (1.5).

By applying (2.2) to the test equation (1.6), we obtain the difference equation

$$\begin{aligned} U_{n,i} &= u_n + \Delta t \sum_{j=1}^{i} a_{ij}\, \lambda U_{n,j} \;+\; \Delta t \sum_{j=1}^{i-1} \widehat{a}_{ij}\, \mu U_{n-m,j}, \\ u_{n+1} &= u_n + \Delta t \sum_{i=1}^{s} b_i\, \lambda U_{n,i} \;+\; \Delta t \sum_{i=1}^{s} \widehat{b}_i\, \mu U_{n-m,i}. \end{aligned} \tag{2.4}$$

By means of

$$\begin{aligned} &A = [a_{ij}]_{i,j=1}^{s},\ \widehat{A} = [\widehat{a}_{ij}]_{i,j=1}^{s},\ b = [b_1, \ldots, b_s]^T,\ \ \widehat{b}^T = [\widehat{b}_1, \ldots, \widehat{b}_s]^T, \\ &U_n = [U_{n,1}, \ldots, U_{n,s}]^T,\ \mathbf{1} = [1, \ldots, 1]^T, \end{aligned}$$

we can rewrite (2.4) in the form

$$
\begin{aligned}
U_n &= u_n \mathbf{1} + \alpha A U_n + \beta \widehat{A}\, U_{n-m}, \\
u_{n+1} &= u_n + \alpha b^T U_n + \beta \widehat{b}^T U_{n-m},
\end{aligned}
\tag{2.5}
$$

where $\alpha := \Delta t \lambda$, $\beta := \Delta t \mu$. The P-stability region of the IMEX method (2.2) is defined as follows.

Definition 2.1 *The P-stability region of the method (2.2) is the set S_P of the pair of complex numbers (α, β), such that* $\det[I - \alpha A] \neq 0$ *and the zero solution of (2.5) is asymptotically stable for any $m \geq 1$.*

Let $r(\alpha)$ denote the stability function of the Runge-Kutta method defined by A and b, and let S_A denote the stability region of the method, i.e.,

$$
r(\alpha) = 1 + \alpha b^T (I - \alpha A)^{-1} \mathbf{1}, \quad S_A = \{\alpha \in \mathbb{C} \,:\, |r(\alpha)| < 1\}.
$$

Moreover, put

$$
P_\alpha(z) = \det[I - \alpha A - z\widehat{A} + \alpha \mathbf{1} b^T + z \mathbf{1} \widehat{b}^T], \quad Q_\alpha = \det[I - \alpha A] \tag{2.6}
$$

and define the set Γ_α by

$$
\Gamma_\alpha = \{z \in \mathbb{C} \,:\, |P_\alpha(z)| = |Q_\alpha|\,\}. \tag{2.7}
$$

We can rewrite Γ_α as

$$
\Gamma_\alpha = \{z \in \mathbb{C} \,:\, P_\alpha(z) - Q_\alpha \mathrm{e}^{\mathrm{i}\vartheta} = 0,\ 0 \leq \vartheta < 2\pi\}.
$$

Hence, Γ_α is a closed curve, and there is a point on Γ_α which achieves the value

$$
\sigma_\alpha = \inf\{|z| \,:\, z \in \Gamma_\alpha\}. \tag{2.8}
$$

Using this value we can characterize P-stability regions as follow.

Theorem 2.2 *Assume that* $\det[I - \alpha A] \neq 0$ *and consider the following three statements:*

(a) $\alpha \in S_A$ *and* $|\beta| < \sigma_\alpha$*;*

(b) $(\alpha, \beta) \in S_P$*;*

(c) $\alpha \in S_A$ *and* $|\beta| \leq \sigma_\alpha$.

Then, we have (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c).

Putting $R_\alpha(z) = P_\alpha(z)/Q_\alpha$, we further rewrite Γ_α as

$$\Gamma_\alpha = \{z \in \mathbb{C} : |R_\alpha(z)| = 1\}.$$

Theorem 2.2, which can be proved by Lemma 8 in [15], may be regarded as a special case of Theorem 13 in [15]. Nevertheless, Theorem 2.2 has a significance especially from a viewpoint of application. IMEX Runge-Kutta methods were not supposed to be an object of study in [15]. Also, we present a proof using a generalization of Lemma 8 in [15] by in 't Hout and Spijker [10], which gives a better perspective as to a variation of Theorem 2.2 described in Sect. 4 (see also [12] on a similar application of the generalization).

PROOF. Since (2.5) is rewritten as

$$\begin{bmatrix} I - \alpha A & 0 \\ -\alpha b^T & 1 \end{bmatrix} \begin{bmatrix} U_n \\ u_{n+1} \end{bmatrix} + \begin{bmatrix} 0 & -\mathbf{1} \\ 0 & -1 \end{bmatrix} \begin{bmatrix} U_{n-1} \\ u_n \end{bmatrix} + \begin{bmatrix} -\beta \widehat{A} & 0 \\ -\beta \widehat{b}^T & 0 \end{bmatrix} \begin{bmatrix} U_{n-m} \\ u_{n-m+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

the characteristic equation of (2.5) is given by

$$\det[\lambda^m \boldsymbol{V}(\lambda) + \boldsymbol{W}(\lambda)] = 0, \tag{2.9}$$

where

$$\boldsymbol{V}(\lambda) = \begin{bmatrix} \lambda(I - \alpha A) & -\mathbf{1} \\ -\lambda \alpha b^T & \lambda - 1 \end{bmatrix}, \quad \boldsymbol{W}(\lambda) = \begin{bmatrix} -\lambda \beta \widehat{A} & 0 \\ -\lambda \beta \widehat{b}^T & 0 \end{bmatrix}.$$

As for the equation (2.9), consider following three statements:

(a)′ $\det \boldsymbol{V}(\lambda) \neq 0$ for any $|\lambda| \geq 1$ and $\sup_{|\lambda|=1} \rho[\boldsymbol{V}(\lambda)^{-1} \boldsymbol{W}(\lambda)] < 1$;

(b)′ all the roots of (2.9) lie inside the unit circle for any $m \geq 1$;

(c)′ $\det \boldsymbol{V}(\lambda) \neq 0$ for any $|\lambda| \geq 1$ and $\sup_{|\lambda|=1} \rho[\boldsymbol{V}(\lambda)^{-1} \boldsymbol{W}(\lambda)] \leq 1$.

Here, $\rho[\,\cdot\,]$ denotes the spectral radius of a matrix.

It is easily verified that

$$\deg\{\det[\lambda^m \boldsymbol{V}(\lambda) + \boldsymbol{W}(\lambda)]\} = \deg\{\det[\lambda^m \boldsymbol{V}(\lambda)] = (m+1)(s+1)$$

for any $m \geq 1$. Hence, by Corollary 1.2 in [10], we have (a)$' \Longrightarrow$ (b)$' \Longrightarrow$ (c)$'$.

Since $\boldsymbol{V}(\lambda)$ is rewritten as

$$\boldsymbol{V}(\lambda) = \begin{bmatrix} \lambda I & 0 \\ -\lambda\alpha b^T(I-\alpha A)^{-1} & \lambda - r(\alpha) \end{bmatrix} \begin{bmatrix} I-\alpha A & -\lambda^{-1}\mathbf{1} \\ 0 & 1 \end{bmatrix},$$

we get

$$\det \boldsymbol{V}(\lambda) = \lambda^s\big(\lambda - r(\alpha)\big)\det[I-\alpha A]. \tag{2.10}$$

Thus, $\det \boldsymbol{V}(\lambda) \neq 0$ for any $|\lambda| \geq 1$ if and only if $\alpha \in S_A$.

Moreover, for $\zeta \in \mathbb{C}$, we have

$$\boldsymbol{V}(\lambda) + \zeta\boldsymbol{W}(\lambda) = \begin{bmatrix} \lambda I & 0 \\ -\lambda L_\alpha(\beta\zeta) & \lambda - R_\alpha(\beta\zeta) \end{bmatrix} \begin{bmatrix} M_\alpha(\beta\zeta) & -\lambda^{-1}\mathbf{1} \\ 0 & 1 \end{bmatrix},$$

where $M_\alpha(z) = I - \alpha A - z\widehat{A},\ L_\alpha(z) = (\alpha b^T + z\widehat{b}^T)(I-\alpha A - z\widehat{A})^{-1}$,

$$R_\alpha(z) = 1 + (\alpha b^T + z\widehat{b}^T)(I - \alpha A - z\widehat{A})^{-1}\mathbf{1},$$

and $R_\alpha(z)$ is rewritten as $R_\alpha(z) = P_\alpha(z)/Q_\alpha$ by Cramer's rule. Hence, we get

$$\det[\boldsymbol{V}(\lambda) + \zeta\boldsymbol{W}(\lambda)] = \lambda^s[Q_\alpha\lambda - P_\alpha(\beta\zeta)], \quad \zeta \in \mathbb{C}, \tag{2.11}$$

which implies that if $\alpha \in S_A$,

$$\rho[\boldsymbol{V}(\lambda)^{-1}\boldsymbol{W}(\lambda)] = \Big(\inf\{|\zeta| \,:\, Q_\alpha\lambda - P_\alpha(\beta\zeta) = 0\}\Big)^{-1} \tag{2.12}$$

for $|\lambda| \geq 1$. From this it is verified that $\sup_{|\lambda|=1} \rho[\boldsymbol{V}(\lambda)^{-1}\boldsymbol{W}(\lambda)] < 1$ if and only if $|\beta| < \sigma_\alpha$ and $\sup_{|\lambda|=1} \rho[\boldsymbol{V}(\lambda)^{-1}\boldsymbol{W}(\lambda)] \leq 1$ if and only if $|\beta| \leq \sigma_\alpha$. $\square$

3 P-stability regions

In this section, we investigate P-stability regions of several IMEX Runge-Kutta methods. First, we consider the IMEX Θ-method, represented by the arrays

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1-\Theta & \Theta \\ \hline & 1-\Theta & \Theta \end{array}\,, \qquad \begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & 1 & 0 \end{array}\,.$$

As is well known, the Θ-method is A-stable; the stability region of the method is the outside of a circle with center $1/(2\Theta-1)$ and radius $1/(2\Theta-1)$ when $\Theta > 1/2$ and the left half plane when $\Theta = 1/2$. Moreover, by the equations (2.6) we get

$$P_\alpha(z) = 1 + (1-\Theta)\alpha + z, \quad Q_\alpha = 1 - \Theta\alpha.$$

Hence, the set Γ_α is a circle with center $-1-(1-\Theta)\alpha$ and radius $|1-\Theta\alpha|$, and we have

$$\sigma_\alpha = \Big| |1-\Theta\alpha| - |1+(1-\Theta)\alpha| \Big|. \tag{3.1}$$

When $\Theta = 1$,

$$\sigma_\alpha = \Big| |1-\alpha| - 1 \Big| > |\alpha| \geq -\operatorname{Re}\alpha$$

holds for any $\operatorname{Re}\alpha < 0$, which implies that the IMEX Θ-method with $\Theta = 1$ is P-stable, i.e., the P-stability region of the method contains the region $\{|\beta| < -\operatorname{Re}\alpha\}$. On the other hand, in the case α is a negative real number, (3.1) is rewritten as

$$\sigma_\alpha = \begin{cases} -\alpha & \big(1+(1-\Theta)\alpha \geq 0\big) \\ 2+(1-2\Theta)\alpha & \big(1+(1-\Theta)\alpha < 0\big) \end{cases}.$$

Hence, the IMEX Θ-method is not P-stable when $\Theta < 1$ (cf. Fig. 1).

In general, it is difficult to find P-stability regions by an analytic method. A certain numerical method is necessary to investigate the regions.

Let N be a positive integer, and divide the interval $[0,\, 2\pi)$ into N equal parts:

$$0 = \vartheta_0 < \vartheta_1 < \cdots < \vartheta_k = k\Delta\vartheta < \cdots < \vartheta_N = 2\pi, \;\; \Delta\vartheta = 2\pi/N.$$

The set Γ_α is approximated by

$$\Gamma_{\alpha,N} = \{ z \in \mathbb{C} \,:\, P_\alpha(z) - Q_\alpha \mathrm{e}^{\mathrm{i}\vartheta_k} = 0, \;\; k = 0,\, 1,\, \dots,\, N-1 \},$$

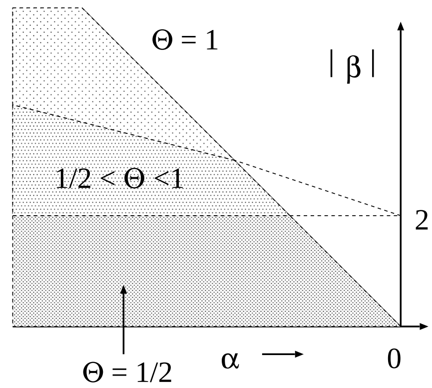

Fig. 1. P-stability regions of the IMEX Θ-methods.

and we have

$$\sigma_\alpha = \lim_{N\to\infty} \sigma_{\alpha,N}, \quad \sigma_{\alpha,N} = \inf\{|z| \,:\, z \in \Gamma_{\alpha,N}\},$$

which follows from the fact that the roots of $P_\alpha(z) - Q_\alpha\, e^{i\vartheta} = 0$ are continuous functions of ϑ. Thus, we can obtain an approximate value of σ_α by taking a sufficiently large N and solving N algebraic equations with some numerical method, e.g., Durand-Kerner's algorithm (see, e.g., [1,13]).

We compare two IMEX methods defined by the arrays

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1/2 & 1/2 \\ \hline & 1/2 & 1/2 \end{array}\;, \quad \begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array} \tag{3.2}$$

and

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 \\ 1 & 0 & 1-\gamma & \gamma \\ \hline & 0 & 1-\gamma & \gamma \end{array}\;, \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 \\ 1 & \delta & 1-\delta & 0 \\ \hline & \delta & 1-\delta & 0 \end{array}\;, \quad \gamma = \frac{2-\sqrt{2}}{2}\;, \quad \delta = 1 - \frac{1}{2\gamma}\;. \tag{3.3}$$

The former is a combination of the trapezoidal rule and Heun's second order method (the explicit trapezoidal rule). The corresponding method is called the IMEX trapezoidal rule in [9]. The latter is a scheme constructed by Ascher, Ruuth and Spiteri [2] on the basis of an L-stable singly diagonally Runge-Kutta method of order 2. Both methods are of order 2 as IMEX method. By the equations (2.6) we have

$$P_\alpha(z) = 1 + \frac{\alpha}{2} + \Big(1 + \frac{\alpha}{2}\Big)z + \frac{z^2}{2}, \quad Q_\alpha = 1 - \frac{\alpha}{2}$$

for the IMEX trapezoidal rule (3.2) and

$$P_\alpha(z) = 1 - \alpha + \sqrt{2}\alpha + (1 - \alpha + \sqrt{2}\alpha)z + \frac{z^2}{2}, \quad Q_\alpha = (1 - \gamma\alpha)^2$$

for the Ascher-Ruuth-Spiteri (ARS) method (3.3).

For both methods, approximate values of σ_α are computed by solving quadratic equations. Fig. 2 shows the P-stability regions of the two IMEX methods in the case α is a negative real number. The boundaries determined by σ_α were drawn by taking $N = 1000$ and solving the quadratic equations by the quadratic formula. In the case of the IMEX trapezoidal rule, the width of the region decreases and tends to zero as $|\alpha|$ increases, whereas the width increases infinitely in the case of the ARS method (3.3). This suggests that the ARS method has a better stability property than the IMEX trapezoidal rule as for DDEs.

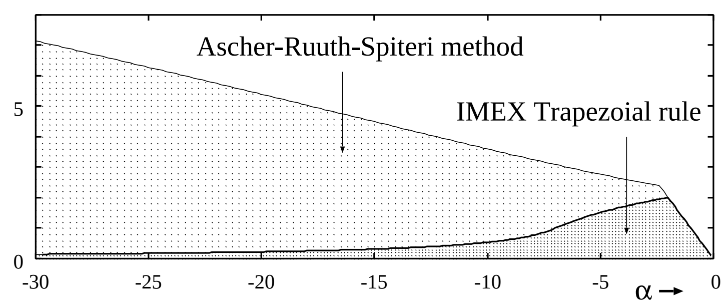

Fig. 2. P-stability regions of the methods (3.2) and (3.3).

The P-stability region of the ARS method is larger than that of the IMEX trapezoidal rule, but it is rather small compared with those of P-stable methods. By increasing the number of stages, we can construct a method with a larger P-stability region.

As an illustration, we consider a 4-stage IMEX Runge-Kutta method defined by the arrays

$$
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 \\
\frac{1}{2} & 0 & -\frac{1}{2} & 1 & 0 \\
1 & 0 & -1 & 1 & 1 \\
\hline
 & 0 & -1 & 1 & 1
\end{array}\ , \qquad
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 \\
\hline
 & 0 & 0 & 1 & 0
\end{array}\ . \tag{3.4}
$$

This method satisfies (2.3), i.e., it is of order 2. The stability function of the diagonally implicit method is

$$
r(\alpha) = \frac{1 - 2\alpha + \alpha^2/2}{(1-\alpha)^3}\ ,
$$

and it is easily verified that the corresponding method is L-stable. Moreover, we have

$$
P_\alpha(z) = 1 - 2\alpha + \frac{\alpha^2}{2} + (1-2\alpha)z + \Big(\frac{1}{2} - \alpha\Big)z^2, \quad Q_\alpha = (1-\alpha)^3, \tag{3.5}
$$

and we can compute approximate values of σ_α by solving quadratic equations. We plot σ_α versus $\operatorname{Re}\alpha$ for some fixed values of $\operatorname{Im}\alpha$ in Fig. 3. Each curve indicates the approximate values of σ_α computed with $N = 1000$. This figure suggests that $\sigma_\alpha \geq -\operatorname{Re}\alpha$, i.e., the method is P-stable, which can be verified analytically in the case α is real.

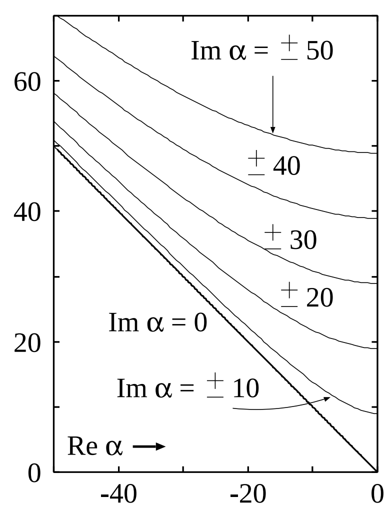

Fig. 3. Values of σ_α for the method (3.4).

Proposition 3.1 *Consider the IMEX method defined by the arrays (3.4). If α is a negative real number, we have $\sigma_\alpha = -\alpha$.*

PROOF. From (3.5) it follows that $P_\alpha(-\alpha) = Q_\alpha$. Hence we have $-\alpha \in \Gamma_\alpha$ and $\sigma_\alpha \le -\alpha$.

To show $\sigma_\alpha \ge -\alpha$, it suffices to verify that $|P_\alpha(z)| \le |Q_\alpha|$ for any $|z| = -\alpha$, since it implies that $|P_\alpha(z)| < |Q_\alpha|$ for any $|z| < -\alpha$ by the maximal modulus principle.

Differentiating $\rho(\vartheta) = |P_\alpha(-\alpha\, \mathrm{e}^{\mathrm{i}\vartheta})|^2$, $0 \le \vartheta < 2\pi$, we obtain

$$\rho'(\vartheta) = 2\alpha^2(1-2\alpha)(2-4\alpha+\alpha^2)\sin(\vartheta)\left(\frac{1-2\alpha+\alpha^2-\alpha^3}{\alpha(2-4\alpha+\alpha^2)} - \cos(\vartheta)\right).$$

When $-1/\sqrt{3} \le \alpha < 0$, we have $(1-2\alpha+\alpha^2-\alpha^3)/[\alpha(2-4\alpha+\alpha^2)] \le -1$, and $\rho(\vartheta)$ has a maximum at $\vartheta = 0$ and a minimum at $\vartheta = \pi$. When $\alpha < -1/\sqrt{3}$, we have $-1 < (1-2\alpha+\alpha^2-\alpha^3)/[\alpha(2-4\alpha+\alpha^2)] < 0$, and $\rho(\vartheta)$ has relative maximums both at $\vartheta = 0$ and $\vartheta = \pi$, but $\rho(\pi) = (1-\alpha-\alpha^2-\alpha^3)^2 < \rho(0) = |Q_\alpha|^2$ for $\alpha < 0$. Hence, in both cases, $\rho(\vartheta) \le \rho(0) = |Q_\alpha|^2$ for any $0 \le \vartheta < 2\pi$, which implies that $|P_\alpha(z)| \le |Q_\alpha|$ for any $|z| = -\alpha$. $\square$

4 Continuous IMEX Runge-Kutta methods

We can define continuous extensions (dense outputs) of IMEX Runge-Kutta methods in the same way as in the case of standard Runge-Kutta methods. Let $w_i(\theta)$, $\widehat{w}_i(\theta)$ be polynomials which satisfy

$$w_i(0) = 0,\;\; w_i(1) = b_i,\;\; \widehat{w}_i(0) = 0,\;\;\; \widehat{w}_i(1) = \widehat{b}_i.$$

For the equation (1.1)

$$\phi(t_n + \Delta t\,\theta) = u_n + \Delta t \sum_{i=1}^{s} w_i(\theta)\, LU_{n,i} \;+\; \Delta t \sum_{i=1}^{s} \widehat{w}_i(\theta)\, g\Big(t_{n,i}, U_{n,i}\Big), \qquad (4.1)$$

with $0 \le \theta \le 1$, gives an approximate solution on the interval $[t_n, t_{n+1}]$. If the polynomials $w_i(\theta)$, $\widehat{w}_i(\theta)$ satisfy suitable order conditions (see, e.g., [11], 2.1.7), the continuous extension (4.1) preserves the order of accuracy of the underlying IMEX Runge-Kutta method (1.4). For example, the IMEX method is of order 2 and $w_i(\theta)$, $\widehat{w}_i(\theta)$ satisfy

$$\sum_{i=1}^{s} w_i(\theta) = \theta, \quad \sum_{i=1}^{s} \widehat{w}_i(\theta) = \theta,$$

the approximate solution is $O(\Delta t^2)$ accurate on the whole integral interval.

By replacing the DDE (1.5) with an ODE in the form

$$\frac{du}{dt} = Lu(t) + g\Big(t, u(t), \phi(t-\tau)\Big) \tag{4.2}$$

in each interval $[t_n,\, t_{n+1}]$, we can apply the continuous IMEX method to the DDE (1.5). In particular, under the constraint (2.1), the method is represented as

$$\begin{aligned}
U_{n,i} &= u_n + \Delta t \sum_{j=1}^{i} a_{ij}\, LU_{n,j} \;+\; \Delta t \sum_{j=1}^{i-1} \widehat{a}_{ij}\, g\Big(t_{n,j}, U_{n,j}, \Phi_{n-m,j}\Big), \\
\Phi_{n,i} &= u_n + \Delta t \sum_{j=1}^{s} w_j(c_i)\, LU_{n,j} \;+\; \Delta t \sum_{j=1}^{s} \widehat{w}_j(c_i)\, g\Big(t_{n,j}, U_{n,j}, \Phi_{n-m,j}\Big), \\
u_{n+1} &= u_n + \Delta t \sum_{i=1}^{s} b_i\, LU_{n,i} \;+\; \Delta t \sum_{i=1}^{s} \widehat{b}_i\, g\Big(t_{n,i}, U_{n,i}, \Phi_{n-m,i}\Big),
\end{aligned} \tag{4.3}$$

where $\Phi_{n,i} := \phi(t_{n,i})$. Moreover, by applying (4.3) to the test equation (1.6), we get

$$\begin{aligned}
U_{n,i} &= u_n + \Delta t \sum_{j=1}^{i} a_{ij}\, \lambda U_{n,j} \;+\; \Delta t \sum_{j=1}^{i-1} \widehat{a}_{ij}\, \mu\, \Phi_{n-m,j}, \\
\Phi_{n,i} &= u_n + \Delta t \sum_{j=1}^{s} w_j(c_i)\, \lambda U_{n,j} \;+\; \Delta t \sum_{j=1}^{s} \widehat{w}_j(c_i)\, \mu\, \Phi_{n-m,j}, \\
u_{n+1} &= u_n + \Delta t \sum_{i=1}^{s} b_i\, \lambda U_{n,i} \;+\; \Delta t \sum_{i=1}^{s} \widehat{b}_i\, \mu\, \Phi_{n-m,i},
\end{aligned}$$

which is rewritten in the vector form

$$\begin{aligned}
U_n &= u_n \mathbf{1} + \alpha A U_n + \beta \widehat{A} \Phi_{n-m}, \\
\Phi_n &= u_n \mathbf{1} + \alpha W U_n + \beta \widehat{W} \Phi_{n-m}, \\
u_{n+1} &= u_n + \alpha b^T U_n + \beta \widehat{b}^T \Phi_{n-m},
\end{aligned} \tag{4.4}$$

where

$$\Phi_n = [\Phi_{n,1},\, \dots,\, \Phi_{n,s}]^T, \quad W = [w_j(c_i)]_{i,j=1}^s, \quad \widehat{W} = [\widehat{w}_j(c_i)]_{i,j=1}^s.$$

Using (4.4) we can define the P-stability region of the continuous IMEX method as follows.

Definition 4.1 *The P-stability region of the continuous IMEX Runge-Kutta method (4.3) is the set S_P of the pair of complex numbers (α, β), such that* $\det[I - \alpha A] \neq 0$ *and the zero solution of (4.4) is asymptotically stable for any* $m \geq 1$.

Put

$$\widehat{P}_\alpha(z) = \det \begin{bmatrix} I - \alpha A + \alpha \mathbf{1} b^T & -z\widehat{A} + z\mathbf{1}\widehat{b}^T \\ -\alpha W + \alpha \mathbf{1} b^T & I - z\widehat{W} + z\mathbf{1}\widehat{b}^T \end{bmatrix}, \tag{4.5}$$

$$\widehat{Q}_\alpha(z) = \det \begin{bmatrix} I - \alpha A & -z\widehat{A} \\ -\alpha W & I - z\widehat{W} \end{bmatrix}. \tag{4.6}$$

Define $\widehat{\Gamma}_\alpha$ and $\widehat{\sigma}_\alpha$ by

$$\widehat{\Gamma}_\alpha = \{z \in \mathbb{C} \,:\, |\widehat{P}_\alpha(z)| = |\widehat{Q}_\alpha(z)|\}, \tag{4.7}$$

and $\widehat{\sigma}_\alpha = \inf\{|z| \,:\, z \in \widehat{\Gamma}_\alpha\}$, respectively. We can prove the following theorem by the same argument as in the proof of Theorem 2.2.

Theorem 4.2 *Assume that* $\det[I - \alpha A] \neq 0$ *and consider the following three statements:*

(a) $\alpha \in S_A$ *and* $|\beta| < \widehat{\sigma}_\alpha$;
(b) $(\alpha, \beta) \in S_P$;
(c) $\alpha \in S_A$ *and* $|\beta| \leq \widehat{\sigma}_\alpha$.

Then, we have (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c).

By way of illustration, we consider the IMEX method (3.4) with the linear interpolant defined by

$$\begin{aligned} &w_1(\theta) = 0,\ w_2(\theta) = -\theta,\ w_3(\theta) = w_4(\theta) = \theta, \\ &\widehat{w}_3(\theta) = \theta,\ \widehat{w}_i(\theta) = 0,\ i = 1,\ 2,\ 4. \end{aligned} \tag{4.8}$$

For the method we have

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & -1/2 & 1/2 & 1/2 \\ 0 & -1 & 1 & 1 \end{bmatrix}, \quad \widehat{W} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and

$$\begin{aligned} \widehat{Q}_\alpha(z) &= \frac{1}{2}(1-\alpha)^2(2-2\alpha-z), \\ \widehat{P}_\alpha(z) &= 1 - 2\alpha + \frac{\alpha^2}{2} + \frac{1}{2}(1-3\alpha+\alpha^2)z. \end{aligned}$$

Since $\widehat{Q}_\alpha(z)$, $\widehat{P}_\alpha(z)$ are of degree 1 with respect to z, the set $\widehat{\Gamma}_\alpha$ becomes a circle, and its center and radius are given by $-\zeta/\kappa$ and $\rho/|\kappa|$, respectively, where

$$\begin{aligned} \kappa &= |1-3\alpha+\alpha^2|^2 - |1-\alpha|^4, \\ \zeta &= (2-4\alpha+\alpha^2)(1-3\,\overline{\alpha}+\overline{\alpha}^2) + 2(1-\alpha)^3(1-\overline{\alpha})^2, \\ \rho &= \big|(2-\alpha)^2(1-\alpha)^2(1-2\alpha)\big|. \end{aligned}$$

Moreover, since κ is rewritten as

$$\kappa = -\operatorname{Re}\alpha(2-\operatorname{Re}\alpha)(1-2\operatorname{Re}\alpha) + (\operatorname{Im}\alpha)^2(5-2\operatorname{Re}\alpha),$$

it is positive if $\operatorname{Re}\alpha < 0$. Hence, $\widehat{\sigma}_\alpha$ is given by

$$\widehat{\sigma}_\alpha = \Big|\rho - |\zeta|\Big|/\kappa \tag{4.9}$$

for $\operatorname{Re}\alpha < 0$. When α is a negative real number, ζ is reduced to $\zeta = (2-\alpha)(1-2\alpha)(2-5\alpha+3\alpha^2-\alpha^3)$ and $\widehat{\sigma}_\alpha$ is reduced to $\widehat{\sigma}_\alpha = -\alpha$. We plot the curves $\widehat{\sigma}_\alpha$ computed by (4.9) in Fig. 4, which suggests that $\widehat{\sigma}_\alpha \geq -\operatorname{Re}\alpha$, i.e., the method (3.4) with the linear interpolant (4.8) is P-stable.

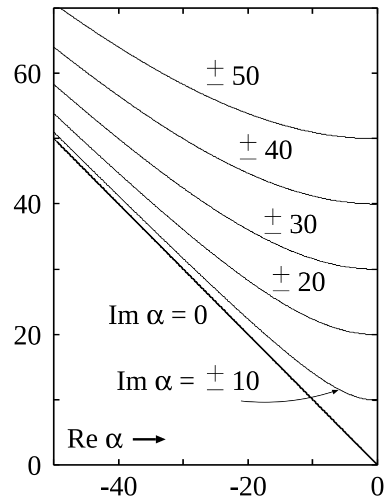

Fig. 4. Values of $\widehat{\sigma}_\alpha$ for the method (3.4) with the linear interpolant.

5 Numerical examples

In this section, we present some numerical examples. We consider the following delayed reaction-diffusion equation (cf. [14], p. 220) on the interval $\Omega = (0, 1)$

$$\frac{\partial v}{\partial t} = D\,\frac{\partial^2 v}{\partial x^2} + \mu v(t-\tau, x)[1 + v(t,x)^2\,], \quad t \geq 0, \quad x \in \Omega, \tag{5.1}$$

under the homogeneous Dirichlet condition

$$v(t,0) = v(t,1) = 0, \quad t \geq 0, \tag{5.2}$$

where D is a positive constant and μ is a real parameter. This model is derived from a sort of delayed logistic equation by taking the effect of spatial diffusion into account.

Let M be an integer, $\Delta x := 1/M$ and define the mesh points $x_0 = 0 < x_1 < \cdots < x_j = j\Delta x < \cdots < x_M = 1$. We use the notation $u^j(t)$ for an approximate function of $v(t, x_j)$. By replacing the second order spatial derivative with the second order centered difference, we obtain an MOL approximation

$$\frac{du}{dt} = Lu(t) + g\Big(u(t), u(t-\tau)\Big), \tag{5.3}$$

where

$$u(t) = \begin{bmatrix} u^1(t) \\ u^2(t) \\ \\ \vdots \\ u^{M-1}(t) \end{bmatrix}, \quad L = \frac{D}{\Delta x^2} \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ 0 & 1 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -2 \end{bmatrix},$$

and the j-th component of $g\big(u(t), u(t-\tau)\big)$ is $\mu u^j(t-\tau)[1 + u^j(t)^2]$.

The eigenvalues of the matrix L are

$$\lambda_k = -\frac{4D}{\Delta x^2} \sin^2\Big(\frac{k\pi\Delta x}{2}\Big), \quad k = 1, 2, \ldots, M-1.$$

By the linearization method (see, e.g., [7], Chap. 10), it is shown that if μ satisfies

$$|\mu| < -\lambda_1 = \frac{4D}{\Delta x^2} \sin^2\Big(\frac{\pi\Delta x}{2}\Big), \tag{5.4}$$

the zero solution of (5.3) is asymptotically stable for any $\tau > 0$. The upper bound $-\lambda_1$ increases as M increases and is contained in the range $9D \leq -\lambda_1 < \pi^2 D \approx 9.8696D$ for $M \geq 3$. Fig. 5 shows a typical behavior of the solution of (5.3) in the case the condition (5.4) is satisfied. The parameter values are

$$D = 1, \ \mu = -8, \ \tau = 1, \ M = 100, \tag{5.5}$$

and the initial condition is given by

$$u^j(t) = x_j(1 - x_j), \ -\tau \leq t \leq 0, \ 1 \leq j \leq M-1. \tag{5.6}$$

The solution oscillates and is damped, as is often seen in the case of the usual delayed logistic equation without diffusion.

P-stable methods may preserve the asymptotic property independent of the stepsize Δt. Fig. 6 shows the numerical results by the Θ-method with $\Theta = 1$ and the method (3.4) in the case where the parameter values are

$$D = 10, \ \mu = -80, \ \tau = 1, \ M = 1000, \tag{5.7}$$

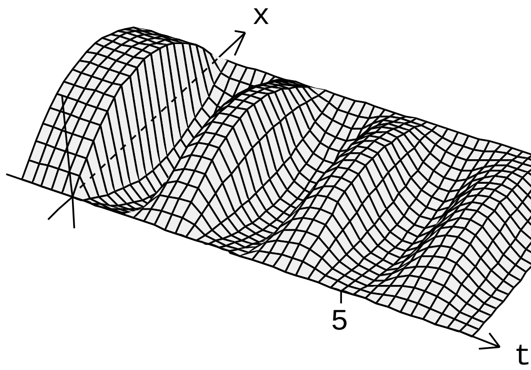

Fig. 5. Solution of (5.3) for $D = 1$, $\mu = -8$, $\tau = 1$, $M = 100$.

and the initial condition is given by (5.6). The approximate values $u_n^{500} \approx v(t_n, 0.5)$ for $m = 1$ are plotted, where m is the positive integer that is used for defining Δt in (2.1). As for these two methods and the method (3.4) with the linear interpolant (4.8), stable numerical solutions are obtained independent of the value of m.

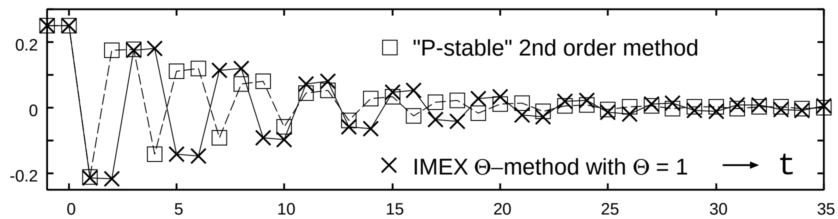

Fig. 6. Numerical results by P-stable methods.

On the other hand, in the case of the IMEX Θ-method with $\Theta < 1$ or the ARS method (3.3), the asymptotic property of the numerical solutions changes at some value of m. For example, the numerical solutions by the IMEX Θ-method with $\Theta = 1/2$ diverge when $m \leq 39$ and tend to zero when $m \geq 40$ for the parameter values (5.7) (Fig. 7). When $m = 40$, $\beta := \mu\Delta t = -80/40 = -2$. This coincides with the bound derived from the P-stability region (Fig. 1). In the case of the method (3.3), such a change occurs between $m = 31$ and $m = 32$.

In the case of the IMEX trapezoidal rule, the divergence of the numerical solutions continues until rather large m. Fig. 8 shows a numerical solution by the method with $m = 275$ for the parameter values (5.5). It is still unstable, whereas stable solutions are obtained even by the IMEX Θ-method with $\Theta = 1/2$ or the ARS method (3.3) with $m = 4$ for these parameter values. The instability of the IMEX trapezoidal rule is explained from the P-stability region of the method, whose width rapidly tends to zero as $|\alpha|$ increases (Fig. 2).

Acknowledgment. I would like to thank Barbara Zubik-Kowal, who gave me encouragement to carry out this study.

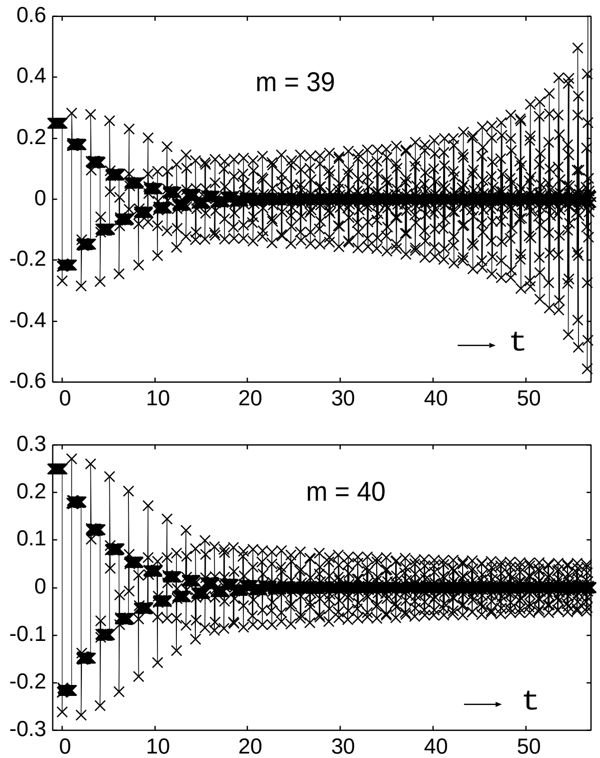

Fig. 7. Approximate values u_n^{500} by the IMEX Θ-method with $\Theta = 1/2$.

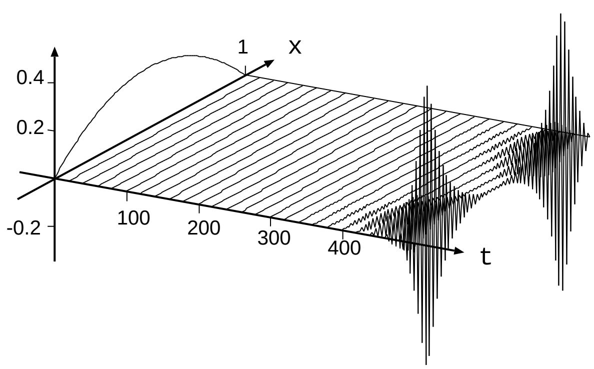

Fig. 8. Numerical solution by the IMEX trapezoidal rule.

References

[1] O. Aberth, Iteration methods for finding all zeros of a polynomial simultaneously, Math. Comp. **27** (1973), 339–344.

[2] U. M. Ascher, S. J. Ruuth, R. J. Spiteri, Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations, Appl. Numer. Math. **25** (1997), 151–167.

[3] V. K. Barwell, Special stability problems for functional differential equations, BIT **15** (1975), 130–135.

[4] A. Bellen, M. Zennaro, Numerical methods for delay differential equations, The Clarendon Press Oxford University Press, New York, 2003.

[5] M. P. Calvo, J. de Frutos, J. Novo, Linearly implicit Runge-Kutta methods for advection-reaction-diffusion equations, Appl. Numer. Math. **37** (2001), 535–549.

[6] E. Hairer, S. P. Nørsett, G. Wanner, Solving ordinary differential equations I, 2nd rev. ed., Springer-Verlag, Berlin, 1993.

[7] J. K. Hale, S. M. Verduyn Lunel, Introduction to functional-differential equations, Springer-Verlag, New York, 1993.

[8] D. Hoff, Stability and convergence of finite difference methods for systems of nonlinear reaction-diffusion equations, SIAM J. Numer. Anal. **15** (1978), 1161–1177.

[9] W. Hundsdorfer, J. Verwer, Numerical solution of time-dependent advection-diffusion-reaction equations, Springer-Verlag, Berlin, 2003.

[10] K. J. in 't Hout, M. N. Spijker, Stability analysis of numerical methods for delay differential equations, Numer. Math. **59** (1991), 807–814.

[11] C. A. Kennedy, M. H. Carpenter, Additive Runge-Kutta schemes for convection-diffusion-reaction equations, Appl. Numer. Math. **44** (2003), 139–181.

[12] T. Koto, NP-stability of Runge-Kutta methods based on classical quadrature, BIT **37** (1997), 870–884.

[13] V. Y. Pan, Solving a polynomial equation: some history and recent progress, SIAM Rev. **39** (1997), 187–220.

[14] J. Wu, Theory and applications of partial functional-differential equations, Springer-Verlag, New York, 1996.

[15] M. Zennaro, P-stability properties of Runge-Kutta methods for delay differential equations, Numer. Math. **49** (1986), 305–318.

[16] X. Zhong, Additive semi-implicit Runge-Kutta methods for computing high-speed nonequilibrium reactive flows, J. Comput. Phys. **128** (1996), 19–31.