Monografías del Seminario Matemático García de Galdeano **33**, 129–136 (2006)

POSITIVITY-PRESERVING AND ENTROPY-DECAYING IMEX METHODS

Inmaculada Higueras and Teo Roldán

Abstract. Ordinary differential equations containing additive terms with different stiffness properties may arise when some time dependent partial differential equations are discretized in space. IMEX Runge-Kutta methods are suitable to treat this kind of problems. Sometimes the solutions to these problems have qualitative properties (norm, energy, entropy, total variation, positivity, etc) that represent important physical features of the problem. In this case, in order to preserve the physical meaning of the numerical solution, it is important to maintain these properties with both the spatial discretization and the time stepping method. IMEX Runge-Kutta methods methods can preserve some qualitative properties of the exact solution under certain stepsize restrictions. In this paper we review some results concerned to these preserving properties and show how they can be used for some problems.

Keywords: IMEX Runge-Kutta, monotonicity, entropy diminishing, positivity.

AMS classification: 65L06, 65L05, 65M20.

§1. Introduction

In this paper we consider the numerical integration of additive Initial Value Problems (IVP) of the form

$$\begin{cases} \dfrac{d}{dt}u(t) = f(u(t)) + \tilde{f}(u(t)), & t \geq t_0, \\ u(t_0) = u_0, \end{cases} \tag{1}$$

where f and $\tilde{f}$ are Lipschitz continuous functions from $\mathbb{R}^m$ to $\mathbb{R}^m$ with different stiffness properties. We assume that the IPV has unique solution $u : [t_0, \infty] \to \mathbb{R}^m$ for each $(t_0, u_0) \in \mathbb{R}^{1+m}$. These kind of problems may arise from semi-discretizations of some evolutionary partial differential equations (PDEs) by method of lines (MOL). In such case the functions f and $\tilde{f}$ often correspond to the spatial discretization of different type terms of the given PDE (e. g. advection and diffusion [1, 2, 11, 15]).

We are interested in problems where there exists a norm or a seminorm or an entropy function such that the solution satisfies the following monotonicity property

$$\|u(t)\| \leq \|u(t_0)\|, \quad \forall t \geq t_0. \tag{2}$$

This question has already been considered by several authors (cf. [11, 12, 16, 17, 18, 19]).

We are also interested in systems whose solutions are non-negative, i. e.,

$$u_0 \geq 0 \;\Rightarrow\; u(t) \geq 0, \quad \forall t \geq t_0, \tag{3}$$

where the vector inequalities must be understood component-wise. For example, problems whose solutions are concentrations or densities of chemical species, must satisty this property. These systems are usually called positive systems ([3, 5, 10]).

Norm monotonicity (2) and positiviy (3) are related to other monotonicity properties, such as maximum principles or total variation diminishing ([4, 6, 11, 13, 16, 17, 18, 19]). In order to obtain properties (2) and (3) for the solution, some conditions must be imposed on the functions f and $\tilde{f}$. In this paper we assume that $(f, \tilde{f}, \|\cdot\|)$ satisfy

$$\|y+\tau f(y)\| \leq \|y\|, \quad \left\|y+\tilde{\tau}\tilde{f}(y)\right\| \leq \|y\| \quad \forall\, y \in \mathbb{R}^m, \tag{4}$$

for some fixed $\tau, \tilde{\tau} > 0$. In such case, it can be proved that the solution of the problem satisfies the monotonicity property (2). We denote this class of problems by $\mathscr{F}(\tau, \tilde{\tau})$.

We also assume that

$$y+\sigma f(y) \geq 0, \quad y+\tilde{\sigma}\tilde{f}(y) \geq 0 \quad \forall\, y \in \mathbb{R}^m,\, y \geq 0, \tag{5}$$

for some fixed $\sigma, \tilde{\sigma} > 0$. In such case, the solution of the problem satisfies (3). We denote this class of positive problems by $\mathscr{F}_+(\sigma, \tilde{\sigma})$.

When f and $\tilde{f}$ have different stiffness properties, a common way to solve numerically the IVP (1) is by means of Additive Runge-Kutta (ARK) methods. An s-stage ARK method $(\mathbb{A}, \tilde{\mathbb{A}})$ is defined by two $(s+1)\times(s+1)$ real matrices

$$\mathbb{A} = \begin{pmatrix} \mathscr{A} & 0 \\ b^t & 0 \end{pmatrix}, \qquad \tilde{\mathbb{A}} = \begin{pmatrix} \tilde{\mathscr{A}} & 0 \\ \tilde{b}^t & 0 \end{pmatrix},$$

where $\mathscr{A}$ and $\tilde{\mathscr{A}}$ are $s \times s$ matrices, and $b, \tilde{b} \in \mathbb{R}^s$. The numerical solution $u_{n+1} \approx u(t_n + \Delta t)$ from $u_n \approx u(t_n)$ is given by

$$u_{n+1} = u_n + \Delta t \sum_{i=1}^{s} b_i\, f(U_{n,i}) + \Delta t \sum_{i=1}^{s} \tilde{b}_i\, \tilde{f}(U_{n,i}),$$

where the internal stages $U_{n,i}$ are given by

$$U_{n,i} = u_n + \Delta t \sum_{j=1}^{s} a_{ij}\, f(U_{n,j}) + \Delta t \sum_{j=1}^{s} \tilde{a}_{ij}\, \tilde{f}(U_{n,j}). \tag{6}$$

The matrices $\mathbb{A}$ and $\tilde{\mathbb{A}}$ are chosen so that the problem (1) is integrated efficiently. If, for instance, f corresponds to the convection term, and $\tilde{f}$ corresponds to the diffusion term, then a suitable way to proceed consists in using an explicit method for f and an implicit one for $\tilde{f}$. ARK methods combining implicit and explicit schemes are usually called IMplicit-EXplicit (IMEX) Runge-Kutta methods ([1], [15]). Many IMEX Runge-Kutta methods satisfy $\mathscr{A}e = \tilde{\mathscr{A}}e$, where e denotes the vector with all ones. In this way, the abscissae of the method $c_i = \sum_{j=1}^{s} a_{ij}$ and $\tilde{c}_i = \sum_{j=1}^{s} \tilde{a}_{ij}$, $i = 1, \ldots, s$, coincide. This property simplifies to a great extent the form of the order conditions of the method.

The internal stages $U_{n,i}$ approximate the exact solution $u(t)$ at $t = t_n + c_i \Delta t$. For many methods it holds that $c_i \geq 0$, and therefore $t_n + c_i \Delta t \geq t_n$. Consequently, a natural monotonicity requirement for both the internal stages and the numerical solution is

$$\|U_{n,i}\| \leq \|u_n\|, \quad i = 1, \ldots, s, \qquad \|u_{n+1}\| \leq \|u_n\|, \tag{7}$$

for all $n \geq 0$, probably under a stepsize restriction $\Delta t \leq \Delta t_{\max}$. In some contexts methods preserving this property (7) are called strong stability preserving (SSP) methods ([6], [7], [17]). In the same way, if the solution satisfies the positivity condition (3), a natural requirement for the numerical solution is

$$U_{n,i} \geq 0, \quad i = 1, \ldots, s, \qquad u_{n+1} \geq 0, \tag{8}$$

for all $n \geq 0$, probably under a stepsize restriction $\Delta t \leq \Delta t_{\max}$.

The aim of this paper is to show how ARK methods must be used in order to preserve monotonicity and positivity properties of the problem (1). The rest of the paper is organized as follows. In §2 we review some results concerned to monotonicity and positivity properties. In §3, we show three ARK methods. Two of these methods are used in §4, where the numerical experiments confirm the results in §2.

§2. Monotonicity and positivity for ARK methods

Positivity and other monotonicity properties for Runge-Kutta methods have been studied by different authors ([11], [12], [18]). In this context the concept of radius of absolute monotonicity plays an important role. This concept was extended to ARK methods in [7].

Definition 1. [7] An s-stage ARK method $(\mathbb{A}, \tilde{\mathbb{A}})$ is said to be absolutely monotonic (a.m.) at a given point $(\xi, \tilde{\xi}) \in \mathbb{R}^2$ if the matrix $I - \xi\mathbb{A} - \tilde{\xi}\tilde{\mathbb{A}}$ is regular and

$$(I - \xi\mathbb{A} - \tilde{\xi}\tilde{\mathbb{A}})^{-1}\,\mathbb{A} \geq 0,$$

$$(I - \xi\mathbb{A} - \tilde{\xi}\tilde{\mathbb{A}})^{-1}\,\tilde{\mathbb{A}} \geq 0,$$

$$(I - \xi\mathbb{A} - \tilde{\xi}\tilde{\mathbb{A}})^{-1}\,e \geq 0.$$

The additive method is said to be a.m. on a given set $\Omega \in \mathbb{R}^2$ if it is a.m. at each point $(\xi, \tilde{\xi}) \in \mathbb{R}^2$. The region of absolute monotonicity, denoted by $\mathscr{R}(\mathbb{A}, \tilde{\mathbb{A}})$, is defined by

$$\mathscr{R}(\mathbb{A}, \tilde{\mathbb{A}}) = \left\{ (r, \tilde{r}) \;\middle|\; r \geq 0, \tilde{r} \geq 0 \text{ and } (\mathbb{A}, \tilde{\mathbb{A}}) \text{ is a.m. on } [-r, 0] \times [-\tilde{r}, 0] \right\}. \tag{9}$$

For more details see [7].

Remark 1. For Runge-Kutta methods the radius of absolute monotonicity is defined by

$$R(\mathbb{A}) = \sup\{\, r \mid r \geq 0 \text{ and } \mathbb{A} \text{ is absolutely monotonic on } [-r, 0] \,\}.$$

The analogous concept for ARK methods is the curve of absolute monotonicity $\partial R(\mathbb{A}, \tilde{\mathbb{A}})$, defined now by the frontier of the region $\mathscr{R}(\mathbb{A}, \tilde{\mathbb{A}})$ in (9), excluding the coordinate axis.

In the following Theorem from [7], monotonicity for an ARK method is ensured under certain stepsize restrictions.

Theorem 1. *Consider the IVP* (1) *with* $(f, \tilde{f}) \in \mathscr{F}(\tau, \tilde{\tau})$. *Assume that the ARK method* $(\mathbb{A}, \tilde{\mathbb{A}})$ *is a.m. at* $(-r, -\tilde{r})$. *Then for*

$$\Delta t \leq r\,\tau, \qquad \Delta t \leq \tilde{r}\,\tilde{\tau},$$

the internal stages and the numerical solution satisfy the monotonicity inequalities (7).

Sufficient conditions to obtain positivity for the numerical solution of the ARK method are given in the following result.

Theorem 2 ([9]). *Consider the IVP* (1) *with* $(f,\tilde{f}) \in \mathscr{F}_+(\sigma,\tilde{\sigma})$. *Let* $U_{n,1},\ldots,U_{n,s},u_{n+1}$, *be the internal stages and the numerical solution obtained with the ARK method* $(\mathbb{A},\tilde{\mathbb{A}})$ *from* u_n, *with* $u_n \geq 0$. *Assume too that for* $\Delta t \leq H$, *the system* (6) *has a unique solution* $U = U(h,u_n)$ *that depends continuously on* h *and* u_n. *If the ARK method is a.m. at* $(-r,-\tilde{r})$ *with* $r\sigma,\tilde{r}\tilde{\sigma} \leq H$, *then, for*

$$\Delta t \leq r\sigma, \qquad \Delta t \leq \tilde{r}\tilde{\sigma},$$

the internal stages and the numerical solution satisfy the positivity inequalities (8).

Observe that in order to have $\Delta t > 0$ in Theorems 1 and 2, we require that the region of a.m. contains values $(r,\tilde{r})$ with $r \neq 0$, $\tilde{r} \neq 0$. An algebraic criteria to check this condition is given in [7]. We remark that the fact that $R(\mathbb{A}) \geq 0$, $R(\tilde{\mathbb{A}}) \geq 0$ does not imply the above condition as some coupling conditions between both methods are required.

§3. Some ARK methods

Depending on the shape of the region of absolute monotonicity $R(\mathbb{A},\tilde{\mathbb{A}})$, sharper stepsize restrictions may occur to maintain monotonicity for the ARK method. In order to have a good monotone ARK method, the Runge-Kutta methods $\mathbb{A}$ and $\tilde{\mathbb{A}}$ not only must have large radious of a.m. but also they must be properly coupled. Next we show three ARK methods. More details and more ARK methods can be viewed in [1, 7].

Example 1. Here we show the forward-backward Euler method padded as an IMEX method [1, 7].

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline \mathbb{A} & 1 & 0 \end{array} \qquad \begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 0 & 1 \\ \hline \tilde{\mathbb{A}} & 0 & 1 \end{array} \tag{10}$$

For this ARK method, $R(\mathbb{A}) = 1$, $R(\tilde{\mathbb{A}}) = +\infty$ and the region of absolute monotonicity $\mathscr{R}(\mathbb{A},\tilde{\mathbb{A}})$ is the biggest one: the cartesian product of $[0,R(\mathbb{A})]$ and $[0,R(\tilde{\mathbb{A}})]$

$$\mathscr{R}(\mathbb{A},\tilde{\mathbb{A}}) = \left\{ (r,\tilde{r}) \;\middle|\; 0 \leq r \leq 1, 0 \leq \tilde{r} \right\}.$$

In this case the two Runge-Kutta methods are perfectly coupled.

Example 2. IMEX order 2 method (Pareshi & Russo [15])

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{3}{2} & \frac{3}{2} & 0 & 0 \\ 1 & \frac{2}{3} & \frac{1}{3} & 0 \\ \hline \mathbb{A} & \frac{2}{3} & \frac{1}{3} & 0 \end{array} \qquad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{3}{2} & \frac{5}{4} & \frac{1}{4} & 0 \\ 1 & \frac{5}{9} & \frac{1}{9} & \frac{1}{3} \\ \hline \tilde{\mathbb{A}} & \frac{5}{9} & \frac{1}{9} & \frac{1}{3} \end{array} \tag{11}$$

For this ARK method $R(\mathbb{A}) = 2/3$, $R(\tilde{\mathbb{A}}) = 4/5$ and

$$R(\mathbb{A},\tilde{\mathbb{A}}) = \left\{ (r,\tilde{r}) \;\middle|\; 0 \le r \le \frac{2}{3}, 0 \le \tilde{r} \le \frac{2}{5}(2-3r) \right\}.$$

Observe that in this case, although the two Runge-Kutta methods are not perfectly coupled, $\mathscr{R}(\mathbb{A},\tilde{\mathbb{A}})$ contents points $(r,\tilde{r})$ with $r \neq 0$, $\tilde{r} \neq 0$. With this ARK method there may be stepsize restrictions to get monotonicity or positivity due to the restriction on the absolute monotonicity region.

Example 3. IMEX SSP2 (Pareshi & Russo [15])

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \mathbb{A} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \qquad \begin{array}{c|ccc} \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \hline \tilde{\mathbb{A}} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \tag{12}$$

For this method we have $R(\mathbb{A}) = 2$, $R(\tilde{\mathbb{A}}) = 12/5$, but $\partial R(\mathbb{A},\tilde{\mathbb{A}}) = (0,0)$ and consequently this is not a good method for monotone problems.

§4. Numerical experiments

In order to check the results shown in section 2 we have considered two problems: the Broadwell model, a hyperbolic system with relaxation studied in [2], and a chemical reaction involving eight reactants.

4.1. Broadwell model

In this section we consider the Broadwell model, a simple velocity kinetic model for a two-dimensional gas [2], and show how the above results can be used to obtain monotonicity for the entropy function. The discrete model after an upwind semmidiscretization in space is as follows

$$\begin{aligned} \partial_t f_j + \frac{f_j - f_{j-1}}{\Delta x} &= \frac{1}{\varepsilon}\left(h_j^2 - f_j g_j\right), \\ \partial_t h_j &= -\frac{1}{\varepsilon}\left(h_j^2 - f_j g_j\right), \\ \partial_t g_j - \frac{g_{j+1} - g_j}{\Delta x} &= \frac{1}{\varepsilon}\left(h_j^2 - f_j g_j\right). \end{aligned} \tag{13}$$

For the continuous problem there exists an entropy function

$$\mathscr{H}(t) = \int f(x,t)\log f(x,t) + 2h(x,t)\log h(x,t) + g(x,t)\log g(x,t)\,dx,$$

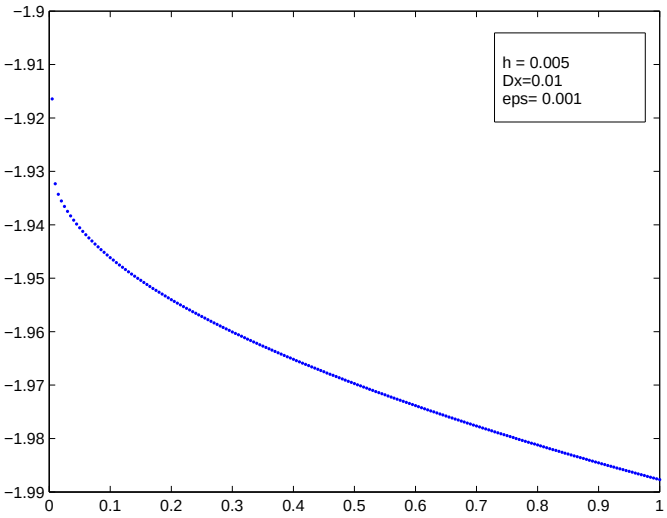

Figure 1: Broadwell model. Graph of the discrete entropy in logarithm scale.

that is monotonically decreasing. This monotonicity property is also preserved for the discrete entropy

$$\mathscr{H}_{\Delta x}(t) = \Delta x \sum_j \big(f_j(t)\log f_j(t) + 2h_j(t)\log h_j(t) + g_j(t)\log g_j(t)\big),$$

correspondig to the semidiscrete problem (13).

We have integrated the equation (13) with the ARK method (11), considering periodic boundary conditions. For this method, Theorem 1 guarantees monotonicity for the discrete entropy function under the stepsize restriction

$$\Delta t \leq \frac{4\Delta x\,\varepsilon}{5K_0 + 6\varepsilon}. \tag{14}$$

where K_0 is a constant which depends on the initial values. For details see [8]. In Figure 1 we show the discrete entropy in logarithm scale for $\Delta t = 0.005$, $\Delta x = 0.01$ and $\varepsilon = 0.001$, according to (14).

4.2. HIRES problem

This problem may be displayed in the form (1) with $u \in \mathbb{R}^8$, and $t \geq 0$. The HIRES problem originates from plant physiology and describes how light is involved in morphogenesis. The solution of the problem is positive ($u_i(t) \geq 0,\ \ i = 1,\ldots,8$). For details see [14].

The linear part of the problem has been considered as f and the non-linear part as $\tilde{f}$. We have integrated this problem with the ARK methods (10) and (12), considering different positive initial values. For example, for $u_0 = (1,0,0,0,0,57,0,57)$, the problem is in the set $\mathscr{F}_+(\sigma,\tilde{\sigma})$ with $\sigma = 0.0997$ and $\tilde{\sigma} = 6.3\cdot 10^{-5}$.

Theorem 2 applied to the *good* method guarantees positivity for the numerical solution under the stepsize restriction

$$\Delta t \leq \min(r\,\sigma, \tilde{r}\,\tilde{\sigma}) = 0.0997.$$

We have confirmed this sufficient condition numerically. The stepsize restriction is not too sharp as we have obtained positivity for the numerical solution for $\Delta t \leq \Delta t_{\text{num}} \approx 0.191$.

It is not possible to obtain a sufficient condition for method (12). In this case Theorem 2 gives $\min(r\,\sigma, \tilde{r}\,\tilde{\sigma}) = 0$. As we expected for the *bad* method, there is a severe stepsize restriction to obtain positivity. This is only possible for $\Delta t \leq \Delta t_{\text{num}} \approx 2.4E-4$.

Acknowledgements

This work has been partially supported by the Ministerio de Ciencia y Tecnología, Project BFM2001-2188.

References

[1] ASCHER, U., RUUTH, S., AND SPITERI, R. Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations. *Appl. Numer. Math. 25* (1997), 151–167.

[2] CAFLISCH, R., JIN, S., AND RUSSO, G. Uniformly accurate schemes for hyperbolic systems with relaxation. *SIAM J. Numer. Anal. 34* (1997), 246–281.

[3] CARRILLO, J., JUNGEL, A., AND TANG, S. Positive entropic schemes for a nonlinear fourth-order parabolic equation. *Discrete and continuous dynamical systems 3* (2003), 1–20.

[4] FERRACINA, L., AND SPIJKER, M. Stepsize restrictions for the TVD property in general Runge-Kutta methods. *SIAM J. Numer. Anal. 42* (2004), 1073–1093.

[5] GERISCH, A., AND WEINER, R. On the positivity of low order explicit Runge-Kutta schemes applied in splitting methods. *Comput. Math. Appl. 45* (2003), 53–67.

[6] GOTTLIEB, S., AND SHU, C. Total variation diminishing Runge-Kutta schemes. *Math. Comp. 67* (1998), 73—85.

[7] HIGUERAS, I. Strong stability for additive Runge-Kutta methods. *SIAM J. Numer. Anal. 44* (2006), 1735–1758

[8] HIGUERAS, I., AND ROLDÁN, T. On strong stability for additive Runge-Kutta methods: entrophy monotonicity for Broadwell model. *Journal of Scientific Computing* (2004). Preprint submitted for publication.

[9] HIGUERAS, I., AND ROLDÁN, T. Positivity for Runge-Kutta and additive Runge-Kutta methods. *Preprint* (2005).

[10] HORVÁTH, Z. Positivity of RK and diagonally split RK methods. *Appl. Numer. Math. 28* (1998), 309–326.

[11] HUNDSDORFER, W., AND VERWER, J. *Numerical solution of time-dependent Advection-Diffusion-Reaction equations*. Springer Series in Computational Mathematics, 2003.

[12] KRAAIJEVANGER, J. Contractivity of Runge-Kutta methods. *BIT 31* (1991), 482–528.

[13] LEVEQUE, R. J. *Numerical Methods for Conservation Laws*. Lectures in Mathematics, ETH Zrich, Birkhuser, Basel, 1992.

[14] MAZZIA, F., AND IAVERNARO, F. *Test Set for Initial Value Problem Solvers*, release 2.2. Department of Mathematics, University of Bari (Italy). Avalaible at `http://pitagora.dm.uniba.it/~testset`. Formerly maintained by CWI Amsterdam.

[15] PARESCHI, L., AND RUSSO, G. Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation. *J. Sci. Comput.* (2004). To apperar in.

[16] RUUTH, S., AND SPITERI, R. Two barriers on strong stability preserving time discretization methods. *J. Sci. Comput. 17* (2002), 211–220.

[17] SHU, C., AND OSHER, S. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. of Computational Phys. 77* (1988), 439–471.

[18] SPIJKER, M. A note on contractivity in the numerical solution of initial value problems. *BIT 27* (1987), 424–437.

[19] SPITERI, R., AND RUUTH, S. A new class of optimal high order strong stability preserving time discretization methods. *SIAM J. Numer. Anal. 40* (2002), 469–491.

Inmaculada Higueras and Teo Roldán
Departamento de Matemática e Informática
Universidad Pública de Navarra
Campus de Arrosadía
31006 Pamplona, Spain
`higueras@unavarra.es` and `teo@unavarra.es`