



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Journal of Computational and Applied Mathematics 190 (2006) 74–98

JOURNAL OF  
COMPUTATIONAL AND  
APPLIED MATHEMATICS

[www.elsevier.com/locate/cam](http://www.elsevier.com/locate/cam)

# Some new additive Runge–Kutta methods and their applications

Hongyu Liu<sup>\*,1</sup>, Jun Zou<sup>2</sup>

*Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong*

Received 4 November 2004; received in revised form 24 January 2005

Dedicated to Roderick S.C. Wong on the occasion of his 60th birthday

---

## Abstract

We propose some new additive Runge–Kutta methods of orders ranging from 2 to 4 that may be used for solving some nonlinear system of ODEs, especially for the temporal discretization of some nonlinear systems of PDEs with constraints. Only linear ODEs or PDEs need to be solved at each time step with these new methods.

© 2005 Elsevier B.V. All rights reserved.

*Keywords:* Additive Runge–Kutta method; A-stability; Stiffly accurate; Constrained PDEs

---

## 1. Preliminaries

In this paper we shall explore some new additive Runge–Kutta methods that may be used for solving some nonlinear system of ODEs, especially for the temporal discretization of some nonlinear systems of PDEs with constraints: for example, the Navier–Stokes equations in incompressible flow, where the velocity satisfies the divergence-free constraint, and the mean-field magnetic induction system in geodynamo modeling, where the magnetic field has to satisfy the divergence-free constraint. Only linear ODEs or PDEs need to be solved at each time step with these new additive Runge–Kutta methods.

---

\* Corresponding author.

*E-mail addresses:* [hylu@math.cuhk.edu.hk](mailto:hylu@math.cuhk.edu.hk) (H. Liu), [zou@math.cuhk.edu.hk](mailto:zou@math.cuhk.edu.hk) (J. Zou).

<sup>1</sup> Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, supported by a Direct Grant of CUHK.

<sup>2</sup> Substantially supported by Hong Kong RGC Grants (Projects 403403 and 4048/02P).

We start with the introduction of some basic notations and results on the existing Runge–Kutta (RK) methods.

As usual, we shall represent a standard s-stage RK method (cf. [7]) by the tableau

$$\begin{array}{c|c} c & A^T \\ \hline & b^T \end{array} = \begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array} \tag{1}$$

where  $A = (a_{ij})_{i,j=1}^s$  is the coefficient matrix,  $b^T = (b_i)_{i=1}^s$  the weight vector, and  $c = (c_i)_{i=1}^s$  a vector used to specify the discrete times. We shall use  $h > 0$  to denote the time stepsize, and  $t_n = t_0 + nh$ ,  $n = 0, 1, 2, \dots$ , for the discrete time points. When applied to the following system of first-order ODEs,

$$y'(t) = f(t, y) \quad \text{for } t > t_0; \quad y(t_0) = y_0, \tag{2}$$

where  $y \in \mathbb{R}^m$  and  $f : (t_0, \infty) \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  is a nonlinear vector-valued function, scheme (1) can be written as

$$y_i^{(n)} = y_{n-1} + h \sum_{j=1}^s a_{ij} f(t_{n-1} + c_j h, y_j^{(n)}), \quad i = 1, 2, \dots, s, \tag{3}$$

$$y_n = y_{n-1} + h \sum_{i=1}^s b_i f(t_{n-1} + c_i h, y_i^{(n)}), \tag{4}$$

where  $y_n$  is an approximation of  $y(t_n)$  and  $y_i^{(n)} \approx y(t_{n-1} + c_i h)$ .

Without loss of generality, it suffices for us to consider only one time step of scheme (3)–(4). So we shall set  $n = 1$  and write the intermediate stage vectors  $y_i^{(n)}$  as  $Y_i \approx y(t_0 + c_i h)$  and  $y_1 \approx y(t_1)$ . Then scheme (3)–(4) can be expressed as

$$Y_i = y_0 + h \sum_{j=1}^s a_{ij} f(t_0 + c_j h, Y_j), \quad i = 1, 2, \dots, s, \tag{5}$$

$$y_1 = y_0 + h \sum_{i=1}^s b_i f(t_0 + c_i h, Y_i). \tag{6}$$

In general, the parameters  $c_i$ 's in (5)–(6) are required to satisfy the conditions

$$c_i = \sum_{j=1}^s a_{ij}, \quad i = 1, \dots, s,$$

in order to essentially simplify order conditions, especially for high-order methods. These conditions indicate that the local truncation error of each approximation in (5) is at least first-order accurate.

Stability will be one of our central issues to be considered when we construct the new RK methods. Recall that the stability function  $R(z)$  with  $z = \lambda h$  of an RK method is the approximate solution generated by one step of the method for the Dahlquist test problem:

$$y'(t) = \lambda y, \quad y_0 = 1, \tag{7}$$

with  $\operatorname{Re}\{\lambda\} < 0$ . Then the stability region of the method is defined to be

$$S = \{z \in \mathbb{C}; |R(z)| \leq 1\}. \quad (8)$$

We know that the stability function of the implicit RK method (1) is given (cf. [7]) by

$$R(z) = \frac{\det(I - zA + z\mathbf{1}b^T)}{\det(I - zA)}. \quad (9)$$

Recall that a method is said to be  $A(\alpha)$ -stable if the sector

$$S_\alpha = \{z; |\arg(-z)| \leq \alpha, z \neq 0\}$$

is contained in the stability region. An  $A(\frac{\pi}{2})$ -stable method is called  $A$ -stable.

It can be seen that an RK method is  $A$ -stable if and only if

$$|R(iy)| \leq 1 \text{ for all real } y, \quad R(z) \text{ is analytic for } \operatorname{Re}\{z\} < 0. \quad (10)$$

Clearly, an  $A$ -stable RK method must be implicit. For very stiff problems, one may need schemes with  $L$ -stability, that is, schemes which are  $A$ -stable and

$$\lim_{z \rightarrow \infty} R(z) = 0. \quad (11)$$

In many stiff situations, the  $L$ -stability condition (11) may be too strong. In fact, it is often sufficient in those very stiff cases to require that  $\lim_{z \rightarrow \infty} |R(z)|$  equals some constant less than, for example,  $\alpha < \frac{1}{2}$ ; this will be used in our subsequent construction of the additive methods.

## 2. Additive Runge–Kutta methods

The focus of our study is on the numerical solution of the following ODEs of additive form:

$$y'(t) = f(t, y) + g(t, y), \quad (12)$$

where  $y \in R^m$ ,  $f : (t_0, \infty) \times R^m \rightarrow R^m$  is stiff, but linear with respect to  $y$ , and  $g(t, y)$  is nonlinear but not stiff. We remark that our subsequent construction techniques can be equally applied to construct similar schemes for the case where  $f(t, y)$  is linear but non-stiff, while  $g(t, y)$  is nonlinear but stiff.

As we will show in Section 10, system (12) can also be PDEs with constraints, where  $f(t, y)$  and  $g(t, y)$  are both differential operators of space variables. In fact, the PDEs with constraints are the major systems our new additive RK methods intend to solve.

The additive methods use the idea of the RK method but aim to provide a more effective way to deal with the ODEs of the additive form (12). One can already find some additive RK methods in the literature, which can achieve favorable results in numerical solutions of certain stiff additive ODEs like (12) (see [1,5,6,9,12]). But all the existing additive schemes are not suitable to the PDEs with constraints, or are applicable but very expensive.

When applied to Eq. (12), an  $s$ -stage additive RK method is a scheme of the form:

$$y_i^{(n)} = y_s^{(n-1)} + h \sum_{j=1}^s a_{ij} f(t_{n-1} + c_j h, y_j^{(n)}) + h \sum_{j=1}^s b_{ij} g(t_{n-1} + c_j h, y_j^{(n)}), \quad (13)$$

where  $i = 1, 2, \dots, s$  and  $n = 1, 2, 3, \dots$ . The following *row conditions* are always assumed

$$c_i = \sum_{j=1}^s a_{ij} = \sum_{j=1}^s b_{ij}, \tag{14}$$

in order to simplify the order conditions so that it is possible to find some reasonable high-order schemes. By putting  $f(t, y) = 0$  or  $g(t, y) = 0$  in (12), we can get from (13) two standard  $s$ -stage RK methods of form (1) with the weighted values  $b_i = a_{si}$  or  $b_{si}$ . Hence, they are *stiffly accurate* in parlance of [11] and one can conveniently represent an  $s$ -stage additive RK method by the triple  $(c, \mathbf{A}, \mathbf{B})$  or the following tableau (with weights as in (1) no longer necessary):

$c_1$	$a_{11}$	$\cdots$	$a_{1s}$	$b_{11}$	$\cdots$	$b_{1s}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
$c_s$	$a_{s1}$	$\cdots$	$a_{ss}$	$b_{s1}$	$\cdots$	$b_{ss}$

(15)

Due to the coupling between blocks  $\mathbf{A}$  and  $\mathbf{B}$  in (15), the order conditions and the stability analysis for the additive RK methods are much more complicated than the standard (1).

### 3. Stability analysis

The general additive RK methods were first introduced by Cooper and Sayfy (cf. [5,6]) for solving the stiff problem

$$y'(t) = F(t, y) \quad \text{for } t > t_0; \quad y(t_0) = y_0. \tag{16}$$

They associated the additive methods with a sequence of decompositions of  $F(t, y)$  of the form

$$F(t, y) = J^{(n)} y + g^{(n)}(t, y), \quad n = 1, 2, 3, \dots, \tag{17}$$

where  $\{J^{(n)}\}$  is chosen to be independent of  $t$  and often as an approximation to the Jacobian of  $F$  evaluated at some sequence of computed values.

To study the stability of the additive methods, let  $y$  be the particular solution of the system

$$y'(t) = Jy + g(t, y), \quad t > t_0, \tag{18}$$

which has the initial value  $y(t_0) = y_0$ . It is known that if the trivial solution of  $y' = Jy$  is exponentially stable, i.e.,  $\text{Re}\{\lambda\} < 0$  for all  $\lambda \in \sigma[J]$ , and that  $\|g(t, y)\| = o(\|y\|)$ , then the trivial solution of (18) is also exponentially stable, which implies that there is an  $\varepsilon > 0$  such that if  $\|y_0\| \leq \varepsilon$ , then  $\|y(t)\|$  has limit zero. Using this model problem, Cooper and Sayfy established the stability for their additive schemes associated with system (18).

**Theorem 1.** *Suppose that  $J$  is a constant  $m \times m$  matrix. The trivial solution of  $y'(t) = Jy$  is exponentially stable and  $\|g(t, y)\| = o(\|y\|)$ . Furthermore, we assume that the  $(c, \mathbf{A}, \mathbf{B})$  RK method is linearly*

implicit, i.e.,

$$y_i^{(n)} = y_s^{(n-1)} + h \sum_{j=1}^i a_{ij} J y_j^{(n)} + h \sum_{j=1}^{i-1} b_{ij} g(t_{n-1} + hc_j, y_j^{(n)}), \quad (19)$$

and the scheme  $(c, \mathbf{A})$  is  $A$ -stable. Then for any fixed positive  $h$  and arbitrary  $y_s^{(0)}$ , method (19) uniquely defines a sequence  $y_s^{(n)}$  for  $i=1, 2, \dots, s$  and  $n=1, 2, 3, \dots$ , and there exists a  $\delta > 0$  such that if  $\|y_s^{(0)}\| \leq \delta$ , the sequence  $\{y_s^{(n)}\}$  has limit zero.

We are interested in the more general additive system (12), instead of system (16) with special Jacobian-type decompositions (17). Clearly the results in Theorem 1 cannot be used for the stability estimates of the additive methods (15), especially the smallness condition, i.e.,  $\|g(t, y)\| = o(\|y\|)$ , which plays a key role in the proof of the above theorem, is usually not true to system (12) of our interest.

For our purpose, we shall take a similar approach to the one in [12] to directly apply the linear stability analysis for standard RK methods to the following test problem associated with (12):

$$\frac{dy}{dt} = \lambda_f y + \lambda_g y, \quad y_0 = 1, \quad (20)$$

where  $\lambda_f$  and  $\lambda_g$  represent the eigenvalues of  $\partial f/\partial y$  and  $\partial g/\partial y$  in (12). They are complex parameters satisfying<sup>3</sup>  $\operatorname{Re}\{\lambda_f\} \leq 0$ ,  $\operatorname{Re}\{\lambda_g\} \leq 0$  and  $|\operatorname{Re}\{\lambda_f\}| \gg |\operatorname{Re}\{\lambda_g\}|$ .

We emphasize that the additive RK schemes proposed in [12] are different from the RK schemes of form (13) to be studied in the current paper, and cannot be used for those PDE systems with constraints, see Section 10.

By applying the additive method (15) to the test problem (20), we get after one time step that

$$y_1 = R(z_f, z_g), \quad z_f = h\lambda_f, \quad z_g = h\lambda_g, \quad (21)$$

where  $R(z_f, z_g)$  is the stability function. Then we introduce our new definition of the  $A(\alpha)$ - and  $L$ -stability for the additive methods.

**Definition 1.** An  $A(\alpha)$ -stability domain of an additive method (15) in the complex plane of  $z_g = h\lambda_g$  is defined as the domain

$$S_\alpha^{f \cap g} = \{z_g \in \mathbb{C}; |R(z_f, z_g)| \leq 1 \text{ for all } z_f \in S_\alpha^f\}, \quad (22)$$

where

$$S_\alpha^f = \{z_f \in \mathbb{C}; |\arg(-z_f)| \leq \alpha, z_f \neq 0\} \cup \{0\}. \quad (23)$$

An additive method (15) is said to be  $L$ -stable if it is  $A$ -stable (i.e.  $A(\frac{\pi}{2})$ -stable) and

$$\lim_{z_f \rightarrow \infty} |R(z_f, z_g)| = 0. \quad (24)$$

<sup>3</sup> Our subsequent analysis can be extended equally to the case  $|\operatorname{Re}\{\lambda_f\}| \ll |\operatorname{Re}\{\lambda_g\}|$ .

If we denote by  $S^g$  the stability domain of the explicit part  $(c, \mathbf{B})$  in (15) as a standard RK method, we have

$$S_\alpha^{f \cap g} \subset S^g. \tag{25}$$

This can be seen by taking  $\lambda_f = 0$  in the test (21), and means that in the stability domain  $S_\alpha^{f \cap g}$ , the explicit part  $(c, \mathbf{B})$  is stable. In the forthcoming discussions, since the term  $f(t, y)$  in (12) is stiff, it is natural for us to require that the semi-explicit part  $(c, \mathbf{A})$  is to be  $A(\alpha)$ -stable, and we should have  $0 \in S_\alpha^{f \cap g}$ .

In case the first part  $(c, \mathbf{A})$  is  $A$ -stable as a standard RK method, then the  $L$ -stability of the additive method (15) implies that its first part  $(c, \mathbf{A})$  is still  $L$ -stable and this can be verified directly.

#### 4. Construction of the additive RK methods

Our construction of additive RK methods is based on the satisfaction of both stability and accuracy conditions. Corresponding to (15), we introduce the following notations for  $i = 1, 2, \dots, s, \sigma = 1, 2, 3, \dots$

$$a_i(\sigma) = c_i^\sigma - \sigma \sum_{j=1}^s a_{ij} c_j^{\sigma-1} \quad b_i(\sigma) = c_i^\sigma - \sigma \sum_{j=1}^s b_{ij} c_j^{\sigma-1}. \tag{26}$$

Then the Taylor expansion gives the order conditions for the general method (15) up to fourth order as follows (see [5]).

An additive RK method (23) is of order  $p \leq 4$  if and only if the conditions

$$a_i(1) = b_i(1) = 0, \quad i = 1, 2, \dots, s, \tag{27}$$

$$a_s(\sigma) = 0, \quad b_s(\sigma) = 0, \quad \sigma \leq p, \tag{28}$$

$$\sum_{i=1}^s a_{si} c_i^{\tau-1} a_i(\sigma) = 0, \quad \sum_{i=1}^s a_{si} c_i^{\tau-1} b_i(\sigma) = 0, \quad \sigma + \tau \leq p, \tag{29}$$

$$\sum_{i=1}^s b_{si} c_i^{\tau-1} a_i(\sigma) = 0, \quad \sum_{i=1}^s b_{si} c_i^{\tau-1} b_i(\sigma) = 0, \quad \sigma + \tau \leq p \tag{30}$$

and the following extra conditions for  $p = 4$ ,

$$\sum_{i=1}^s a_{si} \sum_{j=1}^s a_{ij} a_j(2) = 0, \quad \sum_{i=1}^s a_{si} \sum_{j=1}^s a_{ij} b_j(2) = 0, \tag{31}$$

$$\sum_{i=1}^s a_{si} \sum_{j=1}^s b_{ij} a_j(2) = 0, \quad \sum_{i=1}^s a_{si} \sum_{j=1}^s b_{ij} b_j(2) = 0, \tag{32}$$

$$\sum_{i=1}^s b_{si} \sum_{j=1}^s a_{ij} a_j(2) = 0, \quad \sum_{i=1}^s b_{si} \sum_{j=1}^s a_{ij} b_j(2) = 0, \tag{33}$$

$$\sum_{i=1}^s b_{si} \sum_{j=1}^s b_{ij} a_j(2) = 0, \quad \sum_{i=1}^s b_{si} \sum_{j=1}^s b_{ij} b_j(2) = 0, \quad (34)$$

are satisfied, where  $\sigma$  and  $\tau$  take all possible positive integer values.

## 5. Semi-implicit additive RK methods

As we can easily see, the additive RK schemes of the general form (15) are fully implicit, and will be very expensive when applied for solving ODEs or for the time marching schemes of PDEs. Recall that our target system is of form (12), where the linear part  $f(t, y)$  is stiff but the nonlinear part  $g(t, y)$  is not stiff. It will be more practical to have schemes that are only implicit in terms of  $f(t, y)$  and explicit in terms of  $g(t, y)$  as this needs to solve only linear systems at each intermediate time step. To further reduce the computational complexity, we would like to have schemes that are only semi-implicit in terms of  $f(t, y)$ , that is, the first coefficient matrix  $\mathbf{A}$  is lower triangular while the second coefficient matrix  $\mathbf{B}$  is strictly lower triangular in (15). Besides, in order to take into account certain constraint equations in the PDEs (see Section 10), we require that the computed value at each intermediate step, i.e.,  $y_i^{(n)}$  in (13), is given implicitly and this can be satisfied if  $a_{ii} \neq 0$  for  $i = 2, 3, \dots, s$ . On the other hand, it is quite necessary for every intermediate stage  $y_i^{(n)}$  of the method to give physical meaningful computation, that is,  $y_i^{(n)}$  should be the approximated value of  $y(t)$  at time  $t = t_{n-1} + c_i h$ . Naturally, we shall take  $c_s = 1$  and  $y_s^{(n)}$  as the final computed value of each time step.

In summary, the new additive RK methods we are looking for can now be described by the following tableau:

0	0		0
$c_2$	$a_{21}$	$a_{22}$	$b_{21}$ 0
$\vdots$	$\vdots$	$\ddots$	$\vdots$ $\ddots$ $\ddots$
1	$a_{s1}$	$\cdots$ $\cdots$ $a_{ss}$	$b_{s1}$ $\cdots$ $b_{s,s-1}$ 0

(35)

Before starting our construction, we shall discuss a bit about the difference between our new schemes and the existing ones. The most important existing additive schemes are the ones developed in [6], which take  $a_{ss} = 0$  and can therefore be viewed as the combination of an  $(s - 1)$ -stage diagonally implicit RK (DIRK) method (with weights  $b_i = a_{si}$ ,  $i = 1, 2, \dots, s - 1$ ) and an  $(s - 1)$ -stage explicit RK method (with weights  $b_i = b_{si}$ ,  $i = 1, 2, \dots, s - 1$ ). For the implicit part of those schemes, the diagonal entries take as many zeros as possible in order to make the derived scheme more efficient and ease the construction simultaneously. On the contrary, our new schemes are required to satisfy the conditions that  $a_{ii} \neq 0$  for  $i = 2, 3, \dots, s$  due to the application problems in our mind (see Section 10). If this is difficult, at least we should have  $a_{ss} \neq 0$ . Now, in the semi-implicit additive RK method (35), the  $(c, \mathbf{A})$  method is in fact an  $s$ -stage stiffly accurate DIRK method (with weights  $b_i = a_{si}$ ,  $i = 1, 2, \dots, s$ ) and the  $(c, \mathbf{B})$  method is reduced to an  $(s - 1)$ -stage explicit RK method (with weights  $b_i = b_{si}$ ,  $i = 1, 2, \dots, s - 1$ ). Thus, our new additive RK schemes can be viewed as the combination of an  $s$ -stage stiffly accurate DIRK method and an  $(s - 1)$ -stage explicit RK method. To our best knowledge, there are no such additive methods in the literature. In fact, the construction of such schemes are not straightforward at all and we cannot see

any direct application of the existing techniques for the construction of explicit RK and DIRK methods to achieve our purpose here. Theorem 2 below will play the key role in our subsequent construction of the schemes which meet our needs.

We are now going to construct the additive schemes of form (35) with  $a_{ii} \neq 0$  for  $i = 2, 3, \dots, s$ . In order to deal with the stiffness of the linear part  $f(t, y)$ , we shall require that the part  $(c, \mathbf{A})$  in (15) as a standard RK method is  $A(\alpha)$ -stable. The algebraic conditions for a semi-explicit RK method  $(c, \mathbf{A})$ , where the first diagonal element of  $\mathbf{A}$  is zero, to be  $A$ -stable can be given in terms of two sets of parameters  $\{\alpha_i\}_{i=0}^\infty$  and  $\{\beta_i\}_{i=0}^\infty$  (see [6,4]).  $\{\beta_i\}_{i=0}^\infty$  are defined by

$$\prod_{r=1}^s (1 - \tau a_{rr}) = \beta_0 - \tau \beta_1 + \tau^2 \beta_2 - \dots$$

This gives  $\beta_0 = 1$  and  $\beta_s = \beta_{s+1} = \dots = 0$ . Let  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_s$  be the natural basis for  $\mathbb{R}^s$  and let  $\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2 + \dots + \mathbf{e}_s$  be the vector with unit elements. The terms

$$\mathbf{e}_s^T \mathbf{A}^r \mathbf{e}, \quad r = 1, 2, 3, \dots,$$

are the sums of the elements in rows  $s$  of  $\mathbf{A}, \mathbf{A}^2, \mathbf{A}^3, \dots$  and for a method of order  $p$  it has that

$$\mathbf{e}_s^T \mathbf{A}^r \mathbf{e} = \frac{1}{r!}, \quad r = 1, 2, \dots, p.$$

Define  $\alpha_s = \alpha_{s+1} = \dots = 0$  and

$$\alpha_r = \beta_r - \beta_{r-1} \mathbf{e}_s^T \mathbf{A} \mathbf{e} + \dots + (-1)^r \beta_0 \mathbf{e}_s^T \mathbf{A}^r \mathbf{e}, \quad r = 0, 1, \dots, s - 1,$$

so that  $\alpha_0 = 1$ . Then a method of order  $p$  is  $A$ -stable iff  $a_{rr} \geq 0$  for  $r = 1, 2, \dots, s$  and

$$\sum_{r=\pi}^{s-1} y^r \sum_{j=0}^r * (-1)^{r+j} (\beta_{2r-j} \beta_j - \alpha_{2r-j} \alpha_j) \geq 0 \quad \forall y \geq 0,$$

where  $\pi$  is the integral part of  $p/2 + 1$  and the asterisk denotes that the terms  $j = r$  are halved.

Since in the semi-implicit additive RK method (35), the  $(c, \mathbf{B})$  method is in fact reduced to be an  $(s - 1)$ -stage standard explicit RK method, the order of such methods cannot exceed  $s - 1$ ; if  $s > 5$ , the order cannot exceed  $s - 2$ , see [7]. That is, if we want to construct a fourth-order scheme of this type, it is at least of 5-stage. By observing Eqs. (26)–(34), one can readily see that the order conditions turn out to be too tedious to deal with. Together with the  $A$ -stable conditions, the construction becomes extremely difficult. However, it will be much easier if we have  $a_{si} = b_{si}$  in (35) for  $i = 1, 2, \dots, s$ , as implemented in [6], since the number of the order conditions is almost halved as one can see from (26)–(34). But this conflicts with our requirement that  $a_{ss} \neq 0$  and  $b_{ss} = 0$ . Fortunately, we have the following solution to overcome this dilemma.

**Theorem 2.** Assume that the  $s$ -stage additive RK method (15) is of order  $p$ ,  $c_{s-1} = c_s = 1$ , and the  $(s - 1)$ th intermediate stage is an approximation to  $y(t)$  of order  $p - 1$ , i.e.,  $y_{s-1}^{(n)} = y(t_0 + nh) + \mathcal{O}(h^{p-1})$ . Then the method obtained by replacing  $b_{s,s-1}$  and  $b_{ss}$  with  $(b_{s,s-1} + b_{ss})$  and 0, respectively, is still of order  $p$ .



For the proof of this theorem, we need the following lemma:

**Lemma 1.** Suppose  $Y, Z \in \mathbb{R}^M$  are two vectors and satisfy

$$\|Z - Y\|_\infty \leq h\{\|A[f(Z) - f(Y)]\|_\infty + \|B[g(Z) - g(Y)]\|_\infty\} + \mathcal{O}(h^p), \quad (36)$$

where  $A, B \in \mathbb{R}^{M \times M}$  are two constant matrices,  $p$  is a positive integer and  $f, g : \mathbb{R}^M \mapsto \mathbb{R}^M$  are often differentiable and have bounded first-order derivatives. Then we have

$$\|Z - Y\|_\infty \leq \mathcal{O}(h^p). \quad (37)$$

**Proof.** By Eq. (36), we know that there must exist some constant  $C$  such that

$$\|Z - Y\|_\infty \leq Ch\|Z - Y\|_\infty + \mathcal{O}(h\|Z - Y\|_\infty^2) + \mathcal{O}(h^p). \quad (38)$$

Again, by Eq. (36), we easily see that

$$\|Z - Y\|_\infty \leq \mathcal{O}(h).$$

Using this and Eq. (38), we have

$$\|Z - Y\|_\infty \leq \mathcal{O}(h^2).$$

Repeating this process, we can derive (37).  $\square$

**Proof of Theorem 2.** Setting

$$Y^{(n)} = \begin{bmatrix} y_1^{(n)} \\ \vdots \\ y_s^{(n)} \end{bmatrix}, \quad F(Y^{(n)}) = \begin{bmatrix} f(t_{n-1} + c_1 h, y_1^{(n)}) \\ \vdots \\ f(t_{n-1} + c_s h, y_s^{(n)}) \end{bmatrix}, \quad G(Y^{(n)}) = \begin{bmatrix} g(t_{n-1} + c_1 h, y_1^{(n)}) \\ \vdots \\ g(t_{n-1} + c_s h, y_s^{(n)}) \end{bmatrix},$$

we can rewrite (13) as

$$Y^{(n)} = \mathbb{Y}^{(n-1)} + h(\mathbf{A} \otimes I)F(Y^{(n)}) + h(\mathbf{B} \otimes I)G(Y^{(n)}), \quad (39)$$

where  $\mathbb{Y}^{(k)}$  is a block column vector consisting of  $s$ ,  $m$ -dimension column vector  $y_s^{(k)}$ ,  $\mathbf{A}$  and  $\mathbf{B}$  are the coefficient matrices of the first and second part of method (15), respectively,  $I$  is the  $m \times m$  identity matrix, and  $\otimes$  denotes the Kronecker product.

Next, replacing  $b_{s,s-1}$  and  $b_{ss}$  with  $b_{s,s-1} + b_{ss}$  and 0, respectively, the obtained new method for Eq. (12) is

$$Z^{(n)} = \mathbb{Z}^{(n-1)} + h(\mathbf{A} \otimes I)F(Z^{(n)}) + h(\bar{\mathbf{B}} \otimes I)G(Z^{(n)}), \quad (40)$$

where  $Z^{(n)}$ ,  $\mathbb{Z}^{(n)}$  have the same usages as  $Y^{(n)}$ ,  $\mathbb{Y}^{(n)}$ , and  $\bar{\mathbf{B}}$  is the same as  $\mathbf{B}$  with only its  $(s, s-1)$ - and  $(s, s)$ -elements being  $b_{s,s-1} + b_{s,s}$  and 0, instead of  $b_{s,s-1}$  and  $b_{s,s}$  of  $\mathbf{B}$ . By subtraction of Eq. (39) and (40), we obtain

$$\begin{aligned} Z^{(n)} - Y^{(n)} &= (\mathbb{Z}^{(n-1)} - \mathbb{Y}^{(n-1)}) + h\{(\mathbf{A} \otimes I)[F(Z^{(n)}) - F(Y^{(n)})] \\ &\quad + (\bar{\mathbf{B}} \otimes I)[G(Z^{(n)}) - G(Y^{(n)})]\} + h[(\bar{\mathbf{B}} - \mathbf{B}) \otimes I]G(Y^{(n)}). \end{aligned} \quad (41)$$

In order to prove that the new method is also of order  $p$ , it is sufficient to show that the local truncation error of  $z_s^{(n)}$  is  $\mathcal{O}(h^{p+1})$ . Thus, we take  $n = 1$  and  $z_s^{(0)} = y_s^{(0)} = y_0$ . By noting that  $y_{s-1}^{(n)}$  is an approximation of  $y(t)$  of order  $p - 1$  and  $y_s^{(n)}$  is of order  $p$ , we see

$$\|[(\bar{\mathbf{B}} - \mathbf{B}) \otimes I]G(Y^{(1)})\|_\infty = \|b_{s,s}[G(y_{s-1}^{(1)}) - G(y_s^{(1)})]\|_\infty \leq C \|y_{s-1}^{(1)} - y_s^{(1)}\|_\infty \leq \mathcal{O}(h^p), \tag{42}$$

where  $C = |b_{s,s}|L$  and  $L$  is the bound for  $|g'|$ , and we have made use of the fact that the local truncation errors of  $y_{s-1}^{(1)}$  and  $y_s^{(1)}$  are, respectively,  $\mathcal{O}(h^p)$  and  $\mathcal{O}(h^{p+1})$ , by the assumption of the theorem, i.e.,

$$\|y_{s-1}^{(1)} - y(t_0 + h)\| = \mathcal{O}(h^p), \quad \|y_s^{(1)} - y(t_0 + h)\| = \mathcal{O}(h^{p+1}).$$

Hence, we derive by (41)

$$\begin{aligned} \|Z^{(1)} - Y^{(1)}\|_\infty &\leq h\{\|(\mathbf{A} \otimes I)[F(Z^{(1)}) - F(Y^{(1)})]\|_\infty \\ &\quad + \|(\mathbf{B} \otimes I)[G(Z^{(1)}) - G(Y^{(1)})]\|_\infty\} + \mathcal{O}(h^{p+1}). \end{aligned} \tag{43}$$

This implies, by Lemma 1,

$$\|Z^{(1)} - Y^{(1)}\|_\infty \leq \mathcal{O}(h^{p+1}). \tag{44}$$

Then we can deduce

$$\begin{aligned} \|z_s^{(1)} - y(t_0 + h)\|_\infty &\leq \|z_s^{(1)} - y_s^{(1)}\|_\infty + \|y_s^{(1)} - y(t_0 + h)\| \\ &\leq \|Z^{(1)} - Y^{(1)}\|_\infty + \|y_s^{(1)} - y(t_0 + h)\| \leq \mathcal{O}(h^{p+1}). \end{aligned} \tag{45}$$

The proof is completed.  $\square$

We know from Theorem 2 that for those  $p$ th-order additive RK methods of form (35), where the first  $(s - 1)$  stages form an  $(s - 1)$ -stage additive RK method, i.e.,  $a_{is} = b_{is} = 0$  for  $i = 1, \dots, s - 1$ , Theorem 2 says if the first  $(s - 1)$ -stage is of order  $p - 1$ , then by replacing  $b_{s,s-1}$  and  $b_{ss}$  by  $(b_{s,s-1} + b_{ss})$  and 0, respectively, the resulting method is still of order  $p$ .

This provides an important principle, which we can use to construct the desired additive RK methods:

To construct an additive method of form (35) with  $a_{ss} \neq 0$  but  $b_{ss} = 0$ , we can first construct a  $p$ th order method that satisfies  $a_{si} = b_{si}$  ( $i = 1, 2, \dots, s$ ),  $a_{ss} = b_{ss} \neq 0$ ,  $c_{s-1} = c_s = 1$  and the first  $(s - 1)$ -stage is a  $(p - 1)$ th order linearly implicit additive RK method. Then, replacing  $b_{s,s-1}$  and  $b_{ss}$  with  $(b_{s,s-1} + b_{ss})$  and 0, respectively, we will achieve a new method of order  $p$ , which is linearly implicit of form (35) and satisfies that  $a_{ss} \neq 0$  and  $b_{ss} = 0$ .

All our subsequent constructions of the new additive RK methods are based on the above principle. With this principle and the order conditions in Section 4, we can construct additive RK methods with orders ranging from 2 to 4. We shall ignore all the tedious derivations but present many different examples of such schemes in the subsequent sections. For most of the schemes, the  $A(\frac{\pi}{2})$ -stability region of the additive RK method will be plotted against the stability region of the corresponding explicit part  $(c, \mathbf{B})$  as a standard explicit RK method of form (1) for comparisons. Some numerical experiments with these new RK methods will be presented to demonstrate their accuracies, stabilities and efficiencies.

In the sequel, we shall use the descriptions like additive RK. $m$ . $A$ . $n$  or RK. $m$ . $L$ . $n$  for some positive integers  $m$  and  $n$ . “additive RK. $m$ . $A(L)$ . $n$ ” indicates an additive RK method, which is  $m$ th-order accurate and  $A$ -stable (or  $L$ -stable). The number  $n$  means it is the  $n$ th method of this class listed in this paper.

## 6. Examples of additive RK methods with nonzero diagonal entries

This section presents some new additive RK methods with nonzero diagonal entries  $a_{ii}$  for  $i = 2, 3, \dots, s$ . The following tableau gives a new class of additive RK methods of order 2 which is  $A$ -stable:

2nd Order						
0	0	0	0	0	0	0
c	$c - \alpha$	$\alpha$	0	c	0	0
1	$\frac{c - \beta c - \frac{1}{2} + \beta}{c}$	$\frac{\frac{1}{2} - \beta}{c}$	$\beta$	1 - $\frac{1}{2c}$	$\frac{1}{2c}$	0

where  $0 < c < 1$ ,  $\alpha, \beta$  are positive and satisfy

$$\alpha^2 \beta^2 \geq [\alpha\beta - (\alpha + \beta) + \frac{1}{2}]^2.$$

If we take

$$\alpha = \frac{\frac{1}{2} - \beta}{1 - \beta}, \quad \beta > 1 \text{ or } \beta < \frac{1}{2},$$

then the resulting method is  $L$ -stable. We first take  $c = 1/2$ ,  $\alpha = \beta = 1$  and get the following  $A$ -stable one:

Additive RK.2.A.1						
0	0	0	0	0	0	0
$\frac{1}{2}$	$-\frac{1}{2}$	1	0	$\frac{1}{2}$	0	0
1	1	-1	1	0	1	0

Its stability function  $R(z_f, z_g)$  as defined in (21) is

$$R(z_f, z_g) = \frac{(1 - z_f - \frac{1}{2}z_f^2) + (1 - z_f)z_g + \frac{1}{2}z_g^2}{1 - 2z_f + z_f^2}$$

and its  $A(\frac{\pi}{2})$ -stability domain is plotted in Fig. 1 (left). One can observe that the  $A(\frac{\pi}{2})$ -stability domain of the additive method, i.e.  $S_{\pi/2}^{f \cap g}$ , is slightly smaller than that of the explicit part as a conventional RK method, i.e.  $S^g$ . For additive stiff problem (12), if we only use the explicit part, which is a second-order

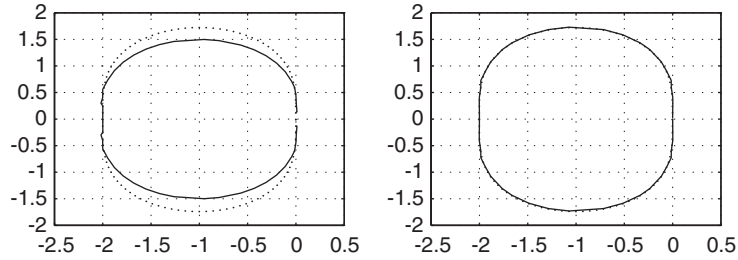


Fig. 1. Stability domain of additive (solid line) and explicit RK (dot) methods.

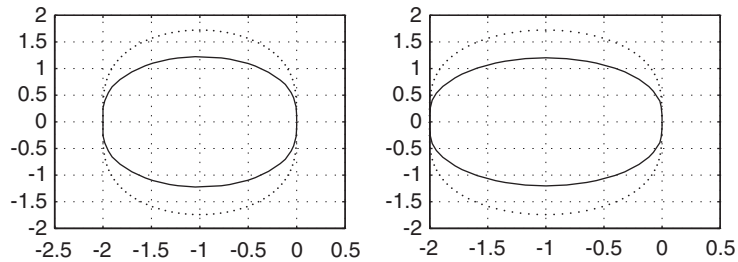


Fig. 2. Stability domain of additive (solid line) and explicit RK (dot) method.

standard RK method, the step length should be chosen to satisfy  $(\lambda_f + \lambda_g)h \in S^g$ , where  $\lambda_f \in \sigma\{\frac{\partial f}{\partial y}\}$  and  $\lambda_g \in \sigma\{\frac{\partial g}{\partial y}\}$ , in order to meet the stability requirement. But as discussed earlier, the additive method needs only to satisfy that  $\lambda_g h \in S_{\pi/2}^{f \cap g}$ . Since  $|\text{Re}\{\lambda_f\}| \ll |\text{Re}\{\lambda_g\}|$ , eventhough the stability domain of the additive method is smaller than that of its explicit part as in the conventional RK method, it allows a relatively much larger range to choose the step length  $h$ . Moreover, considering the efficiency, we can see that such an additive method is definitely superior to the pure implicit method as it does not need to solve any nonlinear equations. So the additive methods of this new type present some advantages over the existing RK methods for solving the additive system (12).

Below are two more examples with A-stability by taking  $\alpha = 1/2, \beta = 1/2, c = 1/2$  and  $\alpha = 1/2, \beta = 1/2, c = 1/4$ , respectively:

Additive RK.2.A.2					Additive RK.2.A.3					
0	0	0	0	0	0	0	0	0	0	0
1/2	0	1/2	0	1/2	0	0	1/4	-1/4	1/2	0
1	1/2	0	1/2	0	1	0	1/2	1/2	0	-1

These two methods have the same stability function  $R(z_f, z_g)$ ,

$$R(z_f, z_g) = \frac{(1 - \frac{1}{4}z_f^2) + z_g + \frac{1}{2}z_g^2}{1 - z_f + \frac{1}{4}z_f^2}$$

and their  $A(\frac{\pi}{2})$ -stability domain is shown in Fig. 1 (right). One can see that these two schemes have a relatively larger stability region (in fact, they almost coincide with that of the explicit RK method), and thus should be more preferable in solving those not too stiff problems.

The following are two  $L$ -stable schemes:

Additive RK.2.L.1						
0	0	0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}(-1 + \sqrt{2})$	$1 - \frac{\sqrt{2}}{2}$	0	$\frac{1}{2}$	0	0
1	$1 - \frac{\sqrt{2}}{2}$	$\sqrt{2} - 1$	$1 - \frac{\sqrt{2}}{2}$	0	1	0

Additive RK.2.L.2						
0	0	0	0	0	0	0
$\frac{1}{4}$	1/20	1/5	0	1/4	0	0
1	1/8	1/2	3/8	-1	2	0

The stability functions of these two schemes are, respectively,

$$R(z_f, z_g) = \frac{[1 + (\sqrt{2} - 1)z_f] + [1 + (\sqrt{2} - 1)z_f]z_g + 1/2z_g^2}{1 - (2 - \sqrt{2})z_f + (3/2 - \sqrt{2})z_f^2},$$

$$R(z_f, z_g) = \frac{(1 + 17/40z_f) + (1 + 17/40z_f)z_g + 1/2z_g^2}{1 - 23/40z_f + 3/40z_f^2}$$

and their  $A(\frac{\pi}{2})$ -stability domains are given in Fig. 2 (left for RK.2.L.1 and right for RK.2.L.2).

Next, we provide some third-order methods that should be at least of 4-stage as we pointed out earlier. But one can show that there exists a 4-stage linearly implicit RK method of order 3 if and only if  $a_{44} = 0$  (the proof is omitted here). This does not meet our requirement that  $a_{ss} \neq 0$  for an  $s$ -stage method. So we can only try to find some 5-stage third-order methods. This is possible, and some  $A$ -stable schemes are given below:

Additive RK.3.A.1										
0	0	0	0	0	0	0	0	0	0	0
$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{5}$	0	0	0	$\frac{1}{2}$	0	0	0	0
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{5}{12}$	$\frac{2}{3}$	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0
1	$\frac{2}{5}$	0	$\frac{1}{5}$	$\frac{2}{5}$	0	0	1	0	0	0
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{5}{6}$	1	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0

with its  $A(\frac{\pi}{2})$ -stability domain as shown in Fig. 3 (left).

**Additive RK.3.A.2**

0	0	0	0	0	0	0	0	0	0
$\frac{2}{9}$	$-\frac{44}{45}$	$\frac{6}{5}$	0	0	0	$\frac{2}{9}$	0	0	0
$\frac{1}{3}$	$-\frac{47}{45}$	$-\frac{71}{100}$	$\frac{6}{5}$	0	0	$\frac{71}{420}$	$\frac{23}{140}$	0	0
$\frac{5}{6}$	$\frac{27}{8}$	$-\frac{13}{4}$	$-\frac{59}{120}$	$\frac{6}{5}$	0	$-\frac{281}{336}$	$\frac{187}{112}$	0	0
1	$\frac{89}{50}$	$-\frac{486}{55}$	$\frac{89}{10}$	$-\frac{562}{275}$	$\frac{6}{5}$	$\frac{1}{10}$	0	$\frac{1}{2}$	$\frac{2}{3}$

with its stability region as shown in Fig. 3 (right).

**Additive RK.3.A.3**

0	0	0	0	0	0	0	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	0	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{5}{12}$	$\frac{2}{3}$	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	0
1	2	$-\frac{7}{2}$	$\frac{1}{3}$	2	0	0	1	0	0
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{5}{6}$	1	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$

with its  $A(\frac{\pi}{2})$ -stability domain given in Fig. 4 (left).

Below is one example of  $L$ -stable schemes,

**Additive RK.3.L.1**

0	0	0	0	0	0	0	0	0	0
$\frac{1}{4}$	$\frac{3}{20}$	$\frac{1}{10}$	0	0	0	$\frac{1}{4}$	0	0	0
$\frac{1}{2}$	$\frac{9}{10}$	$-\frac{13}{10}$	$\frac{9}{10}$	0	0	0	$\frac{1}{2}$	0	0
$\frac{3}{4}$	$\frac{17}{10}$	$-\frac{11}{4}$	$\frac{3}{2}$	$\frac{3}{10}$	0	$-\frac{1}{2}$	0	0	0
1	1	$-\frac{10}{3}$	$\frac{17}{3}$	$-\frac{10}{3}$	1	0	$\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$

with its  $A(\frac{\pi}{2})$ -stability domain given in Fig. 4 (right).

Finally, we present some fourth-order methods. A few 6-stage schemes with  $a_{44}=0$  are given in Section 7, which are applicable in some special cases. Here, we list two 7-stage  $A$ -stable methods.

**Additive RK.4.A.1**

0	0									0
$\frac{1}{3}$	$-\frac{1}{6}$	$\frac{1}{2}$								$\frac{1}{3}$
$\frac{1}{3}$	$\frac{1}{6}$	$-\frac{1}{3}$	$\frac{1}{2}$							$\frac{1}{6}$
$\frac{1}{2}$	$\frac{3}{8}$	0	0	$\frac{1}{2}$						0
$\frac{1}{2}$	$\frac{1}{8}$	0	0	$-\frac{1}{2}$	$\frac{1}{2}$					0
1	$-\frac{1}{2}$	0	0	3	-2	$\frac{1}{2}$				1
1	$\frac{1}{6}$	0	0	0	$\frac{2}{3}$	$-\frac{1}{2}$	$\frac{2}{3}$			$\frac{1}{6}$

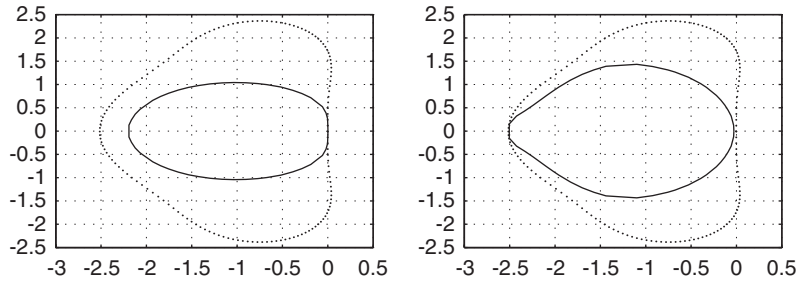


Fig. 3. Stability domain of additive (solid line) and explicit RK (dot) method.

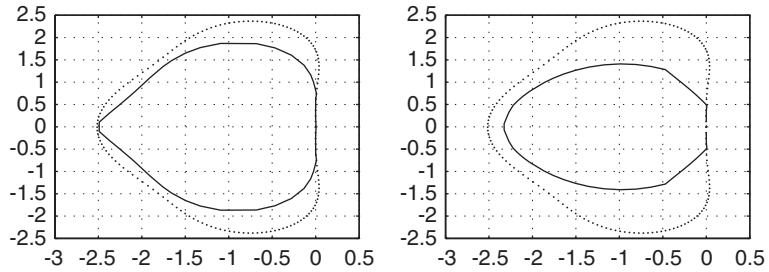


Fig. 4. Stability domain of additive (solid line) and explicit RK (dot) method.

with its  $A(\frac{\pi}{2})$ -stability domain plotted in Fig. 5 (left).

**Additive RK.4.A.2**

0	0							0						
$\frac{1}{3}$	$-\frac{1}{6}$	$\frac{1}{2}$				$\frac{1}{6}$	$\frac{1}{6}$	0					$\frac{1}{6}$	
$-\frac{1}{3}$	$-\frac{1}{6}$	$-\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	$-\frac{1}{2}$	$\frac{1}{2}$	0	0	0	
$-\frac{1}{2}$	$-\frac{1}{6}$	0	$\frac{3}{8}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	0	2	
1	$-\frac{1}{2}$	0	3	-3	1	$-\frac{1}{2}$	$\frac{2}{3}$	0	0	0	$\frac{2}{3}$	$\frac{1}{6}$		
$\frac{1}{6}$	0	0	0	0	$\frac{2}{3}$	$-\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{6}$	0	0	0	$\frac{2}{3}$	$\frac{1}{6}$	

Its  $A(\frac{\pi}{2})$ -stability domain is plotted in Fig. 5 (right).

**7. Numerical experiments**

In this section we carry out some numerical experiments to attest the stability and accuracy of the additive RK methods constructed in Section 6.

Some notations are needed. We shall use  $E(h)$  to denote the discrete  $L_2$ -norm error between the computed solution  $y_h(t)$  by an additive RK method and the exact solution  $y(t)$  to system (12), while  $r(h)$

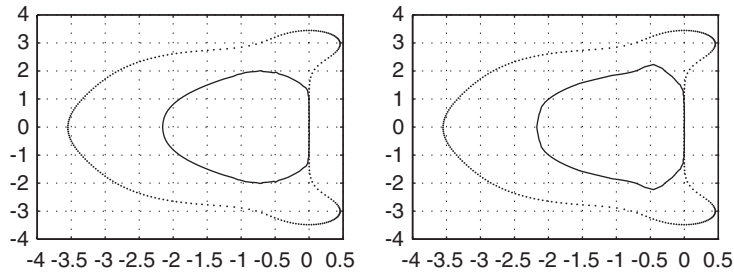


Fig. 5. Stability domain of additive (solid line) and explicit RK (dot) method.

measures the asymptotic convergence rate. The two parameters are given by (with  $h_1 > h_2$ )

$$E(h) = \sqrt{h \sum_i \|y(t_i) - y_h(t_i)\|^2}, \quad r(h) = \ln \frac{E(h_1)}{E(h_2)} / \ln \frac{h_1}{h_2}.$$

In our tests of numerical schemes, we will use the usual practice to compute the numerical solutions in the transient phase where the stiff terms  $f(t, y)$  in (12) contribute to the solutions of the considered stiff problems and outside this phase the stiff terms die out. The numerical solutions are computed with a relatively high-order explicit method using much smaller time steps for accuracy in the transient phase. In all the tests, we take the fourth-order explicit RK methods for computations in the transient phase.

**Example 1.** The first model problem is taken to be

$$y' = Ay + g(y), \quad y(0) = [1, 0, -1]^T, \tag{46}$$

where  $y = [y_1, y_2, y_3]^T$ , and  $A$  and  $g(y)$  are given by

$$A = \begin{bmatrix} -21 & 19 & -20 \\ 19 & -21 & 20 \\ 40 & -40 & -40 \end{bmatrix}, \quad g(y) = \frac{a}{1 + b\|y\|^2} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

where two parameters  $a$  and  $b$  are deliberately introduced to test different situations. Since  $\sigma(A) = \{-2, -40 + 40i, -40 - 40i\}$ , the problem is of type (12). We shall test the cases with different  $a$ 's and  $b$ 's. First, we take  $a = -10, b = 0$  and  $1$ , respectively, to check the stability of both RK.2.A.2 and RK.3.A.3, with the time range taken to be  $[0, 25]$ .

(i)  $b = 0, a = -10$ . The exact solution in this case is

$$\begin{aligned} y_1(t) &= 1/2e^{-50t}(\cos 40t + \sin 40t) + 1/2e^{-12t}, \\ y_2(t) &= -1/2e^{-50t}(\cos 40t + \sin 40t) + 1/2e^{-12t}, \\ y_3(t) &= e^{-50t}(\sin 40t - \cos 40t). \end{aligned}$$

Fig. 6 is the numerical result of RK.2.A.2 to this case with  $h = 0.3, 0.2, 0.01$ .



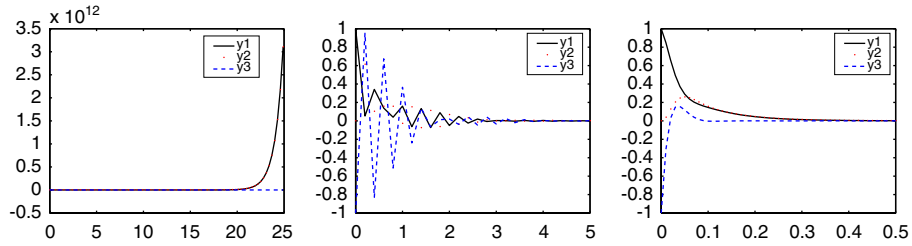


Fig. 6. Numerical results of RK.2.A.2 with  $h = 0.3, 0.2, 0.01$  in turn.

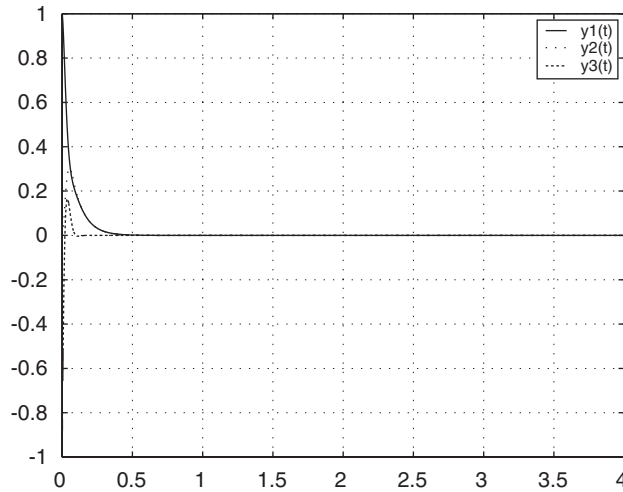


Fig. 7. True solution of Eq. (46) ( $a = -10, b = 1$ ) computed with the fourth-order RK method.

Table 1

Eigenvalues of  $\partial g / \partial y$  at different time points in the case  $a = -10$  and  $b = 1$

$t$	0	0.01	0.02	0.03	0.05	0.1	0.2	0.3	1
$\lambda_g$	-30	-19.112	-15.4069	-13.8988	-12.3178	-10.7254	-10.0694	-10.0063	-10.0000
	-30	-19.112	-15.4069	-13.8988	-12.3178	-10.7254	-10.0694	-10.0063	-10.0000
	10	-0.8878	-4.5931	-6.1012	-7.6822	-9.2746	-9.9306	-9.9937	-10.0000

We observe that if the step length  $h$  gets larger, the numerical solutions vibrate more strongly and even diverge. If only the step length is chosen to satisfy that  $ah$  (here  $a$  is the  $\lambda_g$  in (20)) falls into the  $A(\frac{\pi}{2})$ -stability domain of the method used, it is safe to have the numerical solution to be damped.

(ii)  $b = 1, a = -10$ , where the theoretical solution is computed using the 4-stage and fourth-order explicit RK method (see [8]), which is plotted in Fig. 7.

Below we list the eigenvalues of  $\partial g / \partial y$  at different time points (Table 1):

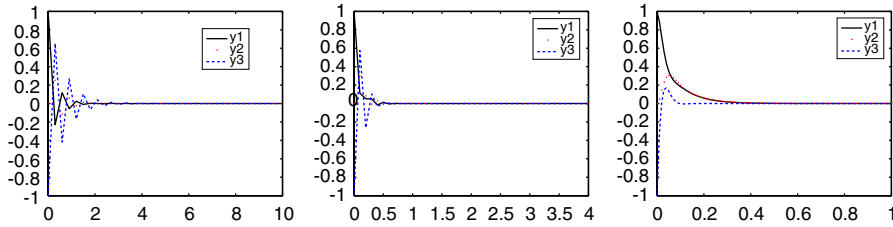


Fig. 8. Numerical results of RK.3.A.3 with  $h = 0.3, 0.1, 0.01$  in turn.

Table 2  
Eigenvalues of  $\partial g/\partial y$  at different time points in the case  $a = -2$  and  $b = 1$

$t$	0.001	0.01	0.1	0.2	0.3	0.4	0.5	0.7	1
$\lambda_g$	-5.6865	-3.9531	-2.5108	-2.2437	-2.1130	-2.0515	-2.0233	-2.0047	-2.0004
	-5.6865	-3.9531	-2.5108	-2.2437	-2.1130	-2.0515	-2.0233	-2.0047	-2.0004
	1.6865	-0.0468	-1.4892	-1.7563	-1.8870	-1.9485	-1.9767	-1.9953	-1.9996

Table 3  
Convergence order of the method RK.3.A.2

$h$	0.04	0.02	0.01	0.005	0.0025
$E(h)$	1.114e-5	1.490e-6	1.931e-7	2.459e-8	3.103e-9
$r(h)$		2.90	2.95	2.97	2.99

Table 4  
Convergence order of the method RK.3.A.3

$h$	0.04	0.02	0.01	0.005	0.0025
$E(h)$	2.171e-5	2.884e-6	3.731e-7	4.750e-8	5.993e-9
$r(h)$		2.91	2.95	2.97	2.99

Fig. 8 shows the numerical behaviors of RK.3.A.3 with  $h = 0.3, 0.1, 0.01$ , respectively. We can also see that if  $\lambda_g h$  falls into the  $A(\frac{\pi}{2})$ -stability domain of the method, it is safe to have the decay result.

(iii)  $b = 1, a = -2$ , the theoretical solution computed as in the previous case. The eigenvalues of  $\partial g/\partial y$  at some time points are listed in Table 2.

Now, with this model, the convergence order of some schemes can be computed as in the following tables. Table 3 is for RK.3.A.2.

Tables 4 and 5 are for RK.3.A.3 and RK.2.A.2, respectively.

To accurately find the exact convergence order of the fourth-order scheme RK.4.A.2, we construct the following model:

$$y' = \lambda y + \alpha y^2, \quad y(0) = 1 \tag{47}$$

Table 5  
Convergence order of the method RK.2.A.2

$h$	0.04	0.02	0.01	0.005	0.0025
$E(h)$	2.408e-5	5.990e-6	1.494e-6	3.731e-7	9.322e-8
$r(h)$		2.01	2.00	2.00	2.00

Table 6  
Convergence order of the method RK.4.A.2

$h$	0.1	0.05	0.02	0.01	0.002	0.001
$E(h)$	3.693e-8	3.896e-9	1.478e-10	1.076e-11	1.962e-14	1.247e-15
$r(h)$		3.25	3.57	3.78	3.92	3.97

Table 7  
Convergence order of the method RK.4.A.1

$h$	0.1	0.05	0.02	0.01	0.002	0.001
$E(h)$	2.10e-8	2.09e-9	7.39e-11	5.20e-12	9.19e-15	5.82e-16
$r(h)$		3.33	3.65	3.83	3.94	3.98

which has the exact solution

$$y(t) = \frac{\lambda e^{\lambda t}}{\alpha(1 - e^{\lambda t}) + \lambda}. \quad (48)$$

In this, we take  $\lambda = -10$ ,  $\alpha = -1$  and since the stiffness is brought in the consideration of the solution at the transient phase (which we may regard as  $[0,1]$ ) is taken as the true solution, which is equivalent to taking  $t_0 = 1$ , and the computation is carried out as in [1,6] and we arrive at results shown in Table 6.

The convergence of RK.4.A.1 is listed in Table 7.

From all the previous tables, one can see that the numerical experiments have clearly verified the actual orders of the corresponding additive RK methods.

## 8. Methods with zero diagonal elements

In this section, we will give a few additive methods of form (35), which allow some diagonal elements  $a_{ii}$  to be zero, but the last entry  $a_{ss} \neq 0$ . The following scheme is the first one listed in Section 6 with  $\alpha = 0$ ,  $\beta = c = \frac{1}{2}$ :

Additive RK.2.A.4					
0	0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0
1	$\frac{1}{2}$	0	$\frac{1}{2}$	0	1

This method can be viewed as a direct combination of the Crank–Nicolson method and the modified mid-point formula [10], and its stability function as defined in (21) is

$$R(z_f, z_g) = \frac{1 + 1/2z_f + (1 + 1/2z_f)z_g + 1/2z_g^2}{1 - 1/2z_f}.$$

The  $A(\frac{\pi}{2})$ -stability domain of the additive method and the stability domain of its explicit part ( $c, \mathbf{B}$ ) as a standard RK method are shown in Fig. 9 (left).

In the remaining part of this section, we present some more such additive methods with order 3 and 4.

**Additive RK.3.A.4**

0	0	0	0	0	0	0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0	$\frac{1}{2}$	0	0	0	0
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4} - a$	a	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0
1	0	c	1-c	0	0	0	d	1-d	0	0
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6} - b$	b	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0

where a,b,c,d are 4 real parameters satisfying

$$a > \frac{1}{2}, \quad b = \frac{3a - 1}{6a - 3}, \quad 2b + 4a + 8ab - \frac{5}{3} = (1 - 6b)c.$$

Taking  $a = 1, b = 2/3, c = -3, d = 1$  and  $a = 2/3, b = 1, c = -5/3, d = 1$ , we get the following two schemes:

**Additive RK.3.A.4.a**

0	0	0	0	0	0	0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0	$\frac{1}{2}$	0	0	0	0
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{3}{4}$	1	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0
1	0	-3	4	0	0	0	1	0	0	0
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0

**Additive RK.3.A.4b**

0	0	0	0	0	0	0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0	$\frac{1}{2}$	0	0	0	0
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{5}{12}$	$\frac{2}{3}$	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0
1	0	$-\frac{5}{3}$	$\frac{8}{3}$	0	0	0	1	0	0	0
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{5}{6}$	1	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0

These two methods have the same stability function and their  $A(\pi/2)$ -stability domain is shown in Fig. 9 (right).

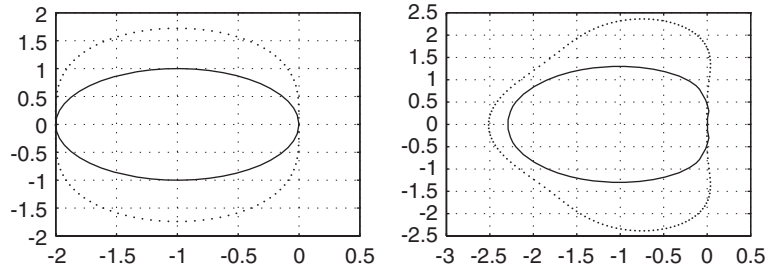


Fig. 9. Stability domain of additive (solid line) and explicit RK (dot) method.

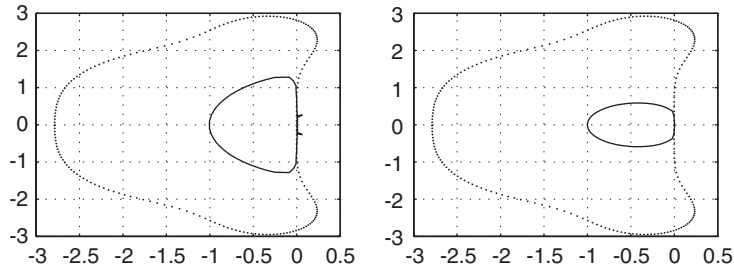


Fig. 10. Stability domain of additive (solid line) and explicit RK (dot) method.

**Additive RK.4.A.3.2**

0	0						0					
$\frac{1}{2}$	0	$\frac{1}{2}$				$\frac{1}{2}$	$\frac{1}{4}$					
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{2}$				0	-1	2			
1	0	1	0	0	$\frac{1}{3}$	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$			
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0	
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	-1	1	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0	

and its  $A(\frac{\pi}{2})$ -stability domain is plotted in Fig. 10 (left).

**Additive RK.4.A.4.2**

0	0						0					
$\frac{1}{2}$	0	$\frac{1}{2}$				$\frac{1}{2}$	$\frac{1}{4}$					
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{2}$				0	-1	2			
1	0	1	0	0	$\frac{3}{2}$	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$			
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{4}{3}$	$\frac{3}{2}$	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0	
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	-2	2	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0	

Its  $A(\frac{\pi}{2})$ -stability domain is shown in Fig. 10 (right).

Table 8  
Convergence order of the method RK.4.nA.5

$h$	0.4	0.2	0.1	0.05	0.02	0.01
$E(h)$	2.108e-5	1.244e-6	7.502e-8	4.606e-9	1.167e-10	7.268e-12
$r(h)$		4.08	4.05	4.026	4.01	4.01

Table 9  
Convergence order of the method RK.4.nA.6

$h$	0.4	0.2	0.1	0.05	0.02	0.01
$E(h)$	7.430e-5	4.351e-6	2.632e-7	1.620e-8	4.110e-10	2.562e-11
$r(h)$		4.08	4.05	4.02	4.01	4.00

### 9. Some not A-stable fourth-order methods

In this section, we present some fourth-order methods that are not A-stable, which may be used for the nonlinear ODEs or PDEs of additive form (12) that are not stiff.

**Additive RK.4.nA.5**

0	0						0				
$\frac{1}{2}$	0	$\frac{1}{2}$					$\frac{1}{2}$				
$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{2}$				$\frac{1}{4}$	$\frac{1}{4}$			
1	0	1	0	0			0	-1	2		
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{2}$		$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	
1	$\frac{1}{6}$	0	$\frac{2}{3}$	$-\frac{1}{6}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	0

**Additive RK.4.nA.6**

0	0						0				
$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$					$\frac{1}{4}$				
$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$\frac{2}{3}$				$\frac{1}{3}$	$\frac{2}{3}$			
$\frac{1}{2}$	$\frac{1}{8}$	0	$\frac{3}{8}$	0			$\frac{1}{8}$	0	$\frac{3}{8}$		
1	$\frac{1}{4}$	0	$\frac{1}{4}$	1	$\frac{1}{8}$		$\frac{1}{2}$	0	$-\frac{1}{2}$	2	
1	$\frac{1}{6}$	0	0	$\frac{2}{3}$	$-\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{6}$	0	0	$\frac{2}{3}$	$\frac{1}{6}$

The following tables verify the fourth-order convergence of additive RK.4.nA.5 (Table 8) and additive RK.4.nA.6 (Table 9) when applied to the model problem (47) in the non-stiff case with  $\lambda = -1$ .

The following are two more such schemes of fourth-order that are not A-stable:

**Additive RK.4.nA.7**

0	0							0									
$\frac{1}{6}$	0	$\frac{1}{6}$						$\frac{1}{6}$		$\frac{1}{3}$							
$\frac{1}{3}$	$\frac{1}{6}$	0	$\frac{1}{6}$					$\frac{1}{6}$	0	0							
$\frac{1}{2}$	$\frac{1}{8}$	0	$\frac{1}{6}$	$\frac{1}{6}$	0			$\frac{1}{6}$	0	0	$\frac{3}{8}$						
1	$\frac{1}{6}$	0	0	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{6}$		$\frac{1}{6}$	0	0	$-\frac{3}{2}$	$\frac{2}{3}$					
1	$\frac{1}{6}$	0	0	$\frac{2}{3}$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	0	0	0	$\frac{2}{3}$	$\frac{1}{6}$				

**Additive RK.4.nA.8**

0	0							0									
$\frac{1}{4}$	0	$\frac{1}{4}$						$\frac{1}{4}$		$\frac{2}{9}$							
$\frac{1}{3}$	$\frac{1}{6}$	0	$\frac{1}{6}$					$\frac{1}{9}$	$\frac{2}{9}$	0							
$\frac{1}{2}$	$\frac{1}{8}$	0	$\frac{1}{6}$	$\frac{1}{6}$	0			$\frac{1}{8}$	0	0	$\frac{3}{8}$						
1	0	0	$\frac{3}{4}$	$\frac{1}{4}$	0	$\frac{1}{4}$		$\frac{1}{2}$	0	$-\frac{3}{2}$	$\frac{2}{3}$						
1	$\frac{1}{6}$	0	0	$\frac{2}{3}$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	0	0	0	$\frac{2}{3}$	$\frac{1}{6}$				

**10. Some applications**

In this section we discuss briefly how to apply our new additive RK methods to solve two important PDE systems with constraints.

*10.1. Kinematic magnetic induction system*

In the kinematic geodynamo modeling, the following mean-field dynamo magnetic induction system is widely used,

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times (\beta(x)\nabla \times \mathbf{B}) + \nabla \times \mathbf{f}(x, t; \mathbf{u}, \mathbf{B}) \quad \text{in } \Omega \times (0, T), \tag{49}$$

$$\nabla \cdot \mathbf{B} = 0 \quad \text{in } \Omega \times (0, T), \tag{50}$$

with appropriate initial and boundary conditions [2,3], where  $\mathbf{B}$  is the magnetic field of the dynamo system,  $\beta(x)$  is the magnetic Reynolds number, and  $\mathbf{f}(x, t; \mathbf{u}, \mathbf{B})$  is a nonlinear vector-valued function of flow  $\mathbf{u}$  and magnetic field  $\mathbf{B}$ , which determines the key dynamo system so that the magnetic field can sustain in the corresponding physical system. Now, as an example, we apply the additive method RK.2.A.2 to the temporal semi-discretization of this system and obtain the following scheme advancing from time  $t_n$  to  $t_{n+1}$ :

$$\mathbf{B}_n^1 = \mathbf{B}_n - \frac{h}{2}\nabla \times (\beta(x)\nabla \times \mathbf{B}_n^1) + \frac{h}{2}\nabla \times \mathbf{f}(x, t_n; \mathbf{u}_n, \mathbf{B}_n), \tag{51}$$

$$\nabla \cdot \mathbf{B}_n^1 = 0, \tag{52}$$

$$\mathbf{B}_{n+1} = \mathbf{B}_n - \frac{h}{2} \nabla \times [\beta(x) \nabla \times (\mathbf{B}_n + \mathbf{B}_{n+1})] + h \nabla \times \mathbf{f}(x, t_{n+1/2}; \mathbf{u}_{n+1/2}, \mathbf{B}_n^1), \tag{53}$$

$$\nabla \cdot \mathbf{B}_{n+1} = 0, \tag{54}$$

where

$$\mathbf{B}_n \approx \mathbf{B}(t_n), \quad \mathbf{B}_n^1 \approx \mathbf{B}(t_{n+1/2}), \quad \mathbf{B}_{n+1} \approx \mathbf{B}(t_{n+1}).$$

Clearly, if explicit schemes are used, then the divergence constraints cannot be enforced for both stage values  $\mathbf{B}_n^1$  and  $\mathbf{B}_{n+1}$ . When implicit schemes as above are used, systems (51)–(52) and (53)–(54) for the two stage values  $\mathbf{B}_n^1$  and  $\mathbf{B}_{n+1}$  are well-posed [2,3].

### 10.2. Navier–Stokes system

The Navier–Stokes system is a general model for incompressible fluid dynamics:

$$\frac{\partial \mathbf{u}}{\partial t} = \nu \Delta \mathbf{u} - \nabla p - (\mathbf{u} \cdot \nabla) \mathbf{u} + f \quad \text{in } \Omega \times (0, T), \tag{55}$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega \times (0, T), \tag{56}$$

where  $\mathbf{u}$  is the fluid velocity,  $p$  is the pressure and  $f$  is an external source, and  $\nu$  is the viscosity of the fluid. Applying the additive method RK.2.L.2 to this system, we obtain the following scheme:

$$\mathbf{u}_n^1 = \mathbf{u}_n + \frac{h}{20} (\nu \Delta \mathbf{u}_n - \nabla p_n) + \frac{h}{5} (\nu \Delta \mathbf{u}_n^1 - \nabla p_n^1) + \frac{h}{4} \{ -(\mathbf{u}_n \cdot \nabla) \mathbf{u}_n + f_n \}, \tag{57}$$

$$\nabla \cdot \mathbf{u}_n^1 = 0, \tag{58}$$

$$\begin{aligned} \mathbf{u}_{n+1} = \mathbf{u}_n + \frac{h}{8} (\nu \Delta \mathbf{u}_n - \nabla p_n) + \frac{h}{2} (\nu \Delta \mathbf{u}_n^1 - \nabla p_n^1) + \frac{3h}{8} (\nu \Delta \mathbf{u}_{n+1} - \nabla p_{n+1}) \\ - h \{ -(\mathbf{u}_n \cdot \nabla) \mathbf{u}_n + f_n \} + 2h \{ -(\mathbf{u}_n^1 \cdot \nabla) \mathbf{u}_n^1 + f_{t_{n+1/4}} \}, \end{aligned} \tag{59}$$

$$\nabla \cdot \mathbf{u}_{n+1} = 0, \tag{60}$$

where

$$(\mathbf{u}_n, p_n) \approx (u(t_n), p(t_n)), \quad (\mathbf{u}_n^1, p_n^1) \approx (u(t_{n+1/4}), p(t_{n+1/4})).$$

Clearly, if explicit schemes are used, then the divergence constraints cannot be enforced for both stage values  $\mathbf{u}_n^1$  and  $\mathbf{u}_{n+1}$ . When implicit schemes as above are used, systems (57)–(58) and (59)–(60) for the two stage values  $(\mathbf{u}_n^1, p_n^1)$  and  $(\mathbf{u}_{n+1}, p_{n+1})$  are well posed.

### Acknowledgements

The authors wish to thank Prof. Geng Sun for some helpful discussions and two anonymous referees for their constructive comments which led to a great improvement of the results and the presentation of the paper.



## References

- [1] A.L. Araújo, A. Murua, J.M. Sanz-Serna, Symplectic methods based on decompositions, *SIAM. J. Numer. Anal.* 34 (1997) 1926–1947.
- [2] K. Chan, K. Zhang, J. Zou, G. Schubert, A nonlinear vacillating dynamo induced by an electrically heterogeneous mantle, *Geophys. Res. Lett.* 28 (2001) 4411–4414.
- [3] K. Chan, K. Zhang, J. Zou, G. Schubert, A nonlinear 3-D, spherical alpha-square dynamo using a finite element method, *Phys. Earth Planet. Int.* 128 (2001) 35–50.
- [4] G.J. Cooper, A. Sayfy, Semiexplicit  $A$ -stable Runge–Kutta methods, *Math. Comp.* 33 (1979) 541–556.
- [5] G.J. Cooper, A. Sayfy, Additive methods for the numerical solution of ordinary differential equations, *Math. Comp.* 35 (1980) 1159–1172.
- [6] G.J. Cooper, A. Sayfy, Additive Runge–Kutta methods for stiff ordinary differential equations, *Math. Comp.* 40 (1983) 207–218.
- [7] E. Hairer, G. Wanner, *Solving Ordinary Differential Equations, II, Stiff and Differential-Algebraic Problems*, Springer, Berlin, Heidelberg, 1991.
- [8] J.D. Lambert, *Computational Methods in Ordinary Differential Equations*, Wiley, New York, 1973.
- [9] O.J. Laurent, Structure preservation for constrained dynamics with super additive Runge–Kutta methods, *SIAM J. Sci. Comput.* 20 (1998) 416–446.
- [10] H. Liu, G. Sun, Symplectic RK methods and symplectic PRK methods with real eigenvalues, *J. Comput. Math.* 22 (5) (2004) 769–776.
- [11] A. Prothero, A. Robinson, On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations, *Math. Comput.* 28 (1974) 145–152.
- [12] X. Zhong, Additive Semi-implicit Runge–Kutta methods for computing high-speed nonequilibrium reactive flows, *J. Comp. Phys.* 128 (1996) 19–31.