



ELSEVIER

Applied Numerical Mathematics 42 (2002) 201–212



APPLIED
NUMERICAL
MATHEMATICS

www.elsevier.com/locate/apnum

On the contractivity of implicit–explicit linear multistep methods [☆]

K.J. in 't Hout

Mathematical Institute, Leiden University, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

Abstract

This paper is concerned with the class of implicit–explicit linear multistep methods for the numerical solution of initial value problems for ordinary differential equations which are composed of stiff and nonstiff parts. We study the contractivity of such methods, with regard to linear autonomous systems of ordinary differential equations and a (scaled) Euclidean norm. In addition, we derive a strong stability result based on the stability regions of these methods. © 2001 IMACS. Published by Elsevier Science B.V. All rights reserved.

Keywords: Initial value problems; Ordinary differential equations; Numerical solution; Implicit–explicit linear multistep methods; Stability analysis; Contractivity

1. Introduction

1.1. Implicit–explicit linear multistep methods

We shall deal with the numerical solution of initial value problems for systems of ordinary differential equations given in decomposed form,

$$U'(t) = F_1(t, U(t)) + F_2(t, U(t)) \quad (t \geq 0), \quad U(0) = u_0. \quad (1.1)$$

Here F_1 and F_2 denote given vector-valued functions, u_0 denotes a given vector, and $U(t)$ (for $t > 0$) is unknown. We assume that the F_1 -part of (1.1) is nonstiff or mildly stiff, whereas the F_2 -part is stiff. By this we mean that in the case of the F_1 -part explicit numerical methods could be applied, while for the F_2 -part implicit methods are desirable.

Initial value problems of the above type frequently arise for example, after semi-discretization of initial-boundary value problems for time-dependent partial differential equations (cf. the references below).

[☆] This work has been supported by the Netherlands Organisation for Scientific Research (NWO) and the Research Council of the K.U. Leuven, Belgium.

E-mail address: hout@math.leidenuniv.nl (K.J. in 't Hout).

For the numerical solution of (1.1), we deal in this paper with the class of *implicit–explicit (IMEX) linear multistep methods*. This class of numerical methods has (recently) been considered by various authors in the literature, see, e.g., [1–3,5,16,17].

Let α_j, β_j (for $j = 0, 1, \dots, k$) and γ_j (for $j = 0, 1, \dots, k-1$) be given, fixed real numbers with $\alpha_k \neq 0, \beta_k \neq 0, |\alpha_0| + |\beta_0| + |\gamma_0| > 0$. The numbers $\alpha_j, \beta_j, \gamma_j$ define an IMEX linear multistep method, which we denote by (α, β, γ) . Let $h > 0$ be a given stepsize, let $t_n = nh$ ($n = 0, 1, 2, \dots$) be the corresponding gridpoints, and suppose initial approximations $u_j \approx U(t_j)$ (for $j = 1, 2, \dots, k-1$) to the exact solution U of (1.1) are known. Then approximations u_{n+k} to $U(t_{n+k})$ for $n = 0, 1, 2, \dots$ are recursively defined by

$$\sum_{j=0}^k \alpha_j u_{n+j} = h \sum_{j=0}^{k-1} \gamma_j F_1(t_{n+j}, u_{n+j}) + h \sum_{j=0}^k \beta_j F_2(t_{n+j}, u_{n+j}). \quad (1.2)$$

It is clearly seen that, in the process (1.2), the implicitness is present only via the function F_2 , and not via F_1 . For several interesting examples of IMEX linear multistep methods we refer to, e.g., [2,5], cf. also Section 2.3 below. We (briefly) mention that a general approach to construct specific IMEX processes (1.2) is to consider first application of a given known implicit linear multistep method to the whole problem (1.1), and then to replace the implicit term which occurs in the case of F_1 by a suitable extrapolation formula, cf., e.g., [5].

1.2. Stability analysis

In this paper we are interested in the stability properties of IMEX linear multistep processes (1.2). In order to investigate the stability of (1.2), we shall deal with the following, linear autonomous *test problem*,

$$U'(t) = LU(t) + MU(t) \quad (t \geq 0), \quad U(0) = u_0. \quad (1.3)$$

Here L, M denote given complex square matrices. The term $LU(t)$ in (1.3) represents the nonstiff (or mildly stiff) part, whereas the term $MU(t)$ represents the stiff part. We subsequently consider the following condition concerning L, M ,

$$L = L_0 \otimes I^{(m)}, \quad M = I^{(l)} \otimes M_0. \quad (1.4)$$

Here L_0 and M_0 are arbitrary complex square matrices, of respective orders l, m (say). Further, \otimes stands for the Kronecker product of two matrices (see, e.g., [7,10]) and $I^{(d)}$ denotes the identity matrix of order d ($d = 1, 2, 3, \dots$). In the following we shall also write I instead of $I^{(d)}$ (whenever the order is clear from the context).

By a basic property of the Kronecker product, it holds that matrices L, M which satisfy the condition (1.4) are commuting. However, we do not require in this paper that L, M are both normal. In other words, L, M are not assumed to be simultaneously unitarily diagonalizable. In view of this, assessing the stability behaviour of processes (1.2) in the case of test problem (1.3) under condition (1.4) is of much interest. We note, furthermore, that problems (1.3) with matrices L, M of the kind (1.4) have practical relevance, as they naturally arise after semi-discretization of initial-boundary value problems for partial differential equations whenever the spatial dimension is greater than one (such as the prototype two-dimensional convection–diffusion equation $u_t + au_x = bu_{yy}$). IMEX linear multistep methods have recently been

considered in the literature in connection with the numerical solution of initial value problems (1.1) obtained in this way, cf. [5,17].

Put $X = hL$, $Y = hM$, and let the vector $U_n = (u_{n+k-1}^T, u_{n+k-2}^T, \dots, u_n^T)^T$. For $j = 0, 1, \dots, k - 1$, define a rational function c_j of two complex variables x, y by

$$c_j(x, y) = \frac{-\alpha_j + \gamma_j x + \beta_j y}{\alpha_k - \beta_k y}.$$

Next, consider the $k \times k$ matrix-valued rational function C of $(x, y) \in \mathbb{C}^2$ given by

$$C(x, y) = \begin{pmatrix} c_{k-1}(x, y) & \dots & c_1(x, y) & c_0(x, y) \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{pmatrix}. \tag{1.5}$$

Then application of the IMEX linear multistep process (1.2) in the case of test problem (1.3) under the condition (1.4) gives rise to the following linear recurrence relation for the vectors U_n (whenever $(\alpha_k I - \beta_k Y)$ is regular):

$$U_{n+1} = C(X, Y)U_n \quad \text{for } n = 0, 1, 2, \dots \tag{1.6}$$

Note that the definition of $C(X, Y)$ is unambiguous, since the matrices X, Y commute. In this paper, we are concerned with *contractivity* of the process (1.6), i.e.,

$$\|C(X, Y)\| \leq 1,$$

where $\|\cdot\|$ denotes a matrix norm subordinate to some given vector norm on \mathbb{C}^{klm} . In addition, we shall also deal with *strong stability*,

$$\|C(X, Y)^n\| \leq K \quad \text{for } n = 1, 2, 3, \dots$$

Here K denotes a positive real constant that is uniformly valid in an interesting (infinite) class of matrices X, Y .

1.3. Review of results from the literature

Varah [16] and Ascher et al. [2] derived stability results for IMEX processes (1.2) which are relevant to specific, scalar cases of the test problem (1.3).

Frank et al. [5] considered the stability behaviour of (1.2) in the general scalar case of test problem (1.3). In [5], results were obtained on the stability regions of several (interesting) IMEX linear multistep methods, cf. also Section 2.4 below. We call the *stability region* of an IMEX method (α, β, γ) the set S consisting of all pairs $(x, y) \in \mathbb{C}^2$ such that the companion matrix $C(x, y)$, given by (1.5), satisfies the following familiar *eigenvalue condition*: if ζ is an eigenvalue of $C(x, y)$, then $|\zeta| \leq 1$, and $|\zeta| < 1$ whenever ζ has an algebraic multiplicity greater than one. For a given matrix $C(x, y)$, this eigenvalue condition is equivalent to the requirement of power boundedness.

Crouzeix [3] (see also Akrivis et al. [1]) established stability and convergence results for (1.2) for initial value problems more general than (1.3), where the nonstiff part is allowed to be nonautonomous (and in [1], nonlinear). Specializing to the situation at hand, the results from [3] concern systems

(1.3) with real matrices L , M that are not assumed to commute, but where M is symmetric and negative definite. Under the conditions on L , M pertinent to [3], strong stability for processes (1.2) in the case of (1.3) was shown, with respect to a scaled Euclidean vector norm depending on h and M .

Finally we note that in the special case of the *IMEX θ -method*, defined by $k = 1$, $\alpha_0 = -1$, $\alpha_1 = 1$, $\beta_0 = 1 - \theta$, $\beta_1 = \theta$, $\gamma_0 = 1$ with real parameter θ , various further stability results can be obtained from the literature, cf. [8,9,13,14].

1.4. Purpose of this paper

As already mentioned in Section 1.2, we deal in this paper with contractivity and strong stability of the process (1.6), given by application of the general IMEX linear multistep process (1.2) in the case of test problem (1.3) under condition (1.4). In this paper, we consider matrix norms that correspond with scaled Euclidean vector norms. Here, our scaling is fixed, i.e., it is completely determined by a given fixed $k \times k$ matrix. An outline of our paper is as follows.

In Section 2.1 we present a general theorem concerning contractivity. This theorem can be regarded as a matrix-valued version of a theorem of von Neumann for two variables.

In Section 2.2 we give the proof of the above theorem.

In Section 2.3 we then apply the general theorem from Section 2.1 in the case of process (1.6), leading to a useful contractivity result. Subsequently, we introduce and analyze contractivity regions for IMEX linear multistep methods.

Finally, in Section 2.4 we study the situation where, in the conditions pertinent to the contractivity result obtained in Section 2.3, the contractivity region is replaced by the (larger) stability region. We show that under this replacement, and an additional relaxation involving the classical numerical range, still strong stability for (1.6) can be concluded.

2. Stability results

2.1. A general contractivity theorem

Let $R = (r_{ij})_{i,j=1}^k$ be any given $k \times k$ matrix-valued function with entries r_{ij} that are complex rational functions on \mathbb{C}^2 . Next, let matrices X , Y be given with

$$X = X_0 \otimes I^{(m)}, \quad Y = I^{(l)} \otimes Y_0, \quad (2.1)$$

where X_0 , Y_0 are arbitrary complex square matrices, of respective orders l , m , say. In this section we consider the matrix $R(X, Y)$ of order klm given by

$$R(X, Y) = (r_{ij}(X, Y))_{i,j=1}^k.$$

We shall formulate a general theorem concerning the contractivity of this matrix.

On any complex vector space \mathbb{C}^d , with positive integer d , let $\langle \cdot, \cdot \rangle$ and $|\cdot|$ denote the Euclidean inner product and Euclidean vector norm, respectively. We denote the matrix norm subordinate to $|\cdot|$ by $\|\cdot\|$.

For any complex $d \times d$ matrix A and any given number $\rho \in [0, \infty)$, we say that A satisfies the *condition* $\mathcal{D}(\rho)$ if

$$\|A + \rho I\| \leq \rho.$$

Furthermore, we say that A satisfies the *condition* $\mathcal{D}(\infty)$ if

$$\operatorname{Re}\langle Av, v \rangle \leq 0 \quad \text{for all } v \in \mathbb{C}^d.$$

Corresponding to the condition $\mathcal{D}(\rho)$, we define a set $D[\rho]$ in \mathbb{C} by

$$D[\rho] = \{z: z \in \mathbb{C}, |z + \rho| \leq \rho\} \quad (\text{whenever } 0 \leq \rho < \infty),$$

and

$$D[\infty] = \{z: z \in \mathbb{C}, \operatorname{Re}(z) \leq 0\}.$$

Next, let G denote any given $k \times k$ matrix which is Hermitian and positive definite. Then on any space \mathbb{C}^{kd} with positive integer d we define an inner product $\langle \cdot, \cdot \rangle_G$ through

$$\langle v, w \rangle_G = \langle \tilde{G}v, w \rangle \quad \text{for all } v, w \in \mathbb{C}^{kd},$$

where $\tilde{G} = G \otimes I^{(d)}$. The vector norm induced by this inner product is denoted by $|\cdot|_G$, and the matrix norm subordinate to $|\cdot|_G$ by $\|\cdot\|_G$.

The following theorem constitutes the main result of this section. It provides a sufficient condition for contractivity of the matrix $R(X, Y)$.

Theorem 2.1. *Let $\rho_1, \rho_2 \in [0, \infty]$ be given. Assume that $\|R(x, y)\|_G \leq 1$ for all $(x, y) \in D[\rho_1] \times D[\rho_2]$. Then $\|R(X, Y)\|_G \leq 1$ whenever (2.1) holds, X satisfies $\mathcal{D}(\rho_1)$ and Y satisfies $\mathcal{D}(\rho_2)$.*

Theorem 2.1 can be viewed as a matrix-valued version of a (well-known) theorem due to von Neumann [11] for two variables. The above theorem is related in a natural way with two generalizations to von Neumann's theorem already obtained in the literature. The first generalization can be found in Sz.-Nagy [15, p. 29], and concerns a scalar ($k = 1$) version of von Neumann's theorem for two variables, cf. also Jorge and Lisbona [9]. The second generalization was derived by Nevanlinna [12], and deals with a matrix-valued version of the theorem of von Neumann for one variable, see also Hairer and Wanner [6, Section V.7]. We mention that the results from [12, 15] are both given in a general Hilbert space setting. Clearly, none of the results from loc. cit. includes the above theorem.

We will prove Theorem 2.1 in the subsequent Section 2.2. We conclude this section with some remarks. Let $\sigma[A]$ denote the spectrum of any matrix A .

Remark 2.2. If X or Y is normal, then the assertion of Theorem 2.1 can be sharpened, analogously to the case of one variable. For example, we have: *Let $\rho \in [0, \infty]$. Assume that X is normal and $\|R(x, y)\|_G \leq 1$ for all $(x, y) \in \sigma[X] \times D[\rho]$. Then $\|R(X, Y)\|_G \leq 1$ whenever (2.1) holds and Y satisfies $\mathcal{D}(\rho)$.*

Remark 2.3. Theorem 2.1 generalizes in a straightforward way to an arbitrary, finite number of variables. If, for instance, R is a matrix-valued rational function of three complex variables (x, y, z) , then one considers matrices X, Y, Z with $X = X_0 \otimes I^{(m)} \otimes I^{(n)}$, $Y = I^{(l)} \otimes Y_0 \otimes I^{(n)}$, $Z = I^{(l)} \otimes I^{(m)} \otimes Z_0$, where X_0, Y_0, Z_0 denote complex square matrices of respective orders l, m, n . The proof of this

generalization follows from a direct modification of the proof given in Section 2.2. We leave the details of this generalization, and its proof, to the reader.

Remark 2.4. The natural question arises as to whether Theorem 2.1 and its generalization of Remark 2.3 remain valid if the condition (2.1) and the corresponding condition in Remark 2.3 are replaced by the weaker requirement that the (relevant) matrices X, Y, Z, \dots are commuting. Some results concerning this question can be obtained from [15], for $k = 1$. If two variables are considered, and $k = 1$, then it easily follows from [15, p. 29] that our result, i.e., the scalar version of Theorem 2.1, is indeed still valid under the above replacement, cf. also [9]. On the other hand, if the (finite) number of variables is three or more, then it is not clear whether our corresponding result (see Remark 2.3) remains valid under the mentioned replacement. In fact, in this case, counterexamples are known in the literature at least when X, Y, Z, \dots are allowed to be operators in a general (infinite-dimensional) Hilbert space, cf. [9,15].

2.2. Proof of Theorem 2.1

Before formulating the actual proof of Theorem 2.1 we derive two lemmas. For any i, j write $r_{ij} = p_{ij}/q_{ij}$ with polynomials p_{ij}, q_{ij} that have no common (nonconstant) divisor. First we have

Lemma 2.5. $R(X, Y)$ is well-defined if and only if $q_{ij}(x, y) \neq 0$ for all $(x, y) \in \sigma[X_0] \times \sigma[Y_0]$ and $1 \leq i, j \leq k$.

Proof. $R(X, Y)$ is well-defined if and only if $q_{ij}(X, Y)$ is nonsingular for all $1 \leq i, j \leq k$. It is easily verified that each $q_{ij}(X, Y)$ is a finite linear combination (over \mathbb{C}) of terms $(X_0)^s \otimes (Y_0)^t$ with integers $s, t \geq 0$. Therefore, according to a well-known property of the Kronecker product (see, e.g., [7,10]), the eigenvalues of $q_{ij}(X, Y)$ are given by $q_{ij}(x, y)$ with $x \in \sigma[X_0], y \in \sigma[Y_0]$. From this observation, the lemma immediately follows. \square

Next, consider for $y \in \mathbb{C}$ the matrix $R_{X_0}(y)$ of order kl given by

$$R_{X_0}(y) = (r_{ij}(X_0, y))_{i,j=1}^k.$$

It is easily seen that each entry of $R_{X_0}(y)$ is a rational expression in y . Thus R_{X_0} is a matrix-valued rational function of one complex variable, y . The subsequent lemma forms the key to our proof of Theorem 2.1.

Lemma 2.6. Assume $R(X, Y)$ is well-defined. Then $R_{X_0}(Y_0)$ is well-defined, and $R_{X_0}(Y_0) = R(X, Y)$.

Proof. Suppose first $k = 1$. For convenience, write $r_{11} = r, p_{11} = p, q_{11} = q$, so that $R = r$ and $R_{X_0}(y) = r(X_0, y) = p(X_0, y)q(X_0, y)^{-1}$. The entries of $R_{X_0}(y)$ are all rational in y and they have a common denominator $d(y)$ given by $d(y) = \det[q(X_0, y)]$, where $\det[\cdot]$ denotes the determinant function. Hence, $R_{X_0}(Y_0)$ is well-defined whenever $d(Y_0)$ is nonsingular. It is readily seen that $d(Y_0)$ is nonsingular if and only if $q(x, y) \neq 0$ for all $(x, y) \in \sigma[X_0] \times \sigma[Y_0]$. By Lemma 2.5, the latter condition is fulfilled, and thus $R_{X_0}(Y_0)$ is well-defined.

Put $P_{X_0}(y) = p(X_0, y), Q_{X_0}(y) = q(X_0, y)$. Then it obviously holds that

$$R_{X_0}(y) = P_{X_0}(y)Q_{X_0}(y)^{-1} \quad \text{for all } y \in \mathbb{C}, d(y) \neq 0. \quad (*)$$

Next, $r(X, Y) = p(X, Y)q(X, Y)^{-1}$. Write the polynomial p as

$$p(x, y) = \sum_{s,t \geq 0} c_{st} x^s y^t \quad \text{whenever } (x, y) \in \mathbb{C}^2,$$

with certain complex coefficients c_{st} . Then

$$p(X, Y) = \sum_{s,t \geq 0} c_{st} X^s Y^t = \sum_{s,t \geq 0} c_{st} (X_0)^s \otimes (Y_0)^t = P_{X_0}(Y_0).$$

Similarly, $q(X, Y) = Q_{X_0}(Y_0)$, and hence, we obtain

$$r(X, Y) = P_{X_0}(Y_0) Q_{X_0}(Y_0)^{-1}. \tag{**}$$

From (*), (**) it follows that $r(X, Y) = R_{X_0}(Y_0)$, which concludes the proof for the case $k = 1$.

Suppose subsequently that $k > 1$. Define $(R_{ij})_{X_0}(y) = r_{ij}(X_0, y)$ for all i, j , so that

$$R_{X_0}(y) = ((R_{ij})_{X_0}(y))_{i,j=1}^k.$$

Analogously as above, we obtain that $(R_{ij})_{X_0}(Y_0) = r_{ij}(X, Y)$ (whenever $1 \leq i, j \leq k$). As a consequence,

$$R_{X_0}(Y_0) = ((R_{ij})_{X_0}(Y_0))_{i,j=1}^k = (r_{ij}(X, Y))_{i,j=1}^k = R(X, Y). \quad \square$$

Actual proof of Theorem 2.1

The assumptions of Theorem 2.1 imply that $q_{ij}(x, y) \neq 0$ for all $(x, y) \in \sigma[X_0] \times \sigma[Y_0]$ and $1 \leq i, j \leq k$. Thus, in view of Lemma 2.5, $R(X, Y)$ is well-defined.

The contractivity of $R(X, Y)$ follows from a combination of Lemma 2.6 and the matrix-valued version of the theorem of von Neumann for one variable:

Let $y \in D[\rho_2]$. Consider the matrix-valued rational function R_y of one complex variable, x , defined by $R_y(x) = R(x, y)$. We have $\|R_y(x)\|_G \leq 1$ for all $x \in D[\rho_1]$. Further, it is easily seen that the matrix X_0 satisfies the condition $\mathcal{D}(\rho_1)$. By virtue of [12, Corollary 3] it therefore holds that $\|R_y(X_0)\|_G \leq 1$. Clearly $R_y(X_0) = R_{X_0}(y)$, and thus we obtain

$$\|R_{X_0}(y)\|_G \leq 1 \quad \text{whenever } y \in D[\rho_2].$$

Next, R_{X_0} is a matrix-valued rational function of one complex variable, and the matrix Y_0 fulfills the condition $\mathcal{D}(\rho_2)$. Hence, we can apply [12, Corollary 3] a second time. This yields $\|R_{X_0}(Y_0)\|_G \leq 1$. Upon invoking Lemma 2.6, the proof of Theorem 2.1 is completed.

2.3. Contractivity regions of IMEX linear multistep methods

We have (cf. (1.5)):

Definition 2.7. The *contractivity region* of an IMEX linear multistep method (α, β, γ) with respect to the norm $\|\cdot\|_G$ is the set $S_{\|\cdot\|_G}$ consisting of all pairs $(x, y) \in \mathbb{C}^2$ such that $\|C(x, y)\|_G \leq 1$.

The above definition forms a natural extension of the corresponding definition in the case of linear multistep methods (cf., e.g., [6, Section V.7]). Clearly, for any norm $\|\cdot\|_G$ there holds $S_{\|\cdot\|_G} \subset S$, where S is the stability region of the method (cf. Section 1.3).

From Theorem 2.1 we immediately arrive at the following useful contractivity result for process (1.6).

Theorem 2.8. *Let $\rho \in [0, \infty)$ be given such that $D[\rho] \times D[\infty] \subset S_{\|\cdot\|_G}$. Then $\|C(X, Y)\|_G \leq 1$ whenever (2.1) holds, X satisfies $\mathcal{D}(\rho)$ and Y satisfies $\mathcal{D}(\infty)$.*

In this section we shall consider contractivity regions for the following two IMEX linear multistep methods,

$$u_{n+2} = \frac{4}{3}u_{n+1} - \frac{1}{3}u_n + \frac{4}{3}hF_{1,n+1} - \frac{2}{3}hF_{1,n} + \frac{2}{3}hF_{2,n+2} \tag{2.2}$$

and

$$u_{n+2} = u_{n+1} + \frac{3}{2}hF_{1,n+1} - \frac{1}{2}hF_{1,n} + \beta_2hF_{2,n+2} + \beta_1hF_{2,n+1} + \beta_0hF_{2,n} \tag{2.3}$$

with $\beta_2 = (1 + \delta)/2$, $\beta_1 = (1 - 2\delta)/2$, $\beta_0 = \delta/2$, where $\delta \in (0, 1]$ denotes a parameter. Further we have (shortly) written $F_{j,n} = F_j(t_n, u_n)$ (for all j, n).

Method (2.2) was considered, e.g., in [1–3,5,16,17]. Method (2.3) was studied in [2,5]. Both methods (2.2), (2.3) are of second order, and possess underlying implicit linear multistep methods that are A -stable (whenever $0 < \delta \leq 1$). Observe that the implicit linear multistep method which underlies (2.2) is the familiar two-step backward differentiation formula.

In order to obtain suitable (real) symmetric, positive definite matrices G for the above IMEX methods, we have considered application of a procedure contained in [6, Section V.6] in the case of their underlying implicit methods, which led us to

$$G_1 = \frac{1}{4} \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}, \quad G_2 = \frac{1}{2} \begin{pmatrix} 2 + \delta & -\delta \\ -\delta & \delta \end{pmatrix}.$$

The following proposition gives a characterization for contractivity regions of the IMEX methods (2.2), (2.3).

Proposition 2.9. *Let $\delta \in (0, 1]$ and $\kappa = 1 + (\delta + \delta^{-1})/2$. Then*

- (a) *Method (2.2) has $S_{\|\cdot\|_{G_1}} = \{(x, y): x \in \mathbb{C}, y \in \mathbb{C}, |1 + 2x|^2 + 4 \operatorname{Re}(y) \leq 1\}$.*
- (b) *Method (2.3) has $S_{\|\cdot\|_{G_2}} = \{(x, y): x \in \mathbb{C}, y \in \mathbb{C}, |1 + \kappa x|^2 + 2\kappa \operatorname{Re}(y) \leq 1\}$.*

Proof. We will show here only part (a); part (b) goes in a similar manner. For method (2.2) we have

$$C(x, y) = \begin{pmatrix} d(y)(4 + 4x) & -d(y)(1 + 2x) \\ 1 & 0 \end{pmatrix},$$

where $d(y) = (3 - 2y)^{-1}$. Next, $G_1 = T^*T$ with matrix T given by

$$T = \frac{1}{2} \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}.$$

Now define $\tilde{C}(x, y) = TC(x, y)T^{-1}$. Then $\|C(x, y)\|_{G_1} = \|\tilde{C}(x, y)\|$ (recall that $\|\cdot\|$ denotes the spectral norm), and a simple computation yields

$$\tilde{C}(x, y) = d(y) \begin{pmatrix} 2 + 4x & 1 + 2y \\ 1 + 2x & 2 \end{pmatrix}.$$

Subsequently, it can be verified that for any complex 2×2 matrix A ,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

there holds: $\|A\| \leq 1$ if and only if

$$|\det[A]| \leq 1 \quad \text{and} \quad |a_{11}|^2 + |a_{12}|^2 + |a_{21}|^2 + |a_{22}|^2 \leq 1 + |\det[A]|^2.$$

Applying this criterion in the case of $\tilde{C}(x, y)$ yields that $\|\tilde{C}(x, y)\| \leq 1$ if and only if

$$|d(y)(1 + 2x)| \leq 1 \quad \text{and} \quad |d(y)|^2 \{4|1 + 2x|^2 + |1 + 2y|^2 + 4\} \leq 1.$$

By noticing that in the latter condition the second inequality implies the first, it easily follows (by computation) that

$$\|\tilde{C}(x, y)\| \leq 1 \iff |1 + 2x|^2 \leq 1 - 4\operatorname{Re}(y). \quad \square$$

As an interesting corollary to Proposition 2.9 we obtain that for each of the IMEX methods (2.2), (2.3) the contractivity region contains as a (strict) subset the Cartesian product of the corresponding contractivity regions of the underlying explicit and implicit linear multistep methods, respectively (obtained upon setting $y = 0$ and $x = 0$, respectively). We note that this positive result does not hold in general. For instance, it can be readily verified that the (one-step) IMEX θ -method (see Section 1.3) forms a counterexample whenever $\theta \neq 1$, cf. also [8]. We note furthermore that the above conclusion for methods (2.2), (2.3) is different from the corresponding conclusion that was obtained in [5] with regard to the stability regions of these methods.

In view of Theorem 2.8, the radius ρ_c is of interest defined by

$$\rho_c = \sup\{\rho: \rho \geq 0, D[\rho] \times D[\infty] \subset S_{\|\cdot\|_G}\}.$$

Note that ρ_c depends on the norm $\|\cdot\|_G$, which we have suppressed however in the notation. From Proposition 2.9 we immediately see for method (2.2) that $\rho_c = 1/2$ and for method (2.3) that $\rho_c = 1/\kappa$. Obviously, for both methods the supremum is a maximum. In the case of method (2.3), the radius ρ_c is strictly increasing as a function of the parameter $\delta \in (0, 1]$. For the interesting values $\delta = 1/8$ and $\delta = 1/2$ (cf., e.g., [5]) the respective values for ρ_c are $16/81$ (≈ 0.20) and $4/9$ (≈ 0.44). If $\delta = 1$ we obtain $\rho_c = 1/2$, i.e., the same value as for method (2.2). In fact, if $\delta = 1$, then the contractivity region $S_{\|\cdot\|_{G_2}}$ of method (2.3) is identical to the contractivity region $S_{\|\cdot\|_{G_1}}$ of method (2.2).

We mention finally that for method (2.3) with $\delta = 0$, there holds $\rho_c = 0$, for any norm $\|\cdot\|_G$. This follows directly from a result on this method in [5].

2.4. Strong stability

The question arises which stability conclusions can be derived for process (1.6) when in the conditions pertinent to Theorem 2.8 the contractivity region $S_{\|\cdot\|_G}$ is replaced by the stability region S of the IMEX method. In this section we shall present a result relevant to this question. Here we simultaneously consider a weaker condition than the condition $\mathcal{D}(\rho)$ (for $\rho < \infty$). Throughout this section, we deal with the spectral matrix norm $\|\cdot\|$.

We call the IMEX linear multistep method (α, β, γ) *stable at infinity* if the matrix

$$\begin{pmatrix} \beta'_{k-1} & \cdots & \beta'_1 & \beta'_0 \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{pmatrix},$$

with $\beta'_j = -\beta_j/\beta_k$ (for $0 \leq j \leq k - 1$), satisfies the eigenvalue condition from Section 1.3. Observe that the above matrix is obtained from $C(x, y)$ by letting $y \rightarrow \infty$ for x bounded. The IMEX methods (2.2), (2.3) which were considered in Section 2.3 are both stable at infinity (whenever $0 < \delta \leq 1$).

The next lemma concerns the uniform power boundedness of $C(x, y)$.

Lemma 2.10. *Let $\rho \in [0, \infty)$. Assume that $D[\rho] \times D[\infty] \subset S$ and that the IMEX linear multistep method is stable at infinity. Then there exists a real constant $K > 0$ such that $\|C(x, y)^n\| \leq K$ for all $(x, y) \in D[\rho] \times D[\infty]$ and $n = 1, 2, 3, \dots$*

Lemma 2.10 forms a variant to a well-known result for linear multistep methods, cf. [6, Section V.7] and the references given therein. As the above lemma can be obtained in a completely analogous way to loc. cit., we shall omit its proof here.

Denote by $\tau[A]$ the classical numerical range (field of values) of any complex $d \times d$ matrix A ,

$$\tau[A] = \{z: z = \langle Av, v \rangle \text{ with } v \in \mathbb{C}^d, \langle v, v \rangle = 1\}.$$

Then for any given $\rho \in [0, \infty)$ we say that A satisfies the *condition $\tilde{D}(\rho)$* if

$$\tau[A] \subset D[\rho].$$

It is easily seen that the condition $\mathcal{D}(\rho)$ implies $\tilde{D}(\rho)$, but that the converse implication is not true (whenever $0 \leq \rho < \infty$). Thus, $\tilde{D}(\rho)$ is a weaker condition on a matrix A than $\mathcal{D}(\rho)$.

Let R be any given $k \times k$ matrix-valued rational function of one complex variable. The following lemma is a direct consequence of a result by van Dorsseleer and Hundsdorfer [4].

Lemma 2.11. *Let $\rho \in [0, \infty)$ and assume that $\|R(z)\| \leq 1$ for all $z \in D[\rho]$. Then it holds that $\|R(A)\| \leq 3k$ for all complex square matrices A satisfying $\tilde{D}(\rho)$.*

Proof. Write $R = (r_{ij})_{i,j=1}^k$. Each r_{ij} is a complex rational function with $|r_{ij}(z)| \leq 1$ whenever $z \in D[\rho]$. By virtue of [4, Theorem 2.4] we then have that $\|r_{ij}(A)\| \leq 3$ for all i, j . The bound of the lemma is obtained from $\|R(A)\| \leq k \cdot \max_{i,j} \|r_{ij}(A)\|$. \square

We now arrive at the main result of this section.

Theorem 2.12. *Let $\rho \in [0, \infty)$. Assume that $D[\rho] \times D[\infty] \subset S$ and that the IMEX linear multistep method is stable at infinity. Then there exists a real constant $K > 0$ such that: $\|C(X, Y)^n\| \leq K$ whenever (2.1) holds, X satisfies $\tilde{D}(\rho)$, Y satisfies $\mathcal{D}(\infty)$ and $n = 1, 2, 3, \dots$*

Proof. By Lemma 2.10, there is a real constant $\hat{K} > 0$ such that $\|C(x, y)^n\| \leq \hat{K}$ for all $(x, y) \in D[\rho] \times D[\infty]$ and $n = 1, 2, 3, \dots$. Let integer $n \geq 1$ be fixed but arbitrary. Define the matrix-valued

rational function R of $(x, y) \in \mathbb{C}^2$ by $R(x, y) = C(x, y)^n$. Similarly to the actual proof of Theorem 2.1 (Section 2.2), let $y \in D[\infty]$, and consider the matrix-valued function R_y of one complex variable x given by $R_y(x) = R(x, y)$. Then $\|R_y(x)\| \leq \widehat{K}$ whenever $x \in D[\rho]$. Further, X_0 satisfies the condition $\mathcal{D}(\rho)$, since $\tau[X] = \tau[X_0]$. Hence, by Lemma 2.11, we have that $\|R_y(X_0)\| \leq K$, with constant $K = 3k\widehat{K}$. Subsequently, we immediately obtain for R_{X_0} defined in Section 2.2,

$$\|R_{X_0}(y)\| \leq K \quad \text{whenever } y \in D[\infty].$$

Identically to Section 2.2, it now follows (by successive application of [12, Corollary 3] and Lemma 2.6) that $\|R(X, Y)\| = \|C(X, Y)^n\| \leq K$. \square

We remark that the above theorem is related to [4, Theorem 2.6], which deals with the case of linear multistep methods.

In view of Theorem 2.12 the radius ρ_s is of interest defined by

$$\rho_s = \sup\{\rho: \rho \geq 0, D[\rho] \times D[\infty] \subset S\}.$$

Clearly, it always holds that $\rho_s \geq \rho_c$ with ρ_c defined in Section 2.3. We consider here the value of ρ_s for the IMEX methods (2.2), (2.3). For each of these methods a useful characterization was derived in Frank et al. [5] for the set

$$\{x: x \in \mathbb{C} \text{ such that } |\zeta| \leq 1 \text{ whenever } \zeta \in \sigma[C(x, y)] \text{ and } y \in D[\infty]\}.$$

Using this characterization, it can be shown for method (2.2) that $\rho_s = 2/3$ and for method (2.3) that $\rho_s \leq 1/2$. Together with a result for ρ_c obtained in Section 2.3, we directly obtain for method (2.3) with $\delta = 1$ that $\rho_s = 1/2$. Next, combining the characterization from [5] with numerical computations, we further find that $\rho_s = 0.50$ (for $0.12 \leq \delta \leq 1$) and $\rho_s < 0.50$ (for $0 < \delta \leq 0.11$). A comparison with the values for ρ_c obtained in Section 2.3 shows that ρ_s is strictly larger than ρ_c for method (2.2) and (very likely) for method (2.3) whenever $0 < \delta < 1$.

Acknowledgement

The author wishes to thank W. Hundsdorfer, C. Palencia and M.N. Spijker for some useful discussions and interesting references related to the topic of this paper.

References

- [1] G. Akrivis, M. Crouzeix, C. Makridakis, Implicit–explicit multistep methods for quasilinear parabolic equations, *Numer. Math.* 82 (1999) 521–541.
- [2] U.M. Ascher, S.J. Ruuth, B.T.R. Wetton, Implicit–explicit methods for time-dependent partial differential equations, *SIAM J. Numer. Anal.* 32 (1995) 797–823.
- [3] M. Crouzeix, Une méthode multipas implicite-explicite pour l'approximation des équations d'évolution paraboliques, *Numer. Math.* 35 (1980) 257–276.
- [4] J.L.M. van Dorsselaer, W. Hundsdorfer, Stability estimates based on numerical ranges with an application to a spectral method, *BIT* 34 (1994) 228–238.
- [5] J. Frank, W. Hundsdorfer, J.G. Verwer, On the stability of implicit–explicit linear multistep methods, *Appl. Numer. Math.* 25 (1997) 193–205.

- [6] E. Hairer, G. Wanner, *Solving Ordinary Differential Equations II*, 2nd edn., Springer, Berlin, 1996.
- [7] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1991.
- [8] W. Hundsdorfer, *Numerical Solution of Advection–Diffusion–Reaction Equations*. Lecture Notes for Ph.D. Course, CWI, Amsterdam (1998).
- [9] J.C. Jorge, F. Lisbona, Contractivity results for alternating direction schemes in Hilbert spaces, *Appl. Numer. Math.* 15 (1994) 65–75.
- [10] P. Lancaster, M. Tismenetsky, *The Theory of Matrices*, 2nd edn., Academic Press, Orlando, FL, 1985.
- [11] J. von Neumann, Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes, *Math. Nachr.* 4 (1951) 258–281.
- [12] O. Nevanlinna, Matrix valued versions of a result of von Neumann with an application to time discretization, *J. Comput. Appl. Math.* 12–13 (1985) 475–489.
- [13] A. Ostermann, M. Thalhammer, Non-smooth data error estimates for linearly implicit Runge–Kutta methods, *IMA J. Numer. Anal.* 20 (2000) 167–184.
- [14] M.N. Spijker, A note on contractivity in the numerical solution of initial value problems, *BIT* 27 (1987) 424–437.
- [15] B. Sz.-Nagy, *Unitary Dilations of Hilbert Space Operators and Related Topics*, CBMS Regional Conf. Ser. in Math., Vol. 19, American Mathematical Society, Providence, RI, 1974.
- [16] J.M. Varah, Stability restrictions on second order, three level finite difference schemes for parabolic equations, *SIAM J. Numer. Anal.* 17 (1980) 300–309.
- [17] J.G. Verwer, J.G. Blom, W. Hundsdorfer, An implicit–explicit approach for atmospheric transport–chemistry problems, *Appl. Numer. Math.* 20 (1996) 191–209.