Highly stable multivalue numerical methods

Raffaele D'Ambrosio, Martina Moccaldi and Beatrice Paternoster

Department of Mathematics, University of Salerno, Via Giovanni Paolo II, 132 - 84084 Fisciano (SA), Italy

Abstract. The numerical solution of partial differential equations discretized along the space variables requires the employ of highly stable methods, due to their intrisic multiscale (thus stiff) nature. The purpose of this paper is then the introduction of some building blocks leading to an efficient and accurate treatment of such stiff problems through highly stable multivalue numerical methods. We present a strategy based on a suitable modification of collocation technique which avoids, unlike classical collocation based Runge-Kutta methods, the order reduction phenomenon. Some novel issues on the error analysis, in view of a combined variable stepsize-variable order implementation, are here presented.

Keywords: Semi-discretized Partial Differential Equations, multivalue numerical methods, collocation methods
PACS: 02.60.Lj

INTRODUCTION

In this paper, we aim to consider the numerical solution of stiff problems of the general class $y' = f(y)$ arising from time-dependent partial differential equations discretized along the space variable. The key point in solving such stiff problems is that of employing suitable highly stable numerical methods and, in the case of stiff differential systems, to avoid order reduction phenomena, typical of classical numerical formulae such as Runge-Kutta methods [1]. Even if there is an extensive bibliography regarding the numerical solution of stiff problems (we refer, for instance, to the monographs [1, 10, 11, 13] and references therein), this issue still deserves some attention.

"*Stiff equations are multiscale problems"*. This sentence, contained in the first pages of the paper [2] by J. R. Cash provides an intuitive idea of the nature of stiffness, very common in mathematical modeling: for instance, solving time-dependent partial differential equations by the method of lines based on the employ of finite elements or finite differences for the spatial discretization leads to stiff systems of ordinary differential equations, due to their instrisic multiscale nature.

Methodology: a brief review on (modified) collocation methods

As announced, we aim to numerically treat stiff problems. The idea we propose is that of employing numerical methods based on modified collocation techniques. Collocation [1, 10, 13, 17] is an extensively applied technique based on the idea of approximating the exact solution of a given functional equation with a continuous approximant belonging to a chosen finite dimensional space (desirably chosen coherently with the qualitative behaviour of the solution). Such an approximant usually satisfies interpolation conditions in the grid points and exactly satisfies the equation on a given set of points. We now briefly recall some basic aspects regarding collocation methods, together with some famous modifications developed in the literature.

- *One-step collocation.* In classical one-step collocation methods (see [1, 10, 13]) the collocation function is given by an algebraic polynomial $P_n(t)$, $t \in [t_n, t_{n+1}]$, satisfying

$$P_n(t_n) = y_n, \qquad P_n'(t_n + c_i h) = f(t_n + c_i h, P_n(t_n + c_i h)), \quad i = 1,2,...,m,$$

  i.e. interpolating the numerical solution in t_n and exactly satisfying the given system in $\{t_n + c_i h,\ i = 1,2,\dots m\}$, where $c_1, c_2, \dots, c_m$ are given real numbers. The solution in t_{n+1} can then be computed from the function evaluation $y_{n+1} = P_n(t_{n+1})$. Guillou and Soule [9] and Wright [17] independently proved that one step collocation methods form a subset of implicit Runge-Kutta methods, whose coefficients are given by certain integrals of the fundamental Lagrange polynomials. The maximum attainable order of such methods is $2m$, and it is obtained by using Gaussian collocation points [10, 13], while the uniform order of convergence over the entire integration interval is only m. As a consequence, they suffer from order reduction showing effective order equal to m [1]. Concerning their linear stability properties, it is known that collocation methods based on Gaussian and Lobatto IIIA nodes are A-stable, while the ones based on Radau IIA points are L-stable [1, 10, 13].

- *Perturbed collocation.* As remarked, only some implicit Runge-Kutta methods are of collocation type. An extension of the collocation idea, the so-called perturbed collocation, is due to Norsett and Wanner [16], and applies to all IRK methods. The authors prove in [16] the equivalence result between the family of perturbed collocation methods and Runge-Kutta methods. The interest of this results is that the properties of collocation methods (e.g. order, linear and nonlinear stability) can be proved in a reasonable short, natural and very elegant way, while it is known that, in general, these properties are very difficult to handle and investigate outside collocation.
- *Multistep collocation.* The idea of multistep collocation was first introduced by Lie and Norsett in [15] (also see [9, 10, 14, 15]) and extends the collocation technique to the family of multistep Runge-Kutta method. The collocation polynomial $P_n(t)$ satisfies the following interpolation and collocation conditions:

$$P_n(t_{n-i}) = y_{n-i} \quad i = 0,1,...,k-1, \qquad P_n'(t_n + c_j h) = f(t_n + c_j h, P(t_n + c_j h)), \quad j = 1,...,m.$$

The numerical solution is then given by $y_{n+1} = P_n(t_{n+1})$. Lie and Norsett [15] proved that the maximum attainable order is $2m+k-1$. They also proved the existence of $\binom{m+k-1}{k-1}$ nodes allowing superconvergence. However, the corresponding methods are not stiffly stable, while in [10] A-stable methods of highest order $2m+k-2$ are introduced.
- *Two-step collocation*. Two-step collocation methods [3] extend the collocation idea to the class of two-step Runge-Kutta methods (introduced by Jackiewicz and Tracogna in [12]), pursuing the aim of deriving highly stable collocation-based methods which do not suffer from order reduction. The continuous approximant is given by

$$P(t_n + sh) = \varphi_0(s) y_{n-1} + \varphi_1(s) y_n + h \sum_{j=1}^{m} \Big(\chi_j(s) f(P(t_{n-1} + c_j h)) + \psi_j(s) f(P(t_n + c_j h)) \Big), \; s \in [0,1]. \qquad (1)$$

The collocation polynomial (1) is expressed as linear combination of the unknown basis functions $\{\varphi_0(s), \varphi_1(s), \chi_j(s), \psi_j(s), j = 1,2,\ldots,m\}$, to be suitably determined. It is required that the polynomial $P(t_n + sh)$ interpolates the solution in the points t_{n-1} and t_n and collocates it in the points $t_{n-1} + c_i h$, $t_n + c_i h$, $i = 1,2,\ldots,m$. As proved in [4], this is equivalent to determine the basis functions as unique solution of the order conditions system

$$\begin{cases} \varphi_0(s) + \varphi_1(s) = 1, \\ \dfrac{(-1)^k}{k!} \varphi_0(s) + \displaystyle\sum_{j=1}^{m} \left(\chi_j(s) \frac{(c_j - 1)^{k-1}}{(k-1)!} + \psi_j(s) \frac{c_j^{k-1}}{(k-1)!} \right) = \dfrac{s^k}{k!}, \end{cases}$$

$s \in [0,1]$, $k = 1,2,\ldots,2m+1$. Thus, the maximum attainable order of convergence is $2m+1$, uniformly on the overall integration interval. However, according to the Daniel-Moore conjecture [1] (i.e. the maximum attainable order of a m-stage A-stable method is $2m$), such methods cannot be A-stable. Hence, a modification of this idea has been proposed to achieve at least A-stability, leading to the so-called family of almost collocation methods.
- *Almost collocation.* In order to fulfill the Daniel-Moore requirement, we are mainly interested in methods of order $p = m + r$, where $r = 1,2,\ldots,m$, obtained by relaxing some order conditions (thus, some interpolation and/or collocation conditions). The formulas obtained by imposing relaxed conditions are known in literature as *two-step almost collocation methods*. In [4, 5, 6, 8] many A-stable and L-stable methods have been introduced: such methods do not suffer from the order reduction phenomenon in the integration of stiff systems (see [1, 10]). This is in contrast to implicit Runge-Kutta methods, whose effective order of convergence is only m, suffering from order reduction.

ERROR ANALYSIS

In [6] an expression of the local truncation error has been provided, i.e.

$$\xi(t_n+sh) = h^{p+1}C_p(s)y^{(p+1)}(t_n) + h^{p+2}C_{p+1}(s)y^{(p+2)}(t_n) + h^{p+2}G_{p+1}(s)\frac{\partial f}{\partial y}\Big(y(t_n)\Big)y^{(p+1)}(t_n) + O(h^{p+3}), \quad (2)$$

where

$$C_\nu(s) = \frac{s^{\nu+1}}{(\nu+1)!} - \frac{(-1)^{\nu+1}}{(\nu+1)!}\varphi_0(s) - \sum_{j=1}^{m}\left(\chi_j(s)\frac{(c_j-1)^\nu}{\nu!} + \psi_j(s)\frac{c_j^\nu}{\nu!}\right), \qquad G_{p+1}(s) = \sum_{j=1}^{m}\eta_j\Big(\chi_j(s)+\psi_j(s)\Big),$$

with $\nu = p,\ p+1$. To get a computable estimation of the local error, useful in a variable stepsize environment, possible choices have been discussed in [4, 6]. In view of a variable order strategy, the terms of order $p+2$ in (2) need to be estimated. In particular, we look for estimates of the type

$$h^{p+2}y^{(p+2)}(t_n) \approx \alpha_0^{(1)}y_{n-1} + \alpha_1^{(1)}y_n + h\sum_{j=1}^{m}\left(\beta_j^{(1)}f\Big(P(t_{n-1}+c_jh)\Big) + \gamma_j^{(1)}f\Big(P(t_n+c_jh)\Big)\right), \quad (3)$$

$$h^{p+2}\frac{\partial f}{\partial y}\Big(y(t_n)\Big)y^{(p+1)}(t_n) \approx \alpha_0^{(2)}y_{n-1} + \alpha_1^{(2)}y_n + h\sum_{j=1}^{m}\left(\beta_j^{(2)}f\Big(P(t_{n-1}+c_jh)\Big) + \gamma_j^{(2)}f\Big(P(t_n+c_jh)\Big)\right), \quad (4)$$

where the real parameters $\alpha_0^{(k)}, \alpha_1^{(k)}, \beta_j^{(k)}, \gamma_j^{(k)}$, with $k = 1,2$ and $j = 1,2,\ldots,m$, can be computed according to the following novel result, whose proof will be reported in [7].

Theorem 1 *Setting*

$$\rho_\ell = \frac{(-1)^{p+\ell}}{(p+\ell)!} - C_p(-1), \qquad \sigma_\ell(t) = \frac{t^p}{p!} - C_{p+\ell}(t),$$

the parameters appearing in the estimates (3) *and* (4) *satisfy the following systems of equations:*

$$\begin{cases} \alpha_0^{(k)} + \alpha_1^{(k)} = 0, \qquad \dfrac{(-1)^i}{i!}\alpha_0^{(k)} + \displaystyle\sum_{j=1}^{m}\left(\beta_j^{(k)}\frac{(c_j-1)^{i-1}}{(i-1)!} + \gamma_j^{(k)}\frac{c_j^{i-1}}{(i-1)!}\right) = 0, \qquad i = 1,2,\ldots,p, \\ \rho_1\alpha_0^{(k)} + \displaystyle\sum_{j=1}^{m}\Big(\beta_j^{(k)}\sigma_0(c_j-1) + \gamma_j^{(k)}\sigma_0(c_j)\Big) = 0, \qquad \rho_2\alpha_0^{(k)} + \displaystyle\sum_{j=1}^{m}\Big(\beta_j^{(k)}\sigma_1(c_j-1) + \gamma_j^{(k)}\sigma_1(c_j)\Big) = 1, \\ G_{p+1}(-1)\alpha_0^{(k)} + \displaystyle\sum_{j=1}^{m}\Big(\beta_j^{(k)}G_{p+1}(c_j-1) + \gamma_j^{(k)}G_{p+1}(c_j)\Big) = 0, \end{cases}$$

for $k = 1,2$, *being* p *the order of the method associated to* (1).

In this paper, we report the values of the parameters for the estimate (3) associated to the methods with $m = 1,2,3,4$ reported in [6]. We observe that, for all the methods in [6], there is no need to provide the estimate (4), since for those methods we have $G_{p+1}(s) = 0$ for construction itself. In the case $m = 1$, the values of the parameters in (3) are

$$\alpha_0^{(1)} = -\frac{9}{5}, \qquad \alpha_1^{(1)} = \frac{9}{5}, \qquad \gamma_1^{(1)} = -\frac{9}{10}, \qquad \beta_1^{(1)} = -\frac{9}{10}.$$

For the method in [6] with $m = 2$, the coefficients in (3) are

$$\alpha_0^{(1)} = \frac{2432}{395}, \qquad \alpha_1^{(1)} = -\frac{2432}{395}, \qquad \gamma_1^{(1)} = \frac{4864}{395}, \qquad \gamma_2^{(1)} = -\frac{9728}{1185}, \qquad \beta_1^{(1)} = \frac{2432}{1185},$$

in correspondence of $\beta_2^{(1)} = 2$. In the case $m = 3$, we have

$$\alpha_0^{(1)} = -\frac{245}{561}, \qquad \alpha_1^{(1)} = \frac{245}{561}, \qquad \gamma_1^{(1)} = -\frac{3117}{79}, \qquad \gamma_2^{(1)} = \frac{3605}{57}, \qquad \gamma_3^{(1)} = -\frac{6125}{232}, \qquad \beta_1^{(1)} = \frac{1959}{901},$$

in correspondence of the values $\beta_2^{(1)} = \beta_3^{(1)} = 0$. Finally, for the method in [6] with $m = 4$, the coefficients in (3) are

$$\alpha_0^{(1)} = -\frac{355}{2356}, \quad \alpha_1^{(1)} = \frac{355}{2356}, \quad \gamma_1^{(1)} = -\frac{199}{28}, \quad \gamma_2^{(1)} = \frac{3351}{178}, \quad \gamma_3^{(1)} = -\frac{1337}{66}, \quad \gamma_4^{(1)} = \frac{5671}{890}, \quad \beta_1^{(1)} = \frac{107}{6565}.$$

in correspondence of the values $\beta_2^{(1)} = \beta_3^{(1)} = \beta_4^{(1)} = 0$.

As a test example, we consider the Prothero-Robinson equation

$$y'(t) = \lambda(y(t) - G(t)) + G'(t), \tag{5}$$

with $G(t) = \sin(t)$, $t \in [0, 2\pi]$, $y(0) = 0$. The implementation strategy we employ combines the stepsize control [6] and the order control [7], by opportunely switching over the family of methods introduced in [6] whose order of convergence are the integers values in the interval $[1,4]$. Figure 1 reports the results for $\lambda = 1\mathrm{e}-6$ and tolerance $1\mathrm{e}-6$: this simple test is only intended to provide a first numerical evidence, obtained by applying the method to a classical problem whose analytical solution is a priori known and, thus, to realize whether basic theoretical expectations are recovered or not. More advanced experiments on a large set of stiff problems will be object of [7]. As expected, the solver tends to increase the order of the method in critical points of the solution.

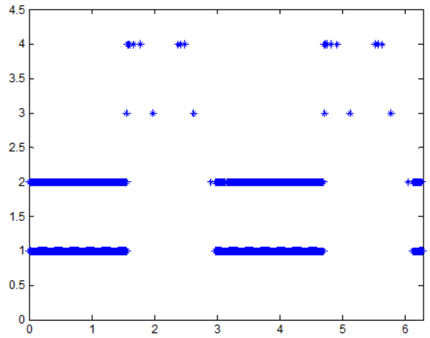

FIGURE 1. Order of convergence of the method employed in each step point for the numerical solution of (5). The implementation strategy is variable stepsize-variable order, involving two-step almost collocation methods [6] of orders from 1 to 4.

ACKNOWLEDGMENTS

This work was supported by INDAM (Istituto Nazionale di Alta Matematica) - GNCS 2014 projects.

REFERENCES

1. J.C. Butcher, *Numerical Methods for Ordinary Differential Equations*, 2nd Edition, John Wiley & Sons, Chichester, 2008.
2. J.R. Cash, *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **459**(2032), 797–815 (2003).
3. R. D'Ambrosio, M. Ferro and B. Paternoster, "Collocation-Based Two-Step Runge-Kutta Methods for Ordinary Differential Equations", in *ICCSA 2008*, Lecture Notes in Computer Science 5073, Springer, 2008, pp. 736–751.
4. R. D'Ambrosio, M. Ferro, Z. Jackiewicz and B. Paternoster, *Numer. Algor.* **53**(2-3), 195–217 (2010).
5. R. D'Ambrosio and Z. Jackiewicz, *Numer. Algor.* **54**(2), 169–193 (2010).
6. R. D'Ambrosio and Z. Jackiewicz, *Math. Comput. Simul.* **81**(9), 1707–1728 (2011).
7. R. D'Ambrosio, M. Moccaldi and B. Paternoster, *in preparation*.
8. R. D'Ambrosio and B. Paternoster, *Appl. Numer. Math.* **62**(10), 1325–1334 (2012).
9. A. Guillou and F.L. Soulé, *RAIRO Anal. Numér. Ser. Rouge* R-3, 17–44 (1969).
10. E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II – Stiff and Differential–Algebraic Problems*, Springer Series in Computational Mathematics 14, Springer–Verlag, Berlin, 2002.
11. Z. Jackiewicz, *General Linear Methods for Ordinary Differential Equations*, John Wiley & Sons, Hoboken, New Jersey, 2009.
12. Z. Jackiewicz and S. Tracogna, *SIAM J. Numer. Anal.* 32, 1390–1427 (1995).
13. J.D. Lambert, *Numerical methods for ordinary differential systems: The initial value problem*, John & Wiley, Chichester, 1991.
14. I. Lie, *Numer. Math.* 57(8), 779–708 (1990).
15. I. Lie and S.P. Norsett, *Math. Comp.* 52(185), 65–79 (1989).
16. S.P. Norsett and G. Wanner, *Numer. Math.* 38(2), 193–208 (1981).
17. K. Wright, *BIT* 10, 217–227 (1970).