UNIVERITÀ DEGLI STUDI DI SALERNO

Dipartimento di Matematica

Dottorato di Ricerca in Matematica

(X Ciclo - Nuova Serie)

Ph.D. Dissertation

Numerical treatment of Special Second Order Ordinary Differential Equations: General and Exponentially Fitted Methods

Elena Esposito

Advisor:
Prof. ssa B. Paternoster

Ph.D. Director:
Prof. ssa P. Longobardi

Anno Accademico 2010 - 2011

To you... who are always with me.

Acknowledgments

It is a pleasure to thank the many people who made this thesis possible.

It is difficult to express my gratitude to my Ph.D. supervisor, Prof. Beatrice Paternoster. With her enthusiasm, her inspiration, and her great efforts to explain things clearly and simply, she helped to make mathematics fun for me. Throughout my thesis-writing period, she provided encouragement, sound advice, good teaching, good company, and lots of good ideas.

I sincerely thank Prof. Liviu Gr. Ixaru, Dr. Dajana Conte and Dr. Raffaele D'Ambrosio, with whom I had the opportunity to closely collaborate in the scientific research and the group of Numerical Analysis of the Department of Mathematics of Salerno.

In particular, the good advice, support and friendship of Dr. Raffaele D'Ambrosio, has been invaluable on both an academic and a personal level, for which I am extremely grateful.

I would like to thank the Ph.D. Director, Prof. P. Longobardi, who allowed me to make this experience wonderful.

A very special thanks goes out to my colleagues of the office number 6 in Fisciano, who have been my traveling companions with whom I have shared many experiences and emotions.

Lastly, and most importantly, I wish to thank my parents. They raised me, supported me, taught me, and loved me.

Contents

1 Introduction **1**

1.1 Problems and motivations 2

1.2 Some recent models 4

1.3 State of the art of numerical methods for $y'' = f(t, y)$ 7

1.4 Stability properties of numerical methods for $y'' = f(t, y)$ 10

1.5 Structure of the thesis 13

I General Linear Methods **15**

2 General Linear Methods for first order ordinary differential equations **17**

2.1 Introduction to General Linear Methods 17

2.2 Representation of GLMs 18

2.3 Classical methods as GLMs 20

2.4 Preliminary properties: preconsistency, consistency, zero-stability, convergence 23

2.5 Order and stage order conditions 27

2.6 Linear stability theory 29

3 General Linear Methods for second order ordinary differential equations 33

3.1 Introduction to GLMs for second order ODEs 33

3.2 Representation of GLMs 34

3.3 Classical methods regarded as GLMs 38

3.3.1 Linear multistep methods 38

3.3.2 Runge-Kutta-Nyström methods 40

3.3.3 Coleman hybrid methods 41

3.3.4 Two-step Runge-Kutta-Nyström methods 41

3.3.5 Parallel-Iterated Pseudo Two-Step Runge-Kutta-Nyström methods 43

3.4 Preconsistency and consistency 44

3.5 Zero-stability 48

3.6 Convergence 53

3.7 Order conditions 61

3.8 Linear stability analysis 69

3.9 Example of new methods 71

II Exponentially Fitted Methods 75

4 Exponentially fitted two-step hybrid methods for second order ordinary differential equations 77

4.1 Introduction 77

4.2 Construction of the methods 79

4.3 Error analysis and estimation of the parameters 85

4.4 Linear stability analysis 86

4.4.1 Methods with constant coefficients 86

4.4.2 Methods with coefficients depending on one and two frequencies 88

4.5 Numerical results 89

5 Parameter estimation for exponentially fitted hybrid methods 93

5.1 Introduction 93

5.2 Exponentially fitted two-step hybrid methods 94

5.3 Parameter selection 98

5.4 Numerical results 101

6 Exponentially fitted two-step Runge-Kutta methods 107

6.1 Introduction 107

6.2 Derivation of the methods 109

6.3 Parameter estimation 119

6.4 Linear stability analysis 121

6.5 Numerical results 121

7 Some new uses of the $\eta_m(Z)$ functions 129

7.1 Introduction 129

7.2 Definition and properties of $\eta_m(Z)$ functions 130

7.3 Description of the method and of the program 134

7.4 Applications 142

III Conclusions and future works 147

Appendix A 165

Appendix B 165

Appendix C 165

List of Tables

4.1 Relative errors corresponding to the solution of the problem (4.5.2), for different values of λ and k. 91

4.2 Relative errors corresponding to the solution of the problem (4.5.3). 91

4.3 Relative errors corresponding to the solution of the problem (4.5.4), with $\nu = 1/10$. 92

5.1 Relative errors corresponding to the solution of the problem (5.4.3), for different values of λ and k. 105

5.2 Relative errors corresponding to the solution of the problem (5.4.4). 105

5.3 Relative errors corresponding to the solution of the problem (5.4.5), with $\nu = 1$. 105

6.1 Numerical results for Problem 1 126

6.2 Numerical results for Problem 2 127

List of Figures

1 Region of stability in the (ν^2, Z)-plane for $s = 2$, with $c_1 = \frac{2}{3}$, $c_2 = \frac{4}{5}$. 154

2 Region of stability in the (ν^2, Z_1, Z_2)-space for $s = 4$, with $c = [0, \frac{1}{3}, \frac{2}{3}, 1]^T$. 155

3 Stability region of the EF5 method in the (Re ω, Im ω, z)-space. . . 160

4 Projection of the stability region in the (Re ω, Im ω)-plane for the adapted TSRK method above reported. 160

1 Introduction

The aim of this research is the construction and the analysis of new families of numerical methods for the integration of special second order Ordinary Differential Equations (ODEs). The modeling of continuous time dynamical systems using second order ODEs is widely used in many fields of applications, as celestial mechanics, seismology, molecular dynamics (see for instance [132, 151] and references therein contained), or in the semidiscretisation of partial differential equations (which leads to high dimensional systems and stiffness, compare [28]). Although the numerical treatment of this problem has been widely discussed in the literature, the interest in this area is still vivid, because such equations generally exhibit typical problems (e.g. stiffness, metastability, periodicity, high oscillations), which must efficiently be overcome by using suitable numerical integrators. The purpose of this research is twofold: on the one hand to construct a general family of numerical methods for special second order ODEs of the type $y'' = f(y(t))$, in order to provide an unifying approach for the analysis of the properties of consistency, zero-stability and convergence; on the other hand to derive special purpose methods, that follow the oscillatory or periodic behaviour of the solution of the problem.

1.1 Problems and motivations

In this work we focus our attention on the initial value problems based on special second order ODEs

$$\begin{cases} y''(t) = f(y(t)), \quad t \in [t_0, T], \\ y(t_0) = y_0 \in \mathbb{R}^d, \\ y'(t_0) = y'_0 \in \mathbb{R}^d, \end{cases} \tag{1.1.1}$$

where the function $f : \mathbb{R}^d \to \mathbb{R}^d$ does not explicitly depend on y' and is supposed to be smooth enough in order to ensure that the corresponding problem (1.1.1) is Hadamard well-posed. We observe that it is not a restriction to consider an autonomous ODE such as the one in (1.1.1), i.e. a differential equation whose right hand side does not explicitly depend on t, since any nonautonomous problem can be regarded as a differential system of the form (1.1.1). Although the problem (1.1.1) could be transformed into a doubled dimensional system of first order ODEs and solved by standard formulae for first order differential systems, the development of numerical methods for its direct integration is more natural and efficient. We are concerned with General Linear Methods (GLMs) for second order ODEs, with the aim to provide an unifying approach for the analysis of the basic properties of numerical methods for ODEs. This class of methods properly includes all the classical methods already considered in the literature, such as linear multistep methods [85], Runge-Kutta-Nyström methods [82], two-step hybrid methods [32] and two-step Runge-Kutta-Nyström methods [127] as special cases. The family of methods that we consider is wider and more general with respect to the ones already considered in the literature: in fact, our new methods depend on more parameters which can be exploited, for instance, in order to provide a better balance between order of convergence and stability properties.

In the context of first order ODEs, the circumlocution "generalized multistep methods" has been used for the first time by Gragg and Stetter [77] in 1964, where

they introduced predictor-corrector schemes depending also on the stage derivative in one "nonstep point", as they mention, i.e. in one internal point. This is one of the first attempts to combine a multivalue strategy (i.e. formulae passing a collection of vector as output and needing a similar collection as input in the successive step point) together with a multistage strategy (i.e. methods depending on the approximation to the solution in some internal points which do not belong to the grid), which is typical of Runge-Kutta methods. A systematic theory concerning GLMs for first order ODEs is due to J. C. Butcher (compare [21, 97] and the references therein contained) to provide an unifying framework for the approach to the basic questions of consistency, convergence and stability of numerical methods for ODEs. It is important to observe that the discovery of a GLM theory "opened the possibility of obtaining essentially new methods which were neither Runge-Kutta nor linear multistep methods nor slight variations of these methods" (compare [157]).

For second order ODEs (1.1.1) no systematic investigation on GLMs has begun till now: even if many linear and nonlinear methods appeared in the literature (see, for instance, [82, 83, 85, 94] and references therein contained) there is not yet a very wide and general class of methods for the numerical solution of the problem (1.1.1) together with a series of general theorems which provide an unifying framework to analyze the properties of such methods. In order to transfer to second order ODEs the same benefits obtained in the case of first order ODEs, the purpose of this work is the foundation of a theory of GLMs for the numerical solution of (1.1.1), starting from a suitable formulation of these methods.

Part 1 of this dissertation is devoted to laying the foundations of the theory of GLMs for second order ODEs. Part 2, instead, concerns with the construction and the theoretical analysis of special purpose methods, that represent an efficient approach to the problem (1.1.1) in presence of periodicity and high oscillations in the solution; in particular, we will concern with the so-called exponentially fitted methods [94]. In fact, classical numerical methods for ODEs may not be well-suited

to follow a prominent periodic or oscillatory behaviour of the solution, because, in order to catch the oscillations, a very small stepsize would be required with corresponding deterioration of the numerical performances, especially in terms of efficiency. For this reason, many classical numerical methods have been adapted in order to efficiently approach the oscillatory behaviour. One of the possible ways to proceed in this direction can be realized by imposing that a numerical method exactly integrate (whitin the round-off error) problems of type (1.1.1) whose solution can be expressed as linear combination of functions other than polynomials.

1.2 Some recent models

In this section we aim to present some of the most recent models, which involve systems of second order ODEs (1.1.1), of interest in Celestial Mechanics, Climatology, Chaos Theory, Ecology, Sensor Dynamics.

1. *The Pleiades problem.* The Pleiades problem (see, for istance, [82]) is a celestial mechanics problem describing the motion and the collisions of seven stars in the plane of coordinates (x_i, y_i) and masses $m_i = i$, $i = 1, 2, \ldots, 7$. By means of some mechanical considerations mainly based on the Newton's second law of motion, the mathematical description for this dynamical system can be formulated according to the following model:

$$\begin{cases} z'' = f(z), \quad z \in \mathbb{R}^{14}, \\ z(0) = z_0, \\ z'(0) = z_0', \end{cases} \tag{1.2.1}$$

where $z = [x, y]^T$, $x, y \in \mathbb{R}^7$, $f(z) = \left[f^{(1)}(z), f^{(2)}(z)\right]^T$, and the functions $f^{(1)}, f^{(2)} : \mathbb{R}^{14} \to \mathbb{R}^7$, assume the form

$$f_i^{(1)}(z) = \sum_{j \neq i} m_j \frac{x_j - x_i}{r_{ij}^{3/2}}, \quad f_i^{(2)}(z) = \sum_{j \neq i} m_j \frac{y_j - y_i}{r_{ij}^{3/2}}.$$

2. *Oscillations and Chaos: the Duffing problem.* The Duffing equation, introduced for the first time by Duffing in [69], is an important model which describes a nonlinear forced damped oscillator. The equation takes the form

$$\ddot{x} + \delta \dot{x} + \alpha x^3 + \beta x = \gamma \cos(\omega t),$$

where $\delta \geq 0$ is the damping constant, αx^3 is the nonlinearity term, γ is the forcing amplitude and ω is the forcing frequency. For $\beta > 0$ and in correspondence of small values of x (see the monography by Thompson and Stewart, 2002), the Duffing oscillator can be interpreted as a forced oscillator with a hardening spring if $\alpha > 0$, or a softening spring if $\alpha < 0$. For $\beta < 0$, the Duffing oscillator describes the dynamics of a point mass in a double well potential, and it can be regarded as a model of a periodically forced steel beam which is deflected toward two magnets (compare Moon and Holmes, 1979; Guckenheimer and Holmes, 1983; Ott, 2002). The study of the response of the system to the periodic forcing is particularly interesting: we discover, in fact, that the system is highly sensitive to initial conditions [145]. Small differences in initial conditions yield widely diverging outcomes, rendering long-term prediction impossible in general. This happens even though the model is deterministic, i.e. its future dynamics is fully determined by the initial conditions, with no random elements involved. In other words, the deterministic nature of these systems does not make them predictable. This behavioue is known as deterministic chaos, or simply *chaos* (compare, for istance, [146]).

3. *A model in Ecology: the height of forest trees.* Height is an important property of forest trees and reveals many important aspects, such as species succession,

forest mensuration and site assessment. An example of model describing the time growth of trees height is given in [113], where the autors provided the analysis of the influence of the first year's height growthin predicting heights for later ages. The model discussed in [113] assumes the following form:

$$\ddot{h}_{ijk} = \dot{j}_{ijk} - 2a\dot{h}_{ijk} - bh_{ijk},$$

where the function h_{ijk} regards the total tree height at age k, for the tree j on the site i of the forest, while the function g_{ijk} is tree height growth at age k, for the tree j on site i. In [113] the function g_{ijk} considered by the autors is such that

$$\dot{g}_{ijk} = c_0 + c_1 V(t) + (a_0 + a_1 V_2(t))\dot{h}_{ijk} + (b_0 + b_1 V_3(t))h_{ijk},$$

where $V_1(t)$ is an environmental variable representing the annual heat status of the site, $V_2(t)$ is an environmental variable representing the annual moisture status of the site, $V_3(t)$ is a site variable representing the nutrient status of the site, assumed to be constant over the life of the tree, but different among the sites.

4. *VDPOL problem.* The VDPOL problem originates from electronics and describes the behaviour of nonlinear vacuum tube circuits. The circuit scheme, designed by Balthazar van der Pol in the 1920's, is an RLC loop, but with the passive resistor of Ohm's Law replaced by an active element which would pump energy into the circuit whenever the amplitude of the current falls below a certain level. In the 1920's this active element was an array of vacuum tubes, now it is a semiconductor device. The voltage drop at the semiconductor (which used to be RI) is given by a nonlinear function $f(I)$ of the current I.

The problem is of the form:

$$\begin{cases} z'' = f(z, z'), \quad z \in \mathbb{R}, \quad t \in [t_0, T], \\ z(0) = z_0, \\ z'(0) = z'_0, \end{cases} \tag{1.2.2}$$

where the function f is given by

$$f(z, z') = \mu(1 - z^2)z' - z, \mu > 0.$$

1.3 State of the art of numerical methods for $y'' = f(t, y)$

In the early '90s, several linear methods for second order ODEs were found in the literature, mainly developed for particular applications, such as the Schrödinger equation and problems of celestial mechanics (from Gautschi, 1961 [75], Stiefel-Bettis, 1969 [142], Raptis-Allison, 1978 [135]). In many papers van der Houwen and others ([150, 151, 153]) propose new methods that take into account the special nature of the solution of the system of ODEs, deriving Runge-Kutta and Runge-Kutta-Nyström methods with high-order phase dispersion: a high order phase dispersion, in fact, ensures that the numerical solution of the differential equation is in phase with the exact solution on the test equation. Indeed, the investigation on the phase dispersion and the stability properties has been carried out on the homogeneous equation test

$$y'' = -\omega^2 y, \qquad \omega \in \mathbb{R},$$

known as simple armonic oscillator. It is not restrictive to shrink the study to the harmonic oscillator, because, for example, each motion which involves mechanical vibrations is attributable to simple harmonic oscillator for small amplitudes of vibration, and also, any periodic motion can be expressed as a series of harmonic

functions of time, which are analogous to the solutions of the used test equation. A natural evolution of the methods with high order of phase dispersion, which ensured a phase error less than that obtained by traditional methods, is represented by the RKN phase-fitted methods [128]. These methods, producing a homogenous phase error, which is zero on linear test equations, are suitable to integrate the second order ODEs, in which high-frequency oscillations are produced by a linear part.
In the context of the numerical integration of (1.1.1), collocation methods [67, 82, 83] deserve to be mentioned, mainly because their implicitness and good stability properties are well suited to numerical integration of stiff systems. Many papers appeared in the literature: we only cite [48], where a general procedure to determine global integration methods is derived, together a complete discussion on the choice of nodes, algorithms and implementation. Linear stability properties have been studied in [152] for classical Runge-Kutta-Nyström methods, and in [26, 27] for mixed collocation methods. We also briefly refer to some highlighted papers regarding numerical methods for (1.1.1) based on geometric integration, i.e. methods that preserve the geometric properties of the flow of the differential equation: we recall here the numerical schemes of Hockbruch, Lubich [86] and Hairer, Lubich [78] for second-order ODEs, where high frequencies are generated by a linear part. It is worth mentioning the numerical treatment of highly oscillatory differential equations, designed to require a new complete function evaluation only after a time step over one or many periods of the fastest oscillations in the system, with the aim to guarantee the numerical long-time conservation of energy. For a complete discussion of these methods, see [79, 96].
An interesting class of methods is the family of two-step Runge-Kutta methods, that have been introduced in the literature by Byrne and Lambert ([13]), to obtain the same order of one-step Runge-Kutta methods, but with a smaller number of stages thus gaining more efficiency. To improve the local accuracy of these methods, Costabile in [43] proposed the pseudo Runge-Kutta methods of second species, and

in [44, 45] the PRKm autostarting or of III species. Nakashima in [119] generalized the PRKm of III species of Costabile and obtained a further reduction of the cost. In the '90s, Jackiewicz *et al.* proposed a generalization of PRKm of first species; Hairer and Wanner have considered multistep and multistage methods, also providing remarkable insights to the multiderivative case [80]. Burrage presented a very general family of *multivalue* methods [15], deriving the order conditions. In [98, 100], the numerical solution depends on the value of the stages in two sequential steps; therefore, the higher number of parameters of the corresponding methods allows to obtain a high order without increasing the computational cost. The analysis of general two-step Runge-Kutta methods was carried out using the Albrecht's approach [100] and then using the theory of B-series [81]. As in the case of classical Runge-Kutta methods, many works have occurred in the literature concerning the derivation of order conditions [100, 81, 23], the construction of embedded two-step Runge-Kutta formulae, and their implementation with variable step size, in order to provide a local error estimation [102, 143, 144]. Messina et al. have considered the possibility of a parallel implementation of multistep Runge-Kutta methods [118]. In [47] a complete survey of pseudo Runge-Kutta appeared, considering also the special second order ODEs.

In the background of two-step nonlinear methods for second order ODEs $y'' = f(t, y)$, there are in the literature few examples of two-step Runge-Kutta-Nyström methods, obtained as indirect methods by two-step Runge-Kutta methods [129], and the existence of highly stable methods within this class has been proved [126].

In the context of the numerical integration of second-order ODEs with periodic or oscillating solutions, exponential-fitting methods [89] play a key role. These numerical methods benefit of the knowledge of the behaviour of the solution, and in particular, of the knowledge of the frequencies that appear in the solution. Within the class of linear multistep methods for second order ODEs, Gautschi [75] and Stiefel-Bettis [142] considered trigonometric functions depending on one or more frequencies, while

Lyche [115] derived methods exactly integrating initial value problems based on ODEs of order r whose solution can be expressed as linear combination of powers and exponentials; Raptis-Allison [135] and Ixaru-Rizea [91] derived special purpose linear multistep methods for the numerical treatment of the radial Schrödinger equation $y'' = (V(x) - E)y$, by means of trigonometric and exponential basis of functions. More recently, in the context of Runge–Kutta-Nyström methods, exponentially-fitted methods have been considered, for instance, by Calvo [24], Franco [72], Simos [137, 138] and Vanden Berghe [147], while their trigonometrically-fitted version has been developed by Paternoster in [125]; mixed-collocation based Runge–Kutta-Nyström methods have been introduced by Coleman and Duxbury in [33]. Recent adaptations of the Numerov method have been provided in [71, 87, 148]. For a more extensive bibliography see [94] and references within.

1.4 Stability properties of numerical methods for $y'' = f(t, y)$

In the numerical integration of stiff oscillatory systems, some special stability properties are required, notably the P–stability, but this desirable feature is not possessed by any class of numerical methods for $y'' = f(t, y)$. The concept of P–stability was first introduced by Lambert and Watson in [110], and it is of particular interest in the numerical treatment of periodic stiffness which is exhibited, for example, by Kramarz's system [107]. In this case two or more frequencies are involved, and the amplitude of the high frequency component is negligible or it is eliminated by the initial conditions. Then, the choice of the step size is governed not only by accuracy demands, but also by stability requirements. P–stability ensures that the choice of the step size is independent of the values of frequencies, but it only depends on the desired accuracy [34, 128]. Only a few numerical methods possess this desirable feature (see for istance [46, 150, 151]). It is worth mentioning that in the

class of linear multistep methods for (1.1.1), P–stability can be obtained only by methods of the second order and that the stability properties gradually deteriorate when the order increases.

We recall that linear stability analysis of numerical methods for the problem (1.1.1) is based on the test equation

$$y'' = -\omega^2 y, \tag{1.4.1}$$

where ω is a real constant. When a linear *p-step* method, such as any of those described by Henrici in [85], is applied to (1.4.1), it gives the difference equation

$$\sum_{j=0}^{p} (\alpha_j + \nu^2 \beta_j) y_{n+j} = 0, \tag{1.4.2}$$

for the set of numerical approximations $\{y_m\}$ for the solution at $\{x_m = x_0 + mh\}$, where h is a fixed steplength, $\nu = \omega h$, and the coefficients $\{\alpha_j\}$ and $\{\beta_j\}$ are indipendent of h. The solutions of (1.4.2) are determined by the roots r_s, for $s = 1, \ldots, p$, of the corresponding characteristic equation. Two of those roots, r_1 and r_2, tend to 1 as $\nu \to 0$; they are called the *principal roots* and the others, if any, are *spurious roots*. The solutions of the test equation (1.4.1) are periodic, with period $2pi/\omega$ for all non-trivial initial conditions on y and y'. The desire that numerical solutions should mimic this behaviour as closely as possible provides the motivation for the following definitions [110]: a linear multistep method has an interval of periodicity $(0, \nu_0^2)$ if, for all $\nu^2 \in (0, \nu_0^2)$, the roots r_s satisfy

$$r_1 = e^{i\phi(\nu)\omega}, \quad r_2 = e^{-i\phi(\nu)\omega}, \quad |r_s| \leq 1, \quad s > 2,$$

where $\phi(\nu)$ is real; a method is said to be *P-stable* if its interval of periodicity is $(0, \infty)$.

Numerov's method is not P-stable, but it has an interval of periodicity $(0, 6)$. In

[110], the authors proved that P–stable linear multistep methods

$$\sum_{j=0}^{p} \alpha_j y_{n+j} = h^2 \sum_{j=0}^{p} \beta_j f_{n+j}$$

can achieve maximum order 2. In [46] a family of two-step fourth order methods had already been derived, showing the existence of P-stable methods.

In the context of RungeKuttaNyström methods

$$\begin{aligned} y_{n+1} &= y_n + h y'_n + h^2 \sum_{i=1}^{m} \overline{b_i} f(t_n + c_i h, Y_i), \\ y'_{n+1} &= y'_n + h \sum_{i=1}^{m} b_i f(t_n + c_i h, Y_i), \\ Y_i &= y_n + c_i h y'_n + h^2 \sum_{j=1}^{m} a_{ij} f(t_n + c_j h, Y_j),\ i = 1, \ldots, m, \end{aligned}$$

many A-stable and P-stable methods exist: see [150, 151].

The methods falling in the subclass of collocation methods, whose coefficients (see [82]) are of the form

$$\begin{aligned} a_{ij} &= \int_0^{C_i} L_j(s) ds, \\ b_i &= \int_0^1 L_i(s) ds, \\ \overline{b}_i &= \int_0^1 (1 - s) L_i(s) ds, \end{aligned}$$

have only bounded stability intervals and are not P–stable [127]. It is also known that symmetric one step polynomial collocation methods can't be P-stable [31], and no P-stable methods were found in the special class of two step collocation methods considered in [126].

P-stability as defined for polynomial-based methods, ensures that the periodicity condition holds for all steplengths h, whatever the value of ω in the test equation. For an exponential-fitting method, Coleman–Ixaru in [34] gave the following definition

Definition 1.4.1. *A family of exponential-fitting methods with the stability function*

$R_{nm}(\nu^2;\vartheta)$, *where* $\nu = \omega h$ *and* $\vartheta = kh$, *is P-stable if, for each value of* k, *the inequality* $|R_{nm}(\nu^2;\vartheta) < 1|$ *holds for all values of* ω *and for all steplengths* h, *except possibly for a discrete set of exceptional values of* h *determined by the chosen value of* k.

1.5 Structure of the thesis

All the cited methods for (1.1.1) are included in the more general class of GLMs that is the object of Part 1 of this dissertation; Part 2 instead, is devoted to the construction of the exponentially fitted version of some classes of methods and their analysis. In particular, Part 1 is divided into 2 chapters. In *Chapter 2* we give, just for completeness, some preliminary notions on General Linear Methods for first order ordinary differential equations. We recall the formulation of the methods, the definitions of the main properties, such as preconsistency, consistency, zero-stability, convergence, order conditions and linear stability. *Chapter 3* is entirely devoted to the construction of General Linear Methods for second order ODEs (1.1.1) and their analysis. First we introduce the formulation of the methods and the representation of classical methods regarded as GLM. We give the definition of zero-stability, and a criterion to analyze zero-stability. Next the convergence analysis of GLMs is treated and, a sufficient condition is derived. Order conditions for GLMs are discussed and the analysis of linear stability is treated. Finally, examples of new methods are provided.

Part 2 of our dissertation concerns with the construction, the theoretical analysis and the numerical comparison with classical solvers of adapted numerical methods to efficiently approach the problem (1.1.1) in presence of periodicity and high oscillations in the solution. This part consists of 4 chapters. In *Chapter 4* we construct the exponentially fitted (EF) version of the class of *Coleman two-step hybrid me-*

thods [32], using the *Six-step flow chart* [94] for the determination of the coefficients of the methods. Next, we analyze the linear stability properties of the derived methods and finally, we provide numerical tests confirming the theoretical expectations. *Chapter 5* treats the problem of the parameter estimation for exponentially fitted hybrid methods, considered in Chapter 3. Here we present our proposal of parameter estimation technique, and we provide some numerical tests. *Chapter 6* is devoted to the derivation of EF version of the class of *two-step Runge-Kutta methods*. We present the constructive technique of adapted TSRK methods. Next, we approach the problem to estimate the unknown parameter on which the coefficients of the methods depend, and we analyze the linear stability properties of the derived methods. Finally, we provide numerical tests confirming the theoretical expectations.

In *Chapter 7* we develop a method for the conversion of formulae obtained in the frame of the exponential fitting for various approximation schemes, to forms expressed in terms of functions $\eta_m(Z)$, introduced in [89]. The new forms secure automatic elimination of $0/0$ behaviour, enable a uniform accuracy in the evaluation and allow an extended area of applicability. We also present a code for this conversion.

Part I

General Linear Methods

2 General Linear Methods for first order ordinary differential equations

2.1 Introduction to General Linear Methods

The construction of a general framework in which numerical methods for ODEs can be placed is certainly an useful tool for their development and analysis. Moreover, wider and more general classes of formulae with respect to classical ones depend on more parameters, which can be exploited to break the order barriers affecting classical methods and obtain higher order methods, or to achieve higher stability properties (e.g. large stability regions for explicit methods, A-stability, L-stability and algebraic stability for implicit methods) in order to approach the solution of ODEs in an efficient and accurate way. The circumlocution generalized multistep methods has been used for the first time by Gragg and Stetter [77] in 1964, where they introduced predictor-corrector schemes depending also on the stage derivative in one nonstep point, as they mention, i.e. in one internal point. This is one of the first attempts to combine a multivalue strategy (i.e. formulae passing a collection of vector as output and needing a similar collection as input in the successive step point, see [21]) together with a multistage strategy (i.e. methods depending on the approximation to the solution in some internal points which do not belong to the grid), which is typical of Runge-Kutta methods. Further important contributions in

the development of a theory of multivaluemultistage integration methods have been provided by J. Butcher from 1965 on (see [21] and references therein contained), Gear [76], Dahlquist [66], Donelson and Hansen [68], Byrne and Lambert [13], Costabile [43, 47], Jackiewicz and Tracogna [100]. In very recent times the first monography totally devoted to GLMs has also been released [97].

2.2 Representation of GLMs

General linear methods are introduced as a middle ground between linear multistep methods and Runge-Kutta methods. Consider the initial value problem for an autonomous system of differential equations, which we write in the form

$$\begin{cases} y'(t) = f(y(t)), \quad t \in [t_0, T], \\ y(t_0) = y_0, \end{cases} \tag{2.2.1}$$

where the function $f\mathbb{R}^m \to \mathbb{R}^m$, $y_0 \in \mathbb{R}^m$. GLMs for (2.2.1) can be represented by the abscissa vector $\mathbf{c} = [c_1, c_2, \ldots, c_s]^T$, and four coefficient matrices $\mathbf{A} = [a_{ij}]$, $\mathbf{U} = [u_{ij}]$, $\mathbf{B} = [b_{ij}]$ and $\mathbf{V} = [v_{ij}]$, where $\mathbf{A} \in \mathbb{R}^{s\times s}$, $\mathbf{U} \in \mathbb{R}^{s\times r}$, $\mathbf{B} \in \mathbb{R}^{r\times s}$, $\mathbf{V} \in \mathbb{R}^{r\times r}$. On the uniform grid $t_n = t_0 + nh$, $n = 0, 1, \ldots, N$, $Nh = T - t_0$, these methods take the form

$$\begin{cases} Y_i^{[n]} = \sum_{j=1}^{s} a_{ij} h f(Y_j^{[n]}) + \sum_{j=1}^{r} u_{ij} y_j^{[n-1]}, \quad i = 1, 2, \ldots, s, \\ y_i^{[n]} = \sum_{j=1}^{s} b_{ij} h f(Y_j^{[n]}) + \sum_{j=1}^{r} v_{ij} y_j^{[n-1]}, \quad i = 1, 2, \ldots, r, \end{cases} \tag{2.2.2}$$

$n = 0, 1, \ldots, N$, where s is the number of internal stages and r is the number of external stages, which propagate from step to step. Here, h is a stepsize, $Y_i^{[n]}$ is an approximation (possibly of low order) to $y(t_{n-1} + c_i h)$, and $y_i^{[n]}$ is an approximation to the linear combination of the derivatives of y at the point t_n. As discussed by Butcher and Burrage [15, 18, 21], method (2.2.2) can be represented conveniently by the abscissa vector $\mathbf{c}$ and a partitioned $(s + r) \times (s + r)$ matrix

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right]. \tag{2.2.3}$$

Introducing the notation

$$Y^{[n]} = \begin{bmatrix} Y_1^{[n]} \\ Y_2^{[n]} \\ \vdots \\ Y_s^{[n]} \end{bmatrix}, \qquad F(Y^{[n]}) = \begin{bmatrix} f(Y_1^{[n]}) \\ f(Y_2^{[n]}) \\ \vdots \\ f(Y_s^{[n]}) \end{bmatrix}, \qquad y^{[n]} = \begin{bmatrix} y_1^{[n]} \\ y_2^{[n]} \\ \vdots \\ y_r^{[n]} \end{bmatrix},$$

(2.2.2) can be written in the vector form

$$\left[\begin{array}{c} Y^{[n]} \\ \hline y^{[n]} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A}\otimes\mathbf{I} & \mathbf{U}\otimes\mathbf{I} \\ \hline \mathbf{B}\otimes\mathbf{I} & \mathbf{V}\otimes\mathbf{I} \end{array}\right] \left[\begin{array}{c} hF(Y^{[n]}) \\ \hline y^{[n-1]} \end{array}\right].$$

Here $\mathbf{I}$ is the identity matrix of dimension m, and the Kronecker product of two matrices $\mathbf{A} \in \mathbb{R}^{m_1 \times n_1}$ and $\mathbf{B} \in \mathbb{R}^{m_2 \times n_2}$ is defined as a block matrix of the form

$$\mathbf{A}\otimes\mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1,n_1}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2,n_1}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{m_1,n_1}\mathbf{B} \end{bmatrix} \in \mathbb{R}^{m_1 m_2 \times n_1 n_2}.$$

These methods include as special cases many known methods for ODEs.

2.3 Classical methods as GLMs

In this section we represent the classical methods as GLMs. For example, Runge-Kutta (RK) methods given by

$$\begin{cases} Y_i^{[n]} = y_{n-1} + h\sum_{j=1}^{s} a_{ij} f(Y_j^{[n]}), \quad i = 1, 2, \ldots, s \\ y_n = y_{n-1} + h\sum_{j=1}^{s} b_j f(Y_j^{[n]}), \end{cases} \tag{2.3.4}$$

$n = 0, 1, \ldots, N$, or by the Butcher tableau

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b^T} \end{array} = \begin{array}{c|ccc} c_1 & a_{11} & \ldots & a_{1s} \\ c_2 & a_{21} & \ldots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & \ldots & a_{ss} \\ \hline & b_1 & \ldots & b_s \end{array}$$

can be represented as GLM (2.2.2) with $r = 1$ in the form

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{e} \\ \hline \mathbf{b^T} & 1 \end{array}\right] = \left[\begin{array}{ccc|c} a_{11} & \ldots & a_{1s} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{s1} & \ldots & a_{ss} & 1 \\ \hline b_1 & \ldots & b_s & 1 \end{array}\right].$$

Similarly, the class of linear multistep methods

$$y_n = \sum_{j=1}^{k} \alpha_j y_{n-j} + h \sum_{j=0}^{k} \beta_j f(y_{n-j}), \tag{2.3.5}$$

$n = k, k+1, \ldots, N$, can be written for scalar ODEs as GLM (2.2.2) with $s = 1$ and $r = 2k$ in the form

$$\left[\begin{array}{c} y_n \\ \hline y_n \\ y_{n-1} \\ \vdots \\ y_{n-k+1} \\ hf(y_n) \\ hf(y_{n-1}) \\ \vdots \\ hf(y_{n-k+1}) \end{array}\right] = \left[\begin{array}{c|cccccccc} \beta_0 & \alpha_1 & \ldots & \alpha_{k-1} & \alpha_k & \beta_1 & \ldots & \beta_{k-1} & \beta_k \\ \hline \beta_0 & \alpha_1 & \ldots & \alpha_{k-1} & \alpha_k & \beta_1 & \ldots & \beta_{k-1} & \beta_k \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1 & 0 \end{array}\right] \left[\begin{array}{c} hf(y_n) \\ \hline y_{n-1} \\ y_{n-2} \\ \vdots \\ y_{n-k} \\ hf(y_{n-1}) \\ hf(y_{n-2}) \\ \vdots \\ hf(y_{n-k}) \end{array}\right].$$

More compact representation is possible for special cases of (2.3.5). For example, Adams methods can be represented with $r = k + 1$ and backward differentiation formulas with $r = k$, compare in [21, 157]. Similarly, all other methods for first order ODEs can be represented as GLM (see [21, 97]).

Consider next the class of two-step Runge-Kutta (TSRK) methods investigated by Jackiewicz and Tracogna in [100]. These methods depend on stage values at two consecutive steps and have the form

$$\begin{cases} Y_i^{[n]} = (1-u_i)y_{n-1} + u_i y_{n-2} + h\sum_{j=1}^{s}\left(a_{ij}f(Y_j^{[n]}) + b_{ij}f(Y_j^{[n-1]})\right), \\ y_n = (1-\vartheta)y_{n-1} + \vartheta y_{n-2} + h\sum_{j=1}^{s}\left(v_j f(Y_j^{[n]}) + w_j f(Y_j^{[n-1]})\right), \end{cases} \tag{2.3.6}$$

for $i = 1, \ldots, s$, where $Y_i^{[n]}$ is an approximation to $y(t_{n-1} + c_i h)$. The method (2.3.6) can be represented by the abscissa vector $\mathbf{c} = [c_1, \ldots, c_s]^T$ and the tableau

$\mathbf{u}$	$\mathbf{A}$	$\mathbf{B}$
ϑ	$\mathbf{v^T}$	$\mathbf{w^T}$

For scalar ODEs these methods can be represented as GLM (2.2.2) with $r = s + 2$, in the form

$$\left[\begin{array}{c} Y^{[n]} \\ \hline y_n \\ y_{n-1} \\ hf(Y^{[n]}) \end{array}\right] = \left[\begin{array}{c|ccc} \mathbf{A} & \mathbf{e}-\mathbf{u} & \mathbf{u} & \mathbf{B} \\ \hline \mathbf{v^T} & 1-\vartheta & \vartheta & \mathbf{w^T} \\ 0 & 1 & 0 & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} hF(Y^{[n]}) \\ \hline y_{n-1} \\ y_{n-2} \\ hf(Y^{[n-1]}) \end{array}\right],$$

compare [157]. Here, $\mathbf{I}$ is the identity matrix of dimension s and $\mathbf{0}$ is zero matrix or vector of appropriate dimensions. Different representation of (2.3.6) as GLMs are considered in [6, 100, 102]. Other examples of methods which can be represented as GLMs include predictor-corrector schemes in various modes, for example $P(EC)^m$ or $P(EC)^m E$ [108, 109], generalized multistep hybrid methods of Butcher [16], Gear [76], and Gragg and Stetter [77], one-leg methods of Dahlquist [64], cyclic composite methods of Donelson and Hansen [68], pseudo Runge-Kutta methods of Byrne

and Lambert [13] and Costabile [43, 44, 45], special cases of two-step Runge-Kutta methods investigated in [136, 98, 103, 99], and 'Almost Runge-Kutta' methods examined in [19, 20, 22].

2.4 Preliminary properties: preconsistency, consistency, zero-stability, convergence

We state the classical properties of numerical methods for ODEs, adapted to GLMs. To identify useful GLMs (2.2.2) we have to impose some accuracy and stability conditions. To find minimal accuracy conditions, we assume that there exist vectors $\mathbf{q_0}$ and $\mathbf{q_1}$

$$\mathbf{q_0} = [q_{1,0}, \ldots, q_{r,0}]^T, \qquad \mathbf{q_1} = [q_{1,1}, \ldots, q_{r,1}]^T,$$

such that the components of the input vector $y^{[n-1]}$ satisfy

$$y_i^{[n-1]} = q_{i,0}y(t_{n-1}) + q_{i,1}hy'(t_{n-1}) + O(h^2), \quad i = 1, 2, \ldots, r.$$

We then request that the components of the stage vector $Y^{[n]}$ and the output vector $y^{[n]}$ satisfy

$$Y_i^{[n]} = y(t_{n-1} + c_ih) + O(h^2), \quad i = 1, 2, \ldots, s,$$

and

$$y_i^{[n]} = q_{i,0}y(t_n) + q_{i,1}hy'(t_n) + O(h^2), \quad i = 1, 2, \ldots, r.$$

Observe that the condition for $Y_i^{[n]}$ is more general than the condition

$$Y_i^{[n]} = y(t_n) + O(h), \quad i = 1, 2, \ldots, s,$$

considered in [14, 21]. Substituting these relations into (2.2.2), we obtain

$$y(t_{n-1}) + hc_i y'(t_{n-1}) = h\sum_{j=1}^{s} a_{ij} y'(t_{n-1}) + \sum_{j=1}^{r} u_{ij}\left(q_{j,0} y(t_{n-1}) + hq_{j,1} y'(t_{n-1})\right) + O(h^2),$$

$i = 1, 2, \ldots, s$, and

$$q_{i,0} y(t_n) + q_{i,1} hc_i y'(t_n) = h\sum_{j=1}^{s} b_{ij} y'(t_n) + \sum_{j=1}^{r} v_{ij}\left(q_{j,0} y(t_{n-1}) + hq_{j,1} y'(t_{n-1})\right) + O(h^2),$$

$i = 1, 2, \ldots, r$. Comparing $O(1)$ and $O(h)$ terms in these relations, we obtain

$$\sum_{j=1}^{r} u_{ij} q_{j,0} = 1, \quad i = 1, 2, \ldots, s, \quad \sum_{j=1}^{r} v_{ij} q_{j,0} = q_{i,0}, \quad i = 1, 2, \ldots, r,$$

and

$$\sum_{j=1}^{s} a_{ij} + \sum_{j=1}^{r} u_{ij} q_{j,1} = c_i, \quad i = 1, 2, \ldots, s,$$
$$\sum_{j=1}^{s} b_{ij} + \sum_{j=1}^{r} v_{ij} q_{j,1} = q_{i,0} + q_{i,1}, \quad i = 1, 2, \ldots, r.$$

The above considerations motivate the following definitions.

Definition 2.4.1. *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is preconsistent if there exists a vector* $q_0 \in \mathbb{R}^r$ *such that*

$$\mathbf{U}\mathbf{q}_0 = \mathbf{e}, \quad \mathbf{V}\mathbf{q}_0 = \mathbf{q}_0,$$

where $\mathbf{e} = [1, \ldots, 1] \in \mathbb{R}^r$. *The vector* $\mathbf{q}_0$ *is called the preconsistency vector.*

Definition 2.4.2. *A preconsistent GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is consistent if there exists a vector* $q_1 \in \mathbb{R}^r$ *such that*

$$\mathbf{Be} + \mathbf{Vq}_1 = \mathbf{q}_0 + \mathbf{q}_1,$$

where $\mathbf{e} = [1, \ldots, 1] \in \mathbb{R}^r$. *The vector* $\mathbf{q}_1$ *is called the consistency vector.*

Definition 2.4.3. *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is stage-consistent if*

$$\mathbf{Ae} + \mathbf{Uq}_1 = \mathbf{c},$$

with $\mathbf{e} = [1, \ldots, 1] \in \mathbb{R}^s$.

It can be verified that for an $s-$stage Runge-Kutta method (2.3.4) $\mathbf{q_0} = 1$, $\mathbf{q_1} = 0$ and the stage-consistency and consistency conditions are equivalent to

$$\mathbf{Ae} = \mathbf{c}, \quad \mathbf{b^T e} = 1.$$

The condition $\mathbf{Ae} = \mathbf{c}$ is not necessary and RK methods which do not satisfy it are investigated in [121, 159], see also [67]. In the case of linear multistep methods (2.3.5)

$$\mathbf{q_0} = [1, \ldots, 1|0, \ldots, 0]^T \in \mathbb{R}^{2k}, \quad \mathbf{q_1} = [0, -1 \ldots, -k+1|1, \ldots, 1]^T \in \mathbb{R}^{2k},$$

and the stage-consistency and consistency conditions take the form

$$\mathbf{c} = 1, \quad \sum_{j=1}^{k} \alpha_j = 1, \quad \sum_{j=1}^{k} j\alpha_j = \sum_{j=0}^{k} \beta_j,$$

compare [108, 109].

For the class of TSRK methods (2.3.6) we have

$$\mathbf{q_0} = [1\,1|0 \ldots 0]^T \in \mathbb{R}^{s+2}, \quad \mathbf{q_1} = [0\ -1|1 \ldots 1]^T \in \mathbb{R}^{s+2},$$

and the stage-consistency and consistency conditions take the form

$$(\mathbf{A}+\mathbf{B})\mathbf{e}-\mathbf{u}=\mathbf{c}, \quad (\mathbf{v}^{\mathbf{T}}+\mathbf{w}^{\mathbf{T}})\mathbf{e}=1+\vartheta,$$

compare [100]. To find minimal stability conditions we apply GLM (2.2.2) to the equation $y'=0$, $t \geq t_0$. This leads to

$$y^{[n]}=\mathbf{V}y^{[n-1]}=\mathbf{V}^n y^{[0]}, \quad n=0,1,\ldots,$$

and motivates the following definition.

Definition 2.4.4. *A GLM* $(\mathbf{c},\mathbf{A},\mathbf{U},\mathbf{B},\mathbf{V})$ *is zero-stable if there exists a constant* C *such that*

$$\|\mathbf{V}^n\| \leq C,$$

for all $n \geq 0$.

It is well known that the condition (2.4.4) is equivalent to the following criterion (compare [18, 14]).

Theorem 2.4.1. *A GLM* $(\mathbf{c},\mathbf{A},\mathbf{U},\mathbf{B},\mathbf{V})$ *is zero-stable if the minimal polynomial of the coefficient matrix* $\mathbf{V}$ *has no zeros with magnitude greater than* 1 *and all zeros with magnitude equal to* 1 *are simple.*

For RK methods (2.3.4) the minimal polynomial is $p(\omega)=\omega-1$ and these methods are always zero-stable. It can be verified that for linear multistep methods (2.3.5), the minimal polynomial is $p(\omega)=\omega^k\rho(\omega)$, where

$$\rho(\omega)=\omega^k-\sum_{j=1}^{k}\alpha_j\omega^{k-j},$$

is the first characteristic polynomial of (2.3.5). Hence, these methods are stable if $\rho(\omega)$ satisfies the root condition. This means that $\rho(\omega)$ has no root with modulus

greater than one, and every root with modulus one is simple, compare [108, 109]. For the TSRK methods (2.3.6) the minimal polynomial is

$$p(\omega) = \omega(\omega^2 - (1 - \vartheta)\omega - \vartheta).$$

The roots of this polynomial are $\omega = 0$, $\omega = 1$ and $\omega = -\vartheta$, and it follows that (2.3.6) is zero-stable if $-1 < \vartheta \leq 1$, compare [100].
Basic requirements in the context of the numerical integration of ODEs are, together with consistency and zero-stability, also convergence, which are defined for GLMs in the following way (see [21, 97]).

Definition 2.4.5. *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is convergent if, for any initial value problem* (2.2.1) *satsfying the Lipschitz condition*

$$\|f(y) - f(z)\| \leq L \|y - z\|,$$

there exists a nonzero vector $\mathbf{q}_0 \in \mathbb{R}^r$ *and a starting procedure* $S : (0, \infty) \to \mathbb{R}^r$ *such that, for every stepsize* $h > 0$, $\lim_{h \to 0} S(h) = \mathbf{q}_0 y(t_0)$, *and such that, for any* $\bar{t} > t_0$, *the sequence of vectors* $y^{[n]}$, *computed using* n *steps with stepsize* $h = (\bar{t} - t_0)/n$ *and using* $y^{[0]} = S(h)$ *in each case, converges to* $\mathbf{q}_0 y(\bar{t})$.

Theorem 2.4.2. *A GLM* $(\mathbf{c}, \mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V})$ *is convergent if and only if it is consistent and zero-stable.*

2.5 Order and stage order conditions

To formulate the stage order and order conditions for GLMs (2.2.2) we assume that the components of the input vector $y_i^{[n-1]}$ for the next step satisfy

$$y_i^{[n]} = \sum_{k=0}^{p} q_{ik} h^k y^{(k)}(t_n) + O(h^{p+1}), \quad i = 1, 2, \ldots, r, \tag{2.5.1}$$

for some real parameters q_{ik}, $i = 1, 2, \ldots, r, \quad k = 0, 1, \ldots, p$. We then request that the components of the internal stages $Y_i^{[n]}$ are approximations of order $q \geq p - 1$ to the solution $y(t)$ of (2.2.1) at the points $t_{n-1} + c_i h$, i.e.,

$$Y_i^{[n]} = y(t_n + c_i h) + O(h^{p+1}), \quad i = 1, 2, \ldots, s, \tag{2.5.2}$$

and that the components of the output vector $y_i^{[n]}$ satisfy

$$y_i^{[n+1]} = \sum_{k=0}^{p} q_{ik} h^k y^{(k)}(t_{n+1}) + O(h^{p+1}), \quad i = 1, 2, \ldots, r. \tag{2.5.3}$$

The integers q and p are called the stage order and order, respectively, of GLM (2.2.2). We collect the parameters q_{ik} appearing in (2.5.1) and (2.5.3) in the vectors $\mathbf{q}_k$ defined by

$$\mathbf{q}_k = [q_{1k}, q_{2k}, \ldots, q_{rk}]^T \in \mathbb{R}^r, \quad k = 0, 1, \ldots, p.$$

We also introduce the notation $e^{cz} = [e^{c_1 z}, e^{c_2 z}, \ldots, e^{c_s z}]$, and define the vector $\mathbf{w}(z)$ by

$$\mathbf{w}(z) = \sum_{k=0}^{p} \mathbf{q}_k z^k, \quad z \in \mathbb{C}.$$

Here, $\mathbb{C}$ is the set of complex numbers. We have the following theorem

Theorem 2.5.1. *(Butcher [21]) Assume that $y^{[n-1]}$ satisfies* (2.5.1)*. Then the GLM* (2.2.2) *of order p and stage order $q = p$ satisfies* (2.5.2) *and* (2.5.3) *if and only if*

$$e^{cz} = z\mathbf{A}e^{cz} + \mathbf{U}\mathbf{w}(z) + O(z^{p+1}), \tag{2.5.4}$$

and

$$e^{z}\mathbf{w} = z\mathbf{B}e^{cz} + \mathbf{V}\mathbf{w}(z) + O(z^{p+1}). \tag{2.5.5}$$

Expanding e^{cz} and e^z in (2.5.4) *and* (2.5.5) *into power series around $z = 0$ and comparing the constant terms in the resulting expressions we obtain the preconsistency*

conditions

$$\mathbf{U}\mathbf{q}_0 = \mathbf{e}, \quad \mathbf{V}\mathbf{q}_0 = \mathbf{q}_0,$$

where $\mathbf{e} = [1, 1, \ldots, 1] \in \mathbb{R}^{s+2}$. *Comparing the terms of order* z^k, $k = 1, 2, \ldots, p$ *in the resulting expressions, the stage order and order conditions can be reformulated in the form*

$$\frac{c^k}{k!} - \frac{\mathbf{A}c^{k-1}}{(k-1)!} - \mathbf{U}\mathbf{q}_k = 0, \quad k = 1, 2, \ldots, p,$$

and

$$\sum_{l=0}^{k} \frac{\mathbf{q}_{k-l}}{l!} - \frac{\mathbf{B}c^{k-1}}{(k-1)!} - \mathbf{V}\mathbf{q}_k = 0, \quad k = 1, 2, \ldots, p.$$

In corrispondence of $k = 1$, *the stage-consistency and consistency conditions* (2.4.3) *and* (2.4.2) *result from* (2.5.1) *and* (2.5.1), *respectively.*

2.6 Linear stability theory

Linear stability properties are classically provided with respect to the scalar linear test equation

$$y' = \xi y, \quad t \geq 0, \tag{2.6.1}$$

where $\xi \in \mathbb{C}$ and $Re(\xi) \neq 0$. The solution of this simple problem remains bounded when t goes to infinity and we need to require that the numerical solution possesses an analogous stability property to that displayed by the exact solution: let us analyse the conditions to be imposed on the numerical method in order to reproduce the same behaviour of the exact solution. Applying the GLM (2.2.2) to the linear test equation (2.6.1), we obtain the following recurrence relation

$$y^{[n]} = \mathbf{M}(z)y^{[n-1]},$$

$n = 1, 2, \ldots, z = h\xi$. Here, $\mathbf{M}(z) \in \mathbb{C}^{r\times r}$ is the so-called stability matrix, which takes the form

$$\mathbf{M}(z) = \mathbf{V} + z\mathbf{B}(\mathbf{I} - z\mathbf{A})^{-1}\mathbf{U}. \tag{2.6.2}$$

The characteristic polynomial of the stability matrix

$$p(\eta, z) = \det(\eta\mathbf{I} - \mathbf{M}(z)), \tag{2.6.3}$$

is said stability polynomial. It is a polynomial of degree r with respect to η. Denote by $\eta_1(z), \eta_2(z), \ldots, \eta_r(z)$ the roots of the stability function $p(\eta, z)$. The following definitions arise.

Definition 2.6.1. *A GLM* (2.2.2) *is absolutely stable if, for a given $z \in \mathbb{C}$, all the roots $\eta_1(z), \eta_2(z), \ldots, \eta_r(z)$ of the stability polynomial* (2.6.3) *lie in the unit circle.*

Definition 2.6.2. *The region $\mathcal{A}$ of absolute stability of* (2.2.2) *is the set*

$$\mathcal{A} = \{z \in \mathbb{C} : |\eta_i(z)| < 1, i = 1, 2, \ldots, r\}.$$

Definition 2.6.3. *The GLM* (2.2.2) *is said to be A-stable if its region of absolute stability includes the negative complex plane, i.e.*

$$\{z \in \mathbb{C} : Re(z) < 0\} \subset \mathcal{A}.$$

The derivation of A-stable methods is, in general, a nontrivial task. However, some helpful tools have been introduced in the literature in order to provide some useful criteria to study A-stability. In order to achieve A-stability, all the roots $\eta_1(z), \eta_2(z), \ldots, \eta_r(z)$ of the polynomial $p(\eta, z)$ defined by (2.6.3) has to lie in the unit circle for all $z \in \mathbb{C}$ with $Re(z) \leq 0$. By the maximum principle this will be the case if the denominator of $p(\eta, z)$ does not have poles in the negative half plane $\mathbb{C}$ and if the roots of $p(\eta, z)$ are in the unit circle for all $y \in \mathbb{R}$. This last condition can be investigated using the Schur criterion. This criterion, for a general k^{th} degree

polynomial, can be formulated as follows. Consider the polynomial

$$\phi(\omega) = d_k\omega^k + d_{k-1}\omega^{k-1} + \ldots + d_1\omega + d_0,$$

where d_i are complex coefficients, with $d_k \neq 0$ and $d_0 \neq 0$. $\phi(\omega)$ is said to be a Schur polynomial if all its roots ω_i, $i = 1, 2, \ldots, k$, are inside of the unit circle. Define

$$\widehat{\phi}(\omega) = \overline{d}_0\omega^k + \overline{d}_1\omega^{k-1} + \ldots + \overline{d}_{k-1}\omega + \overline{d}_k,$$

where $\overline{d}_i$ is the complex conjugate of d_i. Define also the polynomial

$$\phi_1(\omega) = \frac{1}{\omega}\left(\widehat{\phi}(0)\phi(\omega) - \phi(0)\widehat{\phi}(\omega)\right),$$

of degree at most $k-1$. We have the following theorem.

Theorem 2.6.1. *$\phi(\omega)$ is a Schur polynomial if and only if*

$$\left|\widehat{\phi}(0)\right| > |\phi(0)|,$$

and $\phi_1(\omega)$ is a Schur polynomial.

Definition 2.6.4. *A GLM* (2.2.2) *is said to be L-stable if it is A-stable and, in addition,*

$$\lim_{z\to\infty} \rho(\mathbf{M}(z)) = 0, \tag{2.6.4}$$

where $\rho(\mathbf{M}(z))$ denotes the spectral radius of the stability matrix $\mathbf{M}(z)$.

Once A-stability is obtained, L-stability is achieved by imposing that all the roots of the polynomial $p(\eta, z)/p_r(z)$, where $p(\eta, z)$ is given by (2.6.3), tend to zero as $z \to -\infty$.

Therefore, such methods satisfy the nonlinear system of equations

$$\begin{cases} \lim\limits_{z\to-\infty} \dfrac{p_0(z)}{p_r(z)} = 0, \\ \lim\limits_{z\to-\infty} \dfrac{p_{r-1}(z)}{p_r(z)} = 0. \end{cases} \tag{2.6.5}$$

For a complete discussion of GLMs for first order ODEs, see [21, 97].

3

General Linear Methods for second order ordinary differential equations

3.1 Introduction to GLMs for second order ODEs

It is the purpose of this chapter to introduce a general family of numerical methods suited to numerically integrate initial value problems based on special second order Ordinary Differential Equations (1.1.1).

As already discussed in the introduction, for second order ODEs (1.1.1) no systematic investigation on GLMs has begun till now: even if many linear and nonlinear methods appeared in the literature (see, for instance, [82, 83, 85, 94] and references therein contained) there is not yet a very wide and general class of methods for the numerical solution of the problem (1.1.1) together with a series of general theorems which provide an unifying framework to analyze the properties of such methods. In order to transfer to second order ODEs of the same benefits obtained in the case of first order ODEs, the purpose of this work is the foundation of a theory of GLMs for the numerical solution of (1.1.1). Such a theory would allow the introduction of general theorems concerning, for instance, zero-stability, convergence, order conditions, and many other properties which have been studied in the literature on the case by case basis.

The chapter is organized as follows: Section 2 presents the formulation of GLMs for (1.1.1), while the representation of classical methods regarded as GLMs is reported in Section 3; Section 4 is devoted to the introduction of the concept of consistency for GLM; the study of zero-stability for GLMs is carried out in Section 5 where, together with the definition of zero-stability, a criterion to analyze zero-stability is also provided; in Section 6 the convergence analysis of GLMs is treated and a useful result to assert if a certain GLM is convergent or not is provided; in Section 7 order conditions for GLMs are discussed; Section 8 contains the analysis of linear stability, while in Section 9 examples of new methods are provided.

3.2 Representation of GLMs

In this section, we aim to introduce a general representation formula to GLMs for second order ODEs (1.1.1), in order to properly embrace a wide number of classical numerical methods for (1.1.1), but also allowing the introduce wider and more general classes of methods than those already considered in the literature. To this purpose, we consider the uniform grid

$$I_h = \{t_n = t_0 + nh,\ n = 0, 1, ..., N,\ Nh = T - t_0\},$$

which constitutes the discrete counterpart of the interval of the definition I of the problem (1.1.1) and assume as a point of reference of our analysis the family of GLMs introduced in the context of first order ODEs (compare [21, 97] and references therein contained), i.e.

$$\begin{cases} Y_i^{[n]} = \sum_{j=1}^{s} a_{ij} h f(Y_j^{[n]}) + \sum_{j=1}^{r} u_{ij} y_j^{[n-1]}, & i = 1, 2, \ldots, s, \\ y_i^{[n]} = \sum_{j=1}^{s} b_{ij} h f(Y_j^{[n]}) + \sum_{j=1}^{r} v_{ij} y_j^{[n-1]}, & i = 1, 2, \ldots, r. \end{cases} \tag{3.2.1}$$

This formulation was introduced by Burrage and Butcher [15] in 1980. In order to adapt such formulation to second order ODEs (1.1.1) and achieve the mentioned purpose of gaining a very general class of numerical methods to solve this problem, we inherit the same structure as in (3.2.1) but also include the dependence on the approximations to the first derivative of the solution.

In order to carry out our theory of GLMs for second order ODEs (1.1.1), we introduce the abscissa vector $\mathbf{c} = [c_1, c_2, \ldots, c_s]$ and define the following supervectors

$$y^{[n-1]} = \begin{bmatrix} y_1^{[n-1]} \\ y_2^{[n-1]} \\ \vdots \\ y_r^{[n-1]} \end{bmatrix} \in \mathbb{R}^{rd},\ y'^{[n-1]} = \begin{bmatrix} y'^{[n-1]}_1 \\ y'^{[n-1]}_2 \\ \vdots \\ y'^{[n-1]}_{r'} \end{bmatrix} \in \mathbb{R}^{r'd},\ Y^{[n]} = \begin{bmatrix} Y_1^{[n]} \\ Y_2^{[n]} \\ \vdots \\ Y_s^{[n]} \end{bmatrix} \in \mathbb{R}^{sd}.$$

The vector $y^{[n-1]}$ is denoted as *input vector* of the external stages, and contains all the informations we want to transfer advancing from the point t_{n-1} to the point t_n of the grid. It is important to observe that such vector could also contain not only approximations to the solution of the problem in the grid points inherited from the previous steps, but also other informations computed in the past that we want to use in the integration process. The vector $y'^{[n-1]}$ instead contains previous approximations to the first derivative of the solution computed in previous step points, while the values $Y_j^{[n-1]}$, denoted as *internal stage* values, provide an approximation to the solution in the internal points $t_{n-1} + c_j h$, $j = 1, 2, \ldots, s$.

Our formulation of GLMs for second order ODEs then involves nine coefficient matrices $\mathbf{A} \in \mathbb{R}^{\mathbf{s}\times\mathbf{s}}$, $\mathbf{P} \in \mathbb{R}^{\mathbf{s}\times\mathbf{r}'}$, $\mathbf{U} \in \mathbb{R}^{\mathbf{s}\times\mathbf{r}}$, $\mathbf{C} \in \mathbb{R}^{\mathbf{r}'\times\mathbf{s}}$, $\mathbf{R} \in \mathbb{R}^{\mathbf{r}'\times\mathbf{r}'}$, $\mathbf{W} \in \mathbb{R}^{\mathbf{r}'\times\mathbf{r}}$, $\mathbf{B} \in \mathbb{R}^{\mathbf{r}\times\mathbf{s}}$, $\mathbf{Q} \in \mathbb{R}^{\mathbf{r}\times\mathbf{r}'}$, $\mathbf{V} \in \mathbb{R}^{\mathbf{r}\times\mathbf{r}}$, which are put together in the following partitioned

$(s + r' + r) \times (s + r' + r)$ matrix

$$\left[\begin{array}{c|c|c} \mathbf{A} & \mathbf{P} & \mathbf{U} \\ \hline \mathbf{C} & \mathbf{R} & \mathbf{W} \\ \hline \mathbf{B} & \mathbf{Q} & \mathbf{V} \end{array}\right], \tag{3.2.2}$$

which is denoted as the Butcher tableau of the GLM. Using these notations, a GLM for second order ODEs can then be expressed as follows

$$\begin{aligned} Y^{[n]} &= h^2(\mathbf{A} \otimes \mathbf{I})F^{[n]} + h(\mathbf{P} \otimes \mathbf{I})y'^{[n-1]} + (\mathbf{U} \otimes \mathbf{I})y^{[n-1]}, \\ hy'^{[n]} &= h^2(\mathbf{C} \otimes \mathbf{I})F^{[n]} + h(\mathbf{R} \otimes \mathbf{I})y'^{[n-1]} + (\mathbf{W} \otimes \mathbf{I})y^{[n-1]}, \\ y^{[n]} &= h^2(\mathbf{B} \otimes \mathbf{I})F^{[n]} + h(\mathbf{Q} \otimes \mathbf{I})y'^{[n-1]} + (\mathbf{V} \otimes \mathbf{I})y^{[n-1]}, \end{aligned} \tag{3.2.3}$$

where $\otimes$ denotes the usual Kronecker tensor product, $\mathbf{I}$ is the identity matrix in $\mathbb{R}^{d\times d}$ and $F^{[n]} = [f(Y_1^{[n]}), f(Y_2^{[n]}), \ldots, f(Y_s^{[n]})]^T$. Componentwise,

$$\begin{aligned} Y_i^{[n]} &= h^2 \sum_{j=1}^{s} a_{ij} f(Y_j^{[n]}) + h \sum_{j=1}^{r'} p_{ij} y'^{[n-1]}_j + \sum_{j=1}^{r} u_{ij} y_j^{[n-1]} \quad i = 1, ..., s, \\ hy'^{[n]}_i &= h^2 \sum_{j=1}^{s} c_{ij} f(Y_j^{[n]}) + h \sum_{j=1}^{r'} r_{ij} y'^{[n-1]}_j + \sum_{j=1}^{r} w_{ij} y_j^{[n-1]} \quad i = 1, ..., r', \\ y_i^{[n]} &= h^2 \sum_{j=1}^{s} b_{ij} f(Y_j^{[n]}) + h \sum_{j=1}^{r'} q_{ij} y'^{[n-1]}_j + \sum_{j=1}^{r} v_{ij} y_j^{[n-1]} \quad i = 1, ..., r. \end{aligned} \tag{3.2.4}$$

An important role in the application of GLMs is played by the matrix $\mathbf{A}$ of the tableau (3.2.2), whose structure determines the computational cost of the corresponding GLM. In fact, if $\mathbf{A}$ is a full matrix, the first formula in (3.2.3) is a nonlinear system in the internal stages to be solved at each step. However, the solution of linear

and nonlinear systems of equations can be efficiently computed if their coefficient matrix shows a structured shape. In this case, some function evaluations can be avoided or the Jacobian of the system can be stored and re-used for a certain number of iterations or even a fast computation (e.g. in a parallel environment) can be provided. According to this analysis and to the classical literature existing for GLMs solving first order ODEs (compare [21, 97]), we distinguish the following cases:

- the matrix $\mathbf{A}$ is full. In this case, the corresponding GLM is *fully implicit*;
- the matrix $\mathbf{A}$ is lower triangular. In this case, the corresponding GLM is *diagonally implicit*. Diagonally implicit GLMs depending on a matrix $\mathbf{A}$ which has also a one-point spectrum will be denoted as *type 1* GLMs;
- the matrix $\mathbf{A}$ is strictly lower triangular. In this case, the corresponding GLM is *explicit* and will also be denoted as *type 2* GLM.

If, moreover, the matrix $\mathbf{A}$ is diagonal, the corresponding GLM is said to be of *type 3*. A less relevant case occurs when $\mathbf{A}$ is the zero matrix: we denote GLMs having this property as *type 4* GLMs.
As Henrici observed in [85], "if one is not particularly interested in the values of the first derivatives, it seems unnatural to introduce them artificially". For this reason, we also consider as special case the following class of methods, which can be regarded as a subclass of GLMs (3.2.3) under the assumption that the weights of the first derivative appearing in each equation of (3.2.3) are equal to zero, i.e. the matrices $\mathbf{P}, \mathbf{Q}, \mathbf{C}, \mathbf{R}, \mathbf{W}$ are all equal to the zero matrix. As a consequence, the corresponding methods assume the form

$$\begin{aligned} Y_i^{[n]} &= h^2 \sum_{j=1}^{s} a_{ij} f(Y_j^{[n]}) + \sum_{j=1}^{r} u_{ij} y_j^{[n-1]}, \quad i = 1, ..., s, \\ y_i^{[n]} &= h^2 \sum_{j=1}^{s} b_{ij} f(Y_j^{[n]}) + \sum_{j=1}^{r} v_{ij} y_j^{[n-1]}, \quad i = 1, ..., r. \end{aligned} \tag{3.2.5}$$

It is evident from (3.2.5) that, when the method does not explicitly depend on the first derivative approximations, the matrices $\mathbf{P}, \mathbf{Q}, \mathbf{C}, \mathbf{R}, \mathbf{W}$ do not provide any contribution in the computation of the numerical solution to the problem and, therefore, the tableau of this GLM can obviously be reduced to

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right], \tag{3.2.6}$$

where $\mathbf{A} \in \mathbb{R}^{\mathbf{s}\times\mathbf{s}}, \mathbf{U} \in \mathbb{R}^{\mathbf{s}\times\mathbf{r}}, \mathbf{B} \in \mathbb{R}^{\mathbf{r}\times\mathbf{s}}, \mathbf{V} \in \mathbb{R}^{\mathbf{r}\times\mathbf{r}}$. We again observe that such methods form a class of GLMs (3.2.3) and, therefore, all the results we will next prove concerning GLMs also trivially hold for the subclass (3.2.5).

3.3 Classical methods regarded as GLMs

The family of GLMs for second order ODEs properly contains as special cases all the numerical methods for second order ODEs already introduced in the literature. This is made clear in the following examples.

3.3.1 Linear multistep methods

Linear multistep methods (compare [82, 85]), defined by

$$y_n = \sum_{j=1}^{k} \alpha_j y_{n-j} + h^2 \sum_{j=0}^{k} \beta_j f(y_{n-j}), \tag{3.3.1}$$

can be regarded as GLMs (3.2.5) with $r = 2k$, $s = 1$, $Y^{[n]} = [y_n]$,

$$y^{[n-1]} = [y_{n-1}, y_{n-2}, \ldots, y_{n-k}, h^2 f(y_{n-1}), h^2 f(y_{n-2}), \ldots, h^2 f(y_{n-k})]^T,$$

and in correspondence to the reduced tableau (3.2.6)

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{c|cccccccc} \beta_0 & \alpha_1 & \dots & \alpha_{k-1} & \alpha_k & \beta_1 & \dots & \beta_{k-1} & \beta_k \\ \hline \beta_0 & \alpha_1 & \dots & \alpha_{k-1} & \alpha_k & \beta_1 & \dots & \beta_{k-1} & \beta_k \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 1 & \dots & 1 & 0 \end{array}\right],$$

with $\mathbf{c} = [1]$. A famous example of linear multistep method is the Numerov method (see, for instance, [82, 94])

$$y_n = 2y_{n-1} - y_{n-2} + h^2 \left(\frac{1}{12} f(t_n, y_n) + \frac{5}{6} f(t_{n-1}, y_{n-1}) + \frac{1}{12} f(t_{n-2}, y_{n-2}) \right), \quad (3.3.2)$$

which is an order four method corresponding to the GLM (3.2.5) with $r = 4$, $s = 1$, $Y^{[n]} = [y_n]$, and

$$y^{[n-1]} = \left[y_{n-1}, y_{n-2}, h^2 f(y_{n-1}), h^2 f(y_{n-2}) \right]^T,$$

and the tableau (3.2.6)

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{c|cccc} \frac{1}{12} & 2 & -1 & \frac{5}{6} & \frac{1}{12} \\ \hline \frac{1}{12} & 2 & -1 & \frac{5}{6} & \frac{1}{12} \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

3.3.2 Runge-Kutta-Nyström methods

Runge-Kutta-Nyström methods (see [82])

$$\begin{aligned} Y_i &= y_{n-1} + c_i h y'_{n-1} + h^2 \sum_{j=1}^{s} a_{ij} f\left(Y_j\right), \quad i = 1, ..., s, \\ hy'_n &= hy'_{n-1} + h^2 \sum_{j=1}^{s} b'_j f\left(Y_j\right), \\ y_n &= y_{n-1} + hy'_{n-1} + h^2 \sum_{j=1}^{s} b_j f\left(Y_j\right), \end{aligned} \tag{3.3.3}$$

provide an extension to second order ODEs (1.1.1) of Runge–Kutta methods (compare, for instance, [18, 109]) and involve the dependence on the approximation to the first derivative in the current grid point. Such methods can be recasted as GLMs (3.2.3) with $r = 1$, in correspondence to the tableau (3.2.2)

$$\left[\begin{array}{c|c|c} \mathbf{A} & \mathbf{P} & \mathbf{U} \\ \hline \mathbf{C} & \mathbf{R} & \mathbf{W} \\ \hline \mathbf{B} & \mathbf{Q} & \mathbf{V} \end{array}\right] = \left[\begin{array}{c|c|c} \mathbf{A} & \mathbf{c} & \mathbf{e} \\ \hline \mathbf{b}'^T & 1 & 0 \\ \hline \mathbf{b}^T & 1 & 1 \end{array}\right],$$

where $\mathbf{e}$ is the unit vector in $\mathbb{R}^s$, and the input vectors $y^{[n-1]} = [y_{n-1}]$, $y'^{[n-1]} = [y'_{n-1}]$.

3.3.3 Coleman hybrid methods

We now consider the following class of methods

$$\begin{aligned} Y_i &= (1+c_i)y_{n-1} - c_i y_{n-2} + h^2 \sum_{j=1}^{s} a_{ij} f\left(Y_j\right), \quad i = 1, ..., s, \\ y_n &= 2y_{n-1} - y_{n-2} + h^2 \sum_{j=1}^{s} b_j f\left(Y_j\right), \end{aligned} \tag{3.3.4}$$

introduced by Coleman in [32], which are denoted as two-step hybrid methods. Such methods (3.3.4) can be regarded as GLMs (3.2.5), corresponding to the reduced tableau (3.2.6)

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{c|cc} \mathbf{A} & \mathbf{e}+\mathbf{c} & -\mathbf{c} \\ \hline \mathbf{b}^T & 2 & -1 \\ \mathbf{0} & 1 & 0 \end{array}\right],$$

and characterized by the the input vector $y^{[n-1]} = [y_{n-1} \quad y_{n-2}]^T$.

3.3.4 Two-step Runge-Kutta-Nyström methods

Another interesting class of numerical methods for second order ODEs is given by the family of two-step Runge-Kutta-Nyström methods,

$$\begin{aligned} Y_i^{[n-1]} &= y_{n-2} + hc_i y'_{n-2} + h^2 \sum_{j=1}^{s} a_{ij} f(Y_j^{[n-1]}), \quad i = 1, \ldots, s, \\ Y_i^{[n]} &= y_{n-1} + hc_i y'_{n-1} + h^2 \sum_{j=1}^{s} a_{ij} f(Y_j^{[n]}), \quad i = 1, \ldots, s, \end{aligned}$$

$$
\begin{aligned}
hy'_n &= (1-\theta)hy'_{n-1} + \theta hy'_{n-2} + h^2 v'_j f(Y_j^{[n-1]}) + h^2 w'_j f(Y_j^{[n]}),\\
y_n &= (1-\theta)y_{n-1} + \theta y_{n-2} + h\sum_{j=1}^{s} v'_j y'_{n-2} + h\sum_{j=1}^{s} w'_j y'_{n-1}\\
&\quad + h^2\sum_{j=1}^{s} v_j f(Y_j^{[n-1]}) + h^2\sum_{j=1}^{s} w_j f(Y_j^{[n]}),
\end{aligned}
\tag{3.3.5}
$$

introduced and analyzed by Paternoster in [127, 129, 130, 131]. Such methods depend on two consecutive approximations to the solution and its first derivative in the grid points, but also on two consecutive approximations to the stage values (i.e. the ones related to the points $t_{n-2}+c_ih$ and the ones corresponding to the points $t_{n-1}+c_ih$, $i=1,2,\ldots,s$). This choice is in line with the idea introduced by Jackiewicz et al. (compare [97, 98, 100]) in the context of two-step Runge–Kutta methods for first order ODEs. Two-step Runge-Kutta-Nyström methods can be represented as GLMs (3.2.4) with $r=s+2$ and $r'=2$ through the tableau (3.2.2)

$$
\left[\begin{array}{c|c|c}
\mathbf{A} & \mathbf{P} & \mathbf{U}\\ \hline
\mathbf{C} & \mathbf{R} & \mathbf{W}\\ \hline
\mathbf{B} & \mathbf{Q} & \mathbf{V}
\end{array}\right]
=
\left[\begin{array}{c|cc|ccc}
\mathbf{A} & \mathbf{c} & \mathbf{0} & \mathbf{e} & \mathbf{0} & \mathbf{0}\\ \hline
\mathbf{w}'^T & (1-\theta) & \theta & 0 & 0 & \mathbf{v}'^T\\
\mathbf{0} & 1 & 0 & 0 & 0 & \mathbf{0}\\ \hline
\mathbf{w}^T & \mathbf{w}'^T\mathbf{e} & \mathbf{v}'^T\mathbf{e} & (1-\theta) & \theta & \mathbf{v}^T\\
\mathbf{0} & 0 & 0 & 1 & 0 & \mathbf{0}\\
\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}
\end{array}\right],
$$

in correspondence of the input vectors $y^{[n-1]}=[y_{n-1},y_{n-2},h^2f(Y^{[n-1]})]^T$, $y'^{[n-1]}=[y'_{n-1},y'_{n-2}]^T$.

3.3.5 Parallel-Iterated Pseudo Two-Step Runge-Kutta-Nyström methods

The usage of previous stage values has also been used in the context of Parallel-Iterated Pseudo Two-Step Runge-Kutta-Nyström methods (PIPTSRKMs)

$$\begin{aligned}
V_n &= y_{n-1}\mathbf{e}_v + hy'_{n-1}c_v + h^2\mathbf{A}_{vv}f(V_{n-1}) + h^2\mathbf{A}_{vw}f(W_{n-1}),\\
W_n &= y_{n-1}\mathbf{e}_w + hy'_{n-1}c_w + h^2\mathbf{A}_{wv}f(V_n) + h^2\mathbf{A}_{ww}f(W_n),\\
hy'_n &= hy'_{n-1} + h^2\mathbf{d}_v^T f(V_n) + h^2\mathbf{d}_w^T f(W_n),\\
y_n &= y_{n-1} + hy'_{n-1} + h^2\mathbf{b}_v^T f(V_n) + h^2\mathbf{b}_w^T f(W_n),
\end{aligned}$$

introduced by Cong [37]. Also these methods can be reformulated as GLMs with $r = 2s + 1$ and $r' = 1$, in correspondence to the tableau (3.2.2)

$$\left[\begin{array}{c|c|c} \mathbf{A} & \mathbf{P} & \mathbf{U} \\ \hline \mathbf{C} & \mathbf{R} & \mathbf{W} \\ \hline \mathbf{B} & \mathbf{Q} & \mathbf{V} \end{array}\right] = \left[\begin{array}{cc|c|ccc} \mathbf{0} & \mathbf{0} & \mathbf{c}_v & \mathbf{e}_v & \mathbf{A}_{vv} & \mathbf{A}_{vw} \\ \mathbf{A}_{wv} & \mathbf{A}_{ww} & \mathbf{c}_w & \mathbf{e}_w & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{d}_v^T & \mathbf{d}_w^T & 1 & \mathbf{0} & \mathbf{0} & 0 \\ \hline \mathbf{b}_v^T & \mathbf{b}_w^T & 1 & 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right],$$

and the vectors $Y^{[n]} = [V_n, W_n]^T$, $y^{[n-1]} = [y_{n-1}, h^2 f(V_{n-1}), h^2 f(W_{n-1})]^T$ and $y'^{[n-1]} = [y'_{n-1}]$.

3.4 Preconsistency and consistency

General Linear Methods theory is particularly useful in order to create an unifying approach to analyze the properties of a numerical method for ODEs, e.g. convergence, consistency and stability.
We introduce in this section some definitions regarding the properties of GLMs, which are nothing more than all the basic desirable properties which one would require to whatever numerical method for ODEs: the novelty lays in their new general formulation. Once a method is represented as GLM, it automatically inherits such definitions and all the corresponding results characterizing GLMs. In order to satisfy minimal accuracy requirements, we assume that there exist three vectors

$$\mathbf{q}_0 = [q_{1,0},\, q_{2,0},\, \ldots,\, q_{r,0}]^T,$$

$$\mathbf{q}_1 = [q_{1,1},\, q_{2,1},\, \ldots,\, q_{r,1}]^T,$$

$$\mathbf{q}_2 = [q_{1,2},\, q_{2,2},\, \ldots,\, q_{r,2}]^T,$$

such that the components of the input and the output vectors of the external stages respectively satisfy

$$\begin{aligned} y_i^{[n-1]} &= q_{i,0}y(t_{n-1}) + q_{i,1}hy'(t_{n-1}) + q_{i,2}h^2y''(t_{n-1}) + O(h^3), \\ y_i^{[n]} &= q_{i,0}y(t_n) + q_{i,1}hy'(t_n) + q_{i,2}h^2y''(t_n) + O(h^3), \end{aligned}$$

$i = 1, 2, \ldots, r$ and, moreover, that there exist two vectors

$$\mathbf{q}'_1 = [q'_{1,1},\, q'_{2,1},\, \ldots,\, q'_{r',1}]^T, \quad \mathbf{q}'_2 = [q'_{1,2},\, q'_{2,2},\, \ldots,\, q'_{r',2}]^T,$$

such that the components of the input and the output vectors associated to the first derivative approximations satisfy

$$hy_i'^{[n-1]} = q'_{i,1}hy'(t_{n-1}) + q'_{i,2}h^2y''(t_{n-1}) + O(h^3),$$
$$hy_i'^{[n]} = q'_{i,1}hy'(t_n) + q'_{i,2}h^2y''(t_n) + O(h^3),$$

$i = 1, \ldots, r'$. We finally request that the components of the stage vector $Y^{[n]}$ satisfy the condition

$$Y_i^{[n]} = y(t_{n-1} + c_ih) + O(h^3), \quad i = 1, \ldots, s,$$

which, by expanding the right hand side in Taylor series around the point t_{n-1}, leads to the condition

$$Y_i^{[n]} = y(t_{n-1}) + c_ihy'(t_{n-1}) + \frac{(c_ih)^2}{2}y''(t_{n-1}) + O(h^3), \quad i = 1, \ldots, s.$$

Substituting these relations in the GLM (3.2.4), we obtain

$$\begin{aligned}
&y(t_{n-1}) + c_ihy'(t_{n-1}) + \frac{c_i^2h^2}{2!}y''(t_{n-1}) \\
&= h^2\sum_{j=1}^{s} a_{ij}y''(t_{n-1}) + h\sum_{j=1}^{r'} p_{ij}\left(q'_{j,1}y'(t_{n-1}) + q'_{j,2}hy''(t_{n-1})\right) \\
&+ \sum_{j=1}^{r} u_{ij}\left(q_{j,0}y(t_{n-1}) + q_{j,1}hy'(t_{n-1}) + q_{j,2}h^2y''(t_{n-1})\right) + O(h^3),\ i = 1, \ldots, s,
\end{aligned}$$

$$\begin{aligned}
&q'_{i,1}hy'(t_n) + q'_{i,2}h^2y''(t_n) \\
&= h^2\sum_{j=1}^{s} c_{ij}y''(t_{n-1}) + h\sum_{j=1}^{r'} r_{ij}\left(q'_{j,1}y'(t_{n-1}) + q'_{j,2}hy''(t_{n-1})\right) \\
&+ \sum_{j=1}^{r} w_{ij}\left(q_{j,0}y(t_{n-1}) + q_{j,1}hy'(t_{n-1}) + q_{j,2}h^2y''(t_{n-1})\right) + O(h^3),\ i = 1, \ldots, r',
\end{aligned}$$

$$
\begin{aligned}
&q_{i,0}y(t_n) + q_{i,1}hy'(t_n) + q_{i,2}h^2y''(t_n) \\
&= h^2\sum_{j=1}^{s} b_{ij}y''(t_{n-1}) + h\sum_{j=1}^{r'} q_{ij}\left(q'_{j,1}y'(t_{n-1}) + q'_{j,2}hy''(t_{n-1})\right) \\
&+ \sum_{j=1}^{r} v_{ij}\left(q_{j,0}y(t_{n-1}) + q_{j,1}hy'(t_{n-1}) + q_{j,2}h^2 + y''(t_{n-1})\right) + O(h^3),\ i = 1,\dots,r.
\end{aligned}
$$

Comparing $O(1)$, $O(h)$ and $O(h^2)$ terms in these relations, we obtain the equations

$$
\begin{aligned}
\sum_{j=1}^{r} u_{ij}q_{j,0} &= 1, \quad i = 1,\dots,s, \\
\sum_{j=1}^{r} w_{ij}q_{j,0} &= 0, \quad i = 1,\dots,r', \\
\sum_{j=1}^{r} v_{ij}q_{j,0} &= q_{i,0}, \quad i = 1,\dots,r,
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=1}^{r'} p_{ij}q'_{j,1} + \sum_{j=1}^{r} u_{ij}q_{j,1} &= c, \quad i = 1,\dots,s, \\
\sum_{j=1}^{r'} r_{ij}q'_{j,1} + \sum_{j=1}^{r} w_{ij}q_{j,1} &= q'_1, \quad i = 1,\dots,r', \\
\sum_{j=1}^{r'} q_{ij}q'_{j,1} + \sum_{j=1}^{r} v_{ij}q_{j,1} &= q_0 + q_1, \quad i = 1,\dots,r,
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=1}^{s} a_{ij} + \sum_{j=1}^{r'} p_{ij}q'_{j,2} + \sum_{j=1}^{r} u_{ij}q_{j,2} &= \frac{c^2}{2}, \quad i = 1,\dots,s, \\
\sum_{j=1}^{s} c_{ij} + \sum_{j=1}^{r'} r_{ij}q'_{j,2} + \sum_{j=1}^{r} w_{ij}q_{j,2} &= q'_1 + q'_2, \quad i = 1,\dots,r', \\
\sum_{j=1}^{s} b_{ij} + \sum_{j=1}^{r'} q_{ij}q'_{j,2} + \sum_{j=1}^{r} v_{ij}q_{j,2} &= \frac{q_0}{2} + q_1 + q_2, \quad i = 1,\dots,r.
\end{aligned}
$$

The above discussion leads to the following definitions.

Definition 3.4.1. *A GLM (3.2.4) is preconsistent if there exist vectors* $\mathbf{q}_0$, $\mathbf{q}_1$ *and* $\mathbf{q}'_1$ *such that*

$$\mathbf{U}\mathbf{q}_0 = e, \quad \mathbf{W}\mathbf{q}_0 = 0, \quad \mathbf{V}\mathbf{q}_0 = \mathbf{q}_0,$$

$$\mathbf{P}\mathbf{q}'_1 + \mathbf{U}\mathbf{q}_1 = \mathbf{c}, \quad \mathbf{R}\mathbf{q}'_1 + \mathbf{W}\mathbf{q}_1 = \mathbf{q}'_1, \quad \mathbf{Q}\mathbf{q}'_1 + \mathbf{V}\mathbf{q}_1 = \mathbf{q}_0 + \mathbf{q}_1.$$

In the context of GLMs for first order ODEs, Butcher [21] observed that preconsistency is equivalent to the concept of *covariance* of a GLM, which essentially ensures that numerical approximations are appropriately transformed by a shift of origin.

Definition 3.4.2. *A preconsistent GLM (3.2.4) is consistent if exist vectors* $\mathbf{q}_2$ *and* $\mathbf{q}'_2$ *such that*

$$\mathbf{C}e + \mathbf{R}\mathbf{q}'_2 + \mathbf{W}\mathbf{q}_2 = \mathbf{q}'_1 + \mathbf{q}'_2, \quad \mathbf{B}e + \mathbf{Q}\mathbf{q}'_2 + \mathbf{V}\mathbf{q}_2 = \frac{\mathbf{q}_0}{2} + \mathbf{q}_1 + \mathbf{q}_2.$$

Definition 3.4.3. *A consistent GLM (3.2.4) is stage-consistent if*

$$\mathbf{A}e + \mathbf{P}\mathbf{q}'_2 + \mathbf{U}\mathbf{q}_2 = \frac{\mathbf{c}^2}{2}.$$

Using the GLM formalism, according to the above definitions, we can also recover the consistency of the classical numerical methods already considered in the literature. For instance, the Numerov method (3.3.2) is consistent with preconsistency and consistency vectors $\mathbf{q}_0 = [1\ 1\ 0\ 0]^T$, $\mathbf{q}_1 = [0\ -1\ 0\ 0]^T$ and $\mathbf{q}_2 = [0\ 1/2\ 1\ 1]^T$. Runge–Kutta–Nyström methods (3.3.3) are consistent with preconsistency and consistency vectors $\mathbf{q}_0 = [1]$, $\mathbf{q}_1 = \mathbf{q}_2 = [0]$, $\mathbf{q}'_1 = [1]$ and $\mathbf{q}'_2 = [0]$. Coleman hybrid methods (3.3.4) are consistent with preconsistency and consistency vectors $\mathbf{q}_0 = [1\ 1]^T$, $\mathbf{q}_1 = [0\ -1]^T$ and $\mathbf{q}_2 = [0\ 1/2]^T$. Two-step Runge–Kutta–Nyström methods (3.3.5) are consistent with preconsistency and consistency vectors $\mathbf{q}_0 = [1\ 1\ 0\ \ldots\ 0\ 0]^T$, $\mathbf{q}_1 = [0\ -1\ 0\ \ldots\ 0\ 0]^T$, $\mathbf{q}_2 = [0\ 1/2\ 1\ \ldots\ 1\ 1]^T$, $\mathbf{q}'_1 = [1\ 1]^T$ and $\mathbf{q}'_2 = [0\ -1]^T$.

3.5 Zero-stability

Another basic requirement in the context of the numerical integration of ODEs is, together with consistency, also zero-stability. In order to define such minimal stability requirement, we apply the GLM (3.2.3) to the problem

$$y'' = 0,$$

obtaining the recurrence relation

$$\begin{bmatrix} hy'^{[n]} \\ y^{[n]} \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{W} \\ \mathbf{Q} & \mathbf{V} \end{bmatrix} \begin{bmatrix} hy'^{[n-1]} \\ y^{[n-1]} \end{bmatrix}.$$

The matrix

$$\mathbf{M}_0 = \begin{bmatrix} \mathbf{R} & \mathbf{W} \\ \mathbf{Q} & \mathbf{V} \end{bmatrix},$$

is denoted as the *zero-stability matrix* of the GLM (3.2.3). The following definition occurs.

Definition 3.5.1. *A GLM (3.2.3) is zero-stable if there exist two real constants C and D such that*

$$\|\mathbf{M}_0^m\| \leq mC + D, \quad \forall m = 1, 2, \ldots. \tag{3.5.1}$$

Condition (3.5.1) is useless for practical purposes and, therefore, we introduce the following criterion which allows the study of zero-stability by means of easy linear algebra arguments. This result follows the lines drawn by Butcher in [18].

Theorem 3.5.1.

The following statements are equivalent:

(i) $\mathbf{M}_0$ *satisfies the bound (3.5.1);*

(ii) the roots of the minimal polynomial of the matrix $\mathbf{M}_0$ *lie on or within the unit circle and the multiplicity of the zeros on the unit circle is at most two;*

(iii) there exists a matrix B *similar to* $\mathbf{M}_0$ *such that*

$$\sup_m \{\|B^m\|_\infty,\ m \geq 1\} \leq m + 1.$$

Proof. The result holds by proving the following implications: $(i) \Rightarrow (ii)$, $(ii) \Rightarrow (iii)$, $(iii) \Rightarrow (i)$. We first prove that $(i) \Rightarrow (ii)$. Suppose that λ is an eigenvalue of $\mathbf{M}_0$ and denote by v the corresponding eigenvector. As a consequence, we obtain

$$\|\mathbf{M}_0^m\|_\infty = \sup_{x \neq 0} \frac{\|\mathbf{M}_0^m x\|_\infty}{\|x\|_\infty} \geq \frac{\|\mathbf{M}_0^m v\|_\infty}{\|v\|_\infty} = \frac{\|\lambda^m v\|_\infty}{\|v\|_\infty} = |\lambda|^m,$$

and, taking into account that $\mathbf{M}_0$ satisfies the assumption (i), we obtain $|\lambda| \leq 1$. Since λ is an element of the spectrum of $\mathbf{M}_0$, it is also a root of its minimal polynomial. We suppose that λ is a repeated zero of the minimal polynomial with multiplicity $\mu(\lambda) = 3$: then there exist three nonzero vectors u, v and w such that

$$\mathbf{M}_0 w = \lambda w + u, \quad \mathbf{M}_0 u = \lambda u + v, \quad \mathbf{M}_0 v = \lambda v.$$

It is easy to prove by induction that

$$\mathbf{M}_0^m w = \lambda^m w + m\lambda^{m-1} u + \frac{m(m-1)}{2}\lambda^{m-2} v, \quad \text{for any} \ \ m \geq 2.$$

As a consequence, the following bound holds:

$$\begin{aligned} \|\mathbf{M}_0^m\|_\infty &= \sup_{x \neq 0} \frac{\|\mathbf{M}_0^m x\|_\infty}{\|x\|_\infty} \geq \frac{\|\mathbf{M}_0^m w\|_\infty}{\|w\|_\infty} \\ &= \frac{\left\|\lambda^m w + m\lambda^{m-1} u + \frac{m(m-1)}{2}\lambda^{m-2} v\right\|_\infty}{\|w\|_\infty} \\ &= |\lambda|^{m-2}\left(|\lambda|^2 - m|\lambda|\frac{\|u\|_\infty}{\|w\|_\infty} - \frac{m(m-1)}{2}\frac{\|v\|_\infty}{\|w\|_\infty}\right). \end{aligned}$$

If $|\lambda| = 1$, then

$$\|\mathbf{M}_0^m\|_\infty \geq 1 - m\frac{\|u\|_\infty}{\|w\|_\infty} - \frac{m(m-1)}{2}\frac{\|v\|_\infty}{\|w\|_\infty},$$

and, by setting $C := \frac{\|u\|_\infty}{\|w\|_\infty}$ and $D := \frac{\|v\|_\infty}{\|w\|_\infty}$, we obtain

$$\|\mathbf{M}_0^m\|_\infty \geq 1 + mC + \frac{m(m-1)}{2}D,$$

which means that $\|\mathbf{M}_0^m\|$ cannot be linearly bounded as $m \to \infty$, against the hypothesis (i). In conclusion, if $\mu(\lambda) = 3$, then $|\lambda| < 1$. In correspondence of $\mu(\lambda) = 2$, we have

$$\|\mathbf{M}_0^m\|_\infty \geq |\lambda|^{m-1}\left(m\frac{\|v\|_\infty}{\|u\|_\infty} - |\lambda|\right),$$

which, for $|\lambda| = 1$, leads to the bound

$$\|\mathbf{M}_0^m\|_\infty \geq mC - 1,$$

which can coherently be combined with the assumption (i). We next suppose that (ii) holds: then, we can choose the matrix B as the Jordan canonical form of $\mathbf{M}_0$

$$B = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 \end{bmatrix},$$

where the block $\mathbf{J}_1$ assumes the form

$$\mathbf{J}_1 = \begin{bmatrix} \lambda & a \\ 0 & \lambda \end{bmatrix},$$

with $|\lambda| = 1$ and

$$a = \begin{cases} 1, & \text{if } \mu(\lambda) = 2, \\ 0, & \text{if } \mu(\lambda) = 1, \end{cases}$$

while the block $\mathbf{J}_2$ contains the eigenvalues λ_i of modulus less than 1 on the diagonal and $1 - |\lambda_i|$ on the upper co-diagonal. Since B is a block diagonal matrix, we have

$$B^m = \begin{bmatrix} \mathbf{J}_1^m & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2^m \end{bmatrix},$$

with

$$\mathbf{J}_1^m = \begin{bmatrix} \lambda^m & am\lambda^{m-1} \\ 0 & \lambda^m \end{bmatrix}.$$

It follows that

$$\|B^m\|_\infty = \max \left\{ \left\| \begin{pmatrix} \lambda^m & am\lambda^{m-1} \\ 0 & \lambda^m \end{pmatrix} \right\|_\infty, \quad \|J_2^m\|_\infty \right\} \leq m + 1.$$

Finally, if (iii) is true, since B is similar to $\mathbf{M}_0$, then there exists a matrix P such that $B = P^{-1}\mathbf{M}_0 P$. As a consequence,

$$\|\mathbf{M}_0^m\|_\infty = \left\|PB^m P^{-1}\right\|_\infty \leq m + 1,$$

i.e. $\mathbf{M}_0$ satisfies the zero-stability bound (3.5.1). □

Remark Let us note that condition (ii) in the following theorem is peculiar in the numerical solution of second order ODEs (1.1.1). In fact, the notion of zero-stability for GLMs solving first order ODEs (compare [18, 21, 97]) implies that the minimal polynomial of its zero-stability matrix can possess at most one root of

modulus one, while all the others have modulus less than one. Instead, in the case of second order ODEs, two roots of the minimal polynomial of the zero-stability matrix lying on the unit circle are allowed, taking into account also the case of complex conjugate roots of modulus one, as might happen in second order ODEs (1.1.1) in the oscillatory case. This is made clear in [82], where the authors prove the necessity for convergence of such a zero-stability condition in the context of linear multistep methods. After introducing the general concept of zero-stability for GLMs (3.2.4) designed to numerically solve second order ODEs (1.1.1), we also want to verify that the obtained result recovers the classical ones, i.e. that the classical zero-stable methods satisfy the requirement on the minimal polynomial of their zero-stability matrix asserted in Theorem 3.5.1. For instance:

- the minimal polynomial associated to the zero-stability matrix of the Numerov method (3.3.2) is $p(\lambda) = \lambda^2(\lambda - 1)^2$, which satisfies the requirement (ii) in Theorem 3.5.1, i.e. the Numerov method is zero-stable;

- in the case of Runge–Kutta–Nyström methods (3.3.3), the minimal polynomial of the zero-stability matrix is $p(\lambda) = (\lambda - 1)^2$, which satisfies the requirement (ii) in Theorem 3.5.1;

- the same minimal polynomial arises in the case of Coleman two-step hybrid methods (3.3.4) and, therefore, their zero-stability is recovered;

- as regards two-step Runge–Kutta–Nyström methods (3.3.5), the minimal polynomial of their zero-stability matrix is $p(\lambda) = \lambda^2(\lambda^2 - (1-\theta)\lambda - \theta)$ and, therefore, such methods are zero-stable if and only if $-1 < \theta \leq 1$: this restriction on θ recovers the classical result (compare [127]).

3.6 Convergence

In this section we focus our attention on the convergence analysis of GLMs (3.2.4), first extending the ideas introduced by Butcher [21] in order to formulate a rigorous definition of convergence for a GLM (3.2.4). In force of the nature of GLMs, a starting procedure is needed in order to determine the missing starting values $y^{[0]}$ and $y'^{[0]}$ to be used as input for the first step of the integration process: in the context of convergence analysis, we only need to assume that there exists a starting procedure

$$S_h : \mathbb{R}^{2d} \to \mathbb{R}^{dr},$$

associating, for any value of the stepsize h, a starting vector $y^{[0]} = S_h(y_0, y'_0)$ such that

$$\lim_{h\to 0} \frac{S_h(y_0, y'_0) - (\mathbf{q_0} \otimes I)y(t_0)}{h} = (\mathbf{q_1} \otimes I)y'(t_0), \tag{3.6.1}$$

and, moreover, the initial vector $y'^{[0]}$ is provided in order to ensure that

$$\lim_{h\to 0} y'^{[0]} = (\mathbf{q'_1} \otimes I)y'(t_0). \tag{3.6.2}$$

We now introduce the following definition.

Definition 3.6.1. *A preconsistent GLM (3.2.4) is convergent if, for any well-posed initial value problem (1.1.1), there exists a starting procedure S_h satisfying (3.6.1) such that the sequence of vectors $y^{[n]}$, computed using n steps with stepsize $h = (\bar{t} - t_0)/n$ and using $y^{[0]} = S_h(y_0, y'_0)$, converges to $\mathbf{q}_0 y(\bar{t})$, and the sequence of vectors $y'^{[n]}$, computed using n steps with the same stepsize h starting from $y'^{[0]}$ satisfying* (3.6.2)*, converges to $\mathbf{q'}_1 y'(\bar{t})$, for any $\bar{t} \in [t_0, T]$.*

Proving the convergence of a numerical method is generally a quite tedious and nontrivial task: however, the following result creates a very close connection among the concepts of convergence, consistency and zero-stability and allows to prove the convergence of a numerical scheme by checking some algebraic conditions involving

the coefficients of the method. This powerful result has already been proved in the context of GLMs for first order ODEs [21, 97]. We now extend this result and state a sufficient condition for the convergence of GLMs (3.2.4) for second order ODEs.

Theorem 3.6.1. *A GLM (3.2.4) is convergent if it is consistent and zero-stable.*

Proof. We introduce the vectors

$$\widehat{y}^{[n-1]} = \begin{bmatrix} \widehat{y}_1^{[n-1]} \\ \widehat{y}_2^{[n-1]} \\ \vdots \\ \widehat{y}_r^{[n-1]} \end{bmatrix}, \widehat{y}^{[n]} = \begin{bmatrix} \widehat{y}_1^{[n]} \\ \widehat{y}_2^{[n]} \\ \vdots \\ \widehat{y}_r^{[n]} \end{bmatrix}, \widehat{y}'^{[n-1]} = \begin{bmatrix} \widehat{y}'^{[n-1]}_1 \\ \widehat{y}'^{[n-1]}_2 \\ \vdots \\ \widehat{y}'^{[n-1]}_{r'} \end{bmatrix}, \widehat{y}'^{[n]} = \begin{bmatrix} \widehat{y}'^{[n]}_1 \\ \widehat{y}'^{[n]}_2 \\ \vdots \\ \widehat{y}'^{[n]}_{r'} \end{bmatrix},$$

defined by

$$\begin{aligned} \widehat{y}_i^{[n-1]} &= q_{i0}y(t_{n-1}) + q_{i1}hy'(t_{n-1}) + q_{i2}h^2y''(t_{n-1}), \\ \widehat{y}_i^{[n]} &= q_{i0}y(t_n) + q_{i1}hy'(t_n) + q_{i2}h^2y''(t_n), \\ h\widehat{y}'^{[n-1]}_i &= q'_{i1}hy'(t_{n-1}) + q'_{i2}h^2y''(t_{n-1}), \\ h\widehat{y}'^{[n]}_i &= q'_{i1}hy'(t_n) + q'_{i2}h^2y''(t_n), \end{aligned}$$

where $q_{i,0}$, $q_{i,1}$ and $q'_{i,1}$ are the components of the preconsistency vectors $\mathbf{q}_0$, $\mathbf{q}_1$ and $\mathbf{q}'_1$, while $q_{i,2}$ and $q'_{i,2}$ are the components of the consistency vectors $\mathbf{q}_2$ and $\mathbf{q}'_2$. We next denote by

$$\xi_i(h), \quad \eta_i(h), \quad \zeta_i(h),$$

the residua arising after replacing in the GLM (3.2.4) $y_i^{[n-1]}$, $y_i^{[n]}$, $y'^{[n-1]}_i$ and $y'^{[n]}_i$ by $\widehat{y}_i^{[n-1]}$, $\widehat{y}_i^{[n]}$, $\widehat{y}'^{[n-1]}_i$ and $\widehat{y}'^{[n]}_i$ respectively and, moreover, $Y_i^{[n]}$ by $y(t_{n-1} + c_ih)$. The

mentioned replacements lead to the following equations:

$$\begin{aligned} y(t_{n-1}+c_ih) &= h^2\sum_{j=1}^{s} a_{ij}y''(t_{n-1}+c_jh) \\ &+ h\sum_{j=1}^{r'} p_{ij}(q'_{j1}y'(t_{n-1})+hq'_{j2}y''(t_{n-1})) \\ &+ \sum_{j=1}^{r} u_{ij}(q_{j0}y(t_{n-1})+q_{j1}hy'(t_{n-1})+q_{j2}h^2y''(t_{n-1})) \\ &+ \xi_i(h), \qquad i=1,2,\ldots,s, \end{aligned} \tag{3.6.3}$$

$$\begin{aligned} q'_{i1}hy'(t_n) + q'_{i2}h^2y''(t_n) &= h^2\sum_{j=1}^{s} c_{ij}y''(t_{n-1}+c_jh) \\ &+ h\sum_{j=1}^{r'} r_{ij}(q'_{j1}y'(t_{n-1})+hq'_{j2}y''(t_{n-1})) \\ &+ \sum_{j=1}^{r} w_{ij}(q_{j0}y(t_{n-1})+q_{j1}hy'(t_{n-1})+q_{j2}h^2y''(t_{n-1})) \\ &+ \zeta_i(h), \qquad i=1,2,\ldots,r', \end{aligned} \tag{3.6.4}$$

$$\begin{aligned} q_{i0}y(t_n) + q_{i1}hy'(t_n) + q_{i2}h^2y''(t_n) &= h^2\sum_{j=1}^{s} b_{ij}y''(t_{n-1}+c_jh) \\ &+ h\sum_{j=1}^{r'} a_{ij}(q'_{j1}y'(t_{n-1})+hq'_{j2}y''(t_{n-1})) \\ &+ \sum_{j=1}^{r} v_{ij}(q_{j0}y(t_{n-1})+q_{j1}hy'(t_{n-1})+q_{j2}h^2y''(t_{n-1})) \\ &+ \eta_i(h), \qquad i=1,2,\ldots,r. \end{aligned} \tag{3.6.5}$$

By expanding $y(t_{n-1}+c_ih)$, $y''(t_{n-1}+c_jh)$ in Taylor series around t_{n-1}, replacing the obtained expansions in (3.6.3) and using the hypothesis of preconsistency of the

method, we obtain that

$$\xi_i(h) = O(h^2), \quad i = 1, 2, \ldots, s.$$

In particular, we observe that if the method (3.2.4) is also stage consistent, we have

$$\xi_i(h) = O(h^3), \quad i = 1, 2, \ldots, s.$$

Proceeding in analogous way for (3.6.4) and (3.6.5), using in these cases the preconsistency and consistency conditions, we obtain

$$\zeta_i(h) = O(h^3), \quad i = 1, 2, \ldots, r',$$

and

$$\eta_i(h) = O(h^3), \quad i = 1, 2, \ldots, r.$$

Subtracting the equations for $y_i^{[n]}$ and $\widehat{y}_i^{[n]}$, we obtain

$$\begin{aligned} y_i^{[n]} - \widehat{y}_i^{[n]} &= h^2 \sum_{j=1}^{s} b_{ij}(f(Y_j^{[n]}) - f(y(t_{n-1} + c_j h))) \\ &+ h \sum_{j=1}^{r'} q_{ij}(y'^{[n-1]}_j - \widehat{y}'^{[n-1]}_j) \\ &+ \sum_{j=1}^{r} v_{ij}(y_j^{[n-1]} - \widehat{y}_j^{[n-1]}) - \eta_i(h), \end{aligned}$$

or, equivalently, in tensor form

$$\begin{aligned} y^{[n]} - \widehat{y}^{[n]} &= h^2(\mathbf{B} \otimes I)\left(F(Y^{[n]}) - F(y(t_{n-1} + \mathbf{c}h))\right) \\ &+ h(\mathbf{Q} \otimes I)(y'^{[n-1]} - \widehat{y}'^{[n-1]}) \\ &+ (\mathbf{V} \otimes I)(y^{[n-1]} - \widehat{y}^{[n-1]}) - \eta(h). \end{aligned} \tag{3.6.6}$$

By means of analogous arguments, we obtain the following representation of the difference between $hy'^{[n]}$ and $h\widehat{y}'^{[n]}$:

$$\begin{aligned} h(y'^{[n]} - \widehat{y}'^{[n]}) &= h^2(\mathbf{C}\otimes I)\Big(F(Y^{[n]}) - F(y(t_{n-1}+\mathbf{c}h))\Big) \\ &+ h(\mathbf{R}\otimes I)(y'^{[n-1]} - \widehat{y}'^{[n-1]}) \\ &+ (\mathbf{W}\otimes I)(y^{[n-1]} - \widehat{y}^{[n-1]}) - \zeta(h). \end{aligned} \tag{3.6.7}$$

In order to provide a more compact version of formulae (3.6.6) and (3.6.7), we introduce the notations

$$\begin{aligned} u_n &= y^{[n]} - \widehat{y}^{[n]}, \\ v_n &= h(y'^{[n]} - \widehat{y}'^{[n]}), \\ w_n &= h^2(\mathbf{B}\otimes I)\Big(F(Y^{[n]}) - F(y(t_{n-1}+\mathbf{c}h))\Big) - \eta(h), \\ z_n &= h^2(\mathbf{C}\otimes I)\Big(F(Y^{[n]}) - F(y(t_{n-1}+\mathbf{c}h))\Big) - \zeta(h). \end{aligned}$$

With these notations, formulae (3.6.6) and (3.6.7) respectively assume the form

$$u_n = w_n + (\mathbf{Q}\otimes I)v_{n-1} + (\mathbf{V}\otimes I)u_{n-1}, \tag{3.6.8}$$

$$v_n = z_n + (\mathbf{R}\otimes I)v_{n-1} + (\mathbf{W}\otimes I)u_{n-1}. \tag{3.6.9}$$

Observe that, by applying the Lipschitz continuity of the function F, the following bound for w_n arises

$$\|w_n\| \leq h^2 L\|B\|\|Y^{[n]} - y(t_{n-1}+\mathbf{c}h)\| + \|\eta(h)\|, \tag{3.6.10}$$

where L is the Lipschitz constant of F. In order to provide a bound for $\|Y^{[n]} - y(t_{n-1}+\mathbf{c}h)\|$, we use the following representation to the difference inside the norm:

$$Y^{[n]} - y(t_{n-1}+\mathbf{c}h) = h^2(\mathbf{A}\otimes I)\Big(F(Y^{[n]}) - F(y(t_{n-1}+\mathbf{c}h))\Big) + (\mathbf{P}\otimes I)v_{n-1}$$

$$+ \quad (\mathbf{U} \otimes I)u_{n-1} - \xi(h).$$

As a consequence, the following bound holds:

$$\begin{aligned} \|Y^{[n]} - y(t_{n-1} + \mathbf{c}h)\| &\leq h^2 L \|\mathbf{A}\| \|Y^{[n]} - y(t_{n-1} + \mathbf{c}h)\| + \|\mathbf{P}\| \|v_{n-1}\| \\ &+ \|\mathbf{U}\| \|u_{n-1}\| + \|\xi(h)\|. \end{aligned}$$

Assuming that $h < h_0$ and $h_0 L \|A\| < 1$, we obtain

$$\|Y^{[n]} - y(t_{n-1} + \mathbf{c}h)\| \leq \frac{\|P\|}{1 - h_0^2 L \|A\|} \|v_{n-1}\| + \frac{\|U\|}{1 - h_0^2 L \|A\|} \|u_{n-1}\| + \frac{\|\xi(h)\|}{1 - h_0^2 L \|A\|}. \tag{3.6.11}$$

Substituting in (3.6.10), we obtain

$$\|w_n\| \leq h^2 (D\|v_{n-1}\| + E\|u_{n-1}\|) + h^2 \delta(h), \tag{3.6.12}$$

where

$$D = \frac{L\|B\|\|P\|}{1 - h_0^2 L \|A\|}, \qquad E = \frac{L\|B\|\|U\|}{1 - h_0^2 L \|A\|}, \qquad \delta(h) = \frac{L\|B\|\|\xi(h)\|}{1 - h_0^2 L \|A\|} + \|\eta(h)\|.$$

In analogous way, we obtain the following bound for z_n:

$$\|z_n\| \leq h^2 (\overline{D}\|v_{n-1}\| + \overline{E}\|u_{n-1}\|) + h^2 \overline{\delta}(h), \tag{3.6.13}$$

where

$$\overline{D} = \frac{L\|C\|\|P\|}{1 - h_0^2 L \|A\|}, \qquad \overline{E} = \frac{L\|C\|\|U\|}{1 - h_0^2 L \|A\|}, \qquad \overline{\delta}(h) = \frac{L\|C\|\|\xi(h)\|}{1 - h_0^2 L \|A\|}.$$

We put together the two bounds (3.6.12) and (3.6.13) obtaining, in vector form,

$$\|e_n\| \leq h^2 \|\Lambda\| \cdot \|d_{n-1}\| + h^2 \|\sigma\|, \tag{3.6.14}$$

where

$$e_n = \begin{bmatrix} w_n \\ \\ z_n \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} D & E \\ \\ \overline{D} & \overline{E} \end{bmatrix}, \qquad \sigma = \begin{bmatrix} \delta(h) \\ \\ \overline{\delta}(h) \end{bmatrix}, \qquad d_{n-1} = \begin{bmatrix} u_{n-1} \\ \\ v_{n-1} \end{bmatrix}.$$

Proceeding in analogous way for Equations (3.6.8) and (3.6.9), we obtain

$$d_n = \mathbf{M}_0 d_{n-1} + e_n. \tag{3.6.15}$$

Applying Equation (3.6.15) n times, we obtain

$$d_n = \mathbf{M}_0^n d_0 + \sum_{j=1}^{n} \mathbf{M}_0^{n-j} e_j, \quad n \geq 0,$$

and, passing through the norm, we obtain the bound

$$\|d_n\| \leq \|\mathbf{M}_0^n\| \cdot \|d_0\| + \sum_{j=1}^{n} \|\mathbf{M}_0^{n-j}\| \|e_j\|.$$

Since the hypothesis of zero-stability holds, there exist $C_1, D_1 \in \mathbb{R}$ such that $\|\mathbf{M}_0^n\| \leq nC_1 + D_1$. By using this bound and the estimation (3.6.14), we obtain

$$\|d_n\| \leq (n + C_1 + D_1)\|d_0\| + \sum_{j=1}^{n} \left((n-j)C_1 + D_1\right)\left(C_2\|d_{j-1}\| + D_2\right),$$

where $C_2 = h^2\|\Lambda\|$ and $D_2 = h^2\|\sigma\|$. This bound, after some calculations, can be rewritten as

$$\|d_n\| \leq \alpha(n) + \sum_{j=2}^{n} \beta_j(n)\|d_j\|, \tag{3.6.16}$$

where

$$\alpha(n) = (nC_1 + (n-1)C_1C_2 + D_1 + C_2D_1)\|d_0\| + \left(\frac{n(n-1)}{2}C_1 + nD_1\right)D_2,$$

$$\beta_j(n) = ((n-j-1)C_1C_2 + C_2D_1)\,.$$

We set $j = i_1$ and apply the corresponding inequality (3.6.16) for $\|d_{i_1}\|$, i.e.

$$\|d_{i_1}\| \leq \alpha(n) + \sum_{i_2=2}^{i_1} \beta_{i_2}(n)\|d_{i_2}\|.$$

Replacing this inequality in (3.6.16) leads to

$$\|d_n\| \leq \alpha(n) + \sum_{i_1=2}^{n} \beta_j(n)\alpha(j) + \sum_{i_1=2}^{n}\sum_{i_2=2}^{i_1} \beta_{i_1}(n)\beta_{i_2}(i_1)\|d_{i_2}\|.$$

By iterating this process, we obtain

$$\|d_n\| \leq \alpha(n) + \sum_{i_1=2}^{n}\sum_{i_2=2}^{i_1}\sum_{i_3=2}^{i_2}\cdots\sum_{i_N=2}^{2}\prod_{j=1}^{N}\beta_{i_j}(i_{j-1})\alpha(i_j), \tag{3.6.17}$$

under the assumption that $i_0 = n$. We observe that the right hand side of the inequality (3.6.17) is expressed as the summation of $\alpha(n)$, which can be bounded by $D_1\|d_0\|$ as n tends to infinity, plus a series whose principal term behaves as $O(1/n^2)$ and, therefore, it converges. Then, the following bound holds

$$\|d_n\| \leq \left(D_1 + C_1^2\|\Lambda\|(\bar{t}-t_0)\right)\|d_0\| + O(h^2), \tag{3.6.18}$$

which completes the proof. □

Remark This result allows us to recover the convergence of any classical method for (1.1.1) that can be regarded as GLM. For instance, in the case of Runge–Kutta–Nyström methods [82], the matrix Λ and the vector σ respectively assume the form

$$\Lambda = \frac{L}{1-h_0^2L\|A\|}\begin{bmatrix} \|b\|\cdot\|c\| & \|b\| \\ \|b'\|\cdot\|c\| & \|b'\| \end{bmatrix}, \qquad \sigma = \begin{bmatrix} \dfrac{L\|b\|\cdot\|\xi(h)\|}{1-h_0^2L\|A\|} + \|\eta(h)\| \\ \dfrac{L\|b'\|\cdot\|\xi(h)\|}{1-h_0^2L\|A\|} \end{bmatrix},$$

and

$$\alpha(n) = (n + 1 + nC_2)\|d_0\| + \frac{n(n+1)}{2}D_2, \qquad \beta_j(n) = jC_2.$$

By analogous considerations to that in Theorem 3.6.1, we achieve the bound (3.6.18), where $C_1 = D_1 = 1$.

As regards Coleman hybrid methods [32], the matrix Λ and the vector σ respectively assume the form

$$\Lambda = \frac{L\|b\|}{1 - h_0^2 L\|A\|} \begin{bmatrix} 0 & 1 \\ \\ 0 & 0 \end{bmatrix}, \qquad \sigma = \begin{bmatrix} \dfrac{L\|b\| \cdot \|\xi(h)\|}{1 - h_0^2 L\|A\|} + \|\eta(h)\| \\ \\ 0 \end{bmatrix},$$

and

$$\alpha(n) = (2n + 1 + (2n - 1)C_2)\|d_0\| + n^2 D_2, \qquad \beta_j(n) = (2j - 1)C_2.$$

Then, the bound (3.6.18) holds, with $C_1 = 2$ and $D_1 = 1$.

3.7 Order conditions

The derivation of practical conditions to derive nonlinear numerical methods of a certain order is a nontrivial problem which has been successfully treated in the literature by Albrecht [1, 2, 3, 4, 109], Butcher [18, 21] and references therein contained, Hairer and Wanner [82, 84] and references therein, using different approaches. In the context of first order ODEs, the order theory for GLMs has been developed by J. Butcher (see [18, 21] and references therein contained), by considering rooted trees associated to the numerical methods. However, in the case of high stage order methods, a different approach to derive order conditions can be used. This approach has been discussed by Butcher himself in [17], in the context of diagonally implicit multistage integration methods (see [97]), in the cases $q = p$ and $q = p - 1$, where p

is the order of the method and q is its stage order. We use this approach to derive order conditions of GLMs for second order ODEs (1.1.1). As initial case of study, we assume that the order p of the GLM is equal to its stage order q: this choice allows the methods to have a uniform order of convergence and, as a consequence, they would not suffer from order reduction (see [18] as regards first order ODEs) in the integration of stiff differential systems. We first assume that the components of the input and output vectors respectively satisfy

$$y_i^{[n-1]} = \sum_{k=0}^{p} q_{ik} h^k y^{(k)}(t_{n-1}) + O(h^{p+1}), \tag{3.7.1}$$

$$y_i^{[n]} = \sum_{k=0}^{p} q_{ik} h^k y^{(k)}(t_n) + O(h^{p+1}), \tag{3.7.2}$$

for some real parameters q_{ik}, $i = 1, 2, \ldots, r$, $k = 0, 1, \ldots, p$. We will next denote p as the *order* of the method. We then assume that the components of the internal stages $Y_i^{[n]}$ are approximations of order q to the solution of (1.1.1) at the internal points $t_{n-1} + c_i h$, i.e.

$$Y_i^{[n]} = y(t_{n-1} + c_i h) + O(h^{q+1}), \quad i = 1, 2, \ldots, s. \tag{3.7.3}$$

We will next denote q as the *stage order* of the method. We also request that the components of the input and output vectors of the derivatives respectively satisfy

$$h y'^{[n-1]}_i = \sum_{k=1}^{p} q'_{ik} h^k y^{(k)}(t_{n-1}) + O(h^{p+1}), \tag{3.7.4}$$

$$h y'^{[n]}_i = \sum_{k=1}^{p} q'_{ik} h^k y^{(k)}(t_n) + O(h^{p+1}), \tag{3.7.5}$$

for some real parameters q'_{ik}, $i = 1, 2, \ldots, r'$, $k = 1, 2, \ldots, p$.
We introduce the following notation

$$e^{\mathbf{c}z} = [e^{c_1 z}, e^{c_2 z}, \ldots, e^{c_s z}],$$

and define the vectors

$$\mathbf{w} = \mathbf{w}(z) = \sum_{k=0}^{p} \mathbf{q}_k z^k,$$

and

$$\mathbf{w}' = \mathbf{w}'(z) = \sum_{k=1}^{p} \mathbf{q}'_k z^k.$$

We aim to obtain algebraic conditions ensuring that a GLM (3.2.4) has order $p = q$. The following theorem holds.

Theorem 3.7.1. *Assume that $y^{[n-1]}$ and $y'^{[n-1]}$ satisfy respectively (3.7.1) and (3.7.4). Then the GLM (3.2.4) of order p and stage order $q = p$ satisfies (3.7.2), (3.7.3) and (3.7.5) if and only if*

$$e^{\mathbf{c}z} = z^2\mathbf{A}e^{\mathbf{c}z} + \mathbf{P}\mathbf{w}'(z) + \mathbf{U}\mathbf{w}(z) + O(z^{p+1}), \tag{3.7.6}$$

$$e^{z}\mathbf{w}'(z) = z^2\mathbf{C}e^{\mathbf{c}z} + \mathbf{R}\mathbf{w}'(z) + \mathbf{W}\mathbf{w}(z) + O(z^{p+1}), \tag{3.7.7}$$

$$e^{z}\mathbf{w}(z) = z^2\mathbf{B}e^{\mathbf{c}z} + \mathbf{Q}\mathbf{w}'(z) + \mathbf{V}\mathbf{w}(z) + O(z^{p+1}). \tag{3.7.8}$$

Proof. Since $Y_i^{[n]} = y(t_{n-1} + c_i h) + O(h^{q+1})$, $i = 1, 2, \ldots, s$, it follows that

$$h^2 f(Y_i^{[n]}) = h^2 y''(t_{n-1} + c_i h) + O(h^{p+3}) = \sum_{k=2}^{p} \frac{c_i^{k-2}}{(k-2)!} h^k y^{(k)}(t_{n-1}) + O(h^{p+1}).$$

Expanding in Taylor series around t_{n-1}, Equation (3.7.2) can be written in the form

$$y_i^{[n]} = \sum_{k=0}^{p} \left(\sum_{l=0}^{k} \frac{1}{l!} q_{i,k-l} \right) h^k y^{(k)}(t_{n-1}) + O(h^{p+1}). \tag{3.7.9}$$

Substituting the relations (3.7.1), (3.7.2), (3.7.3), (3.7.4) and (3.7.5) in the GLM formulation (3.2.4), we obtain

$$\sum_{k=0}^{p} \left(c_i^k - \sum_{j=1}^{s} k(k-1) a_{ij} c_j^{k-2} - \sum_{j=1}^{r'} k! p_{ij} q'_{jk} - \sum_{j=1}^{r} k! u_{ij} q_{jk} \right) \frac{h^k}{k!} y^{(k)}(t_{n-1}) = O(h^{p+1}),$$

$$\sum_{k=0}^{p}\left(\sum_{l=0}^{k}\frac{k!}{l!}q_{i,k-l}-\sum_{j=1}^{s}k(k-1)b_{ij}c_j^{k-2}-\sum_{j=1}^{r'}k!q_{ij}q'_{jk}\right)\frac{h^k}{k!}y^{(k)}(t_{n-1})=O(h^{p+1}),$$

and

$$\sum_{k=0}^{p}\left(\sum_{l=0}^{k}\frac{k!}{l!}q'_{i,k-l}-\sum_{j=1}^{s}k(k-1)c_{ij}c_j^{k-2}-\sum_{j=1}^{r'}k!r_{ij}q'_{jk}-\sum_{j=1}^{r}k!w_{ij}q_{jk}\right)\frac{h^k}{k!}y^{(k)}(t_{n-1})$$
$$=O(h^{p+1}).$$

Equating to zero the coefficients of $h^k y^{(k)}(t_{n-1})/k!$, $k=0,1,\ldots,p$, multiplying them by $z^k/k!$, and summing them over k from 0 to p, we obtain

$$e^{c_i z}-z^2\sum_{j=1}^{s}a_{ij}e^{c_j z}-\sum_{j=1}^{r'}p_{ij}w'_j-\sum_{j=1}^{r}u_{ij}w_j=O(z^{p+1}),\quad i=1,2,\ldots,s,$$
$$e^{z}w'_i-z^2\sum_{j=1}^{s}c_{ij}e^{c_j z}-\sum_{j=1}^{r'}r_{ij}w'_j-\sum_{j=1}^{r}w_{ij}w_j=O(z^{p+1}),\quad i=1,2,\ldots,r',$$
$$e^{z}w_i-z^2\sum_{j=1}^{s}b_{ij}e^{c_j z}-\sum_{j=1}^{r'}q_{ij}w'_j-\sum_{j=1}^{r}v_{ij}w_j=O(z^{p+1}),\quad i=1,2,\ldots,r.$$

These relations are equivalent to (3.7.6), (3.7.7) and (3.7.8). This completes the proof. □

It follows from the proof of Theorem 3.7.1 that the conditions (3.7.6), (3.7.7) and (3.7.8) are respectively equivalent to

$$\mathbf{c}^k-k(k-1)\mathbf{A}\mathbf{c}^{k-2}-k!\mathbf{P}\mathbf{q}'_k-k!\mathbf{U}\mathbf{q}_k=0, \tag{3.7.10}$$
$$\sum_{l=0}^{k}\frac{k!}{l!}\mathbf{q}'_{k-l}-k(k-1)\mathbf{C}\mathbf{c}^{k-2}-k!\mathbf{R}\mathbf{q}'_k-k!\mathbf{W}\mathbf{q}_k=0, \tag{3.7.11}$$
$$\sum_{l=0}^{k}\frac{k!}{l!}\mathbf{q}_{k-l}-k(k-1)\mathbf{B}\mathbf{c}^{k-2}-k!\mathbf{Q}\mathbf{q}'_k-k!\mathbf{V}\mathbf{q}_k=0, \tag{3.7.12}$$

for $k=2\ldots,p+1$.

These equalities constitute the system of order condition that a GLM has to satisfy in order to achieve order p equal to the stage order q.

We introduce the vectors $\widehat{y}_i^{[n-1]}$, $\widehat{y}_i^{[n]}$, $\widehat{y}_i'^{[n-1]}$ and $\widehat{y}_i'^{[n]}$ defined by the following expansions

$$\widehat{y}_i^{[n-1]} = \sum_{k=0}^{p} \mathbf{q}_{ik} h^k y^{(k)}(t_{n-1}), \qquad \widehat{y}_i^{[n]} = \sum_{k=0}^{p} \mathbf{q}_{ik} h^k y^{(k)}(t_n),$$

$$h\widehat{y}_i'^{[n-1]} = \sum_{k=1}^{p} \mathbf{q}'_{ik} h^k y^{(k)}(t_{n-1}), \qquad h\widehat{y}_i'^{[n]} = \sum_{k=1}^{p} \mathbf{q}'_{ik} h^k y^{(k)}(t_n).$$

The following theorem holds.

Theorem 3.7.2. *Assume that the GLM (3.2.4) has order p and stage order $q = p$. Moreover, also suppose that the starting values satisfy*

$$\left\| y_i^{[0]} - \widehat{y}_i^{[0]} \right\| = O(h^p), \quad i = 1, 2, \ldots, r,$$

and

$$\left\| y_i'^{[0]} - \widehat{y}_i'^{[0]} \right\| = O(h^p), \quad i = 1, 2, \ldots, r',$$

as $h \to 0$. Then,

$$\left\| y_i^{[n]} - \widehat{y}_i^{[n]} \right\| = O(h^p), \quad i = 1, 2, \ldots, r,$$

and

$$\left\| y_i'^{[n]} - \widehat{y}_i'^{[n]} \right\| = O(h^p), \quad i = 1, 2, \ldots, r',$$

as $h \to 0$, $h_n = \overline{t} - t_0$. In addition,

$$\left\| Y_i^{[n]} - y(t_{n-1} + c_i h) \right\| = O(h^p), \quad i = 1, 2, \ldots, s.$$

Proof. Following the lines drawn in Theorem 3.6.1, we introduce the residua $\xi_i(h)$, $\eta_i(h)$ and $\zeta_i(h)$ defined by

$$y(t_{n-1}+c_ih) = h^2\sum_{j=1}^{s} a_{ij}y''(t_{n-1}+c_jh) + h\sum_{j=1}^{r'} p_{ij}\widehat{y}_j'^{[n-1]} + \sum_{j=1}^{r} u_{ij}\widehat{y}_j^{[n-1]} + \xi_i(h),$$

$$h\widehat{y}_i'^{[n]} = h^2\sum_{j=1}^{s} c_{ij}y''(t_{n-1}+c_jh) + h\sum_{j=1}^{r'} r_{ij}\widehat{y}_j'^{[n-1]} + \sum_{j=1}^{r} w_{ij}\widehat{y}_j^{[n-1]} + \zeta_i(h),$$

$$\widehat{y}_i^{[n]} = h^2\sum_{j=1}^{s} b_{ij}y''(t_{n-1}+c_jh) + h\sum_{j=1}^{r'} q_{ij}\widehat{y}_j'^{[n-1]} + \sum_{j=1}^{r} v_{ij}\widehat{y}_j^{[n-1]} + \eta_i(h).$$

Expanding in Taylor series around the point t_{n-1} and using (3.7.10), (3.7.11) and (3.7.12), it follows that

$$\xi_i(h) = O(h^{p+1}), \quad i = 1, 2, \ldots, s,$$

$$\zeta_i(h) = O(h^{p+1}), \quad i = 1, 2, \ldots, r',$$

$$\eta_i(h) = O(h^{p+1}), \quad i = 1, 2, \ldots, r.$$

Proceeding in the same way as in the proof of Theorem 3.6.1, we obtain

$$u_n = (\mathbf{V}\otimes\mathbf{I})u_{n-1} + (\mathbf{Q}\otimes\mathbf{I})v_{n-1} + w_n,$$

$$v_n = (\mathbf{W}\otimes\mathbf{I})u_{n-1} + (\mathbf{R}\otimes\mathbf{I})v_{n-1} + z_n,$$

where now

$$u_n = y^{[n]} - \widehat{y}^{[n]}, \quad w_n = h^2(\mathbf{B}\otimes\mathbf{I})(F(Y^{[n]}) - F(y(t_{n-1}+ch))) - \eta(h),$$

and

$$v_n = y'^{[n]} - \widehat{y}'^{[n]}, \quad z_n = h^2(\mathbf{C}\otimes\mathbf{I})(F(Y^{[n]}) - F(y(t_{n-1}+ch))) - \zeta(h).$$

Moreover, the following bounds hold

$$\|w_n\| \leq h^2(D\,\|v_{n-1}\| + E\,\|u_{n-1}\|) + h^2\delta(h),$$

$$\|z_n\| \leq h^2(\overline{D}\,\|v_{n-1}\| + \overline{E}\,\|u_{n-1}\|) + h^2\overline{\delta}(h),$$

and $D,\ E,\ \overline{D},\ \overline{E},\ \delta(h)$ and $\overline{\delta}(h)$ are defined as in Theorem 3.6.1. From the fact that $\delta(h) = O(h^p)$, $\overline{\delta}(h) = O(h^p)$, we gain the following bound for the vector $d_n = [u_n \quad v_n]^T$,

$$\|d_n\| \leq D_1\,\|d_0\| + C_1^2\,\|d_0\|\,\|\Lambda\|\,(\bar{t} - t_0),$$

by analogous considerations to that in Theorem 3.6.1. By using the hypothesis on the accuracy of the starting values, the first part of the thesis is proved.

As regards the second part of the thesis concerning the accuracy of the stages, it follows from the inequality

$$\left\|Y^{[n]} - y(t_{n-1} + ch)\right\| \leq \frac{\|\mathbf{P}\|}{1 - h_0^2 L\,\|\mathbf{A}\|}\,\|v_{n-1}\| + \frac{\|\mathbf{U}\|}{1 - h_0^2 L\,\|\mathbf{A}\|}\,\|u_{n-1}\| + \frac{\|\xi(h)\|}{1 - h_0^2 L\,\|\mathbf{A}\|},$$

provided in Theorem 3.6.1. □

By using Equations (3.7.10), (3.7.11) and (3.7.12), we are able to recover the order conditions of the classical numerical methods for the solution of the problem (1.1.1). For instance, consider the Numerov method (3.3.2): in this case, the vectors $\mathbf{q}_k$ assume the form

$$\mathbf{q}_0 = [1 \quad 1 \quad 0 \quad 0]^T, \qquad \mathbf{q}_1 = [0 \quad -1 \quad 0 \quad 0],$$

$$\mathbf{q}_2 = \left[0 \quad \frac{1}{2} \quad 1 \quad 1\right]^T, \qquad \mathbf{q}_k = \left[0 \quad -\frac{(-1)^k}{k!} \quad 0 \quad -\frac{(-1)^{k-2}}{(k-2)!}\right], \quad k \geq 3.$$

Applying order conditions (3.7.10) and (3.7.12), we recover that the Numerov method has order 4 (compare, for instance, [94]).

Associated to Runge–Kutta–Nyström methods (3.3.3), we have $\mathbf{q}_k = [\delta_{0k}]$, for $k = 0, 1, \ldots, p$, $\mathbf{q}'_k = [\delta_{1k}]$, for $k = 1, \ldots, p$, where δ_{jk} denotes the usual Kronecker delta. In correspondence to these vectors, conditions (3.7.10), (3.7.11) and (3.7.12) lead to

$$A\mathbf{c}^{k-2} = \frac{\mathbf{c}^k}{k(k-1)},$$

$$\mathbf{b}'^T\mathbf{c}^{k-2} = \frac{1}{k-1},$$

$$\mathbf{b}^T\mathbf{c}^{k-2} = \frac{1}{k(k-1)},$$

$k = 2, 3, \ldots, p+1$, which are the order conditions for (3.3.3) classically derived in the literature (compare [82]).

Concerning Coleman hybrid methods (3.3.4), we have $\mathbf{q}_0 = [1 \quad 1]^T$ and $\mathbf{q}_k = [0 \quad \frac{(-1)^k}{k!}]^T$, $k \geq 1$. Correspondingly, order conditions (3.7.10) and (3.7.12) assume the form

$$A\mathbf{c}^{k-2} = \frac{\mathbf{c}^k + (-1)^k\mathbf{c}}{k(k-1)},$$

$$\mathbf{b}^T\mathbf{c}^{k-2} = \frac{1 + (-1)^k}{k(k-1)},$$

$k = 2, 3, \ldots, p+1$, and these are the same order conditions derived by Coleman in [32].

As regards the family of two-step Runge–Kutta–Nyström methods, the vectors $\mathbf{q}_l$ and $\mathbf{q}'_l$ respectively take the form

$$\mathbf{q}_0 = [1 \quad 1 \mid \mathbf{0}]^T, \qquad \mathbf{q}_1 = [0 \quad -1 \mid \mathbf{0}]^T,$$

$$\mathbf{q}_j = \left[0 \quad \frac{(-1)^j}{j!} \mid \frac{(\mathbf{c}-\mathbf{e})^{j-2}}{(j-2)!}\right]^T, \ j \geq 2,$$

$$\mathbf{q}'_1 = [1 \quad 1]^T, \qquad \mathbf{q}'_k = \left[0 \quad \frac{(-1)^k}{k!}\right]^T, \ k \geq 2.$$

Correspondingly, the order conditions assume the form

$$A\mathbf{c}^{k-2} = \frac{\mathbf{c}^k}{k(k-1)},$$

$$\mathbf{v}'^T(\mathbf{c}-\mathbf{e})^{k-2} + \mathbf{w}'^T\mathbf{c}^{k-2} = \frac{1+(-1)^k\theta}{k-1},$$

$$\mathbf{v}^T(\mathbf{c}-\mathbf{e})^{k-2} + \mathbf{w}^T\mathbf{c}^{k-2} = \frac{1+(-1)^{k+1}\theta}{k(k-1)} + \frac{(-1)^k}{k-1}\mathbf{v}'^T\mathbf{e},$$

$k = 2, 3, \ldots, p+1$, which are the order conditions for (3.3.5) derived in [129].

3.8 Linear stability analysis

We now focus our attention on the basic linear stability requirements that any numerical method for ODEs has to accomplish. The definition of such properties we present in this section are formulated according to the formalism of GLMs. Linear stability properties for numerical methods solving second order ODEs (1.1.1) are classically provided with respect to the scalar linear test equation

$$y'' = -\lambda^2 y. \tag{3.8.1}$$

Applying the GLM (3.2.3) to the test equation (3.8.1), we obtain

$$Y^{[n]} = -\lambda^2 h^2 \mathbf{A} Y^{[n]} + h\mathbf{P}y'^{[n-1]} + \mathbf{U}y^{[n-1]}, \tag{3.8.2}$$

$$y^{[n]} = -\lambda^2 h^2 \mathbf{B} Y^{[n]} + h\mathbf{Q}y'^{[n-1]} + \mathbf{V}y^{[n-1]}, \tag{3.8.3}$$

$$hy'^{[n]} = -\lambda^2 h^2 \mathbf{C} Y^{[n]} + h\mathbf{R}y'^{[n-1]} + \mathbf{W}y^{[n-1]}. \tag{3.8.4}$$

We set $z = \lambda h$, $\Lambda = (\mathbf{I} + z^2\mathbf{A})^{-1}$, assuming that the matrix $\mathbf{I} + z^2\mathbf{A}$ is invertible. Then it follows from (3.8.2) that

$$Y^{[n]} = \Lambda\mathbf{P}hy'^{[n-1]} + \Lambda\mathbf{U}y^{[n-1]},$$

and substituting this into (3.8.3) and (3.8.4), we obtain

$$\begin{bmatrix} y^{[n]} \\ hy'^{[n]} \end{bmatrix} = \mathbf{M}(z) \begin{bmatrix} y^{[n-1]} \\ hy'^{[n-1]} \end{bmatrix},$$

where the matrix $\mathbf{M}(z)$ is defined by

$$\mathbf{M}(z) = \begin{bmatrix} -z^2\mathbf{B}\Lambda\mathbf{U} + \mathbf{V} & -z^2\mathbf{B}\Lambda\mathbf{P} + \mathbf{Q} \\ -z^2\mathbf{C}\Lambda\mathbf{U} + \mathbf{W} & -z^2\mathbf{C}\Lambda\mathbf{P} + \mathbf{R} \end{bmatrix}.$$

The matrix $\mathbf{M}(z) \in \mathbb{R}^{r+r'}$ is the so-called *stability (or amplification) matrix*, while its characteristic polynomial

$$p(\omega, z^2) = \det(\mathbf{M}(z^2) - \omega I_{r+r'}),$$

is denoted as *stability polynomial*, which is polynomial of degree $r + r'$ with respect to ω and its coefficients are rational functions with respect to z^2.
We next consider the following definitions.

Definition 3.8.1. $(0, \beta^2)$ *is a stability interval for the GLM (3.2.4) if,* $\forall z^2 \in (0, \beta^2)$*, the spectral radius* $\rho(\mathbf{M}(z^2))$ *of the matrix* $\mathbf{M}(z^2)$ *satisfies*

$$\rho(\mathbf{M}(z^2)) < 1. \tag{3.8.5}$$

The condition (3.8.5) is equivalent to the fact that the roots of stability polynomial are in modulus less than 1, $\forall z^2 \in (0, \beta^2)$. In particular, setting $S(z^2) = trace(\mathbf{M}^2(z^2))$ and $P(z^2) = det(\mathbf{M}^2(z^2))$, for a one-step Runge–Kutta–Nyström

method, the condition (3.8.5) is equivalent to

$$P(z^2) < 1, \quad |S(z^2)| < P(z^2) + 1, \quad \forall z \in (0, \beta^2).$$

Definition 3.8.2. *A GLM is $A-$ stable if $(0, \beta^2) = (0, +\infty)$.*

If the eigenvalues of the stability matrix (or, equivalently, the roots of the stability polynomial) are on the unit circle, then the interval of stability becomes an interval of periodicity, according to the following definition.

Definition 3.8.3. *$(0, H_0^2)$ is a periodicity interval for the method (3.2.4) if, $\forall z^2 \in (0, H_0^2)$, the stability polynomial $p(\omega, z^2)$ has two complex conjugate roots of modulus 1, while all the others have modulus less than 1.*

For a one-step Runge–Kutta–Nyström method, the interval of periodicity [151] is then defined by

$$(0, H_0^2) := \{z^2 \ : \ P(z^2) \equiv 1, \ |S(z^2)| < 2\}.$$

Definition 3.8.4. *A GLM is $P-$ stable if its periodicity interval is $(0, +\infty)$.*

3.9 Example of new methods

Using GLMs we introduce new classes of methods. We start from a GLM with the following features

$$Y^{[n]} = \begin{bmatrix} \overline{y}_{n+k} \\ \overline{y}_{n+k+1} \\ y_{n+k} \end{bmatrix}, \quad f(Y^{[n]}) = \begin{bmatrix} \overline{f}_{n+k} \\ \overline{f}_{n+k+1} \\ f_{n+k} \end{bmatrix}, \quad y^{[n]} = \begin{bmatrix} y_{n+k} \\ y_{n+k-1} \\ \vdots \\ y_{n+1} \end{bmatrix},$$

and with coefficient matrices

$$\mathbf{A} = \begin{bmatrix} \widehat{\beta}_k & 0 & 0 \\ -\widehat{\alpha}_{k-1}\widehat{\beta}_k & \widehat{\beta}_k & 0 \\ \beta_k - \widehat{\beta}_k & \beta_{k+1} & \widehat{\beta}_k \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \beta_k - \widehat{\beta}_k & \beta_{k+1} & \widehat{\beta}_k \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} -\widehat{\alpha}_{k-1} & -\widehat{\alpha}_{k-2} & \cdots & -\widehat{\alpha}_1 & -\widehat{\alpha}_0 \\ \widehat{\alpha}_{k-1}\widehat{\alpha}_{k-1} - \widehat{\alpha}_{k-2} & \widehat{\alpha}_{k-1}\widehat{\alpha}_{k-2} - \widehat{\alpha}_{k-3} & \cdots & \widehat{\alpha}_{k-1}\widehat{\alpha}_1 - \widehat{\alpha}_0 & \widehat{\alpha}_{k-1}\widehat{\alpha}_0 \\ -\alpha_{k-1} & -\alpha_{k-2} & \cdots & -\alpha_1 & -\alpha_0 \end{bmatrix},$$

$$\mathbf{V} = \begin{bmatrix} -\alpha_{k-1} & -\alpha_{k-2} & \cdots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

We compute $\overline{y}_{n+k}$ as the solution of the conventional BDF method

$$\overline{y}_{n+k} + \sum_{j=0}^{k-1} \widehat{\alpha}_j y_{n+j} = h^2 \widehat{\beta}_k \overline{f}_{n+k}, \qquad \overline{f}_{n+k} = f(t_{n+k}, \overline{y}_{n+k}).$$

Next, we compute $\overline{y}_{n+k+1}$ as the solution of the same BDF advanced one step

$$\overline{y}_{n+k+1} + \widehat{\alpha}_{k-1}\overline{y}_{n+k} + \sum_{j=0}^{k-2} \widehat{\alpha}_j y_{n+j+1} = h^2 \widehat{\beta}_k \overline{f}_{n+k+1},$$

and finally we apply the Corrector:

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h^2 \widehat{\beta}_k f_{n+k} + h^2(\beta_k - \widehat{\beta}_k)\overline{f}_{n+k} + h^2 \beta_{k+1}\overline{f}_{n+k+1}.$$

These methods are the second order version of the modified BDF formulae of the first order case (Cash, 1980)

$$\overline{y}_{n+k} + \sum_{j=0}^{k-1} \widehat{\alpha}_j y_{n+j} = h\widehat{\beta}_k \overline{f}_{n+k}, \qquad \overline{f}_{n+k} = f(t_{n+k}, \overline{y}_{n+k}),$$

$$\overline{y}_{n+k+1} + \widehat{\alpha}_{k-1}\overline{y}_{n+k} + \sum_{j=0}^{k-2} \widehat{\alpha}_j y_{n+j+1} = h\widehat{\beta}_k \overline{f}_{n+k+1},$$

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h\widehat{\beta}_k f_{n+k} + h(\beta_k - \widehat{\beta}_k)\overline{f}_{n+k} + h\beta_{k+1}\overline{f}_{n+k+1}.$$

We study the convergence, in the particular case $k = 2$. First, we derive the preconsistency and consistency vectors, respectively

$$\mathbf{q}_0 = [1, 1]^T, \quad \mathbf{q}_1 = [0, -1]^T, \quad \mathbf{q}_2 = [0, 1/2]^T.$$

Next, we solve the system of order conditions for $p = 1$, and obtain

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -\beta_3 & \beta_3 & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 2 & -1 \\ 3 & -2 \\ 2 & -1 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} -\beta_3 & \beta_3 & 1 \\ \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} 2 & -1 \\ \\ 1 & 0 \end{bmatrix}.$$

Since $\sigma(\mathbf{V}) = \{1\}$, it follows that the methods are zero-stable, for any β_3. In this way we obtain that the methods are consistent and zero-stable and then convergent. It has now become the usefulness of the developed theory on order conditions and convergence property in the analysis of properties of new numerical methods, if we succeed in representing them as GLMs.

Part II

Exponentially Fitted Methods

4

Exponentially fitted two-step hybrid methods for second order ordinary differential equations

4.1 Introduction

In this chapter we derive numerical methods approximating the solution of initial value problems based on second order ordinary differential equations (1.1.1), with $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ smooth enough in order to ensure the existence and uniqueness of the solution, which is assumed to exhibit a periodic/oscillatory behaviour. As mentioned in the introduction, classical numerical methods for ODEs may not be well-suited to follow a prominent periodic or oscillatory behaviour because, in order to accurately catch the oscillations, a very small stepsize would be required with corresponding deterioration of the numerical performances, especially in terms of efficiency. For this reason, many classical numerical methods have been adapted in order to efficiently approach the oscillatory behaviour. One of the possible ways to proceed in this direction can be realized by imposing that a numerical method exactly integrates (within the round-off error) problems of type (1.1.1) whose solution can be expressed as linear combination of functions other than polynomials: this is the spirit of the exponential fitting technique (EF, see [94]), where the adapted numerical method is developed in order to be exact on problems whose solution is

linear combination of

$$\{1, t, \ldots, t^K, \exp(\pm\mu t), t\exp(\pm\mu t), \ldots, t^P \exp(\pm\mu t)\},$$

where K and P are integer numbers. The methods we consider in this section belong to the class of two-step hybrid methods

$$Y_i^{[n]} = (1 + c_i)y_n - c_i y_{n-1} + h^2 \sum_{j=1}^{s} a_{ij} f(Y_j^{[n]}), \quad i = 1, ..., s \tag{4.1.1}$$

$$y_{n+1} = 2y_n - y_{n-1} + h^2 \sum_{i=1}^{s} b_i f(Y_i^{[n]}), \tag{4.1.2}$$

introduced by Coleman in [32], which can also be represented through the Butcher array

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} \tag{4.1.3}$$

with $c = [c_1, c_2, ..., c_s]^T$, $A = (a_{ij})_{i,j=1}^{s}$, $b = [b_1, b_2, ..., b_s]^T$, where s is the number of stages. The interest in this class of methods, as also pointed out by Coleman in [32], lies in their formulation: "many other methods, though not normally written like this, can be expressed in the same way by simple rearrangement". For this reason, they represent one of the first attempts to obtain wider and more general classes of numerical methods for (1.1.1), towards a class of General Linear Methods [21, 53, 62, 97] for this problem.

The aim of this chapter is the derivation of EF-based methods within the class (4.1.1)-(4.1.2) depending on one or two parameters, which we suppose can be estimated in advance. Frequency-dependent methods within the class (4.1.1)-(4.1.2) have already been considered in [154], where phase-fitted and amplification-fitted two-step hybrid methods have been derived, and also in [55], where trigonometrically fitted methods (4.1.1)-(4.1.2) depending on one and two frequencies have been

proposed. In Section 2 we present the constructive technique of EF methods of type (4.1.1)-(4.1.2). Section 3 is devoted to the local error analysis and the parameter estimation, while in Section 4 we analyze the linear stability properties of the derived methods. Finally Section 5 provides numerical tests confirming the theoretical expectations.

4.2 Construction of the methods

We present the constructive technique we used to derive EF methods within the class (4.1.1)-(4.1.2), based on the so-called six-step procedure, introduced by Ixaru and Vanden Berghe in [94] as a constructive tool to derive EF based formulae approaching many problems of Numerical Analysis (e.g. interpolation, numerical quadrature and differentiation, numerical solution of ODEs), especially when their solutions show a prominent periodic/oscillatory behaviour. This procedure provides a general way to derive EF formulae whose coefficients are expressed in a regularized way and, as a consequence, they do not suffer from numerical cancellation. Indeed, coefficients expressed as linear combinations of sine, cosine and exponentials suffer from heavy numerical cancellation and, in the implementation, they are generally replaced by their power series expansion, suitably truncated. On the contrary, the coefficients of EF methods obtained by using the six-step flow chart are expressed by means of the $\eta_k(Z)$ functions introduced by Ixaru (see [88, 94] and references therein contained) and, as a consequence, the effects of numerical cancellation are notably reduced. In agreement with the procedure, we first consider the following set of $s+1$ functional operators

$$\mathcal{L}[h,\mathbf{b}]y(t) = y(t+h) - 2y(t) + y(t-h) - h^2\sum_{i=1}^{s} b_i y''(t+c_i h), \tag{4.2.1}$$

$$\mathcal{L}_i[h,\mathbf{a}]y(t) = y(t+c_i h) - (1+c_i)y(t) + c_i y(t-h) - h^2\sum_{j=1}^{s} a_{ij} y''(t+c_j h), \tag{4.2.2}$$

for $i = 1, \ldots, s$, which are associated to the method (4.1.1)-(4.1.2). We next report the first five steps of the procedure, while the remaining one, i.e. the local error analysis, is reported in Section 3.

- step (i) *Computation of the classical moments.* The reduced classical moments (see [94], p. 42) are defined, in our case, as

$$L^*_m(\mathbf{b}) = h^{-(m+1)}\mathcal{L}[h;\mathbf{b}]t^m,\ m = 0, 1, 2, \ldots,$$
$$L^*_{im}(\mathbf{a}) = h^{-(m+1)}\mathcal{L}_i[h;\mathbf{a}]t^m,\ i = 1, \ldots, s,\ m = 0, 1, 2, \ldots.$$

- step (ii) *Compatibility analysis.* We examine the algebraic systems

$$L^*_m(\mathbf{b}) = 0,\ m = 0, 1, \ldots, M' - 1,$$
$$L^*_{im}(\mathbf{a}) = 0,\ i = 1, \ldots, s,\ m = 0, 1, \ldots, M - 1,$$

  to find out the maximal values of M and M' for which the above systems are compatible. If $s = 2$, we have

$$L^*_0 = 0, \quad L^*_1 = 0, \quad L^*_2 = 2(1 - b_1 - b_2),$$

$$L^*_3 = 6(-b_1c_1 - b_2c_2), \quad L^*_4 = 12(\frac{1}{6} - b_1c_1^2 - b_2c_2^2),$$

$$L^*_{10} = 0, \quad L^*_{11} = 0, \quad L^*_{12} = c_1 + c_1^2 - 2(a_{11} + a_{12}),$$

$$L^*_{13} = -c_1(1 + 6a_{11} - c_1^2) - 6a_{12}c_2, \quad L^*_{14} = c_1 + c_1^4 - 12(a_{11}c_1^2 + a_{12}c_2^2),$$

$$L^*_{20} = 0, \quad L^*_{21} = 0, \quad L^*_{22} = c_2 + c_2^2 - 2(a_{21} + a_{22}),$$

$$L^*_{23} = -c_2(1 + 6a_{22} - c_2^2) - 6a_{21}c_1, \quad L^*_{24} = c_2 + c_2^4 - 12(a_{21}c_1^2 + a_{22}c_2^2),$$

  and, therefore, $M = M' = 4$.

- step (iii) *Computation of the G functions.* In order to derive EF methods, we need to compute the so-called *reduced (or starred) exponential moments* (see [94], p. 42), i.e.

$$E_0^*(\pm z, \mathbf{b}) = \exp(\pm\mu t)\mathcal{L}[h, \mathbf{b}]\exp(\pm\mu t),$$
$$E_{0i}^*(\pm z, \mathbf{a}) = \exp(\pm\mu t)\mathcal{L}_i[h, \mathbf{a}]\exp(\pm\mu t), i = 1, \ldots, s,$$

where $z = \mu h$. Once computed the reduced exponential moments, we can derive the G functions, defined in the following way:

$$G^+(Z, \mathbf{b}) = \frac{1}{2}\Big(E_0^*(z, \mathbf{b}) + E_0^*(-z, \mathbf{b})\Big),$$
$$G^-(Z, \mathbf{b}) = \frac{1}{2z}\Big(E_0^*(z, \mathbf{b}) - E_0^*(-z, \mathbf{b})\Big),$$
$$G_i^+(Z, \mathbf{a}) = \frac{1}{2}\Big(E_{0i}^*(z, \mathbf{a}) + E_{0i}^*(-z, \mathbf{a})\Big), \quad i = 1, \ldots, s,$$
$$G_i^-(Z, \mathbf{a}) = \frac{1}{2z}\Big(E_{0i}^*(z, \mathbf{a}) - E_{0i}^*(-z, \mathbf{a})\Big), \quad i = 1, \ldots, s,$$

where $Z = z^2$. In our case, the G functions take the following form

$$G^+(Z, \mathbf{b}) = 2\eta_{-1}(Z) - 2 - Z\sum_{j=1}^{s} b_j\eta_{-1}(c_j^2 Z),$$
$$G^-(Z, \mathbf{b}) = -Z\sum_{j=1}^{s} b_j c_j\eta_0(c_j^2 Z),$$
$$G_i^+(Z, \mathbf{a}) = \eta_{-1}(c_i^2 Z) + c_i\eta_{-1}(Z) - 2(1 + c_i) - Z\sum_{j=1}^{s} a_{ij}\eta_{-1}(c_j^2 Z),$$
$$G_i^-(Z, \mathbf{a}) = c_i\eta_0(c_i^2 Z) - c_i\eta_0(Z) - 2(1 + c_i) - Z\sum_{j=1}^{s} c_j a_{ij}\eta_0(c_j^2 Z),$$

for $i = 1, \ldots, s$. We observe that the above expressions depend on the functions $\eta_{-1}(Z)$ and $\eta_0(Z)$ (compare [88, 94]), which are defined as follows

$$\eta_{-1}(Z) = \frac{1}{2}[\exp(Z^{1/2}) + \exp(-Z^{1/2})] = \begin{cases} \cos(|Z|^{1/2}) & \text{if } Z \leq 0, \\ \cosh(Z^{1/2}) & \text{if } Z > 0, \end{cases}$$

and

$$\eta_0(Z) = \begin{cases} \dfrac{1}{2Z^{1/2}}[\exp(Z^{1/2}) - \exp(-Z^{1/2})] & \text{if } Z \neq 0, \\ 1 & \text{if } Z = 0, \end{cases}$$

$$= \begin{cases} \sin(|Z|^{1/2})/|Z|^{1/2} & \text{if } Z < 0, \\ 1 & \text{if } Z = 0, \\ \sinh(Z^{1/2})/Z^{1/2} & \text{if } Z > 0. \end{cases}$$

We next compute the p-th derivatives $G^{\pm(p)}$ and $G_i^{\pm(p)}$, taking into account the formula for the p-th derivative of $\eta_k(Z)$ (see [94]), i.e.

$$\eta_k^{(p)}(Z) = \frac{1}{2^p}\eta_{k+p}(Z).$$

We thus obtain

$$G_i^{+(p)}(Z, \mathbf{a}) = \frac{c_i^{2p}}{2^p}\eta_{p-1}(c_i^2 Z) + \frac{c_i}{2^p}\eta_{p-1}(Z) - \sum_{j=1}^{s} a_{ij}\frac{d^p}{dZ^p}\Big(Z\eta_{-1}(c_j^2 Z)\Big),$$

$$G_i^{-(p)}(Z, \mathbf{a}) = \frac{c_i^{2p+1}}{2^p}\eta_p(c_i^2 Z) - \frac{c_i}{2^p}\eta_p(Z) - \sum_{j=1}^{s} a_{ij}c_j\frac{d^p}{dZ^p}\Big(Z\eta_0(c_j^2 Z)\Big),$$

$$G^{+(p)}(Z, \mathbf{b}) = \frac{1}{2^{p-1}}\eta_{p-1}(Z) - \sum_{j=1}^{s} b_j\frac{d^p}{dZ^p}\Big(Z\eta_{-1}(c_j^2 Z)\Big),$$

$$G^{-(p)}(Z, \mathbf{b}) = -\sum_{j=1}^{s} b_j c_j\frac{d^p}{dZ^p}\Big(Z\eta_{-1}(c_j^2 Z)\Big),$$

for $i = 1, \ldots, s$.

- step (iv) *Definition of the function basis.* We next decide the shape of the function basis to take into account: as a consequence, the corresponding method will exactly integrate (i.e. the operator $\mathcal{L}[h, \mathbf{b}]y(t)$ annihilates in correspondence of the function basis) all those problems whose solution is linear combination of the basis functions.

  In general, the set of M functions is a collection of both powers and exponentials, i.e.

$$\{1, t, \ldots, t^K, \exp(\pm\mu t), t\exp(\pm\mu t), \ldots, t^P \exp(\pm\mu t)\}, \tag{4.2.3}$$

  where K and P are integer numbers satisfying the relation

$$K + 2P = M - 3. \tag{4.2.4}$$

  Let us next consider the set of M' functions

$$\{1, t, \ldots, t^{K'}, \exp(\pm\mu t), t\exp(\pm\mu t), \ldots, t^{P'} \exp(\pm\mu t)\} \tag{4.2.5}$$

  annihilating the operators $\mathcal{L}_i[h, \mathbf{a}]y(t)$, $i = 1, 2, \ldots, s$ and assume that $K' = K$ and $P' = P$, i.e. the external stage and the internal ones are exact on the same function basis. We observe that other possible choices can be taken into account: this can be explained by means of the compatibility of the linear systems to be solved in order to derive the parameters of the methods. In fact, the s^2 unknown elements of the matrix A are derived by solving a linear system of $s(K' + 2P' + 3)$ equations, while the s elements of the vector b are the solution of a $K + 2P + 3$ dimensional linear system.

Such systems are compatible if and only if

$$\begin{cases} s^2 = s(K' + 2P' + 3), \\ s = K + 2P + 3, \end{cases}$$

or, equivalently, if $K' + 2P' = K + 2P$. One natural choice which satisfies this requirement is, of course, $K' = K$ and $P' = P$, but other possibilities can be certainly taken into account, even if they are not explored in this section.

- step (v) *Determination of the coefficients.* After a suitable choice of K and P, we next solve the following algebraic systems:

$$G^{\pm(p)}(Z, \mathbf{b}) = 0,\ p = 0, ..., P,$$
$$G_i^{\pm(p)}(Z, \mathbf{a}) = 0,\ i = 1, \ldots, s,\ p = 0, ..., P.$$

This chapter focuses on the complete analysis of two-stage EF methods with $K = -1$ and $P = 1$ within the class (4.1.1)-(4.1.2), whose coefficients have been reported in Appendix A. In correspondence to this choice of K and P, the fitting space assumes the form

$$\{1, t, \exp(\pm\mu t), t\exp(\pm\mu t)\}. \tag{4.2.6}$$

We observe that, even if $K = -1$, the monomial t is present in the basis (4.2.6), because it automatically annihilates the linear operators (4.2.1)-(4.2.2). It is also possible to extend the above procedure in order to derive EF methods belonging to the class (4.1.1)-(4.1.2), in the case of more than one parameter. In particular, Appendix A reports the coefficients of two-parameters EF methods with 4 stages, with respect to the basis of functions

$$\{1, t, \exp(\pm\mu_1 t), \exp(\pm\mu_2 t)\}. \tag{4.2.7}$$

The final step of this procedure, i.e. the error analysis of the derived formulae, is reported in Section 4.3.

4.3 Error analysis and estimation of the parameters

According to the used procedure, the general expression of the local truncation error for an EF method with respect to the basis of functions (4.2.3) takes the form (see [94])

$$lte^{EF}(t) = (-1)^{P+1} h^M \frac{L^*_{K+1}(\mathbf{b}(Z))}{(K+1)! Z^{P+1}} D^{K+1} (D^2 - \mu^2)^{P+1} y(t), \tag{4.3.1}$$

with K, P and M satisfying the condition (4.2.4). Taking into account our choice (4.2.6) for the functional basis, we obtain

$$lte^{EF}(t) = \frac{L^*_2(\mathbf{b}(Z))}{2\mu^4} D^2 (D^2 - \mu^2)^2 y(t). \tag{4.3.2}$$

We next expand lte^{EF} in Taylor series around t, evaluate it in the current point t_n and consider the leading term of the series expansion, obtaining

$$lte^{EF}(t_n) = -\frac{1 + 6c_1c_2}{24\mu^2} \left(\mu^4 y^{(2)}(t_n) - 2\mu^2 y^{(4)}(t_n) + y^{(6)}(t_n) \right) h^4 + \mathcal{O}(h^5). \tag{4.3.3}$$

The local error analysis also constitutes a starting point for the estimation of the unknown parameter μ which is, in general, a nontrivial problem. In fact, up to now, a rigorous theory for the exact computation of the parameter μ has not yet been developed, but several attempts have been done in the literature in order to provide an accurate estimation (see [92, 94] and references therein), generally based on the minimization of the leading term of the local discretization error. For this reason we annihilate the term $\mu^4 y^{(2)}(t_n) - 2\mu^2 y^{(4)}(t_n) + y^{(6)}(t_n)$ and estimate the parameter in the following way:

$$\mu = \sqrt{\frac{y^{(4)}(t_n) + \sqrt{y^{(4)}(t_n)^2 - y''(t_n)y^{(6)}(t_n)}}{y''(t_n)}}. \tag{4.3.4}$$

The expressions for the occurring derivatives can be obtained analytically from the given ODEs (1.1.1). In Chapter 5 we will analyze more in detail the problem of the estimation of the parameters.

4.4 Linear stability analysis

We next analyze the linear stability properties [34, 151, 152] of the resulting methods, taking into account their dependency on the parameters. The following definitions regard both the case of constant coefficients methods (4.1.1)-(4.1.2), and their exponentially fitted version.

4.4.1 Methods with constant coefficients

Following [151], we apply (4.1.1)-(4.1.2), to the test problem

$$y'' = -\lambda^2 y, \qquad \lambda \in \mathbb{R}$$

obtaining the following recurrence relation (see [56])

$$\begin{bmatrix} y_{n+1} \\ \\ y_n \end{bmatrix} = \begin{bmatrix} M_{11}(\nu^2) & M_{12}(\nu^2) \\ \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_n \\ \\ y_{n-1} \end{bmatrix}, \tag{4.4.1}$$

where

$$M_{11}(\nu^2) = 2 - \nu^2 b^T Q(\nu^2)(e + c),$$
$$M_{12}(\nu^2) = -1 + \nu^2 b^T Q(\nu^2)c,$$

and $Q(\nu^2) = (I + \nu^2 A)^{-1}$, with $\nu^2 = h^2\lambda^2$.

The matrix

$$M(\nu^2) = \begin{bmatrix} M_{11}(\nu^2) & M_{12}(\nu^2) \\ 1 & 0 \end{bmatrix}, \tag{4.4.2}$$

is the so-called *stability (or amplification) matrix* [151, 152]. Let us denote its spectral radius by $\rho(M(\nu^2))$. From [151, 152], the following definitions hold.

Definition 4.4.1. $(0, \beta^2)$ *is a stability interval for the method* (4.1.1)-(4.1.2) *if,* $\forall \nu^2 \in (0, \beta^2)$*, it is*

$$\rho(M(\nu^2)) < 1. \tag{4.4.3}$$

The condition (4.4.3) means that both the eigenvalues $\lambda_1(\nu^2)$ and $\lambda_2(\nu^2)$ of $M(\nu^2)$ are in modulus less than 1, $\forall \nu^2 \in (0, \beta^2)$. By setting $S(\nu^2) = \mathrm{Tr}(M^2(\nu^2))$ and $P(\nu^2) = \det(M^2(\nu^2))$, (4.4.3) is equivalent to

$$P(\nu^2) < 1, \quad |S(\nu^2)| < P(\nu^2) + 1, \quad \nu^2 \in (0, \beta^2). \tag{4.4.4}$$

Definition 4.4.2. *The method* (4.1.1)-(4.1.2) *is A-stable if* $(0, \beta^2) = (0, +\infty)$.

If $\lambda_1(\nu^2)$ and $\lambda_2(\nu^2)$ both lie on the unit circle, then the interval of stability becomes an interval of periodicity, according to the following definition.

Definition 4.4.3. $(0, H_0^2)$ *is a periodicity interval if,* $\forall \nu^2 \in (0, H_0^2)$*,* $\lambda_1(\nu^2)$ *and* $\lambda_2(\nu^2)$ *are complex conjugate and have modulus 1.*

Equivalently,

$$P(\nu^2) = 1, \quad |S(\nu^2)| < 2, \quad \forall \nu^2 \in (0, H_0^2). \tag{4.4.5}$$

Definition 4.4.4. *The method* (4.1.1)-(4.1.2) *is P-stable if* $(0, H_0^2) = (0, +\infty)$.

4.4.2 Methods with coefficients depending on one and two frequencies

Coleman and Ixaru discussed in [34] the modifications to introduce in the linear stability analysis for one-parameter depending EF methods. As a consequence of the presence of the parameter μ, the interval of stability becomes a bidimensional stability region for the one parameter family of methods. In order to emphasize the dependency on the fitted parameter $Z = z^2$, we use the notation $M(\nu^2, Z)$, $R(\nu^2, Z) = \frac{1}{2}\mathrm{Tr}(M(\nu^2, Z))$, $P(\nu^2, Z) = \det(M(\nu^2, Z))$ to denote the stability matrix, its halved trace and determinant respectively. The following definition arises:

Definition 4.4.5. *A region of stability* Ω *is a region of the* (ν^2, Z) *plane, such that* $\forall(\nu^2, Z) \in \Omega$

$$P(\nu^2, Z) < 1, \quad |R(\nu^2, Z)| < (P(\nu^2, Z) + 1). \tag{4.4.6}$$

Any closed curve defined by $P(\nu^2, Z) \equiv 1$ *and* $|R(\nu^2, Z)| = \frac{1}{2}(P(\nu^2, Z) + 1)$ *is a stability boundary.*

We next consider the linear stability analysis of methods depending on two frequencies. As stated before, for methods with constant coefficients, the stability region is an interval on the real axis, while methods depending on one frequency have a bidimensional stability region. In the case of methods depending on the values of two parameters μ_1, μ_2 the stability region becomes tridimensional. We now denote the stability matrix of the methods as $M(\nu^2, Z_1, Z_2)$, with $Z_1 = \mu_1^2 h^2$ and $Z_2 = \mu_2^2 h^2$. The definition of stability region for two-parameters depending methods can be adapted as follows [56, 71]:

Definition 4.4.6. *A three dimensional region Ω of the (ν^2, Z_1, Z_2) space is said to be the region of stability of the corresponding two-frequency depending method if,* $\forall(\nu^2, Z_1, Z_2) \in \Omega$,

$$P(\nu^2, Z_1, Z_2) < 1, \quad |R(\nu^2, Z_1, Z_2)| < \frac{1}{2}(P(\nu^2, Z_1, Z_2) + 1). \tag{4.4.7}$$

Any closed curve defined by

$$P(\nu^2, Z_1, Z_2) \equiv 1, \quad |R(\nu^2, Z_1, Z_2)| = \frac{1}{2}(P(\nu^2, Z_1, Z_2) + 1). \tag{4.4.8}$$

is a stability boundary for the method.

Examples of bidimensional and tridimensional stability regions are provided in Appendix A.

4.5 Numerical results

We now perform some numerical experiments confirming the theoretical expectations regarding the methods we have derived. The implemented solvers are based on the following methods:

- COLEM2, two-step hybrid method (4.1.1)-(4.1.2) having constant coefficients (see [56])

$$\begin{array}{c|cc} \frac{1}{2} & -1 & \\ 1 & 2 & -1 \\ \hline & 2 & -1 \end{array} \tag{4.5.1}$$

  with $s = 2$ and order 2;

- EXPCOLEM2, one-parameter depending exponentially fitted method (4.1.1)-(4.1.2), with $s = 2$ and order 2, whose coefficients are reported in Appendix A.

We implement such methods in a fixed stepsize environment, with step $h = \frac{1}{2^k}$, with k positive integer number.

The numerical evidence confirms that EF-based methods within the class (4.1.1)-(4.1.2) are able to exactly integrate, within round-off error, problems whose solution is linear combination of the considered basis functions. This result also holds for large values of the stepsize: on the contrary, for the same values of the step of integration, classical methods (4.1.1)-(4.1.2) are less accurate and efficient, because in order to accurately integrate problems with oscillating solutions, classical methods require a very small stepsize, deteriorating the numerical performances in terms of efficiency.

Problem 1. We consider the following simple test equation

$$\begin{cases} y''(t) = \lambda^2 y(t), \\ y(0) = 1, \\ y'(0) = -\lambda, \end{cases} \tag{4.5.2}$$

with $\lambda > 0$ and $t \in [0, 1]$. The exact solution of this equation is $y(t) = \exp(-\lambda t)$ and, therefore, our exponentially fitted methods can exactly reproduce it, i.e. the numerical solution will be affected by the round-off error only. Table 4.1 shows the results we have obtained by using the above numerical methods.

λ	k	COLEM2	EXPCOLEM2
2	4	8.32e-1	1.09e-14
	5	2.29e-1	3.94e-14
	6	5.96e-2	1.20e-13
3	7	2.71e-1	1.06e-12
	8	6.85e-2	7.96e-12
	9	1.72e-2	5.97e-12
4	8	9.09e-1	1.83e-11
	9	2.29e-1	2.26e-11
	10	5.74e-2	1.64e-10

Table 4.1: Relative errors corresponding to the solution of the problem (4.5.2), for different values of λ and k.

Problem 2. We examine the following linear equation

$$\begin{cases} y''(t) - y(t) = t - 1, \\ y(0) = 2, \\ y'(0) = -2, \end{cases} \tag{4.5.3}$$

with $\lambda > 0$ and $t \in [0, 5]$. The exact solution is $y(t) = 1 - t + \exp(-t)$ and, therefore, it is linear combinations of all the basis functions in (4.2.6). The obtained results are reported in Table 4.2.

k	COLEM2	EXPCOLEM2
5	8.53e-1	1.65e-14
6	2.71e-1	5.16e-14
7	7.26e-2	2.21e-13

Table 4.2: Relative errors corresponding to the solution of the problem (4.5.3).

Problem 3. We next focus on the Prothero-Robinson problem [133]

$$\begin{cases} y''(t) + \nu^2[y(t) - \exp(-\lambda t)]^3 = \lambda^2 y, \\ y(0) = 1, \\ y'(0) = -\lambda, \end{cases} \tag{4.5.4}$$

in $t \in [0, 5]$, which is a nonlinear problem whose exact solution is $y(t) = \exp(-\lambda t)$. The obtained results are reported in Table 4.3.

k	COLEM2	EXPCOLEM2
1	3.65e-1	2.41e-15
2	1.70e-1	3.16e-16
3	2.65e-2	1.21e-15

Table 4.3: Relative errors corresponding to the solution of the problem (4.5.4), with $\nu = 1/10$.

5

Parameter estimation for exponentially fitted hybrid methods

5.1 Introduction

It is the purpose of this chapter to analyse the family of two-step hybrid methods

$$\begin{aligned} Y_i^{[n]} &= (1+c_i)y_n - c_i y_{n-1} + h^2 \sum_{j=1}^{s} a_{ij} f(Y_j^{[n]}), \quad i = 1, ..., s, \\ y_{n+1} &= 2y_n - y_{n-1} + h^2 \sum_{i=1}^{s} b_i f(Y_i^{[n]}), \end{aligned} \tag{5.1.1}$$

considered in the previous chapter, in order to provide a strategy for the estimation of the unknown parameters on which the coefficients of EF methods depend.

The coefficients of classical formulae are constant matrices and vectors, while, on the contrary, the coefficients of EF formulae are matrices and vectors depending on the value of a parameter to be suitably determined. This parameter depends on the solution of the problem and its behaviour: for instance, it could be the value of the frequency of the oscillations when the solution is oscillatory, or the argument of the exponential function describing the exponential decay of a certain phenomenon modelled by (1.1.1). We have introduced in [50] the family of EF-based methods (5.1.1), assuming that the unknown parameter is known in advance.

A rigorous theory for the exact computation of the parameter has not yet been developed. However, some attempts have been done in the literature (see, for instance, [111, 117] and references therein contained) in order to provide an accurate estimation of the parameter, generally based on the minimization of the leading term of the local discretization error. We aim to provide in this chapter an analogous strategy to determine an approximation to the parameter, in such a way that the performances of the corresponding EF methods are not compromised by the missing of the exact value of the parameter.

The chapter is organized as follows: Section 2 reports the constructive strategy to derive EF-based formulae within the class (5.1.1); Section 3 is devoted to the presentation of the parameter estimation technique, while Section 4 provides some numerical tests.

5.2 Exponentially fitted two-step hybrid methods

In this section we recall the constructive technique introduced in Chapter 4 to derive EF-based methods within the class (5.1.1). This strategy is based on the *six-step flow chart* introduced by Ixaru and Vanden Berghe in [94].
We associate to (5.1.1) the following $s+1$ linear operators

$$\mathcal{L}[h,\mathbf{b}]y(t) = y(t+h) - 2y(t) + y(t-h) - h^2\sum_{i=1}^{s} b_i y''(t+c_i h),$$
$$\mathcal{L}_i[h\mathbf{a}]y(t) = y(t+c_i h) - (1+c_i)y(t) + c_i y(t-h) - h^2\sum_{j=1}^{s} a_{ij} y''(t+c_j h),$$

for $i = 1,\ldots,s$, and proceed as follows:

- step (i) We compute the *starred classical moments* (see [94]) by using formulae

$$L^*_{im}(\mathbf{a}) = h^{-(m+1)}\mathcal{L}_i[h;\mathbf{a}]t^m,\ i = 1,\dots,s,\ m = 0,1,2,\dots,$$
$$L^*_{m}(\mathbf{b}) = h^{-(m+1)}\mathcal{L}[h;\mathbf{b}]t^m,\ m = 0,1,2,\dots,.$$

- step (ii) *Compatibility analysis.* We examine the algebraic systems

$$L^*_{im}(\mathbf{a}) = 0,\ i = 1,\dots,s,\ m = 0,1,\dots,M-1,$$
$$L^*_{m}(\mathbf{b}) = 0,\ m = 0,1,\dots,M'-1.$$

  to find out the maximal values of M and M' for which the above systems are compatible. Assuming $s = 2$, we discovered in [50] that such values are $M = M' = 4$.

- step (iii) *Computation of the G functions.* In order to derive EF methods, we need to compute the so-called *starred exponential moments* (see [94], p. 42), i.e.

$$E^*_{0i}(\pm z,\mathbf{a}) = \exp(\pm\mu t)\mathcal{L}_i[h,\mathbf{a}]\exp(\pm\mu t),\ i = 1,\dots,s,$$
$$E^*_{0}(\pm z,\mathbf{b}) = \exp(\pm\mu t)\mathcal{L}[h,\mathbf{b}]\exp(\pm\mu t).$$

  Once computed the reduced exponential moments, we derive the following set of functions:

$$G^+_i(Z,\mathbf{a}) = \frac{1}{2}\Big(E^*_{0i}(z,\mathbf{a}) + E^*_{0i}(-z,\mathbf{a})\Big),\quad i = 1,\dots,s,$$
$$G^-_i(Z,\mathbf{a}) = \frac{1}{2z}\Big(E^*_{0i}(z,\mathbf{a}) - E^*_{0i}(-z,\mathbf{a})\Big),\quad i = 1,\dots,s,$$
$$G^+(Z,\mathbf{b}) = \frac{1}{2}\Big(E^*_{0}(z,\mathbf{b}) + E^*_{0}(-z,\mathbf{b})\Big),$$
$$G^-(Z,\mathbf{b}) = \frac{1}{2z}\Big(E^*_{0}(z,\mathbf{b}) - E^*_{0}(-z,\mathbf{b})\Big),$$

where $Z = z^2$. In our case, the G functions take the following form

$$G_i^+(Z,\mathbf{a}) = \eta_{-1}(c_i^2 Z) + c_i\eta_{-1}(Z) - 2(1+c_i) - Z\sum_{j=1}^{s} a_{ij}\eta_{-1}(c_j^2 Z),$$

$$G_i^-(Z,\mathbf{a}) = c_i\eta_0(c_i^2 Z) - c_i\eta_0(Z) - 2(1+c_i) - Z\sum_{j=1}^{s} c_j a_{ij}\eta_0(c_j^2 Z),$$

$$G^+(Z,\mathbf{b}) = 2\eta_{-1}(Z) - 2 - Z\sum_{j=1}^{s} b_j\eta_{-1}(c_j^2 Z),$$

$$G^-(Z,\mathbf{b}) = -Z\sum_{j=1}^{s} b_j c_j\eta_0(c_j^2 Z).$$

We next compute the p-th derivatives $G^{\pm(p)}$ and $G_i^{\pm(p)}$, taking into account the formula for the p-th derivative of $\eta_k(Z)$ (see [94])

$$\eta_k^{(p)}(Z) = \frac{1}{2^p}\eta_{k+p}(Z),$$

and obtaining

$$G^{+(p)}(Z,\mathbf{b}) = \frac{1}{2^{p-1}}\eta_{p-1}(Z) - \sum_{j=1}^{s} b_j\frac{d^p}{dZ^p}\Big(Z\eta_{-1}(c_j^2 Z)\Big),$$

$$G^{-(p)}(Z,\mathbf{b}) = -\sum_{j=1}^{s} b_j c_j\frac{d^p}{dZ^p}\Big(Z\eta_{-1}(c_j^2 Z)\Big),$$

$$G_i^{+(p)}(Z,\mathbf{b}) = \frac{c_i^{2p}}{2^p}\eta_{p-1}(c_i^2 Z) + \frac{c_i}{2^p}\eta_{p-1}(Z) - \sum_{j=1}^{s} a_{ij}\frac{d^p}{dZ^p}\Big(Z\eta_{-1}(c_j^2 Z)\Big),$$

$$G_i^{-(p)}(Z,\mathbf{b}) = \frac{c_i^{2p+1}}{2^p}\eta_p(c_i^2 Z) - \frac{c_i}{2^p}\eta_p(Z) - \sum_{j=1}^{s} a_{ij}c_j\frac{d^p}{dZ^p}\Big(Z\eta_0(c_j^2 Z)\Big).$$

- step (iv) *Definition of the function basis.* We next decide the shape of the function basis to take into account: as a consequence, the corresponding method will exactly integrate (i.e. the operator $\mathcal{L}[h,\mathbf{b}]y(t)$ annihilates in correspondence of the function basis) all those problems whose solution is linear combination of the basis functions. In the exponential fitting framework, the function

basis (also known as *fitting space*) is a set of M functions of the type

$$\{1, t, \ldots, t^K, \exp(\pm\mu t), t\exp(\pm\mu t), \ldots, t^P \exp(\pm\mu t)\}, \tag{5.2.1}$$

where K and P are integer numbers satisfying the relation

$$K + 2P = M - 3 = 1. \tag{5.2.2}$$

Let us next consider the set of M' functions

$$\{1, t, \ldots, t^{K'}, \exp(\pm\mu t), t\exp(\pm\mu t), \ldots, t^{P'} \exp(\pm\mu t)\},$$

annihilating the operators $\mathcal{L}_i[h, \mathbf{a}]y(t)$, $i = 1, 2, \ldots, s$ and assume that $K' = K$ and $P' = P$, i.e. the external stage and the internal ones are exact on the same function basis.

- step (v) *Determination of the coefficients.* After a suitable choice of K and P, we next solve the following algebraic systems

$$G_i^{\pm(p)}(Z, \mathbf{a}) = 0,\ i = 1, \ldots, s,\ p = 0, ..., P,$$
$$G^{\pm(p)}(Z, \mathbf{b}) = 0,\ p = 0, ..., P,$$

and derive the coefficient of the corresponding EF-based method.

- step (vi) *Error analysis.* According to the used procedure [94], the general expression of the local truncation error for an EF method with respect to the basis of functions (5.4.2) takes the form

$$lte^{EF}(t) = (-1)^{P+1} h^M \frac{L^*_{K+1}(\mathbf{b}(Z))}{(K+1)Z^{P+1}} D^2(D^2 - \mu^2)y(t), \tag{5.2.3}$$

with K, P and M satisfying the condition (5.2.2). For the sake of completeness,

we remark that this expression of the local truncation error can be derived by using the approach of Coleman and Ixaru [35], who provided an adaptation of the theory by Ghizzetti and Ossicini (1970) to the case of EF-based formulae. This approach consists in regarding the error associated to an EF-based formula as

$$E[y] = L[y](\xi) \int_{-h}^{h} \Phi(t) dt,$$

where $\xi \in (-h, h)$ and, in our case, $L[y] = D^{k+1}(D - \mu)^{P+1} y(t)$. We observe that the kernel $\Phi(t)$ is an even function in the null space of L.

The expression of the local truncation error (5.2.3) is our starting point to estimate the unknown parameter μ.

5.3 Parameter selection

Step (vi) of the constructive procedure described above provided us the expression of the local truncation error (5.2.3), with K, P and M satisfying the condition $K + 2P = M - 3$. Taking into account that, in our case, $K = 1$, $P = 0$ and $M = 4$, we obtain

$$lte^{EF}(t) = -h^2 \frac{L_2^*(\mathbf{b}(Z))}{2\mu^2} D^2(D^2 - \mu^2) y(t). \tag{5.3.1}$$

We aim to estimate the value of the parameter μ that annihilates or minimizes the leading term of (5.2.3), by solving the equation

$$D^2(D^2 - \mu_j^2) y(t_j) = 0, \tag{5.3.2}$$

where μ_j is an approximation to the unknown parameter μ in the point t_j of the grid. We observe that the values

$$\mu_j = \pm \sqrt{\frac{y^{(iv)}(t)}{y''(t)}}\Bigg|_{t=t_j}, \quad \mu_j = \pm y''(t)\Big|_{t=t_j},$$

are solutions of (5.3.2). More generally, it can be shown that for any integer K and P, the value

$$\mu_j = \pm y''(t)\Big|_{t=t_j}$$

satisfies the reference differential equation

$$D^{K+1}(D^2 - \mu^2)^{P+1} y(t) = 0, \tag{5.3.3}$$

in every point of the grid. This situation is formalized in the following result.

Proposition 5.3.1. *For any grid point t_j, $\mu_j = \pm y''(t_j)$ is solution of* (5.3.3) *with multiplicity $P+1$, $P \geq 0$.*

Proof. Equation (5.3.3) can be regarded in the form

$$D^{K+1}\left(\sum_{i=0}^{P+1}(-1)^i\binom{P+1}{i}D^{P+1-i}\mu_j^i\right)y(t_j) = 0.$$

Therefore, in correspondence of $\mu_j = y''(t_j)$, we obtain

$$D^{K+2P+3}\left(\sum_{i=0}^{P+1}(-1)^i\binom{P+1}{i}\right)y(t_j) = 0,$$

which is always satisfied because

$$\sum_{i=0}^{P+1}(-1)^i\binom{P+1}{i} = (1+(-1))^{P+1} = 0.$$

□

These preliminary remarks confirm that Equation (5.3.3) leads to different choices to estimate the unknown parameter and, the more P is high, the more the number of possible choices increases. In order to estabilish a suitable strategy for the derivation of an appropriate and reliable estimation to the unknown parameter, we follow the lines drawn in [87] in the case of two-point boundary value problems. In particular,

we first analyze the solutions of (5.3.3) when the solution $y(t)$ belongs to the fitting space: for instance, we assume that

$$y(t) = t^q e^{\mu t}.$$

Then, the following result holds.

Theorem 5.3.1. *Let us assume that $y(t) = t^q e^{\mu t}$ is solution to the differential problem* (1.1.1)*. Then,*

$$\nu^2 = \mu^2$$

is a root of $D^{K+1}(D^2 - \mu^2)^{P+1} y(t)$ with multiplicity $P - q + 1$.

Proof. In correspondence of $y(t) = t^q e^{\mu t} = D_\mu^q e^{\mu t}$, the reference differential equation (5.3.3) assumes the form

$$D_\mu^q D_t^{K+1} (D_t^2 - \nu^2)^{P+1} e^{\mu t} = 0$$

or, equivalently,

$$D_\mu^q \mu^{K+1} (\mu^2 - \nu^2)^{P+1} e^{\mu t} = 0.$$

By using the Leibniz rule for higher derivatives, the previous formula can be expressed in the form

$$\sum_{n=0}^{q} \sum_{r=0}^{q-n} \binom{q}{n} \binom{q-n}{r} \beta_{n,r} \mu^{K+1-q+n+r} (\mu^2 - \nu^2)^{P+1-r} t^n e^{\mu t} = 0, \tag{5.3.4}$$

with

$$\beta_{n,r} = \frac{(K+1)!}{(K+1+n+r-q)!} \cdot \frac{(P+1)!}{(P+1-r)!}.$$

The thesis is obtained by observing that the left hand side of (5.3.4) has a common factor $(\mu^2 - \nu^2)^{P-q+1}$. □

This result can be exploited to estabilish a strategy for the approximation of the unkwown parameter in the coefficients of the methods. We denote by $p(\mu_j)$ the value of $D^{K+1}(D^2 - \mu_j^2)^{P+1} y(t_j)$ and apply Theorem (5.3.1), i.e. we solve at each time step the nonlinear equations $p(\mu) = 0$, $p'(\mu) = 0$, ..., $p^{(P-q+1)}(\mu) = 0$. If there exists a common solution for all these equations which is constant overall the integration interval, then the solution to the problem we are solving belongs to the fitting space and the obtained constant value is chosen as approximation to the unknown parameter μ. On the contrary, if such common solution does not exist and the values of μ_j vary along the integration interval, the solution to the approached differential problem does not belong to the fitting space and the approximation to μ we choose at each time step is the root of smallest modulus among the set of solutions of $p(\mu) = 0$, $p'(\mu) = 0$, ..., $p^{(P-q+1)}(\mu) = 0$, in order to avoid inaccurate results due to numerical instability.
This approach for the estimation of the unknown parameter will next be applied to some test cases reported in Section 5.4.

5.4 Numerical results

We present the numerical results arising from the implementation of the following methods belonging to the same family (5.1.1):

- HYB2, two-step hybrid method (5.1.1) having constant coefficients (see [32])

$$\begin{array}{c|cc} \frac{1}{\sqrt{6}} & \frac{1+\sqrt{6}}{12} & \\ -\frac{1}{\sqrt{6}} & -\frac{\sqrt{6}}{12} & \frac{1}{12} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \tag{5.4.1}$$

  with $s = 2$ and order 4;

- EXP2, one-parameter depending exponentially fitted method (5.1.1), with $s = 2$ and order 2, corresponding to the fitting space (5.2.1) with $K = 1$ and $P = 0$, i.e.

$$\{1, t, \exp(\pm\mu t)\}, \tag{5.4.2}$$

and depending on the following coefficients

$$\begin{aligned}
b_1 &= \frac{2c_2(\eta_1(Z)-1)\eta_0(c_2^2Z)}{-Z(c_1\eta_1(c_2^2Z)\eta_0(c_1^2Z) - c_2\eta_1(c_1^2Z)\eta_0(c_2^2Z))},\\
b_2 &= \frac{2c_1(\eta_1(Z)-1)\eta_0(c_1^2Z)}{Z(c_1\eta_1(c_2^2Z)\eta_0(c_1^2Z) - c_2\eta_1(c_1^2Z)\eta_0(c_2^2Z))},\\
a_{11} &= \frac{c_1(1+c_1)(c_1-3c_2-1)}{6(c_1-c_2)},\\
a_{12} &= \frac{c_1(1+3c_1+2c_1^2)}{6(c_1-c_2)},\\
a_{21} &= \frac{-c_2(1+3c_2+2c_2^2)}{6(c_1-c_2)},\\
a_{22} &= \frac{c_2(1+c_2)(3c_1-c_2+1)}{6(c_1-c_2)},
\end{aligned}$$

where $c = [c_1, c_2]^T$ is the abscissa vector.

Both methods depend on the same number of internal stages, therefore the computational cost due to the solution of the nonlinear system in the stages is the same and the numerical evidence shows their comparison in terms of accuracy. The methods are implemented with fixed stepsize

$$h = \frac{1}{2^k},$$

where k is a positive integer number. The reported experiments aim to confirm the theoretical expectations regarding the derived methods and to test the strategy of parameter estimation above described.

We consider the following problems:

- the scalar linear test equation

$$\begin{cases} y''(t) = \lambda^2 y(t), \\ y(0) = 1, \\ y'(0) = -\lambda, \end{cases} \tag{5.4.3}$$

with $\lambda > 0$ and $t \in [0, 1]$, whose exact solution is $y(t) = \exp(-\lambda t)$;

- the linear problem

$$\begin{cases} y''(t) - y(t) = t - 1, \\ y(0) = 2, \\ y'(0) = -2, \end{cases} \tag{5.4.4}$$

with $\lambda > 0$ and $t \in [0, 5]$ and exact solution $y(t) = 1 - t + \exp(-t)$, which is linear combination of the basis functions in (5.4.2);

- the Prothero-Robinson problem [133]

$$\begin{cases} y''(t) + \nu^2[y(t) - \exp(-\lambda t)]^3 = \lambda^2 y, \\ y(0) = 1, \\ y'(0) = -\lambda, \end{cases} \tag{5.4.5}$$

with $t \in [0, 5]$, whose exact solution is $y(t) = \exp(-\lambda t)$.

As regards EXP2 method, we apply the strategy described in Section 3 for the estimation of the unknown parameter μ. To achieve this purpose, we consider the

reference differential equation

$$p(\mu, t) = D^2(D^2 - \mu^2)y(t) \tag{5.4.6}$$

and, according to the Theorem (5.3.1), we determine the roots of $p(\mu_n, t_n) = 0$ of multiplicity $1 - q$, where t_n is the current step point. Equation (5.4.6) requires the computation of the second and forth derivatives of $y(t)$; we observe that such derivatives can be derived directly from the analytic formulation of the problem, in terms of the partial derivatives of the function f. In fact, $y''(t) = f(t, y(t))$ and $y^{(iv)}(t) = f_{yy}(t, y(t))(y'(t), y'(t)) + f_y(t, y(t))f(t, y(t))$, where the unkwown value of $y'(t)$ can be approximated by the finite difference

$$y'(t) \approx \frac{y(t+h) - y(t)}{h}.$$

We observe that, in order to avoid further function evaluations, we can replace the values of the derivatives appearing in (5.4.6) by the corresponding backward finite differences in the following way

$$y^{(r)}(t_{n+1}) \approx \frac{1}{h^n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} y_{n-i}. \tag{5.4.7}$$

The numerical evidence is reported in Tables 5.1, 5.2 and 5.3. The results confirm that EXP2 method is able to exactly solve the above problems within round-off error, since their solutions belong to the linear space generated by (5.4.2). The superiority of EXP2 method in terms of accuracy is visible from the experiments, which reveal that it outperforms HYB2 method on the considered test problems. Although EXP2 and HYB2 depend on the same number of stages, i.e. $s = 2$, and HYB2 has higher order of convergence, a larger computational effort is necessary for the latter to obtain the same accuracy of EXP2.

λ	k	HYB2	EXP2
2	4	1.10e-5	1.09e-14
	5	7.36e-7	8.45e-14
	6	4.76e-8	1.20e-13
3	4	4.19e-4	2.02e-14
	5	2.89e-5	2.29e-13
	6	1.90e-6	4.02e-13
4	4	9.29e-3	9.49e-14
	5	6.65e-4	6.08e-13
	6	4.43e-5	4.96e-12

Table 5.1: Relative errors corresponding to the solution of the problem (5.4.3), for different values of λ and k.

k	HYB2	EXP2
4	2.65e-4	3.34e-16
5	1.96e-5	1.87e-14
6	1.33e-6	5.16e-14

Table 5.2: Relative errors corresponding to the solution of the problem (5.4.4).

k	HYB2	EXP2
4	9.79e-2	2.43e-12
5	5.36e-3	1.79e-12
6	2.96e-4	1.35e-11

Table 5.3: Relative errors corresponding to the solution of the problem (5.4.5), with $\nu = 1$.

6

Exponentially fitted two-step Runge-Kutta methods

6.1 Introduction

This chapter is devoted to the derivation of exponentially fitted two-step Runge-Kutta (TSRK) methods for the solution of ordinary differential equations

$$\begin{cases} y'(t) = f(t, y(t)), \quad t \in [t_0, T], \\ y(t_0) = y_0 \in \mathbb{R}^d, \end{cases} \tag{6.1.1}$$

where $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a sufficiently smooth function ensuring that the corresponding problem is well posed. The class of methods we aim to consider is the family of two-step Runge-Kutta methods

$$\begin{cases} y_{n+1} = \theta y_{n-1} + (1-\theta) y_n + h \sum_{j=1}^{m} \Big(v_j f(Y_j^{[n-1]}) + w_j f(Y_j^{[n]}) \Big), \\ Y_i^{[n]} = u_i y_{n-1} + (1-u_i) y_n + h \sum_{j=1}^{m} \Big(a_{ij} f(Y_j^{[n-1]}) + b_{ij} f(Y_j^{[n]}) \Big), \end{cases} \tag{6.1.2}$$

with $i = 1, 2, \ldots, s$. In (6.1.2), y_n is an approximation of order p to $y(t_n)$, $t_n = t_0 + nh$, and $Y_i^{[n]}$ are approximations of order q to $y(t_{n-1} + c_i h)$, $i = 1, 2, \ldots, s$, where $y(t)$ is the solution to (6.1.1) and $c = [c_1, \ldots, c_s]^T$ is the abscissa vector.

TSRK methods (6.1.2) can be represented by the abscissa vector c and the table of their coefficients

$$\begin{array}{c|c|c} u & A & B \\ \hline \theta & v^T & w^T \end{array} = \begin{array}{c|ccccccccc} u_1 & a_{11} & a_{12} & \cdots & a_{1s} & b_{11} & b_{12} & \cdots & b_{1s} \\ u_2 & a_{21} & a_{22} & \cdots & a_{2s} & b_{21} & b_{22} & \cdots & b_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ u_s & a_{s1} & a_{s2} & \cdots & a_{ss} & b_{s1} & b_{s2} & \cdots & b_{ss} \\ \hline \theta & v_1 & v_2 & \cdots & v_s & w_1 & w_2 & \cdots & w_s \end{array} .$$

The peculiarity of two-step Runge-Kutta methods (6.1.2) lies in their dependency on the stage derivatives at two consecutive step points: as a consequence, "we gain extra degrees of freedom associated with a two-step scheme without the need for extra function evaluations" (see [100]), because the function evaluations $f(Y_j^{[n-1]})$ are completely inherited from the previous step and, therefore, the computational cost of these formulae only depends on the structure of the matrix B. The achieved degrees of freedom can be used in order to improve the properties of existing one-step methods, especially in terms of order of convergence and stability.

Two-step Runge-Kutta methods (6.1.2), introduced by Byrne and Lambert [13], have been extensively investigated by several authors [43, 44, 45, 100]: there is a rich bibliography on TSRK methods (fully referenced in the monograph [97]) regarding, for instance, the derivation of general order conditions by means of Albrecht approach [100], Butcher trees [23] and B-series [81], the estimation of the local truncation error [8, 143], technical issues for a variable stepsize-variable order implementation of TSRK methods [8, 60, 101, 102, 143], A-, L- and algebraically stable TSRK methods for the numerical treatment of stiff problems [38, 39, 58, 60], derivation of continuous extensions for TSRK methods [9, 11, 101] and the development of

the family of collocation based TSRK methods [39, 54, 59, 60, 61]. These methods have also been introduced in the context of Volterra integral equations [40, 42] and delay differential equations [7, 10]. However, up to now, nothing has been said concerning the development of TSRK methods based on functions other than algebraic polynomials.

The aim of this chapter is the derivation of EF-based methods within the class (6.1.2), depending on the value of a parameter to be suitably determined. Such value would be known only if the analytic expression of the exact solution could be given in advance, which is in general an unrealistic requirement. However, an approximation to the unknown parameter can also be experimentally found: nevertheless, in many cases, even when it would be possible to approximate or measure the value of the parameter through suitable experiments, the derived value would anyway suffer from the presence of errors. If the value of the parameter is not determined with a sufficient level of accuracy, the performances of the corresponding EF-based numerical method would be subject to a relevant deterioration.

The chapter is organized as follows: in Section 2 we present the constructive technique of adapted TSRK methods; Section 3 approaches the problem to estimate the unknown parameter on which the coefficients of the methods depend, while in Section 4 we analyze the linear stability properties of the derived methods. Section 5 provides numerical tests confirming the theoretical expectations. In Appendix B we report a *MATHEMATICA* script for the generation of a family of adapted TSRK methods.

6.2 Derivation of the methods

This section is devoted to the presentation of the constructive technique leading to our class of special purpose TSRK formulae. Since we aim to obtain exponentially fitted TSRK methods, we adapt to our purposes the six-step procedure introduced by Ixaru and Vanden Berghe in [94]. This procedure provides a general way to

derive EF formulae whose coefficients are expressed in a regularized way and, as a consequence, they do not suffer from numerical cancellation.

In agreement with this procedure, we first associate to the method (6.1.2) the following set of $s+1$ functional operators

$$\begin{aligned}
\mathcal{L}[h,\mathbf{a}]y(t) &= y(t+h) - \theta y(t-h) - (1-\theta)y(t) \\
&\quad - h\sum_{i=1}^{s}(v_i y'(t+(c_i-1)h) + w_i y'(t+c_i h)), \\
\mathcal{L}_i[h,\mathbf{b}]y(t) &= y(t+c_i h) - u_i y(t-h) - (1-u_i)y(t) \\
&\quad - h\sum_{j=1}^{s}\big(a_{ij}y'(t+(c_j-1)h) + b_{ij}y'(t+c_j h)\big),\ i=1,\ldots,s,
\end{aligned}$$

where

$$\mathbf{a} = \left[\begin{array}{ccc} \theta & v^T & w^T \end{array}\right], \qquad \mathbf{b} = \left[\begin{array}{ccc} u & A & B \end{array}\right].$$

Then, the constructive procedure consists in the following six steps.

- step (i) *Computation of the classical moments.* We apply the linear operators (4.2.1) and (4.2.2) to the monomials t^q, $q=0,1,\ldots$, obtaining

$$\begin{aligned}
\mathcal{L}[h,\mathbf{a}]t^q &= h^q L_q^*(\mathbf{a}), \\
\mathcal{L}_i[h,\mathbf{b}]t^q &= h^q L_{iq}^*(\mathbf{b}),
\end{aligned}$$

  where

$$\begin{aligned}
L_q^*(\mathbf{a}) &= 1 + (-1)^{q+1}\theta - q\sum_{i=1}^{s}(v_i(c_i-1)^{q-1} + w_i c_i^{q-1}), \\
L_{iq}^*(\mathbf{b}) &= c_i^q + (-1)^{q+1}u_i - q\sum_{j=1}^{s}(a_{ij}(c_j-1)^{q-1} + b_{ij}c_j^{q-1}),
\end{aligned}$$

  with $i=1,\ldots,s$, $q=0,1,\ldots$, are the so-called *reduced classical moments* (compare [94]).

- step (ii) *Compatibility analysis.* We examine the linear systems

$$L_q^*(\mathbf{a}) = 0,\ q = 0, 1, \ldots, M' - 1, \tag{6.2.1}$$

$$L_{iq}^*(\mathbf{b}) = 0,\ i = 1, \ldots, s,\ q = 0, 1, \ldots, M - 1 \tag{6.2.2}$$

to determine the maximal values of the integers M and M' such that the above systems are compatible. By defining

$$d_k = \begin{cases} c_k - 1, & k = 1, 2, \ldots, s, \\ \\ c_{k-s}, & k = s+1, \ldots, 2s, \end{cases} \tag{6.2.3}$$

the following result holds.

Theorem 6.2.1. *Assume $M = M' = 2s + 2$, $d_i \neq d_j$ for $i \neq j$, and*

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \prod_{j=1}^{s} \left((t - c_j)^2 - \frac{1}{4} \right) dt \neq 0. \tag{6.2.4}$$

Then, the linear systems (6.2.1) and (6.2.2) admit a unique solution.

Proof. In correspondence of the value $M = 2s + 3$, the system (6.2.1) in the unknowns θ, v_i and w_i, $i = 1, 2, \ldots, s$, takes the form

$$\begin{cases} L_0^*(\mathbf{a}) = 0, \\ \\ L_1^*(\mathbf{a}) = 1 + \theta - \displaystyle\sum_{i=1}^{s} (v_i + w_i) = 0, \\ \\ L_2^*(\mathbf{a}) = 1 - \theta - 2 \displaystyle\sum_{i=1}^{s} (v_i(c_i - 1) + w_i c_i) = 0, \\ \\ \vdots \\ \\ L_{2s+1}^*(\mathbf{a}) = 1 + \theta - (2s + 1) \displaystyle\sum_{i=1}^{s} (v_i(c_i - 1)^{2s} + w_i c_i^{2s}) = 0, \end{cases}$$

or, equivalently,

$$\begin{cases} \theta - \sum_{i=1}^{s}(v_i + w_i) = -1, \\ -\theta - 2\sum_{i=1}^{s}(v_i(c_i - 1) + w_i c_i) = -1, \\ \vdots \\ \theta - (2s+1)\sum_{i=1}^{s}(v_i(c_i - 1)^{2s} + w_i c_i^{2s}) = -1, \end{cases}$$

which is a linear system with coefficient matrix

$$H = \begin{bmatrix} 1 & -1 & \dots & -1 \\ -1 & -2d_1 & \dots & -2d_{2s} \\ \vdots & \vdots & \vdots & \vdots \\ (-1)^{j-1} & -jd_1^{j-1} & \dots & -jd_{2s}^{j-1} \\ \vdots & \vdots & & \vdots \\ 1 & -(2s+1)d_1^{2s} & \dots & -(2s+1)d_{2s}^{2s} \end{bmatrix},$$

where

$$d_j = \begin{cases} c_j - 1, & 1 \le j \le s, \\ c_j, & s+1 \le j \le 2s. \end{cases}$$

In order to achieve the thesis, we need to prove that the matrix H is nonsingular

by computing its determinant:

$$\det H = -(2s+1)! \det \begin{bmatrix} -1 & 1 & \dots & 1 \\ \frac{1}{2} & d_1 & \dots & d_{2s} \\ \vdots & \vdots & & \vdots \\ -\frac{1}{2s+1} & d_1^{2s} & \dots & d_{2s}^{2s} \end{bmatrix},$$

$$= -(2s+1)! \int_0^{-1} \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ t & d_1 & \dots & d_{2s} \\ \vdots & \vdots & & \vdots \\ t^{2s} & d_1^{2s} & \dots & d_{2s}^{2s} \end{bmatrix} dt,$$

$$= -(2s+1)! \det \begin{bmatrix} 1 & \dots & 1 \\ d_1 & \dots & d_{2s} \\ \vdots & & \vdots \\ d_1^{2s-1} & \dots & d_{2s}^{2s-1} \end{bmatrix} \int_0^{-1} (t-d_1)\cdots(t-d_{2s})dt,$$

which is nonzero when $d_i \neq d_j$ for $i \neq j$ and

$$\int_0^{-1} (t-d_1)\cdots(t-d_{2s})dt = \int_{-\frac{1}{2}}^{\frac{1}{2}} \prod_{j=1}^{s} \left((t-c_j)^2 - \frac{1}{4}\right) dt \neq 0.$$

By analogous arguments, we obtain that the system (6.2.2), $i = 1, 2, \dots, s$,

assumes the form

$$\begin{cases} u_i - \sum_{j=1}^{s}(a_{ij} + b_{ij}) = -c_i, \\ -u_i - 2\sum_{j=1}^{s}(a_{ij}(c_j - 1) + b_{ij}c_j) = -c_i^2, \\ \vdots \\ u_i - (2s+1)\sum_{j=1}^{s}(a_{ij}(c_j - 1)^{2s} + b_{ij}c_j^{2s}) = -c_i^{2s+1}, \end{cases}$$

and its coefficient matrix is the matrix H. Therefore, the same analysis as previous holds. □

We observe that condition (6.2.4) is equivalent to $c \neq \pm\frac{\sqrt{6}}{6}$ when $s = 1$, while in the case $s = 2$ it provides the following constraint

$$c_1 \neq \frac{5c_2 \pm \sqrt{5(30c_2^4 - 6c_2^2 + 1)}}{5 - 30c_2^2}. \tag{6.2.5}$$

- step (iii) *Computation of the G functions.* In order to derive EF methods, we need to compute the so-called *reduced (or starred) exponential moments* (see [94], p. 42), i.e.

$$E_0^*(\pm z, \mathbf{a}) = \exp(\pm\mu t)\mathcal{L}[h, \mathbf{a}]\exp(\pm\mu t),$$
$$E_{0i}^*(\pm z, \mathbf{b}) = \exp(\pm\mu t)\mathcal{L}_i[h, \mathbf{b}]\exp(\pm\mu t), i = 1, \ldots, s.$$

Once the reduced exponential moments have been computed, we can derive the G functions, defined in the following way:

$$G_i^+(Z, \mathbf{b}) = \frac{1}{2}(E_{0i}^*(z, \mathbf{b}) + E_{0i}^*(-z, \mathbf{b})),\ i = 1, \ldots, s,$$
$$G_i^-(Z, \mathbf{b}) = \frac{1}{2z}(E_{0i}^*(z, \mathbf{b}) - E_{0i}^*(-z, \mathbf{b})),\ i = 1, \ldots, s,$$

$$G^+(Z,\mathbf{a}) = \frac{1}{2}(E_0^*(z,\mathbf{a}) + E_0^*(-z,\mathbf{a})),$$
$$G^-(Z,\mathbf{a}) = \frac{1}{2z}(E_0^*(z,\mathbf{a}) - E_0^*(-z,\mathbf{a})),$$

where $Z = z^2$. In our case, the G functions take the following form

$$\begin{aligned}
G^+(Z,\mathbf{a}) &= \eta_{-1}(Z) - \theta\eta_{-1}(Z) - Z\sum_{i=1}^{s}(v_i(c_i-1)\eta_0((c_i-1)^2Z) \\
&\quad + w_i c_i \eta_0(c_i^2 Z)) - (1-\theta), \\
G^-(Z,\mathbf{a}) &= \eta_0(Z) + \theta\eta_0(Z) - \sum_{i=1}^{s}\left(v_i\eta_{-1}((c_i-1)^2Z) + w_i\eta_{-1}(c_i^2Z)\right), \\
G_i^+(Z,\mathbf{b}) &= \eta_{-1}(c_i^2Z) - u_i\eta_{-1}(Z) - Z\sum_{j=1}^{s}(a_{ij}(c_j-1)\eta_0((c_j-1)^2Z) \\
&\quad + b_{ij}c_j\eta_0(c_j^2Z)) - (1-u_i), \\
G_i^-(Z,\mathbf{b}) &= c_i\eta_0(c_i^2Z) + u_i\eta_0(Z) - \sum_{j=1}^{s}(a_{ij}\eta_{-1}((c_j-1)^2Z) + b_{ij}\eta_{-1}(c_j^2Z)).
\end{aligned}$$

We observe that the above expressions depend on the functions $\eta_{-1}(Z)$ and $\eta_0(Z)$ (compare [88, 94]), which are defined as follows: when Z is real, they assume the form

$$\eta_{-1}(Z) = \begin{cases} \cos(|Z|^{1/2}) & \text{if } Z \le 0 \\ \cosh(Z^{1/2}) & \text{if } Z > 0 \end{cases}, \quad \eta_0(Z) = \begin{cases} \sin(|Z|^{1/2})/|Z|^{1/2} & \text{if } Z < 0 \\ 1 & \text{if } Z = 0 \\ \sinh(Z^{1/2})/Z^{1/2} & \text{if } Z > 0 \end{cases},$$

while, when Z is complex, they take the form

$$\eta_{-1}(Z) = \cos(iZ^{1/2}), \quad \eta_0(Z) = \begin{cases} \sin(iZ^{1/2})/Z^{1/2} & \text{if } Z \neq 0 \\ 1 & \text{if } Z = 0 \end{cases},$$

or, equivalently,

$$\begin{aligned}
\eta_{-1}(Z) &= \frac{1}{2}[\exp(Z^{1/2}) + \exp(-Z^{1/2})], \\
\eta_0(Z) &= \begin{cases} \dfrac{1}{2Z^{1/2}}[\exp(Z^{1/2}) - \exp(-Z^{1/2})] & \text{if } Z \neq 0 \\ 1 & \text{if } Z = 0 \end{cases}.
\end{aligned}$$

We next compute the p-th derivatives $G^{\pm(p)}$ and $G_i^{\pm(p)}$, taking into account the formula for the p-th derivative of $\eta_k(Z)$ (see [94])

$$\eta_k^{(p)}(Z) = \frac{1}{2^p}\eta_{k+p}(Z),$$

and obtain

$$\begin{aligned}
G^{+(p)}(Z,\mathbf{a}) &= \frac{(1-\theta)}{2^p}\eta_{p-1}(Z) - \sum_{i=1}^{s}(v_i(c_i-1)\frac{d^p}{dZ^p}(Z\eta_0((c_i-1)^2Z)) \\
&\quad + w_i c_i \frac{d^p}{dZ^p}(Z\eta_0(c_i^2 Z))), \\
G^{-(p)}(Z,\mathbf{a}) &= \frac{(1+\theta)}{2^p}\eta_p(Z) - \sum_{i=1}^{s}(v_i\frac{d^p}{dZ^p}\eta_{-1}((c_i-1)^2Z) + w_i\frac{d^p}{dZ^p}\eta_{-1}(c_i^2Z)), \\
G_i^{+(p)}(Z,\mathbf{b}) &= \frac{c_i^{2p}}{2^p}\eta_{p-1}(c_i^2Z) + \frac{u_i}{2^p}\eta_{p-1}(Z) - \sum_{j=1}^{s}(a_{ij}(c_j-1)\frac{d^p}{dZ^p}(Z\eta_0((c_j-1)^2Z)) \\
&\quad + b_{ij}c_j\frac{d^p}{dZ^p}(Z\eta_0(c_j-1)^2Z))), \\
G_i^{-(p)}(Z,\mathbf{b}) &= \frac{c_i^{2p+1}}{2^p}\eta_p(c_i^2Z) + \frac{u_i}{2^p}\eta_p(Z) - \sum_{j=1}^{s}(a_{ij}\frac{d^p}{dZ^p}\eta_{-1}((c_j-1)^2Z) \\
&\quad + b_{ij}\frac{d^p}{dZ^p}\eta_{-1}(c_j^2Z)).
\end{aligned}$$

- step (iv) *Definition of the function basis.* We now consider the choice of the basis functions to take into account, i.e. we choose the set of functions annihilating the linear operators (4.2.1) and (4.2.2). As a consequence, the corresponding method exactly integrates all those problems whose solution is linear

combination of the chosen basis functions. In general, the set of M functions is a collection of both powers and exponentials, i.e.

$$\{1, t, \dots, t^K, \exp(\pm\mu t), t\exp(\pm\mu t), \dots, t^P \exp(\pm\mu t)\}, \tag{6.2.6}$$

where K and P are integer numbers satisfying the relation

$$K + 2P = M - 3 = 2s - 1. \tag{6.2.7}$$

Let us next consider the set of M' functions

$$\{1, t, \dots, t^{K'}, \exp(\pm\mu t), t\exp(\pm\mu t), \dots, t^{P'} \exp(\pm\mu t)\}, \tag{6.2.8}$$

annihilating the operators $\mathcal{L}_i[h, \mathbf{b}]y(t)$, $i = 1, 2, \dots, s$, and assume that $K' = K$ and $P' = P$, i.e. the external stages and the internal ones are exact for the same basis functions.

- step (v) *Determination of the coefficients.* After a suitable choice of K and P, we next solve the following algebraic systems:

$$\begin{aligned} &G_i^{\pm(p)}(Z, \mathbf{a}) = 0,\ i = 1, \dots, s,\ p = 0, ..., P, \\ &G^{\pm(p)}(Z, \mathbf{b}) = 0,\ p = 0, ..., P. \end{aligned}$$

  In the numerical experiments we will consider EF-based TSRK methods (6.1.2) with $s = 2$, $K = 1$ and $P = 1$, i.e. we choose the function basis

$$\{1, t, \exp(\pm\mu t), t\exp(\pm\mu t)\}. \tag{6.2.9}$$

  The coefficients of the resulting methods are reported in Appendix B.

- step (vi) *Error analysis.* According to the six-step procedure [94], the expres-

sion of the local truncation error for an EF method with respect to the basis of functions (6.2.6) takes the form

$$lte^{EF}(t) = (-1)^{P+1} h^M \frac{L^*_{K+1}(\mathbf{a}(Z))}{(K+1)! Z^{P+1}} D^{k+1}(D-\mu)^{P+1} y(t), \qquad (6.2.10)$$

with K, P and M satisfying the condition (6.2.7).

For the sake of completeness, we remark that this expression of the local truncation error can be derived by using the approach of Coleman and Ixaru [35], who provided an adaptation of the theory by Ghizzetti and Ossicini (1970) to the case of EF-based formulae. This approach consists in regarding the error associated to an EF-based formula as

$$E[y] = L[y](\xi) \int_{-h}^{h} \Phi(t) dt,$$

where $\xi \in (-h, h)$ and, in our case, $L[y] = D^{k+1}(D-\mu)^{P+1} y(t)$. We observe that the kernel $\Phi(t)$ is an even function in the null space of L. The analysis of the local error associated to the developed test case $K = P = 1$, introduced in step (v), is reported in Appendix B.

The local error analysis also provides a starting point for the estimation of the unknown parameter μ which is, in general, a nontrivial problem. In fact, up to now, a rigorous theory for the exact computation of the parameter μ has not yet been developed, but several attempts have been done in the literature in order to provide an accurate estimation (see [94] and references therein), generally based on the minimization of the leading term of the local discretization error. Our attempts to estimate the unknown parameter is described in the following section.

6.3 Parameter estimation

Step (vi) of the constructive strategy described in Section 2 provided us the expression of the local truncation error

$$lte^{EF}(t) = (-1)^{P+1} h^M \frac{L^*_{K+1}(\mathbf{a}(Z))}{(K+1)! Z^{P+1}} D^{k+1}(D-\mu)^{P+1} y(t). \tag{6.3.1}$$

We aim to estimate the value of the parameter μ that annihilates or minimizes the leading term of (6.3.1), by solving the equation

$$D^{K+1}(D-\mu_j)^{P+1} y(t_j) = 0, \tag{6.3.2}$$

where μ_j is an approximation to the unknown parameter μ in the point t_j of the grid. We observe that (6.3.2) is a nonlinear equation of degree $P+1$ in μ_j: if $P = 0$, (6.3.2) assumes the form

$$(D^{K+2} - \mu_j D^{K+1}) y(t_j) = 0,$$

whose unique solution is

$$\mu_j = \frac{D^{K+2} y(t_j)}{D^{K+1} y(t_j)}, \tag{6.3.3}$$

while, for any $P \geq 1$, Equation (6.3.2) admits $P+1$ solutions among whom we aim to search for the best approximation of the unknown parameter. In order to determine such an appropriate and reliable estimation to the unknown parameter, we follow the lines drawn in [87] in the case of two-point boundary value problems. In particular, we first analyze the solutions of (6.3.2) when the solution $y(t)$ belongs to the fitting space: for instance, we assume that

$$y(t) = t^q e^{\mu t}.$$

The following result holds

Theorem 6.3.1. *Assuming that $y(t) = t^q e^{\mu t}$ is solution of the problem* (6.1.1)*, then $\nu = \mu$ is a root of multiplicity $P - q + 1$ of* (6.3.2)*.*

Proof. In correspondence of $y(t) = t^q e^{\mu t} = D_\mu^q e^{\mu t}$, Equation (6.3.2) assumes the form

$$D_\mu^q D_t^{K+1} (D_t - \nu)^{P+1} e^{\mu t} = 0,$$

or, equivalently,

$$D_t^{K+1} D_\mu^q (\mu - \nu)^{P+1} e^{\mu t} = 0. \tag{6.3.4}$$

We observe that $\nu = \mu$ is a root of multiplicity $P + 1$ of (6.3.2). Therefore, after q differentiations, we obtain that $\nu = \mu$ is a root of multiplicity $P - q + 1$ of (6.3.4). □

The above theorem can be interpreted as follows: by denoting the left hand side of (6.3.2) as $p^{(P)}(\mu_j)$, if the solution $y(t)$ to the problem (6.1.1) belongs to the fitting space, then by solving the nonlinear equations $p^{(P)}(\mu) = 0$, $p^{(P+1)}(\mu) = 0$, ..., $p^{(P-q+1)}(\mu) = 0$ we will obtain a constant solution overall the integration interval for each equation, which will next be chosen as approximation to the unknown parameter μ. On the contrary, if we obtain a nonconstant solution for the above equations, i.e. values of μ_j varying along the integration interval, we can next conclude that the function $y(t)$ does not belong to the fitting space and we will assume as approximation to the parameter μ the smallest modulus, in order to avoid inaccurate results due to numerical instability.

This approach for the estimation of the unknown parameter will next be applied to some test cases reported in Section 6.5.

6.4 Linear stability analysis

In this section we aim to carry out the linear stability analysis of the adapted formulae (6.1.2), by considering the linear scalar problem

$$y' = \lambda y,$$

where λ is a complex parameter such that $\mathrm{Re}(\lambda) < 0$. We recall that, for TSRK methods (6.1.2), the stability matrix assumes the form ([97])

$$\mathbf{M}(\omega, z) = \begin{bmatrix} 1-\theta+\omega w^T Q(\omega)(e-u) & \theta+\omega w^T Q(\omega)u & \omega(v^T+\omega w^T Q(\omega)A) \\ 1 & 0 & 0 \\ Q(\omega)(e-u) & Q(\omega)u & \omega Q(\omega)A \end{bmatrix},$$

where $\mathbf{M}(\omega, z) \in \mathbb{R}^{(s+2)\times(s+2)}$, $\omega = \lambda h \in \mathbb{C}$ and $Q(\omega) = (I - \omega B)^{-1}$. This matrix is then used in order to determine the three-dimensional stability region (compare with [55]) in the parameter space ($\mathrm{Re}(\omega)$, $\mathrm{Im}(\omega)$, z), thus extending a concept introduced in [34] for second order differential equations coherently with the following definition.

Definition 6.4.1. *The region of the three-dimensional* $(\mathrm{Re}(\omega), \mathrm{Im}(\omega), z)$ *space on which the inequality*

$$\rho(\mathbf{M}(\omega, z)) < 1, \tag{6.4.1}$$

is satisfied is called a region of stability Ω *for the method (6.1.2).*

Some examples of stability regions are reported in Appendix B.

6.5 Numerical results

We now present some numerical experiments in order to test the performances of

the derived methods on some differential problems (6.1.1). Such numerical evidence is based on the implementation of the following methods:

- EF5: exponentially fitted TSRK method (6.1.2), which can be generated using the MATHEMATICA script, reported in Appendix B. This method has order and stage order 5, and depends on the value of the parameter μ to be estimated;

- TSRK5: two-stage TSRK method (compare with [54]) with abscissa vector $c = [1/2, 1]^T$ and Butcher tableau

$$\begin{array}{c|c|c} u & A & B \\ \hline \theta & v^T & w^T \end{array} = \begin{array}{ccc|cc} -\frac{11}{232} & -\frac{39}{464} & \frac{9}{29} & \frac{111}{464} & -\frac{3}{232} \\ \frac{1}{29} & \frac{4}{87} & \frac{4}{29} & \frac{20}{29} & \frac{14}{87} \\ \hline \frac{1}{29} & \frac{4}{87} & \frac{4}{29} & \frac{20}{29} & \frac{14}{87} \end{array},$$

  of order and stage order 5, with constant coefficient matrices.

As far as EF5 is concerned, with the aim to apply the strategy described in Section 3 for the selection of the parameter μ, we will treat Equation (6.3.2) not only by exactly computing the values of the derivatives appearing in such equation through the required evaluations of the f function, but also, for the sake of comparison, by approximating them through backward finite differences in the following way

$$y^{(r)}(t_{n+1}) \approx \frac{1}{h^r} \sum_{i=0}^{r} (-1)^i \binom{r}{i} y_{n-i}. \tag{6.5.1}$$

We observe that both methods possess the same order and stage order and, due to the equal number of stages, they have the same computational cost. We implement both methods in a fixed stepsize environment, with stepsize $h = 1/2^k$, with k positive integer number. Similarly as in [83], in order to reduce the influence of round-off

errors, we have reformulated the implemented methods as follows

$$\begin{cases} Z_i^{[n]} &= h\sum_{j=1}^{s} a_{ij} f(Z_j^{[n-1]} + \theta y_{n-1} + (1-\theta)y_n) \\ &\quad + h\sum_{j=1}^{s} b_{ij} f(Z_j^{[n]} + \theta y_{n-1} + (1-\theta)y_n) \\ y_{n+1} &= \theta y_{n-1} + (1-\theta)y_n + h\sum_{j=1}^{s} v_j f(Z^{[n-1]} + \theta y_{n-1} + (1-\theta)y_n) \\ &\quad + h\sum_{j=1}^{s} w_j f(Z^{[n]} + \theta y_{n-1} + (1-\theta)y_n), \end{cases} \tag{6.5.2}$$

where

$$Z_i^{[n]} = Y_i^{[n]} - \theta y_{n-1} - (1-\theta)y_n,$$

$i = 1, 2, \ldots, s$, and solved the nonlinear system in $Z_i^{[n]}$ with Newton iterations.

We apply the above methods on the following problems:

- *Problem 1.* The Prothero-Robinson equation [133]

$$\begin{cases} y'(t) = \varepsilon\big(y(t) - F(t)\big) + F'(t), \quad t \in [1,5], \\ y(t_0) = y_0, \end{cases}$$

  where $\mathrm{Re}(\varepsilon) < 0$ and $F(t)$ is a slowly varying function on the integration interval. In our experiments, we have considered $F(t) = te^{-2t}$ and y_0 such that the exact solution is $y(t) = te^{-2t}$. As observed by Hairer and Wanner [83] in the context of Runge-Kutta methods this equation provides much insight into the behaviour of numerical methods for stiff problems. This equation with $F(t) = e^{-2t}$, was also used by Butcher [18] to investigate order reduction for s-stage Runge-Kutta-Gauss methods of order $p = 2s$;

- *Problem 2.* The nonlinear system

$$\begin{cases} y_1'(t) = -y_1(t) + y_2(t)(1 - y_1(t) - y_2(t)), \quad t \in [1, 2], \\ y_2'(t) = y_1(t) - y_2(t)(1 - y_1(t)) - e^{-t}, \\ y_1(t_0) = \dfrac{1}{e}, \quad y_2(t_0) = 0, \end{cases}$$

  whose exact solution is $y(t) = [e^{-t}, 0]^T$. This problem provides the nonstiff version of the system considered in [74].

We observe that the exact solutions of both problems belong to our chosen fitting space (6.2.9) and, therefore, EF5 method is able to exactly integrate them, within the round-off error. Of course, in order to apply such method, an estimation to the parameter μ on which it depends is required: even if in the considered test cases the exact solution is given (and, therefore, the value of the parameter is known), we assume that the exact value of the parameter cannot be a priori recognized. In order to derive an estimate to such value, we apply the approach for the parameter selection reported in Section 3, with the aim to test its effectiveness.

Numerical results are collected in tables which report, for the EF5 method,

- the value $e_h^{EX}(T)$ of the global error at the endpoint of integration, when the exact value of μ is used;
- the value $e_h^{DER}(T)$ of the global error at the endpoint of integration, when the estimation of μ is obtained by solving Equation (6.3.2) using the exact values of the involved derivatives, through the evaluation of the f function;
- the (relative) error $e_\mu^{DER}(T)$ associated to the estimated parameter by using exact derivatives in (6.3.2);
- the global error $e_h^{DIFF}(T)$ in the endpoint of integration, when the derivatives

involved in Equation (6.3.2) are approximated by means of backward finite differences;

- the (relative) error $e_{\mu}^{DIFF}(T)$ associated to the estimation of the parameter by approximating the derivatives in (6.3.2) through backward finite differences.

For the TSRK5 method, the tables report the value $e_h^{TSRK5}(T)$ of the global error at the endpoint of integration.

Concerning Problem 1, taking into account the expression ((.0.4)) of local truncation error associated to the EF5 method (reported in Appendix B), the approach described in Section 3 suggests us to assume as objective functions to be annihilated at each step point the algebraic polynomials

$$\begin{aligned} p^{(1)}(\mu) &= D^2(D-\mu)^2 y(t)\big|_{t=t_n} = (D^4 - 2\mu D^3 + \mu^2 D^2) y(t)\big|_{t=t_n}, \\ p^{(2)}(\mu) &= D^2(D-\mu)^3 y(t)\big|_{t=t_n} = (D^5 - 3\mu D^4 + 3\mu^2 D^3 - \mu^3 D^2) y(t)\big|_{t=t_n}, \end{aligned} \tag{6.5.3}$$

in the unknown μ, assuming that t_n is the current step point. We denote as μ_n the selected value of the parameter at the step point t_n, which is the minimum root among the zeros of $p^{(1)}(\mu)$ and $p^{(2)}(\mu)$ derived by using the Matlab command `roots`. Table 6.1 shows the obtained results, associated to the value $\varepsilon = -10$. The numerical evidence confirms that the method EF5 is able to exactly integrate this problem within round-off error, since its solution falls in the fitting space, and shows better accuracy with respect to the TSRK5 method. The approach described in Section 3 provides a reliable estimation to the parameter μ, which does not deteriorate the performances of the EF5 method. At this stage we want to remark that the knowledge of a reasonably accurate value of the parameter is necessary in order to avoid a dramatical deterioration of the performance of the method. To support this thesis, we focus on the errors reported in Table 6.1 obtained when the derivatives in (6.5.3) are approximated by finite differences. In this case, due to the low accuracy

	EF5					TSRK5
h	$e_h^{EX}(T)$	$e_h^{DER}(T)$	$e_\mu^{DER}(T)$	$e_h^{DIFF}(T)$	$e_\mu^{DIFF}(T)$	$e_h^{TSRK5}(T)$
$1/2^4$	9.74e-13	1.69e-12	1.30e-8	4.07e-4	2.91e-1	3.01e-6
$1/2^5$	7.16e-16	1.19e-15	2.50e-10	9.94e-8	1.43e-1	3.71e-8
$1/2^6$	2.39e-16	2.39e-16	3.26e-12	5.31e-10	6.52e-2	7.51e-10

Table 6.1: Numerical results for Problem 1

in the derivatives approximation, the corresponding inaccurate estimation to the unknown parameter causes a relevant deterioration to the performances of the EF5 method, which are even worse than the ones of the constant coefficient TSRK5 method. This also confirms the importance to reliably estimate the parameter on which the coefficients of an EF-based method depend on.

Regarding Problem 2, according to the approach described in Section 3, the objective functions to be annihilated at each step point are the algebraic polynomials

$$\begin{aligned} p^{(1)}(\mu) &= D^2(D-\mu)^2 y(t) = (D^4 - 2\mu D^3 + \mu^2 D^2) y(t), \\ p^{(2)}(\mu) &= D^2(D-\mu)^3 y(t) = (D^5 - 3\mu D^4 + 3\mu^2 D^3 - \mu^3 D^2) y(t), \\ p^{(3)}(\mu) &= D^2(D-\mu)^4 y(t) = (D^6 - 4D^5\mu + 6D^4\mu^2 - 4D^3\mu^3 + D^2\mu^4) y(t), \end{aligned} \tag{6.5.4}$$

in $t = t_n$.

Table 6.2 shows the obtained results. The numerical evidence confirms the theoretical expectation: the method EF5 exactly integrates Problem 2 within round-off error. Also in this case the superiority of EF5 on TSRK5 is evident from the obtained results. Moreover, we observe that the usage of finite differences in replacement of

	EF5					TSRK5
h	$e_h^{EX}(T)$	$e_h^{DER}(T)$	$e_\mu^{DER}(T)$	$e_h^{DIFF}(T)$	$e_\mu^{DIFF}(T)$	$e_h^{TSRK5}(T)$
$1/2^3$	2.71e-14	2.72e-14	2.20e-5	5.39e-5	16.60	5.66e-9
$1/2^4$	8.69e-15	8.78e-15	1.37e-5	5.55e-5	32.42	1.86e-10
$1/2^5$	6.18e-16	6.03e-16	1.68e-5	3.63e-5	64.18	4.54e-12

Table 6.2: Numerical results for Problem 2

the derivatives appearing in (6.5.4) causes a prominent worsening of the numerical performances, due to the inaccurate parameter selection. In fact, as it is evident from (6.5.1), such an approximation to the derivatives suffers from a severe numerical instability (compare with [134]).

We observe that a system equivalent to (6.5.4) is

$$D^k(D-\mu)^2 y(t)\big|_{t=t_n} = 0, \quad k = 2, 3, 4,$$

which can be obtained by adding μ times the previous equation to each equation. The advantage here is a degree reduction of equations to be solved, and the possibility to solve with respect to μ by computing the nullspace $[1, 2\mu, \mu^2]^T$ of a 3×3 Wronskian, possibly using an SVD approach. This alternative strategy will be the object of future analysis.

7

Some new uses of the $\eta_m(Z)$ functions

7.1 Introduction

The functions $\eta_m(Z)$, $m = -1, 0, 1, \ldots$, have been introduced in [89] to provide a convenient framework when building up CP methods for the Schrödinger equation. However, as observed later on, the area of applications is much larger, including the possibility of a systematic treatment of oscillatory functions or of functions with an exponential behaviour. In particular this set of functions has been used successfully in building up a number of approximation formulae based on the exponential fitting, see [88].

In most applications the argument Z and $\eta_m(Z)$ are real, and in these cases the η_m functions are closely related to the Bessel functions of real/imaginary argument, see property (v) below, but there are also cases when the argument and the function values are complex. Fortran subroutines for these functions are available, e.g., subroutines GEBASE in [90] and CGEBAS in [93] (up to $m = 6$), and GEBASE, GEBASEV, CGEBASE, CGEBASEV in [94]. A matlab version is in [114] and on the web-page

http://www.dmi.unisa.it/people/conte/www/codici.html.

In this work we focus on some new applications when building up formulae in the frame of the exponential fitting procedure. The coefficients in such formulae are

functions of the involved frequencies with the property that they tend to the constant values of the coefficients in the associate classical approximation formulae when the frequencies tend to zero.

To fix the ideas we concentrate on the case when only one frequency μ is involved. Its value is real, $\mu = \lambda$, for functions which behave like $\exp(\pm\lambda x)$ or like $\sinh(\lambda x)$ and $\cosh(\lambda x)$, and imaginary $\mu = i\omega$ for oscillatory functions with $\sin(\omega x)$ and $\cos(\omega x)$. In all these cases the coefficients are functions of the product $z = \mu h$ which is either real or purely imaginary. An alternative notation consists in using one and the same real parameter v defined as $v = |z| = |\mu|h$ in both cases but in this situation we have to take care that either hyperbolic or trigonometric functions are involved.

An unpleasant feature with the expressions of the coefficients in the ef-based formulae is that quite often these exhibit an undeterminacy of the form $0/0$ when $z = 0$ or $v = 0$ and therefore additional expressions consisting in power expansions in z or v must be provided for use when $|z|$ or v is smaller than some threshold value. This is how it is done in many papers, to mention only [148], [149], [24], [25], [140], [105].

In this chapter we show that the functions $\eta_m(Z)$ where $Z = (\mu h)^2$ (or, with the other notation, $Z = -v^2$ / $Z = v^2$ in the trigonometric/hyperbolic case) provide a powerful tool for eliminating the $0/0$ behaviour entirely, and develop a technique to be used for this aim. The new formulae will then cover all z or v, with no need to invoke series. One and the same expression is then enough irrespective of whether Z is positive or negative, small or big. Even more, the new expression can be used also when Z is complex.

7.2 Definition and properties of $\eta_m(Z)$ functions

These functions have been introduced in [89] as real functions of a real variable, and denoted $\overline{\xi}(Z), \overline{\eta}_0(Z), \overline{\eta}_1(Z), \ldots$. The present notation is that from [88] except for $\eta_{-1}(Z)$ which was there denoted $\xi(Z)$. Later on, [93], these functions have been extended for complex argument Z.

The functions $\eta_{-1}(Z)$ and $\eta_0(Z)$ are defined in terms of some standard functions. When Z is real the familiar trigonometric or hyperbolic functions are used:

$$\eta_{-1}(Z) = \begin{cases} \cos(|Z|^{1/2}) & \text{if } Z \leq 0 \\ \\ \cosh(Z^{1/2}) & \text{if } Z > 0 \end{cases}, \eta_0(Z) = \begin{cases} \sin(|Z|^{1/2})/|Z|^{1/2} & \text{if } Z < 0 \\ \\ 1 & \text{if } Z = 0 \\ \\ \sinh(Z^{1/2})/Z^{1/2} & \text{if } Z > 0\,. \end{cases} \tag{7.2.1}$$

Notice that when $Z < 0$ function $\eta_0(Z)$ is closely related to the sinc function, $\eta_0(Z) = \text{sinc}(\sqrt{|Z|})$.

When Z is complex the functions sin and cos of a complex argument are involved, as it follows:

$$\eta_{-1}(Z) = \cos(iZ^{1/2})\,, \quad \eta_0(Z) = \begin{cases} \sin(iZ^{1/2})/iZ^{1/2} & \text{if } Z \neq 0 \\ \\ 1 & \text{if } Z = 0\,. \end{cases} \tag{7.2.2}$$

Finally, an equivalent definition is through exponential functions of a complex argument,

$$\eta_{-1}(Z) = \frac{1}{2}[\exp(Z^{1/2}) + \exp(-Z^{1/2})]\,,$$

$$\eta_0(Z) = \begin{cases} \dfrac{1}{2Z^{1/2}}[\exp(Z^{1/2}) - \exp(-Z^{1/2})] & \text{if } Z \neq 0 \\ \\ 1 & \text{if } Z = 0\,, \end{cases} \tag{7.2.3}$$

as in [94].

The functions $\eta_m(Z)$ with $m > 0$ are further generated by recurrence

$$\eta_m(Z) = [\eta_{m-2}(Z) - (2m-1)\eta_{m-1}(Z)]/Z,\ m = \ 1,\ 2,\ 3, ... \tag{7.2.4}$$

if $Z \neq 0$, and by following values at $Z = 0$:

$$\eta_m(0) = 1/(2m+1)!!,\ m = \ 1,\ 2,\ 3, ... \tag{7.2.5}$$

Some useful properties when Z is real are as follows:

(i) Series expansion:

$$\eta_m(Z) = 2^m \sum_{q=0}^{\infty} \frac{(q+m)!}{q!(2q+2m+1)!}\, Z^q\,,\ m = \ 0,\ 1,\ 2, \ldots \tag{7.2.6}$$

(ii) Asymptotic behaviour at large $|Z|$:

$$\eta_m(Z) \approx \begin{cases} \eta_{-1}(Z)/Z^{(m+1)/2} & \text{for odd } m\,, \\ \\ \eta_0(Z)/Z^{m/2} & \text{for even } m\,. \end{cases} \tag{7.2.7}$$

(iii) Differentiation properties:

$$\eta'_m(Z) = \frac{1}{2}\eta_{m+1}(Z)\,,\ m = -1,\ 0,\ 1,\ 2,\ 3, \ldots \tag{7.2.8}$$

(iv) Generating differential equation: $\eta_m(Z),\ m = 0,\ 1,\ \ldots$ is the regular solution at $Z = 0$ of

$$Zw'' + \frac{1}{2}(2m+3)w' - \frac{1}{4}w = 0. \tag{7.2.9}$$

(v) Relation with the spherical Bessel functions:

$$\eta_m(-x^2) = \ x^{-m} j_m(x),\ m = 0,\ 1,\ \ldots \tag{7.2.10}$$

Most of these, in particular (i) and (iii), remain valid also for complex Z.

The property presented in the following theorem will be crucial for the development of the method described in the next section. It is valid irrespective of whether

Z is real or complex.

Theorem 7.2.1. *The functions $\eta_m(Z)$ satisfy the following relations*

$$\eta_m(Z) = \eta_m(0) + ZD_m(Z),\ m = -1, 0, 1, 2, 3\ldots, \tag{7.2.11}$$

where

$$D_m(Z) = \eta_m(0)\left[\frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right) - \sum_{i=1}^{m+1}(2i-3)!!\eta_i(Z)\right]. \tag{7.2.12}$$

Proof. We at first observe that, from definition (7.2.4),

$$\eta_m(Z) = \frac{\eta_{m-1}(Z) - Z\eta_{m+1}(Z)}{2m+1},\ m = 0, 1, 2, ..., \tag{7.2.13}$$

and proceed by induction on m. For $m = -1$, we have (see book [94])

$$\eta_{-1}(Z) = 1 + \frac{1}{2}Z\eta_0^2\left(\frac{Z}{4}\right) = \eta_{-1}(0) + ZD_{-1}(Z).$$

Let us suppose $m \geq 0$ and let (7.2.11)-(7.2.12) be valid for $m-1$, i. e.

$$\eta_{m-1}(Z) = \eta_{m-1}(0) + ZD_{m-1}(Z), \tag{7.2.14}$$

with

$$D_{m-1}(Z) = \eta_{m-1}(0)\left[\frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right) - \sum_{i=1}^{m}(2i-3)!!\eta_i(Z)\right]. \tag{7.2.15}$$

By substituting (7.2.14) in (7.2.13), and by using (7.2.5), which shows that $\eta_{m+1}(0) = \eta_{m-1}(0)/(2m+1)$, we have

$$\eta_m(Z) = \frac{\eta_{m-1}(0) + Z\left(D_{m-1}(Z) - \eta_{m+1}(Z)\right)}{2m+1} = \eta_{m+1}(0) + Z\frac{D_{m-1}(Z) - \eta_{m+1}(Z)}{2m+1}.$$

From (7.2.15) we have

$$\frac{D_{m-1}(Z)-\eta_{m+1}(Z)}{2m+1}=$$
$$\frac{1}{2m+1}\,\eta_{m-1}(0)\left[\frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right)-\sum_{i=1}^{m}(2i-3)!!\,\eta_i\left(Z\right)-(2m-1)!!\,\eta_{m+1}(Z)\right]=D_m(Z),$$

which concludes the proof. □

7.3 Description of the method and of the program

Let $\phi(v)$ be a linear combination of products of either trigonometric or hyperbolic functions of v (coexistence of these species is not allowed), with the property that $\phi(0)=0$. In this section we develop a method for converting this into a function of the form $v^r Z^k F(Z)$ where r and k are non-negative integers, and $F(Z)$ is a linear combination of products of $\eta(Z)$ functions of the argument $Z=-v^2$ in the trigonometric case and $Z=v^2$ in the hyperbolic case, with the property that $F(0)\neq 0$.

The advantage with such a conversion is that the elements which make the original function $\phi(v)$ vanishing at $v=0$ are now concentrated in the factor $v^r Z^k$. The applicability is straightforward: since the coefficients of the formulae built up in the frame of the exponential fitting procedure are expressed by ratios of two such functions, the use of our procedure analytically eliminates the mentioned unpleasant 0/0 behaviour.

The most general form of $\phi(v)$ to be covered by our procedure is

$$\phi(v)=\sum_{n=1}^{N}\alpha_n(v)\left[\prod_{i=1}^{l_{-1,n}}\psi_{-1}\left(\beta_i^{-1,n}v\right)\right]\left[\prod_{i=1}^{l_{0,n}}\psi_0\left(\beta_i^{0,n}v\right)\right], \tag{7.3.1}$$

where N, $l_{-1,n}$ and $l_{0,n}$ are known integers, the pair ψ_{-1}, ψ_0 is either $\psi_{-1}(v)=\cos(v)$ and $\psi_0(v)=\sin(v)$ or $\psi_{-1}(v)=\cosh(v)$ and $\psi_0(v)=\sin(v)$, $\alpha_n(v)$ are polynomial coefficients, and $\beta_i^{-1,n}$, $\beta_i^{0,n}$ are nonnegative constants.

Indeed, it can be proved that this function can be brought to the announced form,

$$\phi(v) = v^r Z^k F(Z) \quad \text{where} \quad F(0) \neq 0,$$

in which $F(Z)$ is of the form

$$F(Z) = \sum_{n=1}^{M} a_n(Z) \prod_{j=0}^{k} \left[\prod_{i=1}^{l_{j,n}} \eta_j \left(b_i^{j,n} Z \right) \right], \tag{7.3.2}$$

where $M \geq N$, $b_i^{j,n} \geq 0$ and $a_n(Z)$ is a polynomial in Z.

The first, introductory stage of the procedure consists in expressing the function $\phi(v)$ as a linear combination of products of the functions $\eta_{-1}(Z)$ and $\eta_0(Z)$, by using the definition of these functions. This means a direct replacement of $\cos(\beta v)$ or $\cosh(\beta v)$ by $\eta_{-1}(\beta^2 Z)$, and of $\sin(\beta v)$ or $\sinh(\beta v)$ by $\beta v \eta_0(\beta^2 Z)$, which leads to

$$\phi(v) = v^r f(Z), \tag{7.3.3}$$

where $f(Z)$ has the form

$$f(Z) = \sum_{n=1}^{N} a_n(Z) \left[\prod_{i=1}^{l_{-1,n}} \eta_{-1} \left(b_i^{-1,n} Z \right) \right] \left[\prod_{i=1}^{l_{0,n}} \eta_0 \left(b_i^{0,n} Z \right) \right]. \tag{7.3.4}$$

This introductory step is implemented in the Mathematica module "etaTransf" reported in the Appendix C.

Example 7.3.1. Let us consider

$$\begin{aligned} \phi(v) = &- v \cosh(\theta v)^2 + v \cosh(v/2) \cosh(2\theta v) + 2 \cosh(\theta v) \sinh(v/2) + \\ &- 2 \cosh(v/2) \cosh(2\theta v) \sinh(v/2) - \cosh(\theta v) \sinh(v) + \cosh(\theta v)^2 \sinh(v). \end{aligned}$$

The function $\phi(v)$ is of the form (7.3.1). Indeed, it contains only one species of functions (hyperbolic), and corresponds to $N = 6$, $\alpha_1(v) = -v = -\alpha_2(v)$, $\alpha_3(v) =$

$-2 = -\alpha_4(v)$, $\alpha_5(v) = -1 = -\alpha_6(v)$, $l_{-1,n} = 2$ for $n = 1, 2, 4, 6$, $l_{-1,n} = 1$ for $n = 3, 5$, $l_{0,n} = 0$ for $n = 1, 2$, $l_{0,n} = 1$ for $n = 3, 4, 5, 6$,

$$\beta_i^{-1,n} = \begin{cases} \theta & n = 1, 6, \quad i = 1, 2 \quad \textit{and} \quad n = 3, 5, \quad i = 1; \\ 2\theta & n = 2, 4, \quad i = 1, 2; \end{cases}$$

and

$$\beta_i^{0,n} = \begin{cases} 1/2 & n = 3, 4, \quad i = 1; \\ 1 & n = 5, 6, \quad i = 1. \end{cases}$$

Then, by replacing $\cosh(\theta v) = \eta_{-1}(\theta^2 Z)$, $\cosh(v/2) = \eta_{-1}(Z/4)$, $\cosh(2\theta v) = \eta_{-1}(4\theta^2 Z)$, $\sinh(v/2) = \frac{v}{2}\eta_0(Z/4)$, $\sinh(v) = v\eta_0(Z)$, we obtain the expression (7.3.3) with $r = 1$ and

$$\begin{aligned} f(Z) = {} & \eta_{-1}^2(\theta^2 Z)[\eta_0(Z) - 1] + \eta_{-1}(\theta^2 Z)[\eta_0(Z/4) - \eta_0(Z)] \\ & + \eta_{-1}(Z/4)\eta_{-1}(4\theta^2 Z)[1 - \eta_0(Z/4)]. \end{aligned} \tag{7.3.5}$$

Two situations are now possible depending on whether $f(0)$ is vanishing or not. If $f(0) \neq 0$ the procedure is stopped but if $f(0) = 0$ (as is the case also with the function in the above example) it is continued until we can express $f(Z)$ as

$$f(Z) = Z^k F(Z)$$

where $F(0) \neq 0$. The determination in advance of the value of k is important because it helps in conveniently organizing the subsequent steps of the procedure. In fact, the module "etaTransf" has also a section in which this k is evaluated.
Once k is known, an iteration scheme is activated, starting with $f^{(0)}(Z) = f(Z)$ of the form (7.3.4). Specifically, in a finite number k of steps we determine the functions

$f^{(s+1)}(Z)$ such that

$$f^{(s)}(Z) = Zf^{(s+1)}(Z), \quad s = 0, ..., k-1.$$

The final output of this iteration chain is assigned to the desired $F(Z)$, viz.: $F(Z) = f^{(k)}(Z)$. This is of the form (7.3.3) and $F(0) \neq 0$.

As a matter of fact, the form of $F(Z)$ is not unique, and different iteration procedures may result in different forms. All these forms are equivalent, of course, but it makes sense to give preferrence to the one which produces the shortest form of $F(Z)$. After comparing different conversion versions we decided to present below the one which seems the most advantageous from this point of view.

With this scheme we meet two situations:

If $k = 1$, we simply substitute in $f^{(0)}(Z)$ the expression given by (7.2.11) for $\eta_j(Z)$, $j = -1, 0$, thus determining $f^{(1)}(Z)$ with $f^{(1)}(0) \neq 0$ and in this way the conversion is completed. If $k \geq 2$, we care that the last step is slightly different from the previous ones. Thus, at each step $s = 0, ..., k-2$ (we call these regular steps), if $f^{(s)}(0) = 0$, then we define $f^{(s+1)}(Z) = f^{(s)}(Z)/Z$ but, if $f^{(s)}(0) \neq 0$, then we write

$$f^{(s)}(Z) = f_0^{(s)}(Z) + Zf_1^{(s)}(Z) + ... + Z^{M_s} f_{M_s}^{(s)}(Z),$$

where $f_0^{(s)}(Z)$ is a linear combination of products of the functions $\eta_j(Z)$, with $j = -1, 0$ for $s = 0$, and $j = 0, ..., s$ for $s > 0$, and $f_0^{(s)}(0) \neq 0$. Then we substitute in the term $f_0^{(s)}(Z)$ the expression given by (7.2.11) for $\eta_j(Z)$, thus determining the expression of $f^{(s+1)}(Z)$. In particular, at the second-last step $s = k-2$ we will have determined the expression of

$$f^{(k-1)}(Z) = f_0^{(k-1)}(Z) + Zf_1^{(k-1)}(Z) + ... + Z^{M_{k-1}} f_{M_{k-1}}^{(k-1)}(Z),$$

with $f_0^{(k-1)}(Z)$ being a linear combination of products of the functions η_j for $j =$

$1, ..., k-1$, η_0^{2j} for $j = 1, ..., k$. If $f_0^{(k-1)}(0) = 0$, then $f^{(k)}(Z) = f^{(k-1)}(Z)/Z$. If $f_0^{(k-1)}(0) \neq 0$, then at the *last step* $s = k-1$ we substitute in $f_0^{(k-1)}(Z)$ the expression given by (7.2.11) for $\eta_j(Z)$, $j = 1, ..., k-1$ and the following expression for η_0^{2j}

$$\begin{aligned}\eta_0^{2j}(Z) &= 1 + \left(\eta_0^j(Z) - 1\right)\left(\eta_0^j(Z) + 1\right) \\ &= 1 + (\eta_0(Z) - 1)\left(\eta_0^{j-1}(Z) + ... + 1\right)\left(\eta_0^j(Z) + 1\right) \\ &= 1 + ZD_0(Z)\left(\eta_0^{j-1}(Z) + ... + 1\right)\left(\eta_0^j(Z) + 1\right),\end{aligned}$$

thus determining $f^{(k)}(Z)$ with $f^{(k)}(0) \neq 0$. The desired $F(z)$ therefore is $F(Z) = f^{(k)}(Z)$. This scheme is implemented in the Mathematica module "ZpowerTransf", reported in the Appendix C. To make the scheme more transparent we come with details on cases when $f(Z)$ is of the form (7.3.4) where $a_n(Z)$ are simply constants, and $k = 1, 2, 3$. For further simplicity we also assume that the first I terms of the sum over n represent a linear combination of the values of η_{-1} with different arguments, the subsequent J terms are for a linear combination of the values of η_0, and the last term is simply a constant. Thus we have $N = I + J + 1$, with

$$\begin{aligned}&l^{-1,n} = 1, \quad l^{0,n} = 0,\ n = 1, \ldots, I, \\ &l^{-1,n} = 0, \quad l^{0,n} = 1,\ n = I+1, \ldots, I+J \\ &l^{-1,I+J+1} = 0, \quad l^{0,I+J+1} = 0,\end{aligned}$$

which can be briefly written as

$$f(Z) = \sum_{i=1}^{I} a_i \eta_{-1}(b_i Z) + \sum_{j=1}^{J} c_j \eta_0(d_j Z) + e, \tag{7.3.6}$$

where, of course, $\sum_{i=1}^{I} a_i + \sum_{j=1}^{J} c_j + e = 0$ in order to secure that $f(0) = 0$.
This is perhaps the case which is the most frequently met in current evaluations

related to the ef approach.

CASE $k = 1$. One step is only involved here and this is treated as a regular step. By substituting the expressions (7.2.11) for $\eta_{-1}(Z)$ and $\eta_0(Z)$, i.e.

$$\eta_{-1}(Z) = 1 + ZD_{-1}(Z), \quad \eta_0(Z) = 1 + ZD_0(Z), \tag{7.3.7}$$

with

$$D_{-1}(Z) = \frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right), \quad D_0(Z) = \frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right) - \eta_1\left(Z\right),$$

we obtain

$$f(Z) = \sum_{i=1}^{I} a_i + \sum_{j=1}^{J} c_j + e + Z\left(\sum_{i=1}^{I} a_i b_i D_{-1}\left(b_i Z\right) + \sum_{j=1}^{J} c_j d_j D_0\left(d_j Z\right)\right),$$

i.e., $f(Z) = Zf^{(1)}(Z)$ with

$$f^{(1)}(Z) = \sum_{i=1}^{I} \frac{a_i b_i}{2}\eta_0^2\left(\frac{b_i Z}{4}\right) + \sum_{j=1}^{J} \frac{c_j d_j}{2}\eta_0^2\left(\frac{d_j Z}{4}\right) - \sum_{j=1}^{J} c_j d_j \eta_1\left(d_j Z\right). \tag{7.3.8}$$

We then assign $F(Z) = f^{(1)}(Z)$, and this concludes the conversion procedure.

CASE $k = 2$. Here there are two steps, the regular step $s = 0$ and the final step $s = 1$. The output of the regular step is $f^{(1)}(Z)$ of eq.(7.3.8) which we write as

$$f^{(1)}(Z) = \sum_{i=1}^{I} a_i \eta_0^2\left(b_i Z\right) + \sum_{j=1}^{J} c_j \eta_1\left(d_j Z\right), \tag{7.3.9}$$

where, for simplicity of notation, we use the same name for the coefficients a_i, b_i, c_j, d_j. Of course, the coefficients a_i and c_i are related, $\sum_{i=1}^{I} a_i + \sum_{j=1}^{J} c_j = 0$.

The second step is also the last step and therefore, as explained before, we replace in (7.3.9)

$$\eta_0^2(Z) = 1 + ZD_0(Z)\left(\eta_0(Z) + 1\right), \quad \eta_1(Z) = \frac{1}{3} + ZD_1(Z), \tag{7.3.10}$$

with

$$D_0(Z) = \frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right) - \eta_1(Z), \quad D_1(Z) = \frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right) - \eta_1(Z) - \eta_2(Z). \tag{7.3.11}$$

Then we obtain

$$f(Z) = Z^2 f^{(2)}(Z),$$

where

$$\begin{aligned} f^{(2)}(Z) = &\sum_{i=1}^{I} \frac{a_i b_i}{2}\eta_0^2\left(\frac{b_i Z}{4}\right)(\eta_0(b_i Z) + 1) \\ &+ \sum_{j=1}^{J} c_j d_j \left(\frac{1}{2}\eta_0^2\left(\frac{d_j Z}{4}\right) - \eta_1(d_j Z) - \eta_2(d_j Z)\right), \end{aligned} \tag{7.3.12}$$

and this completes the procedure, with $F(Z) = f^{(2)}(Z)$.

CASE $k = 3$. In this case we have two regular steps, $s = 0, 1$. The output (7.3.8) of $s = 0$ is just accepted but the output (7.3.12) of $s = 1$ is inadequate because it has been derived by a technique for the last step. This step must be repeated anew with the technique for a regular step, i.e., we go back to an expression of type (7.3.9) for $f^{(1)}(Z)$ in which we replace

$$\eta_0(Z) = 1 + ZD_0(Z), \quad \eta_1(Z) = \frac{1}{3} + ZD_1(Z), \tag{7.3.13}$$

with $D_0(Z)$, $D_1(Z)$ given by (7.3.11), to obtain

$$f(Z) = Z^2 f^{(2)}(Z),$$

where

$$\begin{aligned} f^{(2)}(Z) &= \sum_{i=1}^{I} a_i \left(1 + b_i Z D_0(b_i Z)\right)^2 + \sum_{j=1}^{J} c_j \left(\frac{1}{3} + d_j Z D_1(d_j Z)\right) \\ &= f_0^{(2)}(Z) + Z f_1^{(2)}(Z) \end{aligned} \tag{7.3.14}$$

with

$$f_0^{(2)}(Z) = \sum_{i=1}^{I} 2a_i b_i D_0(b_i Z) + \sum_{j=1}^{J} c_j d_j D_1\left(d_j Z\right),$$
$$f_1^{(2)}(Z) = \sum_{i=1}^{I} a_i b_i^2 \left(D_0(b_i Z)\right)^2.$$

The last step follows, with its specific technique. It is applied only on $f_0^{(2)}(Z)$, which satisfies $f_0^{(2)}(0) = 0$ and has the form

$$f_0^{(2)}(Z) = \sum_{i=1}^{I} a_i \eta_0^2(b_i Z) + \sum_{j=1}^{J} c_j \eta_1\left(d_j Z\right) + \sum_{j=1}^{J} c_j \eta_2\left(d_j Z\right), \tag{7.3.15}$$

where, as before, for simplicity of notation, we use the same name for the coefficients a_i, b_i, c_j, d_j. The other term in (7.3.14) needs no extra treatment because it already contains a factor Z. We replace the expressions (7.3.10) and

$$\eta_2(Z) = \frac{1}{15} + Z D_2(Z), \tag{7.3.16}$$

$$D_2(Z) = \frac{1}{2}\eta_0^2\left(\frac{Z}{4}\right) - \eta_1\left(Z\right) - \eta_2\left(Z\right) - 3\eta_3\left(Z\right),$$

in (7.3.15) thus obtaining

$$f(Z) = Z^3 f^{(3)}(Z),$$

and this concludes the run, with $F(Z) = f^{(3)}(Z) \neq 0$.

Example 7.3.2. Let us consider the function

$$f(Z) = 2\eta_{-1}\left(\theta^2 Z\right) - 2\eta_{-1}\left(4\theta^2 Z\right),$$

with $f^{(0)}(0) = 0$, which is in the form (7.3.6). By substituting the expression (7.3.7) for $\eta_{-1}(Z)$ we obtain

$$f^{(1)}(Z) = \theta^2\left(\eta_0^2\left(\frac{\theta^2 Z}{4}\right) - 4\eta_0^2\left(\theta^2 Z\right)\right),$$

with $f^{(1)}(0) = -3\theta^2 \neq 0$.

Example 7.3.3. We consider the function of *Example 7.3.1.*

In this case $f(Z)$ is not of the form (7.3.6), then we have to apply the generical procedure. We have $k = 2$ and we start with $f^{(0)}(Z) = f(Z)$ given in (7.3.5). For $s = 0$ we have

$$f_0^{(0)}(Z) = f^{(0)}(Z),$$

and, by substituting the expression (7.2.11) with $\eta_j(Z)$, $j = -1, 0$, we obtain:

$$\begin{aligned} f^{(1)}(Z) = -\frac{Z}{64}\Bigg[&\eta_0^4\left(\frac{Z}{16}\right)\left(1 + 2Z\theta^2\eta_0^2\left(Z\theta^2\right)\right) - 2\eta_0^2\left(\frac{Z}{16}\right) \\ &\cdot\left(\eta_1\left(\frac{Z}{4}\right) + 2\theta^2\left(\eta_0^2\left(\frac{Z\theta^2}{4}\right) + \eta_0^2\left(Z\theta^2\right)\left(-4 + Z\eta_1\left(\frac{Z}{4}\right)\right)\right)\right) \\ &+ 8\theta^2\left(-\eta_0^2\left(\frac{Z}{4}\right)\eta_0^2\left(\frac{Z\theta^2}{4}\right)\left(2 + Z\theta^2\eta_0^2\left(\frac{Z\theta^2}{4}\right)\right) - 4\eta_0^2\left(Z\theta^2\right)\eta_1\left(\frac{Z}{4}\right)\right. \\ &\left.+ 2Z\theta^2\eta_0^4\left(\frac{Z\theta^2}{4}\right)\eta_1\left(Z\right) + \eta_0^2\left(\frac{Z\theta^2}{4}\right)\left(\eta_1\left(\frac{Z}{4}\right) + 4\eta_1\left(Z\right)\right)\right)\Bigg]. \end{aligned}$$

For $s = 1 = k - 1$ we have $f^{(1)}(0) = 0$ and $f^{(2)}(Z)$ is defined by

$$f^{(2)}(Z) = f^{(1)}(Z)/Z. \tag{7.3.17}$$

7.4 Applications

The coefficients and the error factor of any ef-based approximation formula are typically expressed by ratios of two functions of the form (7.3.1) and therefore they exhibit a $0/0$ behaviour at $v = 0$. To eliminate this we apply the conversion procedure described in the previous Section separately on the numerator and denominator. Finally, when evaluating the ratio Den/Num the factor $v^r Z^k$ disappears, and this eliminates the undeterminacy. In the following we report on results obtained with this technique on some coefficients derived in the papers [24], [25], [140]. All these are of the mentioned form, see eqs.(7.4.1-7.4.3) below. In particular, the case considered

in Examples 3.1 and 3.3 is just the numerator of α_3 in (7.4.1) after the mentioned expression of γ_1 has been introduced.

- In the paper [24] some sixth order symmetric and symplectic exponentially fitted modified Runge-Kutta methods of Gauss type were derived. The authors give the formulae of the coefficients in terms of hyperbolic functions. We consider three of them, chosen at random. These are

$$\begin{aligned} b_1 &= \frac{v - 2\sinh(v/2)}{2v(1-\cosh(\theta v))}, \\ \gamma_1 &= \frac{2\sinh(v/2) - v\cosh(2\theta v)}{2\sinh(v/2) - \sinh(v) + (\sinh(v) - v)\cosh(\theta v)}, \\ \alpha_3 &= \frac{\gamma_1\cosh(v/2) - \cosh(\theta v)}{v\sinh(\theta v)}, \end{aligned} \tag{7.4.1}$$

whose series expansions in powers of v are also listed in that paper for $\theta = \sqrt{15}/10$:

$$\begin{aligned} b_1 &= \frac{5}{18} + \frac{v^4}{302400} - \frac{v^6}{62208000} + \frac{17v^8}{212889600000} - \frac{15641v^{10}}{41845579776000000} + \ldots, \\ \gamma_1 &= 1 - \frac{3v^6}{56000} + \frac{649v^8}{44800000} - \frac{983177v^{10}}{275968000000} + \frac{2248000621v^{12}}{2583060480000000} + \ldots, \\ \alpha_3 &= \frac{\sqrt{15}}{30} + \frac{\sqrt{15}v^2}{3600} - \frac{71\sqrt{15}v^4}{1890000} + \frac{1849\sqrt{15}v^6}{302400000} - \frac{47169209\sqrt{15}v^8}{33530112000000} \\ &\quad + \frac{178746672227\sqrt{15}v^{10}}{523069747200000000} + \ldots. \end{aligned}$$

By applying the procedure described in the previous section we obtain the expressions

$$\begin{aligned} b_1 &= \frac{\eta_0^2\left(\frac{Z}{16}\right) - 2\eta_1\left(\frac{Z}{4}\right)}{8\theta^2\eta_0^2\left(\frac{Z\theta^2}{4}\right)}, \\ \gamma_1 &= \frac{\left(1 + 2Z\theta^2\eta_0^2\left(Z\theta^2\right)\right)\left(\eta_0^2\left(\frac{Z}{16}\right) - 2\eta_1\left(\frac{Z}{4}\right)\right)}{\eta_0^2\left(\frac{Z}{16}\right) - 2\eta_1\left(\frac{Z}{4}\right) + 2Z\theta^2\eta_0^2\left(\frac{Z\theta^2}{4}\right)\left[\eta_0^2\left(\frac{Z}{4}\right) - 2\eta_1\left(Z\right)\right]}, \\ \alpha_3 &= \frac{Num(Z)}{Den(Z)}, \end{aligned}$$

where $Num(Z)$ is given by (7.3.17), and

$$Den(Z) = 4\theta\left(\eta_0^2\left(\frac{Z}{16}\right) - 2\eta_1\left(\frac{Z}{4}\right) + 2Z\theta^2\eta_0^2\left(\frac{Z\theta^2}{4}\right)\left(\eta_0^2\left(\frac{Z}{4}\right) - 2\eta_1\left(Z\right)\right)\right)$$
$$\cdot\left(2 + Z\theta^2\left(\eta_0^2\left(\frac{Z\theta^2}{4}\right) - 2\eta_1\left(Z\theta^2\right)\right)\right).$$

Of course, the argument Z associated to v from eq.(7.4.1) is positive, $Z = v^2$. However the new formulae automatically cover also the analog of (7.4.1) for oscillatory functions, that is when the hyperbolic functions are replaced by trigonometric functions; in this case Z will be negative, $Z = -v^2$. We also mention that the new formulae are valid for *any* value of θ. Finally, the new formulae allow computing the coefficients with *uniform accuracy* for any Z.

- In paper [25] some sixth order symmetric and symplectic exponentially fitted Runge-Kutta methods of Gauss type were derived. We consider for example the coefficient

$$b_1 = \frac{\sinh(v) - 2\sinh(v/2)}{2v(\cosh(\theta v) - \cosh(2\theta v))}, \tag{7.4.2}$$

for which the authors report the Taylor expansion when $\theta = \sqrt{15}/10$

$$b_1 = \frac{5}{18} + \frac{v^4}{14400} - \frac{191v^6}{87091200} + \frac{623v^8}{8294400000} - \frac{78713v^{10}}{30656102400000} + \ldots.$$

The new formula for this coefficient, obtained by applying our procedure, is

$$b_1 = \frac{\left[\eta_0^2\left(\frac{Z}{16}\right) - 2\eta_1\left(\frac{Z}{4}\right)\right]}{8\theta^2\eta_0^2\left(\frac{Z\theta^2}{4}\right)}.$$

It has the same practical advantages as in the previous case.

- In [140] a family of four-step trigonometrically fitted methods has been derived. We focus on one of the coefficients reported there, viz.:

$$b_0 = \frac{\sin(2v) - 4v\cos(v) + 4\sin(v) - 2v}{-3v^2\sin(2v) + 4v^3\cos(v) + 2v^3}, \tag{7.4.3}$$

whose Taylor expansion is

$$b_0 = \frac{1}{15} + \frac{17v^2}{1575} + \frac{163v^4}{94500} + \frac{60607v^6}{218295000} + \frac{1697747v^8}{37837800000} + \frac{519335027v^{10}}{71513442000000} + \frac{12254045443v^{12}}{10420530120000000} + \frac{609739626367891v^{14}}{3201499468767600000000} + ...$$

The expression of this coefficient in terms of $\eta_m(Z)$ functions is:

$$b_0 = \frac{1}{3\left(\eta_0^2\left(\frac{Z}{4}\right) - 6\eta_0^2(Z) + 12\eta_1(4Z)\right)}\left(-\eta_0^2\left(\frac{Z}{4}\right)(2 + 3\eta_0(Z)) + 8\eta_0^2(Z) + 4\eta_1(Z) + 6\eta_0(Z)\eta_1(Z) - 16\eta_1(4Z) - 2\eta_2(Z) - 16\eta_2(4Z)\right).$$

The latter covers not only the trigonometric case, as in the original derivation, but also the hyperbolic case. Also the series expansion is no more needed.

In conclusion, we have presented a procedure for the conversion of formulae expressed in terms of the trigonometric functions $\sin(\omega x)$, $\cos(\omega x)$ or hyperbolic functions $\sinh(\lambda x)$, $\cosh(\lambda x)$ to forms expressed in terms of $\eta_m(Z)$ functions, introduced in [89]. The possibility of a such conversion is important in the evaluation of the coefficients of the approximation rules derived in the frame of the exponential fitting. The converted expressions allow, among others, a full elimination of the $0/0$ undeterminacy, uniform accuracy in the computation of the coefficients, and an extended area of validity for the corresponding approximation formulae. Another possible application, mentioned but not detailed in the text, consists in obtaining converted expressions for the corresponding factor in the error formula, thus making possible an evaluation of the accuracy.

Part III

Conclusions and future works

Conclusions and future works

This dissertation is focused on the derivation and the analysis of new classes of methods for the numerical treatment of second order ODEs.
We have addressed our attention on the development of an unifying framework for the numerical solution of special second order ODEs (1.1.1), by considering the family of General Linear Methods (3.2.4) for this problem. In our analysis we have gained benefit from the lines drawn in the development of the theory of GLMs for first order ODEs (compare [21, 97] and references therein contained). The introduced class of methods properly includes all the classical numerical methods already considered in the literature for the solution of (1.1.1) as special cases. We have presented the formulation of GLMs and the re-formulation of classical methods for (1.1.1) as GLMs and the main results regarding consistency, zero-stability, convergence and order conditions. These general results could be exploited, for instance, in order to easily derive new numerical methods, avoiding to define and prove from scratch their properties (e.g. the convergence analysis, which is in general tedious and often nontrivial to be carried out).

Future developments of this research will regard the linear and nonlinear stability analysis of GLMs (3.2.4), together with the practical construction of GLMs which present a better balance between order of convergence and stability properties than that shown by classical numerical methods. To this purpose, we will also deal with

the possibility to inherit the best stability properties already achieved in the literature (for instance, the ones possessed by Runge–Kutta–Nyström methods based on indirect collocation on the nodes of orthogonal polynomials, compare [49, 62, 152]), investigating the possibility to reproduce such stability properties on higher order methods.

Concerning Exponentially Fitted methods, we analyzed the family of two-step hybrid methods introduced by Coleman in [32], in order to provide their adapted versions in order to efficiently approach problems exhibiting an oscillatory behaviour, by means of exponential fitting arguments. These methods take advantage from the knowledge of the qualitative behaviour of the solution, which is supposed to be of exponential type, depending on one or two parameters; we have presented a strategy for the estimation of these parameters. This strategy, based on determining the roots of certain polynomials associated to the truncation error, is tested on some selected problems. The numerical evidence confirm the theoretical expectations on the accuracy of the derived methods and the effectiveness of the parameter estimation technique.

Future works will regard the usage of different basis of functions for the derivation of function fitted formulae belonging to the family. In fact, the only employ of non-negative powers and exponential functions in the chosen functional basis may not be completely satisfactory if the problem under consideration has an asymptotic exponential behaviour which is accompanied by a noninteger power of the independent variable for the infinite interval cases. In such cases some other basis of functions must be accordingly constructed. This construction can be based on the informations coming from the asymptotic behaviour analysis of the ODE when the independent variable goes to infinity, if the integration interval is entirely or semi infinite. For instance, a different basis set can be performed in a such a way to reproduce the same asymptotic behaviour of the exact solution. If the integration interval is finite, such asymptotic analysis should be provided for both ends of the interval. Moreover, we aim to achieve orders of convergence greater than two, as it has been done for

Runge-Kutta methods (compare [21, 82]). This may change the values of the parameters depending on the nature of the ODE under consideration.

We have developed a family of EF-based TSRK methods for the numerical integration of initial value problems (6.1.1). As in the case of hybrid methods, we have particularly focused our attention on the computation of a reasonably accurate approximation to the unknown parameter on which the coefficients of the derived methods depend. The numerical evidence have revealed that the strategy used for the estimation of the parameters allows to obtain an accuracy on the approximation to the solution of (6.1.1) which is comparable with the one we would have when the exact value of the parameter is known. The strategy of parameter selection is essentially based on solving some nonlinear systems involving evaluations of the function f at the right hand side of (6.1.1) and its derivatives. The results of the implementation show that a reasonably good estimation to the parameter can be achieved only if the exact values of such derivatives are used, while the employ of approximate derivatives through finite differences would relevantly worsen the performances of the corresponding method. We are aware of the fact that a better accuracy in the parameter selection and, as a consequence, in the approximation to the solution of (6.1.1), requires an higher computational cost. Anyway, such cost is, according to our opinion, lower than the one of a TSRK method with constant coefficients able to provide the same accuracy of our EF-based solver: in fact, such constant coefficients method should have an higher order of convergence which is inevitably achieved by a larger number of stages, with consequent heightening of the dimension (and the cost) of the nonlinear system in the stages to be solved at each time step. The computational cost of the implementation is closely linked also to the choice of the exponential fitting parameter P: in fact, the more P is increased, the more the degree of the nonlinear equations to be solved at each step becomes higher. It is necessary, in our opinion, to create a reasonable balance between the exponential

fitting parameter P and the computational effort of the solver: for instance, P can be chosen in such a way that the accuracy of the classical TSRK method with constant coefficients is remarkably improved with a comparable computational effort. This is the spirit of the choice of P we have made in Chapter 6.

Further developments of this research will regard the introduction of adapted TSRK formulae based on more general function basis: in fact, this chapter represents the first step in order to consider general function basis, following the spirit of *function fitting* (see, for instance, [123, 124]). We also aim to consider *revised* EF-formulae within the family of TSRK methods, extending the idea introduced in [57].

Appendix A

We report the coefficients of EF methods (4.1.1)-(4.1.2) with $s = 2$ and $s = 4$ with respect to the basis (4.2.6) and (4.2.7) respectively. Two-stage EF methods within the class (4.1.1)-(4.1.2) and exact on the functional basis (4.2.6) have the following coefficients:

$$b_1 = -\frac{2c_2\left(\eta_{-1}(Z)-1\right)\eta_0\left(c_2^2Z\right)}{Z\left(c_1\eta_0\left(c_1^2Z\right)\eta_{-1}\left(c_2^2Z\right)-c_2\eta_{-1}\left(c_1^2Z\right)\eta_0\left(c_2^2Z\right)\right)},$$

$$b_2 = \frac{2c_1\left(\eta_{-1}(Z)-1\right)\eta_0\left(c_1^2Z\right)}{Z\left(c_1\eta_0\left(c_1^2Z\right)\eta_{-1}\left(c_2^2Z\right)-c_2\eta_{-1}\left(c_1^2Z\right)\eta_0\left(c_2^2Z\right)\right)},$$

$$a_{11} = \frac{-c_2\eta_{-1}\left(c_1^2Z\right)+c_1\eta_0\left(c_1^2Z\right)-c_1c_2\eta_{-1}(Z)+2c_1c_2-c_1\eta_0(Z)-2c_1+2c_2-2}{Z(c_1-c_2)\eta_{-1}\left(c_1^2Z\right)},$$

$$a_{12} = \frac{c_1^2\eta_{-1}(Z)+c_1\eta_{-1}\left(c_1^2Z\right)-c_1\eta_0\left(c_1^2Z\right)-2c_1^2+c_1\eta_0(Z)+2}{Z(c_1-c_2)\eta_{-1}\left(c_2^2Z\right)},$$

$$a_{21} = \frac{c_2^2\left(-\eta_{-1}(Z)\right)-c_2\eta_{-1}\left(c_2^2Z\right)+c_2\eta_0\left(c_2^2Z\right)+2c_2^2-c_2\eta_0(Z)-2}{Z(c_1-c_2)\eta_{-1}\left(c_1^2Z\right)},$$

$$a_{22} = \frac{c_1\eta_{-1}\left(c_2^2Z\right)+c_1c_2\eta_{-1}(Z)-2c_1c_2-2c_1-c_2\eta_0\left(c_2^2Z\right)+c_2\eta_0(Z)+2c_2+2}{Z(c_1-c_2)\eta_{-1}\left(c_2^2Z\right)}.$$

It is easy to prove that, for Z tending to 0, these coefficients tend to those of two-step hybrid methods based on algebraic collocation (see [56]): therefore, applying the order conditions derived in [32] for Z tending to 0, we discover that these methods

have order 2. Fig. 1 shows an example of stability region for two-stage one-parameter depending method with $c_1 = \frac{2}{3}$, $c_2 = \frac{4}{5}$. The coefficients of four-stage EF methods

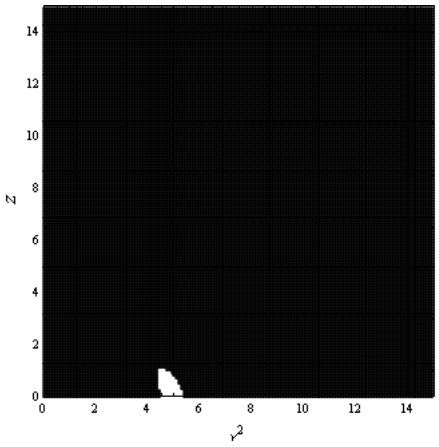

Figure 1: Region of stability in the (ν^2, Z)-plane for $s = 2$, with $c_1 = \frac{2}{3}$, $c_2 = \frac{4}{5}$.

(4.1.1)-(4.1.2) with respect to the functional basis (4.2.7) are too long to be reported in the thesis and, for this reason, we present their truncated power series expansion, in correspondence of the abscissa vector $c = [0, \frac{1}{3}, \frac{2}{3}, 1]^T$:

$$b_1 = \frac{5}{2} + \frac{43\,{Z_2}^2}{360} + \left(\frac{43}{360} + \frac{593\,{Z_2}^2}{272160}\right) {Z_1}^2 + O(Z_1^4) + O(Z_2^4),$$
$$b_2 = -\frac{15}{4} - \frac{37\,{Z_2}^2}{144} - \left(\frac{37}{144} + \frac{9643\,{Z_2}^2}{544320}\right) {Z_1}^2 + O(Z_1^4) + O(Z_2^4),$$
$$b_3 = 3 + \frac{7\,{Z_2}^2}{45} + \left(\frac{7}{45} + \frac{593\,{Z_2}^2}{136080}\right) {Z_1}^2 + O(Z_1^4) + O(Z_2^4),$$
$$b_4 = -\frac{3}{4} - \frac{13\,{Z_2}^2}{720} - \left(\frac{13}{720} + \frac{47\,{Z_2}^2}{544320}\right) {Z_1}^2 + O(Z_1^4) + O(Z_2^4),$$

$$a_{11} = 0, \qquad a_{12} = 0, \qquad a_{13} = 0, \qquad a_{14} = 0,$$
$$a_{21} = \frac{67}{81} + \frac{581 Z_2^2}{14580} + \left(\frac{581}{14580} + \frac{24001 Z_2^2}{33067440}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),$$
$$a_{22} = -\frac{71}{54} - \frac{833 Z_2^2}{9720} + \left(-\frac{833}{9720} - \frac{18607 Z_2^2}{3149280}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),$$
$$a_{23} = \frac{26}{27} + \frac{7 Z_2^2}{135} + \left(\frac{7}{135} + \frac{533 Z_2^2}{367416}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),$$
$$a_{24} = -\frac{41}{162} - \frac{35 Z_2^2}{5832} + \left(-\frac{35}{5832} - \frac{1919 Z_2^2}{66134880}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),$$

$$
\begin{aligned}
a_{31} &= \frac{539}{324} + \frac{929Z_2^2}{11664} + \left(\frac{929}{11664} + \frac{27443Z_2^2}{18895680}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{32} &= -\frac{137}{54} - \frac{37Z_2^2}{216}\left(-\frac{37}{216} - \frac{86801Z_2^2}{7348320}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{33} &= \frac{209}{108} + \frac{403Z_2^2}{3888} + \left(\frac{403}{3888} + \frac{127951Z_2^2}{44089920} + \right) Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{34} &= -\frac{41}{81} - \frac{35Z_2^2}{2916} + \left(-\frac{35}{2916} - \frac{1919Z_2^2}{33067440} + \right) Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{41} &= \frac{5}{2} + \frac{43Z_2^2}{360} + \left(\frac{43}{360} + \frac{593Z_2^2}{272160}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{42} &= -\frac{15}{4} - \frac{37Z_2^2}{144} - \left(\frac{37}{144} + \frac{9643}{139860}Z_1^2\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{43} &= 3 + \frac{7Z_2^2}{45} + \left(\frac{593Z_2^2}{136080} + \frac{7}{45}\right) Z_1^2 + O(Z_1^4) + O(Z_2^4),\\
a_{44} &= -\frac{3}{4} - \frac{13Z_2^2}{720} - \left(\frac{13}{720} + \frac{47}{544320}Z_2^2\right) Z_1^2 + O(Z_1^4) + O(Z_2^4).
\end{aligned}
$$

Also in this case, for Z_1 and Z_2 tending to 0, such coefficients tend to those of two-step hybrid methods based on algebraic collocation and the corresponding method has algebraic order 4. The tridimensional stability region of this method is reported in Fig. 2.

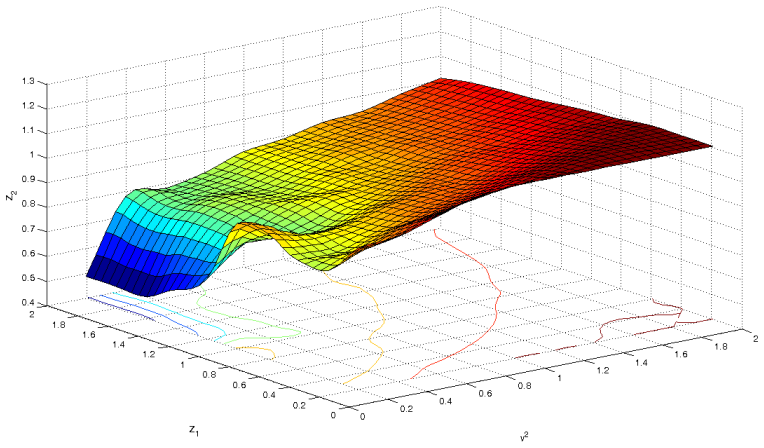

Figure 2: Region of stability in the (ν^2, Z_1, Z_2)-space for $s = 4$, with $c = [0, \frac{1}{3}, \frac{2}{3}, 1]^T$.

Appendix B

We report a MATHEMATICA script for the generation of the coefficients of the EF5 method implemented in Section 5. The method belongs to the family of TSRK formulae (6.1.2) with $s = 2$ and it has been derived in order to exactly integrate each linear combination of the function basis

$$\{1, t, \exp(\pm\mu t), t\exp(\pm\mu t)\},$$

which is a subset of (6.2.6), corresponding to $K = P = 1$. EF5 method is characterized by the abscissa vector $c = [1/2, 3/4]^T$, which provides admissible values for the nodes since they satisfy the constraint (6.2.5).

```
>> L[y[t]] :=(y[t+h]-θ*y[t-h]-(1-θ)*y[t]-h*(v1*y'[t+(c1-1)*h]
           +v2*y'[t+(c2-1)*h]+w1*y'[t+c1*h]+w2*y'[t+c2*h])
           /.t ->0;

>> For[i = 1, i<= 2, Li[y[t]]=(y[t+ci*h]-ui*y[t-h]-(1-ui)*y[t]
   - h*(ai,1*y'[t+(c1-1)*h]+ ai,2*y'[t+(c2-1)*h]+bi,1*y'[t+c1*h]
   + bi,2*y'[t+c2*h]))/.t -> 0; i++]
```

```
>> subs={y[t_] -> t, y'[t_] -> 1};
>> eq1=L[y[t]]/.subs;
>> eq2=L_1[y[t]]/.subs;
>> eq3=L_2[y[t]]/.subs;
>> {θ, u_1, u_2}={θ, u_1, u_2}/.Together[Flatten[Solve
[{eq1== 0, eq2==0, eq3==0},{θ, u_1, u_2}]]];
>> h=z/mu;
>> subs1={y[t_] -> Exp[mu*t], y'[t_] -> mu*Exp[mu*t]};
>> subs2={y[t_] ->  Exp[-mu*t], y'[t_] -> -mu*Exp[-mu*t]};
>> Gp = FullSimplify[((L[y[t]]/.subs1) + (L[y[t]]/.subs2))/2];
>> Gm = FullSimplify[((L[y[t]]/.subs1) - (L[y[t]]/.subs2))/(2*z)];
>> Dp = FullSimplify[D[Gp, z]];
>> Dm = FullSimplify[D[Gm, z]];
>> {v_1, v_2, w_1, w_2} = {v_1, v_2, w_1, w_2}/.Together[Flatten[Solve
[Gp==0, Gm==0, Dp==0, Dm=0, {v_1, v_2, w_1, w_2}]]]

>> Gp1=FullSimplify[((L_1[y[t]]/.subs1)+(L_1[y[t]]/.subs2))/2];
>> Gp2=FullSimplify[((L_2[y[t]]/.subs1)+(L_2[y[t]]/.subs2))/2];
>> Gm1=FullSimplify[((L_1[y[t]]/.subs1)-(L_1[y[t]]/.subs2))/(2*z)];
>> Gm2=FullSimplify[((L_2[y[t]]/.subs1)-(L_2[y[t]]/.subs2))/(2*z)];

>> Dp1=FullSimplify[D[Gp1, z]];
>> Dp2=FullSimplify[D[Gp2, z]];
>> Dm1=FullSimplify[D[Gm1, z]];
>> Dm2=FullSimplify[D[Gm2, z]];
>> {a_{1,1}, a_{1,2}, b_{1,1}, b_{1,2}, a_{2,1}, a_{2,2}, b_{2,1}, b_{2,2}}=
   {a_{1,1}, a_{1,2}, b_{1,1}, b_{1,2}, a_{2,1}, a_{2,2}, b_{2,1}, b_{2,2}} /.
```

```
Together[Flatten[Solve[{Gp1==0, Gm1==0, Gp2==0, Gm2==0, Dp1==0,
Dm1==0, Dp2==0, Dm2==0}, {a1,1, a1,2, b1,1, b1,2, a2,1, a2,2, b2,1, b2,2}]]]
```

The derived coefficients depend on the hyperbolic functions cosh and sinh. In order to convert them in terms of $\eta_k(Z)$-functions, the Mathematica package *formConv* can be used: it is described in [41] and can be downloaded from the web page

`http://cpc.cs.qub.ac.uk/summaries/AEFB_v1_0.html.`

It is also reported in Appendix C.

According to formula (6.3.1), the local error associated to this method is given by

$$lte^{EF}(t) = h^6 \frac{L_2^*(\mathbf{a}(Z))}{2Z^2} D^2(D-\mu)^2 y(t). \tag{.0.4}$$

With the aim to derive the order of convergence of EF5 methods, we study the behaviour of classical order conditions of TSRK methods for z tending to 0. In particular we obtain that the stage order conditions

$$\frac{c^k}{k!} - \frac{(-1)^k}{k!}u - A\frac{(c-e)^{(k-1)}}{(k-1)!} - B\frac{c^{(k-1)}}{(k-1)!},$$

tend to 0 for k up to 5 and the same happens for the order conditions

$$\frac{1}{k!} - \frac{(-1)^k}{k!}\theta - v^T\frac{(c-e)^{(k-1)}}{(k-1)!} - w^T\frac{c^{(k-1)}}{(k-1)!}.$$

Therefore, the derived method has order and stage order both equal to 5. Concerning the linear stability properties of this method, we report in Figure 3 the stability region in the $(\mathrm{Re}\ \omega, \mathrm{Im}\ \omega, z)$-space and in Figure 4 the sections through the stability region by planes $z = -1$, $z = -2$, $z = -3$ and $z = -4$.

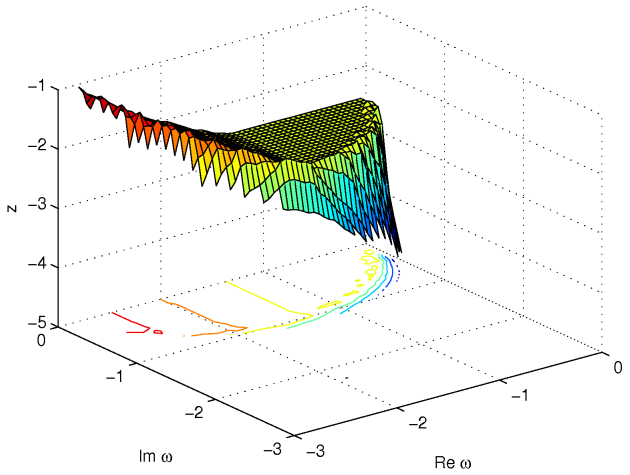

Figure 3: Stability region of the EF5 method in the (Re ω, Im ω, z)-space.

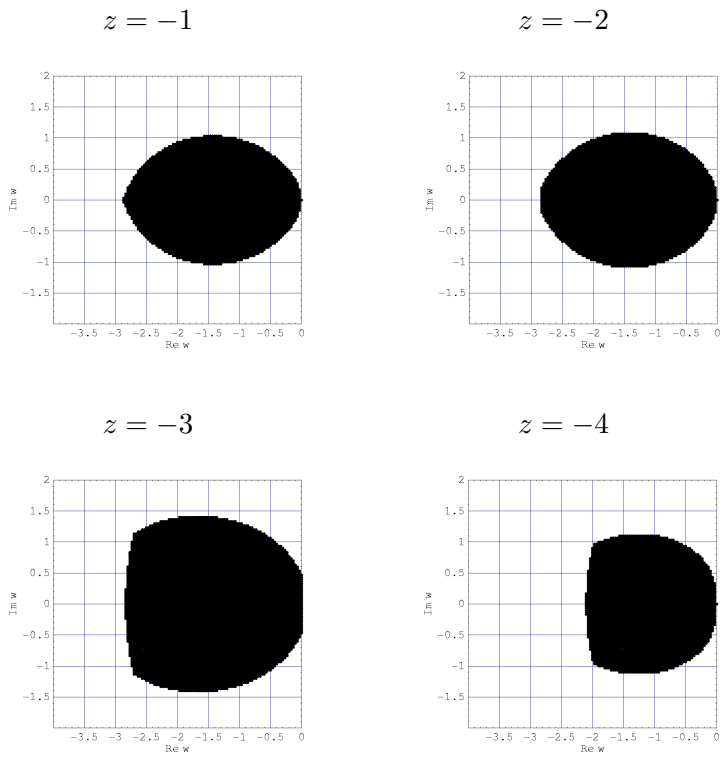

Figure 4: Projection of the stability region in the (Re ω, Im ω)-plane for the adapted TSRK method above reported.

Appendix C

The main program which follows applies the conversion procedure to the coefficient b_1 in (7.4.1).

(* PROGRAM formConv : converts a rational formula, containing oscillatory or hyperbolic functions, in terms of $\eta_m(Z)$ functions and allows a full elimination of the 0/0 undeterminacy *)

Num $= v - 2\sinh(v/2)$;
Den $= 2v(1 - \cosh(\theta v))$;
type = 2 ;
Off[General::spell]
degNum = vDeg[Num]; {NumEta, kNum, rNum} = etaTransf[Num, degNum];
NumNew = ZpowerTransf[NumEta, kNum];
degDen = vDeg[Den]; {DenEta, kDen, rDen} = etaTransf[Den, degDen];
DenNew = ZpowerTransf[DenEta, kDen];
Print["Transformed coefficient:", FullSimplify[$\frac{NumNew}{DenNew}$]]

Below are listed the MATHEMATICA main program *formConv.nb* and modules *vDeg, etaTransf* and *ZpowerTransf.*

Main program: formConv.nb

(* PROGRAM formConv : converts a rational formula, containing oscillatory or hyperbolic functions, in terms of $\eta_m(Z)$ functions and allows a full elimination of the 0/0 undeterminacy *)

Num = (* Please insert here your numerator *) ;

Den =(* Please insert here your denominator *) ;

type = (* Please insert 1 for oscillatory case and 2 for hyperbolic case*) ;

Off[General::spell]

degNum = vDeg[Num]; {NumEta, kNum, rNum} = etaTransf[Num, degNum];

NumNew = ZpowerTransf[NumEta, kNum];

degDen = vDeg[Den]; {DenEta, kDen, rDen} = etaTransf[Den, degDen];

DenNew = ZpowerTransf[DenEta, kDen];

Print["Transformed coefficient:", FullSimplify[$\frac{NumNew}{DenNew}$]]

Module vDeg

(* Function that computes the powers of v to highlight *)

vDeg[funz_] := Module[{Nmax, tmp, deg}, Nmax = 100; tmp = 0;

For[deg = 1, (deg < Nmax && tmp == 0),

tmp = SeriesCoefficient[Series[funz, {v, 0, Nmax}], deg]; deg++];

deg = deg - 1; deg]

Module etaTransf

(* Function that reveals the full power of v and changes in the eta functions*)

etaTransf[funzSt_, deg_] := Module[{degNew, deg1, tmp, Nmax, r, k, funz},

funz = funzSt; tmp = 0;

For[deg1 = 1, (deg1 < Nmax && tmp == 0),

tmp = Coefficient[Denominator[Together[funz/(v^deg1)]], v]; deg1++;];

deg1 = deg1 - 2;

If[deg1 > 0, funz = Together[funz/(v^deg1)]; degNew = deg - deg1,

degNew = deg;];

k = IntegerPart[degNew/2]; r = Mod[degNew, 2];

funz=Together[$\frac{1}{v^r}$ (funz /. {HoldPattern[Cos[a : _]] :> η_{-1}[-a^2],

HoldPattern[Cosh[a : _]] :>η_{-1}[-a^2],

HoldPattern[Sin[a : _]] :> a*η_0[-a^2],

HoldPattern[Sinh[a : _]] :> a*η_0[-a^2]})

/. {v^2 -> ((-1)^type))*Z})];

If[(Mod[(deg1 + r)/2, 2] == 1) && type == 1, funzMod = -funzMod;];

If[deg1 > 0, r = r + deg1;]; {funz, k, r}]

Module ZpowerTransf

ZpowerTransf[funzSt_,kSt_]:=Module[{funz,Nmax,tmp,deg,k,deg1,s},

Nmax=100; funz=funzSt; k=kSt; $\eta_{-1}[0] = 1$;

$\eta_{m\_}[0] = \frac{1}{Factorial2[2*m+1]}$;

$D_{m\_}[Z\_] = \eta_m[0] * \left(\frac{1}{2} * \eta_0\left[\frac{Z}{4}\right] * \eta_0\left[\frac{Z}{4}\right] - \sum_{i=1}^{m+1}(Factorial2[2*i-3] * \eta_i[Z])\right)$;

$\overline{\eta}_{m\_}[Z\_] := \eta_m[0] + Z * D_m[Z]$;

s=0; st=-1;

If[k $\geq$ 0,funz=Together[$\frac{funz/.Table[\eta_i->\overline{\eta}_i,\{j,st,s\}]}{Z}$];];

funz=Simplify[funz]/.Table[$v^{2*j} -> (((-1)^type) * Z)^j, \{j, 1, s+1\}$];

tmp=0;

For[deg1=1,($deg1 < Nmax$&&tmp==0),

tmp=Coefficient[Denominator[Together[funz/(Z^deg1)]],Z]; deg1++];

deg1=deg1-2;

If[deg1>0,funz=Together[funz/(Z^deg1)];k=k-deg1;];

st=0;

For[s=1,s¡k-1,s++,

funzCoeff=CoefficientList[funz,Z]; funzTNot=funzCoeff[[1]];

funzTNot=funzTNot/.Table[$\eta_i -> \overline{\eta}_i, \{j, st, s\}$];

M=Length[funzCoeff];

funz=Together$\left[\left(\sum_{k=2}^{M}\left(funzCoeff[[k]] * Z^{k-1}\right) + funzTNot\right)/Z\right]$;];

If[s==k-1,

funzCoeff=CoefficientList[funz,Z]; funzTNot=funzCoeff[[1]];

funzTNot=funzTNot/.Flatten[Table[$\left\{a : ((\eta_0)\,[\_\_])^{2j} : -> 1 + Factor[a-1]\right\}$, $\{j, 1, s+1\}$]];

funzTNot=funzTNot/.Flatten[{$a : (-1 + (\eta_0)\,[\_\_]) : ->$ $ReplaceAll[a, \eta_0 ->$ $\overline{\eta}_0], Table\,[\eta_i -> \overline{\eta}_i, \{j, st, s\}]$}];

M=Length[funzCoeff];

funz=Together$\left[\left(\sum_{k=2}^{M}\left(funzCoeff[[k]] * Z^{k-1}\right) + funzTNot\right)/Z\right]$;];

funz]

TEST RUN OUTPUT

Transformed coefficient: $\frac{\eta_0^2\left(\frac{Z}{16}\right)-2\eta_1\left(\frac{Z}{4}\right)}{8\theta^2\eta_0^2\left(\frac{Z\theta^2}{4}\right)}$.

Bibliography

[1] Albrecht, P. Numerical treatment of ODEs: the theory of A-methods, *Numer. Math.* 47 (1985), no. 1, 59–87.

[2] Albrecht, P. A new theoretical approach to Runge–Kutta methods, *SIAM J. Numer. Anal.* 24 (1987), no. 2, 391–406.

[3] Albrecht, P. The extension of the theory of A-methods to RK–methods, in *Numerical treatment of differential equations* (Halle, 1987), 8–18, Teubner-Texte Math., 104, Teubner, Leipzig, 1988.

[4] Albrecht, P. The Runge-Kutta theory in a nutshell, *SIAM J. Numer. Anal.* 33 (1996), no. 5, 1712–1735.

[5] Anastassi, Z. A., Simos, T. E. A family of exponentially-fitted Runge–Kutta methods with exponential order up to three for the numerical solution of the Schrodinger equation, *J. Math. Chem.* 41(1), 79–100 (2007).

[6] Bartoszewski, Z., Jackiewicz, Construction of two-step Runge-Kutta methods of high order for ordinary differential equations, *Numer. Algorithms* 18 (1998), 51–70.

[7] Bartoszewski, Z., Jackiewicz, Z. Stability analysis of two-step Runge–Kutta methods for delay differential equations, *Comput. Math. Appl.* 44 (2002), 83–93.

[8] Bartoszewski, Z., Jackiewicz, Z. Nordsieck representation of two-step Runge–Kutta methods for ordinary differential equations, *Appl. Numer. Math.* 53 (2005), 149–163.

[9] Bartoszewski, Z., Jackiewicz, Z. Derivation of continuous explicit two-step Runge–Kutta methods of order three, *J. Comput. Appl. Math.* 205 (2007), 764–776.

[10] Bartoszewski, Z., Jackiewicz, Z. Construction of highly stable parallel two-step Runge–Kutta methods for delay differential equations, *J. Comput. Appl. Math.* 220 (2008), no 1-2, 257–270.

[11] Bartoszewski, Z., Podhaisky, H., Weiner, R. Construction of stiffly accurate two-step Runge–Kutta methods of order three and their continuous extensions using Nordsieck representation, Report No. 01, Martin-Luther-Universität Halle-Wittenberg, Fachbereich Mathematik und Informatik (2007).

[12] Beuter, A., Glass, L., Mackey, M. C., Titcombe, M. S. *Nonlinear Dynamics in Physiology and Medicine*, Springer, 2002.

[13] Byrne, G. D., Lambert, R. J., *Pseudo-Runge–Kutta methods involving two points*, J. Assoc. Comput. Mach. **13**, 114–123, 1966.

[14] Burrage, K., Parallel and Sequention Methods for Ordinary Differential Equations, Oxford Science Publications, Clarendon Press, Oxford 1995.

[15] Burrage, K., Butcher, J. C. Non-linear stability for a general class of differential equation methods, *BIT* 20 (1980), 185-203.

[16] Butcher, J.C.,A modified multistep method for the numerical integration of ordinary differential equations, *J. Assoc. Mach.* 12 (1965), 124–135.

[17] Butcher, J.C. Diagonally-implicit multi-stage integration methods, *Appl. Numer. Math.* 11 (1993), 347–363.

[18] Butcher, J. C. *The Numerical Analysis of Ordinary Differential Equations. Runge–Kutta and General Linear Methods*, John Wiley & Sons, Chichester, New York, 1987.

[19] Butcher, J. C., An introduction to 'Almost Runge-Kutta' methods, *Appl. Numer. Math.* 24 (1997), 331–342.

[20] Butcher, J. C., ARK methods up to order five, *Numer. Algorithms* 17 (1998), 193–221.

[21] Butcher, J.C. *Numerical Methods for Ordinary Differential Equations*, 2nd Edition, John Wiley & Sons, Chichester, 2008.

[22] Butcher, J. C., Moir, N., Experiments with a new fifth order method, *Numer. Algorithms* 33 (2003), 137–151.

[23] Butcher, J. C., Tracogna, S. Order conditions for two-step Runge–Kutta methods, *Appl. Numer. Math.* 24 (1997), 351–364.

[24] Calvo, M., Franco, M. J., Montijano, J. I., Randez, L. Sixth-order symmetric and symplectic exponentially fitted modified Runge–Kutta methods of Gauss type, *Comput. Phys. Commun.* 178 (2008), no. 10, 732–744.

[25] Calvo, M., Franco, J. M., Montijano, J. I., Randez, L. Sixth-order symmetric and symplectic exponentially fitted Runge–Kutta methods of the Gauss type. *J. Comput. Appl. Math.* 223 (2009), no. 1, 387–398.

[26] Carpentieri, M., Paternoster, B. *Stability regions of one step mixed collocation methods for* $y'' = f(x, y)$, Appl. Num. Math., to appear.

[27] Carpentieri, M., Paternoster, B. *Stability analysis of frequency and step length dependent Runge–Kutta-Nyström methods*, Future Generation of Computer Systems, to appear.

[28] Cash, J. R., Efficient numerical methods for the solution of stiff initial-value problems and differential algebraic equations, *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* 459 (2003), no. 2032, 797–815.

[29] Chollom, J., Jackiewicz, Z. Construction of two-step Runge–Kutta methods with large regions of absolute stability, *J. Comput. Appl. Math.* 157 (2003), 125–137.

[30] Coleman, J. P., Numerical methods for $y'' = f(x, y)$ via rational approximations for the cosine, *IMA J. Numer. Anal.* **9**, 145–165.

[31] Coleman, J. P. Rational approximations for the cosine function, P–acceptability and order, *Numerical Algorithms* **3** (1992), 143–158.

[32] Coleman, J. P. Order Conditions for a class of two–step methods for $y'' = f(x, y)$, *IMA J. Numer. Anal.* 23 (2003), 197–220.

[33] Coleman, J. P., Duxbury, S. C. Mixed collocation methods for $y'' = f(x, y)$, *J. Comput. Appl. Math.* 126 (2000), no. 1–2, 47–75.

[34] J. P. Coleman, Liviu Gr. Ixaru, P–stability and exponential–fitting methods for $y" = f(x, y)$, *IMA J. Numer. Anal.* 16, 179–199 (1996).

[35] Coleman, J. P., Ixaru, L.Gr. Truncation errors in exponential fitting for oscillatory problems, *SIAM J. of Numerical Analysis* 44 (2006), 1441-1465.

[36] Cong, N. H. Note on the performance of direct and indirect Runge–Kutta–Nyström methods, *J. Comput. Appl. Math.* 45 (1993), no. 3, 347–355.

[37] Cong, N. H. Parallel-iterated pseudo two-step Runge–Kutta-Nyström methods for nonstiff second-order IVPs, *Comput. Math. Appl.* 44 (2002), no. (1-2), 143-155.

[38] Conte, D., D'Ambrosio, R., Jackiewicz, Z., Two-step Runge–Kutta methods with quadratic stability functions, *J. Sci. Comput.* **44**, 2 (2010), 191–218.

[39] Conte, D., D'Ambrosio, R., Jackiewicz, Z., Paternoster, B. Algebraically stable two-step collocation methods for ordinary differential equations, *submitted*.

[40] Conte, D., D'Ambrosio, R., Paternoster, B. Two–step diagonally–implicit collocation based methods for Volterra Integral Equations, in press on *Appl. Numer. Math.*

[41] Conte, D., Esposito, E., Paternoster, B., Ixaru, L. Gr., Some new uses of the $\eta_m(Z)$ functions, *Comput. Phys. Comm.* 181, no. 1, 128–137 (2010).

[42] Conte, D., Jackiewicz, Z., Paternoster, B. Two-step almost collocation methods for Volterra integral equations, *Appl. Math. Comput.* 204 (2008), 839–853.

[43] Costabile, F., Metodi pseudo Runge-Kutta di seconda specie. (Italian) Calcolo **7** (1970), 305-322.

[44] Costabile, F., Un metodo del terzo ordine per l'integrazione numerica dell'equazione differenziale ordinaria. (Italian) Calcolo **8** (1971), 61-75.

[45] Costabile, F., Un metodo con ascisse gaussiane per l'integrazione numerica dell'equazione differenziale ordinaria. (Italian) Rend. Mat. (6) 6 (1973), 733-748.

[46] Costabile, F., Costabile, C. Two-step fourth order P-stable methods for second order differential equations. *BIT* **22** (1982), no. 3, 384-386.

[47] Costabile, F., A survey of pseudo Runge-Kutta methods. *Rend. Mat. Appl.* (7) 23 (2003), no. 2, 217-234 (2004).

[48] Costabile, F., Napoli, A., A class of collocation methods for numerical integration of initial value problems. *Comput. Math. Appl.*62 (2011), no. 8, 3221–3235.

[49] D'Ambrosio, R., Highly stable multistage numerical methods for Functional Equations: theory and implementation issues, Bi-nationally supervised Ph.D. thesis, University of Salerno-Arizona State University, 2010.

[50] D'Ambrosio, R., Esposito, E., Paternoster, B. Exponentially fitted two-step hybrid methods for $y'' = f(x, y)$. *J. Comput. Appl. Math.* 235 (2011), no. 16, 48884897.

[51] D'Ambrosio, R., Esposito, E., Paternoster, B. Parameter estimation in exponentially fitted hybrid methods for second order differential problems. *J. Math. Chem.* 50 (2012), no. 16, 155-168, doi 10.1007/s10910-011-9903-7.

[52] D'Ambrosio, R., Esposito, E., Paternoster, B. Exponentially fitted two-step Runge-Kutta methods: construction and parameter selection. *Applied Mathematics and Computation* 218 (2012), no. 14, 7468–7480, doi: 10.1016/j.amc.2012.01.014.

[53] D'Ambrosio, R., Esposito, E., Paternoster, B., General Linear Methods for $y'' = f(y(t))$: convergence analysis, *Numerical Algorithms* (2011), (submitted).

[54] D'Ambrosio, R., Ferro, M., Jackiewicz, Z., Paternoster, B. Two-step almost collocation methods for ordinary differential equations, *Numer. Algorithms* 53 (2010), no. 2-3, 195-217.

[55] D'Ambrosio, R., Ferro, M., Paternoster, B. Trigonometrically fitted two-step hybrid methods for second order ordinary differential equations with one or two frequencies, *Math. Comput. Simulat.* 81 (5), 1068–1084 (2011).

[56] D'Ambrosio, R., Ferro, M., Paternoster, B. Two-step hybrid collocation methods for $y'' = f(x, y)$, *Appl. Math. Lett.* 22 (2009), 1076–1080.

[57] R. D'Ambrosio, L. Gr. Ixaru, B. Paternoster, Construction of the EF-based Runge–Kutta methods revisited, *Comput. Phys. Commun.* 182(2), 322–329 (2011).

[58] D'Ambrosio, R., Izzo, G., Jackiewicz, Z. Search for highly stable two-step Runge–Kutta methods for ordinary differential equations, in press on *Appl. Numer. Math.*

[59] D'Ambrosio, R., Jackiewicz, Z. Continuous Two–Step Runge–Kutta Methods for Ordinary Differential Equations, *Numer. Algorithms* 54 (2), 169-193 (2010).

[60] D'Ambrosio, R., Jackiewicz, Z. Construction and implementation of highly stable two-step continuous methods for stiff differential systems, *Math. Comput. Simulat.* 81(9),1707–1728 (2011).

[61] D'Ambrosio, R., Paternoster, B. Two-step modified collocation methods with structured coefficient matrices, in press on *Appl. Numer. Math.*

[62] D'Ambrosio, R., Paternoster, B. Runge–Kutta-Nyström Stability for a Class of General Linear Methods for $y'' = f(x, y)$, in *Numerical Analysis and Applied Mathematics* ed. by T. E. Simos, G. Psihoyios, Ch. Tsitouras, *AIP Conference Proceedings* 1168 (2009), no. 1, 444–447.

[63] Dahlquist, G. A special stability property for linear multistep methods, *BIT* 3 (1963), 27-43.

[64] Dahlquist, G. Error analysis for a class of methods for stiff non-linear initial value problems, in Numerical Analysis: Proc. Dundee Conf. 1975 (G. A. Watson, ed.), *Lecture Notes in Mathematics* 506, 60–72, Springer–Verlag, Berlin 1976.

[65] Dahlquist, G., On accuracy and unconditional stability of linear multistep methods for second order differential equations. *BIT* **18** (1978), no. 2, 133-136.

[66] Dahlquist, G. On one-leg multistep methods, *SIAM J. Numer. Anal.* 20 (1983), no. 6, 1130–1138.

[67] Dekker, K., Verwer, J. G. *Stability of Runge–Kutta methods for stiff nonlinear differential equations*, CWI Monograph , North-Holland, Amsterdam, 1984.

[68] Donelson, J. III, Hansen, E. Cyclic composite multistep predictor-corrector methods, *SIAM J. Numer. Anal.* 8 (1971), 137–157.

[69] Duffing, G. *Erzwungene Schwingungen bei Veränderlicher Eigenfrequenz.*, F. Vieweg u. Sohn, Braunschweig, 1918.

[70] Engl, G. J., Ussif, A. Estimation of Trends Using Ordinary Differential Equations: An Application to Occupational Injuries, *Casualty Actuarial Society Forum*, Fall 2004, pp. 481–496.

[71] Fang, Y., Song, Y., Wu, X. Trigonometrically-fitted explicit Numerov-type method for periodic IVPs with two frequencies, *Comput. Phys. Commun.* 179 (2008), vol. 11, 801-811.

[72] Franco, J. M. Exponentially fitted explicit Runge–Kutta-Nyström methods, *J. Comput. Appl. Math.* 167 (2004), 1–19.

[73] Franco, J. M. A class of explicit two-step hybrid methods for second-order IVPs, *J. Comput. Appl. Math.* 187 (2006), no. 1, 41–57.

[74] Frank, J.E., van der Houwen, P.J. Parallel Iteration of the Extended Backward Differentiation Formulas, *IMA J Numer Anal* 21 (2001), no. 1, 367–385.

[75] Gautschi, W. Numerical integration of ordinary differential equations based on trigonometric polynomial, *Numer. Math.* 3 (1961), 381–397.

[76] Gear, C. W., Hybrid methods for initial value problems in ordinary differential equations, *J. Soc. Indust. Appl. Math. Ser. B Numer. Anal.* 2 (1965), 69–86.

[77] Gragg, W. B., Stetter, H. J. Generalized multistep predictor-corrector methods, *J. Assoc. Comput. Mach.* 11 (1964), 188-209.

[78] Hairer, E., Lubich, C. *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal. 38, 414–441, 2000.

[79] Hairer, E., Lubich, C., Wanner, G. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer Series in Computational Mathematics **31**, Springer, Berlin, 2002.

[80] Hairer, E., Wanner, G. *Multistep-Mutistage-Multiderivative methods for Ordinary Differential Equations*, Computing **11**, 287–303, 1973.

[81] Hairer, E., Wanner, G. Order conditions for general two-step Runge–Kutta methods, *SIAM J. Numer. Anal.* 34(1997), 2087-2089.

[82] Hairer, E., Norsett, S. P., Wanner, G. *Solving Ordinary Differential Equations I – Nonstiff Problems*, Springer Series in Computational Mathematics 8, Springer–Verlag, Berlin, 2000.

[83] Hairer, E., Wanner, G. *Solving Ordinary Differential Equations II – Stiff and Differential–Algebraic Problems*, Springer Series in Computational Mathematics 14, Springer–Verlag, Berlin, 2002.

[84] Hairer, E., Wanner, G. On the Butcher group and general multi-value methods, *Computing* 13 (1974), 1-15.

[85] Henrici, P. *Discrete variable methods in ordinary differential equations*, John Wiley & Sons, New York-London, 1962.

[86] Hockbruck, M., Lubich, Ch., *A Gautschi-type method for oscillatory second order differential equations*, Numer. Math. 83, 403–426, 1999.

[87] Hollevoet, D., Van Daele, M., Vanden Berghe, G. The optimal exponentially-fitted Numerov method for solving two-point boundary value problems, *J. Comput. Appl. Math.* 230 (2009), no. 1, 260–269.

[88] Ixaru, L. Gr. Operations on oscillatory functions, *Comput. Phys. Comm.* 105 (1997), 1–19.

[89] L. Gr. Ixaru, *Numerical Methods for Differential Equations and Applications*, Reidel, Dordrecht, Boston, Lancaster, 1984.

[90] Ixaru, L. Gr., De Meyer, H., Vanden Berghe, G. SLCPM12, A program for solving regular Sturm-Liouville problems, *Comput. Phys. Comm.* 118 (1999), no. 2-3, 259–277.

[91] Ixaru, L. Gr., Rizea, M. A. Numerov-like scheme for the numerical solution of the Schrödinger equation in the deep continuum spectrum of energies, *Comput. Phys. Commun.* 19 (1980), 23–27.

[92] L. Gr. Ixaru, G. Vanden Berghe, H. De Meyer, Frequency evaluation in exponential fitting multistep methods for ODEs, *J. Comput. Appl. Math.* 140, 423–435 (2002).

[93] L. Gr. Ixaru, Comput. Phys. Commun. 147 (2002) 834.

[94] Ixaru, L. Gr., Vanden Berghe, G. *Exponential Fitting*, Kluwer Academic Publishers, Dordrecht, 2004.

[95] Iyengar, R. Quantitative models of Mammalian Cell Signaling Pathways, *Sci. Signal.* 1 (2008), no. 7.

[96] Iserles, A., Munthe-Kaas, H. Z., Norsett, S. P., Zanna, A. *Lie-group methods*, Acta Numerica, 215–365, 2000.

[97] Jackiewicz, Z. *General Linear Methods for Ordinary Differential Equations*, John Wiley & Sons, Hoboken, New Jersey, 2009.

[98] Z. Jackiewicz, R. Renaut, A. Feldstein, Two-step Runge–Kutta methods, *SIAM J. Numer. Anal.* 28 (1991), no. 4, 1165–1182.

[99] Jackiewicz, Z., Renaut, R., Zennaro, M. Explicit two-step Runge-Kutta methods, *Appl. Math.* 40 (1995), 433–456.

[100] Jackiewicz, Z., Tracogna, S. A general class of two-step Runge–Kutta methods for ordinary differential equations, *SIAM J. Numer. Anal.* 32 (1995), 1390–1427.

[101] Jackiewicz, Z., Tracogna, S. Variable stepsize continuous two-step Runge–Kutta methods for ordinary differential equations, *Numer. Algorithms* 12 (1996), 347–368.

[102] Jackiewicz, Z., Verner, J. H. Derivation and implementation of two-step Runge–Kutta pairs, *Japan J. Indust. Appl. Math.* 19 (2002), 227–248.

[103] Jackiewicz, Z., Zennaro, M. Variable stepsize explicit two-step Runge-Kutta methods, *Math. Comput.* 59 (1992), 421–438.

[104] Jamrog, D. C., Szipro A. A. *A Smallpox and an Inhalation Anthrax Model Implemented Using Ordinary Differential Equations*, Technical Report of Lincoln Laboratory, Massachussetts Insitute of Technology, 2006.

[105] Kalogiratou, Z., Monovasilis, Th., Simos, T. E. Computation of the eigenvalues of the Schrödinger equation by exponentially-fitted Runge-Kutta-Nyström methods, *Comput. Phys. Comm.* 180 (2009), no. 2, 167–176.

[106] Konguetsof, A., Simos, T. E. An exponentially-fitted and trigonometrically-fitted method for the numerical solution of periodic initial-value problems, *Comput. Math. Appl.* 45 (2003), no. 1-3, 547–554.

[107] Kramarz, L., Stability of collocation methods for the numerical solution of $y'' = f(t, y)$, BIT **20** (1980) 215–222.

[108] Lambert, J. D., Computational methods in Ordinary Differential Equations, John Wiley & Sons, Chichester, 1973.

[109] Lambert, J. D. *Numerical methods for ordinary differential systems: The initial value problem*, John & Wiley, Chichester, 1991.

[110] Lambert, J. D., Watson, I. A.: Symmetric multistep methods for periodic initial value problems. *J. Inst. Math. Appl.* **18**, 189–202 (1976).

[111] Landau, L. D., and Lifshitz, F. M., *Quantum Mechanics*, Pergamon, New York, (1965).

[112] Lang, S. *Linear Algebra*, Springer, 1987.

[113] Leary, R. A., Nimerfro, K., Holdaway, M., Brand, G., Burk, T., Kolka, R., Wolf, A. *Height growth modeling using second order differential equations and the importance of initial height growth*, *Forest Ecol. Manag.* 57 (1997), no. 2, 165–172.

[114] Ledoux, V., Van Daele, M., Vanden Berghe, G. MATSLISE: a MATLAB package for the numerical solution of Sturm-Liouville and Schrödinger equations, *ACM Trans. Math. Software* 31 (2005), no. 4, 532–554.

[115] Lyche, T. Chebyshevian multistep methods for ordinary differential equations, *Numer. Math.* 19 (1972), 65–75.

[116] Martucci, S., Paternoster, B. Vandermonde-type matrices in two step collocation methods for special second order Ordinary Differential Equations, in *Computational Science ICCS 2004*, ed. by M. Bubak em et al., Lecture Notes in Comput. Sci. 3039, Part IV, Springer Verlag, 2004, 418-425.

[117] A. Messiah *Quantum Mechanics*, Dover Publications, New York, (1999).

[118] Messina, E., de Swart, J. J. B., van der Veen, W. *Parallel iterative linear solvers for multistep Runge–Kutta methods*, J. Comput. Appl. Math. **85**, 145–167, 1997.

[119] Nakashima, Masaharu On a pseudo-Runge-Kutta method of order 6. Proc. Japan Acad. Ser. A Math. Sci. 58 (1982), no. 2, 6668

[120] Norberto, A. M., Benavente, M. A., Eguaras, M. A Model in Differential Equations to Describe the Mite Varroa Destructor Population Dynamic in Apis Mellifera Colonies, *Foro Red. Mat.* 16 (2005).

[121] Oliver, J., A curiosity of low order explicit Runge-Kutta methods, *Math. Comput.* 29 (1975), 1032–1036.

[122] Oppenheim, A. V., Schafer, R. W., Buck, J. R. *Discrete-Time Signal Processing*, 2nd edn., Prentice Hall, Upper Saddle River, 1999.

[123] Ozawa, K. A functional fitting Runge–Kutta method with variable coefficients, *Jpn. J. Ind. Appl. Math.* 18 (2001), no. 1, 107-130.

[124] Ozawa, K. A functional fitting Runge–Kutta-Nyström method with variable coefficients *Jpn. J. Ind. Appl. Math.* 19 (2002), no. 1, 55-85.

[125] Paternoster, B. Runge–Kutta(-Nyström) methods for ODEs with periodic solutions based on trigonometric polynomials, *Appl. Numer. Math.* 28 (1998), 401–412.

[126] Paternoster, B. *General two-step Runge–Kutta methods based on algebraic and trigonometric polynomials*, Int. J. Appl. Math. **6**(4), 347–362, 2001.

[127] Paternoster, B. Two step Runge–Kutta-Nyström methods for $y'' = f(x, y)$ and P–stability, in *Computational Science – ICCS 2002*, ed. by P.M.A.Sloot, C.J.K.Tan, J.J.Dongarra, A.G.Hoekstra, *Lecture Notes in Computer Science* 2331, Part III, 459-466, Springer Verlag, Amsterdam, (2002).

[128] Paternoster, B. A phase-fitted collocation-based Runge–Kutta-Nyström method, *Appl. Numer. Math.* **35**(4) (2000), 239–355.

[129] Paternoster, B. Two step Runge–Kutta-Nystrom methods for oscillatory problems based on mixed polynomials, *Computational Science– ICCS 2003*, ed. by P.M.A.Sloot, D.Abramson, A.V.Bogdanov, J.J.Dongarra, A.Y.Zomaya, Y.E.Gorbachev, *Lecture Notes in Computer Science* 2658, Part II, 131-138, Springer, Berlin Heidelberg, 2003.

[130] Paternoster, B. Two step Runge–Kutta-Nystrom methods based on algebraic polynomials, *Rendiconti di Matematica e sue Applicazioni* 23 (2003), Serie VII, 277–288.

[131] Paternoster, B. A general family of two step Runge–Kutta-Nyström methods for $y'' = f(x, y)$ based on algebraic polynomials, *Computational Science – ICCS 2003* ed. by V.N.Alexandrov, G.D. van Albada, P.M.A.Sloot, J.J.Dongarra, *Lecture Notes in Computer Science* 3994, Part IV, 700-707, Springer Verlag, 2006.

[132] L. R. Petzold, L. O. Jay, J. Yen, Numerical solution of highly oscillatory ordinary differential equations, *Acta Numerica* 6, 437–483 (1997).

[133] A. Prothero and A. Robinson, On the stability and the accuracy of one-step methods for solving stiff systems of ordinary differential equations, *Math. Comput.* 28, 145–162 (1974).

[134] Quarteroni, A., Sacco, R., Saleri, F. Numerical mathematics. Second edition. Texts in Applied Mathematics, 37. Springer-Verlag, Berlin, 2007. xviii+655 pp.

[135] Raptis, A. D., Allison, A. C. Exponential-fitting methods for the numerical solution of the Schrödinger equation, *Computational PHYSICS Communications* 14 (1978), 1–5.

[136] Renaut, R. A. Two-step Runge-Kutta methods and hyperbolic partial differential equations, *Math. Comput.* 55, (1990), 563–579.

[137] Simos, T. E. Dissipative trigonometrically-fitted methods for linear second order IVPs with oscillating solution, *Appl. Math. Lett.* 17, 5 (2004), 601–607.

[138] Simos, T. E. An exponentially-fitted Runge–Kutta method for the numerical integration of initial-value problems with periodic or oscillating solutions, *Comput. Phys. Comm.* 115 (1), 1–8 (1998).

[139] Simos, T. E. Exponentially-fitted Runge–Kutta-Nyström method for the numerical solution of initial-value problems with oscillating solutions, *Appl. Math. Lett.* 15 (2002), no. 2, 217-225.

[140] Simos, T. E. High-order closed Newton-Cotes trigonometrically-fitted formulae for long-time integration of orbital problems, *Comput. Phys. Comm.* 178 (2008), no. 3, 199–207.

[141] Spencer, S. L., Berryman, M. J., Garca, J. A., Abbott, D. An ordinary differential equation model for the multistep transformation to cancer, *J. Theoret. Biol.* 231 (2004), no. 4, 515–524.

[142] Stiefel, E., Bettis, D. G. Stabilization of Cowell's method, *Numer. Math* 13, (1969), 154–175.

[143] Tracogna, S. Implementation of two-step Runge–Kutta methods for ordinary differential equations, *J. Comput. Appl. Math.* 76 (1997), 113–136.

[144] Tracogna, S., Welfert, B. Two-step Runge–Kutta: Theory and practice, *BIT* 40 (2000), 775–799.

[145] Ueda, Y. Explosion of strange attractors exhibited by Duffing's equation, in *Nonlinear Dynamics*, RHG Helleman (ed.), 422–434, New York Academy of Science, 1980.

[146] Ueda, Y. *The Road to Chaos*, Aerial Press, 1992.

[147] Vanden Berghe, G., De Meyer, H., Van Daele, M., Van Hecke, T. Exponentially fitted Runge–Kutta methods, *Comput. Phys. Comm.* 123 (1999), 7–15.

[148] Vanden Berghe, G., Van Daele, M. Exponentially-fitted Numerov methods, *J. Comput. Appl. Math.* 200 (2007), 140–153.

[149] Vanden Berghe, G., Van Daele, M. Exponentially-fitted Obrechkoff methods for second-order differential equations, *Appl. Numer. Math.* 59 (2009), no. 3-4, 815–829.

[150] P.J.van der Houwen, B.P.Sommeijer, Explicit Runge–Kutta-Nyström methods with reduced phase errors for computing oscillating solution, *SIAM J. Num. Anal.* 24, 595–617, (1987).

[151] P. J. van der Houwen, B. P.Sommeijer, Diagonally implicit Runge–Kutta Nyström methods for oscillatory problems, *SIAM J. Num. Anal.* 26, 414–429 (1989).

[152] van der Houwen, P. J., Sommeijer, B. P., Cong, N. H. Stability of collocation–based Runge–Kutta–Nyström methods, *BIT* 31 (1991), 469–481.

[153] P.J.van der Houwen, B.P.Sommeijer, K.Strehmel, R.Weiner, On the numerical integration of second order initial value problems with a periodic forcing function, *Computing* 37, 195–218, 1986.

[154] H. Van de Vyver, Phase-fitted and amplification-fitted two-step hybrid methods for $y'' = f(x, y)$, *J. Comput. Appl. Math.* 209, 33–53 (2007).

[155] Volterra, V. Variations and fluctuations of the number of individuals in animal species living together. Reprinted in R.N. Chapman, *Animal Ecology*, McGraw Hill, New York, 1931.

[156] Williams, P. S., Simos, T. E. A new family of exponentially fitted methods, *Math. Comput. Modelling* 38 (2003), no. 5-6, 571-584.

[157] Wright, W. M. *General linear methods with inherent Runge–Kutta stability*, Ph.D. thesis, The University of Auckland, New Zealand, 2002.

[158] Zhao, X. Q. *Dynamical Systems in Population Biology*, CMS Books in Mathematics, New York, Springer–Verlag, 2003.

[159] Zlatev, Z., Modified diagonally implicit Runge-Kutta methods, *SIAM J. Sci. Stat. Comput.* 2 (1981), 321–334.