(Preprint for)

J. Álvarez and J. Rojo

An improved class of generalized
Runge-Kutta methods for stiff problems.
Part I: The scalar case

Appl. Math. Comput., 130 (2002), pp. 537-560

An improved class of generalized Runge-Kutta methods for stiff problems. Part I: The scalar case

Jorge Álvarez* and Jesús Rojo†

Departamento de Matemática Aplicada a la Ingeniería, E.T.S. de Ingenieros Industriales,

Universidad de Valladolid, E-47011 Valladolid, Spain

E-mail: joralv@wmatem.eis.uva.es, jesroj@wmatem.eis.uva.es

Abstract

A new family of p-stage methods for the numerical integration of some scalar equations and systems of ODEs is proposed. These methods can be seen as a generalization of the explicit p-stage Runge-Kutta ones, while providing better order and stability results. We will show in this first part that, at the cost of losing linearity in the formulas, it is possible to obtain explicit A-stable and L-stable methods for the numerical integration of scalar autonomous ODEs. Scalar autonomous ODEs are of very little interest in current applications. However, be begin studying this kind of problems because most of the work can be easily extended to a more general situation. In fact, we will show in a second part (entitled 'The separated system case'), that it is possible to generalize our methods so that they can be applied to some non-autonomous scalar ODEs and systems. We will obtain linearly implicit L-stable methods which do not require Jacobian evaluations. In both parts, some numerical examples are discussed in order to show the good performance of the new schemes.

Keywords: Generalized RK Methods, Stiff ODEs, Linear Stability, Numerical Experiments.

AMS subject classification: 65L05, 65L07, 65L20.

---

*This work was partially supported by Consejería de Educación y Cultura de la Junta de Castilla y León and Unión Europea (Fondo Social Europeo) under proyect VA19/00B and by Junta de Castilla y León under proyect VA11/99.

†This work was partially supported by Consejería de Educación y Cultura de la Junta de Castilla y León and Unión Europea (Fondo Social Europeo) under proyect VA19/00B and by Spanish DGES under project PB98-0359.

1 Introduction.

During the last decades there has been a considerable amount of research on methods for numerical integration of stiff systems of ODEs, usually looking for better stability properties. Nearly all such methods are implicit in character.

The most widely used algorithms are those based on linear multistep formulas like the BDF methods (see e.g. [1]), because they are very efficient for general stiff problems. However, from a result of Dahlquist it is known that no linear multistep method of order greater than two can be A-stable [2], and so these formulas are not suited to some stiff problems (for example, those with Jacobians whose eigenvalues have large imaginary parts).

Implicit Runge-Kutta formulas [3, 4] have been widely used because of their excellent stability properties (such as A-stability, L-stability and B-stability), but the need for solving nonlinear algebraic equations at each step makes these formulas generally too costly when considering some huge systems of ODEs.

To reduce the amount of computational effort required to solve the nonlinear equations (by Newton-type iterations) when integrating with a fully implicit Runge-Kutta method, some classes of Runge-Kutta formulas have been developed. With the class of diagonally implicit Runge-Kutta (DIRK) methods (also called semi-implicit or semi-explicit Runge-Kutta methods) the algebraic cost of the LU factorization is reduced. By considering the class of singly diagonally implicit Runge-Kutta (SDIRK) methods and also the class of singly implicit Runge-Kutta (SIRK) methods, it is possible to reduce even more the algebraic cost of the LU factorization (see e.g. [5, 6, 3] for more details).

Many other attempts have been made in order to reduce the computation cost per step by considering linearly implicit methods, in this way eliminating the need for solving nonlinear systems which, as pointed before, usually are solved by Newton-type iteration (and hence require additional function evaluations for every iteration at every step). Such formulas have the computational advantage that it is necessary to solve only linear systems of algebraic equations at each step.

Among the many different RK-like methods of this type we have the Rosenbrock methods [7] and the ROW-methods (also called Rosenbrock-Wanner methods and modified Rosenbrock methods) [8, 9, 10]. These formulas, however, require the exact Jacobian at every step. Therefore the computations are costly when the Jacobian matrix is expensive to evaluate. For this reason, extensions of Rosenbrock methods have been considered in which the exact Jacobian is fixed for some number of steps so that the computation cost is reduced (see e.g. [11, 12, 13, 14]). Moreover, Rosenbrock-type methods in which the exact Jacobian is no longer needed have been considered. The so called W-methods [15], the MROW-methods [16] and the generalized Runge-Kutta methods [17] (see also [18]) fall into this class. For an excellent survey of some of these methods the reader is referred to [5].

In our recent papers [19, 20], examples are shown of explicit and linearly implicit two-stage methods of order three for the numerical integration of scalar autonomous ODEs, which do not require Jacobian evaluations, some of them being as well A-stable and L-stable. Some comparisons with Runge-Kutta methods as well as numerical experiments are also reported. In [21] we describe how to construct from a given function R a one-parameter family of explicit (or linearly implicit) two-stage methods, having R as the associated stability function, and illustrate this fact by obtaining a two-stage third order formula whose associated stability function is given by e^z.

Our first aim in the present paper is to introduce the general form of the new explicit p-stage methods for the numerical integration of scalar autonomous ODEs. We begin studying these methods for scalar problems because, as we will see in a second part, most of the work can be easily extended in order to integrate separated systems of ODEs. These methods can be seen as a generalization of the explicit Runge-Kutta methods providing better order and stability results with the same number of stages. In fact, from Butcher's theory we know that an p-stage explicit Runge-Kutta method cannot have order greater than p. Moreover, the stability function of such methods is a polynomial, and so none of them is A-stable. We will show that it is possible to obtain A-stable explicit formulas for scalar autonomous problems of order three and five, with only two and three stages respectively, from our class of methods. By a further generalization of our schemes, we will show in the second part that it is possible to obtain A-stable linearly implicit formulas for some non-autonomous scalar ODEs and systems, which do not require Jacobian evaluations. For example, in [22] we present a two stage third order method for separated systems of ODEs being L-stable, as well as preliminary results using our method to integrate systems that arise when solving some nonlinear parabolic PDEs by the method of lines approach.

Finally we illustrate the efficiency of those schemes by carrying out some numerical experiments.

2 The new family of methods.

In this first part, we will restrict our attention to the scalar autonomous initial value problem

$$y'(x) = f(y(x)) , \quad y(x_0) = y_0 . \tag{1}$$

For this problem, let us consider the family of explicit p-stage methods defined by

$$y_{n+1} = y_n + hF_{p+1}(k_1, k_2, \ldots, k_p) , \tag{2}$$

where the stages are given by

$$\begin{aligned} k_1 &= f(y_n) \\ k_2 &= f(y_n + hF_2(k_1)) \\ k_3 &= f(y_n + hF_3(k_1, k_2)) \\ &\vdots \\ k_p &= f(y_n + hF_p(k_1, k_2, \ldots, k_{p-1})) , \end{aligned} \tag{3}$$

and for each i with $2 \le i \le p+1$, F_i is any homogeneous function of degree one, that is

$$F_i(\alpha\, x_1, \alpha\, x_2, \ldots, \alpha\, x_{i-1}) = \alpha\, F_i(x_1, x_2, \ldots, x_{i-1}) , \tag{4}$$

holds for each $\alpha \in \mathbb{R}$ and $(x_1, x_2, \ldots, x_{i-1})$ in a subset of $\mathbb{R}^{i-1}$.

The family of methods we have just defined, may be shown to be a generalization of the explicit p-stage Runge-Kutta methods for problem (1). In fact, taking $F_i(x_1, x_2, \ldots, x_{i-1}) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ii-1}x_{i-1}$ $(2 \le i \le p)$ in (3) and $F_{p+1}(x_1, x_2, \ldots, x_p) = b_1x_1 + b_2x_2 + \cdots + b_px_p$ in (2), we get all the explicit p-stage Runge-Kutta formulas as a subfamily of our class.

Moreover, our methods can be seen as generalized explicit Runge-Kutta methods whose coefficients depend on the stages and are no longer constants. To show this it is enough to note that from the fact that F_i is an homogeneous function of degree one, by using Euler's theorem for homogeneous functions, we obtain

$$F_i(k_1, k_2, \ldots, k_{i-1}) = \sum_{j=1}^{i-1} \frac{\partial F_i}{\partial x_j}(k_1, k_2, \ldots, k_{i-1})\, k_j , \quad 2 \le i \le p+1 ,$$

from which it is clearly enough to take the Butcher array

$$\begin{array}{c|ccccc}
 & a_{21} & & & & \\
 & a_{31} & a_{32} & & & \\
 & \vdots & \vdots & \ddots & & \\
 & a_{p1} & a_{p2} & \cdots & a_{pp-1} & \\
\hline
 & b_1 & b_2 & \cdots & b_{p-1} & b_p
\end{array}$$

where now the parameters a_{ij} and b_i are given in terms of the stages k_i and the functions F_i through the relations

$$\begin{aligned}
a_{ij} &= \frac{\partial F_i}{\partial x_j}(k_1, k_2, \ldots, k_{i-1}), \quad 1 \leq j < i \leq p, \\
b_i &= \frac{\partial F_{p+1}}{\partial x_i}(k_1, k_2, \ldots, k_p), \quad 1 \leq i \leq p.
\end{aligned}$$

Our new p-stage methods are difficult to study because of the nonlinearities that can arise from the homogeneous functions in (4). For reasons that will be clear later when studying the consistency, order and linear stability properties of the schemes, we will give a better formulation of our methods in order to simplify our analysis.

Note at this point that the quantities given by

$$c_i = F_i(1, 1, \ldots, 1), \quad 2 \leq i \leq p+1, \tag{5}$$

play a similar role to those associated with the classical Runge-Kutta methods. In fact, we have that the stages k_i can be seen as approximations to $y'(x_n + c_i h)$. Now we give the promised better formulation of the methods.

3 A useful formulation.

We introduce the terms

$$s_i = \frac{k_i - k_1}{k_1} = \frac{k_i}{k_1} - 1, \quad 2 \leq i \leq p, \tag{6}$$

where the stages k_i are given by (3). From considerations that will be clear later, we take $s_i = 0$ when $k_1 = 0$ in (6).

It is a simple task to show that $s_i = O(h)$, and we will exploit this property in order to simplify our study. Moreover, the terms s_i can be seen as approximations to $c_i h f_y(y_n)$. In fact $s_i = c_i h f_y(y_n) + O(h^2)$ (here we assume that f has a sufficient number of bounded derivatives), and so we can obtain approximations to the Jacobian $f_y(y_n)$ by taking $s_i/(c_i h)$ (when $c_i \neq 0$).

It is easy to show recursively that the stages k_i (with $2 \le i \le p$) can be obtained from k_1 and s_j with $2 \le j \le i-1$ (see e.g. (8) below). Therefore, in terms of k_1 and s_i, any method of the preceding family takes the form

$$y_{n+1} = y_n + hk_1 G_{p+1}(s_2, s_3, \ldots, s_p)\,, \tag{7}$$

where the s_i are given by (6) in terms of the stages

$$\begin{aligned} k_1 &= f(y_n) \\ k_2 &= f(y_n + hk_1 G_2) \\ k_3 &= f(y_n + hk_1 G_3(s_2)) \\ &\vdots \\ k_p &= f(y_n + hk_1 G_p(s_2, s_3, \ldots, s_{p-1}))\,, \end{aligned} \tag{8}$$

and the functions G_i can be obtained from the homogeneous functions F_i through the relations

$$\begin{aligned} G_i(s_2, s_3, \ldots, s_{i-1}) &= \frac{1}{k_1} F_i(k_1, k_2, \ldots, k_{i-1}) = F_i\left(1, \frac{k_2}{k_1}, \ldots, \frac{k_{i-1}}{k_1}\right) \\ &= F_i(1, 1+s_2, \ldots, 1+s_{i-1})\,, \qquad 2 \le i \le p+1\,, \end{aligned} \tag{9}$$

Note that when $i = 2$ we have that $G_2 = (1/k_1)F_2(k_1) = F_2(1) = c_2$ holds from (5). From the above relations, we can write any method (2) in the form given in (7).

Reciprocally, it is also possible to obtain from a given method in terms of k_1 and s_i the associated expression in terms of the k_i by using the relations

$$F_i(k_1, \ldots, k_{i-1}) = k_1 G_i(s_2, \ldots, s_{i-1}) = k_1 G_i\left(\frac{k_2}{k_1} - 1, \ldots, \frac{k_{i-1}}{k_1} - 1\right)\,, \qquad 2 \le i \le p+1\,.$$

For $i = 2$ we have $F_2(k_1) = k_1 G_2$.

4 Consistency and order of the methods.

In what follows we will assume that $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz in $\mathbb{R}$, i. e. there exists a Lipschitz constant L such that

$$|f(y) - f(y^*)| \le L|y - y^*|\,,$$

for every $y, y^* \in \mathbb{R}$. With the previous assumptions it is well know that for any $y_0 \in \mathbb{R}$, there exists a unique solution $y(x)$ of problem (1) (throughout any interval $[x_0, b]$), where $y(x)$ is continuous and differentiable.

We will investigate the consistency of any p-stage method given by (7). When doing so, we will assume the existence (and continuity) of $y'(x)$ in $[x_0, b]$, but not necessarily that of higher derivatives.

Using Henrici's notation for one step methods, our methods can be expressed as $y_{n+1} = y_n + h\,\Phi(x_n, y_n, h)$, with the increment function Φ (not depending explicitly on x because (1) is autonomous) given by

$$\Phi(x, y, h) = k_1 G_{p+1}(s_2, s_3, \ldots, s_p)\,,$$

through (6–9) (with y in place of y_n).

As is usual with other one-step methods, we define the local truncation error $T(x, h)$ of any formula of our family to be

$$T(x, h) = y(x + h) - y(x) - hk_1 G_{p+1}(s_2, s_3, \ldots, s_p)\,, \quad x \in [x_0, b]\,, \tag{10}$$

where $h > 0$, the stages k_i are given by (8) with the exact solution of (1) $y(x)$ in place of y_n, the terms s_i by (6) and the functions G_i through (9).

Definition 1 *Method (7) is said to be consistent (with (1) satisfying our previous assumptions) if, in the limit as $h \to 0$ we have that $T(x, h)/h \to 0$ uniformly for $x \in [x_0, b]$.*

It is known that consistency is a necessary condition for convergence, and therefore we want to establish the condition that method (7) must satisfy if we require it to be consistent. In order to do so, we need the following lemma:

Lemma 1 *Suppose that for each j with $2 \leq j \leq p + 1$, the functions G_j in (9) are continuous in a neighbourhood W_{j-2} of the point $(0, 0, \ldots, 0) \in I\!R^{j-2}$. Let s_i $(2 \leq i \leq p)$ be given through (6) in terms of the stages k_i in (8) with y in place of y_n. Then for every $\epsilon > 0$ there exist $h_1 > 0$ such that*

$$\begin{aligned} |s_i| &\leq hL(|c_i| + \epsilon)\,, \quad 2 \leq i \leq p\,, \\ |G_j(s_2, s_3, \ldots, s_{j-1})| &\leq |c_j| + \epsilon\,, \quad 2 \leq j \leq p + 1\,, \end{aligned}$$

hold for every $0 < h \leq h_1$ and $y \in I\!R$.

Recall that as we have pointed out before, $c_i = G_i(0, 0, \ldots, 0)$. The proof follows from the Lipschitz character of f, and from the fact that the functions G_j are continuous in a neighbourhood W_{j-2} of the point $(0, 0, \ldots, 0) \in \mathbb{R}^{j-2}$, by a recursive procedure.

From the above lema it is clear why $s_i = O(h)$. Now we have

Theorem 1 *Method (7) is consistent with (1) (under the above assumptions) iff*

$$G_{p+1}(0, 0, \ldots, 0) = c_{p+1} = 1\,. \tag{11}$$

Proof. For any given $x \in [x_0, b]$ and $h > 0$ we get from the mean value theorem that

$$y(x+h) - y(x) = h\, y'(\alpha_x)\,,$$

where $\alpha_x \in (x, x+h)$. It then follows from (10) that

$$\frac{T(x,h)}{h} = y'(\alpha_x) - k_1 G_{p+1}(s_2, s_3, \ldots, s_p)\,,$$

Now in the limit as $h \to 0$ we have that $y'(\alpha_x) \to y'(x)$ uniformly for $x \in [x_0, b]$. From the previous lemma we easily get that as $h \to 0$, $s_i \to 0$ for each i (with $2 \leq i \leq p$). Finally, from the continuity of the function G_{p+1}, and taking into account that $y'(x) = f(y(x))$ and $k_1 = f(y(x))$, we get

$$\lim_{h\to 0} \frac{T(x,h)}{h} = (1 - G_{p+1}(0, 0, \ldots, 0)) f(y(x))\,,$$

from which we obtain the consistency condition (11). $\diamondsuit$

Note at this point that using Henrici's notation for one step methods, the consistency condition reads $\Phi(y, 0) = f(y)$. Obviously this consistency condition takes the form (11).

Now we define the consistency of order q in the usual way, that is,

Definition 2 *Method (7) is said to be consistent (with the differential equation (1)) of order q, if q is the largest integer such that there exists $N \geq 0$ and $h_0 > 0$ with $\sup_{x_0 \leq x \leq b} |T(x,h)| \leq N\, h^{q+1}$ for all $h \in (0, h_0]$.*

If all the partial derivatives of $f(y)$ up to order q exist (and are continuous), then consistency follows from the consistency of order $q \geq 1$.

It can be seen that any consistent method of order q is convergent of order q, but the proof is more complicated that for other one step methods like the Runge-Kutta ones and so we give not the proof here. The main difficulty arises when proving that the increment function $\Phi(y,h) = k_1 G_{p+1}(s_2, s_3, \ldots, s_p)$ satisfies a Lipschitz condition in y for h small enough. Even though this property is nearly automatic for most of the one-step methods from the Lipschitz condition that satisfies function f, for our methods it is not as easy because of their nonlinear structure.

5 Methods of polynomial type.

Now, for every fixed p, we restrict our attention to the family of methods given by (2–3), where now all the F_i $(2 \leq i \leq p+1)$ are assumed to be homogeneous functions (of degree one) of the special form

$$F_i(x_1, x_2, \ldots, x_{i-1}) = \sum_{j_2+j_3+\ldots+j_{i-1}=0}^{r_i} A_{j_2 j_3 \cdots j_{i-1}}\, x_1 \left(\frac{x_2}{x_1}\right)^{j_2} \left(\frac{x_3}{x_1}\right)^{j_3} \cdots \left(\frac{x_{i-1}}{x_1}\right)^{j_{i-1}}, \tag{12}$$

with all r_i being nonnegative integers.

The above family of methods, still contains all explicit p-stage Runge-Kutta methods. In fact, taking $r_i = 1$ in (12) we get from (2) and (3) all explicit p-stage Runge-Kutta methods.

In what follows, we will restrict our attention to the above family of methods because in terms of s_i the associated functions G_i are of polynomial type. This greatly simplifies the study of the order conditions. Moreover, order conditions for the general methods can be easily obtained from the order conditions for methods of polynomial type, by considering the Taylor expansion of the functions G_i in terms of the s_i.

From now on, we will assume that all quantities $c_i = F_i(1, 1, \ldots, 1)$ in (12) are different from zero. Even though we lose a bit of generality with this assumption, we obtain many advantages that will be clear later. In fact, it can be seen that for a given number of stages p (at least for $p = 2, 3, 4$), the highest order is attained only when all $c_i \neq 0$, and so, all interesting methods (from an order point of view) are considered.

With the above assumption we now change our definition of the s_i in (6), hoping that this will not confuse the reader. We define

$$s_i = \frac{k_i - k_1}{c_i\, k_1}, \quad 2 \leq i \leq p\,, \tag{13}$$

with the stages k_i given by (3). Now it is easy to see that $s_i = h f_y(y_n) + O(h^2)$, and this will be useful later when looking for methods with good linear stability properties, since it simplifies the study of the order conditions.

In terms of k_1 and s_i, the preceding method takes the form (7–8), but now the functions G_i are given in terms of the new s_i and the functions F_i in (12) through the relations

$$\begin{aligned} G_i(s_2, s_3, \ldots, s_{i-1}) &= \frac{1}{k_1} F_i(k_1, k_2, \ldots, k_{i-1}) = F_i\left(1, \frac{k_2}{k_1}, \ldots, \frac{k_{i-1}}{k_1}\right) \\ &= F_i(1, 1 + c_2\, s_2, \ldots, 1 + c_{i-1}\, s_{i-1})\,, \quad 2 \leq i \leq p+1\,. \end{aligned} \tag{14}$$

It can be seen that the functions G_i take the form

$$G_i(z_2, z_3, \ldots, z_{i-1}) = c_i \left(1 + \sum_{j_2+j_3+\ldots+j_{i-1}=1}^{r_i} a_{j_2 j_3 \cdots j_{i-1}} \, z_2^{j_2} z_3^{j_3} \cdots z_{i-1}^{j_{i-1}} \right) . \tag{15}$$

Note that for this special class of methods, as we have pointed out before, all G_i are polynomial functions of the s_i. Therefore, in what follows we will refer to this formulas as methods of polynomial type.

Now we study the attainable order, in terms of the number of stages, of the methods given by (7), (8), (13) and (15). From now on, we will suppose that f is smooth enough in order that all the derivatives which occur when considering the Taylor expansion of the local truncation error make sense.

6 Two-stage methods of polynomial type. Attainable order.

To show that we can obtain explicit two-stage methods of polynomial type from (7–8) (together with (13) and (15)) of order three, it is enough to consider the Taylor expansion of the associated local truncation error. Note that when doing so, only the parameters c_2, c_3 and a_i with $1 \leq i \leq 2$ will appear in the order conditions, that is, it suffices to take in (15) $r_3 = 2$ (the other parameters can be arbitrarily chosen). This easily follows from the fact that $s_2 = O(h)$, and therefore $s_2^k = O(h^k)$ for any given $k \in \mathbb{N}$. We obtain the following order conditions

$$\begin{aligned} c_3 &= 1 , \\ c_3 a_1 &= 1/2 , \\ c_3 c_2 a_1 &= 1/3 , \\ c_3 a_2 &= 1/6 , \end{aligned}$$

from which the general form of a third order two-stage method of polynomial type is given by

$$y_{n+1} = y_n + h k_1 G_3(s_2) ,$$

where k_1, k_2 and s_2 are given by

$$k_1 = f(y_n) , \quad k_2 = f\left(y_n + \frac{2}{3} h k_1 \right) , \quad s_2 = \frac{3(k_2 - k_1)}{2k_1} , \tag{16}$$

and G_3 takes the form

$$G_3(s_2) = 1 + \frac{1}{2} s_2 + \frac{1}{6} s_2^2 + \sum_{i=3}^{r_3} a_i s_2^i , \tag{17}$$

that is, it is enough to take $c_3 = 1$ (consistency condition), $c_2 = 2/3$, $a_1 = 1/2$ and $a_2 = 1/6$. Parameters a_i with $i \geq 3$ can be arbitrarily chosen.

The additional order conditions (now with $r_3 = 3$ in (15)) for order four are given by

$$\begin{aligned} c_3 c_2^2 a_1 &= 1/4\,, && (18)\\ c_3 c_2 a_2 &= 1/6\,, && (19)\\ c_3 a_3 &= 1/24\,, && (20) \end{aligned}$$

and it is easy to check that no two-stage method of polynomial type has order greater than three (conditions (18) and (19) cannot be satisfied). However, condition (20) can be satisfied by taking $a_3 = 1/24$, obtaining in this way third order methods that minimize the principal part of the local truncation error. Note at this point that the other terms of the principal part of the local truncation error (those associated with conditions (18) and (19)) are the same for any method of order three.

7 Three-stage methods of polynomial type. Attainable order.

When considering three-stage methods of polynomial type from (7–8) together with (13) and (15), it is possible to get a family of fifth order formulas, depending on many free parameters. As in the two-stage case, only some of this parameters will appear in the order conditions, due to the fact that both s_2 and s_3 are $O(h)$. However, the resulting order conditions are still too cumbersome; hence we define a new term $\tilde{s}_3 = s_3 - s_2$ to make our study easier. Note that from our previous considerations it is clear that $\tilde{s}_3 = O(h^2)$. Now we can describe any three-stage method of polynomial type by

$$y_{n+1} = y_n + hk_1 \tilde{G}_4(s_2, \tilde{s}_3)\,,$$

with $\tilde{s}_3 = s_3 - s_2$, and where s_2 and s_3 are given through (13) in terms of the stages k_i $(1 \leq i \leq 3)$

$$k_1 = f(y_n)\,, \quad k_2 = f(y_n + hk_1 G_2)\,, \quad k_3 = f(y_n + hk_1 G_3(s_2))\,.$$

Now functions G_2, G_3 and $\tilde{G}_4$ are given by

$$\begin{aligned} G_2 &= c_2\,,\\ G_3(s_2) &= c_3 \left(1 + \sum_{i=1}^{r_3} a_i s_2^i\right)\,,\\ \tilde{G}_4(s_2, \tilde{s}_3) &= c_4 \left(1 + \sum_{i+2j=1}^{\tilde{r}_4} a_{ij} s_2^i \tilde{s}_3^j\right)\,, && (21) \end{aligned}$$

and it is easily seen when looking for formulas of order five, that it is enough to consider the parameters in (21) with $r_3 = 3$ and $\tilde{r}_4 = 4$. This means that the order conditions (for order five) are completely determined in terms of the parameters c_i $(2 \leq i \leq 4)$, a_i $(1 \leq i \leq 3)$, and a_{ij} $(1 \leq i + 2j \leq 4)$. This follows from the fact that $s_2 = O(h)$ and $\tilde{s}_3 = O(h^2)$.

Now, any three-stage methods of polynomial type must satisfy the following order conditions in order to be of order five

$$c_4 = 1\,, \qquad (22)$$

$$c_4 a_{10} = 1/2\,, \qquad (23)$$

$$c_4\left(c_2(a_{10} - a_{01}) + c_3 a_{01}\right) = 1/3\,, \qquad (24)$$

$$c_4\left(a_{20} + a_1 a_{01}\right) = 1/6\,, \qquad (25)$$

$$c_4\left(c_2^2(a_{10} - a_{01}) + c_3^2 a_{01}\right) = 1/4\,, \qquad (26)$$

$$c_4\left(c_2(2a_{20} - a_{11} + a_1 a_{01}) + c_3(a_{11} + 2a_1 a_{01})\right) = 1/3\,, \qquad (27)$$

$$c_4\left(a_{30} + a_1 a_{11} + a_2 a_{01}\right) = 1/24\,, \qquad (28)$$

$$c_4\left(c_2^3(a_{10} - a_{01}) + c_3^3 a_{01}\right) = 1/5\,, \qquad (29)$$

$$c_4\left(c_2^2(2a_{20} - a_{11} + a_1 a_{01}) + c_3^2(a_{11} + 3a_1 a_{01})\right) = 7/20\,, \qquad (30)$$

$$c_4\left(c_2^2(a_{20} - a_{11} + a_{02}) + c_2 c_3(a_{11} - 2a_{02} + 2a_1 a_{01}) + c_3^2 a_{02}\right) = 2/15\,, \qquad (31)$$

$$\begin{aligned} c_4\big(c_2(3a_{30} - a_{21} + 2a_1(a_{11} - a_{02}) + 2a_2 a_{01}) \\ + c_3(a_{21} + 2a_1(a_{11} + a_{02}) + (a_1^2 + 2a_2)a_{01})\big) &= 11/60\,, \end{aligned} \qquad (32)$$

$$c_4\left(a_{40} + a_1 a_{21} + a_1^2 a_{02} + a_2 a_{11} + a_3 a_{01}\right) = 1/120\,. \qquad (33)$$

It is easy to check that the above system has many solutions. In fact, we get two doubly infinite families of solutions, taking a_2 and a_3 as free parameters. The remaining coefficients can be computed as follows:

Step 1 Obviously, we have c_4 from (22). Coefficients a_{10}, a_{01}, c_2 and c_3, can be chosen such that (23), (24), (26) and (29) are satisfied. We get a pair of solutions in this way.

Step 2 Now, for each solution of the previous step, we can obtain a_1, a_{20}, a_{11} and a_{02}, by solving the linear system given by (25), (27), (30) and (31).

Step 3 Finally, we obtain a_{40}, a_{30} and a_{21} by solving the remaining linear system (28), (32) and (33) (in terms of the parameters a_2 and a_3).

The result is

$$c_2 = \frac{6 \mp \sqrt{6}}{10}\,, \quad c_3 = \frac{6 \pm \sqrt{6}}{10}\,, \quad c_4 = 1\,, \quad a_1 = \frac{-3 \pm 2\sqrt{6}}{5}\,, \quad a_{10} = \frac{1}{2}\,, \quad a_{01} = \frac{9 \pm \sqrt{6}}{36}\,,$$
$$a_{20} = \frac{3 \mp \sqrt{6}}{12}\,, \quad a_{11} = \frac{3 \pm 2\sqrt{6}}{72}\,, \quad a_{02} = \frac{1 \pm 4\sqrt{6}}{72}\,, \quad a_{30} = \frac{-9 \mp \sqrt{6}}{36}\, a_2\,,$$
$$a_{21} = \frac{-(3 + 11a_2) \pm (1 - 4a_2)\sqrt{6}}{24}\,, \quad a_{40} = \frac{-3\,(3 - 10a_2 + 30a_3) \pm (3 + 20a_2 - 10a_3)\sqrt{6}}{360}\,. \tag{34}$$

The other parameters may be arbitrarily chosen.

For order six we must add to (22–33) the following conditions (now it is enough to take $r_3 = 4$ and $\tilde{r}_4 = 5$ in (21))

$$c_4 \left(c_2^4 (a_{10} - a_{01}) + c_3^4 a_{01} \right) = 1/6\,, \tag{35}$$
$$c_4 \left(c_2^3 (2a_{20} - a_{11} + a_1 a_{01}) + c_3^3 (a_{11} + 4a_1 a_{01}) \right) = 11/30\,, \tag{36}$$
$$\begin{aligned} c_4 \Big(& 2c_2^3 (a_{20} - a_{11} + a_{02}) + c_2^2 c_3 (a_{11} - 2a_{02} + 2a_1 a_{01}) \\ & + c_2 c_3^2 (a_{11} - 2a_{02} + 3a_1 a_{01}) + 2c_3^3 a_{02} \Big) = 1/4\,, \end{aligned} \tag{37}$$
$$\begin{aligned} c_4 \Big(& c_2^2 (3a_{30} - a_{21} + 2a_1 (a_{11} - a_{02}) + 2a_2 a_{01}) \\ & + c_3^2 (a_{21} + a_1 (3a_{11} + 2a_{02}) + 3(a_1^2 + a_2) a_{01}) \Big) = 4/15\,, \end{aligned} \tag{38}$$
$$\begin{aligned} c_4 \Big(& c_2^2 (3a_{30} - 2a_{21} + a_{12} + a_1 (a_{11} - 2a_{02}) + a_2 a_{01}) \\ & + 2c_2 c_3 (a_{21} - a_{12} + a_1 (2a_{11} - a_{02}) + (a_1^2 + 2a_2) a_{01}) + c_3^2 (a_{12} + 4a_1 a_{02}) \Big) = 17/90\,, \end{aligned} \tag{39}$$
$$\begin{aligned} c_4 \Big(& c_2 (4a_{40} - a_{31} + a_1 (3a_{21} - 2a_{12}) + 2(a_1^2 - a_2) a_{02} + 3a_2 a_{11} + 3a_3 a_{01}) \\ & + c_3 (a_{31} + 2(a_1 (a_{21} + a_{12}) + (2a_1^2 + a_2) a_{02} + (a_1 a_2 + a_3) a_{01}) + (a_1^2 + 2a_2) a_{11}) \Big) = 13/180\,, \end{aligned} \tag{40}$$
$$c_4 \left(a_{50} + a_1 (a_{31} + a_1 a_{12} + 2a_2 a_{02}) + a_2 a_{21} + a_3 a_{11} + a_4 a_{01} \right) = 1/720\,. \tag{41}$$

Now from (34) and conditions (35–41), it is not difficult to see that order six cannot be attained with only three stages. In fact, conditions (35–37) cannot be satisfied, and the coefficients of the associated terms in the principal part of the local truncation error are the same for any method of order five. However, conditions (38–41) can be satisfied by taking

$$a_2 = \frac{-519 \pm 226\sqrt{6}}{300}\,, \quad a_{12} = \frac{-103 \pm 42\sqrt{6}}{864}\,, \quad a_{31} = \frac{-3\,(119 + 660a_3) \pm 5\,(25 - 144a_3)\sqrt{6}}{4320}\,,$$
$$a_{50} = \frac{(22597 + 2400a_3 - 7200a_4) \mp 8\,(1143 - 200a_3 + 100a_4)\sqrt{6}}{28800}\,, \tag{42}$$

obtaining in this way three-stage methods of order five that minimize the principal part of the local truncation error. Note that now a_{30}, a_{21} and a_{40} are given in (34) by

$$a_{30} = \frac{221 \mp 101\sqrt{6}}{720}\,, \quad a_{21} = \frac{-123 \mp 22\sqrt{6}}{1440}\,, \quad a_{40} = \frac{(59 - 180a_3) \mp 2\,(9 + 10a_3)\sqrt{6}}{720}\,, \tag{43}$$

8 Methods of rational type.

In the last three sections we have only considered methods of polynomial type. Now we will consider methods of rational type, that is, methods given by (2–3), where now all the F_i with $2 \leq i \leq p+1$ are supposed to be homogeneous functions (of degree one) of rational type. More precisely, we will consider functions F_i given in terms of the quotient of two homogeneous polynomials $\tilde{N}_i(x_1, x_2, \ldots, x_{i-1})$ and $\tilde{D}_i(x_1, x_2, \ldots, x_{i-1})$ with degrees $r_i + 1$ and r_i respectively, for some nonnegative integer r_i, that is

$$F_i(x_1, x_2, \ldots, x_{i-1}) = \frac{\tilde{N}_i(x_1, x_2, \ldots, x_{i-1})}{\tilde{D}_i(x_1, x_2, \ldots, x_{i-1})} = \frac{\displaystyle\sum_{j_1+j_2+\ldots+j_{i-1}=r_i+1} N_{j_1 j_2 \cdots j_{i-1}} x_1^{j_1} x_2^{j_2} \cdots x_{i-1}^{j_{i-1}}}{\displaystyle\sum_{j_1+j_2+\ldots+j_{i-1}=r_i} D_{j_1 j_2 \cdots j_{i-1}} x_1^{j_1} x_2^{j_2} \cdots x_{i-1}^{j_{i-1}}} . \tag{44}$$

It is not difficult to see that the new family of methods contains all the previous methods of polynomial type as a subfamily. Now, assuming as before that all quantities c_i are different from zero, we get from (13) and relation (14) that, in terms of k_1 and s_i, any method takes the form (7–8), with the functions G_i given by

$$G_i(z_2, z_3, \ldots, z_{i-1}) = c_i \left(\frac{1 + \displaystyle\sum_{j_2+j_3+\ldots+j_{i-1}=1}^{n_i^*} n_{j_2 j_3 \cdots j_{i-1}} z_2^{j_2} z_3^{j_3} \cdots z_{i-1}^{j_{i-1}}}{1 + \displaystyle\sum_{j_2+j_3+\ldots+j_{i-1}=1}^{d_i^*} d_{j_2 j_3 \cdots j_{i-1}} z_2^{j_2} z_3^{j_3} \cdots z_{i-1}^{j_{i-1}}} \right). \tag{45}$$

Moreover, n_i^* and d_i^* can be obtained from the functions F_i (in (44)) of the associated method, by means of

$$\begin{aligned} n_i^* &= \max\{j_2 + j_3 + \ldots + j_{i-1} \ / \ N_{j_1 j_2 \cdots j_{i-1}} \neq 0\} \\ d_i^* &= \max\{j_2 + j_3 + \ldots + j_{i-1} \ / \ D_{j_1 j_2 \cdots j_{i-1}} \neq 0\} , \end{aligned}$$

and therefore, are always lower or equal than $r_i + 1$ and r_i respectively.

Now it is a simple task to obtain the order conditions for the rational methods, from the order conditions for the methods of polynomial type. All we need is to consider the Taylor's expansion of the functions $G_i(s_2, \ldots, s_{i-1})$ in (45) (as functions of the s_j with $2 \leq j \leq i-1$), and then compare with the associated expansion of a method of polynomial type with the same number of stages. To show this, we will obtain the order conditions for the two and three-stage methods of rational type from the order conditions of the corresponding methods of polynomial type. Note at this point that it is also possible to obtain the order conditions for general methods (that is, for arbitrarily given G_i) in the same way, because only expansions of the G_i in terms of the s_j are involved. However, in what follows we will consider only rational functions G_i because this suffices for our stability purposes.

9 Two-stage methods of rational type.

To obtain all the explicit two-stage methods (of rational type) of order three, it suffices to note that from (16) and (17) we have that $G_2 = c_2 = 2/3$ and also

$$G_3(s_2) = c_3 \left(\frac{1 + \sum_{i=1}^{n_3^*} n_i s_2^i}{1 + \sum_{i=1}^{d_3^*} d_i s_2^i} \right) = 1 + \frac{1}{2} s_2 + \frac{1}{6} s_2^2 + O(s_2^3) \,, \tag{46}$$

must hold. Obviously it is enough to consider $n_3^* = d_3^* = 2$ in (46), obtaining the following conditions

$$\begin{aligned} c_2 &= 2/3 \,, \\ c_3 &= 1 \,, \\ n_1 &= 1/2 + d_1 \,, \\ n_2 &= 1/6 + (1/2)d_1 + d_2 \,. \end{aligned}$$

If we want to minimize the principal part of the local truncation error, all we need is to take $n_3^* = d_3^* = 3$ in (46) and expand to higher order the second term, obtaining

$$n_3 = 1/24 + (1/6)d_1 + (1/2)d_2 + d_3 \,.$$

The general form of a third order two-stage method of our family is given by

$$y_{n+1} = y_n + hk_1 G_3(s_2) \,,$$

where k_1, k_2 and s_2 are given by

$$k_1 = f(y_n) \,, \quad k_2 = f\left(y_n + \frac{2}{3} h k_1\right) \,, \quad s_2 = \frac{3(k_2 - k_1)}{2k_1} \,,$$

and G_3 takes the form

$$G_3(s_2) = \frac{1 + \frac{1 + 2d_1}{2} s_2 + \frac{1 + 3d_1 + 6d_2}{6} s_2^2 + \sum_{i=3}^{n_3^*} n_i s_2^i}{1 + d_1 s_2 + d_2 s_2^2 + \sum_{i=3}^{d_3^*} d_i s_2^i} \,.$$

Some third order methods with special properties (such as A-stability, L-stability, order four when applied to linear problems, etc) have been developed in [19, 20], where also some numerical experiments can be found.

10 Three-stage methods of rational type.

Now, following our notations in section 7, we can describe any three-stage method of rational type by

$$y_{n+1} = y_n + hk_1\tilde{G}_4(s_2, \tilde{s}_3) \,, \tag{47}$$

with $\tilde{s}_3 = s_3 - s_2$, and where as usually s_2 and s_3 are given through (13) in terms of the stages

$$k_1 = f(y_n) \,, \quad k_2 = f(y_n + hk_1G_2) \,, \quad k_3 = f(y_n + hk_1G_3(s_2)) \,,$$

Now the rational functions are given by

$$\begin{aligned} G_2 &= c_2 \,, \\ G_3(s_2) &= c_3 \left(\frac{1 + \sum_{i=1}^{n_3^*} n_i s_2^i}{1 + \sum_{i=1}^{d_3^*} d_i s_2^i} \right) , \\ \tilde{G}_4(s_2, \tilde{s}_3) &= c_4 \left(\frac{1 + \sum_{i+2j=1}^{\tilde{n}_4^*} n_{ij} s_2^i \tilde{s}_3^j}{1 + \sum_{i+2j=1}^{\tilde{d}_4^*} d_{ij} s_2^i \tilde{s}_3^j} \right) , \end{aligned} \tag{48}$$

and it is easily seen that when looking for fifth order formulas, it is enough to consider the parameters in (48) with $n_3^* = d_3^* = 3$ and $\tilde{n}_4^* = \tilde{d}_4^* = 4$, that is, the order conditions (for order five) are completely given in terms of the parameters c_i $(2 \leq i \leq 4)$, n_i and d_i $(1 \leq i \leq 3)$, and n_{ij} and d_{ij} $(1 \leq i + 2j \leq 4)$. In fact, we can obtain the order conditions for order five, by comparing (as in the two stage case) the expansions of the functions in (48) with those of the functions in (21) with the parameters given by (34). The c_i are given as in (34). The n_i are given in terms of a_1 in (34) and the free parameters a_2, a_3, d_1, d_2 and d_3, through relations

$$n_i = \sum_{j=0}^{i} a_j d_{i-j} \,, \quad 1 \leq i \leq 3 \,,$$

where obviously we take $a_0 = d_0 = 1$. The terms n_{ij} are given by

$$n_{i_1 i_2} = \sum_{j_k=0}^{i_k} a_{j_1 j_2}\, d_{i_1-j_1\, i_2-j_2}\,, \quad 1 \le i_1 + 2i_2 \le 4\,,$$

with $a_{00} = d_{00} = 1$, and where the parameters a_{ij} are those of (21). The d_{ij} can be arbitrarily chosen.

In order to minimize the principal part of the local truncation error, we take $n_3^* = d_3^* = 4$ and $\tilde{n}_4^* = \tilde{d}_4^* = 5$ in (48), obtaining the additional conditions

$$n_4 = \sum_{j=0}^{4} a_j d_{4-j}\,, \quad n_{i_1 i_2} = \sum_{j_k=0}^{i_k} a_{j_1 j_2}\, d_{i_1-j_1\, i_2-j_2}\,, \quad i_1 + 2i_2 = 5\,,$$

where a_i $(1 \le i \le 2)$ and a_{ij} $(1 \le i + 2j \le 5)$ are given as in (34), (42) and (43), and the other parameters are free (as before, we take $a_0 = d_0 = a_{00} = d_{00} = 1$).

11 Linear stability properties of the methods.

Now we are going to study the linear stability properties of the methods. When we apply a p-stage method (7) of our family to the scalar test equation

$$y' = \lambda\, y\,, \quad \lambda \in \mathbb{C}\,,$$

we get

$$y_{n+1} = R(z)\, y_n\,,$$

where $R(z)$ is the associated stability function, with $z = h\lambda$. Moreover, from (6–9) we obtain recursively

$$\begin{aligned} k_1 &= \lambda\, y_n \\ s_2 &= z\, G_2 \\ s_3 &= z\, G_3(z\, G_2) \\ &\vdots \\ s_p &= z\, G_p(z\, G_2, z\, G_3(z\, G_2), \ldots, z\, G_{p-1}(z\, G_2, \ldots, z\, G_{p-2}(\cdots)))\,, \end{aligned} \tag{49}$$

from which we obtain the following expression for the stability function

$$R(z) = 1 + z\, G_{p+1}(s_2, s_3, \ldots, s_p)\,,$$

where the s_i are given in terms of z by the relations (49).

With our change of notation for the s_i in (13), it is not difficult to see that in place of (49) we have that

$$\begin{aligned} k_1 &= \lambda y_n \\ s_2 &= z \\ s_3 &= z\, G_3(z) \\ &\vdots \\ s_p &= z\, G_p(z, z\, G_3(z), \ldots, z\, G_{p-1}(z, \ldots, z\, G_{p-2}(\cdots))) \end{aligned} \tag{50}$$

(with $z = h\,\lambda$), and so the associated stability function is now given by

$$R(z) = 1 + z\, G_{p+1}(s_2, s_3, \ldots, s_p)\,,$$

in terms of z through relations (50).

12 Three-stage methods being A-stable.

From the above section it is clear that the stability function of a method of polynomial type is always a polynomial function. Therefore it is not possible to obtain formulas of polynomial type with good linear stability properties such as A-stability or L-stability. However, it is also clear that we can obtain methods of rational type whose associated stability function is a rational function. Moreover, we can obtain A-stable and L-stable methods of rational type, without losing the highest attainable order for a given number of stages.

Now we will give some three-stage methods of order five being A-stable (or L-stable). For all such methods we consider the c_i given as in (34) (with the upper sign), that is

$$c_2 = \frac{6-\sqrt{6}}{10}\,, \quad c_3 = \frac{6+\sqrt{6}}{10}\,, \quad c_4 = 1\,.$$

We will also take $d_i = 0$ $(i \geq 1)$ in all cases, obtaining that $n_i = a_i$.

For example, taking

$$d_{10} = \frac{-3}{5}\,, \quad d_{01} = \frac{3-7\sqrt{6}}{30}\,, \quad d_{20} = \frac{77-18\sqrt{6}}{100}\,, \quad d_{11} = \frac{153+29\sqrt{6}}{360}\,, \quad d_{30} = \frac{27-73\sqrt{6}}{600}\,,$$
$$d_{21} = \frac{-44+3\sqrt{6}}{120}\,, \quad d_{40} = \frac{-168+97\sqrt{6}}{600}\,, \quad n_1 = \frac{-3+2\sqrt{6}}{5}\,, \quad n_{10} = \frac{-1}{10}\,,$$
$$n_{01} = \frac{63-37\sqrt{6}}{180}\,, \quad n_{20} = \frac{216-79\sqrt{6}}{300}\,, \quad n_{11} = \frac{44-3\sqrt{6}}{120}\,, \quad n_{30} = \frac{168-97\sqrt{6}}{600}\,,$$

and the other parameters equal to zero, we obtain from (47–48) a method whose associated stability function is given by the $(2,3)$-Padé approximation to the exponential function (see e.g. [3]), and thus being L-stable.

Taking

$$d_{10} = \frac{-2}{3}, \quad d_{01} = \frac{3-7\sqrt{6}}{30}, \quad d_{20} = \frac{41-9\sqrt{6}}{50}, \quad d_{11} = \frac{431-59\sqrt{6}}{600}, \quad d_{30} = \frac{1396-619\sqrt{6}}{750},$$
$$d_{21} = \frac{1436-709\sqrt{6}}{3600}, \quad d_{40} = \frac{432353-178017\sqrt{6}}{180000}, \quad d_{31} = \frac{-20769+7966\sqrt{6}}{21600},$$
$$d_{50} = \frac{127698-38147\sqrt{6}}{1080000}, \quad d_{60} = \frac{-7193669+2942716\sqrt{6}}{2160000}, \quad n_1 = \frac{-3+2\sqrt{6}}{5},$$
$$n_2 = \frac{-519+226\sqrt{6}}{300}, \quad n_{10} = \frac{-1}{6}, \quad n_{01} = \frac{63-37\sqrt{6}}{180}, \quad n_{20} = \frac{221-79\sqrt{6}}{300},$$
$$n_{11} = \frac{3474-1111\sqrt{6}}{5400}, \quad n_{30} = \frac{43409-18001\sqrt{6}}{18000}, \quad n_{21} = \frac{20769-7966\sqrt{6}}{21600},$$
$$n_{40} = \frac{1892669-781091\sqrt{6}}{540000}, \quad n_{50} = \frac{7193669-2942716\sqrt{6}}{2160000},$$

(the other parameters are zero) we get another L-stable method, being optimal with respect to the local truncation error. The associated stability function is given by the $(2,4)$-Padé approximation to the exponential function.

Finally, if we take

$$d_{10} = \frac{-1}{2}, \quad d_{01} = \frac{3-7\sqrt{6}}{30}, \quad d_{20} = \frac{36-9\sqrt{6}}{50}, \quad d_{11} = \frac{1323-247\sqrt{6}}{1800},$$
$$d_{30} = \frac{5969-2566\sqrt{6}}{3000}, \quad d_{21} = \frac{1159-486\sqrt{6}}{2400}, \quad d_{40} = \frac{480158-199037\sqrt{6}}{180000},$$
$$d_{31} = \frac{-3729+1411\sqrt{6}}{3600}, \quad d_{50} = \frac{135777-46528\sqrt{6}}{720000}, \quad d_{60} = \frac{-1282889+525021\sqrt{6}}{360000},$$
$$n_1 = \frac{-3+2\sqrt{6}}{5}, \quad n_2 = \frac{-519+226\sqrt{6}}{300}, \quad n_{01} = \frac{63-37\sqrt{6}}{180},$$
$$n_{20} = \frac{216-79\sqrt{6}}{300}, \quad n_{11} = \frac{421-144\sqrt{6}}{600}, \quad n_{30} = \frac{45569-18791\sqrt{6}}{18000},$$
$$n_{21} = \frac{3729-1411\sqrt{6}}{3600}, \quad n_{40} = \frac{694953-286792\sqrt{6}}{180000}, \quad n_{50} = \frac{1282889-525021\sqrt{6}}{360000},$$

the resulting method is A-stable, with the property of being optimal with respect to the local truncation error, and with associated stability function given by the $(3,3)$-Padé approximation to the exponential function.

13 Numerical experiments (the scalar case).

In order to show the behaviour as $h \to 0$ for the methods explained in the last section, we will consider the following simple problem (taken from [3], pp. 134)

$$y'(x) = \frac{y(x)(1-y(x))}{2y(x)-1}, \qquad y(0) = \frac{5}{6},$$

for which the solution is

$$y(x) = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{5}{36}\, e^{-x}}.$$

With fixed step size $h = 2^{-n}$ for various $n = 1, 2, 3, \ldots, 9$ over 2^n steps, the value of $y(1)$ was computed using our methods and two Runge-Kutta methods. The magnitude of the error E for different h and for each of these methods is shown in Figure 1 in logarithmic scale. The fifth-order methods of this section with associated stability functions given by the $(2,3)$, $(2,4)$ and $(3,3)$-Padé approximation to the exponential function are marked $M23$, $M24$ and $M33$ respectively. For comparison purposes we also include in figure 1 a three stage third-order Runge-Kutta method (see e.g. [3], pp. 134) marked $RK3$ and a six stage fifth-order Runge-Kutta method (see e.g. [3], pp. 202) marked $RK5$.

On the logarithmic scale used for this figure, the error for each method is represented very closely by a straight line whose slope equals the order of the method. It can be seen that methods marked $M24$ and $M33$ perform very similarly for this problem, and the slope for the associated graphs is bigger than 5 (in fact ≈ 5.5). It is easy to explain this behaviour by noting that both methods share the property of being optimal with respect to the local truncation error. It is also clear that, at the cost of more arithmetical operations, our methods perform better than the Runge-Kutta method marked $RK3$ with the same number of stages, and than the Runge-Kutta method marked $RK5$ with the same order (and this with less function evaluations).

To show the good behaviour of the methods from a stability point of view, we will also consider the following problem

$$y'(x) = -b\, y(x)\sqrt{c^2 + y^2(x)}\,, \qquad y(0) = a\,, \tag{51}$$

depending on the three parameters a, $b > 0$ and $c > 0$, for which the exact solution is given by

$$y(x) = \frac{a\, c}{c\, \mathrm{ch}(b\, c\, x) + \sqrt{a^2 + c^2}\, \mathrm{sh}(b\, c\, x)}.$$

The derivative with respect to y of the function $f(y) = -b\, y\sqrt{c^2 + y^2}$ in (51) is

$$f_y = -b\, \frac{c^2 + 2y^2}{\sqrt{c^2 + y^2}}\,, \tag{52}$$

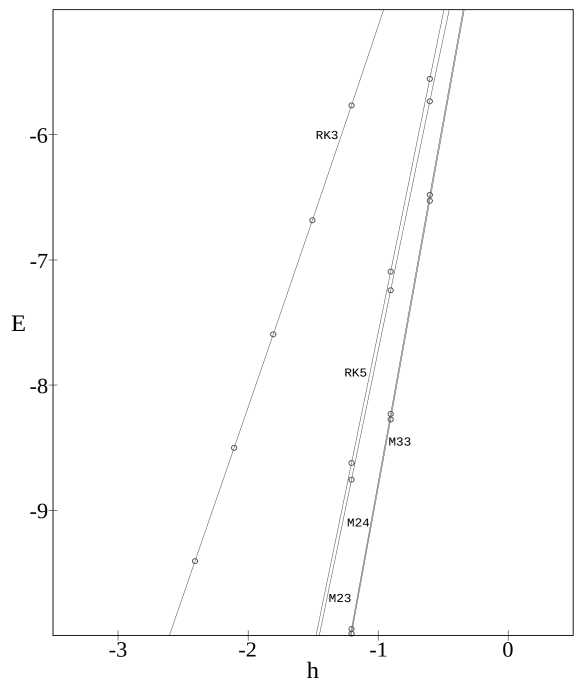

Figure 1: Error versus step size (double logarithmic scale) for various methods.

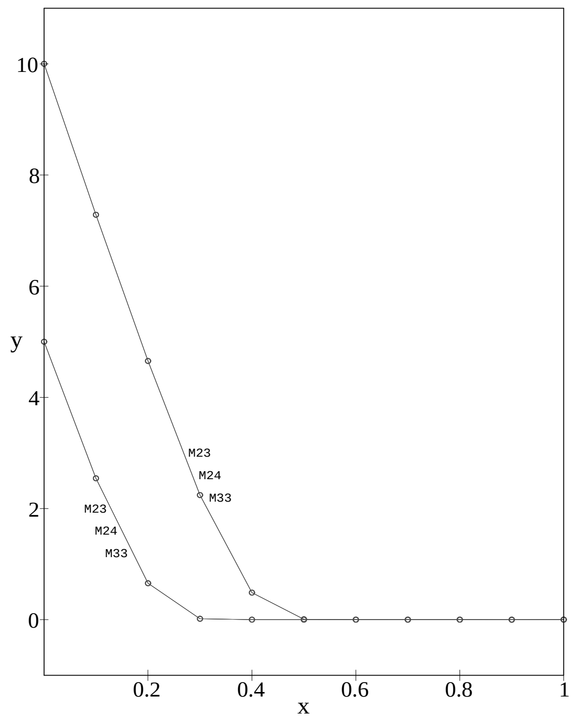

Figure 2: Numerical solutions for $a = 5, 10$, $b = 10$, $c = 3000$, with $h = 0.1$, $x \in [0, 1]$.

and so function f is one-sided Lipschitz continuous ($f_y < 0$) with one-sided Lipschitz constant 0. Therefore, the true solutions of this non-linear problem show a contractive behaviour. In fact, the solutions after a transient are virtually identical to the steady-state solution $y(x) \equiv 0$ (the solution of (51) when $a = 0$). We also have from (52) and from our previous comment that $f_y \approx -b\,c$ along the integration (at least after the transient).

To illustrate the behaviour of the methods marked $M23$, $M24$ and $M33$ when applied to a non-linear stiff problem, we take $b = 10$ and $c = 3000$ in (51), and integrate this problem over $x \in [0, 1]$ with initial conditions $a = 5, 10$ and fixed step size $h = 0.1$. Figure 2 shows the good qualitative behaviour of the numerical solutions we get in this manner. The three fifth-order methods perform very similarly for this problem. Note that for the range of values of the initial condition a we are considering, we have $f_y \approx -30000$ along the integration, and therefore the two explicit Runge-Kutta methods marked $RK3$ and $RK5$ give numerical overflow when applied to this problem with fixed step size $h \geq 0.0001$.

The numerical solutions we get from our methods when applied to problem (51) for a wide range of values of the parameters a, b and c (with $b\,c >> 1$ in order to retain the stiffness of the problem), and fixed step size $h = 0.1$, show that the qualitative behaviour is not always as good.

For example, when f_{yy} is too big (that is, f_y is a rapidly varying function) the numerical solutions tend to the steady-state slowly. However, decreasing the step size the situation is easily solved. In fact, the step size to be used seems to depend much more on the nonlinear character of f that on the stiffness of the problem.

14 Conclusions.

In this first part we have introduced and studied a new family of methods for the numerical integration of scalar autonomous ODEs. This work will help us in a second part because most of the results and properties can be easily extended to system problems. These new methods seem quite promising, for instance in the context of solving some nonlinear parabolic equations (by the method of lines). We will investigate this and other questions in part II: 'The separated system case'.

References

[1] J.D. Lambert, *Numerical Methods for Ordinary Differential Systems. The Initial Value Problem.* Wiley, Chichester, 1991.

[2] G. Dahlquist, A special stability problem for linear multistep methods. *BIT* **3**: 27–43 (1963).

[3] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods.* Wiley, Chichester, 1987.

[4] J.C. Butcher, Implicit Runge-Kutta processes. *Math. Comp.* **18**: 50–64 (1964).

[5] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems.* Springer-Verlag, Berlin, 1996.

[6] R. Alexander, Diagonally implicit Runge-Kutta methods for stiff O.D.E.'s. *SIAM J. Numer. Anal.* **14**: 1006–1021 (1977).

[7] H.H. Rosenbrock, Some general implicit processes for the numerical solution of differential equations. *Comput. J.* **5**: 329–330 (1963).

[8] S.P. Nørsett and A. Wolfbrandt, Order Conditions for Rosenbrock Type Methods. *Numer. Math.* **32**: 1–15 (1979).

[9] P. Kaps and P. Rentrop, Generalized Runge-Kutta Methods of Order Four with Stepsize Control for Stiff Ordinary Differential Equations. *Numer. Math.* **33**: 55–68 (1979).

[10] P. Kaps and G. Wanner, A Study of Rosenbrock-Type Methods of High Order. *Numer. Math.* **38**: 279–298 (1981).

[11] J. G. Verwer and S. Scholz, Rosenbrock methods and time-lagged Jacobian matrices. *Beitr. Numer. Math.* **11**: 173–183 (1983).

[12] J. G. Verwer, S. Scholz, J. G. Blom and M. Louter-Nool, A Class of Runge-Kutta-Rosenbrock Methods for Solving Stiff Differential Equations. *Z. Angew. Math. Mech.* **63**: 13–20 (1983).

[13] P. Kaps and A. Ostermann, Rosenbrock Methods using few LU-Decompositions. *IMA J. Numer. Anal.* **9**: 15–27 (1989).

[14] A.M. Urbani, Metodi Row modificati per problemi stiff. *Calcolo* **27**: 89–102 (1990).

[15] T. Steihaug and A. Wolfbrandt, An Attempt to Avoid Exact Jacobian and Nonlinear Equations in the Numerical Solution of Stiff Differential Equations. *Math. Comp.* **33**: 521–534 (1979).

[16] H. Zedan, Avoiding the exactness of the Jacobian matrix in Rosenbrock formulae. *Comput. Math. Appl.* **19**: 83–89 (1990).

[17] J. G. Verwer, S-Stability Properties for Generalized Runge-Kutta Methods. *Numer. Math.* **27**: 359–370 (1977).

[18] J.D. Day and D.N.P. Murthy, Two Classes of Internally S-Stable Generalized Runge-Kutta Processes Which Remain Consistent With an Inaccurate Jacobian. *Math. Comp.* **39**: 491–509 (1982).

[19] J. Álvarez and J. Rojo, New A-stable explicit two-stage methods of order three for the scalar autonomous IVP, in *Proc. of the 2nd. Meeting on Numerical Methods for Differential Equations, NMDE'98* (P. de Oliveira, F. Oliveira, F. Patrício, J.A. Ferreira, A. Araújo, Eds.), Coimbra, Portugal, 57–66 (1998).

[20] J. Álvarez and J. Rojo, A New Family of Explicit Two-Stage Methods of order Three for the Scalar Autonomous IVP. *Intl. Journal of Applied Sc. & Computations* **5**: 246–251 (1999).

[21] J. Álvarez, Obtaining New Explicit Two-Stage Methods for the Scalar Autonomous IVP with Prefixed Stability Functions. *Intl. Journal of Applied Sc. & Computations* **6**: 39–44 (1999).

[22] J. Álvarez and J. Rojo, Special methods for the numerical integration of some ODE's systems, in *The Third World Congress of Nonlinear Analysts, WCNA-2000, July 19–26, 2000*, Catania, Sicily, Italy.